

## ***CELERRA T2 TECH GUIDE***

### **INTRODUCTION:**

The term “Celerra” generally applies to the entire physical “Celerra Cabinet” or system, the major components of which are the Control Station(s), and Data Movers (Blades). Data Movers, or Blades, are also referred to as Servers, the Celerra File Server [CFS], CIFS Servers, NFS Servers, etc. Data Movers operate independently of each other, and for normal operation, do not depend on the Control Station. The Control Station is used to configure, monitor, execute scripts/tasks, and manage the Servers. The Control Station is needed for communicating, changing, or configuring the system, for rebooting or “failing over” a Server, and for generating events and CallHomes, etc. When securing, shutting down, or rebooting a “server”, this is done using the server\_cpu or nas\_halt commands. The Control Station is usually shutdown or rebooted separately from the Servers [via CLI reboot, init 0, init 6, or other Linux shutdown commands].

### **CELERRA FUNCTION:**

Celerra Servers perform “File Sharing” using NFS, CIFS, iSCSI, and MPFS protocols (network access protocols). *Celerra Servers* run a specialized O/S called “**DART**” [Data Access in Real Time], more commonly referred to as “NAS”. The current NAS family is 6.0.

### **CELERRA PRODUCT LINES:**

#### **NS Series—Gateway:**

Single, dual, or quad Data Mover Gateway configurations connected to Symmetrix and/or CLARiiON CX Storage systems

#### **NS Series—Integrated:**

Single, dual, or quad Data Mover configurations connected to dedicated CLARiiON Storage Systems  
[NS700/NS500/NS600/NS80/NS350/NS40/NS20]

#### **NSX Gen I & II Series:**

High-end NAS platform based on NS design and connecting to either Symmetrix or CLARiiON storage

#### **CX3-Compatible NS Series:**

NS40, NS80, & NS20 (August 2007) are all based on the Clariion CX3 platforms

#### **NX4 low-end Model:**

NX4 is an Integrated-only low-end model based on the AX4-5F8 CLARiiON Mamba array, GA August 2008

#### **NS Mid-range ‘Unified Storage’ Series:**

NS-120, NS-480 Integrated/MPFS series using dedicated CX4 arrays, GA December 2008

#### **NS-960/NS-G8 Series:**

Integrated & Gateway high-end platform, respectively, GA February 2009, first Celerra front-end to be based on CX4 hardware. 2-8 Blades

#### **VG2/VG8 Gateways:** Mid and High-end systems, respectively. GA Sept 2010

#### **2010/2011:**

→Convergence of Clariion/Celerra family

### **LEGACY PRODUCT LINES:**

#### **EMC Celerra Clustered Network Server (CNS), CFS-14, SE, NS:**

→Celerra CNS (CFS-14 with Symm, aka Eagle), CNS-14 Golden Eagle, Celerra SE, Celerra NS600, CNS & Clariion CX600, CNS & Clariion FC4700

**Note:** CNS stands for “Clustered Network Server”, 2-14 DMs, 1-2 CSs

#### **# /nas/sbin/model**

CNS-14

### **CELERRA CONTROL STATION:**

- a.) User Interface for the Celerra system (CLI or GUI)
- b.) Vehicle for making configuration changes (CLI or GUI)
- c.) Automatically initiates Server failover/recovery and handles Callhome events when NAS Services are running
- d.) Logs Celerra Events—Server Log; NAS\_EVENTS; sys\_log; Celerra Manager
- e.) Monitors Health (heartbeat) of Servers via Internal Interfaces [using Ping protocol every 5 seconds]
- f.) Monitors Celerra Cabinet hardware—fans, batteries, temps, voltages, enclosures & blades via backplane/midplane
- g.) Manages power-up sequence

- h.) Used for troubleshooting, diagnostics, and monitoring of System and Datamovers
- i.) Performs hourly NAS database backups of NAS\_DB & DMBS Databases, and other CRON-scheduled functions
- j.) Uses NBS client to query and/or make changes to backend configurations related to the Celerra

## **CONTROL STATION MANAGEMENT:** Access Control, Management, Administration, Statistics, Monitoring

Access Control—ACL support; User & Primary Groups; Owner, Read, Write, Delete Operations; 0 [any used logged in], 2 [admin], 3 [operator], 4 [observer] Access Levels

Management-----Unix-based commands with SNMP support and Scripts; Dual Control Station support; GUI Management; Web. Administration---Installs, Upgrades, Volume & FS management, NIC configuration; NFS & CIFS setup, FS exports, FSCK, Extension of FileSystems; Automatic database backup of nasdb every hour for up to 12 hours.

Statistics-----sysstat; netstat; nfsstat

Monitoring-----Power, Battery, Temp; DataMovers; Environmental events; Uses RPC to communication with DataMovers over TCP/IP; Reporting errors and Call Homes.

**Note:** Dual Control Stations monitor each other using internal network & NAS MCD daemon—Standby CS1 takes over if CS0 stops responding. Control Station monitors each DataMover using ping broadcast every 5 seconds. Data Movers boot and execute “nas.exe” into RAM from the /nas/dos/nas.exe configuration files which are physically located on Hyper #3 [aka, the 3rd symm volume], which both CS & DM “see”—boot.cfg & boot.bat files specifically [Slot info read in /bin]

## **CONTROL STATION MODELS:**

Control Station	GA	Min S/W Rev	ECO #
100-520-581	1-Jun-2006	5.5.21	49235 "Samsung"
100-560-974 (Chivas)	1-Mar-2005	5.4.14	42990 2GB Memory
100-560-688 (Chivas DC)	9 May 2007	5.5.27.5	53241 DC NEBS CS
100-520-665 (Maynard)	22-Aug-2008	5.6.39	62238 2 GB Memory
100-520-606 (Dewars)	14-Aug-2007	5.5.26	52897 "1U Control Station (Falcon)" 2GB Memory
100-520-216	1-Aug-2004	5.2.9.6	39390 "1U Barracuda Control Station" Whalley 512MB Memory
100-520-503	14 Apr 2006	5.5.19	48415 "1U RoHS Barracuda CS" RoHS version of 216
005047772	1-Dec-2002	5.1.9.4	31421 "1U Control Station" 512MB Memory

### **CHIVAS CONTROL STATION:**

--NSX family, 2.8GHz, 2GB memory, 1GB cpu cache

#### **# cat /proc/meminfo**

MemTotal: 2068144 kB

#### **# cat /proc/cpuinfo**

```
processor      : 0
vendor_id     : GenuineIntel
cpu family    : 15
model         : 3
model name    : Intel(R) Pentium(R) 4 CPU 2.80GHz
stepping       : 4
cpu MHz       : 2802.601
cache size    : 1024 KB
```

#### **# /sbin/lspci | grep -i -c "pci bridge"**

4

**Note:** Chivas CS uses PCI bridge with (4) Ethernet slots

### **DEWARS CONTROL STATION:**

Introduced around Napa 5 timeframe, 2007. Used with new NS20 & NS40 variants, and NSX/NS80 platforms.

1U with floppy, CD-ROM, Pentium IV 2.4GHz CPU, 533MHz fsb, 2GB memory, 512MB cpu cache, 80GB SATA disk drive, built on “Falcon” chassis—Whalley Computer. Serial & LAN ports for network connectivity and field access. Bios password protected: “EMCBIO”.

Part Number: 100-520-606, with base 5.5.28.1 software

### **MAYNARD CONTROL STATION:**

See separate Maynard section: 100-520-665

## **CONTROL STATION CONTROL PROCESSES:**

### **NAS MASTER CONTROL DAEMON (MCD):**

#### **# ps -ef |grep nas\_m**

```
root  22394  1  0 Sep11 ?  00:00:00 /nasmcd/nas_mcd -h /nasmcd /nas/sys/nas_mcd.cfg
```

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
Main file located in /nasmcd. Invoked by S95nas script upon bootup. Nas\_mcd configured via /nas/sys/nas\_mcd.cfg. nas\_mcd starts up all other nas processes and monitors Control Station. The NAS Master Control Daemon runs on both Control Stations and monitors each other's heartbeat. The 'heartbeat' timeout threshold is 4 seconds (or loss of 4 consecutive heartbeats). Normal operation is for a heartbeat ping to be sent over each dual internal network once every second. 4 consecutive losses of heartbeats on both internal interfaces will initiate a control station failover.

--nas\_mcd initiates failover  
--nas\_mcd initiates Call Homes  
--nas\_mcd logs Events to /nas/log/sys\_log  
--overwrites crontab on failover /nas/sbin/nasdb\_backup

### **CHANGING CS FAILOVER TIME:**

1. Edit following line in /etc/fc3.d/S95nas and add highlighted value:

(Line 303) export NAS\_DB NBSNAS\_DB; \$MCDHOME/\$NAS\_MCD **-t 30** -h \$MCDHOME \$SYS/\$NAS\_MCD\_CFG > /dev/null 2>&1 &

2. Reboot CS or stop and restart NAS Services

**Note:** Above example changes failover from default of 4 seconds to 30 seconds, which is equivalent to (30) heartbeats.

### **NAS MCD STARTUP FILE:**

#### **/nas/sys/nas\_mcd.cfg**

Contains startup scripts for the following Processes: Event Log; Box Monitor; CallHome; DB\_Backup; Syslog Trimmer; CmdErrlog Trimmer; Cmdlog Trimmer; Smmflog Trimmer; Srdflog Trimmer; HttpAccessLog Trimmer; HttpAgentLog Trimmer; Http Referlog Trimmer; HttpErrorlog Trimmer; JServer; CIFS Server Manager, Usrmapper, APL\_Task Mgr, WebUI, Indications Manager, etc.

### **BOX MONITOR:**

**# ps -ef |grep boxm**

root 25171 22394 0 Sep11 ? 00:00:08 /nas/sbin/nas\_boxmonitor /nas -i

Runs only on the Control Station acting as the Primary. Monitors health of cabinet and invokes DM failover. Spawns sub-processes for database management and configuration changes. Monitors SPB via Comm Boards.

**What do following sys log messages mean?**

May 8 13:33:26 2003 BoxMonitor:6:45 Slot 2 has rebooted. (0)

May 8 13:33:26 2003 BoxMonitor:6:45 Slot 3 has rebooted. (0)

**Jul 11 11:23:07 2003 BoxMonitor:6:45 Slot 3 has rebooted. (1)**

**Comment:** Sys\_log messages that contain a (0) indicate that Box Monitor process has been started & has contacted each DM—it does not mean that they have been rebooted! The (1) status at the end of the line would mean that the Server has been rebooted.

### **LISTING TYPES OF CELERRA CABINET:** NAS 5.0 +

**#/nasmcd/sbin/t2cab -l**

T2cab - Version 5.01.14 - 11/18/02

usage: t2cab -f

Returns the cabinet family and child of the family.

-f : Returns the Cabinet Family only

Where cabinet family is:

**0: Unknown Cabinet**

**1: Eagle Cabinet** [Old CFS-14 Style Cabinet—shoebox style]

**2: Golden Eagle Cabinet** [CNS-14 System]

**3: Celerra IP Cabinet** [Celerra NS systems, Clariion or Symm backend]

Golden Eagle Cabinet Children Include:

Golden Eagle 1

Eagle Cabinet Children Include:

The Spare Lplane connector is present :

No fans present - Raven.

4 fans present - Bobyx - Hawk with Bypass cable.

1,2,3 fans present - Illegal Cabinet configuration.

The Spare Lplane connector is not present :

1 fan present - Anaconda - Eagle.

4 fans present - Bobyx with out a Bypass cable.

No,2,3 fans present - Illegal Cabinet configuration.

**#/nasmcd/sbin/t2slot -f** [Returns a value from the list above that identifies the type of Cabinet]

**\$ /nas/sbin/t2cab**

1 - EAGLE CABINET

The Cabinet child is : 1 - an ANACONDA - EAGLE

**USE T2SLOT COMMAND TO VERIFY CONTROL STATION SLOT:**

**#/nasmcd/sbin/t2slot**

0

**ACCESSING BIOS ON LINUX CONTROL STATION:**

1. At Emulex LP8000 Bios message, press <Alt E>, then <s> to skip Bios
2. Select Emulex device
3. Configure Boot Device
4. Select Unused DID and ensure that starting LUN is set to 00
5. Select Boot Device as WWPN and reboot

**LATER CONTROL STATION MODELS REQUIRE BIOS PASSWORD:** e.g., NS700

1. Press F2 when displayed during initial system bootup

Entering SETUP...

Enter CURRENT Password: **emcbios** [password is not case-sensitive]

2. Available BIOS Menu Items:

Main      Advanced      Security      Boot      System Management      Exit

→To select an item from the Menu, use → arrow key, then enter

→Use esc once or twice to exit BIOS program without making any changes

**CONTROL STATION SYSTEM ERRORS ON STARTUP:**

Check the DMI Event Log for bios-related/hardware messages, aka SMBIOS System Management BIOS, or DMI

Security→Event Log Configuration <enter> select DMI Event Log or other option

**NAS SERVICES STARTUP SCRIPT:** /etc/rc.d/rc3.d

lrwxrwxrwx 1 root root 20 Mar 17 15:49 S95nas -> /etc/rc.d/init.d/nas (Special startup script)

**TRUSS & STRACE:** Traces System Calls, Signals Received, and Machine Faults on the CS for SCO [Linux uses “strace”]

-p process list -f follow child processes -c summarize -a show argument strings -e environment strings -v verbose -o output to a file

**\$truss -p /proc/503 [PID#]**

**Running “truss” Command on Control Station to Identify Errant Processes:**

**\$truss -aef /nas/sbin/getreason** [collect the output]

**\$truss -o /tmp/outfile -f /nas/sbin/getreason** [outputs to a file called ‘outfile’]

**\$strace -o /tmp/aaa -f server\_export server\_2 -p -u -P cifs -n testshare** [Use to output strace to file]

**\$strace -p 1260** [Use this command to attach to a running process and observe output]

**EXAMINING PROCESSES:**

**# cat /proc/4338/status**

Name: java

State: S (sleeping)

Pid: 4338

PPid: 4295

-----abridged-----

**# pstree -pau 1181** →Useful in laying out processes in tree fashion—nas\_mcd in this example

```
nas_mcd,1181) -h /nasmcd /nas/sys/nas_mcd.cfg
|_ch_monitor,2307) /nas/sbin/ch_monitor /nas
| `ch_monitor,2313) /nas/sbin/ch_monitor /nas
|   `sleep,28418) 1
|-cs_res_mon_drv,2443) /nas/sbin/cs_res_mon_driver
|   `sleep,21798) 1800
|-dirsSync,2517) /nas/sbin/dirsSync /nas /nbsnas 180
|   `sleep,27940) 180
|-dirsSync,2518) /nas/sbin/dirsSync -c /nas/sys /nasmcd/CHomeFiles 300
|   `sleep,28133) 300
|-evtClctrMon,2309) /nas/sbin/evntClctrMon
|   `sleep,27448) 300
```

**# ps -elf**

F	S	UID	PID	PPID	C	PRI	NI	ADDR	SZ	WCHAN	STIME	TTY	TIME
CMD													
100	S	root	1	0	0	68	0	-	350	do_sel	Apr14 ?	00:00:10	
init													

**GDB UTILITY:**

**# gdb -pid 385 bt**

**Note:** Linux utility that can be used to view threads of processes on the Control Station

**CONTROL STATION LOGS--TROUBLESHOOTING:**

**History:** \$more /home/nasadmin/.bash\_history [command history log from root shell]

**Operating System Logs & Information:**

\$var/log/messages [great source of info & system config]

\$var/log/boot.log [good source to see what services started or failed on bootup]

\$var/log/dmesg [Log during last boot up]

**TAILING MESSAGES LOG:**

# tail -f /var/log/messages -n0 [Opens logs with no previous lines showing]

**CELERRA DATA MOVERS:**

**NS700 MODEL FILE SERVER:**

Used in CX700, dual 3.1GHz P4 processors, 4GB RAM, 533MHz FSB, Barracuda Motherboard, supports 1 or 2 GBps fibre speeds, two 10/100 management ports using serial/peer-power connection to CS, (6) Broadcom GigE NICs, (2) Agilent 2-GB FC HBA's, Falcon Control Station, NAS 5.3.10.4 and higher

**# server\_sysconfig server\_2 -P**

server\_2 :

Processor = Intel Pentium 4

Processor speed (MHz) = 3100

Total main memory (MB) = 4024

Mother board = Barracuda XP

Bus speed (MHz) = 533

**Note:** Uses Agilent Fibre Channel HBA Controllers

**# server\_sysconfig server\_2 -pci**

server\_2 : PCI DEVICES:

On Board:

Agilent Fibre Channel Controller

0: fcp-0 IRQ: 22 addr: 500601601060101a

0: fcp-1 IRQ: 21 addr: 500601611060101a

0: fcp-2 IRQ: 18 addr: 500601621060101a

0: fcp-3 IRQ: 20 addr: 500601631060101a

Broadcom Gigabit Ethernet Controller

0: cge0 IRQ: 23

speed=auto duplex=auto txflowctl=disable rxfwctrl=disable

0: cge1 IRQ: 25

speed=auto duplex=auto txflowctl=disable rxfwctrl=disable

0: cge2 IRQ: 24

speed=auto duplex=auto txflowctl=disable rxfwctrl=disable

0: cge3 IRQ: 26

speed=auto duplex=auto txflowctl=disable rxfwctrl=disable

0: cge4 IRQ: 27

speed=auto duplex=auto txflowctl=disable rxfwctrl=disable

0: cge5 IRQ: 28

speed=auto duplex=auto txflowctl=disable rxfwctrl=disable

**X-Blade 60 HAMMERHEAD NSX DATA MOVER (XBlade 60):**

NSX platforms, dual 3.4GHz processors, 800MHz FSB, 4-16GB RAM, CMB-Hammerhead Motherboard, two 10/100 management ports using external management switch, supports 1 or 2GB fibre speeds, Chivas Control Station

**Note:** Uses Agilent Fibre Channel HBA Controllers. After introduction of the XBlade65 hardware, and NAS 5.5.22.2, the older XBlade 60 modules were converted to support 4Gb fibre, which, as it turns out, is not supported with any code prior to 5.5.22.2.

**X-Blade 65 HAMMERHEAD 2<sup>nd</sup> GENERATION NSX DATA MOVER (XBlade 65):**

NSX, dual 3.6GHz processors, 800MHz FSB, 4-16GB memory, CMB-H2G2 motherboard, two 10/100 management ports using external switch, supports 1, 2, or 4Gbps fibre speeds, has 8 GbE ports, Chivas Control Station

**NS40 X-Blade 40 CELERRA FILE SERVER:**

Introduced NAPA 2 5.5.22.2 with min. Flare support 03.22.040.5.005

Blade hardware used with NS40, NS20, NX4, NS-120, NS-480, NS-G2 platforms

Dual 2.8GHz P4 processors with 4GB DDR memory and 800MHz fsb, two 10/100 management ports using internal management switch, supports 1/2/4GBps fibre speeds, four GbE ports, Chivas Control Station with 2GB memory, avail. Gateway & Integrated; Supports up to 32TB of storage; Replacement for NS500 and low-end NS700.

**IO MODULES/PORTS:**

(2) 4Gbps FC ports for array connectivity

(2) 2Gbps FC ports for tape backup connectivity

(2) Serial ports

(4) 10/100/1000 BaseT Ethernet ports as one configuration, or (2) Optical GbE ports & (2) 10/100/1000BaseT ports as another config  
→Single Control Station connecting to Blades via Serial and Ethernet interfaces

→1 or 2 Data Mover configurations

→Supports Symm5, DMX, CX300/500/700, CX3-20/40 platforms

**NS-960/NS-G8 Blade:**

Introduced February 2009 with Foxglove platform, first Blades based on CX4 hardware, with slot IO Module architecture, and separate management switches similar to the NSX design.

**CELERRA FILE SYSTEM CAPACITIES:**

NS700, NSX, NS40, NS80, NX4, NS-120, NS-480, NS-960/NS-G8      Maximum file system size is 16TB with NAS 5.4-5.6

**DATA MOVER CAPACITY PER SERVER USING FIBRE CHANNEL STORAGE:**

Data Mover	NAS 5.3	NAS 5.4	NAS 5.5	NAS 5.6
NS700	8TB	16TB	16TB FC	16TB FC
NSX X-Blade 60	---	16TB FC/ATA	24TB FC/128TB MPFS	128TB FC/128TB MPFS
NSX X-Blade 65	---	----	24TB FC	128TB FC/128TB MPFS
NS20	---	----	16TB FC/ATA	32TB FC/48TB ATA/128TB MPFS
NS40	---	----	20TB FC/ATA	64TB FC/128TB MPFS
NS80/Gateway	---	----	24TB FC	128TB FC/128TB MPFS
NX4	---	----	----	16TB FC
NS-120	---	---	---	32TB FC/48TB ATA/96TB MPFS
NS-480	---	---	---	64TB FC/128TB MPFS
NS-960/NS-G8	---	---	---	128TB FC

**Capacity Increases when combining ATA drives with FC drives:**

NS500 & NS700 increases to 32TB per Data Mover if combining FC & ATA drives with NAS 5.6

NS20 increases to 48TB per blade if combining FC & ATA drives with NAS 5.6

NS40 increases to 64TB per blade if combining FC & ATA drives with NAS 5.6

NS80 & NSX increases to 64TB per blade if combining FC & ATA drives with NAS 5.6

**MAX FILE SIZE:**

16TB, but only 4TB if using Quotas with filesize policy, a hard limit

**CELERRA GUIDELINES:**

Max File Systems/Celerra: 2048 file systems/blade, max. of 4096 total file systems per Celerra

Max Files/Directory: 500,000

Max Ckpts/fs: 32

Max Sub-directories/fs: 65,533 per Celerra Release Notes

Max Depth/Shares: 8

**Max File Systems per Blade:**

2048/4096 for NSX

Note: See Primus emc94606 for more details on total number of file systems that can be mounted.

**Max Shares/fs:**

500 to 4000 [2.2.39.1+], depending on Share Name length of 12 characters and 27 character path & 50 character comment field [limitation is based on a 512kb size limit for the boot.cfg file].

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
NAS 4.x & higher, with 510 Data Movers, can now support up to 10,000 CIFS Shares. Each Share normally uses a single inode, which occupies one disk block, or 512 bytes. If a Share's path and comments add up to 80 characters or more, we must allocate a full 8000 bytes for the Share entry on the rootfs of the DM—if this happens, then the overall number of Shares supported may be considerably less than 10,000. Latest known size limitation of the boot.cfg file is 640kb (NAS 5.5 and below).

**Celerra Storage Block Size:**

8k [this is NOT changeable]

**Max Share Length:**

Max. of 255 characters when using I18N, but depends on UTF-8 character set. ASCII mode max. sharename length is 12 characters.

**Max ACE Entries/Share:**

900

**Max File Name Length:**

255 bytes, depending on I18N dialect

**Max Path Length:**

1024 bytes

**Max Directory Name Length:**

256 characters

**Microsoft Windows Limitations:**

NT 4.0 Server Manager=12 Characters; Net Share Command=80 Characters NT 4.0/WIN2K; WIN2K Computer Management Console=80 Characters NT 4.0/WIN2K

**Max Netbios/Compnames/Blade:**

29 (NAS 4.2)

Limit of (10) Netbios Aliases per DataMover Interface

**NAS 5.0+:**

512 theoretical, limited by localgroups file

**Max Netbios/Compname Name Length:**

15 characters for NetBIOS protocol & 63 UTF-8 characters for Compnames, with 48 character comment field limit

**Max Username Length:**

20 characters

**Max Unix Name Length:**

8 characters

**Max UID Value:**

$2^{31} - 1 = 2,147,483,647$

**Max GID Value/Blade (pre 2.2.46):**

65,535

**Max GID Value/Blade (NAS 2.2.46 +):**

2,147,483,647

**Max GIDs/File System:**

65,535

**Max Windows Domains Membership/Blade:**

32 default, 512 max with following parameter enabled: **param cifs lsarpc.maxDomains=512**

**Note:** Represents max number of domains accessible by DM (increase if ACLs contain more than 32 domains)

**Max TCP Connections/Blade:**

Default for tcp.maxStreams has been changed to 65,535 since NAS 5.3, and is also the maximum value

**Max Users/Cifs Users/TCP Connection:**

16,000 default, 65,536 maximum number FID,TID,UIDs [param cifs listBlocks= increments of 256]

**Max Open files:**

10,000 default, tunable to 40,000 tested [TCP maxStreams]

**Max Multiplexed Connections:**

Clients are limited to total of 127 outstanding requests to Server without first receiving replies

**Max Open files/Blade:**

15,000 in file open cache by default DM507; max. 40,000 tested NS700 & 510 DM; 200,000 for NAS 5.3/5.4 Increase by setting param & rebooting

**file initialize nodes=65536 dnlc=262144** [Listed here as an example only, not a recommended blanket setting]

**CIFS BLOCK SIZE NEGOTIATION FOR DM:**

DM Default=64k; param cifs W95BufSz=; param cifs NTBufSz=; param cifs W2KBufSz; --to increase buffer size negotiated

**Max Tree Quotas:**

Max number per file system—2047 Max size of Tree Quota per file system—4TB

**Max VLANs/Blade:**

4094

**Max VDMs/Blade:**

29 per DM

**CELERRA OPERATING SYSTEM LAYERS:**

Layer 1: Operating system with kernel and kernel debugger

Layer 2: Hardware device drivers for media, network, and scsi driver components

Layer 3: I/O layer for UDP, TCP/IP, CAM, Storage, and components

Layer 4: File system layer, consisting of VFS (Virtual File System), Security, and shared files within a cluster

Layer 5: Programming interfaces for RPC, Common File System (CFS/UFS)

Layer 6: Applications such as CIFS, PAX, NDMP, NFS, FTP, ONCRPC, HTTP, NIS, & SNMP

Layer 7: Management, Configuration, and System components

**DART KERNEL:**

--multiple threads run simultaneously for one physical process

--Real Time threads only run on a single processor at a time

--Dart Domains, or threads, include ‘General Purpose’ threads that are preemptive and accept blocking operation, ‘Real Time’ threads, which are single threaded and non-preemptive, and ‘Interrupt’ threads for Interrupt Services Routines.

**STRUCTURE OF DART KERNEL:**

Execution Contexts; Mutexes & Synchronizer; Thread Management; Scheduling Basics; MP—Specific; Time Services; Network Application; Utilities; Memory allocation & management; Boot Process; Debugging

**NAS CODE VERSIONS:**

5.4.20.1 (5 = code generation; 4 = code family within generation; 20 = maintenance release; 1 = [1-2 digits partial spin; 3 digits = emergency release; 4 digits = debug code; 5 digits = specific eng. integration build]

**O/S EXECUTION CONTEXTS:** (3) Types of Domains

**GP** → General Purpose domains includes all system threads, are preemptible, and accept blocking operations

**RT** → Real Time threads are single non-preemptible threads dedicated to stream, network, and timer callbacks [gets cpu every 20 ms]

**Interrupt** → Interrupt domains include all ISR (Interrupt Service Routines)

**SYNCHRONIZER:** (4) Types of Objects

**DART LOCKS:** Mutexes & RWLocks

**Mutexes:** Mutexes coordinate MUTUAL Exclusive access to objects. The Sthread\_Mutex is used to protect data accessed by GP domains and blocks other threads. Sthread\_MutexSpl protects data accessed from GP or RT domains. Sthread\_MutexRT is an exclusive RT spinlock of a CPU and prevents preemption except for interrupts—used for operations that change the network stack, such as streams. Each of the different mutexes has two types of interfaces for locking, unlocking, etc.

**RWLock:** Read threads can share access to resources while Write threads require exclusive access.

**Events:** A mechanism that allows threads to await completion of an action, and upon completion, receive a Boolean status of the action if successful or not.

**Condition Variables:** Allows for certain threads to suspend themselves while waiting for a condition to be met, using different signals to awaken the thread, etc.

**DART THREAD TYPES:**

Single RT Thread, highest priority, executes services routines from I/O interrupts

Single IDLE Thread per processor, runs when no other threads run

GP Threads are used by applications

**STHREAD:** (7) types of threads

Yield, Join, Suicide, Detach, Wait, Signal, Broadcast threads

**SCHEDULING BASICS:**

**Soft:** These threads can run on any CPU

**Hard:** These threads run only on the CPU upon which it was created

**MP SPECIFIC:** [Multi Processor operations]

**Atomic Operations:**

**Parallelization & Optimization Guidelines:** using hashed linked lists and dqueue

**Processor Context:** one context per processor in global memory

**TIME SERVICES:**

Timers and User-specified callbacks

**NETWORK APPLICATION:**

Network applications contain a Stream module and GP module. A ‘Collector’ takes a stream head and directs to a GP thread. A ‘collector’ has a queue of waiting threads and stream heads to manage, used typically for TCP protocol & NFS.

**DQUEUE:**

Macros that define generic doubly linked lists

**PROJECT CLASS:**

Allows groups of threads to collaborate on a project.

**DART KERNEL MEMORY MANAGEMENT:**

**Intro:** DART kernel manages memory in a multi-tiered approach via (4) levels, from lowest level to highest

1. Page Level Allocator (PLA): This tier manages the allocation and freeing of virtually contiguous ranges of pages, maintains ownership list of allocated pages.

2. malloc() /free(): Takes memory requests and obtains free blocks from PLA to build blocks or pools of memory to use

3. new() / delete(): calls malloc() and then the constructor of the object or calls the object destructor and then free()

4. Mempool Allocators: Provides for more efficient repeated allocations and deletions of fixed size objects. In otherwords, is designed to better manage the shrinking and growing of the memory pool requirements for system operation

→DART uses a 32-bit virtual memory implementation based on 4GB logical and physical address space (no paging from memory to disk)

→Page sizes vary from 4KB up to 2MB in size

→Memory addresses are 32-bit in size

→DART boot.cfg lives in the DOS physical memory space, which is the first 640KB, and is loaded at 0x500

→DART image used Gload to load into memory starting at physical address 1MB

## **CELERRA FILESYSTEMS: BEST PRACTICES**

### **SYMMETRIX vs. CELERRA STRIPING**

**SYMMETRIX STRIPING RULES:** 2, 15, 64, or 512 cylinder stripes

Symm striping not normally recommended because uses sequential write operations to a single disk. For example, if you had a 2 cyl stripe, Celerra would have to perform (140) 8k writes before going to next vol.--uses only a single controller. Results=Huge performance drop.

**Exception:** Best MPFS performance use 2 cylinder striping--pre NAS 4.0? (another source says to use Celerra striped at 128 or 256k)

**CELERRA STRIPING RULES:** 8, 16, 24, 32, 128, 512k, etc.

Use Stripe Size 32768 for older NAS Codes across (5) symm devices. NAS 4.2 sets stripe depth to 128 or 512k.

Example, for 32k Celerra Stripe, data would be written to single spindles in (4) writes. Striping can distribute workload across multiple disks simultaneously. Striping reduces cache-hit rate & increases Queue handling. MPFS recommendation is to use 128 or 256k Celerra Striping. NAS 5.5.21.0 to standardize new AVM profile striping of 32768 for clar\_r5\_performance, clar\_r5\_economy, clar\_r1.

**IO QUEUE DEPTH:**

**CLARIION RULES:**

Clariion queue depth is 12 IO's per LUN

**CELERRA QUEUE DEPTH:**

64KiB for each dVol, which translates into 8 I/Os per dVol. Each dVol maps to one CLARiiON LUN.

## **FILE SYSTEM EXTENSION ISSUES:**

**Common Problems:**

1. Extend volume issued, not completed by DART, but CS db updated
2. Extend file system issued, not completed by DART, but CS db updated
3. Multiple extension attempts can leave DART memory and CS db in mis-matched condition

**WHAT OCCURS DURING FILE SYSTEM EXTENSION:**

1. CS database is updated first in the /nas/volume directory when nas\_fs -xtend is run
2. CS sends DART a command to extend. DART creates volume structure in memory, then creates new volume [command may fail if all camdisk paths not available or there is an incorrect volume reference]
3. CS sends DART file system extension command to create cylinder groups on new volume and adjusts Superblock info. DART unmounts & remounts the extended fs.
4. CS updates /nas/server/slot\_x/volume file and /nas/dos/slot\_x/boot.cfg file

**TROUBLESHOOTING STEPS:**

→Check cmd\_log.err for failure msg and size of extension attempt

→Check server\_log to see if volume and file has been extended

→Use following commands to piece file system pieces together and to determine where extension failed

→If Replication is involved, the recovery may be more difficult

→Compare nas\_fs -i and nas\_fs -s to see if volumes and fs sizes are same—if not, then NAS\_DB was updated, but actual extend did not complete

→Find filesystem entry for file system, and corresponding volumes file entry—check last field of volumes entry to see what the last volume was that was used for an extension

→Use .server\_config “volume” commands to see if DM memory was updated with fs extension activity

# nas\_fs -i <fs\_name>

# nas\_fs -E <fs\_name>

# nas\_volume -i

# .server\_config server\_2 -v "volume info 581"

# .server\_config server\_2 -v "volume DBbuild 581"

# .server\_config server\_2 -v "volume tree 581"

# .server\_config server\_2 -v "volume usertree 581"

**Note:** File system extension failures still plague both CLI & Celerra Manager. Especially common are file system extensions to replicating file systems, where the target is extended first, but source fails. Do not use ctrl + Z, ctrl + c, resize SSH or Telnet windows while performing an extension.

#### FS EXTENSION PIECES:

Control Station NAS\_DB

Data Mover “volume extend” and “file extend” in Server Log [i.e., DM memory]

**Note:** It’s possible for either CS NAS\_DB or Data Mover to become inconsistent if, at any point, the extension fails

#### EXAMPLE OF FILE SYSTEM INFORMATION BEFORE 1<sup>st</sup> EXTENSION:

# nas\_fs -I | grep fs20 →From this, we see that File System fs20 is built on volume 88

```
id    inuse type acl  volume   name      server  
23     y     1  0    88      fs20
```

# nas\_fs -i id=23 →We see the volume number, v88, and disks used, d8, by the file system

```
id    = 23  
name  = fs20  
acl    = 0  
in_use = True  
type   = uxf  
worm   = off  
volume = v88  
pool   = clar_r5_performance  
member_of = root_avm_fs_group_3  
rw_servers= server_2  
ro_servers=  
rw_vdms =  
ro_vdms =  
auto_ext = no,virtual_provision=no  
stor_devs = APM00063303725-0011  
disks   = d8  
disk=d8 stor_dev=APM00063303725-0011 addr=c16t111 server=server_2  
disk=d8 stor_dev=APM00063303725-0011 addr=c0t111 server=server_2
```

# nas\_fs -E fs20 →Shows hierarchy of what file system is built upon; meta v88, which is built on slice s55, metav85, & disk d8  
0:befs:23

```
23:fs20::0:1::y:1:88:1:::0::24::0:0:  
88:v88::0:4::y:3:2:23:87:  
87:s55::0:4::y:1:1:88:55:  
55:s55::0:0::y:85:87:0:2000:4096000:  
85:v85::0:4::y:3:0:55:84:  
84:d8::0:5::y:4:1:85:8:  
8:d8::0:3::y:466747:APM00063303725:1,2:84:0011:7:
```

# nas\_fs -s fs20 →Original size of file system in MB and Blockcount

total = 1968 avail = 1968 used = 0 ( 0% ) (sizes in MB) ( blockcount = 4096000 )

volume: total = 2000 (sizes in MB) ( blockcount = 4096000 )

# server\_df server\_2

server\_2 :

Filesystem	kbytes	used	avail	capacity	Mounted on
------------	--------	------	-------	----------	------------

fs20 2015984 600 2015384 0% /fs20

**# nas\_disk -l**

id	inuse	sizeMB	storageID	devID	type	name	servers
8	y	466747	APM00063303725-0011	CLSTD	d8		1,2

**# grep 23: filesystems**

```
23:fs20::0:1::y:1:88:1::::0::24::0:0:  
24:root_avm_fs_group_3::0:0::y:101:0::::0::23:3:  
23:fs20::0:1::y:1:88:1::::0::24::0:0:  
24:root_avm_fs_group_3::0:1::y:101:0::::0::23,25:3:
```

**# cat volumes**

```
85:v85::0:4::y:3:0:55:84:  
86:root_avm_vol_group_3::0:4::y:5:0::85:3:  
87:s55::0:4::y:1:1:88:55:  
88:v88::0:4::y:3:2:23:87:
```

**# cat slices**

55:s55::0:0::y:85:87:0:2000:4096000:

**# .server\_config server\_2 -v "volume info 88"**

```
# Component Volumes:.....0x001  
Component Volume List:....87
```

**# .server\_config server\_2 -v "volume DBbuild 88"** →Reverse hierarchical short listing of volumes

```
volume disk 84 c0t111 disk_id=8  
volume disk 84 c16t111 disk_id=8  
volume hyper 85 1 84  
volume slice 87 0 4096000 85  
volume hyper 88 1 87
```

**# .server\_config server\_2 -v "volume tree 88"**

→Shows hierarchical long list of volumes that make up meta 88

```
**** Hyper Volume 88 : 0xe12afd04 Information: ****
```

```
Total References:.....0x0002  
Total Blocks:.....0x3e8000  
Bytes Per Block:.....0x0200  
# Component Volumes:.....0x001  
Component Volume List:....87
```

```
**** Slice Volume 87 : 0xe12aff04 Information: ****
```

```
Start Block Offset: 0  
Total References:.....0x0003  
Total Blocks:.....0x3e8000  
Bytes Per Block:.....0x0200  
# Component Volumes:.....0x001
```

```
Component Volume List:....85 →Meta for slice 87
```

```
**** Hyper Volume 85 : 0xe12afd84 Information: ****
```

```
Total References:.....0x0004  
Total Blocks:.....0x38f9da1c  
Bytes Per Block:.....0x0200  
# Component Volumes:.....0x001
```

```
Component Volume List:....84 →D Volume as listed in output of nas_volume (=d8)
```

```
**** Basic Volume 84 : 0x6412804 Information: ****
```

```
Total References:.....0x0005  
Total Blocks:.....0x38f9da1c  
Bytes Per Block:.....0x0200
```

**# .server\_config server\_2 -v "volume usertree 88"**

1181139394: STORAGE: 4: ---- Displaying users of volume 88, referencecount=2 ----

1181139394: STORAGE: 4: UserOfVol=88, UserKind=IO\_OBJECT, userPointer=0xcc1e1744, start=0, nBlocks=4096000, referenceCount=1

**EXAMPLE OF FILE SYSTEM INFORMATION AFTER 1<sup>st</sup> EXTENSION:****# nas\_fs -xtend fs20 size=10G**
**# tail /nas/log/cmd\_log** →When troubleshooting extension issues, check cmd\_log.err to see what extension size was attempted

2007-06-06 10:15:27.067 db:0:7446:S: nas\_fs -xtend fs20 size=10G

2007-06-06 10:15:30.011 db:0:7446:E: nas\_fs -xtend fs20 size=10G

**# server\_df server\_2**

```
server_2 :
Filesystem      kbytes      used      avail capacity Mounted on
fs20          12341744       600    12341144   0%   /fs20
```

**# nas\_volume -l**

```
95      y  3  0  v95           0  58
96      y  1  0  s58           1  97
97      y  3  0  v97           1  88
```

**# nas\_fs -i fs20** → We see that v88 is unchanged, but that d7 has been added to disk list

```
volume  = v88
disks  = d8,d7
```

**Note:** /nas/volume/filesys file should be unchanged

**# cat slices** → New slice created to build meta 95, which is used for s58 & v97 (new meta used for the concatenation to volume 88)  
58:s58::0:0::y:95:96:0:10240:20971520:

**# nas\_fs -E fs20**

```
0:befs:23
23:fs20::0:1::y:1:88:1:::0::24::0:0:
88:v88::0:5::y:3:2:23:87,97: → Meta 88 shows the extended meta 97, separated by comma [d7, v95, s58, v97, concatenated to 88]
97:v97::0:0::y:3:1:88:96:
96:s58::0:0::y:1:1:97:58:
58:s58::0:0::y:95:96:0:10240:20971520:
95:v95::0:0::y:3:0:58:83:
83:d7::0:5::y:4:1:95:7:
7:d7::0:3::y:466747:APM00063303725:1,2:83:0010:7:
87:s55::0:4::y:1:1:88:55:
55:s55::0:0::y:85:87:0:2000:4096000:
85:v85::0:5::y:3:0:55,59:84:
84:d8::0:5::y:4:1:85:8:
8:d8::0:3::y:466747:APM00063303725:1,2:84:0011:7:
```

**# server\_log server\_2 -a -s |grep -i extend**

```
2007-06-06 10:17:30: ADMIN: 4: Command succeeded: volume extend 88 97
2007-06-06 10:17:30: ADMIN: 4: Command succeeded: file extend uxf 88=23 s=25067520
```

**# .server\_config server\_2 -v "volume info 88"**

```
# Component Volumes:.....0x002
Component Volume List:....87 97
```

**# .server\_config server\_2 -v "volume info 97"**

```
**** Hyper Volume 97 : 0xe0e19484 Information: ****
Total References:.....0x0003
Total Blocks:.....0x1400000
Bytes Per Block:.....0x0200
# Component Volumes:.....0x001
Component Volume List:....96
```

**# .server\_config server\_2 -v "volume DBbuild 97"**

```
volume disk 83 c0t110 disk_id=7
volume disk 83 c16t110 disk_id=7
volume hyper 95 1 83
volume slice 96 0 20971520 95
volume hyper 97 1 96
```

**# .server\_config server\_2 -v "volume tree 97"**

```
**** Hyper Volume 97 : 0xe0e19484 Information: ****
Total References:.....0x0003
Total Blocks:.....0x1400000
Bytes Per Block:.....0x0200
# Component Volumes:.....0x001
Component Volume List:....96
**** Slice Volume 96 : 0xe0e19404 Information: ****
Start Block Offset: 0
Total References:.....0x0004
```

Total Blocks:.....0x1400000  
Bytes Per Block:.....0x0200  
# Component Volumes:.....0x001  
Component Volume List:....95  
\*\*\*\* Hyper Volume 95 : 0xe0e19304 Information: \*\*\*\*

Total References:.....0x0005  
Total Blocks:.....0x38f9da1c  
Bytes Per Block:.....0x0200  
# Component Volumes:.....0x001  
Component Volume List:....83  
\*\*\*\* Basic Volume 83 : 0x6412c04 Information: \*\*\*\*

Total References:.....0x0006  
Total Blocks:.....0x38f9da1c  
Bytes Per Block:.....0x0200

# .server\_config server\_2 -v "volume usertree 97"  
1181148633: STORAGE: 4: ---- Displaying users of volume 97, referencecount=3 ----

1181148633: STORAGE: 4: UserOfVol=97, UserKind=LVOLUME, userPointer=0xcc1e1894,start=0, nBlocks=20971520, referenceCount=2  
\*\*\*\* Hyper Volume 88 : 0xe12af04 Information: \*\*\*\*

Total References:.....0x0002  
Total Blocks:.....0x17e8000  
Bytes Per Block:.....0x0200  
# Component Volumes:.....0x002

**Component Volume List:....87 97**

1181148633: STORAGE: 4: ---- User list complete for volume 97 ----

1181148633: STORAGE: 4: ---- Displaying users of volume 88, referencecount=2 ----

1181148633: STORAGE: 4: UserOfVol=88, UserKind=IO\_OBJECT, userPointer=0xcc1e1744 , start=0, nBlocks=25067520, referenceCount=1

1181148633: STORAGE: 4: ---- User list complete for volume 88 ----

1181148633: STORAGE: 4: ----- Parents Complete for 88 volume -----

1181148633: STORAGE: 4: ----- Parents Complete for 97 volume -----

## **LARGE DIRECTORY & LARGE NUMBERS OF FILES HASHING PARAMETER TUNING:**

Following params are used in situations where customers have very large directories and may need to use these tunable params:

**param ufs dirBlkHashSize=64** [Directory Block size hash param]

**param ufs inoBlkHashSize=6007**

**param ufs inoHashTableSize=15013**

**Note:** Please do not implement above params without consulting TS2/Eng. Above params are meant to be applied to specific sites that might be experienced high CPU activity due to large directory hashing or to large numbers of files within a file system.

See Primus' emc108564 & emc109199 for CPU idle problems with large directory searches and share enumeration algorithms.

**param ufs scanLimit=32**

**Note:** Param to allow cache hash table to be increased to support large directories and files

**param cifs pathCache=0** [To avoid caching of Path]

**Note:** Windows Operating Systems use B-Tree algorithm, which operates differently than bucket hashing of chunks of data pathCache is used to retrieve full pathname of opened file from FID and it must be prime number

## **COMMENT ON VERY LARGE DIRECTORIES:**

Very large directories, as shown below, can create performance issues due to the CPU usage required to parse. Unfortunately, even after deleting the files and subdirectories that create the very large directory, the directory space is not reclaimed—once expanded, the size is fixed until the actual directory is deleted. So, the fix for such a condition would be to actually migrate any relevant files/folders from the very large directory and then delete the original directory.

```
# ls -l  
total 608  
drwxr-xr-x  2 root  root  603136 Dec 13 10:32 largedirectory  
drwxr-xr-x  2 root  root   80 Dec 13 10:32 largedirectory1
```

## **CELLERRA INODE LIMITATIONS:**

Superblocks impose a ceiling of 257 million inodes for any single file system, meaning that a 2TB vs. 4TB file system would have approximately the same number of total inodes available. This inode limit can be adjusted with the following param for any file system >than 2TB in size:

param ufs inodelimit=<hex value up to 4 billion>

**ufs.inodelimit**      **0x0104402c**    **0x0f600000**    **0x0f600000**

### **FILESYSTEM INODE BEHAVIOR:**

Uxfs filesystem flushes inode entries for each filesystem every 30 seconds, similar to other Unix O/S, via use of sync threads.

UFS: 3: deleteEntry(): wait for sync thread to flush blkno 18147617

UFS: 3: flushentries(): wait for ufsDirty thread to flush blkno 2062525

**Note:** Above Log message is benign and indicates condition where cached disk block and ondisk block are not in sync, a normal condition of conflict between inode flush and customer action in that Uxfs prioritizes changes to inode blocks in these situations.

### **CELERRA FILE SYSTEM TYPES/VOLUME TYPES:**

1=uxfs (default Celerra Universal Extended File System type) (or slice for volume type)

2=stripe volume

3=meta volume

4=sfs (secondary file system)

5=rawfs (unformatted file system, such as ufslog; also a SFS in IP Replication)

6=mirrdfs (mirrored file system, as in BCV Snapshots)

7=ckpt (Checkpoint file system, as in SnapSure)

8=mgfs (Migration file system, as in CDMS migrations)

11=vpfs (volume pool file system, as in Checkpoint & IP replication)

100=group file system

101= root\_avm\_fs\_group

102=nmf file system

**Note:** NAS 5.1.9.4—during Farcopy/Nearcopy refresh of an imported filesystem, will show an “inuse” value of “f”, meaning ‘frozen’ and not mountable

### **CELERRA UFS LOG (UFS Transaction Log):**

The UFSLOG supplements the on-disk file system by recording metadata changes to file systems as changes occur. DART stores this log on a separate disk for all data movers. Metadata changes are first written synchronously to the log, with periodic cache flushes committing the log to the on-disk structures, every 30 seconds. After a system crash, while the on-disk structures may contain stale data, the ufslog has a record of all completed operations and is played back in order to recover file system consistency. In cases where a HWM is hit on the log, I/O to the backend is suspended while data is flushed to on-disk storage. Celerra UFSlog writes metadata to disk in async manner, and to log in sequential fashion. One physical ufslog device is used by all Servers, making all Servers able to read the log. UFSLogs reside on LUN 1, with (16) total UFSLogs at 64MB in size each.

### **UFSLOG STRIPING:**

A method for striping the log across a number of backend volumes to improve write performance for heavily used systems. Log is normally contained on a single Volume and divided into different areas for DM work. Striping is only done after Symmtop data and logs are reviewed between TS2 & EE. New stripe depth for the log is 32k.

### **UFS LOG LOCATION (LUN 1):**

**→UFSLOG is not mapped to the Control Station**

**→UFSLOG is located on LUN 1 for all Data Movers**

**→UFSLOG size is 64MB since NAS 5.1**

**# /nas/sbin/rootnas\_disk -i root\_ldisk**

```
id      = 2
name    = root_ldisk
acl     = 0
in_use  = True
size (MB) = 11499
type    = STD
symm_id = 000284701110
symm_dev = 0003      →Physical Symm device is 03
volume_name = root_ldisk
servers  = server_2,server_3,server_4,server_5,server_6
server = server_2      addr=c2t01      FA=14A  FAport=0 [Highlighted portion shows as LUN 1]
server = server_2      addr=c18t011  FA=03A  FAport=0
```

**# nas\_disk -l**

```

id inuse sizeMB storageID-devID type name servers
1 y 11499 000284701110-0002 STD root_disk 1,2,3,4,5
2 y 11499 000284701110-0003 STD root_ldisk 1,2,3,4,5

```

**UFS LOG COMPONENTS :**

Entry Header : UFS\_LogEntryHdr

Transaction Header: UFS\_LogTransHdr

Data Area

**UFS LOG OPERATION :**

The ufs transaction log is a fixed-size circular log file divided into equal segments that consists of a “header” and a “tail”. Each log entry reflects a change to the file system that are not yet written to disk storage, and are identified by a “record number”, starting from 0 to highest numbered record. The record number identifies the position of the record in the log. Tail entries reflect the highest record numbers, with the head reflecting the lowest numbers. Each Server has a single write synchronous UxFS log for all file systems.

Unlike journaled file systems, data is not written to this log.

**HOW UFS LOG MAINTAINS FILE SYSTEM CONSISTENCY UPON REBOOTS OR PANICS:**

Provided file systems are not marked dirty, after system reboots, the usflog records are replayed for all entries between the Head & Tail, restoring the state of the file system for all transactions prior to the reboot event. Metadata logging writes first, synchronously, to the log, and upon “sync” operations are flushed from cache to disk. Each NFS request creates log entries. UFSlog maintains a list of all completed operations.

**EXAMPLES OF PARAMETER CHANGES THAT MAY HELP UFSLOG FLUSH ISSUES—NAS 5.6.45:**

```

param ufs nFlushDir=64
param ufs nFlushIno=64
param ufs nFlushCyl=32
param ufs dirtyListSz=2048
param ufs syncInterval=15000
param ufs nFlushCyl=32 (default with 5.6.48)
param ufs logThreshold=91 (default with 5.6.48)

```

**TROUBLESHOOTING UFSLOG ISSUES:****\$ .server\_config server\_2 -v "printstats fslog"**

UFS Log Statistics:

```

26589959 duplicate objects in log records
log size -1636157440, nrecs 27738608
0 active descriptors, max active 46
11485 active logholds, max active 2169446
active log size 5429 sectors, peak 129772 sectors
times blocked for log flush: 45
Staging: 27738972 requests, 26851228 immediate, 823725 delayed, 64019 saved
Staging efficiency: 0%

```

Transaction Counts:

```

UPD_inactive 2398452
UPD_setattr 2631689
UPD_create 2521863
UPD_remove 2485790
UPD_rename 7738539
UPD_mkdir 273
UPD_rmdir 274
UPD_setlength 6183
UPD_blockwrite 9936788
UPD_trunc 19122
UPD_createContainer 62

```

**\$ .server\_config server\_2 -v "printstats ufs partial"**

\*\*\*\*\* SUMMARY UFS STATS \*\*\*\*\*

FindNode stats:

```

99537 findNodeCalls 330 findNewNodeCalls
6 nodeAllocations 0 hashChangeRetries
6 newNodeInserted 99204 validNode
0 deadNode 0 hashRemoves
0 nodeReadFailed
findNode hash table: calls 99207, entries searched 99453, ave 1

```

Indirect block cache stats:

```

3606 indirectRead 0 indirectWrite
3607 indirectReadHit 0 indirectInsert
346 invalidate
105 indirectTruncRead 0 indirectTruncWrite

```

FS operations:

Operation	Count	Time (usec)	Ave
COUNT[ 15] readinode	6	23	3

```
COUNT[ 15] writeInode      4493     21900      4
COUNT[ 15] rewriteSuperblock   122     286984    2352
COUNT[ 15] readBlock        78     903331    11581
COUNT[ 15] readHoleBlock     57      792       13
COUNT[ 15] getBlock         122     224       1
COUNT[ 15] writeBlock       708     1564158   2209
COUNT[ 15] forceWrites     122     1198549   9824
COUNT[ 15] allocZeroBuffer  158     1405       8
```

Vnode operations:

Operation	Count	Time (usec)	Ave
COUNT[ 13] lookupComp	3008	61332	20
COUNT[ 13] getHandle	62948	13888	0
COUNT[ 13] getAttr	126659	42486	0
COUNT[ 13] setattr	348	110308	316
COUNT[ 13] create	330	196387	595
COUNT[ 13] remove	320	114741	358
COUNT[ 13] rename	590	261836	443
COUNT[ 13] blockRead	118	794393	6732
COUNT[ 13] blockWrite	12376	4042783	326
COUNT[ 13] blockChainWrite	1339	4044695	3020
COUNT[ 13] setSD	330	95	0

UFS function stats

No module for id 15

UFS dirops stats

No module for id 1

UFS dirty blk stats

No module for id 17

UFS io stats

No module for id 3

UFS Directory Hash: 125163 hashqs (1501956 bytes), 2428 hashInfos (38848 bytes),  
3889128 bytes hashSpace

23423 entries searched

UFS Dirty Blk Cache Stats

nAccess 25261, nScans 25160, ave 0

nHits 18315, hit ratio 72

nLateHits 0, nStaleHits 16, nEntries 5920

inocache: new 448, del 5, tot 443, dirty 102

inoblk hash table: calls 4499, entries searched: 4498, ave 0

getPage 0 calls for 0 pages

freePage 0 calls for 0 pages

pageCount: current 3346, max 3354

## **WHILE TRUE SCRIPT TO SEE UFSLOG SIZES ON CELERRA:**

**# for i in 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16**

> do

> /nas/sbin/rootnas\_volume -s root\_ufslog\_\$i

> done

total = 1 avail = 0 used = 1 ( 100% ) (sizes in MB)

total = 1 avail = 0 used = 1 ( 100% ) (sizes in MB)

total = 1 avail = 0 used = 1 ( 100% ) (sizes in MB)

-----output abridged-----

**Note:** Ufslog is incorrectly sized at 1MB in above example and is full in all the slots indicated

## **CELERRA UXFS(Universal Extended File System) FILE SYSTEM:**

File Systems are created using mkfs utility into standard 8kb blocks. Max. file system size is 2TB and single file size 1TB. File system is divided into a number of 64MB Cylinder Groups (CGs) with bitmaps for inodes and blocks that are of equal size, except for the last one. Cylinder Groups have headers that describe its state, may span more than one disk block, but is predetermined in size at the time of file system creation. Cylinder Groups also spread metadata across entire storage space, rather than just a few cylinders. Header contains an object defined by “struct cg”, followed by free block and inode maps. Uxfs itself is a read-optimized implementation of the BSD Fast File System using write-ahead metadata logging [supporting online maintenance, large files, dynamic directory hashing] and delayed writes for better write performance (for combining I/O’s). UXFS uses indexed allocation where file information is stored within an inode using single, double, or triple indirect block indexing.

### **EXAMPLE:**

```
long cs_ndir; //number of directories
long cs_nbfree; //number of free blocks
long cs_nifree; //number of free inodes
long cs_nffree; //number of free frags
```

**Note:** Superblock contains struct checksum of overall file system free block and inode information, which should equal the sum of the values of each individual Cylinder Group header.

**CELERRA FILE SYSTEM LAYOUT:****Superblock→Alternate Superblock→Cylinder Groups→Inodes→Data**

→**Superblock**—persistent and dynamic info about file system (copies are in each cylinder group), such as the number of blocks, data blocks, cylinder groups, cylinders, free inodes and free data blocks, replicated to all Cylinder Groups. Also contains disk offset for superblock, offset of CG block, start offset for Inode and Data blocks, number of cylinders/inodes per group, number blocks in fs, basic block size, number fragments per block, number of free inodes & free data blocks.

→**Cylinder Groups**—contains bitmap of free/inuse data blocks, free inodes in cylinder group, and bitmap of all inodes for the group—code uses struct cg to maintain information. Each Cylinder Group consists of a Header, bitmap for free/inuse data blocks, bitmap for free inodes, and map of all inodes for CG

→**Inodes**—Inodes are stored after the CG summary block, and are a contiguous section of blocks. Each inode uses 128 bytes, with one inode per data block. Inodes contain UID, GID, File Size in number of 512 byte blocks used, Timestamps, CIFS Flags, ACL ID, Direct and Indirect Blocks, as well as mode and links information. The use of Inodes in this manner for UxFS is called ‘Indexed’ allocation. Block Indexing of ‘data’ is done using 18 direct pointers for first 144k bytes, then indirect blocks to locate the rest of the file [single indirect, double, triple, etc.]

→**Directories**—DART supports regular and inline directories in “dirent structure” [“.” Inode number→”. .” Reclen ramelen→Dirent “name”]. Regular directories are in BSD style and are stored in 512 byte data blocks, same as for files. Inline directories use the inode for directory entries and are the default type of directory used—they do not use data blocks! When an inline directory can no longer fit all of its information into the inode, then it is migrated into a regular directory, using data blocks [dot and dot-dot entries are stored by inode number only]

**→File System limitations**

--max size of mount name—512 bytes, stored in superblock

--max size of file name—255 bytes

--max size of directory entry—512 bytes

--max PathLength—1024 bytes

--Data Blocks are 8k, with one inode created per 8k block at file system creation

**CELERRA FILE SYSTEM STRUCTURE:**

Block→SuperBlock→Alternate SuperBlock→Inode→Storage Block→Direct Block→Indirect Block→Cylinder

Group→Bitmap→Special Files→VFS→V-node→Free list→Replay Log→In-core inode→Volume Manager

**FILE SYSTEM LAYOUT:**

Label	Boot Block	Offset
Super Block		Alternate Super Block
Cylinder Group map		Cylinder Group Map
Inode Blocks		Inode Blocks
Storage Blocks		Storage Blocks
(Cylinder Group 0)		(Cylinder Groups 1-n)

**FILE SYSTEM BLOCKS:**

Basic unit of size for file system (8k Celerra) and is not related to disk block or sector size. Celerra does not allow for fragmented blocks. All blocks are sequential, as presented to DART by the Volume Manager.

**SUPERBLOCK:**

Located at fixed offset on file system disk partition, cylinder group 1, stores max data blocks & max number of files. Also stores dynamic summary information about file system blocks, fragments, inodes, directories, label, size of inode block list, number free blocks & inodes, free block list, free inode list. Most alternate superblocks do not hold dynamic data.

**ALTERNATE SUPERBLOCK:**

Copy of superblock information contained in each cylinder group, at different offsets for redundancy; only first alternate superblock has persistent & dynamic data.

**INODE:**

Index node exists for each file & directory in a filesystem. File type, info, link count, UID/GID, RWX permissions, size in bytes, data block addresses in direct & indirect blocks, A-M-C-time. Celerra inodes are on contiguous blocks starting after CG summary block, and contains 128 bytes. Each 8K block holds 64 inodes.

**CELERRA INODE:**

Type/mode      Link count

UID & permissions

File size in bytes

A, C, M-time

(18) 4-byte direct blocks (Direct Blocks are 8k in size with 4-byte addresses)

1<sup>st</sup> Indirect Block (1<sup>st</sup> Indirect Block size is 2048 \* 8k blocks)

2<sup>nd</sup> Indirect Block (2<sup>nd</sup> indirect block is 2048 \* 2048 \* 8k blocks)

3<sup>rd</sup> Indirect Block (files only)

GID & permissions

8K disk blocks used

Generation number

CIFS Flags

ACL ID

#### **FILE SYSTEM IN-CORE INODE:**

An in-memory version of file inode is maintained once file is opened. Contains all of regular inode information, plus pointers to inode free list, hash queues, lock status, etc.

#### **CYLINDER GROUPS:**

Each CG contains 8192 \* 8k blocks, summary block contains used inode bitmap, free storage block bitmap.

#### **SPECIAL FILES:**

Denoted by inode, used to store critical data such as ACLs, quotas, etc.

#### **DIRECTORY FILE:**

Regular file containing file and sub-directory names

#### **MPD DIRECTORY:**

4-byte inode number

DOS attributes

Reclen (record length)

DOS/Windows/Unix file names

#### **COMPAT DIRECTORY:**

Stores only UNIX names—DOS & Windows long names stored in Shadow file

#### **FILE SYSTEM LAYERS:**

Vfs/v-node layer between User and File System layer—allows system calls, such as mount, quotas, fsck, etc

### **BSD FILE SYSTEM FSCK:**

→ Most fs corruption occurs from improper shutdown or startup, and hardware failures

→ FSCK is a multi-pass program and there are conditions where fsck and/or phases must be repeated to complete the repairs

→ FSCK is automatic & non-interactive when encountering expected corruption issues, but interactive when requiring User action

→ Most common corruption is to Superblock Summary Information, which changes with any block or inode change

#### **Superblock:**

Inconsistencies checked for file system size, number of inodes, free block count, free inode count

#### **Cylinder Group Block Maps:**

Ensures that all blocks marked as free are not claimed by any files, and corrects free-inode count if required

#### **Inode Status:**

Verifies state of each inode, link count of each inode to directories, range of each block claimed by an inode, number data blocks claimed by an inode, checks validity of directory data blocks. Orphaned inodes have a stored link count that is non-zero, but an actual link count of zero, and are placed in lost+found by fsck.

#### **FS Connectivity:**

Directories are checked for proper links into the file system and are moved to lost+found if they are not linked

#### **Cylinder Groups:**

Checks Free blocks vs. used-inode maps

### **SMALL LIST OF TYPICAL FSCK ERRORS & MEANINGS:**

#### **Partially Truncated Inode I:**

Inode size shorter than number of blocks allocated, occurs during system crash. FSCK will truncate to correct size.

#### **Partially Allocated Inode I:**

Inode is neither allocated or unallocated—contents will be zeroed

#### **Incorrect Block Count I:**

Block count is x, but should be y. FSCK changes value to y

#### **Unallocated I:**

Directory or file entry points to unallocated inode. FSCK will remove the entry

#### **CG C Bad Magic Number:**

CG C has corrupt cylinder group map—not usually repairable

#### **Blk Missing in Bit Maps (Salvage):**

Cylinder group block map missing free blocks. FSCK will reconstruct the free block map.

#### **Free Blk Count Wrong in Superblock (Salvage):**

Superblock free block info wrong. FSCK will reconstruct free block info.

### **MAC DB:**

File system is divided into cylinder groups of 8k blocks, with primary and alternate superblocks. Mac\_db tools works in 16 sector chunks. Duplicate superblocks are located at block 2, or offset 32 from start of CG. CG summary is at offset 48, block 3.

#### **CG Magic Numbers:**

--Superblock = 0x011954

--Cylinder Group = 0x90255

--Inodes start at block 4 and end at block 127

--Directories are in ‘inline’ style and regular, with each entry as “inum” (4 bytes), “reclen” (2 bytes, “nameLEN (2 bytes), & name.

#### **STRUCT CG MAPS:**

**Blocks:** Map that records total free blocks per cylinder in each cylinder group

**Rotational Positions:** Method for keeping track of number of free blocks in each rotational position for each cylinder

**Used inodes:** Bitmap with one bit per inode in the Cylinder Group. When bit is set the inode is in use.

**Free Blocks:** Bitmap with one bit per inode in the CG. When bit is set the block is free.

#### **CYLINDER SUMMARIES IN PRACTICE:**

##### **File System Initialization:**

Cylinder summaries are recomputed only at the time of file system mounting, as is the superblock information. This is accomplished by scanning each bit in the used inode and free block maps, which can take time, depending on filesystem size. Also, at mount time, an object called “UFS\_CgBlkEntry” is initialized, and contains pointers to two copies of the Cylinder Groups—an active copy that reflects the current state of the CG, and a committed copy that reflects changes to CG once changes are committed to disk.

##### **File System Synchronization:**

The “real-time” mutex in the UFS\_CgBlkEntry objects controls access to the CGs. Normal CG operations keep the mutex locked, running in real time. CGs are loaded into memory at time of mounting to avoid race conditions that would otherwise occur if reading CG into memory and accessing it in real time.

##### **CG Active and Committed Maps:**

The use of two copies of the CGs requires an accurate transaction-logging filesystem, where metadata that is not committed to the log must not be written to disk. Active, in-memory changes to the CGs are the most important, as numerous operations depending on the accuracy of the current state of the CGs.

‘Allocation Operations’ are those operations that look for free blocks or inodes in the active map, marking them as in use. Once committed to the transaction log, the free block or inode is also marked “in use” in the committed map as well [next flush of the ufs log will write the committed map changes to disk so that only “committed changes” are written to disk].

‘Free Operations’ do not change either CG map until the transaction is committed to the ufslog, then updates the active & committed maps.

##### **Block Allocation:**

Blocks for files are allocated by “UFS\_Vnode::alloc()”, which in turn calls the “UFS\_FileSystem::hashalloc()” to search through the cylinder groups. Other calls are used as well.

##### **Inode Allocation:**

Inodes are allocated by “UFS\_FileSystem::ialloc(), which also uses the hashalloc() to search CGs. (1) inode is allocated for each 8k storage block upon file system creation. Inode density can be increased at creation time, but would only be useful if file system contains large number of symbolic links or zero-length files, directories.

#### **PRESENT CYLINDER GROUP FILE SYSTEM LIMITATIONS:**

--Mount times can be substantial due to the way that CG summaries are computed

--CGs are maintained in memory, which can be substantial [not to mention subject to fragmentation?]

--Cannot allocate or free ranges of blocks in a single operation, causing block chain writes and file deletes to be slow

--Many of the original algorithms written for optimal file system performance were designed for JBOD environments and not Symm!

#### **CELLERRA FILE SYSTEM MEMORY IMPLEMENTATION:**

Celerra memory buffers use 8192 bytes per page

#### **VNODE/VFS File System Layer:**

Contains number of references in the file system. File systems cannot be unmounted until reference counts go down to a default number. Also contains device ID, whether fs is RO, number of free blocks.

##### **Functions of VFS File System:**

--mounting and unmounting file system

--Pausing, Resuming, Freezing, Thawing file system for various operations, such as Replication, Checkpoints, Timefinder, etc.

--Reserving/Unreserving blocks for file system

--Reference Count increment/decrement

--Sync

--Quota support

--ACL related functions

--File System extends

### **Functions of VFS\_Vnode:**

**Purpose:** Maintains data needed for generic vnode. VFS\_Vnode performs operations on files [block read/writes, setattr/getattr, ACL ID, Quota, Block reservations, Directory & Symbolic link operations]

→ Pointers to File\_NamingNode—directory lookups, mkdir, symlink, open, close, data read/write, sync, ioctl, setfl, setattr, locking.  
File Naming Node maintains UnixName, NTName, & pointers to parent node

--Pointer to file system

--Reference count

--Vnode cache

--Vnode freelist

--Block size

--State

### **File Naming Node:**

Base class for the following four classes used in logical grouping of vnode operations such as directory lookups, mkdirs, symlinks, for Setup operations such as opens & closes, for Data Read/Write operations (R, W, Sync), Attribute control operations (ioctl, setfl, setattr), and Locking operations (rwlock, rwunlock, etc.).

File\_FileNode—object class conducting operations such as opens, reopens, locks, ioList, and buffer cache map (blktoBufMap)

File\_DirNode—creates and deletes files, directories, symbolic links, File renaming, Lookuppath, add to DNLC

File\_SymLinkNode—Provides for symbolic link operations

File\_RootDirNode—implements root directory operations

### **METADATA LOGGING**—ufslog—see above

### **DATA I/O OPERATIONS:**

**Buffer Cached Objects**—access to buffers are coordinated by immutable and mutable references (read-only, exclusive access, respectively). All buffers have a streams reference count that is maintained in “db\_ref”. The count number is the number of msgb’s that point to it. Buffers are accessed by DART, Disk, FS, and TCP Layers.

**CFS Layer**—each file contains a bufmap, with an object name of ForkBlk2BufMap, consisting of direct and indirect maps.

**UFS Layer**—Directory Blocks and Indirect Blocks—related to dirty buffer status.

## **CELERRA UFS FILE SYSTEM LAYER:**

### **UFS FileSystem:**

Directory Block Cache—dirty directory blocks and indirect blocks

Inode Cache

Quota Record Hash—hash table for each User quota records, and one hash table for Group quota records

GID 16-32bit Mapping

UFS Hash Table

Sync Threads

### **Cylinder Group Object:**

UFS\_CgBlkEntry object used for block and inode allocation routines for maintaining Cylinder Group information

### **UFS\_Vnode Object:**

Implements virtual functions and maintains inode number, quota descriptors, directory blocks and indirect block hash, inode flags

### **Inode Cache:**

Each UFS FS Object has an inode cache (InoCache), with pointers to hash queue of inodeblocks. Sync thread flushes dirty blocks to the inode cache.

### **UFS\_DirBlkCache:**

Links directory blocks and indirect blocks of file system.

### **Directory and Indirect Block Hashing:**

Hashed list of directory and indirect blocks of a file as UFS\_DirBlkEntry objects.

### **Directory Hashing:**

Uxfs file system creates a complete hash table of the directory in memory.

## **CELERRA JOURNALED FILESYSTEM LAYOUT:**

### **TWO MAIN TYPES OF INFORMATION STORED ON DISKS:**

1. Data Blocks [data blocks & fragments]

2. Metadata Data Structures—designed to speed access to Data Blocks [Superblock, Cylinder Groups, Information Tables, Inodes]

**Note:** Changes to files or directories require metadata updates in RAM. These updates are then flushed to disk to the journal log, or “ufslog”. After system crashes, the system will ‘replay’ the ufslog to bring the filesystem back to a certain consistent point.

**Boot Block:** Block that stores procedures used in booting an OS—left blank if file system not used for booting

**Superblock:** Contains filesystem information, such as size, status, label, size of logical block, date/time of last update, cylinder group size, number data blocks in cylinder group, summary data block, file system state [clean, stable, active]

**Cylinder Group Map:** Blocks marked free that are not used for inodes, indirect address blocks, or storage blocks. Also keeps track of fragments to keep disk fragmentation from occurring.

**Inode:** 128 bytes that contain file information [except name--stored in the directory inode]. Inodes are stored in cylinder information block and consist of file type [regular, directory, block/character special, symbolic link, pipe, socket]; mode of file; number of hard links; UID of Owner; GID ownership, number of bytes in file, date/time created, last access time, last modified time.

**Storage or Data Block:** All other space in filesystem used for data blocks, not meta data. Celerra data blocks are allocated in 8k blocks and 1k fragmentations.

**Directories:** 512 bytes

**FILESYSTEM SUMMARY ISSUE:** ‘Rolling Panics’ due to filesystem summary data being wrong

NAS Codes prior to 4.2.17.0 or 5.0.17.0 → Problems in calculating file system summary to cylinder group information

**param ufs verifySummary=1** → forces summary between cylinder groups and filesystem prior to mounting, instead of after

## **REASONS FOR JOURNALED FILE SYSTEM MAINTENANCE [File System Checking]:**

*Requirements for File System Checking*

**Symmetrix Connectivity/Storage:** Major Backend disruptions or problems could impact Host systems that are direct-attached, such as Celerra. Normal CS procedure after Server Panics, or evidence of Backend Connectivity issues, is to engage Symmetrix PSE Lab to checkout Symmetrix Backend.

**Fibre Channel Fabric:** Fabric issues have the potential to create connectivity issues during Registered State Change Notifications (RSCN), whether intentional or switch-related. Code versions prior to NAS 4.2, 5.0, & 5.1 contain numerous improvements to the way that State Change Notifications, and loss of Back-End connectivity, are handled by the Data Mover. Firmware revisions are also factors.

**Hosts and Applications:** Hosts and Applications have the potential to cause file system corruption and panics, especially if the application or Host has access to Celerra file systems when it should not have access. An example of this might be ECC, VolumeLogix, or some other 3rd Party software that has access to the Symmetrix Backend. It has been observed during certain conditions that Hosts and/or Software such as VolumeLogix had illegal access to volumes on the Symmetrix that should have been exclusive only to the Celerra. Also, some Host operations can create code panics.

**Code-Related Issues:** Outdated NAS code can also be a factor, as certain families are no longer actively patched. The general consensus is that all Operating Systems are subject to potential fs "corruption" and therefore have their own set of tools or utilities to ensure that File Systems are maintained in a "consistent" and usable state. DART file system is based on BSD Unix [Berkeley Fast File System], and relies on ufs journal logging, as well as file system checks & verifications, to ensure that file systems remain consistent. Unfortunately, there are trade-offs between file system "reliability" and "system performance". With Operating System code, changes are often committed to memory, and then later written to disk and the file system--due to performance considerations, these changes are not verified before being written to disk, which does introduce the potential, however small, that file system corruption or problems can occur under unique sets of circumstances. When these problems do occur, Vendors [including EMC] try to ensure that the problem is identified and quickly corrected.

**Previous Server Panics:** Previous panics can introduce the potential for fs damage, depending on the issue, and issues may lay dormant for some time until a certain part of the filesystem is accessed that contains a truncated inode, etc.

## **CELLERRA AND SCSI DISK STORAGE:**

--sector format is 512 bytes (Clariion uses 520 bytes, but when configured with Celerra, presents 512)

--LBA—Logical Block Addressing is used to store & retrieve info

--Celerra writes unique label to every data LUN it sees—can see this with inlines—primus emc68161

--DM supports up to 4096 LUNS down each HBA channel (16 chains \* 16 targets \* 16 LUNS = 4096)

--Celerra uses SCSI-3 standard

## **FA ASSIGNMENTS ON SYMMETRIX:**

Enginuity symm code from 5566-5670 limits the number of devices per SA or FA processor to 2048. The restriction applies to the SA/FA processor, not the port. Since many processors have multiple port assignments, the total must be used when calculating. If all ports were in use on a 4-port processor, then the total logical devices allowed (Celerra Luns) would be 512 per port. But, if only a single port were being used, the total number of Celerra Luns could be 2048—for two ports in use it would be 1024/port, etc. So, the generally stated Celerra limit of 512 Luns/port for DMX systems, is not strictly true.

--5266-5267 code can support 904 devices per processor

--5264-5265 can support only 554 devices per processor (Symm 8000 2-port FA would mean only 256 devices per port)

--5670 & 5671 supports 2048 devices per FA processor

--5771 supports 4096 devices per processor

DMX/DM2 supports 512 devices per FA port

DMX3 supports 2048 devices per FA port

## **CLARIION UNCORRECTABLE SECTOR, PARITY ERRORS:**

### **HOW CLARIION HANDLES MEDIA ERRORS:**

CLARiiON arrays report errors when data cannot be read from a drive, as does Symmetrix. Historically, Symmetrix systems have always reported 0x03/0x11 Errors when hard Read errors occur (Symm uncorrectables remap their bad sectors automatically, though Celerra may panic if trying to read an uncorrectable), and CLARiiON reported hard media errors as 04x00. With the advent of Flare 19 (and backported to Flare 16), CLARiiON will now report Hard Read (i.e., unrecoverable errors) Errors to the Host as scsi error 03x11. Additionally, Clariion arrays continuously run a background sniffer utility (aka Background Verify) that checks for soft errors and other inconsistencies, and tries to fix. When attempts to correct the media or data error fails, and if Flare cannot rebuild data from parity drives (e.g., double-faulted disk drive in same Raid Group), then a dual unrecoverable read failure occurs and the data becomes “uncorrectable”, or “invalidated”, and essentially “unrecoverable”, subject to a hard error to the accessing Host, which for Celerra means a panic. So, the Uncorrectable Error could already exist and be known as a bad sector by the CLARiiON well before a Host tries to perform a Read to the bad sector. Or, the media error could turn into a Hard Uncorrectable Error if the Host is trying to Read data and the CLARiiON cannot honor the READ, essentially marking the sector Invalid.

### **SCSCI CAM ERRORS REPORTED BY HOST:**

0x03 Medium Error  
0x11 Unrecovered Read Error  
cdb: 28 (Read operation)

### **Uncorrectable Errors:**

CLARiiON event codes logged by FLARE whenever data is unable to be read from a particular drive location, and subsequent attempts to reconstruct the data from other drives within the RG have failed.

### **Invalidated Errors:**

CLARiiON event codes are logged by FLARE when media errors become unrecoverable, the data is marked as being void of valid information. At this point, a hard error is returned to the Host and it becomes an Uncorrectable sector error. Both CLARiiON and NAS have developed tools that can identify the exact location and files affected by the Invalidated errors, and then write to the bad sector to “fix” or revalidate the sector—the data obviously is still gone, but the Hard Error on Reads for the given sectors are resolved. The CLARiiON BRT tool was released with FLARE 19 Patch 030. The traditional NAS tool is Volcopy/MapBlock, while Napa 3 introduces the Revector tool to help streamline the process, and GrandNapa provides further enhancements with the cse\_recover tool whereby the backend will actually do the sector repair work, making it much faster than the sector by sector read that the Celerra must perform.

### **Coherency Errors:**

Events on Clariion backends that are related to—  
--Hardware problems with drives  
--Frumon code problems with LCCs  
--Flare bug affecting ATA drives

### **HOW DATA MOVER HANDLES MEDIA ERRORS:**

For DART, when reading data from the backend, and receiving an 04x00 or 03x11 error from the CLARiiON (or 03x11 Error from Symmetrix), will panic with IO failure—can lead to Rolling panics if client keeps reading same sectors. From the Data Mover perspective, the hard backend Read errors are seen at the CAM layer and reported by the Volume layer, as opposed to the File System layer. Since the Errors are at the Volume level, FSCK by itself cannot repair, and hence the need for the VolCopy repair tool, or improved tools that Napa 3 and GrandNapa will provide.

## **DIAGNOSING UNCORRECTABLE SECTOR ERROR ISSUES ON CLARIION BACKENDS:**

### **Primary Symptom are Intermittent IO Failure Panics on Data Mover:**

IO failure despite all retries/failovers  
<![CDATA[Slot 2: 1147480411: I/O Error: c0t1112 Irp 0x70f66b04 CamStatus 0x84 ScsiStatus 0x02 Sense 0x03/0x11/0x00

### **Examples of sys log entries:**

/nas/log/sys\_log -->Clariion passes backend events to Celerra  
May 12 21:00:01 2006 NaviEventMonitor:4:2 Backend Event Number 0x840 Host OEM-XOO25IL9VL9 Storage Array  
APM00041700339 SPA Device  
Bus 1 Enclosure 1 Disk 11 SoftwareRev 6.19.0 (4.14) Unknown Error 2.19.0.701.5.027 Description Data Sector Invalidated  
/nas/log/sys\_log  
May 12 21:00:02 2006 NaviEventMonitor:3:3 Backend Event Number 0x953 Host OEM-XOO25IL9VL9 Storage Array  
APM00041700339 SPA Device  
Bus 1 Enclosure 1 Disk 8 SoftwareRev 6.19.0 (4.14) Unknown Error 2.19.0.701.5.027 Description Uncorrectable Parity Sector

### **Examples of Server Log entries:**

CamStatus 84 ScsiStatus 02 Sense 0400 00

### **Clariion SPCollect Logs:**

|                                    |  |                     |
|------------------------------------|--|---------------------|
| A 05/13/06 00:34:16 Bus1 Enc1 Dsk8 | 956 Parity Invalidated [vr_rd Raid]      | 0 21773bc0 12001000 |
| A 05/13/06 00:34:16 Bus1 Enc1 DskB | 957 Uncorrectable Sector [vr_rd Raid]    | 0 21773bc0 12001000 |
| A 05/13/06 00:34:16 Bus1 Enc1 DskB | 840 Data Sector Invalidated [vr_rd Raid] | 0 21773bc0 12001000 |

**Clariion Background Verify of affected LUN & Viewing results, as run from Control Station:**

**\$ /nas/sbin/navicli -h 192.168.1.200 setsniffer 28 -bv -bvtimer asap** (asap, high, medium, low)

**Note:** Initiates the Background Verify operation for LUN 28

**\$ /nas/sbin/navicli -h 192.168.1.200 getsniffer 28 | -alltot** (Verifies progress & last results of Background Verify)

VERIFY RESULTS FOR UNIT 28

Sniffing state: ENABLED

Sniffing rate (100 ms/IO): 10

Background verify priority: ASAP

Verify State: Background Verify Running

Percent Complete: 81 -->This setting returns to 0 once the Background Verify operation has completed  
Corrected Uncorrectable

Most Recently Completed Full Unit Verify

---

|                    | Corrected | Uncorrectable |   |
|--------------------|-----------|---------------|---|
| Checksum errors    | 0         | 103           | -->Background verify does 102 passes by default |
| Write Stamp errors | 0         | 0             |   |
| Time Stamp errors  | 0         | 0             |   |
| Shed Stamp errors  | 0         | 0             |   |
| Coherency errors   | 0         | 0             |   |

**OBSERVING LUN REBUILDS ON ATA DRIVE:**

**\$ /nasmed/sbin/navicli -h 192.168.1.200 getlun -prb | egrep -A1 "21\$|19\$" ; date**

LOGICAL UNIT NUMBER 19

Prct Rebuilt: 100

LOGICAL UNIT NUMBER 21

Prct Rebuilt: 78

Thu Jun 29 00:48:59 CDT 2006

**BASIC LUN REBUILD QUERY:**

# /nas/sbin/navicli -h 192.1.4.220 getlun -prb

LOGICAL UNIT NUMBER 3

Prct Rebuilt: 100

**VOLCOPY/FSCCK/MAPBLK--TRADITIONAL METHOD FOR HANDLING UNCORRECTABLE SECTOR ISSUES:**

**Note 1:** NAS 5.5.23.2 (Napa 3) introduces a new Revector tool that can be used in place of VolCopy tool for those situations where a Data Mover has panicked with an 0x03/0x11Error (Unrecovered Read Error)—VolCopy will still be required for all other situations where Revector does not apply.

**Note 2:** Unfortunately, a change has been made with NAS 5.5.23.2 and VOLCOPY can only be run against the Basic Volume ID—it cannot be used to run against the metavolume, which is the traditional method for using Volcopy.

**WORKAROUNDS FOR NAS 5.5.23.2:**

**VOLCOPY ON STANDBY SERVER—pre NAS 5.5.23.2:**

1. If Volcopy method is required, upload a pre-5.5.23.2 nas.exe and run the VolCopy on the Standby Server
2. Run FSCK/Builtin Mapblk on Production Server if running 5.5.23.2 or higher.
  - a.) Create “ME\_Hyper\_<metavolID>\_Repaired.txt” file on the rootfs of the production Server and add each Bad Block number, as found in the Server Log, to separate lines in this file.
  - b.) Start FSCK with indicator that DART should use the Repaired.txt file located in the rootfs to record the file information in the Report.txt file:

**\$ .server\_config server\_2 “file fsck uxfs 283=44 mapblock=1”**

**VOLCOPY ON PRODUCTION SERVER—pre NAS 5.5.23.2**

1. Create the empty Repaired.txt file on the rootfs of the Production Server:

# touch ME\_Hyper\_<metavolID>\_Repaired.txt

2. Set the ReportBadBlocks param

3. Run the Volcopy—it will populate the Bad Block information in the Repaired.txt file, rather than having to sort through Server Logs:

**# .server\_config server\_2 “volcopy src=283 threads=100 repair file= ME\_Hyper\_<metavolID>\_Repaired.txt”**

**Note:** Do not use the “file=” syntax with NAS 5.4, it will panic the Server

4. Edit results of this file to leave only Bad Block number on each line

256003932

512005200 (etc.)

5. To use the integrated FSCK/Mapblock feature, run the FSCK from the Production Server using NAS 5.5.23.2 or higher

**\$ .server\_config server\_2 “file fsck uxfs 283=44 mapblock=1”**

**1. Server panics with Uncorrectable Sector error.** may or may not successfully failover, and may or may not develop into a Rolling Panic—if an application, or NFS Client is out there and continually tries to Read the same bad sector each time the DM is back up, then the Server would panic again, etc.

**Panic Header:**

**DART panic message: >>PANIC in file: ./BVolumeIrp.cxx at line: 298 : IO failure despite all retries/failovers**

**Server Log:**

2006-09-04 08:47:32: STORAGE: 3: readBlock() : Volume 118 Bad Block Index ec84faf01157377652: UFS: 3: readBlock failed, err 16

2005-08-06 22:44:02: CAM: 3: I/O Error: c80t117 Irp 0x90e11084 CamStatus 0x84 ScsiStatus 0x02 Sense 0x04/0x00/0x00

2005-08-06 22:44:02: CAM: 3: camFlags 0x50 Addr 0x8d635304 Len 0x1c000

2005-08-06 22:44:02: CAM: 3: cdb: 28 00 00 c9 02 80 00 00 e0 00 00 00

**Integrated Celerra Log Messages:**

Bus 1 Enclosure 1 Disk 8 SoftwareRev 6.19.0 (4.14) Unknown Error 2.19.0.701.5.027 Description Uncorrectable Parity Sector

Bus 1 Enclosure 1 Disk 11 SoftwareRev 6.19.0 (4.14) Unknown Error 2.19.0.701.5.027 Description Data Sector Invalidated

**Gateway System Reports Errors to NaviEventMonitor Windows Client:**

Sep 1 11:00:06 2006 NaviEventMonitor:3:3 Backend Event Number 0x956 Host OEM-XO

O25IL9VL9 Storage Array APM00050800851 SPB Device Bus 1 Enclosure 1 Disk 7 Description Parity Invalidated

Sep 1 11:00:06 2006 NaviEventMonitor:3:3 Backend Event Number 0x957 Host OEM-XO

O25IL9VL9 Storage Array APM00050800851 SPB Device Bus 1 Enclosure 1 Disk 3 Description Uncorrectable Sector

**2. NAS Support stabilizes Data Mover** by identifying and unmounting the offending file system(s)

**3. NAS Support assesses health of backend & may determine cause to be an Uncorrectable Sector**—Panic Header leaves recognizable signature, Integrated Celerras log CLARiiON media errors in sys\_log; etc.

**4. NAS Support engages CLARiiON Support via Subcase to investigate backend**

a.) SP Collects are reviewed on each SP

**SPCollect Logs:**

A 05/13/06 00:34:16 Bus1 Enc1 Dsk8 956 Parity Invalidated [vr\_rd Raid] 0 21773bc0 12001000

A 05/13/06 00:34:16 Bus1 Enc1 DskB 957 Uncorrectable Sector [vr\_rd Raid] 0 21773bc0 12001000

A 05/13/06 00:34:16 Bus1 Enc1 DskB 840 Data Sector Invalidated [vr\_rd Raid] 0 21773bc0 12001000

b.) If LUNs are found with errors, Background Verify is run & then checked for number of errors (sniffer)

**\$ /nas/sbin/navicli -h 192.168.1.201 setsniffer 33 1 -bvtimer HIGH** →BV started on affected LUN 33

**\$ /nas/sbin/navicli -h 192.168.1.201 getsniffer 33**

VERIFY RESULTS FOR UNIT 33

Sniffing state: ENABLED

Sniffing rate (100 ms/IO): 5

Background verify priority: High

Historical Total of all Non-Volatile Recovery Verifies (0 passes)

---

Corrected Uncorrectable

Checksum errors 0 25

c.) Based on whatever else is happening with array, CLARiiON may recommend drive replacements, LCC Card replacements, etc., and once backend is stabilized, and the RG parity rebuilds have completed, will most likely recommend an Unbind and Rebind on the affected LUNs, and a Flare Upgrade to help resolve known Uncorrectables issues with ATA drives.

**5. NAS Support offers recommendation for corrective actions**—VolCopy, FSCK, and Mapblk repair method may be used as a way to correct the Backend Sector problems, restore the file system to a consistent state, and identify whatever files/directories have been affected (when possible).

**Note 1:** Generally speaking, the number of Uncorrectable Errors for any given case should be relatively small (under ten) if systems are running FLARE 19 Patch 030 and higher, while systems running older FLARE versions may have hundreds of Errors per affected LUN. ATA drives are particularly susceptible to the issues involved with Uncorrectables, though on rare occasions, Fibre Channel drives will also have Uncorrectables.

- a.) NAS Support identifies the file systems associated with the LUNs that CLARiiON has identified as having Uncorrectables
- b.) Permanently unmount and delete any associated Checkpoints
- c.) Abort any associated Replication Sessions with the affected production file system (? Not sure of this step)
- d.) Permanently unmount the production file system

**Note 2:** In cases where the bad sectors are verified as “unused”, or are only associated with a Checkpoint file system, the only corrective action that would need to take place would be to delete the checkpoint file system or leave the “unused” blocks alone, as the working assumption is that the “unused” bad blocks will always be written to next, as opposed to being Read from, thereby revalidating the data.

## **6. Perform VolCopy repair work using the Standby Data Mover.**

**Note:** If required, upload the latest nas.exe to perform the VolCopy and subsequent FSCK work, and reboot the Data Mover.

## **7. Set the following parameter to allow for the VolCopy tool to identify and then write to the bad sectors:**

**\$ .server\_config server\_3 “param volume allowReportBadBlock=1”**

**Note:** Setting this param is no longer required with NAS 5.5.23.x and higher

## **8. Associate Vol. Structure & CTL information to the Standby Data Mover using the virtual mount:**

**\$ server\_mount server\_3 -v fs1**

**9. Start the VolCopy repair process on appropriate metavolume** (finds bad sectors on indicated metavolume or dvolume, writes 0’s to revalidate the data, and records LBA block information in Server Log):

**\$ .server\_config server\_3 “volcopy src=<vol\_ID\_of\_fsf> threads=10 repair”**

server\_3 : commands processed: 0

Error 5: server\_3 : Input/output errorRPC: Timed out →Normal output, check Server Log to verify that VolCopy has started

**\$ .server\_config server\_3 “volcopy src=<vol\_ID\_of\_fsf> abort”** →Use to abort a Volcopy session 5.5.23+  
**Server Log:**

2006-09-02 00:16:41: STORAGE: 4: BEGIN:volcopy repair-Vol:118 srcStart:0 threads:10 xferSize:128 nBlks:778240000

**Note:** Save all output from Server Log and record each BadBlock number

## **10. Monitor VolCopy progress using CLI or Server Log:**

**\$ .server\_config server\_3 -v “volcopy display”**

\*\*\* Displaying VolCopy instances \*\*\*

VC 1: Volcopy src:251 dst: command:3 threads:10

numBlocks:32768000000 blocksCopied:30685721728 progress:93% interleaved:yes

### **Server Log:**

2006-09-02 15:49:44: STORAGE: 4: RepThread:VC118\_20 Volume 118 repair is in-progress(Status:5%)

2006-09-02 16:01:53: STORAGE: 4: RepThread:VC118\_20 Volume 118 repair is in-progress(Status:10%)

2006-09-02 16:14:07: STORAGE: 4: RepThread:VC118\_20 Volume 118 repair is in-progress(Status:15%)

2006-09-02 16:26:19: STORAGE: 4: RepThread:VC118\_20 Volume 118 repair is in-progress(Status:20%)

## **11. After completion of VolCopy repairs, reset parameter to 0:**

### **Server Log:**

2006-09-02 19:42:12: STORAGE: 4: END:volcopy repair-Vol:118 srcStart:0 threads: 1 xferSize:128 nBlks:1868605824

2006-09-02 19:42:12: STORAGE: 4: volcopy: Volume repair on 118 succeeded.

2006-09-02 19:42:12: ADMIN: 4: Command succeeded: volcopy src=118 repair

### **Example of Bad Block Found by VolCopy Process in Server Log:**

2006-09-02 00:18:22: STORAGE: 4: **VC\_118:6816047**

**\$ .server\_config server\_3 “param volume allowReportBadBlock=0”**

## **12. Using the Standby Server, run FSCK on the affected file system(s).**

**\$ .server\_config server\_3 "logsys set severity FSTOOLS=LOG\_DBG3"**

**Note:** Do not set LOG\_DBG3 for NAS Versions 5.4 or 5.5 as the proper FSCK logging is automatically done

**\$ .server\_config server\_3 "file fsck ufs 118=23"**

**Note:** Where 118 is Volume ID and 23 is FSID

## **13. Virtually unmount the file system(s) from the Standby Server when FSCK is completed:**

**\$ server\_umount server\_3 -v fs1**

## **14. Remount file system on the Production Server**

**\$ server\_mount server\_2 fs1 /fs1**

**Note:** At this point, the actual Volume and File System repair processes have been completed, and Mapblk is used to find and map affected sectors to actual file system inodes, etc.

## **15. From the Control Station, create mountpoint & mount the DM file system in preparation for MapBlock:**

**# mkdir /mapblk**

**# mount server\_2:/fs1 /mapblk** [Mounting Data Mover file system fs1 to /mapblk mountpoint on CS]

**Note:** May need to NFS export the file system to CS if the mount command fails due to lack of permissions

**# server\_export server\_2 -o anon=0 /fs1**

## **16. Use MapBlock Utility to map each logical bad block found in the Server Log, as seen by VolCopy process, to an actual inode owner (file system object), then run find command to map inode to file name.**

**Note:** Run MapBlock to map each bad block to an inode, & output to file. Run find or ls to map each found inode to a file name.

a.) Upload the mapblk utility from the GTSCentral/dms website [mapblk.gz] to the CS and unzip:

/home/nasadmin/mapblk

# ls -la

-rw----- 1 root root 332461 Oct 12 09:13 mapblk

b.) For older NAS Versions, the badBlock message that was generated by VolCopy (written to the Server Log) process was different than the current message [VC\_147:6816047], yet still requires division by 16 to obtain the true LBA block number to plug into the mapblk utility.

### **Old NAS BadBlock Message in Server Log:**

**Log Msg:** srccvol:282 badBlock: 623520562 size:512

**Divide:** # echo \$(( 623520562 / 16 ))

**Result:** 38970035.125 → Always round down to nearest whole number to get 8k block number for inode = 38970035

### **New NAS BadBlock Message in Server Log:**

2006-09-02 00:18:22: STORAGE: 4: VC\_118:6816047

**Divide:** # echo \$(( 6816047 / 16 ))

**Result:** 426002.9375 → Always round down to nearest whole number to get 8k block number for inode = 426002

**# /home/nasadmin/mapblk/mapblk server\_2 118 426002**

file system size 48640000

data block in inode 71641

**Note:** Optionally, run mapblk in the background and input to a file to capture the inode numbers

**# nohup /home/nasadmin/mapblk/mapblk server\_2 118 426002 >>fs1.mapblkinode.log &**

**Note:** Above command will output an inode number for the block and Volume ID indicated into the file. Repeat these steps until every badblock has an associated inode in the output file.

c.) Match the inode number to file names using the following process while in the /mapblk mountpoint to the production fs:

**# find . -ls > /tmp/inode\_to\_name\_list.log &**

**Note:** Alternatively, run the find in the background for instances where you are only finding one or two inodes

**# nohup find /mapblk -inum <inode\_no> >>fs1.inode\_name\_map.log &**

17. Advise customer of affected filenames for potential backup Restore

## **CELERRA VOLUME COPY REPAIR TOOL & PROCESS:**

There have been recent cases where Clariion backends are recording uncorrectable scsi sector errors. For Celerra, there are several types of blocks that can be affected by these errors: directory, metadata, & indirect blocks. For ‘data blocks’, a new utility called ‘Volume Copy’ has been bundled into a special nvolvopy.exe, and more recently, with updated versions of NAS 5.4 & 5.5.

The Volume Copy utility works at the volume level by basically writing 0’s to the bad sector, which serves to revalidate the sector, though the data itself is lost. After the Volcopy and fsck procedures are run, the file system can generally be remounted and data restored as necessary—volume copy does not understand file system structures. Since the uncorrectable sector errors are issued by the storage system at the volume layer, an FSCK alone will not correct the issue, since FSCK works only at the FS layer. An “I/O failure despite all retries” is an error that occurs between the CAM & Volume Layer, not the file system layer.

### **CAM→Volume Layer→File System Layer**

0x03/0x11 uncorrectable sector errors, as reported by Symmetrix (and now Clariion with Flare 19+)

04/00 soft media errors, as reported by Clariion storage systems

**Note:** The Volcopy Utility can only be run with NAS versions 5.3 or higher, on either Clariion or Symmetrix backends. Volcopy and FSCK are the only two parts of the toolset that correct the problem, with Volcopy finding and writing zeroes to the uncorrectable sectors, and FSCK putting the file system back into a consistent state afterwards. Mapblk is a reporting tool to determine inode file names from the block number listed in Server log. Not all identified bad blocks will have a filename because some bad blocks could be free space or metadata. Some additional concerns before using VolCopy would be to try and ID the file systems involved. If all file systems on an affected LUN can be taken offline, then run the Volcopy tool against the whole lun.

## **STANDARD CELERRA I/O RETRY MECHANISM:**

Each path has a timeout value of 40 secs for a Read or Write IO failure, up to a total of 240 secs max (4 minutes). Be aware that for IO Failure panics that involve NFS clients, NFS clients will automatically retry to Read or Write to the same file when the Data Mover comes up, in essence creating a Rolling Panic situation.

## **VALID NAS VERSIONS FOR VOLCOPY:**

5.3.25.0, 5.4.19.4 and above

## **SYMPTOMS & CAUSE:**

Data Mover may intermittently panic with "IO failure despite all retries/failovers" and log "CamStats 84 ScsiStatus 02 Sense 0400 00" errors. Basic cause would be a backend event such as,double-faulted Raid Group, power outage without battery backup. In reality,

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
many of the Uncorrectable issues emanate from the use of ATA drives on CLARiiON backends. Flare 19 Patch 027 was issued to correct a problem where writes could be successful, but later Reads of the data became media errors. ATA drives also could timeout during error handling and result in faulted drives, and uncorrectable errors. Patch 034 fixes sector remapping issue for RAID 3 if different size ATA drives are used within same RG, which resulted in uncorrectables in upper half of drives. Patch 030 fixes rebuild issue that could cause hard error, uncorrectable.

### **BASIC STEPS INVOLVED WITH VOLCOPY REPAIR:**

1. Provide SPCollects to Clariion and engage—they will need to run manual BV to determine # of uncorrectables—BV does resolve some issues. Often it appears that bad disk drives are the cause of the problem and may require unbind of luns, disk replacement, rebind of luns, and restore from Tape or other means.
2. Identify affected file systems, determine whether to use file system volume or disk volume method
3. Set param volume allowReportBadBlock=1 and reboot Data Mover to enable VolCopy Tool
4. Unmount affected file systems and delete any associated Checkpoints or IP Replication SavVols
5. Conduct virtual fs mount using -v
6. Run volcopy tool
7. Virtually unmount file system(s), turn off the allowReportBadBlock param, and reboot the Data Mover
8. FSCK the affected file systems
9. Run MapBlock to map LBA address from Server Log to an inode number

### **TYPICAL BAD LBA BLOCK BEING REPORTED BY VOLCOPY:**

STORAGE:4: VC\_90:12583135

### **BASIC VOLCOPY PROCEDURE--see Primus emc117301:**

#### **I. CORRECTING THE BAD SECTORS & RESTORING FILE SYSTEMS TO CONSISTENCY:**

**Prerequisites:** Always have Symmetrix or Clariion backend completely checked out prior to running this procedure, and then only use this procedure with the explicit approval of Escalation Engineering, as it requires a special nas.exe version. Volcopy can be run against file systems, as in the procedure below, or against the disk itself. Volcopy reads in 64k chunks, but once it encounters a sector error, will begin reading in 512 bytes, sector-by-sector within the 64k chunk, zeroing out bad sectors and logging the lba information message in the Server Log. The VC output in Server Log is the 512 byte sector within the file system that is corrupt, not an 8k block.

#### **RUNNING VOLCOPY AGAINST THE DISK VOLUME vs. FILE SYSTEM:**

1. Run steps 1-6 in the above Volcopy procedure
2. Record the chain target & lun information for each path to the disk volume in question from Server Log  
Command succeeded: volume disk 1515 c32t3l12 disk\_id=71
3. Virtually unmount the file system from the Standby server
4. Read in each path to the disk using the following:

**\$server\_config server\_2 "volume disk 1515 c32t3l12 disk\_id=71"** (repeat for each path)

5. Run volcopy against the disk volume

**\$server\_config server\_2 "volcopy src=1515 threads=10 repair"**

6. Reboot data mover after completion and run FSCK

#### **Some Server Log messages during VolCopy Repairs:**

2006-04-14 04:03:05: STORAGE: 4: BEGIN:volcopy repair-Vol:4996 srcStart:0 threads:10 xferSize:128 nBlks:6704695296

2006-04-14 05:41:19: STORAGE: 4: RepThread:VC4996\_0 Volume 4996 repair is in-progress(Status:5%)

### **UNCORRECTABLE SECTOR ERRORS GETWELL PLAN:**

Engineering is implementing a phased plan to handle the uncorrectable sector error issues. One of the major problems with Clariion arrays is that Media Errors were not being reported as 03/11 errors, as the Symm has always done, but no different from generic Hardware Errors 04/00. With advent of FLARE 19, arrays will be able to properly report Media errors as 03/11 errors.

#### **PHASE 0:**

Current manual process, whereby Server panics (rolling panic) on the Bad Sector errors, file system becomes unavailable until manual VolCopy recovery procedure is run. If file system involves Checkpoints or Replication, must delete or deactivate. FSCK only corrects for metadata issues, cannot repair data blocks.

#### **PHASE 1:**

Bad Sector media error of 03/11 will be reported if running FLARE 19. Celerra will still panic, initiate the volcopy repair during reboot, will fsck if bad sector is in the pfs, will invalidate SnapShots, will deactivate Replication Sessions, runs Fs\_mapObjects tool to find damaged objects, then reassign bad sectors. DM will not be available during this repair process, and will be in loaded state.

#### **PHASE 2:**

Similar to Phase 1 except that Data Mover will come to contacted state and service all other file systems while repair is being done in background.

#### **PHASE 3:**

Try to prevent panic. Will freeze file system, runs volcopy procedure, fsck's, unfreezes and re-exports as required.

### **PHASE I & PHASE I PRIME UNCORRECTABLE SECTOR CODE FIX PLAN:**

#### **Phase I Napa 3:**

Depends on Flare 19 patch 034 for 03x11 scsi sense key during DM Panic. Server reboots with affected fs unmounted (no more Rolling Panics), use revector tool to repair bad blocks as listed in sys\_log, fsck & mapblk automatically map inodes to file names for bad blocks, remount file system

### **\$ .server\_config server\_2 “revector start vol=xxx”**

After the revector tool is run, the file system can be mounted, at which time FSCK and MapBlock will be run, and upon completion, the file system will be mounted. In practice, though, run the nas\_fsck manually, which also invokes the mapblk tool and will ensure that the File System mount information will be consistent with the Control Station db. For Cognac release, the Data Mover will recover automatically, with the eventual plan being that the Data Mover will be able to log the errors & fix without panicking at all.

#### **Phase I Prime GrandNapa:**

Depends on Flare 19 patch 034 for 03x11 sense key error during DM panic. Uses new cse\_recover script on Control Station to invoke CLARiiON BRT tool to cleanup and report on bad sectors. See section under ‘GrandNapa’ for more details.

## **SEVERAL UNCORRECTABLE REPAIR METHODS:**

### **I. VolCopy, FSCK, MapBlock**

→DART driven repairwork, very slow & manual for each bad sector, dependent on CLARiiON Support’s intervention & analysis of SPCollects & BV

→Before NAS 5.5.23.2 (Napa 3) and Flare 19 Patch 030, the only Celerra method of recovery was VolCopy

### **II. Revector Tool—Phase I Napa 3 [5.5.23]**

→DART driven repairwork, very slow, but much manual work reduced, BV may still need to be redone and VolCopy method used to repair bad sectors that fall outside the file system metavolume associated with the Panic

→DART will panic, but recover with affected file system <unmounted>

### **III. Cse\_recover Tool—Phase I Prime GrandNapa [5.5.27.5]**

→SP driven repairwork, much manual work reduced, very fast, but BV may still be needed to address other Uncorrectables that fall outside the metavolume associated with the triggering Panic event

Keep in mind that at any given time, you may need to use a combination of the above tools on a system

#### **Example:**

Napa 3 & Flare 19 P034 are in place. Server panics, Bad Sector identified, Revector tool used. However, other Uncorrectable events may have affected other metavolumes and luns, especially across the same Raid Group—Run BV to determine if there are other Uncorrectables and then follow the traditional VolCopy repair method to fix

## **COGNAC 5.6.36.4 UNCORRECTABLE UPDATE:**

Due to an oversight, the revector tool and cse\_recover tools were never built into the Cognac family. AR114835 ports the cse\_recover tool into 5.6.37.

## **REVECTOR REPAIR METHODOLOGY FOR NAPA 3:**

### **CAVEATS:**

→Tool must be used with Napa 3 +, Flare 19 patch 034 +, an 03x11 SCSI error in Panic Header, and be run on the Production Server  
→Run manual FSCK using nas\_fsck—do not use server\_mount to invoke FSCK after revector repairwork

### **General Process that occurs:**

→Server panics, recovers gracefully with affected file system <unmounted>--No autofsck allowed in Phase I or Phase I Prime

→Associated Checkpoints & Repl. Sessions become inactive

→FSCopy operations will abort

→Examine crash header to determine Vol ID & block number of affected file system--check Server Log & Sys\_log to confirm entries

→Start the Revector command to correct the bad blocks

→After Revector completes, abort Replication Sessions on affected file system

→Track progress of Revector in Server Log--a list of repaired blocks is written to a “...Repaired.txt” file on the rootfs of the DM

→Run nas\_fsck on Production Server, which runs fsck, runs MapBlock & automatically creates the filename-to-inode-to-path file for customers--replaces “...Repaired.txt” file with a “...Report.txt” file

→Remount file system on Production Server

→Present tidy list to customer & restore files from Backup

### **Panic Handler States for Volume Recovery Situation:**

DISKID\_SET—Panic Handler disk\_id & bad sector address for basic volume set in PH just prior to panic when media error encountered

HYPERVOL\_SET—Panic Handler state after reboot—LUN-based physical blocks are reverse mapped to hyper-based logical block numbers and file written to Data Mover rootfs

REVECTORING\_IN\_PROG—Panic Handler state during the revector recovery process. If interrupted, can be resumed (this state won’t be used for Phase I Prime cse\_recover)

REVECTORING\_DONE—Panic handler state at end of the sector cleaning process. This entry is removed once the FSCK/MapBlock process has been completed, and file system can be normally remounted and accessed

### **Repair Process:**

1. Uncorrectable occurs

2. Data Mover tries to Read invalid sector & generates 03x11 scsi errors
3. Data Mover panics, reboots, comes back online without bad file system

#### # server\_mount server\_2

root\_fs\_2 on / ufs,perm,rw  
sector on /sector ufs,perm,rw

#### **bolsha on /bolsha ufs,perm,rw,<unmounted>**

4. Check Panic Header, Sys\_log, or run “mepanic dump” to obtain Revector information:

#### **DART panic/fault message:**

>>PANIC in file: ../BVolumeIrp.cxx at line: 323 :  
IO failure on Vol:130, blkNo:5248, SKey:0x311

#### **Sys\_log:**

Nov 3 15:57:56 2006 UFS:3:7 Slot 3: 1162587599: Mount failed, Please start Revectoring using "revector start vol=137"

5. Run Revector Tool on Production Server to repair bad blocks

#### **# .server\_config server\_2 "revector start vol=137"**

6. Check progress in Server Log:

2006-11-03 17:01:00: STORAGE: 4: RepThread:VC130\_0 Volume 130 repair is in-progress(Status:91%)  
2006-11-03 17:02:12: STORAGE: 3: 8: Revector Completed:BasicVol:130, 19 blks repaired, total 40959984 nBlks covered

7. Use “mepanic” command to verify if the revector has completed—State 4 indicates that revector process is done

#### **# .server\_config server\_2 -v "mepanic dump"**

1187038890: STORAGE: 4: State: 4 HyperVol: 105 :: LBlkNo: 115102576

8. Revector process creates Repaired.txt file & writes to rootfs

#### **[root@laip2 slot\_2]# cat ME\_Hyper\_137\_Repaired.txt**

5248  
5235

5249 --output abbreviated--

9. Run nas\_fsck on Production Server (Invokes MapBlock):

#### **# nas\_fsck -start bolsha**

id = 74  
name = bolsha  
volume = v137  
fsck\_server = server\_2  
status = In-progress  
2006-11-03 17:05:17: FSTOOLS: 4: 0: FsId: 74 Fsck Started (Manual).  
2006-11-03 17:05:17: STORAGE: 3: Found bad sector 5248 on Vol:137 repair state 4  
2006-11-03 17:05:17: FSTOOLS: 4: starting fsck  
2006-11-03 17:05:17: FSTOOLS: 4: 6: MAPBLOCK:: MapBlock Started  
2006-11-03 17:05:17: FSTOOLS: 4: \*\* Last Mounted on /bolsha  
2006-11-03 17:05:17: FSTOOLS: 4: Phase 1: Validate Inodes  
2006-11-03 17:05:17: FSTOOLS: 4: 0 % complete  
2006-11-03 17:05:25: FSTOOLS: 4: Aclchk fsid 74: 100% complete  
2006-11-03 17:05:26: FSTOOLS: 4: MAPBLOCK:: Finding Pathnames: Completed.

10. After fsck completes, the Repaired.txt file is replaced with the Report.txt that contains File name, Sector No., Block No., Inode No., and path

#### **[root@laip2 slot\_2]# cat ME\_FS\_74\_Report.txt**

-----MapBlock Output-----

| Sector No.           | Block No | Inode No | Path            |
|----------------------|----------|----------|-----------------|
| 5248                 | 328      | 25       | /data1/         |
| 2k3tools/rktools.exe |          |          |                 |
| 5249                 | 328      | 25       | Duplicate Entry |
| 5250                 | 328      | 25       | Duplicate Entry |

#### **Server Panic During FSCK:**

2007-10-18 04:10:26: FSTOOLS: 4: 0: FsId: 2553 Fsck Started.  
2007-10-18 04:10:26: STORAGE: 3: Found bad sector 884688320 on Vol:12552 repair state 2  
2007-10-18 04:10:26: UFS: 3: Fsck failed as revectoring for Media Error affected HyperVol:12552 not done  
2007-10-18 04:10:26: UFS: 3: File fsck failed for fs 2553 with status IO\_Error  
2007-10-18 04:10:26: FSTOOLS: 4: 2: FsId: 2553 Fsck Failed

**Note:** If there is a problem during FSCK, such as a Server panic, and the FSCK needs to be run again, you may need to run the following command in order to reset the proper Panic Handler flags to allow the FSCK to run on the affected file system volumes

#### **\$ .server\_config server\_7 -v “mepanic reset all”**

11. Remount file system and verify:  
**# server\_mount server\_2 -a**
12. Provide customer file list for restore purposes

### **CLARIION TOOL:** Bad Blocks Reporting & Clean-up Tool (BRT)

In certain situations, this tool may be used by Clariion TS2 to clear invalid sectors and map bad blocks to data files

#### **BRT PROCEDURE:**

1. Use naviseclli to obtain list of invalidated sectors (Flare 16 & 19) from Windows or Solaris host using correct agent
2. Use naviseclli to clear bad blocks by zeroing out
3. Map bad blocks to data files
4. Restore identified files from backup source

## **INTEGRATEDS FAIL TO CALL HOME FOR UNCORRECTABLE ERRORS FROM NAVIEVENTMONITOR:**

/nas/log/sys\_log: (AR79015 fixes this issue in 5.4.27, 5.5.25, and 5.6.36 versions)

May 12 21:00:01 2006 NaviEventMonitor:3:3 Backend Event Number 0x957 Host OEM-XOO25IL9VL9 Storage Array  
APM00041700339 SPA Device

Bus 1 Enclosure 1 Disk 12 SoftwareRev 6.19.0 (4.14) Unknown Error 2.19.0.701.5.027 Description Uncorrectable Sector

May 12 21:00:01 2006 NaviEventMonitor:4:2 Backend Event Number 0x840 Host OEM-XOO25IL9VL9 Storage Array  
APM00041700339 SPA Device

Bus 1 Enclosure 1 Disk 12 SoftwareRev 6.19.0 (4.14) Unknown Error 2.19.0.701.5.027 Description Data Sector Invalidated

May 12 21:00:01 2006 NaviEventMonitor:3:3 Backend Event Number 0x953 Host OEM-XOO25IL9VL9 Storage Array  
APM00041700339 SPA Device

Bus 1 Enclosure 1 Disk 8 SoftwareRev 6.19.0 (4.14) Unknown Error 2.19.0.701.5.027 Description Uncorrectable Parity Sector

May 12 21:00:01 2006 NaviEventMonitor:3:3 Backend Event Number 0x956 Host OEM-XOO25IL9VL9 Storage Array  
APM00041700339 SPA Device

Bus 1 Enclosure 1 Disk 8 SoftwareRev 6.19.0 (4.14) Unknown Error 2.19.0.701.5.027 Description Parity Invalidated

**Note:** See AR79015 for more details—fixed in NAS Versions 5.4.27.0 & 5.5.25.0. For Captive or Integrated NS systems, Celerra uses a navilog\_mon entry in the nas\_mcd.cfg file to run the navilog\_mon process, which collects CLARiiON events from NaviEventMonitor [Naviagent] and then posts to the sys\_log using /nas/sbin/postevent. Unfortunately, none of the Uncorrectable errors are delivered with a Severity of CRITICAL, so the Celerra does not CallHome.

#### **# cat /nas/sys/nas\_mcd.cfg**

```
daemon "Navi Event Monitor"  
  executable  "/nas/sbin/navilog_mon"  
  optional    no  
  autorestart yes
```

## **BACKEND EVENTS THAT INTEGRATED CELERRAS WILL CALLHOME FOR:**

**# nas\_event -list -a callhome |grep Navi**

NaviEventMonitor 4 Navi Event with severity CRITICAL was received

NaviEventMonitor 100 A control lun has been cache compromised

**Note:** Navi Events with severity CRITICAL will generate CallHomes, as will the event 100 for Cache Compromised Control Luns

#### **NAVI EVENTS:**

#### **# nas\_event -l -f NaviEventMonitor**

```
id description  
1 Navi Event with severity INFORMATION was received  
2 Navi Event with severity WARNING was received  
3 Navi Event with severity ERROR was received  
4 Navi Event with severity CRITICAL was received  
5 Unknown Navi Event was received  
10 Navi Event Monitor has terminatedF  
100 A control lun has been cache compromised  
101 A control lun has been trespassed  
200 Device group has changed status  
201 A SFP on the CLARiiON is reporting an error status
```

#### **# cat /nas/sys/nas\_eventlog.cfg**

```
#  
# Navi Event Monitor
```

```
#  
facilitypolicy 138, 0  
disposition range=0-1000, logfile "/nas/log/sys_log"  
disposition range=4-4, callhome immediate  
disposition range=4-4, exec "/nas/tools/automaticcollection -getlogs"  
disposition range=100-100, callhome immediate  
disposition range=100-100, exec "/nas/tools/automaticcollection -getlogs"
```

#### **EXAMPLES FROM SYS LOG:**

Aug 27 11:42:51 2007 NaviEventMonitor:**2:4** Backend Event Number 0xa07 Error Host OEM-2MYA58ZGAQT Storage Array APM00062101038 SPB Device Bus 1 Enclosure 0 Disk 1 SoftwareRev 6.24.1 (5.0) BaseRev 2.19.0.500.5.040 Description CRU Powered Down

Aug 27 11:42:51 2007 NaviEventMonitor:**3:3** Backend Event Number 0x906 Host OEM-2MYA58ZGAQT Storage Array APM00062101038 SPB Device Bus 1 Enclosure 0 Disk 1 SoftwareRev 6.24.1

**2:4 →Severity 2 = Critical, Event ID 4**

**3:3 →Severity 3 = Error, Event ID 3**

#### **EXAMPLE BACKEND EVENT THAT GENERATES CELERRA CALLHOME FOR INTEGRATED:**

**Situation:** SP Reboot will generate the 0xa23 event, reported by Clariion as a severity CRITICAL issue, and Celerra reports in sys\_log as a Critical 2, which generates a CallHome

##### **sys log:**

Feb 28 19:48:09 2007 NaviEventMonitor:**2:4** Backend Event Number 0xa23 Error Host SPA Storage Array APM00063303725 SPA Device SP B SoftwareRev 6.24.1 (0.0) BaseRev 3.24.0.40.5.007 Description Peer SP Down.

**2:4 →2 is severity level = Critical, while 4 is Event ID**

##### **getlog from SPA:**

02/28/2007 19:47:01 SP B (a23) Peer SP Down. [0x03] 0 0

##### **# cat /nas/log/ConnectHome/RSC\***

<![CDATA[Backend Event Number 0xa23 Error Host SPA Storage Array APM00063303725 SPA Device SP B SoftwareRev 6.24.1 (0.0) BaseRev 3.24.0.40.5.007 Description Peer SP Down.

#### **CLARIION NAVIEVENTMONITOR:**

Clariion uses an external events file (callhome\_template) hosted on an external Windows management system that monitors events on the array and takes action depending on the Rules. This file is similar to nas\_eventlog.cfg and changes from release to release. For Integrated arrays, this callhome\_template is not configured and NaviEvents are forwarded to Celerra for disposition. NAS & CLARiiON do not coordinate CallHome levels, leading to the problem whereby Uncorrectable Sector errors can occur on the backend but the Celerra does not consider these CallHome worthy. This is fixed with AR79015, NAS 5.4.27.0 & 5.5.25.0—Celerra should now see an Uncorrectable event from NaviEventMonitor as a CRITICAL (4) event from the Backend, and will generate a CallHome.

#### **NAVIAGENT:**

##### **# ps -ef |grep navi**

```
root 25176 21500 0 May28 ? 00:00:01 /usr/bin/perl /nas/sbin/navilog_mon  
root 25181 21500 0 May28 ? 00:00:50 /nas/opt/Navisphere/bin/naviagent -d -f /nas/etc/Navisphere/agent.config -r /nas/log
```

→Naviagent is a daemon that reads events delivered from the CLARiiON, processes the events as defined in the /nas/sys/storage\_eventlog.cfg file, and sends the events to the sys\_log. Events meeting the CallHome criteria then generate a CallHome.

→Naviagent keeps track of which events have been already processed so as not to repost

→Naviagent runs as a daemon controlled by the NAS Master Control Daemon [see /nas/sys/nas\_mcd.cfg]

daemon "Navi Event Monitor"

```
executable "/nas/sbin/navilog_mon"  
optional no  
autorestart yes  
daemon "Naviagent"  
executable "/nas/sbin/start_naviagent"  
optional no  
autorestart yes  
ioaccess no
```

#### **VIEWING CLARIION EVENTS USING NAVISPHERE:**

- a.) Obtain the appropriate Flare template file from Clariion Support—Call\_Home\_Template\_6.24.0.tpl
- b.) Open Navisphere>Monitors tab>Templates>import templates>browse to template file of choice and load

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
c.) Rightclick the imported template file>properties>Advanced: All events are listed under the CRITICAL, ERROR, WARNING, & INFORMATION icons for review

**Note:** Please note that not all events are “enabled” by default

## **MISCELLANEOUS HARDWARE & SETTINGS:**

### **MODEMTEST COMMAND:**

**# /nas/sbin/modemtst**

Modem test V0.7

All ports are busy now, try again later

**Note:** emc74010 documents that a software fix is now available with 5.4.17.0 & 5.3.20.0. It was found that HUP was not calling Login process after being terminated, no matter what the reason for the termination, and therefore could not properly reset ttyS0 modem connection. The fix is built into “mgetty-1.1.30-0.7.EMC.02” and now allows the Linux login process to properly receive a HUP signal upon disconnect and then to properly reset the modem port. See AR53833.

**# rpm -q mgetty**

mgetty-1.1.30-0.7.EMC.02

**/etc/mgetty+sendfax/mgetty.config**

data-only y (this line must not be commented out)

### **VARIOUS NAS 5.6 MODEM ISSUES:**

#### **1. ETA emc188184 Modem Stops Dialing Home after Upgrading:**

The **/etc/mgetty+sendfax/mgetty.config** file did not carryover from 5.5 with the proper string:

##### **Modem Test indicates the issue:**

Error 14504231224: Test failed for the Primary Modem.

##### **Wrong entry:**

fax-id 49 115 xxxxxxxx

##### **Correct entry:**

data-only y

**Note:** See AR118928 & AR118933 for more details. Issue fixed with NAS 5.6.38.2 and higher. A workaround for a system that is not running NAS 5.6.38 or higher is to add the following lines to the /nas/opt/connectemc/modem.cfg file:

# cat /opt/connectemc/modem.cfg

**set dial display on**

**set modem command init-string AT+FCLASS=0\{13}**

#### **2. ETA emc192298 Cannot Dial into Celerra Modem after Upgrade to 5.6.38:**

There is an issue with the Control Station startup scripts where the link file /dev/callhome is not recreated during reboots. This link file is necessary to associate the modem port (usually ttyS0) with the /dev/callhome service. The engineering solution is to place the following command inside the /nasmcd/sbin/ch\_mgetty\_monitor script in order to ensure that the link is always present for Dial In service. The Dial Out CallHome service is not affected because it always runs /nas/sbin/modemtst, which automatically recreates the /dev/callhome link if necessary.

##### **FIX:**

**# ln -sf /dev/ttyS0 /dev/callhome**

##### **Example of output when modemtst is run and the /dev/callhome link does not exist (modemtst will recreate the link):**

# /nas/sbin/modemtst

modemtst: version 1.0

modemtst: No callhome port /dev/callhome

modemtst: Callhome port: /dev/callhome --> /dev/ttyS0

modemtst: Modem /dev/ttyS0 OK

modemtst: Deleted symlink /dev/callhome --> /dev/ttyS0

modemtst: Callhome port: /dev/callhome --> /dev/ttyS0

##### **/nasmcd/sbin/ch\_mgetty monitor script:**

→Script ensures that mgetty process is started and monitoring the /dev/callhome device in order to deliver CallHome events

→Monitors for ConnectHome event files and delivers CallHomes—replaces ch\_exec function

→Provides for capability to Dial into modems even without NAS Services running

# ps -eafl |grep mgetty

4 S root 8359 1 0 81 0 - 922 wait Jul14 ? 00:00:00 /bin/sh /nasmcd/sbin/ch\_mgetty\_monitor

### **NORMAL MODEMTEST OUTPUT:**

**# /nas/sbin/modemtst**

Modem test V0.7

Modem OK.

**Note:** Typical test reinitializes modem by testing for Carrier, Send, Receive signals before returning TR light to on status

## **PROBLEM WITH FAX LINES CALLING CONTROL STATION:**

**Note:** NAS 5.4.20.1 will prevent mgetty from monopolizing the modem due to fax calls

### **# rpm -q mgetty**

mgetty-1.1.30-0.7.EMC.02 (new version)

### **WORKAROUND FIX FOR OTHER VERSIONS:**

1. Edit /etc/mgetty+sendfax/mgetty.config to add following statement just above the ----port specific section----

### **data-only y**

----- port specific section -----

2. Grep for mgetty callhome process and kill the process—process will auto restart

# ps -ef |grep mgetty

root 1753 1 0 2005 ? 00:00:02 /bin/sh /nasmcd/sbin/ch\_mgetty\_m

#kill 1753

## **OTHER TROUBLESHOOTING COMMANDS:**

### **# ps -ef |grep -i mget**

```
root 1191 1 0 17:31 ? 00:00:00 /sbin/mgetty -x 0 ttyS1
root 6582 1 0 17:44 ? 00:00:00 /sbin/mgetty -x 0 ttyS0
```

### **# /usr/sbin/lsof /dev/ttYS\*** [Shows info about serial ports opened on CS0]

| COMMAND | PID  | USER | FD | TYPE | DEVICE | SIZE  | NODE NAME  |
|---------|------|------|----|------|--------|-------|------------|
| mgetty  | 1191 | root | 0u | CHR  | 4,65   | 70965 | /dev/ttYS1 |
| mgetty  | 1191 | root | 1u | CHR  | 4,65   | 70965 | /dev/ttYS1 |
| mgetty  | 1191 | root | 2u | CHR  | 4,65   | 70965 | /dev/ttYS1 |
| mgetty  | 6736 | root | 0u | CHR  | 4,64   | 70964 | /dev/ttYS0 |
| mgetty  | 6736 | root | 1u | CHR  | 4,64   | 70964 | /dev/ttYS0 |
| mgetty  | 6736 | root | 2u | CHR  | 4,64   | 70964 | /dev/ttYS0 |

### **# setserial**

setserial version 2.17, 27-Jan-2000

```
usage: setserial serial-device -abqvVWz [cmd1 [arg]] ...
       setserial -g [-abGv] device1 ...
-----abridged output-----
```

### **# ls -la /var/run/mgetty\***

```
-rw-r--r-- 1 root root 5 Oct 8 17:51 /var/run/mgetty.pid.ttyS0
-rw-r--r-- 1 root root 5 Oct 8 17:31 /var/run/mgetty.pid.ttyS1
```

### **# /sbin/fuser -u -v /dev/ttYS\***

| USER       | PID  | ACCESS | COMMAND      |
|------------|------|--------|--------------|
| /dev/ttYS0 | root | 7666   | f.... mgetty |
| /dev/ttYS1 | root | 1191   | f.... mgetty |

## **STOPPING AND RESTARTING MGETTY PROCESS:**

1. Verify mgetty process:

### **# ps -ef |grep mget**

```
root 1353 1 0 09:51 ? 00:00:00 [mgetty]
root 10598 1 0 10:49 ? 00:00:00 /sbin/mgetty -x 0 ttYS0
```

2. Vi and comment out Serial Port line "S0:3:respawn"

### **#vi /etc/inittab**

# Run mgetty to monitor serial port

#S0:3:respawn:+/sbin/mgetty -x 0 ttYS0 [Comment out this line]

3. Shutdown mgetty process by running:

### **# /sbin/telinit q**

4. Verify that mgetty stops:

### **# ps -ef |grep mget**

```
root 1353 1 0 09:51 ? 00:00:00 [mgetty]
root 31111 20844 0 11:18 pts/0 00:00:00 grep mget
```

5. To restart, uncomment the line in /etc/inittab and start the mgetty process, then verify that it runs:

### **#vi /etc/inittab**

```
S0:3:respawn:+/sbin/mgetty -x 0 ttyS0
# /sbin/telinit q [Starts mgetty process by reexamining /etc/inittab file]
# ps -ef |grep mget
root    1353   1  0 09:51 ?    00:00:00 [mgetty]
root    7065   1  0 11:23 ?    00:00:00 /sbin/mgetty -x 0 ttyS0
```

## **CONFIGURING MODEM ACCESS TO LINUX CS WITHOUT NAS OPERATIONAL:**

```
#rm /etc/nologin.ttyS0
#/sbin/mgetty -x -0 ttyS0
```

## **HOOKING UP KEYBOARD & MONITOR DIRECTLY TO DATAMOVER:** Keybd = lower 'db9' connection

- Step 1. Connect keyboard to lower db9 port on DM
- Step 2. Connect monitor to db15 video port
- Step 3. Plug monitor into a/c
- Step 4. Type following message: Console>logsys add output console
- Step 5. Run commands on datamover

## **CONNECTING DIRECTLY TO NX4, NS-120, NS-480, NS40, NS20 BLADE CONSOLE:**

1. Connect serial cable from Laptop (COM1 or COM2) to upper serial maintenance port (indicated by wrench symbol) on the back of the Blade using mini-DB9 cable
2. Open and configure Hyperterminal session using the following settings:  
Name of connection: Connect using: COM1 or COM2; Bits per second: 9600 Data bits: 8 Parity: None Stop bits: 1 Flow Control: None (verify that Autodetect & ANSI are selected for terminal emulation)
3. Click on 'Call' icon to connect to Data Mover console:  
CONSOLE>
4. Set output to the console and run direct commands to the Server:

**CONSOLE>logsys add output console**

1233070880: ADMIN: 6: Command succeeded: logsys add output console

**CONSOLE> version**

Product = EMC Celerra File Server

Version = T5.6.43.4 (CMB\_SLEDGEHAMMER 32-bit)

Debug = DEBUGOFF

BIOS = 3.58

POST = Rev. 01.50

1233070896: ADMIN: 6: Command succeeded: version

CONSOLE>ifconfig

**Note:** Version command for DART is available only with 5.6.43 and later

5. Reset the console output to log:

**CONSOLE>logsys delete output console**

## **CELERRA PERFORMANCE TROUBLESHOOTING:**

### **Two Performance Measurements:**

1. Response Times (User's Perspective): Timespan between User request and actual response
2. Throughput: Number of transactions accomplished in a given time by the Server (truest picture of actual work being done)

### **Two Types of Performance Problems:**

1. System Timeouts--usually indicated with application, client, or Celerra logs or error messages
2. Performance Decreases--without error messages; Requires Network Traces & Careful analysis

### **Criteria for Busy Systems:**

- Greater than 70% utility of any component
- FC IOPS exceeding 100 I/O per second
- ATA IOPS exceeding 50
- Number of dirty pages exceeds the High Watermark
- Disk Queues greater than 10

### **Basic Systems Involved:**

1. Disk Subsystems--I/O Reads & Writes. UFS filesystems consist of 'superblocks', 'Inode Tables', & 512kb 'Data Blocks'. Superblocks contain filesystem information such as total number of free data blocks, number of free inodes, etc. Inodes contain all information about a file, except for its name, which is located in the directory. Data is addressed in 512 blocks. The physical disk is further divided into cylinder groups, which are collections of tracks over a number of multiple platters. The concept of 'cylinder groups' was introduced to improve disk performance, as files are stored in cylinder groups to reduce access times. Each cylinder

group also contains a copy of the primary superblock for the filesystem. Data is stored in “blocks” and “fragments” (8kb/1kb). In summary, information is stored on disk as ‘data’ and as ‘metadata’. When files are being created, deleted, or updated, updates are made to metadata in memory, then flushed to disk in a ‘metadata’ journal log fashion (Celerra ufslog). Only entire transactions are logged to this ufslog, which allows for complete transaction replaying of filesystem information after system crashes, etc.

### **Important I/O Factors:**

- Sequential I/O optimizes disk subsystems by using less random access and more sequential access with fewer disk movements, and can optimize I/O further with prefetch caching mechanisms
- Read vs. Write operations are factors
- I/O request size affects I/O time—request size is size of I/O used to transfer data from disk to memory, and vice versa, and is usually a function of file system block size
- Average customer file system is 70% Reads & 30% Writes

### **Typical Situations Using Random I/O Patterns:**

- Home Directories
- Transactional Databases
- SnapSure SavVol Reads
- IP Replication activity

### **Typical Situations Using Sequential Read I/O:**

- Backup to disk Restores
- Policy Engines and Indexing
- Imaging

### **Typical Situations Using Sequential Write I/O:**

- SnapSure SavVol writes
- Backup to Disk backups
- Imaging

### **Disk Bottleneck Indicators:**

High disk utilization; large disk queue lengths; larger % time waiting for disk I/O; large I/O rates; low buffer cache hit ratio; large run queues with idle CPU

2. Network--Lookups, connections, session builds/teardowns, Reads, Writes, Locks, etc.
3. CPU—CPU’s run one instruction per cycle. CPU’s utilize high speed cache to maintain data and instructions to keep the CPU busy. High speed cache can be accessed in one CPU cycle. TLB cache is used to speed up translation of virtual memory addresses to physical memory addresses. Users affect CPU usage in two ways: Application processes and System Kernel calls that result from User operations.

### **CPU Bottleneck Indicators:**

0% Idle; Large Run Queues over time; Slow response times; High % system calls

4. Code—Code is executed as “processes” running in RAM and executed in two modes--User & Kernel Level. Kernel level is the only level allowed to access hardware directly, such as I/O, IPC, Network communications, etc. (this is why such things as NIC drivers are also kernel-level drivers). Unix is a multi-user and multitasking Operating System where User and Kernel modes are measured by time.

**Processes:** Processes are started from code stored on disk and are created by a parent from a fork system call, which are then placed into memory to begin execution of code. Paging & Swapping operations only apply to the “data” pages of a running process. If processes have to wait before being executed, they may be written to swap space. Processes can be stopped by signals issued from Hardware (Processor), Software (Kernel O/S or Application), or User (ctrl + c on keyboard). CPU Scheduler takes processes and schedules them with a finite amount of CPU time (quantum or time slices, usually 100ms), depending on priority. (3) Types of CPU priorities are “real-time” (1-127), “system” (128-177), and “user” (178-251)—the latter two makeup “timeshare processes”. When processes have to wait to be executed, their priority increases. When processes are being executed, their priority lowers.

**Buffer Cache:** Data read from or written to disk are stored in “buffer cache”, a RAM pool designed to minimize physical I/O from disk. Both Reads and Writes are copied to buffer cache—writes are written periodically to disk by the syncer daemon. Buffer cache contains “data” and “metadata” (superblocks or mounted file systems, inodes of open files, cylinder group information tables, etc), and are usually written synchronously to disk, which is the quickest method. Asynchronous writes are safer, but not generally used due to longer time and overhead involved.

### **Static Buffer Cache vs. Dynamic Buffer Cache:**

Static buffers are the traditional kernel parameters “nbuf” and “bufpages”, usually 10% of physical RAM in size. Nbuf defines the number of files that can have a buffer cache entry, while “bufpages” defines the actual size of buffer cache in 4kb pages.

**Caution:** Increasing “buffer cache” will reduce amount of memory available for other processes and can result in excessive paging.

### **How to Analyze a Customer's Environment:** BPS=(Block Size \* Parallel Requests) /latency

1. Block Size--increasing this value may help
2. Parallel Requests--add to this as well
3. Latency--Reduce this [usually this is the primary Networking variable that can be influenced]

### **Difference Between Direct-Attached and Network-Attached-Storage:**

1. Networking requires more overhead for Users or Host access for Authentication & Communication

2. Direct Attached Storage uses only (1) Host to (1) Disk Device & requires no Authentication or Access Protocols

**Other Considerations Regarding Performance Issues:**

1. Is the Application designed for network operations? [operational latencies are inherent with NAS]
2. NFS v. CIFS Operating Systems handle networking & I/O differently

**Built-In Network Performance Enhancers:**

1. Cache--Memory used to store data for future operations
2. Buffers--collection points for data
3. Queues--executes I/O's in parallel
4. Locking--several locking mechanisms in use

**File Locking Modes: Usually handled by O/S**

1. Exclusive--Read/Write access exclusively granted to (1) client at a time; Write buffers used Read Ahead & Write Back buffers  
Shared--Read only access to multiple clients. No Write buffers involved.

Nolock--Concurrent Read/Write operations controlled by O/S. No Read Ahead or Write Back buffers used.

**NFS LOCKING & I/O BEHAVIOR:** Stateless protocol using LOCKD & STATD daemons for locking; Server does not cache data; Clients can retry after 'failures' without ill effect; Session builds not required, making this fault tolerant; Very fast Write operations to disk; NFS negotiates I/O size during mount operation for NFSv3 but with NFSv2 only 8k I/O transfer sizes used for Reads/Writes. Journal filesystem uses (1) device & can be taxed by NFS operations--writes much faster to disk than CIFS.

**Note:** If a Server crashes or loses connectivity while Clients have established locks, Server keeps track of these locks and will notify each Client of their previous locks, allowing a 'grace period' of 45sec for reclaiming locks.

**CIFS LOCKING & I/O BEHAVIOR:** Stateful protocol with 'oplocks' builtin; Requires overhead for Session builds, no fault tolerance; Server caches I/O ops.; CIFS is sequential in behavior and does not allow parallel sessions for single files; I/O varies from 4 - 64k and is negotiated during Session Setup.

**ETHERNET NETWORKS:**

Switches & NICs autonegotiate Speed & Duplexing via Hardware controls [Limit this through SW commands]

Switches support Full & Half Duplex, while Hubs are purely Half Duplex

**Rule of Thumb:** Always set Speed & Duplexing uniformly throughout the network

**Exception:** Datamover can be auto full if connected to Switch port that is also auto full, and communicate to a Half-Duplex client on a Hub if the Hub port connected to the Switch is set correctly to half duplex [Switch, not Hub, will handle buffering]. This is called the concept of "Link Partners", where the Host setting & Switch Port settings must match, Host-to-Host does not matter.

**Gigabit Ethernet:**

Data travels at light speed, same as 100 Base-T, but because clock times are 10x faster, can place more data on wire in less time. This mechanism requires 'flow control' for both sides to communicate properly--ensure Flow Control is either disabled on both ends or enabled.

**CELERRA PERFORMANCE MECHANISMS:**

Memory on DMs used to cache metadata for Inode & Session information, ACL's & Auth. Info [Do not Cache data--Symm does this] Async Threshold--Number of FBA block changes collected prior to sending I/O to disk subsystem. By default, Celerra does 32 FBA block changes (16kb). Max. Async Threshold for NFSv3 should not exceed Write Size.

**Note:** Celerra does not conduct automatic load balancing between FA ports

**FileSystem:** Celerra uses 8k block sizes. Max I/O size for NAS NFSv2 is 8k.

**Quotas:** Consider not setting asynch threshold to 1; Using async=1 means that the User's quota is recalculated every 1k change in file size to provide for more accuracy--yet gives less performance. If performance is paramount, do not set this value to 1--performance drops to 512 bytes per I/O, not 8k.

**Device Striping:** In general, use Celerra striping of 32k or more for best file system performance across a min. of (5) devices. If Symm Striping is required, use 2 cyl striping.

**CELERRA PERFORMANCE TESTING NUMBERS:** [www.spec.org/sfs97r1](http://www.spec.org/sfs97r1)

**510 DM with CX600 and DART 5.0**

NFS throughput 16014 ops/sec with 11.1 msec response using Jumbo GigE Alteon NIC with Cisco 6509 GigE Switch & Seagate 73GB 10K drives; 2GB Write & 1.47GB Read cache

**510 DM with Symm 8430 and DART 4.0**

NFS throughput 16772 ops/sec with 9.4 msec response using Jumbo GigE Alteon NIC & Seagate 73GB 10K drives; 2GB Write & 16GB Symm Cache

**514 DM with DMX1000 and Dart 5.1**

NFS throughput 25156 ops/sec with 6.1 msec using Jumbo GigE Intel card and Cisco 6509 Switch, Cheetah 146GB 10K drives, 32GB Cache

**3.06GHz DM with CX700/NS700G and Dart 5.2**

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
NFS throughput 71482 ops/sec [dual DM] with 11.2 msec using Jumbo GigE Intel card and Cisco 6509 Switch, Seagate 146GB 10K drives, 3 GB Mirrored Write Cache  
**2 GHz DM with NS600 and Dart 5.1**  
NFS throughput 25656 ops/sec (dual DM) with 6.5 msec using Jumbo GigE and Cisco 6509 Switch, Seagate 73GB drives, 2048 Mirrored write cache

## **CELERRA SYSTEM PERFORMANCE TOOLS:**

**SAR (System Activity Reporter) COMMAND:** CLI Statistics for Processes, System Calls, File Systems, Hard Disks, Processors, Memory, Kernel, I/O  
**\$/usr/sbin/sar -u 5 5** or \$sar -u 1 5 [Provides STATS on CS0 CPU activity, et al for SCO 2.1.3]

## **CLARIION PERFORMANCE/TROUBLESHOOTING CLARIION PERFORMANCE:**

### **I. KTRACES:**

Use “ktraces” to capture real-time host traffic on the Clariion FrontEnd SP’s using Symm Remote—collect one report for each SP. This capture is non-intrusive and does not impact SP’s. Ideally, run a 3-minute capture, analyze, then collect a larger timespan if required. Typically this tool would be used in conjunction with the Navisphere Analyzer capture to review BackEnd stats. This tool is built-into each Clariion platform—files with .ktr extension. These traces are analyzed by Clariion Engineers.

### **II. KTCONS:**

--Tool used to look at Ktraces:  
#ktcons -il  
>!load fdb  
>!spstat (or other commands)  
Ktrace:  
>!load fdb  
>!redirect c:\ktrace.txt  
>!ktrace -T -a -rstd

### **III. NAVISPHERE ANALYZER (NaviAnalyzer):**

This tool requires software installation on each Array. Use this tool to evaluate Clariion BackEnd performance—analyzed by Clariion Engineers. Peak SP iops on CX600 platform 19,300 (I/O Operations). High Cache hit rates will lead to more SP cpu utilization as they have to work harder since there are less backend calls. SP CPU of 85% or higher is nearly maximum. BackEnd iops of 2000 is maximum for this platform.

→On Array and Off Array versions, produces .nar files

**nas/opt/Navisphere/bin/# java -jar archiveretrieve.jar -User nasadmin -Password nasadmin -Address 10.241.168.57 -File <filename> -Location /tmp**

**Note:** Use above syntax to pull .nar files off Celerra for Integrated systems—may need to define path “c:\analyzer\xxxx”

**Response Times (variable) = Service Time (Static) \* Queue Depth (variable)**

#### **Retrieving NAR files on Flare 26 Systems:**

**# /nas/sbin/navisecli -h ss1\_spa analyzer -archiveretrieve -file archive.nar -location c:\Temp\ -overwrite y -retry 3**

#### **RUNNING NAVIANALYZER FROM CELERRA:**

```
$ /nas/sbin/navisecli -h A_APM00074402116 analyzer -archive -new
$ /nas/sbin/navisecli -h B_APM00074402116 analyzer -archive -new
$ /nas/sbin/navisecli -h A_APM00074402116 analyzer -status
Running. Started on 12/22/2009 19:48:51
$ /nas/sbin/navisecli -h B_APM00074402116 analyzer -status
Running. Started on 12/22/2009 19:48:51
$ /nas/sbin/navisecli -h A_APM00074402116 analyzer -stop
$ /nas/sbin/navisecli -h B_APM00074402116 analyzer -stop
$ /nas/sbin/navisecli -h A_APM00074402116 analyzer -archive -list
32 540 12/22/2009 21:02:33 APM00074402116_SPA_2009-12-23_02-02-32-GMT_M05-00.naz
$ /nas/sbin/navisecli -h B_APM00074402116 analyzer -archive -list
0 551 12/22/2009 21:05:10 APM00074402116_SPB_2009-12-23_02-05-09-GMT_M05-00.naz
$ /nas/sbin/navisecli -h A_APM00074402116 analyzer -archive -file
$ /nas/sbin/navisecli -h B_APM00074402116 analyzer -archive -file
```

### **IV. OTHER TOOLS/CONSIDERATIONS:**

→FBI or RLS Collector: Backend loop troubleshooting, runs automatically with Flare 19  
→Background Verify (Navicli getsniffer runs in background continuously looking for marginal or bad sectors—see emc32911)  
→Cache settings are important—look for High Cache Utilization, Flush rartes  
→Look for high LUN Utilization  
→Look for high I/O queue lengths and response times

## **V. SPCOLLECTS:** Flare 13 +

→Run collection scripts to gather data for TS2 Clariion analysis in order to determine hardware issues, takes about 5 minutes, run navicli spcollect –messner command, no info returned at prompt to indicate success, command is running. SPCollects are written to the /dump directory on the XP SP system, with \_data.zip suffix and current date. Most useful info generally in the \_sus.zip files.  
→Contents of zip are evt.zip (XP event logs, read only from Clariion), fbi.zip (Backend issues), isc.zip (iSCSI), ktd.zip, nav.zip (navicli), psm.zip, rtp.zip (applications), run.zip, sus.zip (Array info & KTfiles, most important logs when looking at backend)

#/nas/sbin/navicli -h spa spcollect –messner [password=messner]

#/nas/sbin/navicli -h spa managefiles –list

C:\dumps\xxx.zip

# /nas/sbin/navicli -h 155.33.25.81 managefiles -retrieve

-file SPA\_APM00044700202\_d6368\_12-14-2005\_20-24-00\_data.zip

→FLARE 19 allows automatic collection of SPCollects, but only one current copy is retained (spcollect –set –auto on)

→SPCollects can be gathered by GUI, FLARE 19, or via CLI, or direct SP Access via EMCRemote connection

## **VI. TOOLS TO ANALYZE SPCOLLECTS & CLARIION INFO:**

**Triage:** DOS bat file tool, creates triage-notes.txt & merged Triage\_logs.txt from SPA & SPB (Triage requires both SPCollects)

**SPLAT:** SP Log Analysis Tool—used to analyze logs & can also look at SPCollects

**Note:** Can download the above tools from Clariion website. CAP2 contains SPLAT

## **CAP2:**

Main purpose of tool is to analyze and report on SPCollects in XML format

→XML formatted report, array configuration and health assessment, launches Splat or FBI Remote—upload/download from SYR Input files for CAP are SPCollects, XML reports, getall files, Connectrix Switch logs, etc.

→Lays out data by Tabs, such as Issues, Hosts, LUN Info, SPInfo, Analysis, etc.

## **VII. REMOTE ACCESS TOOLS FOR CLARIION:**

a. EMCRemote running on a Windows system

b. Primary method now is for Webex session

c. Serial connection with laptop plugged into SP Ethernet port

→Navicli talks directly to agent running on SPs

→GUI or Jar commands go to a cimom process running on the SPs, using HTTP requests to the Navisphere Mgmt Port

## **VIII. OTHER CLARIION TOOLS:**

### **NST (Navisphere Service Taskbar):**

As of Flare 24, the NST can be used by both Customers and CE's to perform various hardware and software tasks on the CLARiiON, such as: Add DAE; Add disks; Replace disk (DRU); Update Flare software via Software Assistant (NDU); Update drive firmware; Register the storage system; Send hardware change information to SYR system; Verifies & prepares Storage System healthcheck report; Latest version 06.28.53 supports FC Solid State Drives (SSD).

**Note:** A network firewall may affect the ability for the NST tool to provide updates, though the tool will say that it is the latest version. In reality, a User would need to check and download the latest .exe file from CLARiiON.

#### **Installation & Use of the NST:**

1. Download the setup\_NST.exe from Powerlink>Products>Hardware/Platforms>CLARiiON CX3 Series>Interoperability/Documentation Tools>CLARiiON Tools

\* Download the Navisphere Service Taskbar

2. Execute setup\_NST.exe and accept defaults

3. Launch NST>Do you want to launch Navisphere Service Taskbar Now? Select yes to launch.

**Navisphere Service Taskbar** [Utilities highlighted in blue]

File Tools Help

**Hardware Registration**>Register Storage System

**Hardware Installation**>Install Disk Array Enclosure & Install Disk

**Software Assistant**>Prepare for Installation & Install Software>(3) Options for Prepare for Installation section: Download software and verify storage environment; Download software only; Verify storage environment only

**Hardware Maintenance**>Verify Storage System & Replace Disk

4. Connect to CLARiiON using File>Connect>Navisphere CONNECTION:

Enter IP Address of SP and login with proper Username and Password if Global security is in place

5. Once logged in, select a Utility. Reports are logged on the local harddrive: c:\emc\repository\APM00030600872

**Comments:** System produces healthcheck reports of the Storage System, validates existing software, assists in downloading and upgrading software, including Rules Checking, provides assistance in adding or replacing disks, adding DAEs, etc.

## **SPQ—Storage Processor Qualifier:**

HealthCheck tool for SPs, using SPCollects as input files

## **TRIAGE:**

## **CELERRA BEST PRACTICES/CELERRA PERFORMANCE (Updated NAS 5.5):**

**Note:** See the 5.5 Best Practices for Performance white paper as found on PowerLink for more detailed and accurate information. With NAS 5.6, there was no overall Best Practices white paper published. Check Celerra forum for related information.

### **I. STORAGE ARRAY RECOMMENDATIONS:**

1. Dedicate separate spindles for NAS & SAN workloads & maximize spindle count whenever possible
- Note:** NAS environments typically have high thread counts & random I/O. SAN environments are characterized by sequential I/O
2. Do not build file systems on single Symm metavolumes or Clariion MetaLuns—if absolutely required, build on sets of metaluns
- Note:** Symm metavolume of (8) hypers will only queue (1) outstanding I/O for Celerra, as in the case of synchronous SRDF, but if built on Celerra metas over (8) striped volumes, will increase outstanding queue depth to (8) I/O's. Clariion queue depth based on (8) I/O's per disk volumes, so spread the lun across (8) disk volumes would increase queue depth to (64) I/O's. A key concept here is that I/O parallelism is best obtained when striping file systems over multiple hypers.
3. Do not mix different RAID group types when creating file system LUNs (AVM follows this rule)—an exception is that Clariion R3 allows file systems to span 4+1 & 8+1 RGs. For (2) luns in a RG for FC drives, set one to SPA, one to SPB. For (2) luns in a RG with ATA drives, set all to one SP or the other.
4. For CLARiiON arrays with more than 15 FC drives, create first shelf with R5 4+1 & Hot Spare, then 4+1, HS, HS, HS, HS, all other shelves would be R5 4+1, 4+1, 4+1
5. Balance data Luns across SPs, bus, and DAE's (ignore control luns), keep in mind that all LUNs in same ATA Raid Group must be owned by the same SP
6. Avoid using ATA drives for random I/O or AVM clar\_r1 volume pools
7. For CX & CX3 arrays, assign minimum Read cache of 32MB and the balance to Write Cache [max 3072 for CX; max 1053 for CX3-20; max 2500MB for CX3-40; max 3072MB for CX3-80]; most Clariion Read cache is used for prefetch, while DMs have their own Read cache.

**Note:** Change cache settings on captive NS600/700 systems using [`#/nas/sbin/setup\_backend/nas\_raid cache configure`](#)

Some systems may require more Read cache if Reads are extensive and performance is being impacted

8. ATA-only Gateway systems, set Clariion write cache HWM to 80 & LWM to 60
  9. For CX3 backends, R5 [clarata\_archive 4+1] or R3 will provide good sequential IO writes, while CX series should use R3 [clarata\_r3]. Max. Clariion disk I/O size is 2MB.
  10. Attributes of a busy system are if any component in system is over 70% utilized; throughput per FC spindle exceeds 100 IOPS, 50 IOPS for ATA; number of dirty pages exceeds memory HWM; Disk queue greater than 10
  11. Attributes of sequential IO: average Read size 64KB; average Write size 512KB; Read throughput=Read Cache hits/sec; Write throughput=Full Stripe Writes/sec
  12. Default cache page setting for Clariion is 8kb for Integrated array, which matches Celerra 8kb I/O size
  13. Default Clariion LUN stripe element size for RAID5 luns is 64kb (Navicli reports stripe size element in 512-byte blocks)
  14. Storage templates are used only for Captive Systems and will automatically bind (2) luns per Raid group (manually create 2 luns/raid group if storage template is not used)
  15. Consider disabling prefetching if small random access is being done by clients (param file prefetch=0)
  16. Not recommended to rely on Celerra auto file system extensions if performance is key as this will tend to fragment the file system if small extensions are being done with AVM
- Fibre Channel Drives support CTQ (Command Tag Queuing) for up to 24 simultaneous commands  
--ATA Drives do not support CTQ (more missed spins around the platter)

#### **Symm LUN Addressing Support:**

Symm 8000 supports Luns 0-255 (00-FF) per FA port (processor); DMX2 supports 512 Luns per port--Luns 0-511 & 512-1023 (Hex 00-1FF & 200-3FF); DMX3 supports 2048 LUNs per port.

### **II. AVM STORAGE RULES:**

1. Allow AVM to create file systems on system-defined pools whenever possible (use user-defined pools with AVM if user intervention required)

**Note:** AVM will create LUNs from same storage array; create LUNs from different RAID groups to max. spindle count; use same RAID group types and sizes when creating new LUNs; chooses least utilized LUNs first, then to achieve SP balancing, then to achieve bus balancing, then dvolumes with the highest ID numbers first. AVM will also create pool volumes of (4) disk volumes first, then (3) volume pool, (2) volume pool, and single metavolume last. AVM will use same pool volume types first when creating LUNs.

--nas\_disk -l shows order that luns are discovered on the bus, with SPA first, then SPB

--Do not span LUNs across different storage platforms [AVM could do this if multiple pool volumes are used—consider manually creating user-defined pools for each storage array—AVM will then build file systems from correct pools]

--Do not extend file systems with same disk spindles, use same number of spindles for extension as for original, if possible

--Do not use AVM system-defined pools on top of Symmetrix parity RAID

--File Systems can be created from only one Storage Pool, but an extension can occur from a storage pool of different disk types

--Do not span file systems across more than one storage array

### **CELERRA STRIPE SIZE ELEMENT RULES:**

--Use 32KB stripe size for NFS/CIFS for both Clariion & Symm

--Use 256KB stripe size for MPFS on both Clariion & Symm

--Use 64KB for ATA Backup2Disk on CLARiiON

### **CLARIION AVM RULES:**

--Balance spindles, RAID groups, LUNs per dvolume, SPs, and Storage arrays

--File Systems are created on min. of (4) dvols when Fibre Channel drives (clar\_r5\_performance), with each dvol in separate RG, striped together, and then sliced

--File Systems are created on min. of (2) dvols when ATA drives (clarata\_r3)

--RAID3 designed for ATA drives, highly sequential I/O

### **SYMMETRIX AVM RULES:**

--(8) consecutive dvols are used when creating volumes

--Relies on BIN file ordering to balance I/O across spindles and directors

### **SYMMETRIX STORAGE POOLS:**

--When creating pools, first 8 unused disk volumes of the same size are striped together in 32k stripes

--If volumes are not of the same size, then they are concatenated together [i.e., will not stripe dissimilar disk sizes together]

--If 8 disk volumes cannot be found, then logic looks for 4 unused disks, then 2, then 1

--Symmetrix profile is not greedy in that all pool space is used up before a new pool is created

--AVM not so commonly used for Symm layouts because it does not load balance on the FA's

--But, AVM's can stripe and slice, or stripe and not slice (using slice=n)

SYMM\_STD & SYMM\_STD\_RDF\_SRC

### **CLARIION FC & ATA DRIVE STORAGE POOLS:**

--LUNs must start with 4 identically sized volumes

--LUNs must be from different RAID groups

--LUNs must match storage profile in use

--Balance LUNs between SPA and SPB, and balance on Bus 0 and 1

--Storage Templates can only be used on Captive Clariion arrays

--Original ATA template was RAID5 6+1 NAS 5.1, targeted for sequential I/O—Clariion did not yet support RAID3 [RAID3 added NAS 5.4] CLAR\_R5\_PERFORMANCE (5 drives); CLAR\_R5\_ECONOMY (9 drives); CLAR\_R1 (2 drives FC); CLARATA\_ARCHIVE (6+1R5)

--NAS 5.4 introduces RAID3 support, using Raid3 4+1 and 8+1 for ATA drives

--When creating pools for file systems, AVM will search out highest available dvolume on SPB first, and work its way down to the lowest dvolume number, then does same thing with SPA luns

### **NAS 5.1 AVM:**

→Uses (2) LUNs per pool entry as minimum

→SPA highest to lowest dvolumes first, then SPB highest to lowest

### **NAS 5.2 AVM:**

→Uses (4) LUNs per pool entry as minimum

→SPB highest to lowest dvolumes first, then SPA...

### **NAS 5.5.21.0:**

→AVM profile for clar\_r5\_performance, clar\_r5\_economy, clar\_r1, will be updated to have a stripe size of 32768K (a change from 8192K)

**Note:** If space remains available in used disk volumes (existing pool members) of the storage pool, the stripe size remains at 8 KB. When allocating space from unused disk volumes (new pool members), the strip element size automatically defaults to 32 KB.

### **NAS 5.5.26.2:**

→RAID-5 6+1 support for ATA/LCFC drives added for CX3 arrays, as well as RAID-5 4+1 support, and traditional RAID-3 4+1 & 8+1. RAID-5 support is not recommended for CX platforms with ATA/LCFC drives.

→AVM profile for clarata\_archive stripe size default changes from 8KB to 64KB

### **FILE SYSTEM RULES:**

--do not span storage arrays

--use most spindles possible when building fs

--build most important file systems first (will use more spindles, hence better I/O performance)

--largest file system size 2TB based on SCSI address limitation for a single volume

--up to 16TB file system possible if striped across enough volumes to maintain 2TB SCSI limitation per single volume

### **III. NAS NETWORK PROTOCOL RULES:**

--on dirty networks consider enabling flow control for rxflowctrl and txflowctrl

--ensure speed & duplex settings are set the same on all Host & Switch ports

--Do not set tcp.ackpush unless legacy clients require

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!

- consider using sndcwnd param to limit DM window size when switch buffers are limited (seen with DM gigabit vs. client 100MB)
- max. performance by using ports in following order: cge0, cge3, cge1, cge4, cge2, cge5 (ports 0, 1, 2 are on one bus, 3, 4, 5 are on another bus)
- Use TOE or copper interfaces when possible (DM 514 uses optical interfaces)
- consider using cifssyncwrite mount option if workload is database related for both CIFS and NFS (will degrade performance for nondatabase workloads). Guarantees that writes to file server are done synchronously
- Also consider disabling oplocks in write-intensive database environments
- Set cifs.prealloc parameter from default 0 to 6

**Note:** prealloc size is size allocated to client for writes between requests to write more data. For example, default 0=8kb block size. Setting to 6 increases block size to 512KB and client will not need to request to write data so often.

- Do not increase CIFS threads to more than 256 (take CAVA into consideration when calculating threads used)
- Do not change file.asyncthreshold param from 32 (max number blocks cached by NFSv3 for asynchronous writes). Changing the value downward will flush data to disk more frequently.
- Safely increase NFS threads from 256 to 512 for 510DM and higher systems

## **IV. I/O CHARACTERISTICS:**

### **CHARACTERISTICS OF SEQUENTIAL I/O:**

- average read size >=64kb
- average write size >=512kb
- Read iops/sec=Read cache hits/sec
- Write iops/sec=Full Stripe Writes/sec (write I/O's are being coalesced into full stripe writes)

### **WORKLOAD TYPES:**

- transactional databases tend to have small, multi-threaded I/Os 8k or less that are randomly distributed
- Checkpoint SavVols are highly sequential & write-intensive by design, while reads are seldom and random
- Backup-to-Disk uses large, sequential, single-threaded write I/O's and few reads
- Home Directories are generally highly random and small 8k size
- Imaging consists of large, sequential, single-streamed reads and writes

## **V. HARDWARE—DISK DRIVES:**

### **ATA vs. FIBRE CHANNEL DRIVES:**

- 5400-7200rpm vs. 10,000rpm FC drives
- ATA drives can only handle single command at a time while FC drives support command tag queuing for up to 24 simultaneous commands
- ATA drives best for sequential I/O such as for backups

## **VI. iSCSI:**

- DM can be CPU-constrained when <than 16kb sequential I/O
- BackEnd FC interconnect become bottleneck for large >16kb sequential I/O and random reads
- BackEnd disks are spindle-bound with small <16kb random reads and writes

## **VII. CONFIGURING ETHERNET & FIBRE CHANNEL PORTS:**

- Most basic rule-of-thumb is to leave Celerra Fibre Channel & Ethernet ports set to auto-negotiate, if the respective Switches also auto-negotiate
- If a FC or Ethernet switch port is hard-coded to a certain speed, then the respective Celerra port also needs to be hard-coded e.g., Ethernet switch set to 100 Full Duplex, then Data Mover or Control Station interface should also be set to 100 Full Duplex
- Leave Internal network interfaces to auto-negotiate at all times

### **Exceptions:**

- For FC AUX ports, if the Switch is hard-coded, it is still recommended to let the Celerra autonegotiate speed and topology.
- For NS500 AUX port, which only runs at 2GB, meaning that both DM and Switch should be hard-coded to 2GB
- For 10/100 Fast Ethernet ports on Celerra, the recommendation would be to have the ports set to Full Duplex on DM and Switch
- For GbE ports on Celerra, let the ports autonegotiate
- For Control Station external NIC, which is also Fast Ethernet, hard-code to 100 Full Duplex on CS & Switch

## **LINUX CONTROL STATION STATISTICS:**

**#top**

**Outputting TOP to File:**

**#top -b -n1 >top.out &**

**#vmstat 5 15**

**#free | # free -m**

**# free -mt** (free memory in Megabytes with Totals)

|      | total | used | free | shared | buffers | cached |
|------|-------|------|------|--------|---------|--------|
| Mem: | 502   | 205  | 297  | 0      | 82      | 46     |

```
-/+ buffers/cache:    76    426
```

```
Swap:      509      0    509
```

```
Total:    1012    205    807
```

**# cat /proc/meminfo**

**# vmstat 5 15**

| procs | memory | swap  | io     | system | cpu                        |
|-------|--------|-------|--------|--------|----------------------------|
| r b w | swpd   | free  | buff   | cache  | si so bi bo in cs us sy id |
| 0 0 0 | 296    | 71960 | 128568 | 211876 | 0 0 1 10 16 15 1 1 10      |
| 0 0 0 | 296    | 71928 | 128568 | 211876 | 0 0 0 18 115 58 0 0 99     |

## **DEBUGGING LINUX MEMORY ISSUES:**

**#top** [Running Top shows low Control Station memory available]

**#top -b -n1 >top.out &**

**#ps -axuwW |more** [sift through processes to see what might be tying up resources]

## **\$top COMMAND ON LINUX 7.2 CONTROL STATION:**

3:28pm up 53 min, 4 users, load average: 0.01, 0.02, 0.00  
 121 processes: 120 sleeping, 1 running, 0 zombie, 0 stopped  
 CPU states: 0.5% user, 1.3% system, 0.0% nice, 98.0% idle  
 Mem: 513088K av, 176048K used, 337040K free, 52K shrd, 59916K buff  
 Swap: 523396K av, 0K used, 523396K free 77544K cached

**PID USER PRI NI SIZE RSS SHARE STAT %CPU %MEM TIME COMMAND**

| PID  | USER     | PRI | NI  | SIZE | RSS  | SHARE | STAT | %CPU | %MEM | TIME | COMMAND        |
|------|----------|-----|-----|------|------|-------|------|------|------|------|----------------|
| 1220 | root     | 17  | 0   | 800  | 800  | 656   | S    | 0.7  | 0.1  | 0:08 | nas_mcd        |
| 6434 | nasadmin | 19  | 0   | 1080 | 1080 | 824   | R    | 0.7  | 0.2  | 0:00 | top            |
| 1391 | root     | 12  | 0   | 1160 | 1160 | 880   | S    | 0.3  | 0.2  | 0:07 | nas_boxmonitor |
| 1    | root     | 9   | 0   | 524  | 524  | 456   | S    | 0.0  | 0.1  | 0:04 | init           |
| 3    | root     | 9   | 0   | 0    | 0    | SW    | 0.0  | 0.0  | 0.0  | 0:00 | keventd        |
| 4    | root     | 19  | 19  | 0    | 0    | SWN   | 0.0  | 0.0  | 0.0  | 0:00 | ksoftirqd_CPU0 |
| 5    | root     | 9   | 0   | 0    | 0    | SW    | 0.0  | 0.0  | 0.0  | 0:00 | kswapd         |
| 6    | root     | 9   | 0   | 0    | 0    | SW    | 0.0  | 0.0  | 0.0  | 0:00 | kreclaimd      |
| 7    | root     | 9   | 0   | 0    | 0    | SW    | 0.0  | 0.0  | 0.0  | 0:00 | bdflush        |
| 8    | root     | 9   | 0   | 0    | 0    | SW    | 0.0  | 0.0  | 0.0  | 0:00 | kupdated       |
| 9    | root     | -1  | -20 | 0    | 0    | SW<   | 0.0  | 0.0  | 0.0  | 0:00 | mdrecoveryd    |
| 15   | root     | 9   | 0   | 0    | 0    | SW    | 0.0  | 0.0  | 0.0  | 0:00 | qla2100_dpc_0  |
| 16   | root     | 9   | 0   | 0    | 0    | SW    | 0.0  | 0.0  | 0.0  | 0:00 | qla2100_dpc_1  |
| 19   | root     | 9   | 0   | 0    | 0    | SW    | 0.0  | 0.0  | 0.0  | 0:00 | kjournald      |
| 125  | root     | 9   | 0   | 0    | 0    | SW    | 0.0  | 0.0  | 0.0  | 0:00 | kjournald      |
| 678  | root     | 9   | 0   | 616  | 616  | 512   | S    | 0.0  | 0.1  | 0:00 | syslogd        |
| 683  | root     | 9   | 0   | 1196 | 1196 | 448   | S    | 0.0  | 0.2  | 0:00 | klogd          |

[Super Daemon for Linux processes]

[Daemon that writes NAS\_EVENTS to Sys\_Log]

[Kernel daemon that writes entries to Server Logs]

**Note:** If the “load average: values in the top output exceeds 5.0, that indicates an extreme memory load on the CS. JServer is allowed to use up to 230MB memory and WebUI up to 96MB.

## **USEFUL LOG FOR CONTROL STATION RESOURCE ALLOCATION ISSUES:**

**/nas/log/cs\_res\_mon.log**

20060608.170840:apache:summary: npids=12, cpusec=0, cpu%=0.0, memK=40920, mem%=3.4  
 20060608.170840:apl\_sched:summary: npids=1, cpusec=0, cpu%=0.0, memK=12926, mem%=0.1  
 20060608.170840:jserver:summary: npids=60, cpusec=1712, cpu%=0.8, memK=165036, mem%=8.4  
 20060608.170840:nasdb\_backup:summary: npids=0, cpusec=0, cpu%=0.0, memK=0, mem%=0.0  
 20060608.170840:tomcat:summary: npids=43, cpusec=111, cpu%=0.0, memK=79506, mem%=8.8

**Note:** Log could be useful in viewing a realtime CS resource issue, or for detecting trends on what processes are taking up the most CS resources. Typically, nas\_cmd or APL processes can use up to 25% of total memory, while Tomcat & Apache are usually limited to 10% of the total memory, and JServer 16%.

## **USING IOSTAT TO VERIFY ACTIVITY ON CONTROL STATION PARTITIONS:**

**# /usr/bin/iotstat -d 5 -x /dev/nde1**

Linux 2.4.9-34.5307.EMC (celerra-cs0) 02/01/2006

| Device | rrqm/s | wrqm/s | r/s  | w/s  | rsec/s | wsec/s | avgrq-sz | avgqu-sz | await     | svctm     | %util |
|--------|--------|--------|------|------|--------|--------|----------|----------|-----------|-----------|-------|
| nde1   | 0.21   | 0.02   | 0.03 | 0.00 | 0.25   | 0.02   | 8.08     | 5.16     | 151229.17 | 143016.67 | 47.57 |

**Note:** iostat monitors system input/output to devices, and provides statistics for CPU & DEVICE usage since last reboot. –c switch used to output CPU utilization; -d switch used to output device utilization –x is followed by the device to query

## **CELERRA DATAMOVER TROUBLESHOOTING:**

**\$server\_config server\_6 “help <topic>” “wins help” “ds help”**

**Note:** Use above syntax to help in finding syntax and useful commands by subject

```
# .server_config server_2 -v "wins help"  
1144097232: SMB: 4: wins addr=<WINS IP addr> name=<host>  
1144097232: SMB: 4: [type=0x00 .. 0xFF] [timeout=n(sec)] [broadcast]  
1144097232: SMB: 4: [who (force name to *, smb_hostQueryName() test)]  
# .server_config server_2 -v "help fc"  
# .server_config server_2 -v "help ds"
```

### **RUNNING LIST ON DATA MOVER:**

```
$ .server_config server_2 -v "ls .etc"  
$ .server_config server_2 -v "ls"  
.etc  
.etc_common  
fs1  
security.evt  
system.evt -----abridged-----
```

### **SERVER MOUNTS:**

```
# .server_config server_2 -v "file mountdisplay"  
# .server_config server_2 -v "unexportall"
```

### **SERVER CPU & RAM ISSUES:**

```
$server_sysstat server_2 [% cpu idle]  
$server_config server_2 -v "printstats cpu" [CPU Histogram of Usage—no longer valid for 4.x code]  
$server_config server_2 -v "cpu display"
```

Note: This command seems to start a counter that logs a CPU Idle entry every (5) minutes

```
$server_config server_2 -v "printstats mem"
```

### **INCREASING I/O PERFORMANCE ON DATAMOVERS:**

Increase Queue Length on Tachyon Fibre Channel Cards in NS600 in NAS 5.1.9.4 or lower:

```
param fcTach device_q_length=256
```

```
param fcTach per_target_q_length=256
```

Allowing Parallel I/O Operations to BackEnd in NAS 5.1.9.4 or lower:

```
param file prefetch=1
```

Note: In some cases, upgrading from 4.2 to 5.1 may result in poor performance if I/O's are random. An ‘enhancement’ was made in 5.1 to address primarily sequential I/O performance. Set following param to return values to 4.1 level:

```
param file prefetch=0
```

### **SERVER PROCESSES:**

```
$/nas/sbin/server_profile server_x [Use to profile CPU Utilization for a DM--dumps to Server Log]
```

Note: NAS 4.0 – 4.2, & <5.0.14 has potential for panicking DM if run against 507 as it looks for dual processors.

### **REFINING SERVER PROFILE FOR SPECIFIC MEMORY RANGES:**

- 1.) #.server\_config server\_x “profile zero”
- 2.) #.server\_config server\_x “profile on start=4a2f68 end=4b8d54” [Run for 3 minutes before running ‘profile off’]
- 3.) #.server\_config server\_x “profile off”
- 4.) #.server\_config server\_x “profile print”
- 5.) #server\_log server\_x -s >slogx\_profile

### **RUNNING SERVER PROFILES ON NAS 5.2 AND HIGHER:**

- 1.) #.server\_config server\_x “profile zero”
- 2.) #.server\_config server\_x “profile on all function” [Resets all functions-let profile run for about 3 minutes]
- 3.) #.server\_config server\_x “profile off”
- 4.) #.server\_config server\_x “profile print”
- 5.) #server\_log server\_x -a -s >slogx.0617

Note: With 5.2, we no longer need to specify “start=...”ranges as exact function names are now printed in the Server Log.

**Caution:** Do not run “profile on function” command on any system prior to NAS 5.2.20.x as it may cause the DM to panic! This **problem has been fixed with NAS 5.3.19.2, 5.2.20.0, & 5.4.17.5**

## CELERRA NFS/RPC STATISTICS:

\$server\_nfsstat server\_2 -rpc | -nfs | -z { zeroes out all stats} -c [client] -s [server] -r [rpc stats only] -n [nfs only]

\$rpcinfo -p server\_2 [Prints port map and RPC programs]

\$server\_config server\_2 -v “printstats rpc” [Various RPC Calls & Statistics]

\$server\_nfsstat server\_2 -s

server\_2 :

Read/write size and alignment distribution (v3):

| size          | read    | crossed | write    | crossed | →Crossed means an I/O exceeded 8k boundary, forcing DM to do extra buffer read |
|---------------|---------|---------|----------|---------|--|
| 1 - 1         | 0       | 0       | 183      | 0       |  |
| 2 - 3         | 3       | 0       | 400586   | 0       |  |
| 4 - 7         | 39      | 0       | 12430    | 0       |  |
| 8 - 15        | 102     | 0       | 18140    | 0       |  |
| 16 - 31       | 100     | 0       | 4281     | 0       |  |
| 32 - 63       | 191     | 0       | 12464    | 0       |  |
| 64 - 127      | 330     | 0       | 94552    | 1125    |  |
| 128 - 255     | 538926  | 9992    | 751543   | 5791    |  |
| 256 - 511     | 490406  | 88200   | 10726464 | 6775    |  |
| 512 - 1023    | 3666    | 0       | 2078183  | 0       |  |
| 1024 - 2047   | 271057  | 15049   | 541734   | 9618    |  |
| 2048 - 4095   | 181237  | 0       | 1110809  | 0       |  |
| 4096 - 8191   | 5210239 | 0       | 1587799  | 32689   |  |
| 8192 - 16383  | 1552627 | 21820   | 407415   | 32985   |  |
| 16384 - 32767 | 79206   | 620     | 60851    | 5       |  |
| 32768 - 65535 | 3024210 | 981     | 3607389  | 1       |  |

**Example:**

# /usr/sbin/rpcinfo -p server\_2 legrep "portmap|nfs|mount"

```
100003 2 udp 2049 nfs
100003 3 udp 2049 nfs
100003 2 tcp 2049 nfs
100003 3 tcp 2049 nfs
100005 3 tcp 1234 mountd
100005 2 tcp 1234 mountd
100005 1 tcp 1234 mountd
100005 3 udp 1234 mountd
100005 2 udp 1234 mountd
100005 1 udp 1234 mountd
100000 2 udp 111 portmapper
100000 2 tcp 111 portmapper
```

## TROUBLESHOOTING NETWORK ISSUES:

\$server\_netstat server\_2 -i [interface stats] -a [udp/tcp sockets] -r [Routing Table] -s [Per protocol stats] -p [Protocol stats for either tcp | udp | ip ]

\$server\_netstat server\_2 -s [checking for Data Mover retransmissions in slow Read performance cases]

\$server\_config server\_x -v “ace ana0 clearstat” [Clears netstat statistics]

\$server\_config server\_2 -v “ana ana0 stat” [Interface settings, Pkts In/Out, Collisions, Errors]

\$server\_config server\_5 -v “trunk trk0 stat” [Trunk statistics on the ports in the trunk]

# .server\_config server\_3 -v “trunk trk0 showcfg”

1139001535: DRIVERS: 4: \*\*\* Trunk trk0: Link is Up \*\*\*

1139001535: DRIVERS: 4: \*\*\* Trunk trk0: Timeout is Short \*\*\*

1139001535: DRIVERS: 4: \*\*\* Trunk trk0: Statistical Load Balancing is IP \*\*\*

1139001535: DRIVERS: 4: Device Local Grp Remote Grp Link LACP Duplex Speed

1139001535: DRIVERS: 4: -----

1139001535: DRIVERS: 4: cge0 10000 37381 Up Up Full 1000 Mbs

1139001535: DRIVERS: 4: cge1 10000 37381 Up Up Full 1000 Mbs

1139001535: DRIVERS: 4: cge3 10000 37381 Up Up Full 1000 Mbs

1139001535: DRIVERS: 4: cge4 10000 37381 Up Up Full 1000 Mbs

**\$ .server\_config server\_2 -v "trunk LACP0 stat"**

**\$ server\_sysconfig server\_2 -v -i trunk02**

**\$server\_config server\_2 -v "ace ace0 stat"**

**\$server\_config server\_x -v "bcm cgeX stat"** [Stats for Broadcom NIC]

**\$server\_config server\_2 -v "bcm cgeN stop"**

**\$server\_config server\_2 -v "bcm cgeN start"**

**\$server\_config server\_2 -v "param tcp"** [Open TCP streams]

**\$server\_config server\_2 -v "printstats ipstat"**

**\$server\_config server\_2 -v "tcp audit"**

**\$server\_config server\_2 -v "tcp connstats"**

**\$server\_config server\_2 -v "tcp stat"**

**\$server\_config server\_2 -v "ip stat"**

**\$server\_config server\_2 -v "ip allstat"** (good summary of IP statistics)

**\$server\_config server\_2 -v "ace ace0 ifstat"**

**# .server\_config server\_2 -v "ifconfig" "ifconfig showall" "ifconfig <name> delete"**

Devices:

fpx0 dmtu=1500, dmac=8:0:1b:43:2f:b

fsn0 dmtu=9000, dmac=8:0:1b:42:24:6a

Interfaces:(4)

166\_89\_24\_90 on fsn0 l=166.89.24.90 n=255.255.252.0 b=166.89.27.255 DNIF UP

mtu=1500, dmtu=9000, vlid=0, mac=8:0:1b:42:24:6a dmac=8:0:1b:42:24:6a

---

Interfaces:(1)

loop6 on loop address= ::1/128 DNIF UP →IPv6 Information

mtu=32768, dmtu=32768, vlid=0, mac=0:0:0:0:0:0 dmac=0:0:0:0:0:0 if\_index=4

**\$server\_config server\_x -v "amgr showarp"**

192.168.2.100 at 0:50:4:47:bb:f6 valid, via el31/el31, use=0 retry histogram ( 0 )

10.241.169.16 at 0:8:74:40:f3:d valid, via fsn0\_30/fsn0, use=1 retry histogram ( 293 1 2 )

10.241.169.2 at 0:4:c1:de:dd:2 valid, via fsn0\_50/fsn0, use=0 retry histogram ( 0 )

**\$server\_config server\_x -v "amgr showroute"**

Host routing table:

127.0.0.1 ---> 127.0.0.1 via loop/loop (permanent), use = 1

Net routing table:

192.168.2.0 ---> 192.168.2.2 via el31/el31 (permanent), use = 0

192.168.1.0 ---> 192.168.1.2 via el30/el30 (permanent), use = 0

10.241.169.0 ---> 10.241.169.30 via fsn0\_30/fsn0 (permanent), use = 2

**# .server\_config server\_2 -v "amgr stat"**

**# .server\_config server\_3 -v "link"**

1111511777: IP: 4: List of linked stream modules:(7)

1111511777: IP: 4: ip <=> el30, handle = 48010, id = 1

1111511777: IP: 4: amgr <=> el30, handle = 48108, id = 2

1111511777: IP: 4: ip <=> el31, handle = 48010, id = 3

1111511777: IP: 4: amgr <=> el31, handle = 48108, id = 4

1111511777: IP: 4: ip <=> fsn0, handle = 48010, id = 5

1111511777: IP: 4: amgr <=> fsn0, handle = 48108, id = 6

1111511777: IP: 4: ip <=> loop, handle = 48010, id = 7

## **TROUBLESHOOTING CIFS ISSUES:**

**# /nas/tools/cifsdebug server\_2** (NAS 5.4 & 5.5, but not aware of this being actively used)

It is possible that this script would fill up the server log of the server if executed, thus causing the server log to wrap around. The script will save the server log prior to running this script so important info is not lost after executing the script. In this script, the severity of KERBEROS, SMB, LDAP, and param NTsec will be altered to capture additional messages if there are errors. At the end of the script, those will be set back to their default values. If you have set those variables to other values prior to the execution of the script, remember to set them back to the desired values. The script will tar up the directory where you specified to save the output files, and it will name the file result.tar. Please submit that file to Engineering for analysis.

Do you wish to continue [yes or no]: yes

Please enter the Celerra CIFS server name: mview\_dm2

Please enter the interface name (not device name such as fxp, cge, etc, but something like 10\_6\_3\_200): foo

Please enter the share name you are connecting: Bogus

Please enter the path name to the directory or file you having issue with (it should look like /filesystemmountpoint/<dir or file>): /m3

Please enter the user name you connect the share with (please don't include domain name): tmatta

Please enter the IP address of the DC: 192.1.4.217

Please enter the Fully Qualified Domain name (eg. test.emc.com): 2k3.pvt.dns

Please enter the Not Fully Qualified Domain name (eg. test): 2k3

\*\*\*\*\*

Please obtain a network trace of the issue for analysis if possible.

Please force a panic if CIFS activities seem to be hung.

Other thing you can do that's not appropriate in this script is to change the DM password with the DC, and see if the process can complete successfully

.server\_config server\_2 -v 16384 "srwpwd compname=<server on Celerra> change"

If you suspect one of the DC is causing an issue, deactivate it by this command:

.server\_config server\_2 -v "pdc invalidate=<IP of the DC>"

To re-enable the DC, do:

.server\_config server\_2 -v "pdc validate=<IP of the DC>"

Material Collection File /tmp/cifs\_materials\_060922\_0908.tar.gz

Please include this file with materials submitted to EMC for problem investigation.

[root@nyip1 tools]# cd /tmp

[root@nyip1 tmp]# ls -la

-rw-r--r-- 1 root root 461674 Sep 22 09:12 cifs\_materials\_060922\_0908.tar.gz

**\$server\_cifsstat server\_3 -full** [-summary; -z to zero] { 'State Info'--SMB Statistics }

**\$server\_config server\_3 -v “printstats cifs”** [Read & Write I/O traffic, and CIFS load statistics]

Cifs Load statistics (for 51155 samples )

Estimate time between 2 samples 5000 ms

Sample duration 255775 s, lasts samples 100 s

|                    | Current | Average | Max | Lasts | Avg | Lasts | Max |
|--------------------|---------|---------|-----|-------|-----|-------|-----|
| Thread usage       | :       | 0       | 0   | 3     | 0   | 1     |     |
| File/dir opened    | :       | 2       | 0   | 9     | 2   | 2     |     |
| Connection in use: |         | 1       | 0   | 3     | 1   | 1     |     |
| User connected     | :       | 1       | 0   | 2     | 1   | 1     |     |
| Stream opened      | :       | 1       | 0   | 2     | 1   | 1     |     |

**\$server\_config server\_3 -v “share”**

→Verbose Share output—use “sharedb info” for summary

Note: Default Share permissions are Everyone FC, though unless manually configured by an Admin from Windows Snapin, a dump of Share permissions may show “No SD or No SD found”—this is code default.

**\$ .server\_config dal1 -v "sharedb info"** →Sharedb output for VDM called “dal1” NAS 5.3.15

1120136996: SMB: 4: Share Database dal1 info (2)

1120136996: SMB: 4:

1120136996: SMB: 4: Mounted on /root\_vdm\_1

1120136996: SMB: 4:

1120136996: SMB: 4: Database Tables:

1120136996: SMB: 4: eb79d904 - @system (1) = 0

1120136996: SMB: 4: eb79db04 - @global (2) = 0

1120136996: SMB: 4: eb79dc04 - **NDAL19P20003** (3) = 0 →Compname

1120136996: SMB: 4: ecd37e04 - NDAL19P23001 (3) = 0

1120136996: SMB: 4: dfe6c104 - NDAL19P20001 (3) = 0

1120136996: SMB: 4:

1120136996: SMB: 4: No active cursors

1120136996: SMB: 4:

1120136996: SMB: 4: 40 Active shares:

1120136996: SMB: 4: C\$@system ( 1) \*n /

1120136996: SMB: 4: IPC\$@system ( 3) \*n /.etc

1120136996: SMB: 4: amildfel\$NDAL19P20001 ( 1) \*n /dal\_fs1/Home/PCS/amildfel

1120136996: SMB: 4: **ISG\$NDAL19P20001** ( 26) \*n /dal\_fs1/Shared/ISG →Sharename, Compname, & Open Conn.

1120136996: SMB: 4: sharri2\$NDAL19P20001 ( 2) \*n /dal\_fs1/Home/ISG/sharri2

1120136996: SMB: 4: obbdalisp01\$NDAL19P20001 ( 1) \*n /dal\_fs1/Home/Generic/obbdalisp01

1120136996: SMB: 4: jnadali5\$NDAL19P20001 ( 1) \*n /dal\_fs1/Home/ISG/jnadali5

**\$server\_config server\_2 -v “sharedb info”** →less verbose than “share”

1115134151: SMB: 4: Share Database server\_2 info (2)

1115134151: SMB: 4:

1115134151: SMB: 4: Mounted on /

1115134151: SMB: 4:

1115134151: SMB: 4: Database Tables:

1115134151: SMB: 4: bc70404 - @system (1) = 0

1115134151: SMB: 4: bd6b504 - @global (2) = 0

1115134151: SMB: 4: bd6b404 - NETAPP1 (3) = 0 →Compname

1115134151: SMB: 4: bd6b304 - NETAPP4A (3) = 0

1115134151: SMB: 4: bd6b204 - NETAPP4 (3) = 0

1115134151: SMB: 4: bd6b104 - NETAPP3 (3) = 0

1115134151: SMB: 4:

1115134151: SMB: 4: No active cursors

1115134151: SMB: 4:

1115134151: SMB: 4: 37 Active shares:

1115134151: SMB: 4: CHECK\$@system ( 1) n /

1115134151: SMB: 4: C\$@system ( 1) n /

1115134151: SMB: 4: IPC\$@system ( 33) \*n /.etc

1115134151: SMB: 4: ReservationsNETAPP4 ( 31) \*n /fs01/data/proj-reservations

1115134151: SMB: 4: ITNETAPP1 (132) \*n /fs02/data/ntprod/deptshares/it →Shows Share,Compname &

### Open Connections

1115134151: SMB: 4: gatewayNETAPP1 ( 48) \*n /fs01/data/gateway

1115134151: SMB: 4: TeleserviceNETAPP1 (307) \*n /fs02/data/teleservices

**\$server\_config server\_2 -v “sharedb asc”** [Dumps DM admin shares and not file system Shares]

1120132392: SMB: 4: 3 Active shares:

1120132392: SMB: 4: CHECK\$@system ( 1) \*n /

1120132392: SMB: 4: C\$@system ( 1) \*n /

1120132392: SMB: 4: IPC\$@system ( 1) \*n /.etc

**\$server\_config server\_3 -v “file mountdisplay”**

Current Mounted File Systems are:

uxfs ro /.etc\_common 160=16 ro

uxfs rw /fs03 535=24 rw,noscan,accesspolicy=MIXED\_COMPAT

uxfs rw /fs06 538=27 rw,noscan,accesspolicy=MIXED\_COMPAT

uxfs rw / 138=5 rw

**\$ .server\_config server\_3 -v "dumpmountdb"** [Dumps NFS mount info]

1115706062: NFS: 4: 020

1115706062: NFS: 4: +---uccpk001 ccp.br.hsbc @ 167.2.6.20

1115706062: NFS: 4: 1 +---/usr/apl/AplAcf

1115706062: NFS: 4: 106

1115706062: NFS: 4: +---ucchx002 ccp.br.hsbc @ 167.2.63.106

1115706062: NFS: 4: 1 +---/usr/apl/Acfhomol

1115706062: NFS: 4: 130

**\$server\_config server\_3 -v “dumpexportdb”** [Dumps NFS exports & security info from DM memory]

1135860645: NFS: 4: Export /fs06 @ c1497004, rootuid ffffffe

Status: valid

Handle: 0x1b 0x1 0x2

worldAccess: RdWr,

needToProcessOptions: FALSE

security = 1,secFlag = 0x2

ROflag = 0,

options: root=172.0.0.0/255.0.0.0

Root Hosts:

172.0.0.0/255.0.0.0,Sec = 1

**\$server\_config server\_3 “export”** [Command will output all exports into Server log without truncation]

**# .server\_config server\_2 -v “shadow readdir .etcPIPE”**

1090012533: CFS: 4: name: samr SAMR key:20

1090012533: CFS: 4: name: lsarpc LSARPC key:52

```
1090012533: CFS: 4: name: srvsvc SRVSVC key:88  
1090012533: CFS: 4: name: svccctl SVCCTL key:124  
1090012533: CFS: 4: name: eventlog EVENTLOG key:160  
1090012533: CFS: 4: name: winreg WINREG key:196  
1090012533: CFS: 4: name: wkssvc WKSSVC key:232  
1090012533: CFS: 4: name: netdfs NETDFS key:268  
1090012533: CFS: 4: name: celerra CELERRA key:304  
1090012533: CFS: 4: name: netlogon NETLOGON key:340
```

**Note:** In some cases, the command will not run because of spaces or other peculiarities in the path—when all else fails, try to determine DOS 8.3 name for the next directory in the path and substitute that name:

```
# .server_config server_7 -v "shadow readdir '\data\userdata\SFISHE3\Sandy's files\Action Plans'"
```

```
1131570657: CFS: 3: shadow fix: getAlternateName failed: NotFound
```

```
# .server_config server_7 -v "shadow readdir \data\userdata\SFISHE3\SANDY'~1\ACTION~1"
```

```
$ .server_config server_2 -v "acl database=.etc dumpslot=1"
```

```
Acl Database of 88 fs locked in shared
```

```
Acl Database of 88 fs unlocked
```

```
Dump of slot 0x1
```

```
-----  
refCount=1 gen=24 size=1 slots 106 bytes CRC=OK
```

```
Owner=USER 0x0 S-1-5-12-1-0
```

```
Group=GROUP GID=0x1 S-1-5-12-2-1
```

```
DACL
```

```
Owner=ALL S-1-1-0
```

```
ALLOWED Flags=3 Mask=1f01ff Rights RWXPDO
```

```
# .server_config server_2 -v "acl dump=/.etc/PIPE/srvsvc"
```

```
Dump of rights of /.etc/PIPE/srvsvc
```

```
===== UNIX =====
```

```
USER 0x0 GROUP 0x1 mode=rw-rw-rw-
```

```
===== NT =====
```

```
acId=5
```

```
Owner=USER 0x0 S-1-5-12-1-0
```

```
Group=GROUP GID=0x1 S-1-5-12-2-1
```

```
DACL
```

```
Owner=ALL S-1-1-0
```

```
ALLOWED Flags=0 Mask=1f01ff Rights RWXPDO
```

```
$ .server_config server_3 -v "acl if=ana0 cache=/mike"
```

```
Acl cache statistiques
```

```
in:1
```

```
hits:0
```

```
miss:1
```

```
stale:0
```

```
# .server_config server_2 -v "acl cache=\.etc" [NAS 5.5]
```

```
Acl cache statistiques
```

```
in:7
```

```
hits:49140760
```

```
miss:8
```

```
stale:0
```

```
# .server_config server_2 -v "acl cache=fs_quota"
```

```
Acl cache statistiques
```

```
in:1
```

```
hits:0
```

```
miss:1
```

```
stale:0
```

```
# .server_config server_2 -v "acl if=ace0 share=shared1" [Dumping Share ACLs]
```

```
Share Shared1 (EAND-FS001)
```

```
===== UNIX =====
```

```
USER ftpuser-emc:0 GROUP 1 mode=rw-----
```

```
===== NT =====
```

```
Owner=USER ftpuser-emc:0 UNIX UID=0x0 'ftpuser-emc':S-1-5-12-1-0
```

```
Group=GROUP 1 UNIX GID=0x1 ".:S-1-5-12-2-1
```

```
Owner=ALL Everyone.:S-1-1-0
```

```
ALLOWED Flags=0 Mask=1f01ff Rights=RWXPDO
```

```
No SACL
```

```
# .server_config server_2 -v "acl share=C$" (NAS 5.5 output)
```

```
Share C$ ()
```

```
===== UNIX =====
```

USER 0x0 GROUP 0x1 mode=rwxr-xr-x

===== NT =====

aclId=0x1

controlSummary = 0x8c04

Owner=USER 0x0 S-1-5-12-1-0

Group=GROUP GID=0x1 S-1-5-12-2-1

DACL

Owner=GROUP GID=0x1 S-1-5-12-2-1

ALLOWED Flags=0x0 Mask=0x1301bf Rights RWX-D-

# .server\_config server\_5 -v "sharedb backup"

Note: Creates “shares.bak” file in /etc/shares/@import

\$ .server\_config server\_3 -v "acl if=ana0 dump=/home/grinch"

# .server\_config vdm1 -v "acl dump=/root\_vdm\_1/fs\_privacy"

1095792316: SMB: 3: Dump with no ThreadCtx

Dump of rights of /root\_vdm\_1/fs\_privacy

===== UNIX =====

USER 0x0 GROUP 0x0 mode=rwxr-xr-x

===== NT =====

aclId=1754

Owner=USER 0x0 S-1-5-12-1-0

Group=GROUP GID=0x0 S-1-5-12-2-0

DACL

Owner=GROUP GID=0x8003 S-1-5-15-79057000-1e31512e-2e75ae2-201

ALLOWED Flags=3 Mask=1200a9 Rights R-X---

Owner=WELLKNOWN S-1-5-12

ALLOWED Flags=3 Mask=1f01ff Rights RWXPDO

Owner=GROUP GID=0x8002 S-1-5-15-79057000-1e31512e-2e75ae2-200

ALLOWED Flags=3 Mask=1f01ff Rights RWXPDO

Note: Syntax for conducting ACL DUMP on VDM with NAS 5.2 & 5.3—no interface required!

# .server\_config server\_2 -v "acl dump=\"/huawei/MIKE HOGAN's ANALYSIS 30AUG06\""

Note: Example of Acl Dump when path contains spaces & apostrophes in directory name

# .server\_config server\_5 -v "cifsThrd"

1113497035: SMB: 4: cifsThrd: checking blocked threads

minimum: 0/96 threads, 5 seconds → Shows if any CIFS threads have been blocked for 5secs or more (none have in this example)

# .server\_config server\_2 -v "THREAD -help" (new Thread command in NAS 5.5.22+)

1151593571: KERNEL: 4: THREAD

THREAD list [service=<service\_name> | pool=<pool\_name> ] [all]

THREAD report [verbose [service=<service\_name> | pool=<pool\_name>]]

THREAD stats

1151593571: KERNEL: 4: Services&Threads dumpOptions

DUMP\_SERVICES =0x00000001

DUMP\_POOLS =0x00000002

DUMP\_THREAD =0x00000004

DUMP\_BLOCKED\_THREADS =0x00000008

PROCESS\_BLOCKED\_STATE =0x00000010

DUMP\_STACK =0x00000020

XML =0x08000000

REPORT =0x10000000

LIST =0x20000000

STATS =0x40000000

INTERNAL\_COMMAND =0x80000000

# .server\_config server\_2 -v "THREAD list all" [output abridged]

1151593594: KERNEL: 4: Service: MAC

| THREAD NAME | STATE | TIME (s) | LAST ACTION             |
|-------------|-------|----------|-------------------------|
| mac00       | idle  | 17.959   |                         |
| mac01       | idle  | 7.967    | volume delete c16t0l0s2 |
| mac02       | idle  | 17.959   |                         |

1151593594: KERNEL: 4: mac00 idle 17.959

1151593594: KERNEL: 4: mac01 idle 7.967 volume delete c16t0l0s2

1151593594: KERNEL: 4: mac02 idle 17.959

1151593594: KERNEL: 4: Service: HTTPD

| THREAD NAME | STATE | TIME (s) | LAST ACTION            |
|-------------|-------|----------|------------------------|
| Httpd0      | idle  | 41.218   | usrmapsvc export       |
| Httpd1      | idle  | 41.190   | usrmapsvc export erase |
| Httpd2      | idle  | 39.941   | usrmapsvc display      |

1151593594: KERNEL: 4: Httpd0 idle 41.218 usrmapsvc export

1151593594: KERNEL: 4: Httpd1 idle 41.190 usrmapsvc export erase

1151593594: KERNEL: 4: Httpd2 idle 39.941 usrmapsvc display

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!

1151593594: KERNEL: 4: NFSD NFSD\_Exec 256 360 0 0  
1151593594: KERNEL: 4: THREAD NAME STATE TIME (s) LAST ACTION  
1151593594: KERNEL: 4: nfs00 idle 83153.168  
1151593594: KERNEL: 4: nfs01 idle 83153.168  
1151593594: KERNEL: 4: nfs02 idle 83153.168  
1151593594: KERNEL: 4: nfs255 idle 83153.176  
1151593594: KERNEL: 4: THREAD NAME STATE TIME (s) LAST ACTION  
1151593594: KERNEL: 4: statd00 idle 607222.761  
1151593594: KERNEL: 4: statd01 idle 607222.761  
1151593594: KERNEL: 4: Service: CIFS  
1151593594: KERNEL: 4: SERVICE NAME POOL NAME THREADS TIME TO PANIC OPTIONS CRITICALS  
1151593594: KERNEL: 4: CIFS SMBVC 3 360 1 0  
1151593594: KERNEL: 4: THREAD NAME STATE TIME (s) LAST ACTION  
1151593594: KERNEL: 4: SMB000 idle for 607221.656 s ctx=0x0 lastCmd=  
1151593594: KERNEL: 4: SMB002 idle for 607221.656 s ctx=0x0 lastCmd=  
1151593594: KERNEL: 4: SMB001 idle for 607221.656 s ctx=0x0 lastCmd=  
1151593594: KERNEL: 4: SERVICE NAME POOL NAME THREADS TIME TO PANIC OPTIONS CRITICALS  
1151593594: KERNEL: 4: CIFS SMB 253 360 0 0  
1151593594: KERNEL: 4: THREAD NAME STATE TIME (s) LAST ACTION  
1151593594: KERNEL: 4: SMB003 idle for 8155.848 s ctx=0xd7ec4804 lastCmd=Echo  
1151593594: KERNEL: 4: SMB016 idle for 8123.070 s ctx=0xd7ec4804 lastCmd=Echo  
1151593594: KERNEL: 4: SMB029 idle for 7205.322 s ctx=0xd7ec4804 lastCmd=Echo

**\$ .server\_config server\_3 "cifs audit"** [User connections on datamover]

**\$ .server\_config server\_2 -v "cifscache"** [Stats on Open Connections & Open Files]

1054585487: SMB: 4: ! Ctxp(0x14cc7604), Name=LNGDAYD-4129872

1054585487: SMB: 4: ! Cnxp(0x98614804), Name=xmatdev, Tid=63

1054585487: SMB: 4: ! miss(0) hit(0) opened(0) maxOpened(0)

1054585487: SMB: 4: Current state: 12 connections 1 file opened

**# .server\_config server\_2 -v "clientOS stats"** →NAS 5.4+

1140743089: SMB: 4: Wellknown OS[16]:

1140743089: SMB: 4: 04:Id=W2K3 OS=Windows Server 2003 3790 LM=Windows Serv

er 2003 5.2 Extra= Grant=1 Count=21

1140743089: SMB: 4: Unknown OS logons: 0 Policy=GrantAll

**# .server\_config server\_2 -v "dfs audit"**

1160754073: SMB: 4: DFS is running

1160754073: SMB: 4: DFS is enabled in the registry

1160754073: SMB: 4: Widelink is not enabled in the registry

1160754073: SMB: 4: Widelink share: none

1160754073: SMB: 4: Dump of Dfs share Programs

1160754073: SMB: 4: -----

1160754073: SMB: 4: Absolute path: /root\_vdm\_2/Programs/Programs

1160754073: SMB: 4: Share type: global share

1160754073: SMB: 4: \* Volvo\2932\_P1X\_Underbody

1160754073: SMB: 4: State: open

1160754073: SMB: 4: Comment:

1160754073: SMB: 4: Timeout: 1800 s

1160754073: SMB: 4: State: DFS\_VOLUME\_STATE\_OK

1160754073: SMB: 4: Target \\uuklutemc002OEM\Volvo\2932\_P1X\_Underbody State=DFS\_STORAGE\_STATE\_ONLINE (2)

**# server\_kerberos server\_2 -ccache** (Data Mover's credential cache)

server\_2 :

Dumping credential cache

Names:

Client: WBCROOT\$@STJOSEPHSWB.COM

Service: WBC-DC.STJOSEPHSWB.COM

Target: HOST/WBC-DC.STJOSEPHSWB.COM@STJOSEPHSWB.COM

Times:

Auth: 04/07/2006 04:38:12 GMT

Start: 04/07/2006 04:38:12 GMT

End: 04/07/2006 14:37:53 GMT

Flags: PRE\_AUTH,OK\_AS\_DELEGATE

Encryption Types:

Key: rc4-hmac-md5

Ticket: rc4-hmac-md5

**# server\_kerberos server\_2 -keytab**

server\_2 :

Dumping keytab file

keytab file major version = 5, minor version 2

-- Entry number 1 --

principal: TESTCIFS\$(NULL)@STJOSEPHSWB.COM

realm: STJOSEPHSWB.COM

encryption type: rc4-hmac-md5

principal type 1, key version: 1

**# server\_kerberos server\_2 -list**

server\_2 :

Kerberos common attributes section:

Supported TGS encryption types: rc4-hmac-md5 des-cbc-md5 des-cbc-crc

Supported TKT encryption types: rc4-hmac-md5 des-cbc-md5 des-cbc-crc

Use DNS locator: yes

**# server\_security server\_2 -i -p gpo**

server\_2 :

Server compname: home

Server NetBIOS: HOME

Domain: stjosephswb.com

Kerberos Max Clock Skew (minutes): 5

LAN Manager Auth Level: Not defined

Digitally sign client communications (always): Not defined

Digitally sign client communications (if server agrees): Not defined

Digitally sign server communications (always): Not defined

Digitally sign server communications (if client agrees): Not defined

Send unencrypted password to connect to third-party SMB servers: Not defined

**\$server\_netstat server\_3 lwc -l** [Open tcp connections]

**\$server\_netstat server\_3 -s -p tcp** [Lingered TCP connections=active]

**\$server\_config server\_3 -v "param tcp"** [Observe TCP maxStreams Value]

**\$server\_config server\_3 -v "pdc dump"** [Domain Controller listing]

**Note:** With NAS 5.4.16.0, Data Mover will open additional secure channel sessions to the DC for other CIFS servers if Clients use NTLMv2 authentication. See AR59648.

**\$server\_config server\_3 -v "pdc trace=1"** [PDC Tracing | trace=0 to turn off]

**\$server\_config server\_3 -v "ntcred user=watson"**

**\$server\_config server\_2 -v "ntcred client=144.20.5.45"**

**#server\_config server\_2 -v "lsarpc if=ace0 trust"**

**Note:** Useful in verifying Domain Trusts & checking domain SIDs. Does not return domain that server is Joined to in the output.

**\$server\_config server\_2 -v "samlogon if=IF\_02 ip=148.95.127.190"** [While this command does not prove that a Secure Channel connection exists, it does prove that the basic domain communications to the DC are accepted and replied to]

1122298722: SMB: 4: Resolved IP=148.95.127.190 to AUFBDS0AUS04 (using smb\_hostQueryName)

**\$server\_config server\_2 -v "ntcred help"**

**Note:** Run this command in NAS 5.3 to get more verbose output

**\$server\_config server\_5 -v "srwpwd dump"** [Checks machine account trust relationship between Server and Domain]

1071157918: SMB: 4: srwpwd V=1 fileV=1 9 records

1071157918: SMB: 4: Entry:11[cel1dm5anb0] @ 13[WIN2KT2.LOCAL]

1071157918: SMB: 4: Time: 0x3ec21417 Now=0x3fd8929e D=0x1167e87

1071157918: SMB: 4: Account: 12[CEL1DM5ANB0\$]

1071157918: SMB: 4: Password:15[8IY9BzafTyhDO7z]

1071157918: SMB: 4: GUID: 3d9da5dd-9b62-41e6-a61a-6932087e601f

1071157918: SMB: 4: Entry:6[SOURCE] @ 6[TS\_NAS]

1071157918: SMB: 4: Time: 0x3ec267b5 Now=0x3fd8929e D=0x1162ae9

1071157918: SMB: 4: Account: 7[SOURCE\$]

1071157918: SMB: 4: Password:14[emcmEDIASERVER]

1071157918: SMB: 4: GUID: 00000000-0000-0000-0000-000000000000

-----abridged-----

**OPTIONS:**

srpwd  
1071157808: SMB: 4: {netbios=...lcompname=...}

1071157808: SMB: 4: {  
    reset lchange lminutes=... ldisplay ldump lremove

### # .server\_config server\_2 -v "srpwd display" | "srpwd dump"

1080226076: SMB: 4: srpwd V=1 fileV=1 2 records

1080226076: SMB: 4: srvPwdUpdate for MOUSE: Next in 657 minutes (period=10020)

1080226076: ADMIN: 4: Command succeeded: srpwd display

**Note:** Above output shows that CIFS Server ‘Mouse’ has a valid Password with the Domain and will be changed in 657 minutes. Notice that the value 10020 is in minutes and translates into one week’s time—this is the default for changing Server passwd with the Domain Controllers with NAS 5.2. Please note that NAS 5.5 rearchitected the entire machine account password implementation, and it is now similar to Windows 2000.

### # .server\_config server\_5 -v "srpwd compname=name change"

**Note:** Do NOT run above command unless specified by TS2 or Eng. as this will change the Server’s password in the Domain and may be disruptive to customer. NAS 5.2 changes its Server password every 7 days.

#### Command Says it Fails, but Actually is Successful in Changing Password:

1080226308: SMB: 3: srpwd unknown action=0

1080226308: ADMIN: 3: Command failed: srpwd compname=mouse

#### Output of “srpwd display” after running password change command:

1080226354: SMB: 4: srpwd V=1 fileV=1 2 records

1080226354: SMB: 4: srvPwdUpdate for MOUSE: Next in **10020** minutes (period=10020)

1080226354: ADMIN: 4: Command succeeded: srpwd display

#### NAS 5.5 Values for Machine Account Password Feature:

**param cifs srpwd maxHistory=2** (stores (2) passwords by default)

**param cifs srpwd.updtMinutes=0** (In minutes--disabled by default)

### # server\_param server\_2 -facility cifs -i srpwd.maxHistory

server\_2 :

```
name      = srpwd.maxHistory
facility_name = cifs
default_value = 2
current_value = 2
configured_value =
user_action = none
change_effective = immediate
range     = (1,10)
description = Number of password kept when password changes
```

### # server\_param server\_2 -facility cifs -i srpwd.updtMinutes

server\_2 :

```
name      = srpwd.updtMinutes
facility_name = cifs
default_value = 0
current_value = 0
configured_value =
user_action = reboot DataMover
change_effective = reboot DataMover
range     = (0,4294967295)
description = Time interval between password changes
```

### \$ .server\_config server\_3 -v "ds dump" [Directory Services dump]

### \$ .server\_config server\_3 -v "ds query=network.lan service=kdc" lpdc lfcl gcl kpcl ldap

### \$ .server\_config server\_3 -v "ds query=network.lan host=dc" lg

### \$ .server\_config server\_2 -v "ds domain=mouse.com getdc"

### \$ .server\_config server\_3 -v "adl list" [Active Directory Listing]

```
1060204910: SMB: 4: ActiveDirectoryServer: @ = 0x73f3c04
1060204910: SMB: 4: domain name: w2khop1.local
1060204910: SMB: 4: netbios domain name: W2KHOP1
1060204910: SMB: 4: domain DN: DC=w2khop1,DC=local
1060204910: SMB: 4: domain controller: w2khop1.w2khop1.local (192.10.0.19)
1060204910: SMB: 4: DC LDAP port: 389
1060204910: SMB: 4: CIFS Server: user/chump
1060204910: SMB: 4: CIFS Server realm: W2KHOP1.LOCAL
```

```

1060204910: SMB: 4: Global Catalog: w2khop1.local
1060204910: SMB: 4: GC DN: DC=w2khop1,DC=local
1060204910: SMB: 4: GC controller: 0
1060204910: SMB: 4: GC LDAP port 3268
1060204910: SMB: 4: Configuration DN: CN=Configuration,DC
=w2khop1,DC=local
1060204910: SMB: 4: next ActiveDirectoryServer 0x0
$ .server_config server_2 -v "dns dump" | stop | start [DNS Information & SRV records]
$ .server_config server_2 -v "dns flush"
# .server_config server_2 -v "dns service dump=w2k.pvt.dns"
1222703888: LIB: 6: CIFS domain:w2k.pvt.dns
1222703888: LIB: 6: Domain GUID:39c57b31-e9de-43e6-a5a3-8893eda99476
1222703888: LIB: 6: Use LdapPing:TRUE
1222703888: LIB: 6: Oversite:FALSE
1222703888: LIB: 6: Forest name:w2k.pvt.dns
1222703888: LIB: 6: Netbios Domain name:W2K Site name:Default-First-Site-Name
# .server_config server_2 -v "dns service domain=w2k.pvt.dns get=kdc"
# .server_config server_2 -v "dns updateAll"
1222703983: LIB: 6: Dns interface(s) will be updated in several minutes
$ .server_config server_2 -v "dns help"
dns domain=<domainname> server=<IP> [server=<IP>] [cache=<size>] [tcpludp]
dns delete=<domainname>
dns flush [cache=<newsize>]
dns dump
dns start
dns stop
dns stop
dns updateAll
dns query ...
dns service ...
where:
domain: domain to add for query
server: DNS server(s) to query
cache: set size of the cache
tcpludp: force tcp or udp mode, defaulted to udp
dump: dump the content of the cache
flush: flush the cache
query: query resource records A, PTR, SOA, NS or SRV
service: query services

```

### **QUERYING FOR ACTIVE DIRECTORY SERVICE RECORDS FROM DNS:**

```

$ .server_config server_2 -v "dns query SRV=_ldap._tcp.dc._msdcs.mouse.com" [Service to gather AD List]
1109859450: LIB: 4: _ldap._tcp.dc._msdcs.mouse.com
1109859450: LIB: 4: Type:SRV TTL=139 s dataCount:1
1109859450: LIB: 4: priority:0 weight:100 port:389 server:mickey.mouse.com
$ .server_config server_2 -v "dns query SRV=_ldap._tcp.mouse.com" [Service for LDAP queries]
1109859648: LIB: 4: _ldap._tcp.mouse.com
1109859648: LIB: 4: Type:SRV TTL=600 s dataCount:1
1109859648: LIB: 4: priority:0 weight:100 port:389 server:mickey.mouse.com
$ .server_config server_2 -v "dns query SRV=_kerberos._tcp.mouse.com" [Kerberos Services]
1109859812: LIB: 4: _kerberos._tcp.mouse.com
1109859812: LIB: 4: Type:SRV TTL=600 s dataCount:1
1109859812: LIB: 4: priority:0 weight:100 port:88 server:mickey.mouse.com
1109859812: LIB: 4: ---
1109859812: ADMIN: 4: Command succeeded: dns query SRV=_kerberos._tcp.mouse.com
$ .server_config server_2 -v "dns query SRV=_kerberos._udp.mouse.com"
1109859820: LIB: 4: _kerberos._udp.mouse.com
1109859820: LIB: 4: Type:SRV TTL=600 s dataCount:1
1109859820: LIB: 4: priority:0 weight:100 port:88 server:mickey.mouse.com
1109859820: LIB: 4: ---
$ .server_config server_2 -v "dns query SRV=_kpasswd._udp.mouse.com" [Passwd Servers]
1109859927: LIB: 4: _kpasswd._udp.mouse.com
1109859927: LIB: 4: Type:SRV TTL=562 s dataCount:1
1109859927: LIB: 4: priority:0 weight:100 port:464 server:mickey.mouse.com
1109859927: LIB: 4: ---

```

1109859927: ADMIN: 4: Command succeeded: dns query SRV=\_kpasswd.\*

**\$ .server\_config server\_2 -v "dns query SRV=\_kpasswd.\_tcp.mouse.com"**

1109859931: LIB: 4: \_kpasswd.\_tcp.mouse.com

1109859931: LIB: 4: Type:SRV TTL=558 s dataCount:1

1109859931: LIB: 4: priority:0 weight:100 port:464 server:mickey.mouse.com

### **QUERY FOR DNS HOST (A) AND REVERSE LOOKUP RECORDS (PTR):**

**\$ .server\_config server\_2 -v "dns query A=merck"**

1109860079: LIB: 4: merck.mouse.com

1109860079: LIB: 4: Type:A TTL=1200 s dataCount:1

1109860079: LIB: 4: 10.241.169.58 (local subnet)

1109860079: LIB: 4: ---

1109860079: ADMIN: 4: Command succeeded: dns query A=merck

**\$ .server\_config server\_2 -v "dns query PTR=10.241.169.16"**

1109860134: LIB: 4: 10.241.169.16

1109860134: LIB: 4: Type:PTR TTL=1200 s dataCount:1

1109860134: LIB: 4: mickey.mouse.com

1109860134: LIB: 4: ---

1109860134: ADMIN: 4: Command succeeded: dns query PTR=10.241.169.16

### **QUERYING WINS SERVERS FOR WINS RECORDS:**

**\$ .server\_config server\_2 -v "wins"**

1054585379: SMB: 4: wins addr=<WINS IP addr> name=<host>

1054585379: SMB: 4: [type=0x00 .. 0xFF] [timeout=n(sec)]

1054585379: SMB: 4: [who (use name=\*, smb\_hostQueryName() test)]

**\$ .server\_config server\_3 -v "wins addr=131.99.75.60 name=compucom type=0x1c"**

**Note:** Wins Address = Wins Server; Name=Domain Name; Type=Wins Record Type [Wins Records for DCs]

**Note:** Plug in different WINS records for more info; 0x1b; 0x1d, etc.

**\$ .server\_config server\_3 -v "wins addr=192.10.4.5 type=0 name=win2kbdc"**

**# .server\_config server\_2 -v "wins addr=3.130.163.239 name=dm2"**

1068609728: SMB: 4: DM2 <00> 0000 (Unique NetBios Name, B Node) Addr=3.130.163.101

1068609728: ADMIN: 4: Command succeeded: wins addr=3.130.163.239 name=dm2

**Note:** Wins lookup of datamover netbios name “DM2” on Wins server

**# .server\_config server\_2 -v "wins addr=3.130.163.239 name=ap3dm2"**

1068609755: SMB: 4: AP3DM2 <00> 4000 (Unique NetBios Name, M Node) Addr=3.130.163.95

**Note:** Wins lookup of netbios name “ap3dm2”

**# .server\_config server\_2 -v "wins addr=3.130.163.239 name=\*"'**

**Caution:** This command can cause DM panic! Resolved with 5.2.16.x and higher.

1068610129: SMB: 4: AP3WINS1APPLGE <20> 4400 (Unique NetBios Name, M Node)

1068610129: SMB: 4: AP3WINS1APPLGE <00> 4400 (Unique NetBios Name, M Node)

1068610129: SMB: 4: USERSAPPLGE <00> c400 (Group NetBios Name, M Node)

1068610129: SMB: 4: USERSAPPLGE <1e> c400 (Group NetBios Name, M Node)

1068610129: SMB: 4: AP3WINS1APPLGE <00> 4400 (Unique NetBios Name, M Node)

**Note:** Wins entries for Domain and Wins Server listed

**\$ .server\_config server\_3 -v "param NTsec logonTraces=6"**

**Note:** With 5.6, you may need to use \$ server\_log server\_x -i option to see trace output in the server log

**\$ .server\_config server\_3 -v "param NTsec logonTraces=3"** [Turn off]

**\$ .server\_config server\_x -v 16384 'lg list'** [Local Group list]

**\$ .server\_config server\_2 "lg remove vs=NTAG3 forever"** [example of removing VDM CIFS server from LGDB]

**\$ .server\_config server\_3 -v "lg help"**

1153482058: LGDB: 4: lg init

1153482058: LGDB: 4: lg check

1153482058: LGDB: 4: lg restore

1153482058: LGDB: 4: lg remove [vs=ntserver]

1153482058: LGDB: 4: lg remove DB

1153482058: LGDB: 4: lg update [vs=ntserver] [force]

1153482058: LGDB: 4: lg list [acclgrplusr] [vs=ntserver][lg=localGroup]

1153482058: LGDB: 4: lg members [vs=ntserver]

1153482058: LGDB: 4: lg groups [vs=ntserver]

1153482058: LGDB: 4: lg flush

```

1153482058: LGDB: 4: lg debug
1153482058: LGDB: 4: lg migsid server domsrc=domain (ifsrc=interface | nbsrc=netbios) domdst=domain (ifdst=interface | nbdst=netbios)
1153482058: LGDB: 4: lg historysid server (if=interface | nb=netbios)
1153482058: LGDB: 4: lg user [add|del|reset][vs=<netbios>]
1153482058: LGDB: 4: lg admin [name=<name>|passwd=<pass>|enable][vs=<netbios>]
1153482058: LGDB: 4: lg freeze
1153482058: LGDB: 4: lg stop
#.server_config server_2 -v "usermapper display" | {dom=} [user=] [group=] [usid=S1-5-...] [gsid=S1-5-...] [uid=][gid=]
$ .server_config server_2 -v "usermapper display"
1047312952: SMB: 4: Usermapper[0] = [167.150.36.56] last access 0
$ .server_config server_2 -v "usermapper dom=biw_nt_master user=supass if=ace0"
1047313358: SMB: 4: UserName0='NT_BIW_MASTER\supass' (0) use=1
S-1-5-15-8331b4b-11eb14cf-792321ae-c7c supass
1047313358: SMB: 4: nt_biw_master\supass UID=33215

```

#### VERIFYING SID CACHE PARAMS:

```
# .server_config server_7 -v "param cifs.sidcache"
```

| Name                            | Location   | Current     | Default     |
|---------------------------------|------------|-------------|-------------|
| cifs.sidcache.globalSidCacheSiz | 0x014353ac | 0x00000076d | 0x000000191 |
| cifs.sidcache.size              | 0x014353a4 | 0x000000035 | 0x000000035 |
| cifs.sidcache.enable            | 0x014353a0 | 0x000000001 | 0x000000001 |

```
$ .server_config server_5 -v "sidcache if=ana0 global dump" or "sidcache if=ana0 dump"
```

```
Sid cache dump
66=>GROUP 2004 Siemens_GG_2.TS_NAS:S-1-5-15-209b5669-55a118ba-7e4b2f8f-9cc
6f=>GROUP 15001 Domain Admins.WIN2KEMC:S-1-5-15-6b635f23-66417ccd-28a68b82-200
6f=>GROUP 2000 Domain Admins.TS_NAS:S-1-5-15-209b5669-55a118ba-7e4b2f8f-200
70=>GROUP 15004 Domain Users.WIN2KEMC:S-1-5-15-6b635f23-66417ccd-28a68b82-201
70=>GROUP 2001 Domain Users.TS_NAS:S-1-5-15-209b5669-55a118ba-7e4b2f8f-201
71=>GROUP 2002 Domain Guests.TS_NAS:S-1-5-15-209b5669-55a118ba-7e4b2f8f-202
```

```
$ .server_config server_5 -v "sidcache global status" [SID Cache Hit ratio]
```

#### NSLOOKUP TOOL:

```
#nslookup
```

```
>set type=SRV
> _ldap._tcp.dc._msdcs.nas.us.dg.com      [Shows DC List, IP Addresses, Port 389]
> _ldap._tcp.nas.us.dg.com                  [Shows LDAP Servers]
> _kerberos._udp.nas.us.dg.com            [Shows list of Kerberos Services & Port 88]
> _kpasswd._tcp.nas.us.dg.com             [Shows list of Password Servers Port 464]
> set type=PTR
> 10.240.16.113
> set type=a
> 10.240.16.113 [or Hostname 'cpc233120']
```

#### NBTSTAT TOOL:

```
c:>nbtstat -n | -a |IP address for remote server [Good for observing registered WINS entries for systems]
```

#### MAC THREADS:

NAS 5.4 and above uses 28 MAC threads (CPU Threads) to perform work and service processes, delivering payload in XML formatted messages. Prior to 5.4, 16 threads were used. MAC threads are synchronous in that a command is issued, and the MAC thread is tied up until it receives a reply.

#### CS-to-DART MAC COMMUNICATIONS—default 20 sec. timeout for MAC requests:

Certain Celerra services and communications use MAC, such as # server\_mount. The default timeout value for these commands is 20secs., which can cause problems during failover & other situations where the file systems are being remounted. See AR81156. Fix will entail extending the timeout values of the ‘build’ function for the HTTP client when mounting large file systems. What happens is that the ‘build’ function takes time, during which the CS is supposed to wait before issuing the Mount command so as to not incur the 20sec. MAC request timeout if it takes >than 20 seconds to mount. The ‘build’ formula is “fs\_size/10 + 120secs”.

#### USING MAC DB TO DETERMINE HWM AND NUMBER OF CIFS THREADS IN USE:

1. Create link to MAC\_DB Tool

**Note:** Run the Mac Tool from the /nas/tools directory

**#/nas/tools/ln -s mac\_db \_mac\_db**

2. Run MAC Commands on Server\_2:

**#/nas/tools/\_mac\_db server\_2**

**macdb> sym/nActiveCifsThreads**

\_nActiveCifsThreads: abef68 [Number of actual CIFS threads in use at time command was run]

\_nActiveCifsThreadsAbsolute: abef6c [Total number of CIFS threads used since last initialization of CIFS]

**macdb> x/d 0xabef68**

00abef68: 00000002 [Decimal counter value represents number of actual CIFS threads in use at time of command]

**macdb> sym/maxActiveCi**

\_maxActiveCifsThreads: abef64 [High Water Mark number of CIFS threads used at one time since last reboot]

**macdb> x/d 0xabef64**

00abef64: 00000053 [Decimal counter represents HWM for simultaneous CIFS threads in use since last reboot]

3. Exiting the Mac interface:

**macdb>quit**

**Other CIFS Counters in NAS Code:**

**# strings - /nas/dos/bin/nas.exe | grep -i cifsthread**

\_maxActiveCifsThreads

\_nActiveCifsThreads

\_nActiveCifsThreadsAbsolute

\_cifsThreadCounts

## **WINS LOOKUP EXAMPLES:**

**# .server\_config server\_2 -v "wins addr=172.16.1.3 name=df type=0x1c"**

1053649027: SMB: 4: DF <1c> 8000 (Group NetBios Name, B Node) Addr=172.17.24.46

1053649027: SMB: 4: DF <1c> 8000 (Group NetBios Name, B Node) Addr=172.17.24.8

**Comment:** Above output are DC records found in Wins database

## **KERBEROS AUTHENTICATION FOR LOGONS** (*rc4-hmac encryption algorithm*):

HMAC is Hash Message Authentication Code—besides hashing a variable input into a set hash output, a Shared secret is also hashed to help authenticate the sender—used in IPSEC.

**TWO TYPES OF WINDOWS LOGONS:** Network Logon and Interactive Logon

### **CLIENT SYSTEM NETWORK LOGON (system):**

The network logon involves the computer during machine startup, or ‘discovery’. After NIC is initialized and provided an IP address by DHCP, the workstation uses DNS query to locate the LDAP service (default-first-site-name). Client then uses LDAP query to locate DC, and negotiates SMB dialect with Server.

Client then requires Secure Channel with DC and negotiates this via the Netlogon service using a NULL SMB session to the DC’s IPC\$ share. An RPC pipe is then opened and the Client & DC exchange authentication challenges to construct a session key. The session key is then updated with the Client’s computer account credentials and forwarded back to the DC. After the DC calculates and compares the credentials to its session key, the Secure Channel is established, but must wait for the Authentication protocol to complete its work. Once the Authentication Package completes its processing (Using NTLM, NTLMv2, or KERBEROS), the DC will reply and send a copy of the SAM or AD to the Netlogon process for the Client.

Once the Secure Channel is setup, client uses DNS query to locate the KDC service to obtain a logon session key, or TGS with the KDC—Kerberos authentication occurs.

Next, Client will connect to IPC\$ share on DC to start DFS referral process, if applicable. Client then does RPC to DC to convert its name into a DN (Distinguished Name). Armed with the DN, the client can then download any updated GPOs that might apply. If NetBIOS is enabled, the browser election process begins. Then, time synchronization occurs. Client finally launches DNS query for SOA and does Dynamic DNS update.

### **CLIENT USER INTERACTIVE LOGON (user):**

Essentially, when a user initiates the Ctrl + Alt + Delete key sequence (aka, SAS—Secure Attention Sequence), and types in their username and password, this information is encrypted and passed to the WinLogon service, which in turn creates a unique logon SID, and calls the LSA service, which in turn calls the appropriate Authentication package, which may already have a cached set of credentials in its local SAM database. Whether the Authentication package retrieves cached User credentials or has to connect to the DC to obtain credentials, the username, hashed password, user SID, and group SIDs are retrieved and stored by the LSA.

The LSA then checks privileges of the user logging in to determine ability to log in, then constructs a default ACL for the User’s primary token. The token is passed to WinLogon for use in creating the User ACL in the DACL of the workstation. WinLogon creates an application and screensaver Desktop to contain a DACL to allow both the User and WinLogon to access.

After Kerberos authentication sequence, an RPC call is made to the DC to convert User's name to a DN. Using the DN, the User's client system uses LDAP to query DC for Group Policies to be applied to User, as well as any DFS referral information.

### **WINDOWS LSA (LOCAL SECURITY AUTHORITY):**

Windows subsystem that manages local security & audit policies & provides user authentication services, such as when a User logs on. LSA RPC calls are named pipes calls between a local security authority and the domain security authority. The LSA stores local policy information in a set of (4) LSA Policy Objects:

- Global Policy Information
- Trusted Domain Information
- User, Group, Local account information
- Private Data such as Server account passwords

### **NAMED PIPES:**

Named Pipes are IPC (InterProcess) Communications between Client & Server as a mechanism for exchanging data between a local process and a remote process on separate computers over the network. For the Windows realm, IPC communications are built into the SMB protocol using DCE/RPC. When in use, Named Pipes are mounted in the \PIPE directory. The SMB IPC\$ share uses Named Pipes when using DCE/RPC services. For an NT 4.0 network, the DOSConnectNamedPipe API is called when using the "net use" command [\\server\IPC\\$](#). NT 4.0 uses SMB and MSRPC services for IPC communications. IPC is used between applications and RPC enables applications to call functions from a local system to a remote system (Remote Procedure Call).

### **MS MSRPC is an implementation of DCE-RPC transports:**

ncacn\_np: DCE-RPC over SMB (named pipes IPC\$ share)—used during Join & Unjoin process

ncacn\_ip\_tcp: DCE-RPC over TCP (Endpoint mapper service (EPM) to discover dynamic TCP points, port 135), intrasite AD replication

### **MSRPC ACTIVE DIRECTORY lsass.exe RPC INTERFACES:**

samr: Security Account Manager RPC service \pipe\lsarpc

lsarpc: Local Security Authority RPC service \pipe\lsarpc

netlogon: netlogon RPC service \pipe\netlogon

drsuci: Active Directory RPC access service—queries to AD using MSRPC vs. LDAP; used AD database replication

frsrpc: AD file replication service between DC's. repadmin.exe & replmon.exe are AD replication admin tools—multi-master replication for sysvol shares

### **OTHER MSRPC INTERFACES:**

epm: endpoint mapper port 135 over TCP, discovers dynamically allocated ports

### **TYPICAL RPC SERVICES FOR MSRPC ACTIVE DIRECTORY INTERFACES:**

SAMR→SAM Security Account Manager RPC service

LSARPC→LSA Local Security Authority RPC service

NETLOGON→Netlogon RPC service

DRSUAPI→AD RPC access service

These services typically use the Endpoint Mapper Service (EPM) to discover TCP ports to use when establishing communications to Named Pipes. ncacn\_np MSRPC transport is used for SMB sessions to IPC\$ share for Joins and Unjoins.

\pipe\lsarpc→lsarpc interface

\pipe\samr→samr interface

\pipe\netlogon→netlogon interface

### **USING NIS CALLS FROM DATA MOVER:**

#### **Typical map names are:**

group.byname

group.bygid

passwd.byname

passwd.byuid

hosts.byname

hosts.byaddr

netgroup

netgroup.byhost

netgroup.byuser

# .server\_config server\_2 -v "yp status"

# .server\_config server\_2 -v "yp match atd.gmeds.com tzn2y3 passwd.byname"

\$ .server\_config server\_2 -v "yp match mynisdomain 5003 group.bygid"

\$ .server\_config server\_2 -v "yp match mynisdomain corp002 hosts.byname"

#### **SAMPLE OUTPUT:**

# .server\_config server\_2 -v "yp match FRL 5068-rcs groupbyname"

1115232118: NETLIB: 4: 5068-rcs::31314:ljones10,bdokter,ddardari,dlenardo,fball,gdamian,harris8,khaley

# .server\_config server\_2 -v "yp match FRL 1044394 group.bygid"

1115231795: NETLIB: 4: 5068-vttest::1044394:jsczepan,khaley,twelsh

\$ .server\_config server\_2 -v "yp match wfg.com wfrantz passwdbyname"

1109684391: NETLIB: 4: wfrantz:\*\*\*\*IX8Q8U/mdTAwE:4208:11:Walter Frantz:/home/wfrantz:/usr/bin/ksh

# .server\_config server\_2 -v "yp match FRL 18237 passwd.byuid"

1115231977: NETLIB: 4: jsczepan:\*LK\*:18237:108191:Jeffry Szczapanski:/ford/srlusr

01/u/jsczepan:/bin/csh

\$ .server\_config server\_2 -v "yp match nas-nis 102793 passwd.byuid"

1121274502: NETLIB: 4: brunoc.umc-users::!:102793:1001:Chris Bruno:/dev/null:/dev/null

1121274502: ADMIN: 4: Command succeeded: yp match nas-nis 102793 passwd.\*

# .server\_config server\_2 -v "yp match 171dc domain=20admins.eng groupbyname"

1129229947: NETLIB: 4: domain=20admins.eng::\*:25001:

**Note:** Must use fully qualified groupname when resolving groups by name if CIFS resolver=0 is set, including the “=20” for a space between names

## **TYPES OF LSARPC QUERIES:**

### **I. RESOLVING NAMES TO SIDS for USERS & GROUPS:**

#### **USER EXAMPLES:**

\$ .server\_config server\_2 -v "lsarpc if=laip1-2a user=thomas"

Finding User SID from Name=SUCCESS

Interface 'laip1-2a' Address=192.1.6.202

User0='2K3\thomas' (0) use=1 nameType=0

S-1-5-15-bfc56af5-d6cf8701-5f67b1a3-4b4 thomas

**UNIX ID=32779 Type=0 →Example of successful SID lookup and mapping by Usermapper**

\$ .server\_config server\_2 -v "lsarpc if=laip1-2a user=mac"

1145628663: SMB: 4: Local Well Known SID set for BUILTIN domain:

1145628663: SMB: 4: SID[2]=20.200.

1145628663: SMB: 5: ExtractDomain>No Domain in reply

1145628663: SMB: 5: ExtractSIDs:Usr='(NULL)\mac' RID=ffffffff U=8 D=fffffff UID=0 T=-1 (LookupSIDs\_Domain)

1145628663: SMB: 5: lookupNames:bad reply=NONE\_MAPPED

1145628663: SMB: 5: sendLookupNames: MsError=LookupNames\_BadReply NTStatus=NONE\_MAPPED

Finding User SID from Name=NONE\_MAPPED

Interface 'elv0' Address=192.1.6.202

Unknown0='(NULL)\mac' (0) use=8 nameType=0

S-?-???????? mac

**UNIX ID=-2 Type=3 →Example of User not found in Active Directory, therefore cannot be mapped**

\$ .server\_config server\_2 -v "lsarpc if=laip1-2a user=thomas"

1145629016: USRMAP: 7: Broadcast internal usermapper addresses

1145629016: USRMAP: 7: Interface broadcast addr #0: 192.168.2.255

1145629016: USRMAP: 7: Interface broadcast addr #1: 192.168.1.255

1145629016: USRMAP: 7: Send the usrmapper broadcast request

1145629018: USRMAP: 4: Broadcast timeout, No answer received

1145629018: USRMAP: 7: Usrmapper broadcast completed: 0/2

1145629018: SMB: 5: Unix user 'thomas' unknown

1145629018: SMB: 5: dom='2K3' (d114)

1145629018: SMB: 5: ExtractSIDs:Usr='2K3\thomas' RID=4b4 U=1 D=0 UID=-2 T=3 (SID Auth\_NoUserPermission)

1145629018: SMB: 5: sendLookupNames: MsError=InvalidName NTStatus=SUCCESS

Finding User SID from Name=SUCCESS

Interface 'laip1-2a' Address=192.1.6.202

User0='2K3\thomas' (0) use=1 nameType=0

S-1-5-15-bfc56af5-d6cf8701-5f67b1a3-4b4 thomas

**UNIX ID=-2 Type=3 →User exists in AD (see SID), but cannot be mapped (Usermapper not running)**

#### **GROUP EXAMPLES:**

\$ .server\_config server\_2 -v "lsarpc if=laip1-2a user=friends"

Finding User SID from Name=SUCCESS

Interface 'laip1-2a' Address=192.1.6.202

Group0='2K3\friends' (0) use=2 nameType=0  
S-1-5-15-bfc56af5-d6cf8701-5f67b1a3-4b5 friends

**UNIX ID=32857 Type=1 →Example of successful mapping**

**\$ .server\_config server\_2 -v "lsarpc if=laip1-2a user=friends"**

1145629620: USRMAP: 7: Broadcast internal usermapper addresses  
1145629620: USRMAP: 7: Interface broadcast addr #0: 192.168.2.255  
1145629620: USRMAP: 7: Interface broadcast addr #1: 192.168.1.255  
1145629620: USRMAP: 7: Send the usrmapper broadcast request  
1145629622: USRMAP: 4: Broadcast timeout, No answer received  
1145629622: USRMAP: 7: Usrmapper broadcast completed: 0/2  
1145629622: SMB: 5: Unix group 'friends' unknown  
1145629622: SMB: 5: dom='2K3' (d524)

1145629622: SMB: 5: ExtractSIDs:Usr='2K3\friends' RID=4b5 U=2 D=0 UID=65534 T=3  
(SIDAuth\_NoGroupPermission)

1145629622: SMB: 5: sendLookupNames: MsError=InvalidName NTStatus=SUCCESS

Finding User SID from Name=SUCCESS

Interface 'laip1-2a' Address=192.1.6.202

Group0='2K3\friends' (0) use=2 nameType=0

S-1-5-15-bfc56af5-d6cf8701-5f67b1a3-4b5 friends

**UNIX ID=65534 Type=3 →Group exists in AD (see SID), but cannot be mapped (Usermapper not running)**

**\$ .server\_config server\_2 -v "lsarpc if=laip1-2a user=monkeys"**

1145629815: SMB: 5: ExtractDomain:No Domain in reply

1145629815: SMB: 5: ExtractSIDs:Usr='(NULL)\monkeys' RID=ffffffff U=8 D=ffffffff  
UID=0 T=-1 (LookupSIDs\_Domain)

1145629815: SMB: 5: lookupNames:bad reply=NONE\_MAPPED

1145629815: SMB: 5: sendLookupNames: MsError=LookupNames\_BadReply NTStatus=NONE\_MAPPED

Finding User SID from Name=NONE\_MAPPED

Interface 'laip1-2a' Address=192.1.6.202

Unknown0='(NULL)\monkeys' (0) use=8 nameType=0

S-?-????????? monkeys

**UNIX ID=-2 Type=3 →Example of Group that does not exist in AD, therefore cannot be mapped**

**# .server\_config server\_2 -v 32768 "lsarpc sid=15-850152b-20db1f76-5bd73fc6=201"**

1145035671: SMB: 5: mapSid Usr='AMERICA\Unknown' use=3 (InvalidAccount)

1145035671: SMB: 5: sendLookupSIDs: MsError=InvalidAccount NTStatus=SUCCESS

Finding User Name from SID=0

Interface '40\_152\_0\_47' Address=40.152.0.47

UserName0='AMERICA\S-1-5-15-850152b-99999999-5bd73fc6' (0) use=8 nameType=0

S-1-5-15-850152b-99999999-5bd73fc6 S-1-5-15-850152b-99999999-5bd73fc6

**UNIX ID=0 Type=3 →Group found in Trusted Domain but no Domain Range has been defined in Usermapper**

## II. RESOLVING SIDS TO NAMES:

**\$ .server\_config server\_5 -v "lsarpc if=ana0 sid=15-52dabe9d-5b5506ca-28a68b82-201"**

Finding User Name from SID=0

Interface 'ana0' Address=192.10.0.103

UserName0='SEINFELD\Domain Users' (0) use=2 nameType=0

S-1-5-15-52dabe9d-5b5506ca-28a68b82-201 Domain Users

UNIX ID=201 Type=1

1068056402: VC: 5: abortCheckWait(smb\_share=0x0)

1068056402: ADMIN: 4: Command succeeded: lsarpc if=ana0 sid=15-52dabe9d-5b5506ca-28a68b82-201

**Note:** Be aware of the syntax used to resolve SIDs—drop the “S-1-5-“ or else the syntax will fail!

## III. RESOLVING BY RSIDS [Relative ID]:

**\$ .server\_config server\_5 -v "lsarpc if=ana0 rsid=201"**

Finding User Name from SID=0

Interface 'ana0' Address=192.10.0.103

UserName0='SEINFELD\Domain Users' (0) use=2 nameType=0

S-1-5-15-52dabe9d-5b5506ca-28a68b82-201 Domain Users

UNIX ID=201 Type=1

1068056363: VC: 5: abortCheckWait(smb\_share=0x0)

1068056363: ADMIN: 4: Command succeeded: lsarpc if=ana0 rsid=201

## IV. VERIFYING DOMAIN TRUSTS BETWEEN DOMAINS:

## # .server\_config server\_2 -v "lsarpc if=136\_141\_250\_52 trust"

```
1071778331: SMB: 4:  
Default values:  
Interface 'elv0' Addr='136.141.250.52'  
1071778331: SMB: 4: Local Well Known SID set for BUILTIN domain:  
1071778331: SMB: 4: SID[2]=20.200.  
1071778331: SMB: 4: Trusted domain 0=RHROOT  
1071778331: SMB: 4: Trusted domain 1=NARNTL  
1071778331: SMB: 4: SID=S-1-5-15-52dabe9d-75932ae1-2b3be507  
1071778331: SMB: 4: Trusted domain 2=EM  
1071778331: SMB: 4: SID=S-1-5-15-66417ccd-74ba50f4-2b3be507
```

### Lookup trusted domains OK

```
1071778331: VC: 5: abortCheckWait(smb_share=0x0)  
1071778331: ADMIN: 4: Command succeeded: lsarpc if=136_141_250_52 trust
```

### NAS 5.5.29 ISSUE WITH TRUSTED DOMAINS THAT ARE QUARANTINED:

NAS code changes behavior so that DART does not query for SIDs from trusted domains that have the quarantine bit set, resulting in access & CIFS outage issues (DM interprets quarantine bit as an unhealthy domain and won't try SID lookups to the domain). Param value was changed to 4. Starting with NAS 5.5.31.4, the default behavior will return to 0. Set the following param to 0 to avoid this behavior:

## # .server\_config server\_2 -v "param cifs lsarpc.queries.trustAttributes=0"

Mask of unsupported and excluded trusted domains for lookupSID queries

### Output from PDC Dump shows Quarantine bit Set:

```
SMB: 4: Trusted domain:domain1.corp [DOMAIN1]  
GUID:00000000-0000-0000-0000-000000000000  
SMB: 4: Flags=0x22 Ix=0 Type=0x2 Attr=0x4 <----- Quarantined  
SMB: 4: SID=S-1-5-15-f2d83cd-591ede04-3fca17b0
```

### V. RESOLVING GROUPS USING LSARPC CALLS:

## # .server\_config server\_2 -v "lsarpc if=10\_43\_3\_46 user='domain admins'"

```
40c6194d DCf2b210c TTNDAPACDC2[APAC] 9 setCurrentDC Ctx=f74d404 Old=0 New=f2b2104  
Finding User SID from Name=0  
Interface '10_43_3_46' Address=10.43.3.46  
UserName0='APAC\domain admins' (0) use=2 nameType=0  
S-1-5-15-53fce379-23e444b6-6b635f23-200 domain admins  
UNIX ID=32770 Type=1  
1086724429: VC: 5: abortCheckWait(smb_share=0x0)  
40c6194d DCf2b210c TTNDAPACDC2[APAC] 8 setCurrentDC Ctx=f74d404 Old=f2b2104 New=0  
1086724429: ADMIN: 4: Command succeeded: lsarpc if=10_43_3_46 user='domain admins'
```

### VI. PERFORMING SAMR CALLS:

## # .server\_config server\_2 -v "samr if=10\_241\_169\_49 user=nas1"

```
1099322602: SECURITY: 3: /etc/group does not exists and NIS not started  
1099322602: SMB: 4: NT_Access_Credential::RequestFromSID User=MOUSE\nas1  
S-1-5-15-42f831d9-66417ccd-28a68b82-45d  
1099322602: SMB: 4: RID=0201 GID=2002 A:7 U:2 ='Domain Users'  
1099322602: SMB: 4: RID=0462 GID=2009 A:7 U:2 ='NAS Users'  
1099322602: SMB: 4: Primary=201 Nb=2 isValid=1 isAdmin=0  
Finding User Groups from UserName=0
```

```
1099322602: VC: 5: abortCheckWait(smb_share=0x0)  
1099322602: ADMIN: 4: Command succeeded: samr if=10_241_169_49 user=nas1
```

### VII. CONDUCTING USER SID-TO-UID/GID LOOKUP ON DM:

## \$ .server\_config server\_7 -v "lsarpc if=ana0 user=nasadmin"

```
1040941811: SMB: 4:  
Default values:  
Interface 'elv0' Addr='192.10.2.30'  
1040941811: SMB: 4: Local Well Known SID set for BUILTIN domain:  
1040941811: SMB: 4: SID[2]=20.200.  
Finding User SID from Name=0  
Interface 'elv0' Address=192.10.2.30  
UserName0='T2DOM3\nasadmin' (0) use=1 nameType=0  
S-1-5-15-74b49ff8-2a1f6232-7ff7ed83-46a nasadmin  
UNIX ID=602 Type=0  
1040941811: ADMIN: 4: Command succeeded: lsarpc if=ana0 user=nasadmin
```

**Note:** DM will always interrogate Domain Controller when this command is used. Usermapper or local passwd/group files must be available for this lookup to succeed. If the User/Group account is valid and a previous mapping has not been made, a new mapping will be issued as a result of this query. Potentially, given a list of User accounts for a Domain, you could script this command to query every account and build a new Usrmapper database. Or, after conducting a 2.0 to 3.0 Usrmapper upgrade, you could forcibly populate SIDs into the database by running lsarpc calls, though this would have a large timeout drawback with accounts that no longer had a valid SID in the Domain—i.e., were deleted.

### **UNSUCCESSFUL USER/GROUP LOOKUP--ACCOUNT DOES NOT EXIST IN DOMAIN:**

# .server\_config server\_5 -v "lsarpc if=ana0 user=nas29"

```
1068044799: SMB: 5: ExtractDomain:No Domain in reply  
1068044799: SMB: 5: ExtractSIDs:Usr='(NULL)\nas29' RID=ffffffff U=8 D=ffffffff UID=0 T=-1 (17)  
1068044799: SMB: 5: lookupNames:bad reply=c0000073  
1068044799: SMB: 5: MsError sendLookupNames=21 NTStatus=c0000073
```

Finding User SID from Name=c0000073

```
1068044799: VC: 5: abortCheckWait(smb_share=0x0)  
1068044799: ADMIN: 3: Command failed: lsarpc if=ana0 user=nas29
```

Error 4020: server\_5 : failed to complete command

**Note:** We cannot retrieve a valid SID for the User or Group, meaning they do not exist in the Domain--see following translation of the NT Status Error:

C:>err c0000073

# for hex 0xc0000073 / decimal -1073741709 :

STATUS\_NONE\_MAPPED ntstatus.h

### **UNSUCCESSFUL USER/GROUP LOOKUP:**

# .server\_config server\_5 -v "lsarpc if=ana0 user=nas23"

server\_5 : commands processed: 0

Error 5: server\_5 : Input/output errorRPC: Timed out

**Note:** If the account is a valid User/Group in the Windows Domain, this error could indicate that Usermapper is not running or is unable to map the account. It might also indicate that local passwd/group files are not resolving the User/Group either.

Use the LSARPC command when troubleshooting User and Group accounts. All LSARPC calls go to the Domain Controllers to be resolved and are never resolved locally from cache!

### **LSARPC LOOKUP WHEN DM RESOLVES GROUP TO DC BUT CANNOT MAP:**

\$ .server\_config server\_2 -v "lsarpc if=128\_206\_8\_41 user='research admins'"

```
1121275488: USRMAP: 3: Usermapper[127.0.0.1] unreachable  
1121275488: RPC: 5: CLIENT::CLIENT 29dda604: Port acquisition failed for Rpc service 536870919 version 3 on host 28.206.8.40  
1121275488: USRMAP: 3: Usermapper[128.206.8.40] unreachable  
1121275488: SMB: 5: Unix group 'research=20admins' unknown  
1121275488: SMB: 5: dom='UMC-USERS' (34c03354)  
1121275488: SMB: 5: ExtractSIDs:Usr='UMC-USERS\research admins' RID=347a U=4 D=0 UID=65534 T=3 (4)  
1121275488: SMB: 5: MsError sendLookupNames=6 NTStatus=0
```

Finding User SID from Name=0

Interface '128\_206\_8\_41' Address=128.206.8.41

UserName0='UMC-USERS\research admins' (0) use=4 nameType=0

S-1-5-15-bfc2566-26bd6a90-49c7643a-347a research admins

**UNIX ID=65534 Type=3**

**Note:** Above output happens when Group cannot be mapped from NIS. “Type=3” refers to unknown sid.

### **TRUSTED DOMAIN GROUPS THAT ARE UNRESOLVABLE DUE TO MISSING RANGE FROM USRMAP.CFG FILE:**

**Server Log shows error trying to map Domain Users group for a particular User:**

```
2006-04-14 10:29:42: SMB: 3: Usr='EMC\c036907':NLreadGroupNames: Domain Users SID not mapped=S-1-5-15-850152b-99999999-5bd73fc6-201 (SID for AMERICA)
```

```
2006-04-14 10:31:08: SMB: 3: Usermapper[127.0.0.1](map2unix): error: 28 for group gid request
```

# .server\_config server\_2 -v 32768 "lsarpc sid=15-850152b-20db1f76-5bd73fc6=201"

```
1145035671: SMB: 5: mapSid Usr='LATINAMERICA\Unknown' use=3 (InvalidAccount)
```

```
1145035671: SMB: 5: sendLookupSIDs: MsError=InvalidAccount NTStatus=SUCCESS
```

Finding User Name from SID=0

Interface '40\_152\_0\_47' Address=40.152.0.47

UserName0='LATINAMERICAS-1-5-15-850152b-20db1f76-5bd73fc6' (0) use=8 nameType=0

S-1-5-15-850152b-20db1f76-5bd73fc6 S-1-5-15-850152b-20db1f76-5bd73fc6

**UNIX ID=0 Type=3 →Shows that Group is not mappable because Domain Range has not been configured in Usermapper**

Use Lsarpc Trust command to see if Domain Users group is a trusted Domain:

```
# .server_config server_2 -v 32768 "lsarpc trust"
```

```
1145035488: SMB: 4: Trusted domain: AMERICA
1145035488: SMB: 4: Flags=0x22 Ix=0 Type=0x1 Attr=0x1000000
1145035488: SMB: 4: SID=S-1-5-15-850152b-99999999-5bd73fc6
```

Internal Usermapper in CONFIG MODE not UNIVERSAL MODE:

```
# .server_config server_2 -v "usrmapsvc info"
```

```
1145034041: USRMAP: 4: Usrmapper service: Enabled
1145034041: USRMAP: 4: Service Class: Primary
1145034041: USRMAP: 4: Database File system: Root
1145034041: USRMAP: 4: Configuration = usrmap.cfg
1145034041: USRMAP: 4: Mapping mode: config
```

No range defined for the AMERICA Domain:

```
# cat usrmap.cfg
```

```
emc,emc.example.com:100000:100001:170000:100001:170000
ogre,ogre.beast.com:600000:600001:670000:600001:670000
indy:700000:700001:870000:700001:870000
_history_sid_range_:900000:900001:970000:900001:970000
```

Use NTCRED to Show Credentials for User:

```
# .server_config server_2 -v 32768 "ntcred user=c036907" |grep 201
```

```
1145035867: SMB: 4: PRIMARY = S-1-5-315b0c1e-88888888-4b0672e8-201
1145035867: SMB: 4: gid=0x493e2 S-1-5-315b0c1e-88888888-4b0672e8-201
1145035867: SMB: 4: gid=0xaee61 S-1-5-15-7a1736d2-77777777-27dd07ca-201 →0xaee61 (700001 decimal) example of Domain
Users group that is mapped
1145035867: SMB: 4: gid=0xffffe S-1-5-15-850152b-99999999-5bd73fc6-201 →0xffffe means Domain Users group for
AMERICA could not be resolved
```

**Comment:** Impact to Users is hard to gauge, but there were reports of performance issues when certain Users accessed the Celerra, perhaps related to timeout delays associated with SIDs that the Data Mover tries to resolve and map, but could not. In this example, the customer had many trusted domains, and Users with Group membership to trusted domains that were not mapped in the usrmap.cfg file. The correct resolution to the problem would be to add the trusted domains to the usrmap.cfg file.

DEBUGGING MEMORY LEAKS ON DATA MOVER:

```
$ .server_config server_3 -v "meminfo"
```

**Note:** Below output represents various memory buckets that DART uses  
1716162599 calls to malloc, 1715768442 to free, 419990763 to realloc

| Size  | In Use | Free   | Total nallocs | nfree       |
|-------|--------|--------|---------------|-------------|
| 8     | 96916  | 4460   | 101376        | 322641187   |
| 16    | 16981  | 2987   | 19968         | 2088927135  |
| 32    | 111315 | 463405 | 574720        | 332848722   |
| 64    | 56401  | 30255  | 86656         | -1616796495 |
| 128   | 82403  | 89885  | 172288        | 788517774   |
| 256   | 5986   | 4734   | 10720         | 365901931   |
| 512   | 16911  | 10017  | 26928         | 2066974133  |
| 1024  | 722    | 110    | 832           | 998396235   |
| 2048  | 4483   | 3629   | 8112          | 331977950   |
| 4096  | 538    | 1150   | 1688          | 19042112    |
| 8192  | 1302   | 0      | 1302          | 312666483   |
| 16384 | 129    | 0      | 129           | 32657       |
| 32768 | 15     | 0      | 15            | 18          |
| 65536 | 37     | 0      | 37            | 37          |

```
$ .server_config server_5 -v "memdata"
```

```
===== Unfreed Objects For Bucket 3 =====
1 times from 1b2f8a: __9logObject+0x86
1 times from 730551: _recvDgram__15smb_winsInitial+0x11d
1 times from 72a8f9: _recvDgram__18smb_browserInitial+0x139
4 times from 19fc18: _addEntry__8dnsCacheRP8dnsEntryi+0x258
4 times from 1985d2: _decodeResourceRecord__16dns_queryMessage17dnsMessageSectionRP8dnsEntry+0x462
```

**Note:** This ouput will list the ‘callers’ or processes that are using the Memory Buckets outputted with “meminfo”

```
$ .server_config server_3 -v "memowners"
```

```
1109625481: KERNEL: 4: Memory_Owner dump.
```

nRegistered = 34, maxOwners = 64  
1109625481: KERNEL: 4: 0 ( 13 frames) No owner, Dump priority 6  
1109625481: KERNEL: 4: 1 ( 119 frames) Free list, Dump priority 0  
1109625481: KERNEL: 4: 2 ( 8 frames) frame table, Dump priority 7  
1109625481: KERNEL: 4: 3 ( 0 frames) Rmode isr vectors, Dump priority 7  
1109625481: KERNEL: 4: 4 ( 48 frames) ROM BIOS + devices, Dump priority 0  
1109625481: KERNEL: 4: 5 ( 0 frames) Sparse dump TOC, Dump priority 7  
1109625481: KERNEL: 4: 6 (1214 frames) paging tables, Dump priority 7  
1109625481: KERNEL: 4: 7 ( 2 frames) Configuration data, Dump priority 7  
1109625481: KERNEL: 4: 8 (3780 frames) image code and data, Dump priority 7  
1109625481: KERNEL: 4: 9 (3262 frames) PooledMemory, Dump priority 6  
1109625481: KERNEL: 4: 10 ( 0 frames) reserved, Dump priority 7  
1109625481: KERNEL: 4: 11 ( 941 frames) Sthread Stack, Dump priority 7  
1109625481: KERNEL: 4: 12 (1536 frames) Message Block List, Dump priority 7  
1109625481: KERNEL: 4: 13 (2189 frames) Buffer Uncached Pools, Dump priority 6  
1109625481: KERNEL: 4: 14 (1041 frames) Small Memory Pool, Dump priority 6  
1109625481: KERNEL: 4: 15 ( 0 frames) Large Memory Pool, Dump priority 6  
1109625481: KERNEL: 4: 16 ( 0 frames) sparse malloc, Dump priority 5  
1109625481: KERNEL: 4: 17 ( 15 frames) Open File Cache, Dump priority 7  
1109625481: KERNEL: 4: 18 ( 64 frames) DNLC Cache, Dump priority 7  
1109625481: KERNEL: 4: 19 ( 0 frames) UFS Cg page allocator, Dump priority 2  
1109625481: KERNEL: 4: 20 ( 0 frames) GIDMAP, Dump priority 5  
1109625481: KERNEL: 4: 21 ( 0 frames) acl command page allocator, Dump priority 2  
1109625481: KERNEL: 4: 22 ( 0 frames) FSTOOLS, Dump priority 5  
1109625481: KERNEL: 4: 23 ( 0 frames) ftspage allocator, Dump priority 2  
1109625481: KERNEL: 4: 24 (13430 frames) malloc heap, Dump priority 6  
1109625481: KERNEL: 4: 25 ( 812 frames) Buffer Cached Descriptors, Dump priority 6  
1109625481: KERNEL: 4: 26 (26625 frames) Buffer Cached Pool, Dump priority 0  
1109625481: KERNEL: 4: 27 (3696 frames) vnode/filenode cache, Dump priority 4  
1109625481: KERNEL: 4: 28 (1790 frames) File System, Dump priority 4  
1109625481: KERNEL: 4: 29 ( 0 frames) vpool page allocator, Dump priority 2  
1109625481: KERNEL: 4: 30 ( 61 frames) ckpt page allocator, Dump priority 1  
1109625481: KERNEL: 4: 31 ( 210 frames) BlockMap page allocator, Dump priority 2  
1109625481: KERNEL: 4: 32 (4360 frames) pax page allocator, Dump priority 3  
1109625481: KERNEL: 4: 33 ( 320 frames) ndmp page allocator, Dump priority 3  
1109625481: ADMIN: 4: Command succeeded: memowners

**Note:** Kernel memory allocation

## **DIRECTORY LISTINGS:**

**\$ .server\_config server\_2 -v "ls /mnt25/IDS\_Citrix/Profiles/LPDtest1/SendTo"**

/mnt25/IDS\_Citrix/Profiles/LPDtest1/SendTo

/mnt25/IDS\_Citrix/Profiles/LPDtest1/SendTo/3B= Floppy (A).lnk

/mnt25/IDS\_Citrix/Profiles/LPDtest1/SendTo/Mail Recipient.MAPIMail

/mnt25/IDS\_Citrix/Profiles/LPDtest1/SendTo/My Documents.mydocs

## **CELLERRA MONITOR--WEB BROWSER INTERFACE:**

**File System Statistics:** File System I/O performance history, HWMarks, Alerts, Inode Usage, etc.

**DataMover Statistics:** DM Configuration, NFS, CIFS, TCP Stats, CPU & Memory Usage, alerts.

**Storage Backend Stats:** Director Performance Stats

**Location of Statistic Databases:**

**/nas/jserver/sdb.old/control\_station/data\_movers/1**

## **SYMMETRIX SCSI STATS PER SERVER:**

**\$ .server\_config server\_2 -v “printstats scsi full”** [Statistics by Controller Bus and Histogram of Activity in ms. Maxqueued are I/O requests from front-end that are queued to CAM layer for processing; Maxpend are I/O requests from back-end awaiting CAM layer.]

**\$ .server\_config server\_2 -v “printstats scsi”** [Most important stat is %Idle on each Controller]

**\$ .server\_config server\_2 -v “printstats filewrite”** [Total aver. usec or secs are most important—2<sup>nd</sup> line from bottom—should be under 20msecs]

**\$ .server\_config server\_2 -v “printstats all”**

**#/nas/symcli/bin/symstat -i 5 -c 9996 -type PORT -dir ALL**

**Note:** Use command to obtain I/O stats on SCSI or Fibre Directors. I/O/sec of 3000 is extremely busy.

## **CAMDISK INFORMATION:**

**# .server\_config server\_2 -v 1024 "camshowconfig"**

CAM Devices on scsi-0:

TID 00: 0:d0+ 1:d1+ 2:d2+ 3:d3+ 4:d4+ 5:d5+

TID 01: 0:d6+ 1:d7- 2:d8+ 3:d9- 4:d10+ 5:d11- 6:d12+ 7:d13- 8:d14+ 9:d15- 10:d16+ 11:d17- 12:d18+ 13:d19- 14:d20+ 15:d21-

TID 02: 0:d22+ 1:d23- 2:d24+ 3:d25- 4:d26+ 5:d27- 6:d28+ 7:d29- 8:d30+ 9:d31- 10:d32+ 11:d33-

**Note:** This command does a probe to verify

**# .server\_config server\_3 -v "camshowconfig c=0"**

CAM Devices on scsi-0:

TID 00: 0:d0+ 1:d1+ 2:d2+ 3:d3+ 4:d4+ 5:d5+ 6:d6+ 7:d7+ 9:d8+

TID 03: 2:d9+ 3:d10- 4:d11+ 5:d12-

TID 05: 10:d13- 11:d14+ 12:d15- 13:d16+ 14:d17- 15:d18+

**# .server\_config server\_3 -v "camgetserial c0t0l0"**

1156366280: CAM: 4: DGC RAID 5 02191D0000A06BCL

1156366280: CAM: 4: APM00030600872 - 001D →Device ID on backend

**# .server\_config server\_2 -v "caminquiry c0t0l0"**

Inquiry Page 00 Evpd 0

0008: 44 47 43 20 20 20 20 20 DGC

0010: 52 41 49 44 20 35 20 20 RAID 5

0018: 20 20 20 20 20 20 20 20

0020: 30 32 31 39 30 30 30 30 02190000

0028: 30 30 42 42 30 41 43 4c 00BB0ACL

0030: 00 00 00 00 00 00 00 00 .....

0038: 00 00 00 3c 02 60 01 9c ...<`..

0040: 08 c0 13 40 00 00 00 00 ...@....

0048: 00 00 00 00 00 00 00 00 .....

0050: 00 00 00 00 00 00 00 00 .....

0058: 00 00 00 00 00 00 00 00 .....

0060: 00 a4 00 10 02 13 00 1b .....

0068: 00 00 00 00 25 00 07 00 ....%...

0070: 00 03 03 00 00 00 00 00 .....

0078: 00 00 00 00 02 00 00 00 .....

0080: 00 00 00 00 00 00 00 00 .....

0088: 00 00 00 00 c0 a8 01 c8 .....

0090: 60 06 01 60 1b 5c 0a 00 `..\..\..

0098: 0d 5f be 2b 37 29 da 11 ..+\_..

00a0: 10 41 50 4d 30 30 30 32 .APM0002

00a8: 33 38 30 31 30 34 30 00 3801040.

00b0: 00 43 58 36 30 30 20 20 .CX600

00b8: 20 20 20 20 20 20 20 20

00c0: 20 41 5f 01 A\_-

**Note:** Useful to see how a specific Target & Lun are resolved by DM

**# .server\_config server\_2 -v "camscanbus"**

1161613699: CAM: 6: Scan All Buses

**# .server\_config server\_2 -v "displayby full"**

---

Name: **logdisk**: NBS1: 1: Disk\_ID: 1 Owner: SP-A blocks: 23068544 paths: 2 IOs pending: 0 IOs in-progress: 0

---

c0t0l0: DGC RAID 5 SN: 600601607eb50900d1206ceed844db11 SP-A IOs: 0

c16t0l0: DGC RAID 5 SN: 600601607eb50900d1206ceed844db11 SP-B IOs: 0

**Note:** Good command to see IO's pending on Disk devices

## **READDING LOGDSK VOLUMES BACK TO DM:**

**\$ .server\_config server\_x “volume disk 1 c0t0l0 disk\_id=1 size=11263”**

## **IMPORTANCE OF LOGDSK NBS1 VOLUME:**

It is important that the logdisk volume has at least one communication path to SPA and SPB, as seen by this error message that resulted during a recent NAS 5.5 Upgrade that runs the automatic post-upgrade healthchecks:

### **PUHC FAILS WITH:**

**# nas\_storage -check -all**

Discovering storage (may take several minutes)

Error 5017: storage health check failed

server\_2 one or more system volumes are HA compromised

# ./dbchk -p

Error: storage system name does not match previous entry, had 'APM00041202945' now have 'APM00041202944' for 'DGC RAID 5 02190600100006NI'.

Error: val does not match previous entry, had '7' now have '31' for 'DGC RAID 502190600100006NT'.

**Resolution:**

Manually add paths for the logdsk volume, or failover and fallback server to rebuild paths

# server\_log server\_3 -s |grep -i logdsk

2007-02-20 14:08:57: ADMIN: 4: Command succeeded: volume disk logdsk c0t0l0

2007-02-20 14:08:57: ADMIN: 4: Command succeeded: volume disk logdsk c16t0l0

**COMPARING VOLUME INFORMATION ON CS TO DM:**

#nas\_volume -i -size id=&lt;id number disk volume&gt;

\$ .server\_config server\_x -v "volume disk &lt;id number disk volume&gt;"

\$ .server\_config server\_2 -v "volume display"

List of all volumes :

| Volume           | In Use           |
|------------------|------------------|
| NBS1 0xdd70ff04  | 2NBS5 0xdd70fe04 |
| 3NBS6 0xdd70fd04 | 391 0x2be1004    |

# .server\_config server\_2 -v "displaybv full"

---

Name: logdsk: 1: NBS1: Disk\_ID: 1 CEL\_ID: APM000716005140000 Owner: SP-A blocks: 23068544 paths: 2 IOs pending: 0 IOs in-progress: 0

---

|                     |  |
|---------------------|--|
| c0t0l0: DGC RAID 5  | SN: 6006016015b21c00e571f3d2331cdd11 SP-A IOs: 0 |
| c16t0l0: DGC RAID 5 | SN: 6006016015b21c00e571f3d2331cdd11 SP-B IOs: 0 |

\$ .server\_config server\_x -v "volume info &lt;id number disk volume&gt;"

\$ .server\_config server\_x -v "volume delete &lt;id number disk volume&gt;"

\$ .server\_config server\_2 -v "volume tree 94" → Volume tree for VolID 94 for file system fs02

\*\*\*\* Hyper Volume 94 : 0x3836a84 Information: \*\*\*\*

Total References:.....0x0002

Total Blocks:.....0x8000000

Bytes Per Block:.....0x0200

Filter Volume:.....Sn94

# Component Volumes:.....0x001

Component Volume List:....93

\*\*\*\* Filter Volume Sn94 : 0xe109d104 Information: \*\*\*\*

Total References:.....0x0003

Total Blocks:.....0x8000000

Bytes Per Block:.....0x0200

\*\*\*\* Hyper Volume ClSn94 : 0x3836184 Information: \*\*\*\*

Total References:.....0x0003

Total Blocks:.....0x8000000

Bytes Per Block:.....0x0200

# Component Volumes:.....0x001

Component Volume List:....93

\*\*\*\* Slice Volume 93 : 0x3836b04 Information: \*\*\*\*

Start Block Offset: 0

Total References:.....0x0004

Total Blocks:.....0x8000000

Bytes Per Block:.....0x0200

# Component Volumes:.....0x001

Component Volume List:....91

\*\*\*\* Stripe Volume 91 : 0x3836b84 Information: \*\*\*\*

Disk Blocks Per Stripe:....0x0010

Total References:.....0x000a

Total Blocks:.....0xd522f00

Bytes Per Block:.....0x0200

# Component Volumes:.....0x002

Component Volume List:....90 83

\*\*\*\* Basic Volume 90 : 0xf763ae04 Information: \*\*\*\* →Represents d14

Total References:.....0x000b

Total Blocks:.....0x6a91780

Bytes Per Block:.....0x0200

\*\*\*\* Basic Volume 83 : 0xf763c604 Information: \*\*\*\* →Represents d7

Total References:.....0x000b

Total Blocks:.....0x6a91780

Bytes Per Block:.....0x0200

### \$ .server\_config server\_2 -v "volume usertree 94"

1174391338: STORAGE: 4: ---- Displaying users of volume 94, referencecount=2 ---

1174391338: STORAGE: 4: UserOfVol=94, UserKind=IO\_OBJECT, userPointer=0x2954, start=0, nBlocks=134217728, referenceCount=1

1174391338: STORAGE: 4: ----- User list complete for volume 94 -----

1174391338: STORAGE: 4: ----- Parents Complete for 94 volume -----

### \$ .server\_config server\_2 -v "volume DBbuild 94"

volume disk 90 c0t117 disk\_id=14

volume disk 90 c16t117 disk\_id=14

volume disk 83 c0t111 disk\_id=7

volume disk 83 c16t111 disk\_id=7

volume stripe 91 16 2 90 83

volume slice 93 0 134217728 91

volume hyper 94 1 93

### \$ .server\_config server\_2 -v "volume info 94"

\*\*\*\* Hyper Volume 94 : 0x3836a84 Information: \*\*\*\*

Total References:.....0x0002

Total Blocks:.....0x8000000

Bytes Per Block:.....0x0200

Filter Volume:.....Sn94

# Component Volumes:.....0x001

Component Volume List:....93

### \$ .server\_config server\_7 -v "volume reversemap 9351 1422594784"

\$ /nas/tools/whereisfs -all (/nas/tools/whereisfs with 5.6.44)

Clariion APM00030600872 Devices and file system(s)

RG FileSystems with total FS count

-----

APM00030600872-0003 Count: 5 fs02 (d14)

APM00030600872-0004 Count: 5 fs02 (d7)

# /nas/bin/nas\_fs -query:\* -format:"\n%-40s%q" -fields:name,disks -query:\* -format:"%s" -fields:name

>/tmp/filesystems.log

## IDENTIFYING COMPONENTS OF META VOLUME:

### # nas\_fs -i m3

volume = v133 →File System m3 is built on top of metavolume v133, which is Volume 133 in the nas\_volume –list

### # .server\_config server\_2 -v "volume info 133"

\*\*\*\* Hyper Volume 133 : 0xddfed604 Information: \*\*\*\*

Total References:.....0x0002

Total Blocks:.....0x3a98000

Bytes Per Block:.....0x0200

# Component Volumes:.....0x003

Component Volume List:....130 131 132 →These are slices that makeup v133, as seen in nas\_volume –list output

### # nas\_volume -list

130 y 1 0 s68 1 133

131 y 1 0 s69 1 133

132 y 1 0 s70 1 133

### # .server\_config server\_2 -v "volume DBbuild 133"

volume disk 87 c0t3l7 disk\_id=9 →d9 volume used, with target & lun info

volume disk 87 c16t3l7 disk\_id=9

volume disk 87 c32t3l7 disk\_id=9

volume disk 87 c48t3l7 disk\_id=9

```
volume hyper 127 1 87
volume slice 130 0 20969472 127
volume disk 86 c0t3l5 disk_id=8 →d8 volume used
volume disk 86 c16t3l5 disk_id=8
volume disk 86 c32t3l5 disk_id=8
volume disk 86 c48t3l5 disk_id=8
volume hyper 128 1 86
volume slice 131 0 20969472 128
volume disk 85 c0t3l3 disk_id=7 →d7 volume used
volume disk 85 c16t3l3 disk_id=7
volume disk 85 c32t3l3 disk_id=7
volume disk 85 c48t3l3 disk_id=7
volume hyper 129 1 85
volume slice 132 0 19501056 129
volume hyper 133 3 130 131 132
# .server_config server_2 -v "camshowconfig"
```

CAM Devices on scsi-0:

TID 00: 0:d0+ 1:d1+ 2:d2+ 3:d3+ 4:d4+ 5:d5+ →Shows which volumes have been diskmarked down this chain, indicated by +

TID 03: 2:d6- 3:d7+ 4:d8- 5:d9+ 6:d10- 7:d11+ →Shows diskmarked volumes (d9+), and non-diskmarked volumes (d10-)

CAM Devices on scsi-16:

TID 00: 0:d12- 1:d13- 2:d14- 3:d15- 4:d16- 5:d17-

TID 03: 2:d18+ 3:d19- 4:d20+ 5:d21- 6:d22+ 7:d23-

## **USING VOLTEST VOLUME PERFORMANCE CAPTURE FOR I/O:**

**Note:** Voltest can provide Write and Read performance statistics on specific meta volumes (such as Celerra File Systems), related to Block I/O size, random or sequential I/O, and number of threads for the Write or Read Operation.

**# .server\_config server\_2 -v "voltest begin 167 rd xfersize=16 sequential threads=10 seconds=5"**

**# .server\_config server\_2 -v "voltest kill 167"**

**Note:** 167 is the ‘meta’ volume id of the file system, rd=read, wr=write, xfersize=I/O block size to use.

### **SAMPLE OUTPUT:**

| Volume | RdOps    | WrOps    | nBytes   | Elapsed  | Bytes/sec | Ops/sec  |
|--------|----------|----------|----------|----------|-----------|----------|
| 167    | 00019f01 | 00000000 | 33e02000 | 00:04:25 | 174061977 | 00021247 |
| 167    | 00019fef | 00000000 | 33dde000 | 00:04:30 | 174035763 | 00021244 |
| 167    | 000194ef | 00000000 | 329de000 | 00:04:35 | 169841459 | 00020732 |

**WARNING:** *Do not use this tool for performance analysis on live production systems as there are a lot of issues—causes panics and systems to go down!*

## **PROFILING CELERRA 'FRONT-END' SCSI PERFORMANCE WITH SYMMETRIX:**

1. **.server\_config ALL -v "printstats scsi reset"** [Zeroes out stats]
2. **.server\_config ALL -v "printstats filewrite reset"** [Zeroes out stats]
3. Rerun “printstats scsi” and “printstats filewrite” command at collection intervals
4. **\$server\_nfsstat ALL -a**
5. **\$server\_sysstat ALL**
6. **\$server\_nfsstat ALL -zero** [Zero out nfsstats after each collection period; Total ncalls are important under RPC Totals]

**Methodology:** Set up a collection interval of 15, 30, or 60 minutes. Idle % is important from SCSI channels of each datamover.

Total average miliseconds for filewrites is important. Normal performance levels should be <20msec. Total “ncalls” for each Server under the “RPC” totals should be divided by the frequency in minutes of your collection period, and divided again by 60 to determine total “nfs” calls to the DataMover’s FrontEnd per second. Important Note here is that several 506/507 datamovers have the capability of overloading a 4.8 Symmetrix.

### **"printstats scsi reset"**

ctr0: maxpend 13, byteCount 521237181480, busy 191547461675546, **idle 227473190304343 = 54%**

ctr1: maxpend 13, byteCount 521344578600, busy 191314482426539, **idle 227704907759274 = 54%**

### **"printstats filewrite reset"**

total 71014793 blocks in 494958755 ticks, **ave 27 msec**

### **"server nfsstat ALL -a"**

**Server rpc:** [Server\_2]

| ncalls | nBadRpcData | nDuplicates | nResends | nBadAuths |
|--------|-------------|-------------|----------|-----------|
|--------|-------------|-------------|----------|-----------|

**341399** 9 0 0 0

**What to Look For on Symmetrix Side:** Engage PSE III or Level 2 Symmetrix

- Channel I/O on Front-End SCSI HBA's
- Total Write Pending on Symmetrix
- Format Pending on Symmetrix
- Cache Fall-Through Time (Read Cache)
- Read/Write Ratio on Channels

### **USING PRINTSTATS TO INVESTIGATE VOLUME PERFORMANCE STATS:**

**\$ .server\_config server\_2 -v "printstats volume full"**

Volume Rd Reqs Wr Reqs Rd Ops Wr Ops

**\$ .server\_config server\_2 -v "printstats volume reset"**

**\$ .server\_config server\_2 "printstats basic\_volume full"**

| Volume | nRead   | nWrite     | Max-IOS    | Busy(ms) | Idle(ms)   | Busy% |
|--------|---------|------------|------------|----------|------------|-------|
| nMerge | MaxQLen | Q-Busy(ms) | Q-Idle(ms) | Q-Busy%  |            |       |
| NBS1:  | 13154   | 412137     | 11         | 1961262  | 1688137696 | 0l    |
| 105    | 2       | 635        | 1690098323 | 0%       |            |       |

**\$ .server\_config server\_2 "printstats basic\_volume reset"**

### **USEFUL PROFILING TOOLS:**

- server\_profile utility
- tcp\_dump utility or other network trace tool
- server\_panic command to induce panic dump for analysis

### **CONDUCTING NETWORK TRACES ON DATAMOVER USING TCPDUMP:** NAS 4.x +

**Step 1. Link server mgr to server tcpdump on CS:**

**# ln -s /nas/bin/server\_mgr /nas/sbin/server\_tcpdump**

**Step 2. Start network capture:**

**# /nas/sbin/server\_tcpdump server\_2 -start ace0 -w /mnt01/tcpdump.log -max 1000** (size in kb) **-s 1514**

Or **# /nas/sbin/server\_tcpdump server\_2 -start fsn0 -w /artwkgrp/artwkgrp/dmp.tst** (NAS 5.4)

Or **# /nas/sbin/server\_tcpdump server\_2 -start cge1 -w /fs\_quota/tcpdump.log -max 1000 -s 1514**

server\_2 :

Packet capturing started.

**Step 3. Monitor capture:**

**# /nas/sbin/server\_tcpdump server\_2 -display**

server\_2 :

Packet capturing OK on device: cge1 , to file: /fs\_quota/tcpdump.log-1  
pkts captured: 38 filtered out: 0 dropped: 0

**Step 4. Stopping the capture:**

**# /nas/sbin/server\_tcpdump server\_2 -stop ace0**

server\_2 :

Packet capturing stopped.

**Step 5. Review trace on CS or download and review using Ethereal, etc.:**

**# /usr/sbin/tcpdump -r /nasmcd/quota/slot\_2/fs\_quota/dump\_nosnap.log -n |more**

**Switches:** -display | -stop ace0 | -host [IP Address only, & don't abbreviate] | -s snaplen [packet size bytes]

**Note 1:**

Try to use -s 1514 whenever possible to prevent truncation of SMB packets

**Note 2:**

\$ /nas/sbin/server\_tcpdump server\_2 -start cge5 -w /tmp/dmp.out → Cannot use tcpdump from DM to write to Control Station  
PacketCapture: failed to initialize asynclog object

### **Engineering Facility to Run TCPDUMP:**

**\$ .server\_config server\_2 -v “netcap action=start device=ace0 filename=/dump/dump.log”** [root of DM]

**\$ .server\_config server\_2 -v “netcap action=display”**

**\$ /nas/sbin/server\_tcpdump server\_2 -display** [Info regarding current tcpdump]

**\$ .server\_config server\_2 -v “netcap action=stop device=ace0”**

**\$/usr/sbin/tcpdump -r /nas/rootfs/slot\_2/dump/dump.log -n** [Reading the capture from Control Station?]

**Note:** TCPDUMP requires the following passwd file entry in /etc/shadow [/etc/passwd] in order to start up:

pcap::!12475:0:99999:7:::

## **OBTAINING TCPDUMP OF DATAMOVER DURING AN NFS OR CIFS STARTUP:**

1. Edit netd file and place following at top of file: netcap action=start device=ana0 filename=/tcpdump.cap
2. Conduct data mover failover
3. Stop tcpdump and analyze (remove netd file entry when done with activity)

## **LINUX CONTROL STATION TCPDUMP:**

Step 1. #/usr/sbin/tcpdump -s 2000 -w /tmp/dump1 host 192.1.5.44 and 192.1.5.23

Step 2. Starts capture between the two nodes; Use “ctrl + c” to stop capture

Step 3. Display Capture: #/usr/sbin/tcpdump -vvex -r /tmp/dump1 |more

## **CAPTURING TRAFFIC ON CS INTERFACE:**

**#/usr/sbin/tcpdump -i eth2 -c 1000 -w /nas/log/toftp2/tcpdump.log &**

## **TESTING I/O PERFORMANCE TO DATAMOVER FROM SUN SOLARIS CLIENT:**

#dd if=/dev/zero of=/mnt/test1 skip=8 bs=1000 count=10000 [Creates write activity]

#/nas/symcli/bin/symdev show 002 [Run command on Control Station to monitor activity]

## **CIFS PERFORMANCE ISSUES & TUNING VARIABLES:**

### **CLIENT SIDE PARAMETERS:**

RW Threads    TCP/IP Parameters    FileSystem Tuning (DNLC & Nodes)    Driver (Jumbo)

**DNLC CACHE (Directory Name Lookup Cache):** Used to speed up name-to-inode mappings

The Directory Name Lookup Cache caches directory/file name pairs whenever a file name is requested from a directory, when doing an NFS Lookup operation. Dart DNLC is expected to be a 100% reliable cache of lookup hits and misses. This means that if an positive match is found, the file must exist and the handle not be stale; if a negative match is found, the file must definitely not exist; and if no match is found, the file may or may not exist, and a further lookup is necessary. The DNLC applies to Unix names only; CIFS lookups went through the Shadow, however, with MPD, we now allow unicode CIFS names to also be cached. The CFS layer must keep the DNLC accurate. This requires that any remove, rename, rmdir, or equivalent operation must make sure to remove the corresponding entry(ies) from the DNLC. It is desired, but not critical, that any lookup add the entry to the DNLC cache. Also, any unmount or freeze must delete all DNLC entries for the corresponding file system. Entries are cached in the DNLC cache as hashed nfs filehandles and are designed to enhance performance as entries can be retrieved from cache much faster than doing another Lookup operation. Negative entries are created if the file lookup fails (ENOENT). Subsequent lookups for the same file name that has a NOENT entry will be dismissed since the name is known to not exist, preventing a duplicate directory Lookup for the same file name. NFS creates will check this cache first to make sure the name does not already exist.

### **DNLC CACHE BUG LEADS TO STALE FILEHANDLES:**

DNLC::add does a lock drop when the lookup fails. That allows two entries to be created at once during multiple lookups for the same entry. However, only one entry is removed by remove() or rename(), the other is left behind, and becomes a stale filehandle. Because the cache mistakenly allows two entries to be created, when NFS directories are deleted and then recreated elsewhere in the same file system tree, the DNLC cache still contains an invalid entry and the ERR\_STALE filehandle condition occurs. AR47569.

### **CLEARING DNLC CACHE ENTRIES FOR SPECIFIC FILE SYSTEM:**

**\$server\_config server\_2 -v "file cleardnlc uxfs fs=115"**

**Nodes:** Max number of Files & Directories that can be opened in a System

**file initialize nodes=65536 dnlc=262144**

### **Ofcache--Open Files Cache:**

→Used for very large numbers of files being opened and reused—default value=15360. Important consideration here is that this value must be smaller than the “nodes” value in the /nas/server/slot\_3/file—long NFS filenames can cause panics and other performance problems:

/nas/server/slot\_3/netd

**nfs start openfiles=15360 nfsd=96**

**# .server\_config server\_x -v "printstats ofcache"** →CIFS small file dirty cache flushing

Open File Cache Statistics

28 access, 10128 checked

8 hits, 20 misses

20 opens, 6 closes, 0 reopens

OfCache hit ratio 28%

→Beginning with NAS 5.6.36, the Open Files Cache is now used by the CIFS protocol, in addition to the traditional NFS user. This enhancement was put into place for CIFS to provide a safeguard against data loss on the CIFS side during system crashes. By default, data is flushed from cache to disk every 10 minutes for open CIFS files. This feature is also called CIFS data trickle sync. See emc208635 for more information and a regression issue resolved at 5.6.44 that caused high CPU utilization and poor system performance.

#### **Disabling of Cache for CIFS:**

**“param cifs ofCache=0”**

#### **Example Of DNLC Server Log Entry:**

##### **UFS: 4: expanding Offset chunk list to 4076 offsets**

Log Entry indicates that Data Mover is busy expanding the dnlc (directory name lookup cache) entries

**Note:** This can happen if directories are being searched in a linear fashion, such as from Tape Backup programs, or Unix find or ls commands. Applications such as viruschecker are also dependent on these parameters.

\$more /nas/server/slot\_2/file

**file initialize nodes=22528 dnlc=65536**

### **EXAMPLE OF FILE NODES & DNLC SYMPTOMS:**

#### **Windows Client Cannot Access Share:**

Error message from Windows client when trying to access the Celerra share: The disk in drive G is not formatted. Do you want to format it now?

#### **Server Log Has Following Messages:**

SHADOW: 3: Couldn't get node from handle Status:3

SHADOW: 4: forceRebuild

**Solution:** Increase nodes & dnlc values per primus emc53040

### **STALE NFS HANDLES ISSUE:**

Due to problem in code, DNLC can have multiple entries for file handles. Deleting a file handle would leave the other handle, resulting in stale file handle error, or complaints that file already exists, etc. AR47720. Fixed 5.3.12.x, 5.2.16.2, though similar issues seen as late as 5.4.15 code.

### **FLUSHING DNLC CACHE ON DATA MOVER:**

**#.server\_config server\_2 -v "file cleardnlc uxfss fs=<fsid>"**

**Note:** Use the cleardnlc command to flush the cache during the time of an access problem to see if access is restored.

**# .server\_config server\_2 -v "file cleardnlc uxfss fs=33"**

1143648779: CFS: 5: DNLC: erased 187023 entries

### **SERVER SIDE PARAMETERS:**

Caching      Storage Load Balancing      Processes (Threads)      Hash Tables

NFS Threads--sometimes adding more will help performance

CIFS Threads--Configure more when many simultaneous Clients connected--Default=32; Not uncommon to see 500 threads

Pathcache--to improve response times when Clients request list of open files from Server

SID Caching--recursive tree copies with ACLs

Oplock Timeouts

Groupcache

### **NETWORK PARAMETERS:**

Jumbo Packets      Switches

### **STORAGE PARAMETERS:**

Stripe Size      Number of Disks      RAID      Caching

### **INCREASING DATAMOVER CIFS THREADS:**

#### **Step 1. Stop CIFS service**

**Step 2. Increase threads: # server\_setup server\_2 -P cifs -o start=200**

**Note:** Default number of threads was 32 [NAS 5.1 increased to 96 threads if resident DM memory allowed]; Maximum number of threads=999. Big problem with this fix is that it is not persistent. If you set threads to a higher level, then stop and start cifs without specifying threads, it will revert back to original value. But since the values are written to the netd and boot.cfg file, reboots will retain the value.

#### **NETD & BOOT.CFG ENTRY:**

**cifs start=96**

## **INCREASING GLOBAL SID CACHING CAPABILITY OF DATAMOVER:** [4.1 Code]

**Purpose:** Used to increase connection performance when many Users are connecting with large numbers of Group memberships. Parameter specifies max number of Global Group SIDs that Celerra caches. Requires less communications with AD Servers to obtain multiple SIDs, especially as it pertains to SIDHistory and multiple Global Group lookups

**param cifs sidcache.globalSidCacheSize=1901** [Should be a prime number]

**Note:** Default Global SID Cache is 401 with all NAS versions above 4.1, but there is a sidcache.size param where the default value is really 53 entries. The cache table will fill up to 53 entries and then use LRU mechanism to delete oldest used entry before adding a new entry.

**# .server\_config server\_2 -v "param cifs sidcache.size"**

cifs.sidcache.size INT 0x010b12fc 53 53 (0,201650) TRUE RESTART 'Per connection SID cache size for ACL mapping'

### **NAS 5.1 +:**

**param cifs sidcache.globalSidCacheSize=401**

**Note:** Parameter name has changed but default is still 401

## **DATA MOVER CACHING:**

Data Mover does not cache User or Group names, but does cache SID GID mappings.

## **INCREASING THE NUMBER OF CIFS SESSIONS PER TCP CONNECTION:**

**Symptoms in Server Log:** "SMB: 3 : addUser: too many sessions (128/128) for FR9203687D"

Datamover is limited to a maximum of 128 CIFS Sessions per TCP connection by default. In very busy CIFS environments, or those that use Terminal Servers (Citrix Metaframe), it's possible to exhaust the number of CIFS Sessions.

**Solution:** Increase "SessUsers" value from default 128 to 512 [New default recently raised to 2048]

**param cifs LanmanServer.SessUsers=2048** [New default 5.2.16.2 & higher]

**Note:** Defined as max number of SMB sessions a single client can open on DM over a single TCP connection. Closest Microsoft Windows setting is MaxWorkItems & MaxMpxCt. Engineering param document implies that there are multiple TCP connections involved, which is wrong. Terminal Server would be one example of a single system using one TCP connection to the data mover and servicing many CIFS sessions from different Users over that connection. Normal PC connection uses no more than two users for the connection (Anonymous + currently logged in user)

**Comment:** Typical environment requiring this change would be Terminal Servers, Citrix Metaframe, or IIS/FTP Servers, where many User SMB Sessions can be multiplexed over the same TCP connection.

## **PARAMETER TO CLEANUP IDLE SMB SESSIONS ON DATA MOVER(see emc115144):**

**param cifs LanmanServer.IdleUserAutoLogoff=720** [minutes] [NAS 5.1.20.4 & 5.2.7.0]

**param cifs LanmanServer.IdleUserAutoLogoffCnxToo=1** [new with 5.3.17.1—not in 5.4 or 5.5]

**Note:** With multiplexed SMB Sessions over a single TCP connection, as in the case of IIS Servers, Citrix MetaFrame, and Terminal Servers, sessions may not cleanup properly, eventually reaching the ceiling in the number of "LanmanServer.SessUsers" sessions that can be opened for a single TCP connection. Though you can increase the Lanman Sessions allowed, the above parameter will allow the data mover to clean-up orphaned SMB Sessions after the stipulated time period, in this case 720 minutes = 12 hours. To try and determine who is failing to breakdown the Session, conduct network trace and look for Session Setups followed by SMB Logoff after the SMB Session is completed. Not known at this time whether this is strictly a MS issue or a code issue. See ARs 35973 & 39039. Starting in 5.3.17, a new param (cifs.LanmanServer.IdleUserAutoLogoffCnxToo) controls whether idle connections are disconnected when idle users are logged off—by default, when a user is idle for the period set by IdleUserAutoLogoff (and has no open files) the user session and its associated connections will be logged off/disconnected.

## **CITRIX OR TERMINAL SERVER ‘CIFS Farm’ PARAMS:**

**param cifs farm.MinUsers (default set to 2)**

**param cifs farm.IdleUserAutoLogoff (default 15 minutes)**

**Note:** New parameters put into place with NAS 5.4.16.1 & 5.3.18.2. See emc115144. No longer sure if LanmanServer.IdleUserAutoLogoff param is valid any longer.

**param cifs farm.MinUsers** (set to 2 by default)

**Note:** Dart will now detect that user is running a multiuser application when more than two connections are established from same client. Default behavior also enforces cleanup of idle cifs connections every 15 minutes.

**# .server\_config server\_2 -v "param fulldescription cifs farm.IdleUserAutoLogoff"**

cifs.farm.IdleUserAutoLogoff 0x0157809c 0x0000000f 0x0000000f

1157121478: ADMIN: 4: Command succeeded: param fulldescription cifs farm.IdleUserAutoLogoff

**Note:** Default is on, set to 15 minutes

**# .server\_config server\_2 -v "param cifs farm.MinUsers"**

cifs.farm.MinUsers INT 0x01578098 2 2 (0,4294967295) FALSE REBOOT 'NA'

# .server\_config server\_2 -v "param cifs LanmanServer.IdleUserAutoLogoff"

cifs.LanmanServer.IdleUserAutoLogoff INT 0x015780a4 4294967295 4294967295 (0,4294967295) TRUE NONE 'Minutes until an idle user is logged off (default:off)'

1157121337: ADMIN: 4: Command succeeded: param cifs LanmanServer.IdleUserAutoLogoff

**Note:** Default is off

# .server\_config server\_2 -v "param cifs farm.IdleUserAutoLogoff"

cifs.farm.IdleUserAutoLogoff INT 0x0157809c 15 15 (0,4294967295) FALSE REBOOT 'NA'

# server\_param server\_2 -facility cifs -info LanmanServer.IdleUserAutoLogoff -v

server\_2 :

```
name      = LanmanServer.IdleUserAutoLogoff
facility_name = cifs
default_value = 4294967295
current_value = 4294967295
configured_value =
user_action = none
change_effective = immediate
range      = (0,4294967295)
description = Minutes until an idle user is logged off (default:off)
detailed_description
```

This parameter sets the number of minutes after which an idle user having no open files is automatically logged off the server. The default behavior is for the user to be logged off when it is explicitly requested by the client or when the TCP connection is reset. Setting this parameter causes the Celerra to free resources associated with user sessions which have been orphaned by the client. Minimum value = 0 Maximum value = 0xFFFFFFF Setting param cifs.LanmanServer.IdleUserAutoLogoff=15 results in the idle user's credential being automatically logged off the server after 15 minutes.

## **TERMINAL SERVICES THINCLIENT MAX OPEN FILES LIMITATION:**

Windows NT multiplexes all file requests sent to a single server over a single virtual circuit. Since ‘SMB’ “shares” this virtual circuit with each additional User, the max number of Open File Handles limit could be reached.

**NT 4.0:** Maximum number of Open File Handles on a single Virtual Circuit=2048

**Windows 2000:** Limit=8192

**Increase this Value with Registry Change:**

HKLM>System>CurrentControlSet>Services>Rdr>Parameters>Add Value>MultipleUsersOnConnection REG\_DWORD Data: 0

**Note:** This situation would only occur if there is a single Terminal Server having all users connect via SMB to a single File Server

## **Enhancing Performance for Very Large CIFS Directory Structures:**

UFS Hashing Table Algorithms and CIFS Path Caching can be tweaked by Engineering, but requires very specific customer tailoring:

**param ufs inoHashTableSize=10069**

**param cifs pathCache=317**

**Note:** pathCache is used to retrieve full pathname of opened file from FID and it must be prime number

## **NEWER PARAM TWEAK SETTINGS:**

**param ufs dirBlkHashSize=64**

**param ufs inoBlkHashSize=6007**

**param ufs inoHashTableSize=15013**

## **Max Number Client TCP Connections to the DataMover:**

Max number of connections/concurrent sessions per Datamover using 2.2.35.4+: 20,000/3000

Max number of concurrent TCP connections prior to 2.2.35.4: 2000

## **CELERRA & TCP RETRANSMISSIONS:** Circumstances in which the Celerra will retransmit packet to Client

1. Fast Retransmits: Done anytime (3) consecutive ACKS are out of order

2. Slow Retransmit: Default wait time is set at (1) second [most NT clients wait much less, 200ms]

## **Calculating TCP Retransmission Rate:**

Conduct Server\_netstat -s over a period of time and observe ‘packets sent’ and ‘data packets retransmitted’.

(Retransmissions/packets sent) \* 100 = % retransmissions

Anything above .1% is considered high. Useful when debugging problems that are related to Clients “reading” from the Data Mover. For Slow Writes to the Data Mover from a Client, examine the Client for TCP retransmissions.

## **TCP STREAMS PERFORMANCE EXAMPLE:**

**Symptom in Server Log:** 1018712517: TCP: 3: OpenStream() fails

### **Resolution:**

With NAS 5.1.20.4 and higher, the following parameter values are now the default. Notice that these params are dependent on each other and must be increased together. The old defaults were 2048 for maxStreams and 9973 for pcbMaxCache.

**param tcp pcbCacheSize=16993** Hex Value=0x00004261 Decimal Value=16993

**param tcp pcbMaxCacheSize=39839** Hex= 0x00009b9f Decimal=39839 [Max cache allocated for TCP connections]

**param tcp maxStreams=16384** Hex Value=0x00004000 Decimal Value=16384

**Note:** Current guidance for systems requiring changes to their current maxStreams values are to apply the values referenced above. NAS 5.3 & higher makes default maxStreams value 65,535 and no longer uses pcbMaxCacheSize or pcbCacheSize values—instead, now use a hashed list instead of cache.

## **TCP MAXSTREAMS ON DATAMOVER: Old Default=2048 New Default=16384**

Can have 2048 concurrent TCP dialogues taking place on the Celerra server [reads—writes—logons]

This is *not* the maximum number of connections that can be in place, however.

**\$server\_config server\_2 -v “param tcp”** [Current Open TCP Streams for a DataMover—“tcp.maxStreams”]

**Note:** TCP maxStreams represents the number of TCP streams on the system. With CIFS, this usually relates to the total number of Clients connecting separately to the data mover—streams can also be used by FTP, HTTP, and NFS access.

## **INCREASING TCP MAXSTREAMS VALUE:**

### **1. Edit Param File for Single Server or Globally:**

**\$/nas/server/slot\_2/param** or **\$/nas/site/slot\_param**

**param tcp pcbCacheSize=16993** Hex Value=0x00004261 Decimal Value=16993

**param tcp pcbMaxCacheSize=39839** Hex Value= 0x00009b9f Decimal Value=39839

**param tcp maxStreams=16384** Hex Value=0x00004000 Decimal Value=16384

**Note:** pcbCacheSize and MaxCacheSize values are replaced with a hash list with NAS 5.3. These values only represent caching capability and is meant as a performance function.

### **2. Reboot Datamover or Failover/Fallback**

**Note:** The above values are not hardware dependent, but may be load-dependent, depending on the growth and size of the Customer's User community. In otherwords, a data mover with 512MB or 3GB of RAM can handle the new defaults with no problem as memory is not allocated until the active connection is made. As usual, validate that the system load is not exceeding the ability of the current hardware to handle the load.

## **TCP KEEPALIVE MECHANISM:**

In some cases it might be necessary to change the default values so that fewer datamover resources are consumed for transient clients. In otherwords, we maintain the “keepalive” TCP connection even after a client closes connection without breaking it down properly.

**/nas/server/slot\_x/param**

**param cifs tcpkeepalive=0xff01030a**

[Default: 01=1 minute keepalive interval message; 03=3 retries before assuming connection lost; 0a=10 minutes between retries]

## **OBSERVING TCP PROTOCOL STATS ON SERVER:**

**\$server\_netstat server\_2 -s -p tcp**

## **DISABLING NFS OVER TCP ON DM:**

1. #vi /nas/server/slot\_x/netd

2. Change following line to read:

nfs start openfiles=15360 nfsd=96 tcp=0

3. Reboot DM

**Note:** If using NFS over UDP, the timeo retransmission timer may estimate values that are unreasonably small

## **TCP/NFS/UFS/CIFS PERFORMANCE TWEAKING:**

**Example:** Changing parameters to allow for increased "fs" memory caching for users pertaining to the "NFS" parameter called "openfiles". Allowing for 16,000 simultaneous User TCP connections and increased "openfiles" capability at two files each:

### **I. TCP PARAMETERS:**

**param tcp maxStreams=8192** [old values—current code sets to max 65535 as default]

### **TCP MAXSTREAMS DEFINED:**

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
Controls the number of concurrent TCP Streams that can be in use at any given instant on the DataMover. Default is 2048. This does not equate to 2048 Users, just 2048 active TCP Sessions. Increases to this value are often necessary when the Server Log records “TCP: 3: OpenStream() fails” or “TCP: 3: Too many connections in the listen queue.”

**param tcp pcbCacheSize=5077** [this param is obsolete in newer NAS versions]

#### **TCP PCBCACHESIZE DEFINED: 401-9973**

Buffer that caches information about current TCP connections. Allows for quick use of ‘lingering’ connections. Default is 401. But, if increased, should be set to a Prime Number. Cache entries receive a ttl, but if entries fill the buffer before old cache expires, then TCP requests are handled in a serial manner, which in turn creates a heavy CPU Utilization, often symptomatic of system resources being starved. Often, “server\_profiling” should be conducted before adjusting this value. Protocol Control Blocks (PCB) Maximum value is 9973.

#### **REVISED MAXSTREAMS AND PCBCACHE TABLES:**

**param tcp pcbCacheSize=16993** Hex Value=0x00004261 Decimal Value=16993

**param tcp pcbMaxCacheSize=39839** Hex Value= 0x00009b9f Decimal Value=39839

**param tcp maxStreams=16384** Hex Value=0x00004000 Decimal Value=16384

**Note:** Current guidance for systems requiring changes to their current maxStreams values are to apply the values referenced above. The above values should always be set concurrently across all three parameters.

#### **LISTEN QUEUE LOGON ERRORS FOR CIFS:**

“Too many connections in port 139’s listen queue”

“Too many connections in port 445’s listen queue”

**Note:** DataMover by default only queues 100 backlogged User Logon requests at any given time. For NT 4.0 environments, an overflow results in above message for Port 139 and for Windows 2000, for Port 445. Increasing following parameter will queue more logon requests—default =64 Hex or 100 decimal:

**param tcp backlog=256**

#### **II. NFS PROTOCOL PARAMETERS:**

##### **a.) Increasing NFS Threads (daemons) Running on Celerra:**

/nas/server/slot\_x/netd

**nfs start openfiles=15360 nfsd=96**

**Note:** Default Values above--Increase nfsd to 250 or 512 if more NFS threads required! Decrease or Increase the number of “openfiles” depending on the Client load for NFS. Values of 44032 are not uncommon if there are a large number of Clients and Openfiles to support.

##### **b.) Increasing NFS Port 2049 Listen Queue:**

/nas/site/slot\_param or /nas/server/slot\_x/param

**param tcp backlog=256**

Default Value = 96; can increase to 256 and 384, but probably should not go higher without specific Eng. recommendation

**Server Log Symptom:** 2002-06-19 18:49:32: TCP: 3: Too many connections in port 2049’s listen queue

**Client Symptom:** Major NFS Timeout logged in /var/adm/messages: “Server Not Responding”

#### **III. UFS/CIFS FILE SYSTEM PARAMETERS: /nas/server/slot\_x/file**

**file initialize nodes=22528 dnlc=65536** [default values]

**file initialize nodes=45056 dnlc=131072**

##### **NODES & DNLC DEFINED:**

Nodes by default is set to 22,528, which tracks open files. Both parameters together are used to track open files. In cases where there are very large numbers of open files, this value, along with DNLC may need to be increased. Where there are linear filesystem searches or backup jobs, this also may impact these parameters. DNLC by default is set to 65,536. Please note that these values are constantly changing and are different in the latest NAS code and hardware models.

##### **NAS VERSION 5.5:**

# cat file

**file initialize nodes=256000 dnlc=384000**

**Note:** DNLC=Directory Name Lookup Cache

#### **FILE NODES AND DNLC VALUES ARE DETERMINED DURING SETUP SLOT:**

DM Memory > than 512 sets values to "file initialize nodes=65536 dnlc=262144"

DM Memory = 512 sets values to "file initialize nodes=22528 dnlc=65536"

DM Memory < than 512 sets values to "file initialize nodes=10240 dnlc=8192"

**Note:** Normally, memory values are determined during setup\_slot using server\_sysconfig -P at getreason code 4 or 5, but on occasion, if DM is still at state “3” then DM will be set to minimize parameters. This may occur during setup\_slot and also during replacement of nas.exe files. One side effect is that virus checker may be impacted—see “vnodes”.

#### **IV. LOST DELAYED WRITES OVER NETWORK:**

This error is commonly seen when using applications over wan networks, such as Oracle Databases, SQL Server, Terminal Services and ThinClients, and Internet Information Server (IIS), where network latency, and resource issues exist between Client/Server.

**Windows Client PopUp Message:** "Windows was unable to save all the data for the file xxx. The data has been lost. This error may be caused by a failure of your computer hardware or network connection. Please try to save this file elsewhere."

#### **INCREASING CELERRA TCP CONGESTION WINDOW SIZE:**

**param tcp sndcwnd=16384 | 32768 | 65536 | 131072**

**Note:** Default=0, meaning that we use the size of receive window from client

Congestion Window is the max amount of data Celerra can send to client without waiting for ACK

**param tcp sndcwnd=1**

**Note:** Setting value to 1 translates into min. of 2 TCP segments of 1460 bytes—useful for troubleshooting some client issues

#### **REDUCING WINDOW SIZE TO AVOID CONGESTION WINDOW TCP TIMEOUTS:**

**param RCP tcpwindow=128000** (do not know if this syntax is correct)

**Note:** Typical values would be 128 or 64k, purpose is to try and prevent the Congestion Window being hit as this slows TCP down due to timeout, retransmit, and restart values. Seems to be most applicable to Remote IP Replication environments. Current NAS 5.4 default is 0.

#### **RCP SESSION TABLE ISSUE:**

An issue with RPC where we get connections half opened that eventually fill the RCP session table--for revisions prior to 5.4.18, the solution is to reboot the DM. With 5.4.18, these connections are cleared by the DM.

**Possible Workaround is to Stop & Restart RCP Daemon:**

**\$server\_config server\_x -v "rcpstop"**

**\$server\_config server\_x -v "rep"**

Sample Server Log Message→2005-10-02 02:10:58: RCPD: 4: RCP connection from 10.17.155.71:38013

**\$ .server\_config server\_2 -v "volmcast display interface"**

1154954222: RCPD: 4: ip:192.1.6.202

1154954222: RCPD: 4: ip:192.1.6.204

#### **EXAMPLE OF TUNING ON MS CLIENTS & CELERRA SERVERS:**

##### **WINDOWS 2000 SERVERS:** [Changes made based on Microsoft TechNet]

MaxCmds--default setting is 50 [Max value is 65,535]; Value increased to 2048

maxMpxCt--default is 125 [Max value is 64,535]; Value increased to 2048

maxWorkItems--default 500 [Max value is 65,535]; Value increased to 8192

##### **CELERRA NAS SERVERS:**

**Windows Value=MaxCmds:**

**param cifs LanmanServer.MaxMpxCt=2048** Celerra Default Value=128; Increased to 2048

**Note:** This parameter appears to be the same as the cifs maxMpxCount param

**Windows Value=MaxMpxCt:**

**# server\_param server\_2 -facility cifs -info maxMpxCount**

server\_2 :

```
name      = maxMpxCount
facility_name = cifs
default_value = 127
current_value = 127
configured_value =
user_action = restart Service
change_effective = restart Service
range     = (5,2048)
description = Unacknowledged command count max (default:127)
```

**# server\_param server\_2 -facility cifs -modify maxMpxCount -value 2048**

**param cifs maxMpxCount**

**Note:** Use the server\_param command to change the default value from 127 to 2048. Please note that the CIFS service must be restarted before the new settings go into effect. The maxMpxCount default is 127 and represents max number of CIFS client commands allowed without acknowledgement by DM. Value is negotiated via SMB sessions with the Client [from the Server] and then enforced by the Client Redirector. Whenever the “maxMpxCt” values are changed, the “MaxWorkItems” must also be increased by a value of 4x the maxMpxCt value.

#### **Windows Value=MaxWorkItems:**

# server\_param server\_2 -facility cifs -info LanmanServer.SessUsers

server\_2 :

```
name      = LanmanServer.SessUsers
facility_name = cifs
default_value = 2048
current_value = 2048
configured_value =
user_action = none
change_effective = immediate
range     = (1,16383)
description = Maximum sessions from a single client (default:128)
```

**param cifs LanmanServer.SessUsers=2048** (old default 128, new default 2048)

**Note:** Represents maximum number of TCP sessions a Client can open with the DM

#### **DEFINITIONS:**

**MaxWorkItems:** Max. number receive buffers that Server Service allocates. Symptoms are that buffer limit is reached and TCP must initiate Flow Control--this decreases network performance. Extreme cases, no new client connections will be allowed.

MaxMpxCt is related to MaxWorkItems—together they influence the number of outstanding concurrent network requests for SMB Connections that a NAS Device or Server can service at any given time.

**MaxMpxCt:** Max. number of simultaneous client requests to a Server. During SMB negotiation, this value is passed to Redirector Service, which must enforce the limit. Value needs to be raised with maxWorkItems. See TechNet: Q271148; Q191370; Q232890

#### **Increasing Values on Windows 2000 Clients/Servers:**

**MaxCmds:** Run: regedit32>HKEY\_LOCAL\_MACHINE>System>CurrentControlSet>Services>lanmanworkstation>parameters>MaxCmds

**MaxMpxCt:** Run: regedit32>HKEY\_LOCAL\_MACHINE>System>CurrentControlSet>Services>lanmanserver>parameters>MaxMpxCt

#### **Monitoring Windows 2000 Systems (MaxCmds & MaxMpxCt):**

Monitor values by using Performance Monitor and adding the "Current Commands" counter for the "Redirector"--this allows tracking of number of concurrent outstanding SMB network requests on an IIS Server.

#### **THEORETICAL ETHERNET CAPACITY (Mbits vs. Mbytes):**

10Mbps → 1.25MBps

100Mbps → 12.5MBps

1000Mbps → 125MBps

#### **THREE BASIC PERFORMANCE FACTORS FOR TCP/IP:**

I. TCP/IP WINDOW SIZE—limited to 32-bits when using the “Sliding Scale”

II. TCP/IP PACKET-LOSS RECOVERY MECHANISMS

III. ROUND-TRIP CALCULATIONS TO DETERMINE RTO (Retransmission Timeout Interval)

#### **I. TCP/IP WINDOW SIZE:**

Original TCP Window Size limited to what the TCP Header (16-bits, or max 65K bytes) could report

New ‘Sliding Window’ extension expands this limitation to 32-bit Windows, used only in SYN segments during Session Setup.

#### **MICROSOFT WINDOWS TCP WINDOW SIZE:**

Win9x – NT 4.0 = 8760

Win2k = 16k

Win2k SP3, XP SP1 = 65k

#### **TCP SLIDING WINDOW:**

**Note:** TCP Sliding Window is the amount of data that the Sender can expect to send before receiving an acknowledgement from Receiver. The ‘Sliding Window’ serves as a means for flow control & congestion control. The “max receive window” is the amount of “receive buffer space” that a Host has.

**Example:** If Sliding Window is 64Kb, Sender must wait for ACK, after sending 64Kb of data, prior to sending more data.

#### **Port Numbers:**

Endpoint ports that combined with TCP form a “connect-oriented” socket

Well-Known Ports—0-1023

Registered Ports—1024-49151

Dynamic/Private Ports—49152-65535

**ACK:**

Acknowledge—Receiver sends an ACK that equals a sequence number plus amount of data in bytes. TCP Sessions agree to initial sequence numbers.

## **II. TCP/IP PACKET-LOSS RECOVERY MECHANISMS:**

**Intro:** TCP/IP performance is dictated by the “transfer rate” (bandwidth=mbps) and round-trip delay (network latency) of packets over the network. Traditional TCP/IP networks used a “slow-start” mechanism whenever more than (1) packet was lost in a single Window Transmission—this necessitated a retransmission timeout and slow re-start & greatly affected data in pipeline.

New algorithms, called Fast Retransmits/Fast Recovery, can better handle multiple packet loss per Window, using SACK (Selective Acknowledgements—Receiver keeps Sender informed of status of Receive Buffer Queue).

### **(4) BASIC TCP ALGORITHMS:**

#### **SLOW START:** (SSTHRESH) All systems should now implement ‘slow start threshold size’ (ssthresh) by default

Seeks to avoid flooding router buffers and slowing internet links by ramping up new packets at a rate determined by acknowledgments returned. From Sender’s machine, this is the ‘congestion window’ [cwnd], as opposed to the ‘sliding tcp advertised window’ from the Receiver. So, in theory, each time an ACK is received, the congestion window index can be increased, but never greater than the “advertised window”. The “cwnd” window starts out at a value of “one segment” for ‘slow start.’ So, capacity is reached when intermediate routers begin discarding packets and then the congestion window value is lowered decrementally.

#### **CONGESTION AVOIDANCE:** (CWND) Method for dealing with packet loss

Timeouts and duplicate ACKs are prime indicators of packet loss on the internetwork. Basically, can occur whenever a large data pipe flows into a smaller pipe, or when there are multiple streams going to a Router, all meaning that a bottleneck & packet loss are occurring. In reality, both “slow start” and “congestion avoidance” are implemented in parallel, although they are different algorithms. Bottomline, whenever timeouts are detected, TCP must slow down using either congestion avoidance or “slow start” to ramp back up. If “cwnd” is less than or equal to ssthresh, TCP is in ‘slow start’. Typically, if timeouts occur, the “cwnd” is set to one segment, which is the same as a “slow start”.

**FAST RETRANSMIT:** Theory here is that duplicate ACKs are being sent by one end, which may be indicative of lost segments or reordered packets. In general, up to (3) consecutive ACKs can be received before a retransmission is requested, and usually done so prior to the expiration of a ‘slow retransmission’ timer.

**Note:** Windows NT systems use 200-250ms RTO. Windows 2000 & Solaris use much smaller RTO retransmission timers of around 50ms. Celerra uses a 500ms fastRTO timer and does not use SACK (Selective Acknowledgements). Therefore, on networks where multiple packets are being dropped for same TCP window, Celerra performs slow retransmit [i.e., waits before sending retransmit request to Client, giving appearance that Celerra is slower to respond].

**FAST RECOVERY:** An algorithm that allows TCP to use ‘congestion avoidance’ instead of ‘slow start’ whenever duplicate ACKs are received and requested for retransmits. As analogy, the data stream only slows down rather than stops, picking up steam again as things settle out on the network. Fast Retransmit (RTO) & Fast Recovery usually work together. Generally, during retransmits, the congestion avoidance rate goes down to  $\frac{1}{2}$  of what the initial rate was—then builds up from there.

## **III. ROUND-TRIP CALCULATIONS TO DETERMINE RTO**

TCP/IP implements ‘reliable’ delivery because it retransmits segments that are not acknowledged before a RTO value expires. RTO is determined by estimating mean and variance of a measured RTT (Round-Trip Time) between Sender & Receiver.

### **CELLERRA TCP FAST RETRANSMIT:**

Fast retransmissions are completely separate and unique from the Celerra fastRTO feature, and are triggered by duplicate ACKs or RTOs which can be triggered when receiving TCP connection times out while waiting for a segment ACK. FastRTO only determines whether the retransmission timeout value is subject to 3 clock ticks of 500ms or a single clock tick of 500ms.

### **CELLERRA fastRTO (Retransmission Timeout) vs. WINDOWS NT/2000:**

**Symptom:** Celerra appears to function slower than Windows 2000/NT 4.0 Servers with respect to TCP Fast Retransmissions for lost or out-of-order network packets. Over large WAN's, this difference in RTO functionality may become noticeable. Celerra's default RTO (Retransmission Timeout) is 1.5secs—setting fastRTO to 1 changes timeout value to .5secs, more closely resembling Windows.

**Cause:** Celerra Implements TCP Retransmits (RTO) differently than native Windows NT or 2000 Operating Systems by using a Slow TCP Clock mechanism of 500ms before implementing a retransmit request to the Client. Windows implements a Fast TCP Clock of 200ms before requesting retransmits for "lost" or "out-of-order" packets. Both Windows & Celerra will request retransmission after receiving (3) or more duplicate ACKs, because it is assumed that (3) or more duplicate ACKs indicates packet loss. In some cases involving numerous dropped packets, Celerra performance may be affected and require following RTO tuning:  
/nas/site/slot\_param or /nas/server/slot\_x/param

**param tcp fastRTO=1**

**Note:** Changes Retransmission Timeout value from 1.5secs to .5secs. Celerra uses two clocks for TCP, a 200ms clock and a 500ms clock. The 200ms clock is used for TCP Delayed ACK, while the slow 500ms clock is used by the fastRTO feature. When changing RTO from 0 to 1, Celerra will now use a single slow clock of 500ms before initiating retransmission request for lost or missing packets. At its default value, Celerra uses (3) slow clock ticks of 500ms before RTO (hence the 1.5sec delay).

**OTHER O/S BEHAVIOR:**

Linux & Solaris tend to use fastRTO under 100ms

**SCSI NEGOTIATION/CONFIGURATION:**

**\$server\_config server\_2 -v "chimconfig" [SCSI negotiation]**

**Output:**

TID 0: Ultra = 1 (1), WIDE, speed 200, offset 8 Tagged 1 Qdepth 128  
TID 1: Ultra = 1 (1), WIDE, speed 200, offset 8 Tagged 1 Qdepth 128  
TID 2: Ultra = 1 (1), WIDE, speed 200, offset 8 Tagged 1 Qdepth 128  
TID 7: Ultra = 0 (0), NARROW, speed 100, offset 0 Tagged 0 Qdepth 0

**SERVER LOG CATEGORY PRECEDENCE NUMBERS:**

2003-02-07 14:17:41: ADMIN: 4: Command succeeded: rquotad action=start

2003-02-07 14:17:41: NDMP: 3: RecBuf in NDMP Pool (count:0), NDMPMAXBUFSIZ is 131072

**0 = Emergency**

**1 = Alert**

**2 = Critical**

**3 = Error**

**4 = Warning**

**5 = Notice**

**6 = info**

**7 = Debug**

**HOW TO SET INCREASED LOGGING LEVELS ON DATAMOVERS:**

**SMB PROTOCOL:**

**\$server\_config server\_2 "logsys set severity SMB=LOG\_DBG3"**

**\$server\_config server\_2 "logsys set severity SMB=LOG\_ERR** [Turn off]

**\$server\_config server\_3 "logsys set severity SMB=LOG\_DEBUG"**

**Note:** LOG\_DEBUG can be used against various modules

**\$server\_config server\_3 "logsys set severity SMB=LOG\_PRINTF"** [Turn off]

**HOW TO DETERMINE WHAT LOGGING LEVEL IS SET ON A FEATURE:**

**# .server\_config server\_2 -v "logsys get severity SMB"**

1143253668: LIB: 4: Server log severity for facility SMB is 4

**Note:** Log level 4 is normal

**TURNING OFF ALL SMB LOG MESSAGES EXCEPT FOR EMERGENCY, ALERT, CRITICAL:**

1. Add following line to the bottom of /nas/server/slot\_x/start file

**logsys set severity SMB=LOG\_CRIT**

2. Reboot Data Mover

**FSCK/FS MAINT:**

**\$server\_config server\_x "logsys set severity FSTOOLS=LOG\_DBG3"**

**NDMP LOGGING:**

**\$server\_config server\_2 "logsys set severity NDMP=LOG\_DBG3"**

**\$server\_config server\_2 "logsys set severity PAX=LOG\_DBG3"**

**\$server\_config server\_2 "logsys set severity NDMP=LOG\_PRINTF** [Turn off]

**\$server\_config server\_2 "logsys set severity PAX=LOG\_PRINTF** [Turn off]

**CIFS LOGON's:**

**\$server\_config server\_3 -v "param NTsec logonTraces=6"**

**\$server\_config server\_3 -v "param NTsec logonTraces=3"** [Turn off]

**SETTING LOGGING FOR ONLY CRITICAL EVENTS:**

# .server\_config server\_2 -v "logsys set severity UFS=LOG\_CRIT"

### **VERIFYING CURRENT LOGGING LEVEL FOR CATEGORIES:**

# .server\_config server\_2 -v "logsys get severity SMB"

1158591814: LIB: 4: Server log severity for facility SMB is 4

### **SYMMTOP SYMMETRIX TOOL:**

Tool to look at SA Port stats, Caching Issues, “Disk Hotspots”.

### **SOLARIS SERVER STATISTICS:**

\$iostat -d [BSD disk activity tool--variety of possible switches for memory, processes, CPU, etc.]

/var/adm/messages dmesg

showrev -p vxdisk list vxprint vxprint -ht [Veritas configuration information]

## **II. CELERRA OPERATING SYSTEM:**

### **CELERRA VOLUME & SERVER MANAGER:**

→"NAS" commands are executed by /nas/bin/volume\_mngr, which links to nas\_cmd

lrwxrwxrwx 1 nasadmin nasadmin 9 Oct 2 14:13 volume\_mngr -> ./nas\_cmd

→"SERVER" commands are executed by /nas/bin/server\_mngr, which links to nas\_cmd

lrwxrwxrwx 1 nasadmin nasadmin 9 Oct 2 14:13 server\_mngr -> ./nas\_cmd

### **SIX TYPES OF CELERRA VOLUMES:**

Basic Volume →Physical volume from which other types are created

Slice Volume →Small volume created from another logical volume for easier management

Stripe Volume →Arrangement of volumes that together appear as a single volume, from basic or slice volumes, with data layered across the volumes in a striping fashion

Hyper Volume →Larger capacity volume created from any type, used for file systems [better known as a metavolume]

Filter Volume →Volume that provides snapshot support

RamDisk Volume →Logical volume created in memory

### **EMC NAS File System:**

After NAS install, new directory created at Linux root, /nas

/nas/dos> directory where DOS partition is mapped for DM's to T&L00; contains Boot & Config files for each Server

/nas/dos/bin> sibpost, gload [loads DART sw], nas [DART SW] executables for servers

/nas/dos/slot\_1 [thru16]; Boot.bat and Boot.cfg files for each DataMover server

/nas/bin> File System and Administration commands

/nas/sbin> Utilities, System files and services; /nas/sbin/getboxmask -r [slots that are powered up]

/nas/sbin/getreason [status of each DM & CS]

0=BIOS check, boot sequence 1=SIB Post 2=Fail SIB 3=DM Booting DOS 4=DM running NAS 5=DM normal &

BoxMon up

6=CS only [singleuser mode] 7=DM crash 9=Reboot 10=Primary CS only; 11=Secondary CS only

/nas/server>Files for each DM configuration; \$nas\_server\_1 - 16

/nas/volume> DB files for Volume manager;

/nas/log > NAS Log Files: install.log/upgrade.log/sys\_log.log/cmd\_log/cmd\_log.err

/nas/log/nas\_log.al.trace →File contains Eiffel dump outputs from scripts which fail to run

### **PROBLEM WITH /NAS AT 100% FULL:**

If nas goes to 100%, cannot create new Shares from either NT Server Manager GUI or CLI as Nasadmin.

#### **SYMPTOMS:**

##### **Server Log Error:**

2003-07-02 13:17:35: SMB: 3: srvsvcAddShare:shareReq2CS no share created

/nas:

# df -k .

Filesystem 1k-blocks Used Available Use% Mounted on

/dev/sde1 1818352 1797024 0 100% /nas

##### **Server export Fails as Nasadmin:**

\$ server\_export server\_10 -P cifs -n aaustin86 /fs-21/Home/dFBnycBUC110/aaustin6

server\_10 : unable to acquire lock(s), try later

**Server Export as Root Succeeds:**

```
# server_export server_10 -P cifs -n aaustin86 /fs-21//Home/dFBnycBUC110/aaustin6  
server_10 : done
```

**Strace of Server Export Shows Problem:**

```
$strace -o /tmp/aaa -f server_export server_2 -p -u -P cifs -n testshare  
server_2 : unable to acquire lock(s), try later
```

```
-----  
14365 getpid() = 14365  
14365 open("/nas/lock/db/1_1_0__", O_WRONLY|O_CREAT|O_EXCL|O_SYNC, 0644) = 6  
14365 write(6, "\0358\0\0X\2\0\0", 8) = -1 ENOSPC (No space left on device)
```

**Note:** Root can still write to fs as long as there are blocks left to write to, whereas all other Users will be denied writes at 100%

**OBTAINING SLOT STATUS IF BOX MONITOR SHUT OFF:**

**#/nasmcd/getreason**

**Note:** This command will return slot status even if Box Monitor is disabled or if you are telneted into CS1 running as Secondary

**UNIX TIME VALUES:** Number of seconds since epoch—Epoch is defined as 00:00:00 January 1, 1970 UTC

*st\_atim* Time when file data was last accessed. Changed by the following system calls: *creat, mknod, pipe, utime, and read*. The seconds portion of *st\_atim* is available as *st\_atime*.  
*st\_mtim* Time when data was last modified. Changed by the following system calls: *creat, mknod, pipe, utime, and write*. The seconds portion of *st\_mtim* is available as *st\_mtime*.  
*st\_ctim* Time when file status was last changed. Changed by the following system calls: *chmod, chown, creat, link, mknod, pipe, unlink, utime, and write*. The seconds portion of *st\_ctim* is available as *st\_ctime*.

**DISABLING DART A-Time UPDATES DURING READS:**

**param ufs updateAccTime=0**

**Note:** Do not apply this change without considering applications that depend on atimes, like Cava, DHSM, etc.

**DART FILE SYSTEM TIMES (3) STD UNIX TIMES & (1) SMB CIFS TIME:**

**atime (access)** - modified when file is opened for read. Last time file accessed. **#ls -lut** shows atime.

**mtime (modified)** - modified when file is opened for write. Last time file modified. **#ls -l** shows mtime.

**ctime (change)** - modified when file is opened for write, or when metadata is changed (renamed, linked, utimed, chowned), which is equivalent to SMB last access, last write, and change in perms, respectively. Last inode change time for the file. **#ls -lc** shows ctime. Ctime represents time since Epoch (00:00:00 UTC, January 1, 1970)

**(create time)** is in Shadow File until 5.2 & the MPD, then moves to the directory entry. SMB adds the creation time.

**Note:** On Linux CS, if a file is read, only the atime changes. But, if the file is changed by writing to it, all three times are changed—atime, mtime, ctime.

**VIEWING ACCESS, MODIFY, & CHANGE TIMES FOR UNIX FILES: Use “stat” command!!**

```
$ stat test.txt
```

**File:** "test.txt"

Access: Mon Jun 17 02:49:34 2002(00003.09:21:30) “atime” = last file access time [read, utime, pipe, mknod, creat]

Modify: Fri Mar 22 18:49:22 2002(00089.16:21:42) “mtime” = last file data modification [creat,mknod,pipe,utime,write]

Change: Fri Mar 22 18:49:11 2002(00089.16:21:53) “ctime” = last file status change

[chmod,chown,creat,link,mknod,pipe,unlink,utime,write]

**How to Determine if /nas and/or /nas/dos Directory is Intact and O.K. on the Symmetrix:**

Step 1. Run **#df -k** to locate device per example: **\$/dev/dsk/c0b0t0d1p1 /nas/dos**

Step 2. Run **#dd if=/dev/dsk/c0b0t0d1p1 of=/dev/null count=1**

Step 3. Device is corrupt if you see:      “UX:dd:ERROR: Cannot open /dev/dsk/c0b0t0d1p1: No such device”  
Device o.k. if you see this:      “1+0 records in 1+0 records out”

**Difference Between /nas/server/slot\_x Directory and /nas/server/server\_x Directory:**

/nas/server/server\_x directory is created based on the order that a DataMover is setup, not by slot number

/nas/server/slot\_x conversely is based on actual slot numbering and does not change

**\$nas\_server -l** [this command will list out the Index values, which are used to create the “server” directory]

**# nas\_server -l**

id type acl slot groupID state name

```

1   1  1000 2      0  server_2
2   1  1000 3      0  server_3
4   4  1000 5      0  server_5

```

**# nas\_server -l id=1**

```

id  type acl slot groupID state name
1   1  1000 2      0  server_2
2   1  1000 3      0  server_3
4   4  1000 5      0  server_5

```

**CELERRA REASON CODES FOR DATA MOVERS AND CONTROL STATIONS:****# /nas/sbin/getreason**

10 - slot\_0 primary control station  
 11 - slot\_1 secondary control station  
 5 - slot\_2 contacted  
 5 - slot\_3 contacted

**# /nas/sbin/getreason -s 3**

5

**TABLES OF REASON CODES:**

|       |   |
|-------|---|
| 0     | Reset (or unknown state)  |
| 1     | DOS boot phase, BIOS check, boot sequence   |
| 2     | SIB POST failures (i.e., hardware failures)   |
| 3     | DART is loaded on data mover, DOS boot and execution of boot.bat, boot.cfg  |
| 4     | DART is ready on data mover, running, and mac threads started   |
| 5     | DART is in contact with Control Station box monitor   |
| 6     | Control Station is ready, but is not running NAS service  |
| 7     | DART is in panic state  |
| 9     | DART reboot is pending or in halted state   |
| 10    | Primary control station reason code   |
| 11    | Secondary control station reason code   |
| 13    | DART panicked and completed memory dump (single DM configurations only, same as code 7, but done with dump)   |
| 14    | This reason code can be set for the Blade for any of the following:<br><ul style="list-style-type: none"> <li>• Data Mover enclosure-ID was not found at boot time</li> <li>• Data Mover's local network interface MAC address is different from MAC address in configuration file</li> <li>• Data Mover's serial number is different from serial number in configuration file</li> <li>• Data Mover was PXE booted with install configuration</li> <li>• SLIC IO Module configuration mismatch (Foxglove systems)</li> </ul> |
| 15    | Data Mover is flashing firmware. DART is flashing BIOS and/or POST firmware. DM cannot be reset   |
| 17    | Data Mover Hardware fault detected  |
| 18    | DM Memory Test Failure. BIOS detected memory error  |
| 19    | DM POST Test Failure. General POST error  |
| 20    | DM POST NVRAM test failure. Invalid NVRAM content error (checksum, WWN, etc.)   |
| 21    | DM POST invalid peer DM type  |
| 22    | DM POST invalid DM part number  |
| 23    | DM POST Fibre Channel test failure. Error in Blade Fibre connection (controller, Fibre discovery, etc)  |
| 24    | DM POST network test failure. Error in Ethernet controller  |
| 25    | DM T2NET Error. Unable to get blade reason code due to management switch problems   |
| Error | Failed To Get Reason Code. DM or CS may not be present in the slot or NS DM might be powered off  |

**While True Scripts:**

```

$while true          $while true
>do                >do
>/nas/sbin/getreason >/nas/sbin/getreason -s 4 [specific slot]
>sleep 8            >sleep 8
>done               >done

```

**Example of Another While True Loop:**

```

$while true
>do server_netstat server_4 -a -s -i |grep ana0
>sleep 10

```

>done

### **Another While True Loop to Monitor CPU Usages:**

```
$while true  
>do /nas/bin/server_sysstat ALL |grep idle  
>sleep 5  
>echo 5 seconds  
>done
```

## **UNIX FILE SYSTEMS:**

Collection of hierarchical directories & files; (2) Types of Owners: Users & Groups: >ls -l r-w-x

**File Access Protection:** \$chmod [access-string, divided into Access Class; Operator; & Access Type]

**Access Class:** u; g; o; a    **Operator:** +; -; =    **Access Type:** r; w; x; ...

>chmod a +w lead [all users have write access to file called 'lead']

\$chmod 777 \* [symbolic mode, indicating full access for all users to directory applied]

#chmod g+r filetest [grants Read access to 'filetest' to group 'g']

## **FILE PERMISSIONS WITH UNIX:**

UNIX contains permissions for files that control access, also called its “access mode.”

### **Users can have (3) types of permissions for files:**

Permission to Read [Read file contents or list directories]

Permission to Write [Write changes to a file or create & remove files in directories]

Permission to Execute [run a program, or access files in a directory]

### **Users can also fall into (3) categories, which affect how the above (3) permissions are given:**

Owner—Other Users in file’s Group—Everyone else

Files usually get UID of user that creates file and GID of the Group that owns the directory

### **Files also have (3) special components for executable files:**

Setuid bit—this sets the process’ UID to that of the file that is being executed

Setgid bit—this sets the process’ GID to that of the file that is being executed

Sticky bit—saves a program’s text image on a swap device so as to load quicker when run. For directories, prevents Users from removing files from a directory when they are not the owner

### **SEARCHING DIRECTORIES FOR STICKY BIT FILES:**

**#ls -l /usr/bin |grep '^.....s'**

## **CHMOD COMMAND:**

Chmod is used to assign or change the “mode” of a file based on a person’s or other’s attributes. Mode is expressed in absolute (Octals) or Symbolic (g+s; g-s; =)

## **ADD VALUES FOR EACH CATEGORY (UGO):**

|   |                                      |
|---|--------------------------------------|
| 0 | = no permissions                     |
| 1 | = execute only (1)                   |
| 2 | = write only (2)                     |
| 3 | = write and execute (2+1)            |
| 4 | = read only (4)                      |
| 5 | = read and execute (4+1)             |
| 6 | = read and write (4+2)               |
| 7 | = read and write and execute (4+2+1) |

## **ACCESS PERMS EXPRESSED IN 3-DIGITS:**

|                 | user | group | others |
|-----------------|------|-------|--------|
| chmod 640 file1 | rw-  | r--   | ---    |
| chmod 754 file1 | rwx  | r-x   | r--    |
| chmod 664 file1 | rw-  | rw-   | r--    |

## **OCTAL VALUES:** Range from 0-7

#chmod 0400 file [Read by Owner]

#chmod 0200 file [Write by Owner]

#chmod 0100 file [Execute by Owner]

#chmod 0700 file [Read-Write-Execute by Owner]

#chmod 0040 file [Read by Group, etc...]

#chmod 0004 file [Read by Other, etc....]

**EXAMPLES:**

#chmod 444 file [read perms to Everyone]  
#chmod 066 file [Read-Write to Group & Other]

**SYMBOLIC VALUES:** +, -, =, u,g,o,a; r, w, x, l, s, t

+add perms - removes perms =assign perms absolutely

u=user's perms

g=group's

o=other's

a=all persons

r=read

w=write

x=execute

l=mandatory locking

s=user or group set-ID

t=sticky bit

u, g, o [User, Group, Other]

**EXAMPLES:**

#chmod -R g-s\* [Recursively removes the set group id bit [setgid, sbits, etc] from a parent directory down  
drwxr-sr-x

#ls -ld →outputs l-bit for locking or sbit for setgid in following position: -----s---

#chmod a-x file [Deny execute perms to everyone]

#chmod go+rw file [Read-Write to Group & Other]

#chmod +l file [Locks file when accessed]

**SET GROUP ID (SGID) BIT:**

When the sgid bit is set, means that a User that executes a program will have the same privileges as the group owner of the program.

**STICKY BIT:**

When the sbit (sticky bit) is set on a directory it means that all files will inherit the default directory GID and not that of the process used to create the file itself. For certain directories that may be open to the category OTHER, when the sticky bit has been set on the directory, the files within can only be deleted by the Owner of the File or Directory, or Root. If the SetGid bit is set on a directory where “satom” is owner and group is “sysadmin”, then all files created in this directory will inherit the group “sysadmin” ownership & override the primary Unix group ownership of other Users’ groups. Sbit cannot be set or unset in ‘absolute’ mode, only via ‘symbolic’ mode.

g-s →removes sbit setting

g+s →sets the sbit

# chmod sato 1000 [Turns sticky bit on]

**Example:**

```
drwxrwsrwx 2 satom sysadmin 512 Apr 21 14:17 .
drwxrwxrwx 40 root users 4608 Apr 21 14:14 ..
-rw-r--r-- 1 satom sysadmin 0 Apr 21 14:15 satom.file
```

**Note:** Celerra now employs the SGID bit as the default in NAS 5.1 and higher for files created in directories using UNIX rules

**SET USER ID (suid) BIT:**

When the suid attribute is set on the access permissions for a program, a user executing the program has the same privileges as the owner of the program—if the owner were root, that would mean root privileges—only applies to executable programs or scripts.

**Note:** The SUID/SGID attributes are cleared when files are written to, meaning that an application should do a SETATTR after writing to insure that the appropriate attributes are put back—this is behavior in line with POSIX standards.

**Octal Permission Values:**

4=100 Binary = Read [View a File or list a Directory]

2=010 Binary = Write [Change File or Directory]

1=001 Binary = Execute [Execute a file or cd to a directory]

**Default Directory Permissions are 755**

**Note:** Add 4+2+1 to obtain the 7, the highest possible value

**Default File Permissions are 644**

**Note:** Values from Left to Right: User—Group—Other

Default CIFS Umask is 022 [Octal Value 22 = Hexvalue 0x12], which is UGO Octal Value 755 for any files created from NT for Unix. By default, UNIX creates files with a perm of 666 and directories with 777. Celerra uses a default Umask of 022, which equates to default file perms of 644 and directories of 755.

### **Standard Unix Directories Permissions:**

READ—Allows users to list names of files in the directory only

EXECUTE—Users can utilize files that can be explicitly named

READ/EXECUTE—Users can List & Utilize files

READ/WRITE/EXECUTE—Users can List, Utilize, and Delete/Add files in the directory

**Note:** For Unix over NFS, new files derive permissions from client's umask

For CIFS created files, Unix permissions are derived from a configurable “umask” value set on the File System!

For CIFS created files, default is to have a UGO of 755 applied [i.e., a umask of 022]

### **UNIX PERMISSIONS:** Traditional Unix uses (3) bits each to define USER, GROUP, and OTHER “rwx” permissions

Read—list files only, in directories

Execute—utilize files that you can name

Read/execute—list and utilize files

Read/write/execute—list files in directories, utilize files, create/delete files

Unix file permission sets contains Unix UIDs and GIDs

### **LEGAL/ILLEGAL UNIX CHARACTERS IN FILE NAMES:**

**Note:** Though the rules don't seem to be hard and fast across the unix front, here is some information related to characters that are allowed, or not allowed, or not recommended, for use with unix names:

#### **Legal Characters:**

**A thru Z**

**a thru z**

**0-9**

**. - -**

#### **Illegal Characters:**

**spaces \* ? ! | \ / ‘ “ { } < > ; , ^ ( ) \$ ~**

Also recommended that hyphen – is not used at beginning of a name, nor period . , as it indicates a special file

### **LEGAL INTERNET HOSTNAME CHARACTERS:**

RFC 952 defines what a valid Internet Host name should consist of for characters—it's common for Unix hostnames to contain an underscore in the name, for example, but this is not a legal Internet Hostname, so the correct approach is for Unix hostnames to be renamed to conform to the Internet Hostname requirements

→A name is defined as a Net, Host, Gateway, or Domain name

→A name is a text string up to 24 characters in length, consisting of alpha-numeric A-Z (upper or lower case, does not matter), 0-9, and can include a minus sign (-) or a period (.), except that periods are only allowed as delimiters for domain style names

→First character must be an alpha character

→Last character of the Hostname must not be a minus (-) or period (.)

→No blanks or spaces are allowed as character entries for the name

### **SUPERUSER ACCOUNT & NFS EXPORTS [aka root]: Default is for no hosts allowed 'root' access**

**root=hostname | netgroup | subnet | IP address** [used to specify access for User "root"]

When accessing remote file systems over NFS, the "root" user's UID is normally mapped to the anonymous user account called "nobody", which has a UID of -2 and a "nogroup" GID of -2 [For non-negative unix systems, these values translate into a 16-bit unsigned representation of 65534].

**Note:** NFS handles requests for Users without valid credentials by mapping them also to the "anonymous" user account. Linux versions have a no\_root\_squash export option that will allow NFS clients to mount an export with Root privileges.

\$ server\_export server\_5

**export "/fs06" root=172.0.0.0/255.0.0.0**

**Note:** This is an example of a file system exported to allow superuser root to connect from a remote system, to this Celerra file system, and retain root privileges on the Celerra file system. It does not mean that regular unix users will have root privileges when connecting to this file system—all other user accounts will still be governed by normal UNIX mode permission bits on files & dirs.

### **STANDARD UNIX SHELLS:**

Bourne shell (sh) original shell found on all unix systems, used for simple programs

C shell (csh) uses C-like programming syntax for writing scripts

TC shell (tcsh) Default C-Shell & emacs-style editing

Korn shell (ksh) more features than Bourne, considered the standard, combines C & TC shells

Bourne Again shell (bash) GNU, uses C & K shells and is compatible with Bourne

**Note:** #ps -s [Shell processes running and commands that are running]

## **DETERMINING RUNNING SHELL:**

#echo \$SHELL

/bin/bash

## **SHELL CAPABILITIES:**

|                      | Bourne | C    | TC  | Korn | BASH |
|----------------------|--------|------|-----|------|------|
| command history      | No     | Yes  | Yes | Yes  | Yes  |
| command alias        | No     | Yes  | Yes | Yes  | Yes  |
| shell scripts        | Yes    | Yes  | Yes | Yes  | Yes  |
| filename completion  | No     | Yes* | Yes | Yes* | Yes  |
| command line editing | No     | No   | Yes | Yes* | Yes  |
| job control          | No     | Yes  | Yes | Yes  | Yes  |

## **BINARY COMPUTER LANGUAGE:**

--Binary is a contraction of “binary digit”

--4-bits is a single hex digit while 8-bits form “bytes” or ‘characters’

--Words are 16-bits long and are called Short Integers, often signed +32768 or -32768

--Double-Words are Long Integers at 32-bits in length

--Kilobytes are expressed in powers of 2. A single K is 1024 bytes ( $2^{10}$  bytes). 512K is actually  $512 \times 1024$ , or 523,288 bytes.

## **HEXADECIMAL BASE 16 NUMBERING SYSTEM:**

### **HEX    BINARY**

|   |      |
|---|------|
| 0 | 0000 |
| 1 | 0001 |
| 2 | 0010 |
| 3 | 0011 |
| 4 | 0100 |
| 5 | 0101 |
| 6 | 0110 |
| 7 | 0111 |
| 8 | 1000 |
| 9 | 1001 |
| A | 1010 |
| B | 1011 |
| C | 1100 |
| D | 1101 |
| E | 1110 |
| F | 1111 |

## **LINUX USER ENVIRONMENT & STARTUP SCRIPTS:**

### **BOURNE SHELL :** sh /bin/bash

/etc/skel → startup scripts for basic shells ; .bashrc = BASH config file ; .bash\_logout ; .bash\_profile for startup environment

/etc/profile → Executed when User logs in, defines HISTSIZE, MAIL, and system environmental [.bashrc & .bash\_profile]

/etc/profile.d → contains scripts

/etc/bashrc → global aliases and system settings

### **C-SHELL:** csh /bin/csh

/etc/login [.login & .cshrc]

#chsh [Changing Shell]

#env

## **HANDLING NFS ALIAS NAMES CELERRA REMOTE MOUNTS:**

Normally, if a Server is rebooted, and Clients have locks on files, the Server will notify the client using statd.hostname that the file is available to reclaim the lock. With alias names, however, we must add the name of the data mover interface to the list of names in statd.hostname so that we properly notify the Client using our “alias”.

## **USING ANON EXPORT FOR NFS:** /etc/exports   Default is to allow UID 'nobody' or value 65534 access

anon=uid Use this option to specify access by UID: Example: anon=100

Set NFS access mappings to specific UID's using the "anon=100" syntax when "exporting" file systems for NFS

Set NFS access mappings to specific HOSTS using "access=taco,root=taco"

Further, if you export NFS to a specific HOST as ReadWrite [rw=taco:mountpoint,access=100] then all others will mount RO

Using anon=0 has the effect of forcing all Anonymous Users to become the “nobody” user -2 for the remote mount  
root= allows a remotely connecting user to have the net effect of being local “root” on the remote file system

### **CLIENT ACCESS FOR NFS EXPORTS: Default is to allow mount access to all machines**

**Resolution Order:** /etc/hosts → /etc/netgroup → NIS → DNS

Provide mount access to clients listed, either by **Hostname | IP Address | subnet | netgroup**

#### **Standard NT Directory Permissions:**

R—Read a File or Directory

W—Write to a File or create File in a Directory

X—Execute a File or enter a Directory

P—Owner [right to change ACL's on a File or Directory]

D—Delete a File or Directory

O—Take ownership of a File or Directory

**Note:** For NT, new files inherit the ACL's of the parent directory

### **OUTPUTTING DIFFERENCE BETWEEN FILES:**

\$diff file1 file2 > diff.txt

#### **Calculating Total Number of Files Possible for Our Default 8k FileSystem:**

**Note:** This is based on (1)inode per every 8k block, which is our FileSystem data block size

File System Size: 8.613GB

Default Inode: 1/8k block

Solution: 8,613,000,000/8000bytes = 1,076,625 max files for this FileSystem [Only if every single data file was exactly 8k or less]

### **VIEWING FREE INODES ON DATAMOVER FILESYSTEM:**

**\$server\_df server\_4 -i fs06**

fs06 1023358 8 1023350 0% /NFS6

**Specifying Inode Density for a new File System:** #nas\_fs -n mv10fs -type ufs -c mv10 -o nbpi=4096

**Verify:** #server\_df server\_2 -i mv10 or \$nas\_fs mv10fs -i

#### **Two Ways to Change Inode Densities:**

I. By “Extending” a File System, which adds more inodes

II. Back-up File System, create new FS with higher Inode density, then Restore File System to new.

**Variables:** \$echo \$VAR \$export VAR{\$PATH; \$SHELL} #set #env \$echo ~ [outputs Home directory to screen]

### **COMPARISION OF UNIX COMMANDS/DIRECTORIES/FUNCTIONS:**

|                     | <b><i>Linux</i></b>              | <b><i>Solaris</i></b>        | <b><i>AIX</i></b>    | <b><i>HP-UX</i></b>      |
|---------------------|----------------------------------|------------------------------|----------------------|--------------------------|
| Password File:      | /etc/passwd                      | same                         | same                 | same                     |
|                     | /etc/shadow                      | same                         | /etc/security/passwd | /tcb/files/auth/r/root   |
| Max UID's:          | 65535                            | 2147483647                   | 4294967295           | 2147483647               |
| Nobody's ID:        | 99                               | 60001/65534                  | 4294967294           | -2                       |
| HostID:             | hostid                           | same                         | same                 | uname -i                 |
| Perf. Monitor:      | top                              | top                          | top/monitor          | top/glance               |
| Memory Stats:       | vmstat                           | same                         | same                 | same                     |
| I/O Stats:          | iostat                           | same                         | same                 | same                     |
| Network IP Config:  | /etc/sysconfig/ /network-scripts | /etc/inet /etc/defaultrouter | lsattr -E -l inet0   | /etc/rc.config.d/netconf |
| Name Service:       | /etc/nsswitch.conf               | same                         | /etc/netsvc.conf     | /etc/nsswitch.conf       |
| Network Params:     | sysctl -a  grep net              | ndd /dev/tcp                 | no -a                | ndd -h                   |
| NIC Config:         | ifconfig -a                      | same                         | same                 | lanscan -v               |
| Exports:            | /etc/exports                     | /etc/dfs/dfstab              | /etc/exports         | same                     |
| File System Table:  | /etc/fstab                       | /etc/vfstab                  | /etc/filesystems     | /etc/fstab               |
| Free Disk Blocks:   | df -k                            | same                         | same                 | bdf                      |
| Device Listing:     | cat /proc/devices                | sysdef                       | lsdev -C             | /sbin/ioscan             |
| Backup:             | tar cvf /dev/rst0/               | ufsdump                      | savevg -i rootvg     | fbackup                  |
| Startup Script:     | /etc/rc.d/rc                     | /sbin/init.d                 | /etc/rc              | /sbin/rc                 |
| Trace System Calls: | strace                           | truss                        | syscalls             | tusc                     |
| OS Level:           | uname -r                         | same                         | oslevel              | uname -r                 |

|            |            |                   |              |                 |
|------------|------------|-------------------|--------------|-----------------|
| Run Level: | runlevel   | who -r            | same         | same            |
| List SW:   | rpm -qa    | pkginfo           | lslpp -L all | swlist          |
| CD-ROM:    | /dev/cdrom | /dev/dsk/c#t6d0s2 | /dev/cd0     | /dev/dsk/c#t2d0 |
| Floppy:    | /dev/fd0   | /dev/diskette     | /dev/rfd0    | --              |

## UNIX Vi EDITOR:

# cat -etv filename [displays exact characters within text file!]

## DETERMINING SHELL & SETTING BOURNE SHELL EDITOR:

1. #echo \$\$SHELL →/bin/bash
2. #export TERM
3. #TERM=vt100
4. #tset [Use ctrl + U to kill shell process; Use ctrl + C to Interrupt Shell processes]

## CHANGING SHELL:

#chsh

**Insert:** i inserts before cursor; I inserts at beginning of line; a appends after cursor; A appends at end of line; o Opens line below; O opens line above; c change operation; C change to end of line; R overwrite text; s sub character; S sub entire line

**Basic Operators:** c Begin a change; d Begin a deletion; y Begin a copy; See cc; dd; yy usages

**Moving around a File:** H top of screen; M goes to middle; L goes to last line; Ctrl + F scrolls fwd one screen; Ctrl + B back Ctrl + G to end of file

## Down & Up One Page:

**ctrl F = Page Down; ctrl B = Page Up ; Shift H top of page; Shift M middle of page; Shift L last line of page**

**Moving across a Line:** Use “f” and character string you wish to go to to move to that section; w to move forward a whole word

**Searching:** /text will go forward to find instances of “text”; n repeats search fwd; N repeats backwards; G to bottom of file;

**Editing:** [a =append x=delete esc=edit r=replace] Input v. Command mode :wq! :q! ZZ=quit program & saves file; R replace with new text; r replaces character; x deletes character;

## Deleting With Vi:

dd deletes current line; D deletes remainder of line

2000dd → Deletes first two thousand lines in the file, etc.

## JOINING TWO LINES TOGETHER IN VI:

Go to end of line above the other; esc J

## COPYING A LINE IN VI EDITOR:

esc : set nu :t 12 [Copies Line 12 into new line 13]

## Copying Contents of File1 to Specific Line in File2:

#vi file1 & go to line to import; !!cat file2 [pulls contents of file2 to line number below cursor]

#vi file1 & go to line to import; :r file2

## Copying & Pasting Entries within File:

Y—yanks current line; paste to appropriate line using ‘P’ for above cursor and ‘p’ for below cursor

15yy—yanks 15 lines underneath cursor; paste to destination using P or p

## Copying Line below Cursor:

:co or :t150 [Copies line 150 & inserts to line 151]

## Opening File and Saving buffer as New File Name:

# vi file

# esc :w <newname>

# :q

## OTHER MOVE AND COPY EXAMPLES:

:line1mdestline

Move (cut) line number, line1, to the line number, destline

:line1, line2mdestline

Move (cut) lines between and including line1 and line2 below line number, destline

:line1tdestline

Transfer (copy) line number, line1, to the line just below line number, destline

:line1, line2tdestline

Transfer (copy) lines between and including line1 and line2 below line number, destline

## MOVING LINES IN VI EDITOR:

:m 1 2 [Moves Line 1 to line 2]

**Using Vi With Line Numbers:** Open Vi session, then “esc” + : set nu

**Other Editors:** vuepad emacs

X deletes backwards; YY copies; p pastes below cursor; . repeats command

ex Editor: Enter : when in Vi to enter the “ex” mode: Substituting characters within a whole file :%s/sh/ksh/

^ beginning of file \$ end of file . carriage return s substitution cmd

#### **EDITING IN BINARY MODE:**

# vi file esc : set binary

#### **SAVING PORTION OF FILE TO NEW FILE IN Vi EDITOR:**

**ma** (mark text at the top of the region to be saved)

**mb** (mark text at the bottom of the region to be saved)

**:a,'b w filename**

#### **CONVERTING DOS FILES TO UNIX:**

# /usr/bin/dos2unix -n dos\_file unix\_file

# sed -e 's/\$//' dosfile.txt > unixfile.txt

**Note:** Different switches available, above converts file to unix format and writes to a new filename. Unix text files end with “LINE FEED” character “\$”, but dos files end with “CARRIAGE RETURN + LINE FEED”.

In some cases, may also need to open file in vi after converting from dos2unix, then save as new filename. Rename file as needed.

# vi file # esc :w <newfile> # :q

#### **USING STRINGS TO CONDUCT CHARACTER SEARCHES:**

# for i in \*.db;do echo \$i; strings \$ilegrep -i "\.ab"; done

**Note:** Searches through all files with .db suffix and greps for any lines with “.ab” characters

#### **READING BINARY FILES:**

1. One tool to try is “strings”, as in this example: #strings usrmappc.db

2. Alternate tool to try is “od -c” or “od -a” command against a file

#### **USING TEE SWITCH TO OUTPUT TO FILE:**

\$ server\_netstat server\_4 |wc -l | tee -a /tmp/chk.out [Appends output to file called "chk.out"]

#### **SETTING VI EDITOR IN PRIVILEGED & NONPRIVILEGED MODES:**

\$ set o -vi Press the Esc key; then use ‘k’ to move forward, ‘j’ to move back; r replace; l to move across line; Shift + \$ to end j and k keys for down and up respectively; h and l keys for backwards and forward on a single line

# ksh -o vi Press Esc key, then use ‘k’ and ‘j’ options

#### **USEFUL COMMANDS:**

wc -l [number of lines in a file] wc -c [word count]

#### **OUTPUTTING DIRECTORY LISTING TO FILE:**

# ls |more >directory.list

#### **LISTING DIRECTORY CONTENTS IN ORDER OF CREATION TIME:**

\$ ls -lt

-rw-rw-r-- 1 nasadmin nasadmin 1043 Oct 1 17:46 test

-rwxrwxrwx 1 root root 243 Sep 30 20:49 loop

#### **LISTING DIRECTORY CONTENTS BY FILE TYPE:**

# ls -laF

drwxrws--- 3 nasadmin nasadmin 4096 Dec 13 12:56 ..

-rwxrwxr-x 1 nasadmin nasadmin 151 Dec 8 12:28 .old\_schedules.tar.gz\*

-rw-r--r-- 1 root root 0 Dec 13 12:55 regular\_file

srwxr-xr-x 1 nasadmin root 0 Dec 9 16:48 .scheduler\_socket=

**Note:** -F switch appends a character to each file name based on file type. \*=executable; / = dirs; @=symbolic links; l = FIFOs; = for sockets; nothing for regular files

#### **DETERMINING PATH LENGTHS:**

# less /home/nasadmin/ts2/dan\_fs.log.gz | awk '{print length, \$0}'|sort -nr|head -10

#### **CONDUCTING COMPLETE LISTING OF NAS SERVER SLOTS:**

**\$ ls -l /nas/server/slot\_\*/\***

**Note:** Useful in that it will list out all files in each slot for SCCS

### **COMPARING DIFFERENCES IN TEXT FILES:**

**# diff file1 file2**

**# sdiff file1 file2** (side by side comparison of files—useful if you have a monitor that can display both lines across the screen)

### **COMPARING DIFFERENCES IN BINARY FILES:**

**# cmp file1 file2 or # cmp file\*** [compares all files named ‘file’]

## **VARIOUS METHODS FOR REMOVING DOS CONTROL CHARACTERS FROM ASCII FILES**

### **USING VI & LINUX COMMAND:**

**Method 1:** Open file with vi and use **# esc :1,\$ s'^M'g**

**Note:** To type '^M, press ctrl key while typing "v" and the "M". Also, type *two single apostrophes*, not 'quotes'!

**Method 2:** vi file; **# esc :% s<ctrl v><ctrl M> //g**

**Method 3:** Open file with vi editor: **#esc :%s<ctrl + v><enter> //g**

**Method 4:** vi file **# esc :1,\$ s<ctrl + v ctrl + m> //g**

**Method 5:** Use Linux DOS2UNIX conversion command

**# /usr/bin/dos2unix -n passwd.bad passwd.new**

### **HOW TO REMOVE LINES WITH ONLY WHITE SPACE FROM A FILE:**

Open file with vi editor: **:g/^<ctrl + v><tab> \*\$/d**

### **How to Ensure that there are no Extra Lines or Spaces in a File Using Vi Editor:**

**esc : set list**

**esc : set nu**

**esc : 100** [Brings you to line 100]

### **SUBSTITUTING WORDS IN A FILE:**

**: % s /old/new/g** [Will replace the word 'old' with 'new']

### **REPLACING “SPACES” WITH LETTERS USING VI:**

**#vi usrmap.db**

**:,\$s/ /xxx/g**

**Note:** Replaces “spaces” with “xxx” throughout file

### **REMOVING LETTERS “XXX” FROM FILE:**

**:.,\$s/XXX//g**

**:1,\$s/@//g** (removes all @ from file)

### **ADDING COLON TO END OF EACH LINE:**

**esc : 1,\$s/\$:/ or :g/\$s//:/g**

### **REMOVING & REPLACING ITEMS USING VI [Substitutions]:**

Group File: S-1-5-15-139d2e78-56b177fd-5475b975-2d61e.\*:58393:ats.cs.mgmt.dir:

1. **: 1, \$s\\$.dir//** [Removes the dir from end of sentence on each line in file]

2. **: 1, \$s\\$.//g** [Removes all periods throughout file]

3. **: 1,\$s/:\$.dir:/ (enter)** [Places “.dir.” at end of each line in file]

### **TRANSLATING CONTENTS OF FILES FROM ONE LETTER TO ANOTHER:**

**# cat file1 | tr 1 2** [Translates all instances of the numeral “1” and puts a “2”]

### **SUBSTITUTING PATTERNS:**

**Example:** Substituting all instances of “info” with “information”

**esc : 1,\$s/info/information /g** [drop “g” to substitute single instances]

## **PASSWORD & GROUP FILE RULES FOR CELERRA:**

1. All characters must be in lowercase
2. The Windows domain must be appended to the user or group name unless the CIFS resolver parameter is set to one
3. All non-alphanumeric characters must be substituted with their hexadecimal equivalent

## **VALID GROUP FILE FORMAT:**

S-1-5-15-242a3a09-1f6d0078-5fc894f0-203\*:32770:domain=20computers.csfb:

S-1-5-15-242a3a09-1f6d0078-5fc894f0-64c25\*:32771:appocdefaultr.csfb:

S-1-5-15-242a3a09-1f6d0078-5fc894f0-201\*:32772:domain=20users.csfb:

## **VALID PASSWORD FILE FORMAT:**

S-1-5-15-242a3a09-1f6d0078-5fc894f0-28a7\*:32768:32768:user mreilly from domain csfb:/usr/S-1-5-15-242a3a09-1f6d0078-5fc894f0-28a7:/bin/sh

S-1-5-15-74581049-1930354d-375b3a1a-172c\*:32769:32768:user mreilly from domain \_history\_sid\_range\_:/usr/S-1-5-15-74581049-1930354d-375b3a1a-172c:/bin/sh

S-1-5-15-242a3a09-1f6d0078-5fc894f0-60720\*:32770:32768:user jwilensk from domain csfb:/usr/S-1-5-15-242a3a09-1f6d0078-5fc894f0-60720:/bin/sh

## **USERMAPPER EDITS—PRODUCING A PASSWD FILE FROM USRMAPUSRC.DB FILE:**

```
1. #strings usrmapusrc.db >users.strings
2. #cat users.strings |grep :user |sort -t: -k4,4 -k3,3 -k1,1 -u >users.strings.sorted
3. Run following script to convert the Version 3 format to Unix Passwd file format
#!/convertv1users.sh users.strings.sorted >passwd
# $1 is assumed to be a file in the dumpfilesall 3 format of
# SID:*:uid:domid:user xxx from domain yyy/usr/SID:/bin/sh
# This format is converted to the /etc/passwd format of
# xxx.yyy:*:uid:domid:SID:/usr/SID:/bin/sh
# sed converts the string ":user xxx from domain yyy:" to :xxx.yyy"
# awk uses : as a delimiter and prints the fields in the correct order
#
if [ -f "$1" ]; then
cat $1 | sed 's/^(.*:*)user \(.*\) from domain \([^\:]*[^\:]*\)\(.*\)\(\.\*)/\1\2.\3\4/' | awk -F : '{print $5":*:$3:$4:$1:$6:$7;}' 
else
    echo "syntax: $0 <input file name>"
    exit
fi
```

## **PRODUCING GROUP FILE FROM USRMAPGRPC.DB FILE:**

1. #strings usrmapgrpc.db | egrep ".\*:\*:\*:\$" | sort -t : -k 3,3 -k 1,1 -u > groups.strings.sorted

2. Run following script to convert the Version 3 format to Unix Group file format

```
#!/convertv1groups.sh groups.strings.sorted >group
# $1 is assumed to be a group file in dumpfilesall 3 format of
# SID:*:domid:name.domain
# This is converted to the /etc/group format of
# name.domain:*:domid:SID

if [ -f "$1" ]; then
cat $1 | awk -F : '{print $4":*:$3:$1;}' 
else
    echo "usage: $0 <input file name>"j
    exit
fi
```

**Caution:** The above scripts work, just remember that the 4<sup>th</sup> colon is not added to the group file and would need to be done manually—just use the following command in vi to add the colon! Also, inspect the beginning and ends of each file as there may be a junk line or two that would need removal:

**esc :1,\$s/\$:/**

## **REMOVING PERIODS FROM GROUP NAMES BUT RETAINING “.” BEFORE DOMAIN NAME:**

**# cat usrmap.group | sed ‘s/\..\*\.\./ > usrmapgrp.db**

## **MANUALLY VERIFYING GAPS IN UID/GID RANGES IN GROUP & PASSWD FILES:**

```
# cat usrmap.passwd |awk -F : '{print $3;}'
```

## **CHECKING PASSWD/GROUP FILES FOR DUPLICATE UID/GIDs:**

```
# cat usrmap.passwd |cut -d ":" -f3 |sort |uniq -c |awk ' $1 != "1" {print $1, $2}'  
# cat usrmap.group |cut -d ":" -f3 |sort |uniq -c |awk ' $1 != "1" {print $1, $2}'
```

Note: Duplicates will print to screen like the following example

```
2 32900
```

```
2 32901
```

## **STRIPPING OUT DUPLICATES FROM FILES USING UNIQ SWITCH:**

```
# cat dump.group [produced from dump 1] | sort -k 3,3 -g -t: -r -o group.rsort  
# cat group.rsort |sort -k 1,1 -t: -u -o group.unique  
# cat group.unique |sort -k 3,3 -g -t: -o group.final.sorted
```

## **Displaying Exact Contents of a File to Include Octal, Hex, Character Formats:**

```
$ od -c data_file
```

## **IMPORTING USRMAPPER PASSWORD/GROUP FILES INTO EXCEL FOR NEW CONVERSION:**

1. Conduct usrmap\_control dumpfilesall 3
2. Download usrmap.passwd and usrmap.group to desktop
3. Open Excel program>File>Open usrmap.passwd
  - a.) Text Import Wizard appears--\*Delimited should be only thing checked
  - b.) Select 'Next'
  - c.) Keep 'Tab' and select 'Other' and enter colon ':' →Next
  - d.) Click 'Finish'
4. Save as a new name—‘passwd.imported’
5. Make edits to the ‘passwd.imported’ & ‘group.imported’ files as necessary
6. Push files back to Linux CS and convert from Text Delimited to Colon Delimited

```
#vi usrmap.db
```

```
esc :%s/<enter tab key once to create tab character>/:/g [This globally substitutes tab character with : ]
```

7. If group file is missing colon at end of each line, use following command to add it back:

```
#vi usrmapgrp.db
```

```
esc :1,$$/:/
```

## **REMOVING EXCEL TAB DELIMITERS FROM A FILE FOR LINUX:**

```
esc :%s/<enter tab key once to create tab character>/:/g
```

## **BEST PRACTICES TO CONSIDER WHEN COPYING OR TARRING FILES:**

```
# cp -p * conagra [Maintains Unix permissions, ownership, and dates of files copied from one directory to another]  
# tar -cpf newfiles.tar * [Maintains Unix permissions, ownership, and dates when tarring up files]  
# tar -xpf newfiles.tar [Same as above, but when ‘untarring’ files]  
# tar -zcpf newfiles.tar.gz * [Tar and zip all files in one operation]  
# tar -ztvf newfiles.tar.gz [Look at contents of a zipped tar file]  
# tar -chf ~/sys_logs.tar sys_log* [Handy cmd to tar up all sys_logs, including the symbol. linked file]  
# tar -zcpf /home/nasadmin/bak/usrmap.bak.tar.gz usrmapper [Tars & zips ./etc/usrmapper directory to a file and location specified--Run from ./etc directory]  
# (cd /etc;tar -cvf - .) | tar -xvf -
```

## **UNTARRING SINGLE FILE FROM ZIPPED TARBALL:**

1. Identify required file:

```
# tar -tzvf usrmap.tar.gz |more
```

2. Extract file:

```
# tar -xvf usrmap.tar.gz nas/cifs/usrmapperV3/linux/dbcfg/sidname.db
```

Caution: Restores to path indicated

## **USEFUL LINUX COMMANDS:**

```
# stat [command shows file creation, modification, and last access times]
# file * [shows all the files and file-types in a directory]
# locate portmap [Shows locations and path for the file or executable]
# expr 2 + 2 + 2 [Math operations]
```

6

## **SEARCHING MAN PAGES:**

**\$ man -k nfs** [Useful for searching man pages by topic]

## **FINDING APPROPRIATE MAN PAGE TOPIC IF YOU DON'T KNOW COMMAND NAME:**

### **\$ apropos copy**

```
bcopy      (3) - copy byte strings
cp        (1) - copy files and directories
cpio      (1) - copy files to and from archives
dd        (1) - convert and copy a file
-----abridged-----
```

## **NAS & SERVER COMMAND MAN PAGES LOCATION:**

/nbsnas/man/cat1

## **UNIX FIND COMMAND:**

```
>find used with switches, powerful tool; >find -print filename [displays] ; >find -atime 5 [find files accessed 5 days ago]
# find / -print -name updategroup [searches entire root file system for file called "updategroup"]
# find -name "newfile" -print #find -user terry #find -mtime 1 [find files modified within last 1 day]
# find /usr/ron -size +10 -atime +30 -print [find files >than 10 blocks & accessed over 30 days ago]
# find /usr/ron -size -20 -mtime +15 -print [find files less than 20 blocks in size modified over 15 days ago]
# find /usr -name "pg5" -exec rm {} \; [finds file called 'pg5' in /usr directory and removes by executing "rm" command]
# find . -mtime -1 [Locating files that have changed timestamps within last day in current directory]
# find . -print |egrep httpd
# find /nas/cifs -name '*.db' -size +100kb -print [Finds files in /nas/cifs directory with .db extension and 100kb in size or more]
```

## **USING FIND TO FIND SPECIFIC UID IN FILE SYSTEM AND PRINT PATH TO FILE:**

**# /nasmcd/quota/slot\_2/fs1/find . -uid 32768 -print >/home/nasadmin/tm/32768.out &**

## **USING FIND COMMAND TO SEARCH FILES BY SIZE:**

**# find /nas -size +5000 -print**

Displays files over 5000 blocks(2.5MB) on the /nas filesystem

## **Using FIND to List Directory Structure in a Home Directory Folder and Output to File:**

**# find /nas/rootfs/slot\_3/engineering/raylau -exec ls {} \; >/home/nasadmin/raylau1 &**

**# find . -mtime -1 -print0 -exec cp -parents {} /dest \;**

One purpose of doing this is for checking total path lengths of a particular filesystem, especially for Tape Backup situations

## **USING FIND TO FIND INODE & OUTPUTTING PATH TO INODE:**

**# find . -inum <inode#> -print**

**# find /emcfind -inum 125228 -print >/inode1 &**

## **USING FIND TO LIST MTIMES:**

**# find . -mtime -1 -print0 -exec cp -parents {} /dest \;**

## **USING FIND TO FIND A NAME & OUTPUTTING PATH TO FILE:**

**# find /emcfind -iname 'WRL0516.tmp' -print >name2 &**

## **USING FIND TO LOCATE ALL FILES WITH ~ IN NAMES:**

**# find . ! -type d -depth -exec ls -al {} \; | egrep "\~" > /path/tilde.log**

**# find . -name \\*~\* -ls >/home/nasadmin/tilde.out &**

**Note:** Above command will output all directories & files with a tilde, as well as the timestamp from ls

## **FINDING SPECIFIC FILENAME FROM LINUX CS:**

**# find -iname edw\_prod\_vsaf\_vpi\_tlog\_200410110701.BAK**

## **FINDING FILES CREATED WITH FUTURE DATE & UPDATE WITH CURRENT TIMESTAMP:**

**PROBLEM:** Numerous files created or modified since a certain date are getting a year 2020 timestamp

```
$ ls -la *
-rw-r--r-- 1 root bin 19932947 May 10 2020 tcp.dump
-rw-r--r-- 1 root root 5 May 10 2020 temp
```

1. Create a template file with a future date in a directory on the CS where the find command will be run from—for example, a test file "date\_2019" with a December 12, 2019 timestamp:

```
# touch -t 201912121212.12 date_2019
```

```
# ls -la
-rw-r--r-- 1 root root 0 Dec 12 2019 date_2019
```

2. Use the template file as a time reference when running the find & use touch to update timestamp:

```
# nohup find /nas/quota/slot_2/root_vdm_3/Avr -newer date_2019 -exec touch {} \; &
```

3. Verify using ls -la of sample directory on a problem file system:

**/nas/quota/slot\_2/root\_vdm\_3/Avr:**

```
-rw-r--r-- 1 root bin 19932947 Oct 11 15:20 tcp.dump
-rw-r--r-- 1 root root 5 Oct 11 15:20 temp
```

## **RESETTING TIMES ON FILES:**

1. Start in directory of FS to check
2. Run command for the "creation" dates where X is the number of days BEFORE today  

```
# find . -atime X -fprint /tmp/creation_date_listing
```
3. Run command for the "modification" dates where X is the number of days BEFORE today  

```
# find . -mtime X -fprint /tmp/modification_date_listing
```
4. Examine the two files' dates:  

```
# cat /tmp/creation_date_listing
# cat /tmp/modification_date_listing
```
5. To change the dates:  

```
# find . -atime X -exec /bin/touch {} \;
# find . -mtime X -exec /bin/touch {} \;
```
6. Alternatively, update time to current time using following:  

```
# touch my_file.doc
```

**Note:** ‘touch’ command only updates inode information on file, does not change directory information

## **FINDING LARGE INODES:**

--Run “ls –ailR” in file system directory from control station mount to dm

--Find Location of Inode on File System using: 

```
#find . -inum 16823119 -exec ls -al {} \;
```

## **CREATING RECURSIVE LISTING OF ENTIRE DIRECTORY:**

```
# /eng/raylau/ls -Rail & [Creates Recursive listing of all directories and files]
```

## **USING LOCATE COMMAND TO FIND PATTERNS:**

```
# locate *[Jj]server' [Prints names of all files that end with Jserver]
```

## **SEARCHING A FILE LIST FOR WORDS:**

```
# xargs grep ndmp <file-list [Searches for all instances of ‘ndmp’ in the “file-list” file]
```

## **SEARCHING PASSWD/GROUP FILES FOR INSTANCES OF SIDs IN FIRST COLUMN:**

```
grep ^S-1-5-15 /nas/rootfs/slot_X/.etc/passwd
```

```
grep ^S-1-5-15 /nas/rootfs/slot_X/.etc/group
```

## **USING GREP TO SEARCH MULTIPLE FILES WITH MULTIPLE STRING VALUES:**

```
# egrep ‘tom | mary | sue’ /etc/{passwd,group,shadow} [Lists out all instances of Tom, Mary, and Sue]
```

```
$ /nas/server/slot_4 egrep ^514 camdisk*
```

```
camdisk:514:c49t2l0+556857258190+,c65t2l0+556857258300+:
```

```
$ server_devconfig server_3 -l -s -a | egrep "d52"
```

```
d52 c0t4l1 04B FA 0 On 000185704957 005B
```

```
d52      c16t4l1   13B FA 0 On 000185704957 005B  
d520     c49t2l6   03B FA 0 On 000185704957 025E  
d520     c65t2l6   14B FA 0 On 000185704957 025E  
d521     c49t2l7   03B FA 0 On 000185704957 025F  
d521     c65t2l7   14B FA 0 On 000185704957 025F
```

### **USING EGREP TO FIND MULTIPLE VALUES IN SINGLE FILE:**

```
$ server_devconfig server_4 -p -s -a | egrep "chain|val= 52"
```

```
chain= 49, scsi-49  
tid/lun= 2/6 type= disk sz= 13943 val= 520 info= 55685725E190  
tid/lun= 2/7 type= disk sz= 13943 val= 521 info= 55685725F190  
tid/lun= 2/8 type= disk sz= 13943 val= 522 info= 556857260190  
tid/lun= 2/9 type= disk sz= 13943 val= 523 info= 556857261190  
tid/lun= 2/10 type= disk sz= 13943 val= 524 info= 556857262190  
tid/lun= 2/11 type= disk sz= 13943 val= 525 info= 556857263190
```

```
# cat filesystems | grep "profile | 75: | ^9:"
```

**DELETING FILES/DIRECTORIES:** #rm [remove files] or #rmdir -p [remove directories & subdirectories] #rm test\* [same filenames]  
#rm -f -i -r [-f to force removal of write-protected files; -i for Y or N prompt; -r for all subdirectories; -R Solaris]  
#rm -i \* [delete all files in directory but prompts first!!] #rm -f [force deletion] #rm -R [remove directory & contents]  
#rm -Rf [super delete command]

### **VNON Errors:** 2.1.30.2 & below

Usually only occurs after a file has been deleted from Celerra FS. File is deleted but Directory entry was not properly cleaned up.

### **COPYING A FILE WHILE EMPTYING THE CONTENTS:**

```
# cp /dev/null export [copies over the file called ‘export’ with 0 bytes of data]
```

**CREATING FILES/DIRECTORIES:** #touch newfile #vi newfile #cat > newfile #cat >> newfile #cat newfile #more newfile #cat -n newfile [numbers each line] #cat test1 > test2 [copies file 1 to two] #file /etc/hosts

**Note:** ctrl + d will save a file created by “cat”

```
$cp -p old new [copies all permissions from file ‘old’ to file ‘new’] #cp -R old/ junk/
```

### **Combining Several Text Files into One:**

```
# cat file1 file2 file3 >concatenated.file [pipes results into new file]
```

### **APPENDING ONE FILE TO END OF ANOTHER:**

```
# cat bbb >> aaa
```

**SORTING FILES:** \$sort names [alphabetical sort of contents of file called ‘names’]

### **COMPARING FILE CONTENTS IN TWO DIFFERENT PATHS:**

**Example:** Comparing “netd” file of Primary Datamover & Standby Datamover

```
$ cat -n /nas/server/slot_7/netd
```

```
$cat -n /nas/server/slot_4/netd
```

### **UNIX TIMESTAMPS FOR FILES:**

\$ls -tl Lists out files in order of modification time

\$ls -lu Shows date that files were last opened or accessed

\$ls -la Shows last modification timestamp for files

\$ls -lc Shows creation date of file

### **BASH SCRIPT TO LOCATE TIMESTAMPS > THAN A SPECIFIC YEAR:** 2004 IN THIS EXAMPLE

```
#!/bin/sh  
i=0  
currentyear=2004  
$find . ! -type d -depth -exec ls -ul {} \; > /tmp/findlist num=`wc -l /tmp/findlist | awk '{print $1}'` while [ $i -le $num ] do  
    i=`expr $i + 1`  
    info=`head -$i /tmp/findlist | tail -1`  
    year=`echo $info | awk '{print $8}'`  
    filename=`echo $info | awk '{print $9}'`
```

```
echo $year | egrep ":" > /dev/null
if [ $? -ne 0 ]; then
    echo $filename
    if [ "$year" -gt "$currentyear" ]; then
        head $filename > /dev/null
    fi
fi
done
```

## **USING LIST COMMAND TO RECURSIVELY LIST OUT ALL INODES ON FILE SYSTEM:**

# ls -iaR

## **USING LIST COMMAND TO ASSESS HEALTH OF FS:**

# cd /nas; find . -ls >/dev/null

**Note:** I/O errors may be an indicator of file system issues

## **CREATING LISTING OF FILES IN VERY LARGE DIRECTORIES:**

1. Verify that file system "fs1" has available space (server\_df server\_2)

2. cd to / on CS and make a directory called "find"

#cd / #mkdir find

3. Mount fs1 to "find" mountpoint

#mount server\_2:/fs1 /find

4. cd to the directory path in question on the fs:

#cd /find/Valco/mars/cache/cache

5. Generate directory contents listing while gzipping to file:

# nohup ls -la | gzip >/find/list.out.gz &

## **CHARACTERS ALLOWED IN DOS NAMING CONVENTIONS:**

Letters A-Z; Numbers 0-9; ~ ! @ # \$ ^ & ( ) - \_ { } '

## **CHARACTERS ALLOWED IN UNIX NAMING CONVENTIONS:**

All ASCII characters except *ASCII Nul and Slash /*

Recommendation is to avoid the following characters within names: ! # & @ \$ ^ ( ) ' " ; | <> { } \* ? \[and spaces]

## **CREATING FILES OF SPECIFIC SIZES FOR TESTING:**

#mkfile 20m {k,b,m} file1 file2 file3 [makes a null file called "file1" at 20MB in size—b=bytes, k=kilobytes, etc]

### **Creating New Files Using the Line Editor [Stream Editor]: #ed newfile**

|      |                                  |  |
|------|----------------------------------|--|
| \$ed | \$a [to append to editor buffer] | \$ctrl + v to paste text or file into buffer |
| \$.  | [exits append mode]              | \$w newfile                                  |
|      |                                  | \$q [quit editor]                            |

## **SETTING ENVIRONMENTAL VARIABLES FOR LINUX CS—Red Hat 7.2:**

**Problem:** Might see an error when trying to execute commands—"NAS\_DB environment not defined"

**Fix:** Issue following command & then verify .bash\_profile file settings

1. # **NAS\_DB=/nas;export NAS\_DB**
2. Verify File: # cat /home/nasadmin/.bash\_profile
- # .bash\_profile

```
# Get the aliases and functions
if [ -f ~/.bashrc ]; then
    . ~/.bashrc
fi

# User specific environment and startup programs

PATH=$PATH:$HOME/bin

export PATH
```

unset USERNAME

**NAS\_DB=/nas**

**export NAS\_DB**

MANPATH=/usr/share/man:/usr/man:\$NAS\_DB/man

export MANPATH

PATH=\$PATH:\$NAS\_DB/bin

export PATH

/nasmcd/.emc\_login

## **APPENDING PATH STATEMENTS TO ENVIRONMENTAL VARIABLES:**

**# echo \$PATH**

/usr/local/bin:/bin:/usr/bin:/usr/X11R6/bin:/home/nasadmin/bin:/nas/bin:/usr/bin/env

**# PATH=\$PATH:/usr/bin/perl**

**Remote DialUp Session to Celerra:** Set your environmental variables to: #TERM=vt100;export TERM

**SETTING VI EDITOR:**

**# EDITOR=vi;export EDITOR**

## **LOGGING TERMINAL SESSION TO FILE:**

**\$ dtterm -l -If newlog &**

## **LINUX SUPPORTS IPCHAINS AS IP FIREWALL:**

IP Firewall Chains are enabled by default and can be configured or disabled via CLI.

**# /sbin/chkconfig --list ipchains**

ipchains 0:off 1:off 2:on 3:on 4:on 5:on 6:off

**# /sbin/service ipchains status**

Chain input (policy ACCEPT):

Chain forward (policy ACCEPT):

Chain output (policy ACCEPT):

**# /sbin/service ipchains stop | start | restart**

**# /sbin/service ipchains stop** →Temporarily stop service

Flushing all chains: [ OK ]

Removing user defined chains: [ OK ]

Resetting built-in chains to the default ACCEPT policy: [ OK ]

**# /sbin/chkconfig --level 2345 ipchains off | on** [to perm disable ipchains]

**# /sbin/ipchains -L** [Listing current firewall rules]

**# /sbin/ipchains -F** [Flushing current firewall rules that may be configured]

## **CONFIGURING IPCCHAIN RULES:**

**# /sbin/ipchains -A input -p icmp -i eth3 -j REJECT**

# /sbin/ipchains -L

Chain input (policy ACCEPT):

|        |      |       |          |             |            |
|--------|------|-------|----------|-------------|------------|
| target | prot | opt   | source   | destination | ports      |
| REJECT | icmp | ----- | anywhere | anywhere    | any -> any |

## **IPCHAIN FILES:**

**# cat /proc/net/ip\_fwchains**

|       |                                      |         |     |     |          |   |   |        |
|-------|--------------------------------------|---------|-----|-----|----------|---|---|--------|
| input | 00000000/00000000->00000000/00000000 | eth3    | 0   | 0   | 1        | 0 | 0 | 0      |
| 0     | 0-65535                              | 0-65535 | AFF | X00 | 00000000 | 0 | 0 | REJECT |

**# cat /proc/net/ip\_fnames**

/etc/sysconfig/ipchains

## **APPENDING OR INSERTING IPCCHAIN RULES:**

**#/sbin/ipchains -A input -p icmp -i eth1 -j REJECT** [Appends to end of ipchain rules]

**#/sbin/ipchains -I input 10 -p tcp -s 10.241.169.0/24 -dport 22 -i eth1 -j ACCEPT** [Inserts rule at line 10]

## **SAVING COPY OF FIREWALL CONFIG:**

**# /sbin/ipchains-save >/home/nasadmin/tm/fwall.txt**

## **RESTORING COPY OF FIREWALL CONFIG:**

**# /sbin/ipchains-restore </home/nasadmin/tm/fwall.txt**

**Important Note:** When configuring rules with IPChains, entries are parsed from top down and first match becomes effective rule.  
Place Global rules at the end of the file

**IPCHAINS INFO:** # man ipfw #man ipchains

### **SNARE [System iNtrusion Analysis and Reporting Environment]:**

Linux module that performs Control Station auditing—off by default for Celerra CS—kernel module, user-space audit daemon, and GUI reporting Tool (Snare)

# /sbin/chkconfig --list audit

audit 0:off 1:off 2:off 3:off 4:off 5:off 6:off

# /sbin/chkconfig audit on

# /sbin/service audit start

Installing Audit Module: Using /lib/modules/2.4.9-34.5406.EMC/kernel/drivers/add on/audit/auditmodule.o

Starting /usr/sbin/auditd:

SNARE audit daemon: version 0.90 starting up

### **EMC FTP Site:**

#ftp 168.159.216.19 Login: ftp Password: none needed for new FTP server

**Comments:** For “binary” files, always type “bin” at prompt. Type “ascii” before transferring text files. To watch evidence of transfer in progress, type “hash” at prompt. To ignore interactive mode, type “prompt” at command line. Use “put” or “get” or “mget \*” or “mput \*” commands to transfer up or down from FTP Server. Old FTP Server IP address was 168.159.4.20.

**FTP Commands:** cdup delete filename dir get remote lcd ls mkdir put *filename* pwd fmdir status mput mget

**Osm log:** Records FTP logins to Control Station. 40-50 FTP streams on a Celerra.

**Note:** If you are getting a “553 Permission Denied” error when trying to FTP from an NT client to the Control Station, set your transfer to “bin”; Also, do not FTP from the ‘root’ directory but from a subdirectory.

### **ALTERNATE FTP TOOL:**

# /usr/bin/ncftp 168.159.216.19

NcFTP 3.0.3 (April 15, 2001) by Mike Gleason ([ncftp@ncftp.com](mailto:ncftp@ncftp.com))

-----abridged-----

>cd /incoming/todd

> put \*

slog2.feb14: 2.20 MB 8.23 MB/s

slog2.gid32rebuild: 2.15 MB 8.90 MB/s

**Note:** Tool logs into ftp site automatically and also gives you transfer stats

### **VERIFYING FILE INTEGRITY USING VARIOUS CHECKSUM TOOLS:**

\$ cksum NAS3010.exe

3961016753 23199814 NAS3010.exe

\$ sum NAS3010.exe

14926 22657

# md5sum <filename> → Use to verify that file is the same if in different locations on the CS, for example

# md5sum clarion\_mgmt

1f682e1ea870e57b39a28d0db423dcaa clarion\_mgmt

# md5sum /nasmcd/sbin/clarion\_mgmt

1f682e1ea870e57b39a28d0db423dcaa /nasmcd/sbin/clarion\_mgmt

### **COMPRESSING DUMPS:**

\$ gzip slot3.dump.0211121824 [LINUX]

-rwxr-xr-x 1 nasadmin nasadmin 145053685 Nov 12 18:26 slot3.dump.0211121824.gz

### **USEFUL GZIP COMMAND TO VERIFY SIZE OF DUMP FILES:**

\$ gzip -l n5120102.exe.gz

|            |              |       |                   |
|------------|--------------|-------|-------------------|
| compressed | uncompressed | ratio | uncompressed_name |
| 5410362    | 18774464     | 71.1% | n5120102.exe      |

\$ gunzip -l 1781.gz

### **UNCOMPRESSING/COMPRESSING GZIPPED & TAR FILES:**

# gzip.bin -d filename.tar.Z # uncompress filename.tar.z #compress file.tar

### **UNZIPPING & UNTARRING A FILE:**

# /bin/gzip -d server\_profile.lin.tar.Z

server\_profile.lin.tar

### **UNCOMPRESSING ".gz" FILES IN LINUX:**

#/bin/gunzip compressed.gz

### **TARRING/UNTARRING FILES:**

# tar cvf tarmenow morefiles stillmore # tar xvf untar.me.tar  
\$ compress file1 \$ uncompress file1

### **EXAMPLE TAR & ZIPPING FILES TO DIRECTORY:**

\$ tar zcvf /nas/var/usrmap.tar.gz . [tars & zips all files in current directory to path specified]  
\$ tar zcvf /nas/var/usrmap.tar.gz usrmapping.cfg usrmapping\_control usrmapping\_svc \*.db

### **USING BZIP2 TO COMPRESS OR UNCOMPRESS FILES ON LINUX CS (similar to GZIP):**

# /usr/bin/bzip2 -d c.tar.bz | c.tar.bz2 [Use to uncompress files]

c.tar

# /usr/bin/bzip2 -z c.tar [Use to compress files]

c.tar.bz2

### **NAS XML:**

# nas\_xml -info:server -level:3

**Note:** Good command to obtain overall Celerra, CS, and DM hardware and software status information

# nas\_xml -info:volume -level:2

### **OTHER SOURCES OF CELERRA COMPONENT INVENTORY:**

Celerra Manager>Inventory page

# /nas/tools/.factory\_check get\_inventory

### **NAS XML COMMANDS:**

\$ /nas/bin/nas\_xml -info:fs

\$ nas\_xml -info:server

\$ nas\_xml -info:ALL

\$ nas\_xml -info:cifs

\$ nas\_xml -info:storage

\$ nas\_xml -info:volume

ALL yields an XML description of the entire system

server yields an XML description of Data Mover configurations

fs yields an XML description of Filesystem configurations

cifs yields an XML description of CIFS configurations

volume yields an XML description of Volume configurations

storage yields an XML description of Storage configurations

**\$ nas\_xml -table:filesystems**

1:root\_fs\_1:0:n:1:10::::0:::

2:root\_fs\_2:0:y:1:12:1::::0:::

----abridged, returns contents of filesystem file-----

**# date;nas\_xml -info:export -level:3;date** (to obtain list of server exports—quicker than server\_export)

Mon Mar 27 20:25:24 PST 2006

<CELERRA SRC='controlstation'>

ON>

</EXPORT>

<SHARE PATH="/restoretst" MTIME="1143489730" MODSTAMP="1143489730:62294"

IS\_SHARE="True" SHARE\_NAME="Admin\$" ALTERNATE\_NAME="" >

<OPTION>netbios=CNRESTORE011</OPTION>

<OPTION>maxusr=4294967295</OPTION>

<OPTION>umask=22</OPTION>

</SHARE>

<SHARE PATH="/restoretst/restore" MTIME="1143489730" MODSTAMP="1143489730:62294"

### **SETTING NAS XML DEBUG MODE FOR TROUBLESHOOTING:**

**# export NAS\_XML\_DEBUG=1****Note:** Dumps XML communications between DART & CS to screen and to log**VERIFYING CONTROL STATION RESOURCES:****# top shift + M** [sorts by amount of memory consumed by process]

9:50am up 4 days, 15:36, 1 user, load average: 0.00, 0.00, 0.00

119 processes: 118 sleeping, 1 running, 0 zombie, 0 stopped

CPU states: 0.7% user, 0.5% system, 0.0% nice, 98.6% idle

Mem: 513088K av, 425152K used, 87936K free, 56K shrd, 124284K buff

**Swap: 528776K av, 0K used, 528776K free 251504K cached**

PID USER PRI NI SIZE RSS SHARE STAT %CPU %MEM TIME COMMAND

932 root 9 0 3568 3568 1304 S 0.0 0.6 0:49 usrmap\_svc

-----abridged-----

**Note:** If Swap “0K used” is “528776K used” or some other high value, then Control Station will begin shutting down processes and may act erratically, failover, etc.**UNIX PROCESSES--viewing & controlling:**

Unix uses two unique processes, 0 &amp; 1. 0 is a special process created whenever the system boots &amp; holds the data structure together as the swap process. Process 1 is spawned by 0 and is the init process. All other processes originate from process 1.

\$ps -ef [lists out processes] \$ps -ef | grep usrm \$ps -ef | grep 14688 [look at specific PID]

#kill -9 PID [to hard kill a process] \$kill PID \$ps -f [Full process listing] \$ps -u 1001 [or Username—lists processes run by that User or Process ID]

#fuser -u [shows users and processes that they are running] #finger /filesystem [shows who is using FS]

#who -r [current run level] -p [current active processes] -d [displays dead processes]

#rwho [login name, name of host, login time] #who -u [user, time, process, and port logged in to]

#who #w [Linux Redhat] #id

**# pkill -9 httpd** [Useful command--kills all httpd processes]**RUNNING JOBS IN BACKGROUND AFTER EXECUTING COMMAND:**

Step 1. Execute the script that you want to run

Step 2. To Place Job in Background: **#ctrl + Z** key then run **#bg** [places job in background]Step 3. To return Job to Foreground: \$jobs [identify which job # you wish to return to foreground]  
\$fg %1 [\$kill %1 --terminates job]**VIEWING PROCESSES ON CS:****# ps -auxmore**

| USER    | PID | %CPU | %MEM | VSZ  | RSS | TTY | STAT | START | TIME | COMMAND        |
|---------|-----|------|------|------|-----|-----|------|-------|------|----------------|
| root    | 1   | 0.0  | 0.0  | 1388 | 472 | ?   | S    | Apr24 | 0:09 | init [3]       |
| rpc     | 615 | 0.0  | 0.1  | 1540 | 588 | ?   | S    | Apr24 | 0:01 | portmap        |
| rpcuser | 643 | 0.0  | 0.1  | 1584 | 624 | ?   | S    | Apr24 | 0:00 | rpc.statd      |
| root    | 827 | 0.0  | 0.1  | 2668 | 896 | ?   | S    | Apr24 | 0:03 | /usr/sbin/sshd |

**Note:** As an example, the Secure Shell Process Daemon (sshd) is using 1/10<sup>th</sup> of 1% of physical RAM (RSS=896) and 2.668MB of Virtual Memory overall**VIEWING ONLY SINGLE INSTANCE OF PROCESSES RUNNING ON CONTROL STATION:****# ps -e lawk '{print \$4}' | sort | uniq | column**

|               |              |                 |              |                 |
|---------------|--------------|-----------------|--------------|-----------------|
| apl_task_mngr | identd       | ksoftirqd_CPU0  | nas_mcd      | sort            |
| atd           | init         | kswapd          | nas_watchdog | sshd            |
| automount     | in.telnetd   | kupdated        | nviagent     | start_apl_task_ |
| awk           | java         | login           | navilog_mon  | su              |
| bash          | jexec        | log_slot        | nd-clnt      | syslogd         |
| bdfflush      | jserver_tail | log_trimmer     | portmap      | uniq            |
| cat           | js_start     | macstat         | ps           | usrmap_svc      |
| CMD           | kapm-idled   | mdrecoveryd     | rpc.statd    | xinetd          |
| column        | keventd      | mgetty          | run_jserver  |                 |
| crond         | kjournald    | nas_boxmonitor  | servmgr_svc  |                 |
| csh           | klogd        | nas_eventcollec | sleep        |                 |
| httpd         | kreclaimd    | nas_eventlog    | snmpd        |                 |

**EXTENDING CONTROL STATION VOLUME LUNS 00 & 01:**

**Note:** NAS versions prior to 2.2.35.x had 2GB partitions, and 4GB partitions after 2.2.35.x. NAS 5.2.7.0 and higher installs create 11GB partitions for LUNs 00 & 01. There is currently no approved procedure to extend LUNS 0 & 1 to 11GB, though Eng is working on the creation of a procedure.

**INSTALLING SECONDARY CS IF RUNNING NAS 5.4 BUT LUNS 0 & 1 ARE STILL AT 4GB:**

See primus emc129785

1. Shutdown Secondary Control Station (if inserted) and remove from backplane

2. Login to the primary Control Station as nasadmin then su to root.

3. # mkdir /emc

```
# mount /dev/sdd3 /emc
# cd /emc/etc/rc.d/rc3.d
# cp -ip S95nas /root
```

4. # vi S95nas and go to line that shows "check\_control\_size" string--should reside around line number 1423 (this line number could vary by the NAS code). Comment out the line, then save and quit the vi session.

```
1421 celerra_S95 ()
```

```
1422 {
```

```
1423 # check_controllun_size
```

```
1424
```

5. # umount /emc

6. Reinsert CS1 into backplane and let boot

**Note:** If a previous attempt to install CS1 had caused the following error messages, then these messages should no longer be present:  
“Control LUN 0 is not large enough. It must be at least 11GBs”

7. Answer the questions to configure root password, keyboard type, timezone, date/time, primary internal network interface, backup internal network interface and external network interface.

8. After "Installing the EMC NAS standby package..." message will be displayed. Enter username (nasadmin by default) and password. Let the installation script finishes.

9. Run /nasmed/getreason and verify CS1 is running as a secondary Control Station.

10. If possible, run /nasmcd/sbin/cs\_standby -takeover to confirm the emcnassby functionality.

**CONTROL VOLUME PARTITION NAMES & LUNS:** Symmetrix Backends

|           |                    |          |         |
|-----------|--------------------|----------|---------|
| /dev/sda1 | 133M 79M 54M 60%   | /nas/dos | →LUN 00 |
| /dev/sdb  |                    |          | →LUN 01 |
| /dev/sdc1 | 29M 2.6M 25M 10%   | /boot    | →LUN 02 |
| /dev/sdc3 | 1.4G 1.3G 146M 90% | /        | →LUN 02 |
| /dev/sdd  |                    |          | →LUN 03 |
| /dev/sde1 | 1.7G 846M 836M 51% | /nas     | →LUN 04 |
| /dev/sdf1 | 1.7G 64M 1.5G 4%   | /nas/var | →LUN 05 |

**CONTROL VOLUME PARTITION NAMES & LUNS:** Clariion Backends

|           |                    |          |   |
|-----------|--------------------|----------|---|
| /dev/nda1 | 133M 79M 54M 59%   | /nas/dos | →LUN 00 Dart Image and DM config. files |
| /dev/hda1 | 30M 11M 18M 35%    | /boot    | →LUN 02                                 |
| /dev/hda3 | 2.0G 1.1G 795M 59% | /        | →LUN 02                                 |
| /dev/nde1 | 1.7G 688M 998M 41% | /nbsnas  | →LUN 04 CS config files                 |
| /dev/hda5 | 2.0G 555M 1.3G 29% | /nas     |   |
| /dev/ndf1 | 1.7G 96M 1.5G 6%   | /nas/var | →LUN 05 Backup files, dumps, logs       |

**CELLERRA CONTROL LUNS:**

LUN 00 →contains the DOS partition (NAS image & DM Config files), /nbsnas/dos & /nas/dos, located on /dev/nda1 on the backend—Server Logs, Dumpfiles, and DOS partition are located here. LUN 00 size now 11263

LUN 01 →ufslog partition, size 11263

LUN 02 →Linux CS0 partition, size 2047

LUN 03 → Linux CS1 partition, all systems except NS Series, size 2047

LUN 04 →contains /nbsnas partition (NASDB CS config files), located on /dev/nde1, size 2047

LUN 05 →contains /nbsnas/var partition (NASDB backups, dumps, log files), located on /dev/ndf1, size 2047

**INFORMATION STORED ON LUN0:**

→DOS boot partition and NAS image

→Server Logs

→Dumpfiles

## **BOOT AND ROOT DEVICE NAMES FOR CS0 & CS1:**

**CS0=/dev/sdc1 for boot & /dev/sdc3 for root**  
**CS1=/dev/sdd1 for boot & /dev/sdd3 for root**

## **COMMANDS TO VIEW FILES IN DIRECTORIES/SUBDIRECTORIES:**

**# df -h**

| Filesystem | Size | Used | Avail | Use% | Mounted on |
|------------|------|------|-------|------|------------|
| /dev/sdc3  | 1.9G | 985M | 817M  | 55%  | /          |
| /dev/sdc1  | 7.6M | 5.7M | 1.5M  | 79%  | /boot      |
| none       | 250M | 0    | 250M  | 0%   | /dev/shm   |
| /dev/sde1  | 1.7G | 559M | 1.1G  | 34%  | /nas       |
| /dev/sda1  | 133M | 58M  | 75M   | 44%  | /nas/dos   |
| /dev/sdf1  | 1.7G | 797M | 888M  | 48%  | /nas/var   |

**# /nas/var du -sk \* |sort -g**

|         |            |
|---------|------------|
| 4       | log        |
| 16      | lost+found |
| 15604   | backup     |
| 1049648 | dump       |

**# du -sxk \* |sort -n**

|        |            |
|--------|------------|
| 12     | lost+found |
| 13     | auditing   |
| 31     | log        |
| 3592   | dump       |
| 33892  | backup     |
| 214044 | emcsupport |

**# /nas/var du -sh ./\* or |grep M**

|      |              |
|------|--------------|
| 16M  | ./backup     |
| 1.1G | ./dump       |
| 4.0k | ./log        |
| 16k  | ./lost+found |

**# du -s /etc**

**# /nas/var du -h --max-depth=1 -x |grep -i M** [grep only Megabytes]

|      |              |
|------|--------------|
| 16k  | ./lost+found |
| 16M  | ./backup     |
| 1.1G | ./dump       |
| 4.0k | ./log        |
| 1.1G | .            |

**# ls -larS**

|            |   |      |      |       |              |               |
|------------|---|------|------|-------|--------------|---------------|
| -rw-r--r-- | 1 | root | root | 7770  | Sep 19 12:16 | install.log   |
| -rw-----   | 1 | root | root | 7890  | Sep 26 09:50 | cron.1.gz     |
| -rw-r--r-- | 1 | root | root | 11624 | Sep 26 09:48 | messages.1.gz |
| -rw-r--r-- | 1 | root | root | 18666 | Sep 26 09:26 | lastlog.1.gz  |

**Note:** Sorts directory list by ascending Size

### **Linux Disk Usage:**

#du -sk foldername

## **CLEANING UP LINUX ROOT FS & SETTING LOGROTATION RULES:**

**SITUATION:** 97% full on root—cannot run compress command as it needs more space than is present to compress the /var/log/messages file [337MB]

**# df -h**

| Filesystem | Size | Used | Avail | Use% | Mounted on |
|------------|------|------|-------|------|------------|
| /dev/sdc3  | 1.9G | 1.7G | 67M   | 97%  | /          |
| /dev/sdc1  | 7.6M | 5.7M | 1.5M  | 79%  | /boot      |
| none       | 250M | 0    | 250M  | 0%   | /dev/shm   |
| /dev/sde1  | 1.7G | 1.6G | 84M   | 95%  | /nas       |
| /dev/sda1  | 133M | 58M  | 75M   | 44%  | /nas/dos   |

```
/dev/sdf1      1.7G 884M 802M 53% /nas/var  
server_6:/    126M 13M 113M 10% /nasmcd/rootfs/slot_6  
server_12:/    1   1   0   0% /nasmcd/rootfs/slot_12
```

# ls -la messages

-rw-r--r-- 1 root root 337661092 Sep 26 07:22 messages

**SOLUTION—COPY FILES TO NULL & SETUP LOGROTATION RULES:**

**1. cp -i /dev/null messages**

```
cp: overwrite `messages'? y
```

**Note:** Root went from 97% to 78% full immediately. Repeat for other logs:

**# cp -i /dev/null pacct.27.gz**

```
# cp /dev/null lastlog
```

→To reduce /nas, deleted a 1GB dumpslot in /nas/log/toftp

**2. Edit /etc/logrotate.conf:**

--Changed default logrotate from 4 weeks to 3 weeks

--Enabled compressing of log files during the logrotate

--Implemented a logrotate scheme based on size as well and set this value to 1M

--tested the new scheme by running new logrotate via cron job

**#crontab -e**

**35 \* \* \* \* /etc/cron.daily/logrotate**

**3. Edit /nas/sbin/nasdb backup and comment out following three lines:**

```
# if [ "$dir" = "$NAS_DB/var/backup" ]; then  
#   cp -p ${1}.gz $dest >/dev/null 2>&1  
# fi
```

**Note:** Lines 484-486 on test system. This will stop nasdb\_backups from going into /home/nasadmin directory, but still leaves nsdb\_backups going to [/nas/var/backup](#)

**FILESYSTEM SIZES AFTER CHANGES:**

```
# df -h  
Filesystem      Size  Used Avail Use% Mounted on  
/dev/sdc3       1.9G  971M  832M  54% /  
/dev/sdc1       7.6M  5.7M  1.5M  79% /boot  
none            250M   0  250M  0% /dev/shm  
/dev/sde1       1.7G  550M  1.1G  33% /nas  
/dev/sda1       133M  58M  75M  44% /nas/dos  
/dev/sdf1       1.7G  768M  917M  46% /nas/var  
server_6:/     126M  13M  113M  10% /nasmcd/rootfs/slot_6  
server_12:/     1   1   0   0% /nasmcd/rootfs/slot_12
```

**LOGROTATE.CONF FILE:**

```
# cat /etc/logrotate.conf  
# see "man logrotate" for details  
# rotate log files weekly  
weekly  
# rotate logs if size grows to 1 megabyte  
size 1M  
# keep 3 weeks worth of backlogs  
rotate 3  
# create new (empty) log files after rotating old ones  
create  
# uncomment this if you want your log files compressed  
compress  
# RPM packages drop log rotation information into this directory  
include /etc/logrotate.d  
# no packages own lastlog or wtmp -- we'll rotate them here  
/var/log/wtmp {  
    monthly  
    create 0664 root utmp  
    rotate 1  
}  
# system-specific logs may be also be configured here.
```

## **PACCT FILES:**

Stop Process Accounting logging temporarily by running #/sbin/accton.

Permanently stop PACCT by editing /etc/rc.d/rc.sysinit as follows, deleting the /var/log/pacct\* files, and rebooting the CS:

```
# Turn on process accounting  
#if [ -x /sbin/accton ] ; then  
    action $"Turning on process accounting" /sbin/accton /var/log/pacct  
#fi
```

## **VIEWING & REMOVING NAS DATABASE LOCKS:**

**# od -d /nas/lock/db/\***

**# ps -ef |grep -i <pid>**

**Note:** Grep the first PID from the lefthand side of the output to find out what processes might be running that are locking the database. Be aware that NASDB backups also lock the database and can affect command completion, but DO NOT seem to take a lock on the db directory. Below example is when there are no processes in the db directory:

```
# od -d /nas/lock/db/*
```

od: /nas/lock/db/\*: No such file or directory

Step 1. Identify Processes Locking DB: **# od -d /nas/lock/db/\*** or **#cd /nas/lock/db #ls -la**

**Note:** Will list out processes locking NAS\_DB, if any. However, note that NAS\_DB Backups will also lock the database at different times during the backup, but without placing a lock on the /nas/lock/db directory

Step 2. Grep on Process ID: **# ps -ef |grep 'PID#'**

**Note:** Process will be in 2<sup>nd</sup> column of output from step 1

Step 3. Determine if Process can be killed, then kill it: **# kill -9 <pid#>**

## **FINDING CALLHOME LOCKS WITH XML SITES (NAS 5.3 & higher):**

**# ls -R |grep -i Lck**

```
get_backend_status.APM00042801790.lck
```

db..LCK

./nbsnas/lock/db..LCK:

dir..LCK

ls: ./proc/4436/fd: No such file or directory

dirsync.1716.lck

dirsyncch.1717.lck

**LCK..callhome**

tier2.lck

sympainlck

sympapislck

**Note:** With introduction of XML, running the /nas/sbin/ch\_stop will delete any callhome files in /nas/log/ConnectHome, but will not remove the lock identified above

**# cd /var/lock; ls -la**

```
-rw-r--r-- 1 root root      11 May  2 15:09 LCK..callhome  
drwxr-xr-x  2 root root    4096 May  2 12:32 subsys  
-rwxr-xr-x  1 root root      0 May  2 12:31 tier2.lck
```

## **PATTERN SEARCHES:**

**#ps -ef | grep root #grep -n user1 /etc/passwd #grep 'ksh\$' /etc/passwd #grep '^s' /etc/passwd**

**#find -print |egrep "trv" [finding instances of 'trv' using search string with find and egrep]**

**\$grep symm symmpdf [searches for string 'symm' in file "symmpdf"] \$grep -x homedirs technotes [exact matches only]**

**\$grep -i cifs cifsdoc [ignores case to search for "cifs" in file 'cifsdoc']**

**\$grep string filename [Greps for specified pattern in the file called 'filename'] \$grep -c hayes bankletter [Prints number matches to screen]**

**\$grep -v string filename [Prints all lines in filename that do not contain the specified 'string']**

**\$grep -i jan b\_days > my\_month [Pipe output of grep into a file]**

**\$server\_cifsstat server\_2 legrep -e [cpu]\*\*\*\*[server] [Greps for patterns within the command output]**

## **LIST COMMAND:**

**\$ls -laR [Lists a=all l=long list format R=list subdirectories recursively--Useful to touch all files in a filesystem, etc]**

\$ls -b Show Nonprinting characters in the output  
\$ls -tail  
ls -S List output of files by Size  
ls -R List recursively through subdirectories  
ls -nl List ownership of files by numeric UID/GID values rather than default names

**\$dir /home/nasadmin**

### **SORTING LIST OUTPUT BY NEWEST FILES LAST:**

**\$ls -ltr**

### **LISTING CONTENTS OF CURRENT DIRECTORY AND OUTPUTTING TO FILE:**

**#ls -l |awk '{print \$9}' >cwd.contents**

### **SYMBOLIC LINKS:**

Symbolic, or Soft Links, are special files whose contents are a portion of the name of another file—i.e., are pointers to another file. Hard Links have multiple directory entries for a single file inode number. Celerra supports only Symbolic Links that work “down” the directory structure, not “up”. Also, it cannot link to an absolute path.

### **SYMBOLIC LINK RULES FOR CELERRA:**

- Symbolic links resolved for Windows clients only if target is within same share
- Symbolic links that refer to target paths with parent directories using ‘..’ must use param shadow followdotdot=1
- In order to backup symbolic links for Windows, set param cifs acl.extacl=8
- Celerra does not support symbolic links that contain full pathnames or absolute links
- Celerra cannot refer up the directory path from a symbolic link

**param shadow followdotdot=1**

### **SYMBOLIC LINKS WITH NAS 5.2.13:**

Celerra now has the capability to access multiple file systems on a single DM through the use of Symbolic Links created on a single file system to other file systems. To allow Celerra to traverse symbolic links between file systems, enable the following parameter:

**param shadow followabsolutpath=1**

### **SUN SOLARIS:**

\$ls \*.c [lists files ending in ‘c’]  
\$ls -F [lists files]; \$ls/ [all files & Directories]; \$ls -ail {-la hidden files} {-lu last time accessed} {-li inodes}  
**ACL's:** #getfacl [displays ACL's] filename [or ls -l to show ACL info, reflected by the + sign]; \$ls -s [for size]  
#setfacl [Solaris only and not Celerra!]  
\$ls -R [recursive] #ls -F /etc/more or #ls -F |more \$ll bin [long list of files in bin directory]  
\$ls -r [lists files & directories in reverse order] \$ls -t [lists files in order last modified] \$ls -l |more  
**Wildcards:** \* Any string of characters ? Any one character [ ] Any character within brackets

### **First Character of ls -l output for each listing:**

-- = Regular File [Regular Files hold data in ASCII, binary, image, database, or other formats]  
d = Directory [Contains only one type of data—listings of file names with inode numbers]  
l = Symbolic Link [pointers or pathnames to regular files, directories, other links, device files; #ln -s]  
b = Block special device file [contain numbers that refer to major or minor devices; I/O by block size of 8k]  
c = Character special device file [I/O operations based on sectors of 512bytes]  
**Remote Logins:** #find .rhosts [/etc/hosts.equiv; /etc/ftpusers; & \$HOME/.rhosts files] #last [last user logged in]  
#cat /etc/nodename [file shows hostname] #csh also will provide hostname & brings up C-shell  
#./install [command used to install drivers]

### **VARIOUS SOLARIS COMMANDS:**

**# df -n**

**# uname -X**

System = SunOS

Node = eagle

Release = 5.9

KernelID = Generic\_117171-12

Machine = sun4u

BusType = <unknown>

Serial = <unknown>

Users = <unknown>

OEM# = 0

Origin#= 1

NumCPU = 1

# **prstat** (similar to Top)

### **SUN SOLARIS SYSTEM MANAGEMENT COMMANDS:**

```
# prtconf -v [Prints out the System configuration] # showrev -p [Current patches for the Solaris Version]
# pkginfo -i [Lists installed software packages] # pkgadd/pkgrm [Install/Uninstall software packages]
# devinfo [disk device information] # prvtoc <devicename> [Volume Table of Contents of device]
```

### **SUN SOLARIS LOGS:**

|                     |  |
|---------------------|--|
| /var/adm/messages   | [System Messages]                        |
| /etc/system         | [System kernel parameters]               |
| /etc/vfstab         | [List of mountable file systems]         |
| /kernel/drv/sd.conf | [list of available target ID's and LUNs] |

### **SUN SOLARIS PERFORMANCE:**

```
# truss -p <pid> -t <system call>
$ iostat -d
$ sar -d 15 15
```

### **USING DD TO CREATE SYSTEM LOAD:**

**#dd -f=/dev/zero of=/dev/rdsckntdns2 skip=8 bs=512 count=1000**

**PASSWD FILE ON SOLARIS SYSTEMS:** Comprised of Seven Fields [/etc/passwd /etc/group /etc/shadow]

**CommandLine:** #useradd                   **GUI:** #admintool [invokes GUI interface]

**EXAMPLE:** user1 : 2050 : 1005 : 100 : Comments Here : HomeDir : login\_shell

**Description:** user   /etc/shadow placeholder   UID   GID   comment field   Home Dir.   Login shell

**SUN SOLARIS STATISTICS:** \$iostat -d [with a variety of possible switches]

**Output of Extended Disk Statistics:** r/s=reads/sec w/s=writes/sec Kr/s=Kbytes read/sec Kw/s=Kbytes written/sec

Wait=aver # trans. waiting svc [queue length] actv=aver. # trans. being serviced svc\_t=aver.sve time in miliseconds

%w=percentage of time queue is not empty %b=percentage time disk is busy

\$sar -d 15 100 [reports disk activity every 15 seconds for 100 times]

### **Using Ping Command to Debug a Network Route Issue:**

Will record roundtrips on a network. Greater than several milliseconds could indicate a busy network.

\$ping sRv servername [s=1datagram per second R=record the route inside IP header v=verbose, lists any ICMP packets]

### **SETTING SPEED AND DUPLEX SETTINGS ON SUN SOLARIS HME NIC CARDS:**

**Step 1. Run the following Commands to Set 100MB Full Duplex on the HME Card:**

**Note:** This example assumes that “auto-negotiation” is “on”

```
#ndd -set /dev/hme instance 0
#ndd -set /dev/hme adv_autoneg_cap 0
#ndd -set /dev/hme adv_100T4_cap 0
#ndd -set /dev/hme adv_100fdx_cap 1
#ndd -set /dev/hme adv_100hdx_cap 0
#ndd -set /dev/hme adv_10fdx_cap 0
#ndd -set /dev/hme adv_10hdx_cap 0
```

**Step 2. Set the following Value on the Switch from the CommandLine:**

```
#ndd /dev/hme lp_autoneg_cap
```

**Step 3. Setting HME Card Speed and Duplex Values:**

```
#ndd /dev/hme link_speed 1 | 0 [1=100MB 0=10MB]
#ndd /dev/hme link_mode 1 | 0 [1=Full Duplex 0=Half Duplex]
```

### **CHECKING SUN SOLARIS HME CARD SPEED & DUPLEX SETTINGS:**

```
# ndd /dev/hme link_speed
# ndd /dev/hme link_mode
```

### **Configuring IP Address on Sun Solaris:**

```
# ifconfig hme0 192.10.3.8 netmask 255.255.255.0 broadcast 192.10.3.255
# ifconfig hme0 up | down [Interface Up or Down]
```

### **Configuring Gateway Address for Solaris: [Two Methods for Doing This]**

1. Commandline:                   #route add default 192.10.3.254
2. Create File:                   #vi /etc/defaultrouter [add IP address of Gateway—becomes permanent Gateway]

## **Verifying and Changing HME Interface Settings/Mode on the Sun Solaris 2.x Server:**

# ifconfig -a

```
lo0: flags=1000849<LOOPBACK,RUNNING,MULTICAST,Ipv4> mtu 8232 index 1 inet 127.0.0.1 netmask ff000000
hme0: flags=1000843<UP,BROADCAST,RUNNING,MULTICAST,Ipv4> mtu 1500 index 2 inet 152.62.37.193
netmask ffffff00 broadcast 152.62.37.255 ether 8:0:20:f5:41:3a
```

### **To Obtain HME Link Status:**

```
# ndd -get /dev/hme link_status [0:link down 1:link up]
```

### **To Obtain HME Link Speed:**

```
# ndd -get /dev/hme link_speed [0:10m 1:100m]
```

### **To Obtain HME Link Mode:**

```
# ndd -get /dev/hme link_mode [0:half duplex 1:full duplex]
```

### **To Change Duplex Settings Manually on Solaris:** /etc/system

```
#set hme:hme_adv_autoneg_cap=0
#set hme:hme_adv_100T4_cap=0
#set hme:hme_adv_100fdx_cap=0
#set hme:hme_adv_100hdx_cap=0
#set hme:hme_adv_10fdx_cap=0
#set hme:hme_adv_10hdx_cap=0
```

### **To Change Speed Settings Manually on Solaris:** /etc/rc2.d/S69inet

```
#ndd -set /dev/hme adv_autoneg_cap 0
#ndd -set /dev/hme adv_100T4_cap 0
#ndd -set /dev/hme adv_100fdx_cap 0
#ndd -set /dev/hme adv_100hdx_cap 0
#ndd -set /dev/hme adv_10fdx_cap 0
#ndd -set /dev/hme adv_10hdx_cap 0
```

## **SOLARIS 7 TCP ACK PARAMTER:**

tcp\_xmit\_lowat [Minimum value of TCP transmit buffers, default = 2048 (2k)]

## **SOLARIS NFS SHARES:**

Troubleshoot using the “share” command

|                   |  |
|-------------------|--|
| /etc/dfs/dfstab   | List of files to export for the Server   |
| /etc/dfs/sharetab | List of FileSystems exported   |
| /etc/rmtab        | FileSystems remotely mounted on the system, maintained by rpc.mountd           |
| /etc/vfstab       | List of files to mount on a Client   |
| /etc/mnttab       | List of FileSystems mounted on a Client [modified by mount and umount command] |

## **DEFAULT MOUNT FOR SOLARIS IS NFSv3 OVER TCP/IP:**

To change to NFSv2 Over UDP:

**\$mount -o vers=2,proto=udp server:/ts /mnt**

**Mounting Sun Client to DataMover:** Don’t forget to mount Celerra’s FS to a mountpoint that you can use!!!

1. First ensure that the File System is “exported” for NFS!

#mkdir /share1

From / mount the remote file system: # mount 193.1.21.173:/g1ufs1 /share1 [remote mntpoint & local mntpoint]

**Unmounting Sun from Celerra:** # umount -a /mntpoint or # umount /mntpoint

**Logging Onto Local Unix Machine as another user:** #exec login

## **RULE ON HARD vs. SOFT MOUNTS: If you are conducting Read-Writes to a Celerra FS, use Hardmounts!**

**Adding Devices in Solaris:** #devfsadm -c disk -c tape -c audio [etc] #devfsadm -i st #devfsadm -i dad #devfsadm -i sd

**Adding a New Disk Device:** #drvconfig -i dad [sd] #disks dad=ide sd=scsi

**Adding a New Tape Device:** #drvconfig -i st #tapes

**Logging In to Solaris from Root with Different Account:** #exec login login: todd password: nasadmin

**Continuous Ping from Solaris for Testing:** #ping -s 193.1.21.172

**Viewing Users on File Systems from Sun Client:**

```
#fuser -cu mnt_point [lists active processes] #fuser -ck mnt_point [Kill all processes accessing File System]
```

## **SETTING SUN SOLARIS CLIENTS TO NOT CACHE NEGATIVE GETATTR HITS:**

**param nfs:nfs3\_lookup\_neg\_cache=0**

**Note:** Functionality is similar to Celerra’s DNLC behavior in that if file lookups are not found in cache, a negative entry will be put in cache. Disable this behavior for NFSv3 on Solaris by using the above param. May help with issues related to Checkpoints.

## **TUNING SUN SOLARIS SYSTEMS FOR ENHANCED PERFORMANCE:**

## **GREATER READ PERFORMANCE:**

Step 1. Add the following lines to /etc/system and then reboot Server

```
set nfs:nfs_nra = 10  
set nfs:nfs3_nra = 10  
set nfs:nfs_max_threads = 24  
set nfs:nfs3_max_threads = 24  
set sq_max_size={ val }  
where { val } = 25 for each 64MB memory of client
```

## **GREATER STREAMS PERFORMANCE:** Change the following TCP/UDP parameters

```
ndd -set /dev/udp udp_recv_hiwat 65535  
ndd -set /dev/udp udp_xmit_hiwat 65535
```

### **ndd -set /dev/tcp tcp\_recv\_hiwat 65535 [Solaris TCP Window Size]**

```
ndd -set /dev/tcp tcp_xmit_hiwat 65535
```

**Note:** To make permanent, add to /etc/rc2.d script. Also, the sq\_max\_size and ndd settings are recommended for purely sequential workloads!

## **TUNING SOLARIS 2.8 CLIENTS FOR IMPROVED PERFORMANCE:**

```
tcp_xmit_hiwat=65536  
tcp_recv_hiwat=65536  
tcp_max_buf=4194304  
tcp_cwnd_max=2097152  
/etc/system:
```

Set maximum Read Ahead threads to 10 per process and up to 80 per client

Set queue buffer size to 200

## **SUN SOLARIS ETHERNET PARAMS:**

```
$pkginfo -l |grep -i gig [info about GB Ethernet card in use]  
$ndd -get /dev/ge \? [this will show you what settings are available]  
#/etc/power.conf [This file can be changed to automatically shutdown the Solaris]
```

## **IBM AIX SYSTEMS:**

```
#lscfg -pv [Displays configuration and vital stats]  
#lsps -a [snapshot of pagefile utilization levels]  
/etc/security/audit/config [Audit Files]
```

## **HP-UX 11.0:**

Unix Platform based on Unix System V Release 4

Network Traces: Use the 'nettl' utility

```
#sysdef [Analyzes current system and configuration parameters]  
#swlist -l PH* [displays installed HP-UX patches] #swlist -l bundle [OS Version & type]  
#ioscan [Scans system hardware] #lsdev -C disk [lists device drivers in system]  
/var/adm/syslog.log /stand/system
```

## **HP-UX TCP WINDOW SIZE:**

```
#ndd -get /dev/tcp $i  
tcp_recv_hwater_def 32768 [Max Receive Window Size]  
tcp_recv_hwater_lfp 65536 [Max Receive Window Size for fast links]  
tcp_recv_hwater_lnp 8192 [Max Receive Window Size for slow links]
```

## **CELERRA FTP/DATA MOVER FTP SERVICE:**

→Data Mover supports both NFS and CIFS for FTP, using either local /etc/hosts or NIS for NFS authentication, or CIFS authentication for CIFS. If a user logs in without the domain name format (domain\user), then the login uses Unix authentication. If the user logs in with domain\user account, and CIFS is running, then the login will attempt to use CIFS authentication.

→Default directory is root “/” for users logging into FTP, but access will not occur if they do not have permission to “/”

→Use defaultdir param to change working directory to a defined mountpoint on a file system

→Use the Celerra Homedir service to allow Users to ftp to their Home Directory

→For FTP CIFS Users using Celerra Homedir service, the following entry must also be made to the netd file:

**ftpd homedir**

**ftpd umask=027** (use this entry to change default umask for ftp users)

## **SERVER FTP & NAS 5.6.42:**

→Introduces the server\_ftp command and options, which allow for much easier configuration and changes to the FTP service on the Blade. Please note that FTP service changes will require that you first stop the service, do the change, then start the service.

**Restricting Users to Home Directory:**

# server\_ftp server\_2 -service -stop

server\_2 : done

# server\_ftp server\_2 -modify -homedir disable

server\_2 : done

# server\_ftp server\_2 -service -start

# server\_ftp server\_2 -service -status

server\_2 : done

State : running

→Default timeout for an active ftp session is 15 minutes

→Access-checking depends on the access-policy in effect, and whether a user is coming in from Unix/FTP client, or CIFS/FTP

→FTP also supports login with Anonymous or FTP user account, and requires that the “ftp” username be resolved from the local /etc/passwd file or from NIS—password is generally the username or the email address of the user—the passwd field for the entry must be disabled! (i.e., \*). The “home” directory of the FTP user will be what is defined in the passwd entry, or if not defined, and the ftpd defaultdir param is not changed, the default directory is “/” root. If FTP user has permissions to root, login will succeed—if not, login will fail.

**Note:** Normal behavior should be that ‘anonymous’, or ‘ftp’ user, or a cifs user, will not be able to log into “/” root unless perms are defined to allow it

**FTP vs. FTPS vs. SFTP (5.6.42):**

FTP authentication goes across the network unencrypted. FTPS authentication uses SSL security to pass user and password in encrypted format. SFTP uses the SSH protocol to encrypt data over the transport mechanism, and uses Public Key authentication and compression. The latter feature is available on the Control Station Linux O/S. Data Movers support only FTP and FTPS.

**CONFIGURING THE DATA MOVER FTP SERVICE FOR CIFS USERS (NAS 5.6.42+):**

**Note:** Setting up CIFS FTP is a little more complicated than NFS or Anonymous FTP. You would probably want to make sure the CIFS user is locked into a particular directory tree path, which requires setting up the Celerra HomeDir service with the autocreate user directory option, and setting the FTP HomeDir feature to “enable”. This will serve to map each CIFS FTP User to their own home directory, and they will not be able to ‘cd’ up the file system tree.

**Steps for configuring FTP for CIFS Users:**

1. Install the Celerra Home Directory Management snapin:

a) From the Celerra Application & Tools CD, install the "Celerra CIFS Management MMC Snapins" on a Windows platform [CelerraCifsMgmt.exe]

**Note:** Accept the default paths during the install, but on the "Select Components" screen, uncheck all boxes except for the "Home Directory Management" service

2. Enable the Home Directory service:

a) From the Programs or Administrative Tools section, open the "Celerra Management" program

b) Rightclick "Data Mover Management" and connect to the appropriate Data Mover using "Connect to Data Mover"

c) Rightclick the "HomeDir" section and "Enable" the service

**Note:** This starts the Home Directory service and creates a blank Home Directory map file /etc/homedir file

# server\_cifs server\_2 |grep Home

Home Directory Shares ENABLED, map=/etc/homedir

d) Use the Celerra Management HomeDir interface to add entries to the map file

**Example:**

# cat /etc/homedir

test.ftp.emc.com:\*:/fs3/ftp/cifs:create

**Note:** The first colon-delimited field represents the Windows Domain FQDN, the second field represents all “users” for the specified domain when using the “\*” wildcard, the third field defines the FTP Home Directory path, and the last field “create” serves to automatically create the User’s FTP directory upon first login to the FTP service

3. To lock CIFS Users into their own Home Directory, enable the FTP Home dir option:

# server\_ftp server\_2 -service -stop

server\_2 : done

# server\_ftp server\_2 -modify -homedir enable

server\_2 : done

FTPD CONFIGURATION

=====

State : stopped

Control Port : 21  
Data Port : 20  
Default dir : /  
Home dir : enable  
----output abbreviated-----

# server\_ftp server\_2 -service -start

server\_2 : done

**Note:** Setting -homedir enable makes sure that CIFS Users will not be able to navigate out of their own Home Directory. If sub-directories exist, the User would be able to navigate down the tree, provided they had NT permissions. Without the Homedir enable option set, a CIFS user could navigate up the file system path and even to the rootfs of the Data Mover, provided the User had the requisite NTFS permissions.

#### 4. Log into the Data Mover's FTP service using a fully qualified user name and password:

c:>ftp 192.1.10.27  
Connected to 192.1.10.27.  
220 server\_2 FTP server (EMC-SNAS: 5.6.43.8) ready.  
User (192.1.10.27:(none)): user@test.ftp.emc.com  
331 Password required for [user@test.ftp.emc.com](mailto:user@test.ftp.emc.com).  
Password:

230 NT User [user@test.ftp.emc.com](mailto:user@test.ftp.emc.com) logged in.

ftp >

#### 5. Configuring Celerra FTPS (FTP Secure, using SSL):

a) Create Server Certificate:

# server\_certificate server\_2 -persona -generate default -key\_size 4096 -cs\_sign\_duration 8 -cn "w2k.pvt.dns"  
server\_2 :

Starting key generation. This could take a long time...

done

### **CONFIGURING THE DATA MOVER FTP SERVICE FOR NFS USERS(Anonymous, etc.):**

1. Create an ftp user “ftp” with UID and GID values, but no password (i.e., :x:), by editing ./etc/passwd file on the Server, or by using a NIS entry

# /nas/sbin/server\_user server\_2 -add ftp

Creating new user ftp

User ID: 400

Group ID: 400

Comment:

Home directory: /fs3/ftp

Shell:

# cat passwd

SL7E10817000220000\_SL7E10817000220000:EFkc/WtYmB.S.:9000:9000:fDCZfJrevWsulnaQco0lkph9OV::ndmp\_md5

**ftp::!400:400::/fs3/ftp:**

2. Disable the password field for the user “ftp” either manually using a vi editor, or the following command:

# /nas/sbin/server\_user server\_2 -passwd -disable ftp

Changing password for user ftp

Removing password for user ftp

# cat passwd

**ftp::400:400::/fs3/ftp:**

**Note:** In testing NFS, it didn’t matter that the passwd value for ftp user was \*, !!, or just blank. The user was allowed to ftp in.

3. Define a home directory path in the “passwd” file for the FTP user as shown above, or by specifying a default directory with the FTP service

# server\_ftp server\_2 -modify -defaultdir /fs3/ftp

**Note:** A Home Directory can be specified either in the passwd/NIS file or using the –defaultdir command. If no home directory is defined, the working directory would be “/” on the Data Mover. If ftp has not been granted access permissions to “/”, then the login would fail, otherwise, the User would gain access to root. Normally, the FTP User will only be able to navigate downward on the file system, not upward, so having the default directory set is important.

4. Having the homedir value “disable” or “enable” made to difference to NFS user

### **FTP PARAMS:**

\$ .server\_config server\_2 -v "param ftpd"

| Name           | Location   | Current    | Default    |
|----------------|------------|------------|------------|
| ftpd.autostart | 0x03500458 | 0x00000001 | 0x00000001 |

```
ftpd.bounceAttackChk      0x035006d8 0x00000001 0x00000001
ftpd.cifsbypass          0x03500604 0x00000001 0x00000001
ftpd.defaultdir           0x035004d8 '/' '/'
ftpd.forceBinXfer         0x03500758 0x00000000 0x00000000
ftpd.longDirA             0x035408b8 0x00000001 0x00000001
ftpd.shortpathdir         0x03500698 0x00000000 0x00000000
ftpd.wildcharsInDir       0x03500658 0x00000000 0x00000000
```

### # server\_param server\_2 -f ftpd -l

```
server_2 :
param_name          facility default current configured
shortpathdir         ftpd     0      0
defaultdir           ftpd     /      /
wildcharsInDir       ftpd     0      0
bounceAttackChk     ftpd     1      1
```

### USING SERVER FTP:

#### # server\_ftp server\_2 -info

##### FTPD CONFIGURATION

```
=====
State          : running
Control Port   : 21
Data Port      : 20
Default dir    : /
Home dir       : disable
Keepalive      : 1
High watermark : 65536
Low watermark  : 32768
Timeout        : 900
Max timeout    : 7200
Read size      : 8192
Write size     : 49152
Umask          : 27
Max connection : 65535
```

##### SSL CONFIGURATION

```
=====
Control channel mode : disable
Data channel mode   : disable
Persona            : default
Protocol           : default
Cipher              : default
Control port       : 990
Data port          : 989
```

### CHANGING DEFAULT DIRECTORY FOR USERS LOGGING INTO FTP SERVICE:

#### # server\_param server\_2 -f ftpd -modify defaultdir -value /fs1/ftp

### DISABLING FTP SERVICE ON A DATAMOVER:

#### # server\_ftp server\_2

```
server_2 :
Error 2100: usage: server_ftp { <movername> | ALL }
               -info
               | -service { -status | -start | -stop }
               | -service -stats [-full | -reset]
               | -modify
```

### Verify that FTP service is running on Server:

#### # server\_netstat server\_2 -a | grep ftp

```
tcp *.ftp          0.0.0.0.0      LISTEN
```

Step 1. \$cd /nas/server/slot\_x

Step 2. Comment out the “ftpd” line in the “netd” file and reboot Data Mover, then verify

# server\_netstat server\_2 -a |grep ftp

### **CHANGING DEFAULT UMASK FOR FTP:**

1. #vi /nas/server/slot\_x/netd [Add following line to netd file]

**ftp umask=022**

### **FORCING DM TO TRANSFER IN BINARY MODE:**

**param ftpd forceBinXfer=1**

**Note:** Reboot DM after setting param. Default=0

### **SETTING DATA MOVER FTP TO LONG DIRECTORY LISTING MODE:**

**param ftpd listDirA=1**

**Note:** See AR47406

### **TROUBLESHOOTING TELNET/FTP:**

/etc/services [Linux--look for telnet entry--23/tcp]

/etc/inetd.conf [SCO--look for entry for telnet or ftp and make sure that it is uncommented]

/usr/sbin/in.telnetd [verify the file exists & permissions o.k.]

**Problem:** Some firewall implementations use an /etc/hosts.allow file to restricted Host connections from FTP or Telnet

### **TCP WRAPPER SERVICE:**

Available and configurable on Control Station using /etc/hosts.allow and /etc/hosts.deny files, works with inetd service. Always use at least one backup external IP address and both internal IP addresses in the hosts.allow file to prevent from being locked out.

### **EXAMINING TELNET SESSIONS ON A PARTICULAR HOST:**

**# netstat -a -n inet**

**Note:** Shows TCP connections and status of

### **DISABLING LINUX SERVICES ON DM:**

--Commenting out /nas/server/slot\_x/netd/#routed [Shuts off RIP service on DM]

--pax & ndmp port=10000 [Shut off if no local tape backup or ndmp in use]

--xattrp [Shutoff if not using EDM to backup CIFS attributes]

--rquotad action=start [Shutdown if quotas not in use]

--ftpd

--sysman [Shuts down snmp]

### **BEST WAY TO DISABLE SNMP DAEMON ON DATA MOVER:**

1. Remove “sysman” line from /nas/server/slot\_x/netd file

2. Reboot Data Mover

**Note:** There are sideaffects to running with SNMP disabled on the Data Mover. It disables the server\_netstat command, but more importantly, will prevent success NAS Upgrades, therefore, the daemon MUST be re-enabled prior to any NAS Upgrade.

### **IMPLEMENTING QUOTAS ON LINUX:**

1. #rpm -qa | grep quota

quota-3.09-1

2. #vi /etc/fstab and change ‘defaults’ to ‘usrquota,grpquota’

3. #mount -o remount /data

4. #touch /data/aquota.{user,group}

5. #quotacheck -cm /data

6. #quotaon /data

7. #edquota user1 [Using prototype: #edquota -p user1 user2]

8. Checking a User’s Quota: #quota user1

9. #edquota -t [establishing grace periods]

### **APACHE SYNTAX CHECKER:**

**# httpd -t**

### **TURNING ON TELNET/FTP SERVICES ON LINUX:**

1. Check /etc/services file and ensure Telnet/FTP are not commented out

2. **# /sbinchkconfig --list**

**# /sbin/chkconfig --list telnet**

- telnet off
- 3. Turn Telnet Service on:

**# /sbin/chkconfig telnet on**

- 4. Start Inet Services:

**# /sbin/service xinetd restart**

Stopping xinetd: [ OK ]

Starting xinetd: [ OK ]

**Note:** Using chkconfig method will configure service to restart automatically after reboots

## **VARIOUS SERVICES THAT CAN BE DISABLED ON LINUX:**

**Note:** /etc/services directory based on Red Hat 7.2

### **# Network services, Internet style**

```
systat      11/tcp    users
systat      11/udp    users
ftp-data    20/tcp
ftp-data    20/udp
ftp         21/tcp
ftp         21/udp
ssh          22/tcp    # SSH Remote Login Protocol
ssh          22/udp    # SSH Remote Login Protocol
telnet      23/tcp
telnet      23/udp
smtp         25/tcp    mail
smtp         25/udp    mail
domain      53/tcp    nameserver # name-domain server
domain      53/udp    nameserver
http         80/tcp    www www-http # WorldWideWeb HTTP
http         80/udp    www www-http # HyperText Transfer Protocol
kerberos    88/tcp    kerberos5 krb5 # Kerberos v5
kerberos    88/udp    kerberos5 krb5 # Kerberos v5
ntp          123/tcp
ntp          123/udp   # Network Time Protocol
netbios-ns  137/tcp
netbios-ns  137/udp   # NETBIOS Name Service
netbios-dgm 138/tcp
netbios-dgm 138/udp   # NETBIOS Datagram Service
netbios-ssn  139/tcp
netbios-ssn  139/udp   # NETBIOS session service
snmp        161/tcp
snmp        161/udp   # Simple Net Mgmt Proto
snmptrap   162/udp   snmp-trap # Traps for SNMP
ldap         389/tcp
ldap         389/udp
microsoft-ds 445/tcp
microsoft-ds 445/udp
kpasswd     464/tcp   kpwd     # Kerberos "passwd"
kpasswd     464/udp   kpwd     # Kerberos "passwd"
```

### **# UNIX specific services**

```
exec        512/tcp
talk        517/udp
```

**Note:** Above listing is just a sampling of contents listed in the /etc/services file

## **CELLERRA PORTS:**

### **IMPORTANT LINUX SERVICES AND PORT NUMBERS:**

|         |                  |                             |   |
|---------|------------------|-----------------------------|---|
| ssh     | Port 22/tcp      | Port state open             | Use /sbin/service sshd stop to stop service               |
| http    | Port 80/tcp      | Port state open             | Edit /nas/http/conf/httpd.conf file                       |
| http    | Port 8000/tcp    | Port state open             | Core Service for Celerra Monitor, WebUI—do not stop       |
| http    | Port 5080/tcp    | Port open                   | Core service, do not stop                                 |
| rpcbind | Port 111/tcp     | Port state open, aka sunrpc | Mandatory NFS Portmapper service—do not stop              |
| smux    | Port 199/tcp/udp | Port state open             | Mandatory snmpd service & ECC—do not stop                 |
| java    | Port 8009/tcp    | Port state open             | Core Java Service & ECC—do not stop                       |
| java    | Port 8010/tcp    | Port state open             | Core Java link Service & ECC, Celerra Monitor—do not stop |
| https   | Port 443/tcp     | Port open                   | Mandatory core service—do not stop                        |
| nas_mcd | Port 9823/tcp    | Port open                   | Master Control Daemon—Core Service—do not stop            |
| SMB     | Port 139/tcp     | Port open                   | Netbios Port 139—CIFS over netbios—can be stopped         |

|            |                      |                              |  |
|------------|----------------------|------------------------------|--|
| SMB        | Port 445/tcp         | Port open                    | Direct Hosted SMB over TCP on Port 445—CIFS services               |
| Nbt/wins   | Port 137/tcp         | Port open                    | Windows netbios name service, can be stopped                       |
| nbs        | Port 5033/tcp        | Port open                    | Core interconnection NS600 service, do not stop                    |
| rquotad    | Port 1024/tcp        | Port open                    | Quota daemon service, core service, do not stop                    |
| rpc.statd  | Port >1024/tcp/udp   | Port open                    | Dynamically allocated port. Close using /sbin/service nfslock stop |
| NIS        | Port 1038/tcp        |                              |  |
| Nlockmgr   | Port 1039/tcp        | [nfs & cifs lock management] |  |
| Mountd     | Port 1234/tcp        | Port open                    | Core service, mount daemon for nfs, do not stop                    |
| NFS        | Port 2059/tcp or udp | Port open                    | Core service, nfs daemon, do not stop                              |
| PAX        | Port 4658/tcp        | [Pax operation]              |  |
| NDMP       | Port 10000/tcp       | Port open                    | NDMP operation, core service, do not stop                          |
| Snmp       | Port 161/udp         | Port open                    | Core Celerra snmp service, do not stop                             |
| Mac        | Port >1024           | Port open                    | Core interconnection service, do not stop                          |
| Lockd      | Port >1024           | Port open                    | Core Celerra service, do not stop                                  |
| Statd      | Port > 1024          | Port open                    | Core Celerra service, do not stop                                  |
| Rcp        | Port 8888/tcp        | Port open                    | Celerra replicator service on Secondary side—can be closed         |
| Mgfs nfs   | Port 1020/udp        | Port open                    | Mgfs NFS Migration service, core service, do not stop              |
| Mgfs mount | Port 1021/tcp or udp | Port open                    | Core NFS Mgfs migration service, do not stop                       |
| Replicator | Port 8887/tcp        | Port open                    | Celerra IP ‘DR’ Replication service                                |
| iSCSI      | Port 3260/tcp        | Port open                    |  |

## **SETTING UP RSH REMOTE SHELL:**

Step 1. Create the “.rhosts” file in /home/nasadmin and enter client IP addresses:

```
#vi /home/nasadmin/.rhosts
192.10.4.4      root
192.10.4.5      nasadmin
localhost        nasadmin
```

Step 2. # chmod 600 .rhosts

Step 3. Turn on ‘rsh’ service: #/sbin/chkconfig rsh on

## **CREATING SSH KEYS IN USER’s HOME DIRECTORY:**

1. #su - user1
2. #ssh-keygen -t rsa -b 1024
3. #ssh-keygen -t dsa -b 1024

## **CREATING PUBLIC/PRIVATE ENCRYPTION KEYS ON LINUX CS:**

1. Generate the encryption key  
# ssh-keygen -t rsa  
Generating public/private rsa key pair  
Enter file in which to save the key ...
  2. Do not enter a passphrase, just hit enter key  
Enter passphrase (empty for no passphrase):
  3. Your identification has been saved in ...  
Your public key has been saved in ...
- Note:** Note the name and location of the public key that was generated—key ends in .pub
4. Copy appropriate public key to location on remote systems

## **CELERRA LINUX CS AND SECURE SHELL SERVICES (SSH—aka Secure Socket Shell):**

### **Note 1:** sshd runs on NAS 4.1 + by default

Celerra uses OpenSSH on the Linux Control Station. SSH uses RSA public key cryptography for connections and authentication and provides for secure encrypted passwords and communications between a Local and Remote Host using the SSHD Daemon. SSH works much the same as Telnet but is secure & encrypted (Telnet & FTP use plaintext communications) but is designed to replace “rlogin” and “rhosts”. SSH also automatically forwards the local host’s DISPLAY variable to the Client system so as to be able to execute an X term or other such program. Linux Redhat 6.2 & 7.2 Control Stations are configured with SSH by default.

### **Note 2:** rlogin and rlogind are similar to telnet and telnetd, with exception that rlogin does not require password to connect to host.

Rsh and remsh are similar to rlogin, requiring setup of .rhosts and hosts.equiv files.

→Default port for SSH is Port 22 (required for Celerra Manager, so do not change)

→/etc/ssh/sshd\_config is configuration file

→Root access via SSH is enabled by default, but can be disabled

**Step 1. Verify Default System Configuration File for SSH:**

# more /etc/ssh/sshd\_config

# \$OpenBSD: sshd\_config,v 1.48 2002/02/19 02:50:59 deraadt Exp \$

# This is the sshd server system-wide configuration file.

**Step 2. Verify SSHD Daemon (should be running by default):**

# ps -ef |grep sshd

root 904 1 0 Mar21 ? 00:00:00 /usr/sbin/sshd

# /sbin/chkconfig --list sshd

sshd 0:off 1:off 2:on 3:on 4:on 5:on 6:off

# /sbin/service sshd status

sshd (pid 904) is running...

**Step 3. Starting & Stopping the SSH Daemon/Service:**

# /sbin/service sshd start | stop | restart

Starting sshd: [ OK ]

**Step 4. Add Remote Hosts to Local "/etc/hosts" File:**

#vi /etc/hosts [Add entries to end of file!]

172.24.80.13 artmis

**Step 5. Login to Remote Host from Prompt :**

# ssh artmis

# ssh nasadmin@artmis

[Login as a different User: #ssh -l root artmis]

The authenticity of host 'artmis (172.24.80.13)' can't be established.

RSA key fingerprint is 88:ec:66:76:a1:13:ea:0f:d3:7e:18:ad:2f:a5:c9:3a.

Are you sure you want to continue connecting (yes/no)? Yes

Warning: Permanently added 'artmis,172.24.80.13' (RSA) to the list of known hosts.

root@artmis's password:\*\*\*\*\* [Enter Remote Host's Password here]

Last login: Wed Oct 30 14:44:47 2002

EMC Celerra / RedHat Linux Wed Oct 30 14:09:29 UTC 2002

server\_2 -

server\_3 -

**LOGGING INTO REMOTE HOST USING SSH:**

# ssh -l root 193.1.21.200

# ssh nasadmin@192.168.25

**DEFINING NAS\_DB ENV WHEN RUNNING CELERRA CMDS ON REMOTE CS USING SSH:**

# ssh cs0 "declare -x NAS\_DB="/nas"; /nas/bin/nas\_quotas -r -fs fs01

**CHANGING DEFAULT SSH-1 TO SSH-2 PROTOCOL:**

1. Edit /etc/ssh/sshd\_config file, changing following line:

#Protocol 2,1

2. Uncomment above line and change to this:

Protocol 2

3. Restart ssh: #/sbin/service sshd restart

**NAS 5.6 LONG LOGON TIMES USING SSH CLIENT:**

**Note:** emc221091 has an effective workaround for long delays when logging into the Control Station from a Remote SSH client, but the problem is really a remote host DNS resolution issue

**Workarounds for Long Logon Times:**

1. vi edit /etc/resolv.conf and remove domain & nameserver specific search entries (leave only “domain localdomain” on one line)

2. Or, vi edit /etc/nsswitch.conf and remove DNS from following line

#hosts: db files nisplus nis dns

hosts: files dns → Remove dns from this line

3. Or, vi edit /etc/ssh/sshd\_config file, change the following line, restart the sshd service:

#UseDNS yes → Uncomment the entry and edit as follows:

**UseDNS no**

# /sbin/service sshd restart

**KNOWN CELERRA LIMITATIONS:**

--boot.cfg file is limited in size to 512kb [Pre NAS 4.x code]

--UFSLOGSIZE is limited to (2) billion transactions; If log reaches more than 80% full, NFS threads are suspended until log reduced to 60% full—if it cannot flush these log entries within 60 seconds, DataMover will panic!

**Note:** Limitation fixed in 4.0**Cause:** High I/O conditions and/or poorly designed File Systems!UFSLOG "*Race Condition*"—occurs when there is thread contention in File System between one process and the single-threaded process that cleans up the ufslog entries.**"ufs log full" Panics:** This can occur when DART cannot flush metadata to file system due to extreme load on Celerra—by default, there are 8 threads allocated to flushing ufslog entries. To increase the number of threads for ufslog activity, change the following (3) params from (8) threads to (16) threads as a start, and possibly as high as (32) if required—consult with TS2 or Eng. before changing these values on your own, as this condition may no longer apply with latest NAS versions:

ufs.nFlushIno 0x01144654 0x00000008 0x00000008 →change to 16

ufs.nFlushDir 0x01144658 0x00000008 0x00000008 →change to 16

ufs.nFlushCyl 0x0114465c 0x00000008 0x00000008 →change to 16

**UFSLOG FLUSH CASE EXAMPLE OF PARAM CHANGES:**

param ufs syncInterval=10000

param ufs nFlushCyl=32

param ufs nFlushDir=32

param ufs nFlushIno=32

**CELERRA UID/GID LIMITATIONS:** 4 billion UIDs/65,534 GIDs on a single file system**PASSWD File:** 64k NIS Unix limitation**16 Cylinders per Cylinder Group:** Master Control SuperBlock is at offset 16, with alternate SuperBlock at offset 32**Basic Celerra Vol's:** Stripe 4-8 volumes in multiples of 512 bytes, at a width of 32,768 bytes, evenly across the Disk Directors.**Running Jobs In Background or Foreground in SCO or LINUX:**

Step 1. Execute the script that you want to run

Step 2. Place Job in Background: # ctrl + Z &amp; # bg [places job in background]

Step 3. Return Job to Foreground: # jobs [identify which job you wish to return to foreground] &amp; # fg %1 [# kill %1 kills job]

**CRONTAB UTILITY:**

System Daemon that executes commands at specific times

# crontab -l -u nasadmin

# /var/spool/cron/crontabs

# /var/spool/cron/nasadmin

# cat nasadmin

# DO NOT EDIT THIS FILE - edit the master and reinstall.

# (/tmp/crontab.4696 installed on Fri Feb 27 11:32:04 2004)

# (Cron version -- \$Id: crontab.c,v 2.13 1994/01/17 03:20:37 vixie Exp \$)

15 \* \* \* \* /home/nasadmin/scripts/ckpt\_refresh.sh

**Note:** Above is an example of where Cronjob resides for 'nasadmin'**Setup in Six columns:** \*Asterisks are used as Wildcards

| #minute to run  | hour | day of mnth | month of year | day of the week    | Command path         |
|-----------------|------|-------------|---------------|--------------------|----------------------|
| 0-59            | 0-23 | 1-31        | 1-12          | 0-7 [0 & 7=Sunday] |                      |
| <b>Example:</b> | 2    | *           | *             | *                  | /home/jar/bin/backup |

**Explanation:** Backup job will run at 2:30 every day of week all year**Example of NAS BackUp Cron Job:**

#var/spool/cron/crontabs/nasadmin

1 \* \* \* \* /nas/sbin/nasdb\_backup /nas /home/nasadmin 1 &gt;/dev/null 2&gt;&amp;1

**Explanation:** NAS Database Backup will occur at 1 minute after each hour, every day, all year**OTHER EXAMPLES:**

0,15,30,45 \* \* \* \* [Job runs every 15 minutes every hour of every day]

1 1 \* \* 0 [Job runs weekly on Sundays]

```
* */2 * * * [Run every 2 hours]  
* 0-23/2 * * * [Run every other hour]  
1-30 * * * * [runs only every minute 1-30]
```

#### **Running a Cron Job on Celerra:** Create and Name your Script File

```
$crontab -l [lists out cron jobs]  
$crontab -l > thm [Creates cron script into a file called ‘thm’]  
$vi the file ‘thm’ and add your scripting “15**** /home/nasadmin/chkdc.rel>/dev/null 2>&1” [15 = 15 minutes after each hour]  
$crontab thm [starts your scripted cron job]
```

#### **OTHER LOCATIONS FOR CRON JOBS/SCRIPTS:**

```
# cat /nas/site/cron.d/nas_sys [abridged]  
1 * * * * root /nas/sbin/root_cron >/dev/null 2>&1  
0-59/5 * * * * nasadmin /nas/sbin/get_backend_status APM00033402672 /nas/log/backend_status.APM00033402672.xml >/dev/null  
2>&1  
1 * * * * nasadmin /nas/sbin/nasdb_backup /nas /home/nasadmin 1 &>/dev/null  
0 2 * * 3 root /nas/sbin/nas_config -alertview -get >/dev/null 2>&1  
# cat /nas/site/cron.d/nas_user [abridged]  
10 6,10,14 * * 1-5 nasadmin /nas/RC_SCRIPTS/create_ckpts.wrapper -q >/dev/null
```

**Note:** This file contains scheduled checkpoint scripts as configured from WebUI

#### **CLARIION SYSTEM SCRIPTS:**

```
/nas/site/cron.d  
# cat nas_sys  
0-59/5 * * * * nasadmin /nas/sbin/get_data_mover_status server_2 5080 /nas/log/data_mover_status.server_2.xml >/dev/null 2>&1  
1 0-23/2 * * * * nasadmin /nas/sbin/get_data_mover_status -resume server_2 5080 /nas/log/data_mover_resume.server_2.xml  
>/dev/null 2>&1  
0-59/5 * * * * nasadmin /nas/sbin/get_backend_status APM00033500569 /nas/log/backend_status.APM00033500569.xml >/dev/null  
2>&1 (script uses /nas/opt/Navisphere/bin/navicli to make backend queries)  
2 0-23/2 * * * * nasadmin /nas/sbin/get_backend_status -resume APM00033500569 /nas/log/backend_resume.APM00033500569.xml  
>/dev/null 2>&1  
Note: Output files are in /nas/log
```

#### **NS SYSTEMS BACKEND CRON JOBS:** /nbsnas/site/cron.d

```
# cat nas_sys  
0-59/5 * * * * nasadmin /nas/sbin/get_backend_status APM00040303779 /nas/log/backend_status.APM00040303779.xml  
>/dev/null 2>&1  
2 0-23/2 * * * * nasadmin /nas/sbin/get_backend_status -resume APM00040303779 /nas/log/backend_resume.APM00040303779.xml  
>/dev/null 2>&1
```

**Note:** The above cronjobs are scheduled to run every 5 minutes on NS systems, the results of which are diff'ed to see if any changes have occurred—if changes have occurred, the script then invokes a sync command to sync up the Celerra and backend databases—the XML files are posted to the /nbsnas/log directory. Errors and events, such as trespassed luns, are posted to the sys\_log.

#### **LISTING CRON JOBS ON LINUX CONTROL STATION:**

```
# crontab -l [Lists Cron Jobs scheduled by Root]  
$ crontab -l [Lists Cron Jobs scheduled by Nasadmin]  
# crontab -e [Brings up Edit window—add new Cron Job entry here or comment out existing Job]  
# crontab -u nasadmin -l [Output cronjobs for User nasadmin or any other user specified]  
# /nas/sbin/nas_cron -l  
# ls /var/spool/cron [Lists User cronjobs that have been configured]  
# /usr/bin/atq [Run this to see if any Crontab actions are queued to run]
```

nasadmin root

**EXAMPLE:**

```
0-59 * * * * /home/nasadmin/dmstats.cron server_2 >/dev/null 2>&1
```

**Note:** Above Job runs every minute, every day of the week, and calls script “dmstats.cron”

#### **UPDATING CRON AFTER VI EDIT CHANGE:**

```
# touch /etc/crontab
```

## **CONFIGURING CRONJOBS TO SURVIVE NAS UPGRADES, ETC.:**

**Note:** Use the crontab file that runs the NAS DB Backups on the Celerra. Vi edit nas\_sys and add Cron job.

### **/nas/site/cron.d**

# cat nas\_sys

**1 \* \* \* \* nasadmin /nas/sbin/nasdb\_backup /nas /home/nasadmin 1 &>/dev/null**

0 4 1 \* \* root /nas/sbin/log\_config -d -c >/dev/null 2>&1

1 \* \* \* \* root /nas/sbin/root\_cron >/dev/null 2>&1

## **HOW TO SET DATAMOVER TO DEFAULT TO BINARY MODE: 2.2+ & above**

/nas/site/slot\_param [or by individual datamover /nas/server/slot\_x/param]

**param ftpd forceBinXfer=1** [Add this to the parameter file & reboot datamover]

## **DISABLING FTP ON DATA MOVER:**

Step 1. \$cd /nas/server/slot\_x

Step 2. Comment out the “ftpd” line in the “netd” file [this will work until next code upgrade]

## **Default Umask for Users Accessing Celerra Using FTP Service: umask=740**

**COPYING FILES FROM FLOPPY:** #doscp a:\script.txt / [copies script.txt file to default /home/nasadmin directory]

#doscp a:driver.tar. or #doscp a:.\*

**COPYING FILES TO LINUX:**

#mcopy -t msdosfile /home/nasadmin/unixfile

**COPYING FILES TO FLOPPY:**

#mcopy -t a:c4.cmd

## **REMOTE ACCESS TO CELERRA CONTROL STATION:**

### **USING RSH REMOTE SHELL TO CONNECT TO LINUX 7.2 CONTROL STATION:**

**Note:** The “rsh” service is no longer enabled by default for this version of Linux

Step 1. Using vi, create “.rhosts” file in /home/nasadmin and enter the client addresses:

192.10.20.3 root

192.10.20.4 root

localhost nasadmin

Step 2. Chmod file: #chmod 600 /home/nasadmin/.rhosts [-rw----- 1 nasadmin nasadmin]

Step 3. Turn on the ‘rsh’ Service: #chkconfig rsh on

**Note:** See man pages for more details: #man ruserok

## **CONFIGURING RSH/RCP FOR LINUX 7.2 CONTROL STATION:**

1. Turn RSH Service on:

**# /sbin/chkconfig --list rsh**

rsh off

**#/sbin/chkconfig rsh on**

2. #vi /etc/security & add “rsh” to last line of file or use # echo rsh >> /etc/security

3. #vi /etc/pam.d/rsh & comment out the last “auth” line:

**#auth required /lib/security/pam\_rhosts\_auth.so**

4. Connect to remote Control Station and copy Pre-Upgrade Check Script to local directory:

**# rcp -p 10.241.168.50:/home/nasadmin/upgrd-ckv9-9g.bin .**

## **INSTALLING SENDMAIL FOR LINUX CONTROL STATION 6.2 & 7.2:**

**Note:** Sendmail is turned off by default. EMC does not encourage the use of this Linux RPM on the Celerra CS.

### **Linux 6.2:**

1. Copy RPMs to Control Station’s root directory

sendmail-8.11.6-1.62.2.i386.rpm

sendmail-cf-8.11.6-1.62.2.i386.rpm

### **Linux 7.2:**

1. Copy RPMs to Control Station’s root directory

sendmail-8.11.6-23.72.i386.rpm

sendmail-cf-8.11.6-23.72.i386.rpm

2. Run command:

**#rpm -Uvh sendmail\*.rpm**

### **VERIFYING SENDMAIL SERVICE:**

**# ps -ef |grep sendmail**

```
root 2754 1 0 Feb26 ? 00:00:00 sendmail: accepting connections
smmsp 2762 1 0 Feb26 ? 00:00:00 sendmail: Queue runner@01:00:00 for /var/spool/clientmqueue
root 28576 14533 0 12:30 pts/0 00:00:00 grep sendmail
```

**# /sbin/chkconfig --list sendmail | --add sendmail** (to add service)

```
sendmail 0:off 1:off 2:on 3:on 4:on 5:on 6:off
```

**# /sbin/service sendmail**

Usage: /etc/init.d/sendmail {start|stop|restart|condrestart|status}

**# /sbin/service sendmail status**

```
sendmail (pid 2762 2754) is running...
```

### **VERIFYING RPM PACKAGES ON NAS CD-ROM:**

1. Mount CD-ROM: # mount /dev/cdrom /mnt/cdrom

2. Run following command in the /mnt/cdrom/RedHat/RPMS directory: **rpm -qp -filesbypkg \*.rpm**

### **FINDING INFORMATION INSIDE AN RPM PACKAGE:**

**# rpm -ql emcnas --dbpath /var/sadm/pkg/emcnas/ | grep "version"**

### **LINUX 7.2 INSTALL:**

/dev/sde →NAS directory

/dev/sdf →VAR directory, both created using fdisk during install

### **CELERRA CONTROL STATION:**

**LINUX RED HAT 7.2:** Introduced with NAS 4.1.4.0, 2.4.7 kernel

--Services are controlled by /etc/rc.d/init.d scripts [init.d services]

--Services also controlled by xinetd [xinetd services]

--Also note that XWindows support has been removed [startx no longer opens an Xwindows session]

### **VERIFYING KERNEL VERSION:**

**# uname -r**

2.4.9-34.5405.EMC

**# cat /proc/version**

```
Linux version 2.6.9-42.5610.EMC (peacej@nasbuild9) (gcc version 3.4.4 20050721 (Red Hat 3.4.4-2)) #1 Tue Jan 8 18:36:46 EST 2008
```

**# nas\_version**

5.6.37-6

### **LINUX TRIVIA:**

# ch [Use tab + tab key to list out all commands starting with ch, etc.]

### **Changing Run Level at Startup to Level 5:**

# vi /etc/inittab [id : 5 : initdefault:]

# /usr/sbin/ntsysv #redhat-config-services [GUI to startup Linux Services] #redhat-config-xfree86 [Display & Video settings]

### **Determining if Partitions are EXT3 or EXT2:**

**# df -T \* #cat /proc/mounts #cat /etc/mtab**

**# df -T**

| Filesystem | Type  | 1k-blocks | Used    | Available | Use% | Mounted on |
|------------|-------|-----------|---------|-----------|------|------------|
| /dev/sda3  | ext3  | 1944904   | 1360268 | 485840    | 74%  | /          |
| /dev/sda1  | ext3  | 7746      | 3530    | 3816      | 49%  | /boot      |
| none       | tmpfs | 256812    | 0       | 256812    | 0%   | /dev/shm   |
| /dev/sdc1  | ext3  | 1818352   | 664896  | 1061084   | 39%  | /nas       |
| /dev/sdb1  | msdos | 136368    | 98348   | 38020     | 73%  | /nas/dos   |
| /dev/sdd1  | ext3  | 31079     | 17282   | 12193     | 59%  | /nas/var   |

| Filesystem | Type  | 1k-blocks | Used    | Available | Use% | Mounted on |
|------------|-------|-----------|---------|-----------|------|------------|
| /dev/sda3  | ext3  | 1944904   | 1360268 | 485840    | 74%  | /          |
| /dev/sda1  | ext3  | 7746      | 3530    | 3816      | 49%  | /boot      |
| none       | tmpfs | 256812    | 0       | 256812    | 0%   | /dev/shm   |
| /dev/sdc1  | ext3  | 1818352   | 664896  | 1061084   | 39%  | /nas       |
| /dev/sdb1  | msdos | 136368    | 98348   | 38020     | 73%  | /nas/dos   |
| /dev/sdd1  | ext3  | 31079     | 17282   | 12193     | 59%  | /nas/var   |

| Filesystem | Type  | 1k-blocks | Used    | Available | Use% | Mounted on |
|------------|-------|-----------|---------|-----------|------|------------|
| /dev/sda3  | ext3  | 1944904   | 1360268 | 485840    | 74%  | /          |
| /dev/sda1  | ext3  | 7746      | 3530    | 3816      | 49%  | /boot      |
| none       | tmpfs | 256812    | 0       | 256812    | 0%   | /dev/shm   |
| /dev/sdc1  | ext3  | 1818352   | 664896  | 1061084   | 39%  | /nas       |
| /dev/sdb1  | msdos | 136368    | 98348   | 38020     | 73%  | /nas/dos   |
| /dev/sdd1  | ext3  | 31079     | 17282   | 12193     | 59%  | /nas/var   |

| Filesystem | Type  | 1k-blocks | Used    | Available | Use% | Mounted on |
|------------|-------|-----------|---------|-----------|------|------------|
| /dev/sda3  | ext3  | 1944904   | 1360268 | 485840    | 74%  | /          |
| /dev/sda1  | ext3  | 7746      | 3530    | 3816      | 49%  | /boot      |
| none       | tmpfs | 256812    | 0       | 256812    | 0%   | /dev/shm   |
| /dev/sdc1  | ext3  | 1818352   | 664896  | 1061084   | 39%  | /nas       |
| /dev/sdb1  | msdos | 136368    | 98348   | 38020     | 73%  | /nas/dos   |
| /dev/sdd1  | ext3  | 31079     | 17282   | 12193     | 59%  | /nas/var   |

| Filesystem | Type  | 1k-blocks | Used    | Available | Use% | Mounted on |
|------------|-------|-----------|---------|-----------|------|------------|
| /dev/sda3  | ext3  | 1944904   | 1360268 | 485840    | 74%  | /          |
| /dev/sda1  | ext3  | 7746      | 3530    | 3816      | 49%  | /boot      |
| none       | tmpfs | 256812    | 0       | 256812    | 0%   | /dev/shm   |
| /dev/sdc1  | ext3  | 1818352   | 664896  | 1061084   | 39%  | /nas       |
| /dev/sdb1  | msdos | 136368    | 98348   | 38020     | 73%  | /nas/dos   |
| /dev/sdd1  | ext3  | 31079     | 17282   | 12193     | 59%  | /nas/var   |

| Filesystem | Type  | 1k-blocks | Used    | Available | Use% | Mounted on |
|------------|-------|-----------|---------|-----------|------|------------|
| /dev/sda3  | ext3  | 1944904   | 1360268 | 485840    | 74%  | /          |
| /dev/sda1  | ext3  | 7746      | 3530    | 3816      | 49%  | /boot      |
| none       | tmpfs | 256812    | 0       | 256812    | 0%   | /dev/shm   |
| /dev/sdc1  | ext3  | 1818352   | 664896  | 1061084   | 39%  | /nas       |
| /dev/sdb1  | msdos | 136368    | 98348   | 38020     | 73%  | /nas/dos   |
| /dev/sdd1  | ext3  | 31079     | 17282   | 12193     | 59%  | /nas/var   |

| Filesystem | Type  | 1k-blocks | Used    | Available | Use% | Mounted on |
|------------|-------|-----------|---------|-----------|------|------------|
| /dev/sda3  | ext3  | 1944904   | 1360268 | 485840    | 74%  | /          |
| /dev/sda1  | ext3  | 7746      | 3530    | 3816      | 49%  | /boot      |
| none       | tmpfs | 256812    | 0       | 256812    | 0%   | /dev/shm   |
| /dev/sdc1  | ext3  | 1818352   | 664896  | 1061084   | 39%  | /nas       |
| /dev/sdb1  | msdos | 136368    | 98348   | 38020     | 73%  | /nas/dos   |
| /dev/sdd1  | ext3  | 31079     | 17282   | 12193     | 59%  | /nas/var   |

| Filesystem | Type  | 1k-blocks | Used    | Available | Use% | Mounted on |
|------------|-------|-----------|---------|-----------|------|------------|
| /dev/sda3  | ext3  | 1944904   | 1360268 | 485840    | 74%  | /          |
| /dev/sda1  | ext3  | 7746      | 3530    | 3816      | 49%  | /boot      |
| none       | tmpfs | 256812    | 0       | 256812    | 0%   | /dev/shm   |
| /dev/sdc1  | ext3  | 1818352   | 664896  | 1061084   | 39%  | /nas       |
| /dev/sdb1  | msdos | 136368    | 98348   | 38020     | 73%  | /nas/dos   |
| /dev/sdd1  | ext3  | 31079     | 17282   | 12193     | 59%  | /nas/var   |

| Filesystem | Type  | 1k-blocks | Used    | Available | Use% | Mounted on |
|------------|-------|-----------|---------|-----------|------|------------|
| /dev/sda3  | ext3  | 1944904   | 1360268 | 485840    | 74%  | /          |
| /dev/sda1  | ext3  | 7746      | 3530    | 3816      | 49%  | /boot      |
| none       | tmpfs | 256812    | 0       | 256812    | 0%   | /dev/shm   |
| /dev/sdc1  | ext3  | 1818352   | 664896  | 1061084   | 39%  | /nas       |
| /dev/sdb1  | msdos | 136368    | 98348   | 38020     | 73%  | /nas/dos   |
| /dev/sdd1  | ext3  | 31079     | 17282   | 12193     | 59%  | /nas/var   |

| Filesystem | Type  | 1k-blocks | Used    | Available | Use% | Mounted on |
|------------|-------|-----------|---------|-----------|------|------------|
| /dev/sda3  | ext3  | 1944904   | 1360268 | 485840    | 74%  | /          |
| /dev/sda1  | ext3  | 7746      | 3530    | 3816      | 49%  | /boot      |
| none       | tmpfs | 256812    | 0       | 256812    | 0%   | /dev/shm   |
| /dev/sdc1  | ext3  | 1818352   | 664896  | 1061084   | 39%  | /nas       |
| /dev/sdb1  | msdos | 136368    | 98348   | 38020     | 73%  | /nas/dos   |
| /dev/sdd1  | ext3  | 31079     | 17282   | 12193     | 59%  | /nas/var   |

| Filesystem | Type  | 1k-blocks | Used    | Available | Use% | Mounted on |
|------------|-------|-----------|---------|-----------|------|------------|
| /dev/sda3  | ext3  | 1944904   | 1360268 | 485840    | 74%  | /          |
| /dev/sda1  | ext3  | 7746      | 3530    | 3816      | 49%  | /boot      |
| none       | tmpfs | 256812    | 0       | 256812    | 0%   | /dev/shm   |
| /dev/sdc1  | ext3  | 1818352   | 664896  | 1061084   | 39%  | /nas       |
| /dev/sdb1  | msdos | 136368    | 98348   | 38020     | 73%  | /nas/dos   |
| /dev/sdd1  | ext3  | 31079     | 17282   | 12193     | 59%  | /nas/var   |

| Filesystem | Type  | 1k-blocks | Used    | Available | Use% | Mounted on |
|------------|-------|-----------|---------|-----------|------|------------|
| /dev/sda3  | ext3  | 1944904   | 1360268 | 485840    | 74%  | /          |
| /dev/sda1  | ext3  | 7746      | 3530    | 3816      | 49%  | /boot      |
| none       | tmpfs | 256812    | 0       | 256812    | 0%   | /dev/shm   |
| /dev/sdc1  | ext3  | 1818352   | 664896  | 1061084   | 39%  | /nas       |
| /dev/sdb1  | msdos | 136368    | 98348   | 38020     | 73%  | /nas/dos   |
| /dev/sdd1  | ext3  | 31079     | 17282   | 12193     | 59%  | /nas/var   |

| Filesystem | Type  | 1k-blocks | Used    | Available | Use% | Mounted on |
|------------|-------|-----------|---------|-----------|------|------------|
| /dev/sda3  | ext3  | 1944904   | 1360268 | 485840    | 74%  | /          |
| /dev/sda1  | ext3  | 7746      | 3530    | 3816      | 49%  | /boot      |
| none       | tmpfs | 256812    | 0       | 256812    | 0%   | /dev/shm   |
| /dev/sdc1  | ext3  | 1818352   | 664896  | 1061084   | 39%  | /nas       |
| /dev/sdb1  | msdos | 136368    | 98348   | 38020     | 73%  | /nas/dos   |
| /dev/sdd1  | ext3  | 31079     | 17282   | 12193     | 59%  | /nas/var   |

| Filesystem | Type  | 1k-blocks | Used    | Available | Use% | Mounted on |
|------------|-------|-----------|---------|-----------|------|------------|
| /dev/sda3  | ext3  | 1944904   | 1360268 | 485840    | 74%  | /          |
| /dev/sda1  | ext3  | 7746      | 3530    | 3816      | 49%  | /boot      |
| none       | tmpfs | 256812    | 0       | 256812    | 0%   | /dev/shm   |
| /dev/sdc1  | ext3  | 1818352   | 664896  | 1061084   | 39%  | /nas       |
| /dev/sdb1  | msdos | 136368    | 98348   | 38020     | 73%  | /nas/dos   |
| /dev/sdd1  | ext3  | 31079     | 17282   | 12193     | 59%  | /nas/var   |

| Filesystem | Type  | 1k-blocks | Used    | Available | Use% | Mounted on |
|------------|-------|-----------|---------|-----------|------|------------|
| /dev/sda3  | ext3  | 1944904   | 1360268 | 485840    | 74%  | /          |
| /dev/sda1  | ext3  | 7746      | 3530    | 3816      | 49%  | /boot      |
| none       | tmpfs | 256812    | 0       | 256812    | 0%   | /dev/shm   |
| /dev/sdc1  | ext3  | 1818352   | 664896  | 1061084   | 39%  | /nas       |
| /dev/sdb1  | msdos | 136368    | 98348   | 38020     | 73%  | /nas/dos   |
| /dev/sdd1  | ext3  | 31079     | 17282   | 12193     | 59%  | /nas/var   |

| Filesystem | Type  | 1k-blocks | Used    | Available | Use% | Mounted on |
|------------|-------|-----------|---------|-----------|------|------------|
| /dev/sda3  | ext3  | 1944904   | 1360268 | 485840    | 74%  | /          |
| /dev/sda1  | ext3  | 7746      | 3530    | 3816      | 49%  | /boot      |
| none       | tmpfs | 256812    | 0       | 256812    | 0%   | /dev/shm   |
| /dev/sdc1  | ext3  | 1818352   | 664896  | 1061084   | 39%  | /nas       |
| /dev/sdb1  | msdos | 136368    | 98348   | 38020     | 73%  | /nas/dos   |
| /dev/sdd1  | ext3  | 31079     | 17282   | 12193     | 59%  | /nas/var   |

| Filesystem | Type  | 1k-blocks | Used    | Available | Use% | Mounted on |
|------------|-------|-----------|---------|-----------|------|------------|
| /dev/sda3  | ext3  | 1944904   | 1360268 | 485840    | 74%  | /          |
| /dev/sda1  | ext3  | 7746      | 3530    | 3816      | 49%  | /boot      |
| none       | tmpfs | 256812    | 0       | 256812    | 0%   | /dev/shm   |
| /dev/sdc1  | ext3  | 1818352   | 664896  | 1061084   | 39%  | /nas       |
| /dev/sdb1  | msdos | 136368    | 98348   | 38020     | 73%  | /nas/dos   |
| /dev/sdd1  | ext3  | 31079     | 17282   | 12193     | 59%  | /nas/var   |

| Filesystem | Type  | 1k-blocks | Used    | Available | Use% | Mounted on |
|------------|-------|-----------|---------|-----------|------|------------|
| /dev/sda3  | ext3  | 1944904   | 1360268 | 485840    | 74%  | /          |
| /dev/sda1  | ext3  | 7746      | 3530    | 3816      | 49%  | /boot      |
| none       | tmpfs | 256812    | 0       | 256812    | 0%   | /dev/shm   |
| /dev/sdc1  | ext3  | 1818352   | 664896  | 1061084   | 39%  | /nas       |
| /dev/sdb1  | msdos | 136368    | 98348   | 38020     | 73%  | /nas/dos   |
| /dev/sdd1  | ext3  | 31079     | 17282   | 12193     | 59%  | /nas/var   |

| Filesystem | Type  | 1k-blocks | Used    | Available | Use% | Mounted on |
|------------|-------|-----------|---------|-----------|------|------------|
| /dev/sda3  | ext3  | 1944904   | 1360268 | 485840    | 74%  | /          |
| /dev/sda1  | ext3  | 7746      | 3530    | 3816      | 49%  | /boot      |
| none       | tmpfs | 256812    | 0       | 256812    | 0%   | /dev/shm   |
| /dev/sdc1  | ext3  | 1818352   | 664896  | 1061084   | 39%  | /nas       |
| /dev/sdb1  | msdos | 136368    | 98348   | 38020     | 73%  | /nas/dos   |
| /dev/sdd1  | ext3  | 31079     | 17282   | 12193     | 59%  | /nas/var   |

| Filesystem | Type  | 1k-blocks | Used    | Available | Use% | Mounted on |
|------------|-------|-----------|---------|-----------|------|------------|
| /dev/sda3  | ext3  | 1944904   | 1360268 | 485840    | 74%  | /          |
| /dev/sda1  | ext3  | 7746      | 3530    | 3816      | 49%  | /boot      |
| none       | tmpfs |           |         |           |      |            |

### **Drivers:**

Kernel drivers loading during bootup

Modules are loaded dynamically as needed [see /etc/modules.conf]

#sbin/lspci [List of controllers on system] #ls /proc/bus/pci #cat /proc/bus/pci/devices

### **Red Hat Installer:**

Bootable CD-ROM contains boot.iso file | 4 Boot Floppies contain bootdisk.img, drvnet.img, drvblock.img & pcmciaadd.img

### **Copying 1<sup>st</sup> Stage Installer from CD-ROM to Floppy:**

#cat bootdisk.img > /dev/fd0 #dd if=bootdisk.img of = /dev/fd0

### **EXTRACTING FILES FROM ISO IMAGE:**

```
#mkdir /iso  
#mount -o loop boot.iso /iso  
#ls /iso
```

### **LINUX NAMING CONVENTION FOR PARTITIONS AND MULTIPLE PHYSICAL DISKS:**

IDE: 0 /dev/hda /dev/hda1 /dev/hda2, etc.

IDE: 1 /dev/hdb /dev/hdb1 /dev/hdb2, etc.

SCSI: 0 /dev/sda /dev/sda1 /dev/sda2, etc.

SCSI: 1 /dev/sdb /dev/sdb1 /dev/sdb2, etc.

### **USING DISK DRUID DURING INSTALL:** Automatically configure or use manual Disk Druid utility

```
/boot 100M  
/ 256M  
/usr 1256M  
/var 400M  
swap 512M
```

### **GRUB LOADER [Grand Unified Bootloader]:**

/boot/grub/grub.conf

### **RESTORING MBR IF DAMAGED:**

**#/sbin/grub-install /dev/hda**

### **BOOTING TO SINGLE USER MODE ON LINUX RED HAT 7.3:**

1. At Grub screen where “linux” is the only word displayed, type “e” for edit, then select the middle line out of the three listed, and type “e” for edit:

**kernel /vmlinuz ro root=/dev/hda3 console=ttyS1,19200**

2. Add a single space and then the word “single” without quotes after 19200

3. After edit has been made, press ‘enter’ key, then “b” for boot

4. System will come up to Single User prompt:

Sh-2.05#

**Note 1:** Older versions of Linux may require following edit at end of kernel line: “s” →root=LABEL=/ s

**Note 2:** The above edit to get to Single User mode is only temporary and after reboot comes up to normal Run Level 3

5. Verifying runlevel on Linux: # /sbin/runlevel

### **ALTERNATE SYNTAX EDIT:**

**kernel /vmlinuz ro root=/dev/hda3 init 1 console=ttyS1,19200**

### **BOOTING SINGLE USER MODE ON CELERRA NS600 TO CHANGE ROOT PASSWD:**

At Lilo Prompt type: **linux single**

If root passwd needs changing, cd to bin directory and type: #passwd root

### **CONVERTING FILE SYSTEM FROM EXT2 TO EXT3 JOURNALING:**

1. Edit /etc/fstab to ‘ext3’

**#tune2fs -j /dev/hda6**

3. Reboot

4. Verify by running #df -T \*

### **FORMATTING PARTITIONS:**

**#mke2fs -T news /dev/hda8**

### **DISPLAYING AND SETTING THE IMMUTABLE ATTRIBUTE ON FILES:**

```
#lsattr file  
#chattr +i passwd [Setting this attribute allows no one to remove file] #chattr -i passwd
```

### **CREATING A SWAP FILE:**

#fuser -k /data [Removes any processes connected to /data]

#umount /data

#dd if=/dev/zero of=/tmp/nuswap bs=512 count=245678 [12MB file]

```
#mkswap /tmp/nuswap  
#swapon /tmp/nuswap  
#swapon -s  
#swapoff /tmp/nuswap  
#mkfs -j /dev/hda6  
#e2label /dev/hda6 /data  
#mount /dev/hda6 /data
```

### **CREATING USER AND GROUP ACCOUNTS:** (not allowed beginning with NAS 5.6)

```
#useradd user1; echo 'user1:password' |chpasswd [Creates and sets default password of 'password' for user1]  
#usermod -G group1, group2, group3 user1 [Adds User1 to the indicated Groups]  
#/nas/sbin/check_user [Enter passwd for user as test, returns 'Authenticated' if successful]  
#userdel -r accntname [deletes accnt from passwd and shadow files]  
#chage -l user1 [Lists out account properties for User1—use #chage user1 to change]  
#groupadd sales -g 104 [Adds new Group called Sales with GID of 104]  
#groupmod -g 105 sales [Changes GID of group 'Sales' to 105]
```

### **# chage tom**

```
Minimum Password Age [0]:  
Maximum Password Age [99999]:  
Last Password Change (YYYY-MM-DD) [2004-02-11]:  
Password Expiration Warning [7]:  
Password Inactive [-1]:  
Account Expiration Date (YYYY-MM-DD) [1969-12-31]:
```

### **CREATING USER PRIVATE GROUPS:**

1. #mkdir /data/{sales,mis}
2. #chgrp sales /data/sales;chgrp mis /data/mis
3. #chmod g+s /data/\* [Sets the setgid bit without affecting current permissions]
4. #chmod 2770 /data/\* [Sets setgid bit and gives RWX permissions to Owner and Group members and access denied to Other]

**Note:** Turns on Group Sticky bit for /data/\* directories. Files or subdirectories created will belong to user but assigned Group ownership of directory.

### **CREATING FILES FOR TEST PURPOSES:**

**#for i in \$(seq 1 100); do echo -n “file\${i}“; touch file\${i} 2>&1; done | less** [Loop script creates 100 files for testing User quotas]

**#dd if=/dev/zero of=bigfile bs=1M count=3** [Script to create large files quickly]

### **PRINTING SERVICE:**

```
#lpr -P printer [BSD] #lp -d printer [System V]  
#/sbin/service cups status [Common Unified Print Service]  
#printconf-gui #/etc/cups/cupsd.conf | printers.conf  
#lpadmin -p printer1 -E -v Dev [Adding printer]  
#lpadmin -d printer1 [Setting default printer]  
#lpadmin -x printer1 [Deleting printer]  
#redhat-config-printer  
http://localhost:631/
```

### **SETTING UP SMB SHARE:**

```
#smbmount //192.168.1.200/joker /joker -o mark=smbuser [Sharename Joker defined in /etc/samba/smb.conf]  
#smbclient //192.168.1.200/joker -u mark  
Password: redhat  
Smb:>cd get put [commands]  
#chkconfig --level 345 smb on  
#touch /etc/samba/smbpasswd #smbpasswd -a mark [User must exist in /etc/passwd first]  
#testparm
```

### **VERIFYING LIST OF AVAILABLE SERVICES WITH RED HAT 7.2:**

```
#/sbin/chkconfig --list [Run Level 3 = CLI Run Level 5 = GUI]  
#/sbin/service --status-all [Useful for Status of services and Mounts on the Control Station]
```

### **VERIFYING A SPECIFIC SERVICE WITH LINUX:**

**\$ /sbin/chkconfig --list sendmail**

```
sendmail 0:off 1:off 2:off 3:off 4:off 5:off 6:off
```

**\$ /sbin/service sendmail status**

sendmail is stopped

**#/sbin/chkconfig sendmail on**

**#/sbin/chkconfig sendmail restart**

Shutting down sendmail: [ OK ]

Starting sendmail: [ OK ]

#### **DEBUGGING SENDMAIL:**

**# sendmail -d0 </dev/null**

#### **HOW TO ENABLE "init.d" SERVICES ON RED HAT 7.2 CONTROL STATION:**

1. Turn the Service On: #/sbin/chkconfig atd on [Scheduler service for Jobs; /var/spool/st]

2. Start the Service: #/sbin/service atd start

#### **HOW TO ENABLE "xinetd" SERVICES ON RED HAT 7.2:**

1. Turn Telnet On: #/sbin/chkconfig telnet on

2. Start Service: #/sbin/service xinetd start

#### **HOW TO ENABLE SNMP SERVICE ON LINUX 4.0.18 CONTROL STATION:**

1. Turn SNMPD On: #/sbin/chkconfig snmpd on

2. Start Service: #/sbin/service snmpd start

#### **LIST OF LINUX SERVICES:**

**# /sbin/chkconfig --list**

#### **LINUX CONTROL STATION VOLUMES & BOOTUP PROCESS Red Hat 6.2:**

| Vol/Part. | Size     | Offset | Device/Mount | Mountpoint         |
|-----------|----------|--------|--------------|--------------------|
| TLUN 00 = | root_dos | 134MB  | 0            | /dev/sda1 /nas/dos |
| TLUN 02 = | sdc1     | 7MB    | 0            | /boot [CS0]        |
|           | sdc2     | 133MB  | 7mb          | /swap              |
|           | sdc3     | 1929MB | 140mb        | /                  |

TLUN 03 = same as above for CS1

**BootUp Process:** BIOS → Boot [tlun 00] → autoexec.bat [/nas/dos] → exec t2slot → cd ./slot\_0 (1) → boot exec boot.bat  
→ loadlin [boot dev=/dev/sdc1 root dev=/dev/sdc3] → Start Linux [/etc/rc.d/init.d from /etc/inittab]

#### **VERIFYING NAS VERSION USED IN BOOTUP ON DMs:**

**\$ cat /nas/dos[slot\_\*/boot.bat |grep gload**

gload \bin\nas.exe boot.cfg

gload \bin\518804.exe boot.cfg

#### **NAS DOS PARTITION: /dev/sda1 Lun 00**

Prior to NAS 2.2.25.6, the NAS DOS partition was mapped to /dev/sdb1

All subsequent NAS versions map NAS DOS to /dev/sda1

#### **CONDUCTING BYTE-for-BYTE READ ON LINUX PARTITION:**

**# dd if=/nas/rdf/5001 of=/dev/null** → might be useful to validate data integrity

#### **CELLERRA BOOT UP PROCESS EXPLAINED:**

→ Control Stations and Data Movers all boot initially into DOS using tlun t0l0.

→ After DOS loads, the /nas/dos/autoexec.bat file uses t2slot.exe -f to determine which slot is booting.

→ Autoexec.bat then runs the appropriate boot.bat file for the slot.

→ Data Mover slots load DART based on boot.bat:

\$ cat /nas/dos[slot\_2/boot.bat

gload \bin\nas15400.exe boot.cfg

→ Control Station slots load LINUX based on boot.bat:

\$ cat boot.bat

\bin\loadlin \bin\vmlinuz root=/dev/sdc3 boot=/dev/sdc1 initrd=/bin/initrd.img panic=1 max\_scsi\_luns=16

→ Both control stations will use same initrd.img & vmlinuz files, though use different root & boot partitions

/boot

lrwxrwxrwx 1 root root 25 Jun 4 10:21 initrd.img -> initrd-2.4.9-34.8.EMC.img

lrwxrwxrwx 1 root root 22 Jun 4 10:21 vmlinuz -> vmlinuz-2.4.9-34.8.EMC

-rw-r--r-- 1 root root 641942 Sep 19 2002 vmlinuz-2.4.9-34.5.EMC

### **IMPORTANT BOOT FILES IN /nas/dos/bin:**

```
gload.exe: MS-DOS executable (EXE)
initrd.img: gzip compressed data, deflated, last modified: Wed Jun 4 10:21:47 2003, max compression, os: Unix
loadlin.exe: MS-DOS executable (EXE)
nas.exe: 386 demand paged pure executable not stripped
t2slot.exe: MS-DOS executable (EXE)
vmlinuz: x86 boot sector
```

### **SAMPLE LILO.CONF FILE ON 5.1.15.3 CELERRA :**

**\$ cat /etc/lilo.conf**

```
boot=/dev/sda
disk=/dev/sda
bios=0x80
map=/boot/map
install=/boot/boot.b
prompt
timeout=50
default=linux
image=/boot/vmlinuz
    label=linux
    root=/dev/sda3
    initrd=/boot/initrd.img
    read-only
```

### **UPDATING THE LILO.CONF FILE:**

**#/sbin/lilo**

### **CHANGES WITH LINUX 7.2:**

Introduces the “ext3” Journaling FileSystem for Linux, which provides for better & more reliable recoveries after unclean shutdowns

#### **Ctrl + Alt + Delete:**

Performs clean shutdown & reboot equivalent to doing #/sbin/shutdown -t3 -r now

#### **BootUp Process:**

/dev/initrd [RO Block Device] takes over from vmlinuz kernel image and begins startup of Linux Kernel. If missing, Linux cannot boot and gives kernel panic error message. If present, /sbin/init will then execute and control rest of bootup process.

Printing [lp commands] functions no longer configured by default on Linux Control Station

### **LINUX BOOT.BAT FOR NAS 5.1 EAGLE & 5.5 NS SERIES:**

**# cat /nas/dos/slot\_0/boot.bat**

**\bin\loadlin \bin\vmlinuz root=/dev/sdc3 boot=/dev/sdc1 initrd=/bin/initrd.img panic=1 max\_scsi\_luns=16**  
(NAS 5.1)

**\bin\loadlin \bin\vmlinuz root=/dev/hda3 boot=/dev/hda1 initrd=/bin/initrd.img panic=1**  
**max\_scsi\_luns=16** (NAS 5.5)

### **LINUX GRUB BOOT LOADER:**

**Note:** Used by NAS 5.4? and higher. Replaces the older lilo.conf boot loader.

**# cat /boot/grub/grub.conf**

```
default=0
timeout=10
serial .unit=1 .speed=19200
terminal .timeout=10 serial console
title linux
    root (hd0,0)
    kernel /vmlinuz ro root=/dev/hda3 console=ttyS1,19200
    initrd /initrd.img
```

### **TMPWATCH UTILITY:**

Linux Control Station runs ‘tmpwatch’ utility each day at 4:02 a.m. that removes content of /tmp directory > than 10 days old. Run from /etc/cron.daily/tmpwatch

**Caution:** Do not mount Celerra Server to this directory on the Control Station!

## **LINUX CS0 INFORMATION ON CD-ROM DRIVE:**

1. Grep the “dmesg” file for Toshiba: **#cat /var/log/dmesg | grep -i toshiba**  
hda: TOSHIBA CD-ROM XM-6602B, ATAPI CD/DVD-ROM drive
2. **#/sbin/ide\_info /dev/hda**  
MODEL="TOSHIBA CD-ROM XM-6602B"  
FW\_REV="1017"  
SERIAL\_NO=""
3. #more /etc/fstab  
/dev/cdrom /mnt/cdrom iso9660 noauto,owner,kudzu,ro 0 0

**Note:** If no reference found in “dmesg” or “ide\_info”, the drive is not being seen by Linux and may be defective

## **NS40 CONTROL STATION CD-ROM:**

**# dmesg |grep CD**

hdc: DW-224E-C, ATAPI CD/DVD-ROM drive

**Note:** This is a TEAC CD/DVD Combo drive, as seen by the Model number, “DW-224E-C”, 24X Recordable, Readable, Rewriteable, 8X DVD-ROM

## **VIEWING CELERRA CONTROL VOLUMES:**

**Example: missing /nas/var directory**

**#/sbin/fdisk -l** [Partial output listed below]

```
Disk /dev/sda: 255 heads, 63 sectors, 529 cylinders
Units = cylinders of 16065 * 512 bytes
Device Boot Start End Blocks Id System
/dev/sda1 * 1 17 136521 6 FAT16
Device Boot Start End Blocks Id System
/dev/sdc1 * 1 4 32098+ 83 Linux
/dev/sdc2 5 69 522112+ 82 Linux swap
/dev/sdc3 70 264 1566337+ 83 Linux
```

**\$ more fstab**

```
LABEL=/ ext3 defaults 1 1
LABEL=boot /boot ext3 defaults 1 2
none /dev/pts devpts gid=5,mode=620 0 0
none /proc proc defaults 0 0
none /dev/shm tmpfs defaults 0 0
/dev/sdc2 swap swap defaults 0 0
/dev/cdrom /mnt/cdrom iso9660 noauto,owner,kudzu,ro 0 0
/dev/fd0 /mnt/floppy auto noauto,owner,kudzu 0 0
/dev/sde1 /nas ext3 noauto,rw,sync 0 0
/dev/sda1 /nas/dos msdos umask=002,noauto,rw,sync,gid=2010 0
/dev/sdf1 /nas/var ext3 noauto,rw 0 0 [directory missing on customer box]
```

**\$ more mtab**

```
/dev/sdc3 / ext3 rw 0 0
none /proc proc rw 0 0
/dev/sdc1 /boot ext3 rw 0 0
none /dev/pts devpts rw,gid=5,mode=620 0 0
none /dev/shm tmpfs rw 0 0
/dev/sde1 /nas ext3 rw,sync 0 0
/dev/sdf1 /nas/var ext3 rw 0 0
/dev/sda1 /nas/dos msdos rw,sync,umask=002,gid=201 0 0
automount(pid14500) /nasmcd/quota autofs rw,fd=5,pgrp=14500,minproto=2,maxproto=3 0 0
automount(pid14474) /nasmcd/rootfs autofs rw,fd=5,pgrp=14474,minproto=2,maxproto=3 0 0
betacs0:(pid14807) /net nfs intr,rw,port=1023,timeo=8,retrans=110,indirect,map=/
etc/amd.net,dev=0000000a 0 0
```

## VIEWING SERVICES ON CS :

# /sbin/service --status-all

atd (pid 1186) is running...

Active Mount Points:

-----  
/usr/sbin/automount --timeout 1 /nasmcd/rootfs file /etc/auto.nas ro,,soft,intr,  
nosuid,noac

/usr/sbin/automount --timeout 1 /nasmcd/quota file /etc/auto.nas rw,,soft,intr,n  
osuid,noac

/usr/sbin/automount --timeout 1 /net/500 file /etc/auto.500 ,ro

crond (pid 1107) is running...

gpm (pid 1089) is running...

-----output abridged-----

## SERVER VOLUME COMMANDS:

\$.[server\\_config server\\_2 “volume info 239” “volume stripe 234 567”](#)

### How to Debug Services that Fail during Bootup on Linux or Solaris:

\$[var/log/dmesg](#) [Reboot and then look at the log created during bootup to inspect what services come up, etc]

#[dmesg |less](#)

#[tail -f /var/log/{dmesg,messages} -n0 \[nolines\]](#)

## LINUX OPERATION:

### LOCKING-DOWN LINUX CONTROL STATION SECURITY:

Disable the following Services by commenting out the desired line in /etc/services & rebooting CS:

|           |        |     |        |            |                |            |                                       |
|-----------|--------|-----|--------|------------|----------------|------------|---------------------------------------|
| #ftp      | stream | tcp | nowait | root       | /usr/sbin/tcpd | in.ftpd    | -l -a                                 |
| telnet    | stream | tcp | nowait | root       | /usr/sbin/tcpd | in.telnetd | <b>(Warning: Will disable telnet)</b> |
| #shell    | stream | tcp | nowait | root       | /usr/sbin/tcpd | in.rshd    |                                       |
| login     | stream | tcp | nowait | root       | /usr/sbin/tcpd | in.rlogind |                                       |
| talk      | dgram  | udp | wait   | nobody.tty | /usr/sbin/tcpd | in.talkd   |                                       |
| ntalk     | dgram  | udp | wait   | nobody.tty | /usr/sbin/tcpd | in.ntalkd  |                                       |
| #finger   | stream | tcp | nowait | nobody     | /usr/sbin/tcpd | in.fingerd |                                       |
| linuxconf | stream | tcp | wait   | root       | /bin/linuxconf | linuxconf  | --http                                |

Disable the following by changing capital “S” to small “s” in /etc/rc.d/rc3.d & rebooting CS:

|              |   |              |
|--------------|---|--------------|
| S80sendmail  | → | s80sendmail  |
| S99linuxconf | → | s99linuxconf |
| S14nfslock   | → | s14nfslock   |

## LINUX NETWORK INTERFACES:

### Linux Interface Statistics:

#[netstat -i # netstat -taupe](#)

Linux Interfaces:      #[sbin/ifconfig -a](#)

## RESTRICTING NUMBER OF CONCURRENT LOGINS TO LINUX CS:

Modify following file to set maximum logins: 0 = 1 user, 1 = 2 users, etc.

/etc/security/limits.conf

naslock hard maxlogins 0

## RED HAT INTERFACE TROUBLESHOOTING:

### LINUX CONTROL STATION FILES RELATED TO EXTERNAL NIC:

/etc/sysconfig/network-scripts/ifcfg-eth2 [Network Device param file]

/etc/inittab [Verify that “id :3 :initdefault” set to bring system up to Run Level 3]

#[cat /etc/inittab |grep -i initdefault](#)

# 0 - halt (Do NOT set initdefault to this)

# 6 - reboot (Do NOT set initdefault to this)

id:3:initdefault:

/etc/rc3.d/S10network [Verify that this script is set to startup NIC devices—reads ‘ifcfg-eth2’ file to set params]

lrwxrwxrwx 1 root root 17 Feb 23 20:23 S10network -> ./init.d/network

/etc/modules.conf [Verify spd/duplexing on bootup ; options de4x5 io=0→100; options de4x5 args='eth2:fdx' →full duplex]

/var/log/messages [Grep ‘eth2:media’ message\* to see how interfaces actually came up ; “100Mb/s full duplex”]

→Use ping and /sbin/route to verify connectivity and gateway

#### **INTERNAL NICs:**

Similar to above: ifcfg-eth0/ifcfg-eth1 files—use ping and /sbin/route and /nas/sbin/getboxmask –r to troubleshoot

### **USING NETCONFIG UTILITY:**

# netconfig --device-eth2

# /sbin/service network status | stop | start

### **VERIFYING WHAT MODULES ARE LOADED ON LINUX:**

# /sbin/lsmod

| Module        | Size   | Used by                   | Not tainted |
|---------------|--------|---------------------------|-------------|
| nfs           | 71680  | 0 (autoclean)             |             |
| lockd         | 48800  | 0 (autoclean) [nfs]       |             |
| sunrpc        | 60400  | 0 (autoclean) [nfs lockd] |             |
| nls_iso8859-1 | 2816   | 1 (autoclean)             |             |
| nls_cp437     | 4352   | 1 (autoclean)             |             |
| msdos         | 4892   | 1 (autoclean)             |             |
| fat           | 30424  | 0 (autoclean) [msdos]     |             |
| sg            | 24260  | 0 (autoclean)             |             |
| autofs        | 9060   | 2 (autoclean)             |             |
| tulip         | 35008  | 1 →External Interface     |             |
| 3c509         | 6976   | 2 →Internal Interfaces    |             |
| ext3          | 57888  | 4                         |             |
| jbd           | 35972  | 4 [ext3]                  |             |
| aic7xxx       | 112288 | 7                         |             |
| sd_mod        | 10460  | 7                         |             |
| scsi_mod      | 54284  | 3 [sg aic7xxx sd_mod]     |             |

**Note:** In this case, shows NIC drivers that are loaded

### **SETTING LINUX 7.2 CS EXTERNAL NIC (NAS 4.1-5.3) TO FULL DUPLEX:**

Step 1. Verify the current driver installed by examining the /etc/modules.conf file, looking for the entry for “eth2”

\$ more modules.conf

```
alias parport_lowlevel parport_pc
options scsi_mod max_scsi_luns=16
alias scsi_hostadapter aic7xxx
alias eth0 tulip
alias eth0 3c509
alias eth1 3c509
alias eth2 de4x5      [new driver that replaces the tulip driver]
options de4x5 io=0
```

**Note:** Check /var/log/messages file for following entry: “betacs0 kernel:eth2:media is 100Mb/s”

Step 2. Replace the “options de4x5 io=0” line with the following:

**options de4x5 args='eth2:fdx' de4x5\_debug=1 [Full Duplex setting, debug setting]**

Step 3. Reboot Control Station and check /var/log/messages file for following entry:

“betacs0 kernel:eth2:media is 100Mb/s full duplex”

**Note:** This now indicates that the Control Station NIC is operating at 100 Full Duplex

### **NS CELERRA CS SHOULD HAVE INTERNAL INTERFACES SET AUTO-NEGOTIATE:**

#### **EXAMPLE OF AUTO-NEGOTIATE:**

eth0: negotiated 100baseTx-FD flow-control, link ok

eth1: negotiated 100baseTx-FD, link ok

### **ONE METHOD OF FORCING FULL DUPLEX ON CONTROL STATION EXT. INTERFACE:**

Step 1. Add following line to /etc/modules.conf file:

**alias eth2 tulip options=5 full\_duplex=1**

Step 2. Reboot

### **STOPPING & STARTING EXTERNAL INTERFACE FOR LINUX CS:**

#sbin/service network stop

#sbin/service network start

**Caution:** Executing the “network stop” option may induce CS Failover & CallHome if CS1 is configured

#ifdown eth2 ifup eth2

### **VERIFYING EXTERNAL NIC DRIVER FOR NS600/NS700 SYSTEMS:**

# /sbin/lsmod

eepro100 20592 2

# cat /etc/modules.conf

alias eth0 eepro100

alias eth0:0 eepro100

alias eth1 eepro100 →External interface for NS systems

# /sbin/mii-tool

eth0: negotiated 100baseTx-FD flow-control, link ok

eth1: no autonegotiation, 100baseTx-HD, link ok

### **VERIFYING OR CHANGING SPEED & DUPLEX SETTINGS ON NS600/700—ETH0/ETH1:**

# /sbin/mii-tool -V

mii-tool.c 1.9 2000/04/28 00:56:08 (David Hinds)

**eth0: negotiated 100baseTx-FD flow-control, link ok**

**eth1: negotiated 100baseTx-FD flow-control, link ok**

#/sbin/mii-tool -v

**eth0: negotiated** 100baseTx-FD flow-control, link ok

product info: Intel 82555 rev 4

basic mode: autonegotiation enabled

basic status: autonegotiation complete, link ok

capabilities: 100baseTx-FD 100baseTx-HD 10baseT-FD 10baseT-HD

advertising: 100baseTx-FD 100baseTx-HD 10baseT-FD 10baseT-HD flow-control

link partner: 100baseTx-FD 100baseTx-HD 10baseT-FD 10baseT-HD flow-control

**eth1: negotiated** 100baseTx-FD flow-control, link ok

product info: Intel 82555 rev 4

basic mode: autonegotiation enabled

basic status: autonegotiation complete, link ok

capabilities: 100baseTx-FD 100baseTx-HD 10baseT-FD 10baseT-HD

advertising: 100baseTx-FD 100baseTx-HD 10baseT-FD 10baseT-HD flow-control

link partner: 100baseTx-FD 100baseTx-HD 10baseT-FD 10baseT-HD flow-control

# cat /etc/modules.conf

alias eth0 eepro100

alias eth0:0 eepro100

alias eth1 eepro100

# /sbin/ifconfig -a

eth0 Link encap:Ethernet HWaddr 00:02:B3:EC:72:CF

→Primary Internal Network

inet addr:192.168.1.100 Bcast:192.168.1.255 Mask:255.255.255.0

UP BROADCAST RUNNING MULTICAST MTU:1500 Metric:1

RX packets:27909342 errors:0 dropped:0 overruns:0 frame:0

TX packets:48706157 errors:0 dropped:0 overruns:0 carrier:0

collisions:0 txqueuelen:100

RX bytes:4227932213 (4032.0 Mb) TX bytes:426098779 (406.3 Mb)

Interrupt:10 Base address:0xc000

eth0:0 Link encap:Ethernet HWaddr 00:02:B3:EC:72:CF

→Backup Internal Network

inet addr:192.168.2.100 Bcast:192.168.2.255 Mask:255.255.255.0

UP BROADCAST RUNNING MULTICAST MTU:1500 Metric:1

Interrupt:10 Base address:0xc000

eth1 Link encap:Ethernet HWaddr 00:02:B3:EC:72:D0

→External Network

inet addr:10.241.169.57 Bcast:10.241.169.255 Mask:255.255.255.0

UP BROADCAST RUNNING MULTICAST MTU:1500 Metric:1

RX packets:22758115 errors:0 dropped:0 overruns:0 frame:0

TX packets:15887600 errors:0 dropped:0 overruns:0 carrier:0

collisions:0 txqueuelen:100  
RX bytes:2083332748 (1986.8 Mb) TX bytes:1160481442 (1106.7 Mb)  
Interrupt:10 Base address:0xe000

**# cat /nas/site/nas\_param**

primary\_if:el30:192.168.1.0:255.255.255.0:192.168.1.255:fxp0: →Primary network NS600G with Intel eepro 100 Interface (fxp0)  
backup\_if:el31:192.168.2.0:255.255.255.0:192.168.2.255:fxp0: →Secondary network

**Note:** File configuration will be different if Secondary Control Station is configured, as in NS700 systems  
cluster\_if:xxxxxx

**SETTING FULL DUPLEX LINUX CS ETH0/ETH1/ETH3 INTERFACE:**

1. **# /sbin/mii-tool -F 100baseTx-FD eth1**

**# /sbin/mii-tool -V**

eth1: 100 Mbit, full duplex, link ok  
2. Add following line to end of /etc/rc.d/rc.local

**/sbin/mii-tool -F 100baseTx-FD eth1**

**Note:** If having issues setting eth3 to FDX, see primus emc124296, and add following to /etc/modules.conf file:  
alias eth3 e1000

**options e1000 Speed=0,0,100 Duplex=0,0,2**

**RESETTING AUTO NEGOTIATE & FLOW CONTROL FOR ETH0/ETH1:**

**# /sbin/mii-tool -A -r**

Invalid media specification '-r'.

restarting autonegotiation...

restarting autonegotiation...

**# /sbin/mii-tool -V**

**eth0: negotiated 100baseTx-FD flow-control, link ok**

eth1: negotiated 100baseTx-FD flow-control, link ok

**# /sbin/mii-tool -A 100baseTX-FD eth1** (Setting Autonegotiate without Flow Control)

restarting autonegotiation...

**# /sbin/mii-tool -V**

mii-tool.c 1.9 2000/04/28 00:56:08 (David Hinds)

eth0: negotiated 100baseTx-FD flow-control, link ok

**eth1: negotiated 100baseTx-FD, link ok**

**NAS 5.6 CS INTERFACE UTILITY:**

**# /sbin/ethtool eth3**

Settings for eth3:

Supported ports: [ TP ]

Supported link modes: 10baseT/Half 10baseT/Full  
100baseT/Half 100baseT/Full  
1000baseT/Full

Supports auto-negotiation: Yes

Advertised link modes: 10baseT/Half 10baseT/Full  
100baseT/Half 100baseT/Full  
1000baseT/Full

Advertised auto-negotiation: Yes

Speed: 100Mb/s

Duplex: Full

Port: Twisted Pair

PHYAD: 0

Transceiver: internal

Auto-negotiation: on

Supports Wake-on: umbg

Wake-on: d

Current message level: 0x00000007 (7)

Link detected: yes

**# /sbin/mii-tool eth3 -v**

eth3: negotiated 100baseTx-FD, link ok

product info: vendor 00:aa:00, model 56 rev 0

basic mode: autonegotiation enabled

basic status: autonegotiation complete, link ok

capabilities: 100baseTx-FD 100baseTx-HD 10baseT-FD 10baseT-HD

advertising: 100baseTx-FD 100baseTx-HD 10baseT-FD 10baseT-HD flow-control

link partner: 100baseTx-FD 100baseTx-HD 10baseT-FD 10baseT-HD

**Note:** Above is a comparison of the two tools. Ethtool allows one to make very specific and extensive param changes.

## **PROBLEM WITH NS CONTROL STATIONS WITH INCORRECT INTERNAL NIC SETTING:**

If Control Station internal settings are not set as in the following example, then it's possible that NAS Services, and NBS partitions, will not be properly started and mounted:

# /sbin/mii-tool -V

eth0: negotiated 100baseTx-FD flow-control, link ok

eth1: negotiated 100baseTx-FD, link ok → Normal and desired to have CS auto-negotiate

eth0: 100 Mbit, full duplex, link ok → This setting is not o.k.

## **SETTING LINUX 6.2 CS EXTERNAL NIC (NAS 2.2) TO FULL DUPLEX:**

Step 1. vi /etc/modules.conf file and add following entry for Tulip Card

**alias eth2 tulip options=3 full\_duplex=1 [100 Full Duplex]**

0=Auto; 1=10M; 3=100M

### **Configuring External Linux CS Interface for Auto:**

#vi /etc/conf.modules >>> alias eth2 tulip full\_duplex=1

### **Forcing Full Duplex 100baseTX on External Linux CS Interface:**

#vi /etc/conf.modules >>> alias eth2 tulip options=3 full\_duplex=1

## **SETTING DEFAULT GATEWAY ON LINUX:**

#/sbin/route add default gw 192.10.3.254 [good until next reboot]

# vi /etc/sysconfig/network GATEWAY=192.10.3.254 (example of default route line)

# /sbin/route delete default gw 192.1.4.1 [to delete a default gw]

## **SETTING OR CHANGING CONTROL STATION IP, HOSTNAME, etc.**

### **I. SET-UP TEMPORARY CONTROL STATION IP & ROUTE:**

# /sbin/ifconfig eth3 192.1.4.250 netmask 255.255.255.0 broadcast 192.1.4.255

# /sbin/route add default gw 192.1.4.254

### **II. USE CELERRA MANAGER TO MAKE NETWORK SETTINGS & HOSTNAME PERMANENT:**

1. Go to Celerra Home>Control Station Properties: Set Hostname, IP Address, Netmask, Gateway, and any other pertinent network information, then click Apply

2. Reboot Control Station and login via SSH

## **USING CLI TO CONFIGURE INTERNAL IP ON LINUX:**

Step 1. Determine Interface Name for Linux: #cat /etc/sysconfig/network

“Networking=Yes Forward\_IPv4=False Hostname: c7cs0

Gateway: 193.1.21.254 Gateway\_Dev=eth2 [this is CS0] NSO=193.1.21.1”

Step 2. #cd /etc/sysconfig/network-scripts #more ifcfg-eth2 [configure accordingly]

“DEVICE=eth2 IPAddr=193.1.21.170 Netmask=255.255.255.0 Broadcast=193.1.21.255 Onboot=Yes”

## **USING SYSCTL TO TUNE LINUX:**

The ‘System Control’ interface can be used to tweak system variables found in /proc/sys

/proc/sys/net → Networking

/proc/sys/fs → File Systems

/proc/sys/vm → Virtual Memory

## **SETTING UP LINUX CS FOR ROUTING/GATEWAY:**

By default, we do not enable ip\_forward prior to NAS 5.5.30, meaning that for multi-homed systems, packets will not be routed between physical interfaces. In order to route or create a gateway between different physical interfaces on the same system, enable IP Forwarding using the following:

### **Method I.**

# echo 1 > /proc/sys/net/ipv4/ip\_forward

**Method II.**

**# /sbin/sysctl -w net.ipv4.ip\_forward=1**

**Note:** The above methods do not make the change persistent on reboots, however.

**MAKING IP FORWARD PERSISTENT:**

1. Add following string to /etc/sysctl.conf file

net.ipv4.ip\_forward = 1

2. # /sbin/sysctl -p /etc/sysctl.conf or # /sbin/service network restart

3. Alternatively, edit the following file and set IP Forwarding to “true”, and reboot system:

/etc/sysconfig/network

FORWARD\_IPV4=true

**Note:** With NAS 5.5.30 and higher, we now enable ip\_forwarding by default

**CHANGING DEFAULT PRIVATE LAN IP ADDRESSES ON NS600/700 INTEGRATED or GATEWAY NS600G/700G:**

**Note:** Internal IP Address change would not be applicable for Gateway models for SPA or SPB as a Public IP Address is assigned for web management, but Internal IP Address scheme may need change on Control Station and Data Mover subnets.

**RULES & LIMITATIONS FOR CHANGING INTERNAL IP ADDRESSES:**

→Must change the private network for Celerra systems if the customer is already using the default subnet for public network

→Default subnets are 192.168.1 and 192.168.2

→New subnet with NS700G 4-Way 192.168.3.xxx for IPMI [Intelligent Platform Management Interface] connection between dual Control Stations only

→Cannot change last octet of CS or DM Internal IP addresses, only first 3 octets [i.e., 100, 2, 3, 200, or 201]

→Must use standard Class C subnet masks—255.255.255.0

→All Primary private IP addresses need to be on the same subnet

**Note:** External IP Address of CS must be on same subnet as SPA/SPB and set to auto-negotiate

**DEFAULT PRIVATE NETWORK ADDRESSES**

emcnas\_i0 192.168.1.100

emcnas\_i1 192.168.2.100

server\_2 192.168.1.2

server\_2b 192.168.2.2

server\_3 192.168.1.3

server\_3b 192.168.2.3

SPA 192.168.1.200

SPB 192.168.1.201

1. Stop NAS Services

2. #mount /nbsnas /nas /nas/dos /nas/var

3. Edit following files to change CS Subnet address: /nas/site/nas\_param site\_param cs\_aliases

4. Edit following files to change CS & DM addresses: /nas/server/slot\_2/ifconfig eof nbs.cs.ro nbs.cs.rw export

5. Regenerate Data Mover Boot.cfg File on DM\_2:

**#/nas/sbin/build\_config /nas/server/slot\_2 /nas/dos/slot\_2**

6. Unmount /nas/dos /nas/var /nas /nbsnas

7. Stop NBS Service: **#/sbin/service nbs stop**

8. #mount /nas

9. Change IP Address of SP-A

**# /nasmdc/sbin/navicli -h 10.241.168.52 networkadmin -get**

Storage Processor: SP A

Storage Processor Network Name: nyip2spa

Storage Processor IP Address: 10.241.168.52

Storage Processor Subnet Mask: 255.255.255.0

Storage Processor Gateway Address: 10.241.168.128

**#/nas/sbin/navicli -h SPA networkadmin -set -name <new name> -address <new IP> -gateway <new gateway> -subnetmask <new mask>**

**Note:** Storage System reboots

10. Wait for SPA to come up, then ping new address to verify, repeat steps for changing IP on SPB

11. Edit CS Network Files to change IP Addresses;

**#vi /etc/sysconfig/network-scripts/ifcfg-eth0 | ifcfg-eth0:0**

**Note:** Broadcast address must end with 255

12. Restart Network Services to put new Address Scheme into effect: **#/sbin/service network restart**

13. Edit /etc/hosts to change all addresses to new values

14. Update CS with New IP of DM\_2:

**# /nasmcd/sbin/t2tty -c 2 "ifconfig el30 dev=fxp0l=<new IP> b=<new broadcast> n=255.255.255.0"**

**# /nasmcd/sbin/t2tty -c 2 "ifconfig el31 dev=fxp0l=<new IP> b=<new broadcast> n=255.255.255.0"**

15. Set HTTP port of DM\_2:

**# /nasmcd/sbin/t2tty -c 2 "mac allow=<cs IP> : <cs backup> httpport=5080"**

16. Ping SP-A, SP-B, Server\_2, and Server\_2b to verify changes

17. #umount /nas

18. Repeat Steps for Slot\_3, if required

19. Synchronizes Control Station view of backend with the Storage Array Subsystem:

**# nas\_storage -sync id=1**

**Note:** If this command fails, try moving the symapi\_db.bin file aside, run nas\_diskmark -all, and then nas\_storage -sync

20. Reboot data movers and verify by pinging

21. Restart NAS Services: **#/sbin/service nas start**

**Note:** Changing Server output to local console, and restoring logging to Server Log

**# /nasmcd/sbin/t2tty -c 2 "logsys add output console"**

**# /nasmcd/sbin/t2tty -c 2 "logsys delete output console"**

### **RUNNING SCRIPT FOR SCREEN CAPTURE ON LINUX CS:**

**#cd; script or #script filename** [Begins file capture of screen session to file ‘typescript’} –Use ctrl + d to end script session

### **EXAMINING SW PACKAGES INSTALLED ON LINUX CS:**

**# rpm -qa |more**

**# rpm -qa --last** [Displays latest rpm packages installed on CS in order]

ypserv-2.8-0.72E Fri 27 Feb 2004 09:19:40 AM EST

yp-tools-2.5-1 Fri 27 Feb 2004 09:19:39 AM EST

ypbind-1.8-1 Fri 27 Feb 2004 09:19:39 AM EST

tftp-server-0.17-14 Fri 27 Feb 2004 09:19:39 AM EST

### **NAS AND LINUX VERSION INFORMATION:**

| NAS Version | RedHat Version    | Kernel Version |
|-------------|-------------------|----------------|
| 2.2         | 6.2               | ?              |
| 3.0         | 6.2               | 2.2.16-0       |
| 4.0         | ?                 | ?              |
| 4.1         | ?                 | ?              |
| 4.2         | ?                 | ?              |
| 5.0         | ?                 | ?              |
| 5.1         | 7.1               | 2.4.9-34       |
| 5.2         | 7.1               | 2.4.9-34       |
| 5.3         | 7.1               | 2.4.9-34       |
| 5.4         | 7.1               | 2.4.9-34       |
| 5.5         | 7.2               | 2.4.20-28      |
| 5.6         | 3.4.4-2 (RHEL4.2) | 2.6.9-42       |

### **DETERMINING CONTROL STATION INFO:**

**\$cat /proc/cpuinfo**

**\$free**

```
total     used     free   shared  buffers   cached
Mem:      513624    473660    39964     156    111348    116536
-/+ buffers/cache:  245776   267848
Swap:     522104    11184   510920
```

**\$netstat -i \$netstat -rn \$/sbin/ifconfig -a**

**\$ls -la /proc**

**\$cat /proc/meminfo**

**#cat /proc/filesystems** represents all processes running in memory by filename

**# cat /proc/version**

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
Linux version 2.4.9-34.5403.EMC (root@hungry) (gcc version 2.96 20000731 (Red Hat Linux 7.1 2.96-98)) #1 Fri Dec 3 11:15:12  
EST 2004

**# cat /proc/cmdline**

root=/dev/sdc3 boot=/dev/sdc1 panic=1 max\_scsi\_luns=16

BOOT\_IMAGE=vmlinuz

**# cat /etc/sysconfig/hwconf (abridged)**

class: NETWORK

bus: PCI

detached: 0

device: eth

**driver: tulip**

desc: "DECIDEChip 21142/43"

vendorId: 1011

deviceId: 0019

subVendorId: 1109

subDeviceId: 2a00

pciType: 1

-

class: SCSI

bus: PCI

detached: 0

driver: unknown

**desc: "Emulex Corporation|LP8000 Fibre Channel Host Adapter"**

vendorId: 10df

deviceId: f800

subVendorId: 10df

subDeviceId: f800

pciType: 1

-----

class: CDROM

bus: IDE

detached: 0

device: hda

driver: ignore

**desc: "TOSHIBA CD-ROM XM-6602B"**

-

class: HD

bus: SCSI

detached: 0

**device: sda**

driver: ignore

desc: "Dgc RAID 5"

host: 0

id: 1

channel: 0

**lun: 0**

-

class: HD

bus: SCSI

detached: 0

**device: sdb**

driver: ignore

desc: "Dgc RAID 5"

host: 0

id: 1

channel: 0

**lun: 1**

-

class: HD

bus: SCSI

detached: 0

**device: sdc**

driver: ignore

desc: "Dgc RAID 5"

host: 0

id: 1

channel: 0

**lun: 2**

-

class: HD

bus: SCSI

detached: 0

**device: sdd**

driver: ignore

desc: "Dgc RAID 5"

host: 0

id: 1

channel: 0

**lun: 3**

-

class: HD

bus: SCSI

detached: 0

**device: sde**

driver: ignore

desc: "Dgc RAID 5"

host: 0

id: 1

channel: 0

**lun: 4**

-

class: HD

bus: SCSI

detached: 0

**device: sdf**

driver: ignore

desc: "Dgc RAID 5"

host: 0

id: 1

channel: 0

**lun: 5**

**CONTROL STATION HARDWARE CLOCK INFO:**

# /sbin/hwclock --show      /etc/sysconfig/clock      /etc/rc.d/rc.sysinit [startup script file]      /var/log/boot.log

**LINUX CONTROL STATION ‘SEGMENTATION FAULT’:**

Error occurs when trying to execute commands on CS due to low memory available:

# cat /proc/meminfo

total: used: free: shared: buffers: cached:

Mem: 525950976 518840320 7110656 69632 75505664 211230720

Swap: 541466624 9531392 531935232

MemTotal: 513624 kB

MemFree: 6944 kB----output abridged-----

**VERIFYING SCSI CONTROL STATION MAPPINGS:**

#cat /proc/scsi/scsi

Attached devices:

Host: scsi0 Channel: 00 Id: 00 Lun: 00

Vendor: EMC Model: SYMMETRIX Rev: 5670

Type: Direct-Access ANSI SCSI revision: 02

Host: scsi0 Channel: 00 Id: 00 Lun: 01

Vendor: EMC Model: SYMMETRIX Rev: 5670

Type: Direct-Access ANSI SCSI revision: 02

```
Host: scsi0 Channel: 00 Id: 00 Lun: 02
  Vendor: EMC  Model: SYMMETRIX  Rev: 5670
  Type: Direct-Access      ANSI SCSI revision: 02
Host: scsi0 Channel: 00 Id: 00 Lun: 03
  Vendor: EMC  Model: SYMMETRIX  Rev: 5670
  Type: Direct-Access      ANSI SCSI revision: 02
Host: scsi0 Channel: 00 Id: 00 Lun: 04
  Vendor: EMC  Model: SYMMETRIX  Rev: 5670
  Type: Direct-Access      ANSI SCSI revision: 02
Host: scsi0 Channel: 00 Id: 00 Lun: 05
  Vendor: EMC  Model: SYMMETRIX  Rev: 5670
  Type: Direct-Access      ANSI SCSI revision: 02
```

### **LISTING KERNEL MODULES ON CS:**

```
#/sbin/lsmod
#/sbin/modprobe -c
# Generated by modprobe -c (2.4.18)
path[boot]=/lib/modules/boot
path[toplevel]=/lib/modules/2.4.9-34.5309.EMC
path[toplevel]=/lib/modules/2.4
path[kernel]=/lib/modules/kernel
-----abridged-----
```

### **CREATING LINUX BOOT FLOPPY ON CONTROL STATION:**

```
#/sbin/mkbootdisk `uname -r`
```

Linux Chat Windows: talk; write; wall

### **Using Stat Command for File & Inode Information:**

Example: # stat /etc/motd

### **DETERMINING SUPERBLOCK INFORMATION ON LINUX CS VOL's—Red Hat 7.2:**

Note: Example is for /nas/var directory & must be unmounted to run command

```
# /sbin/mke2fs -n /dev/sdf1
```

```
mke2fs 1.26 (3-Feb-2002)
Filesystem label=
OS type: Linux
Block size=4096 (log=2)
Fragment size=4096 (log=2)
231360 inodes, 461860 blocks
23093 blocks (5.00%) reserved for the super user
First data block=0
15 block groups
32768 blocks per group, 32768 fragments per group
15424 inodes per group
Superblock backups stored on blocks:
         32768, 98304, 163840, 229376, 294912
```

Comment: Mke2fs command is used to create an ext2 file system

### **DUMPING INFORMATION ON PARTITIONS:**

```
# /sbin/dumpe2fs /dev/sda1
```

### **EASY METHOD FOR RUNNING FSCK ON ROOT PARTITION:**

```
# touch /forcefsck
```

```
# reboot
```

### **RUNNING FSCK ON LINUX CONTROL STATION ROOT VOLUME:**

1. In Dual Control Station environments, log into CS1, unmount partition, and run command

```
# /sbin/e2fsck -f -y -v /dev/sdc3
```

Caution: Do not run this command without seeking a 2<sup>nd</sup> opinion!

2. Alternatively, if single Control Station, boot to single user mode, or unmount volume and run e2fsck

Comment: e2fsck is used to FS Check a linux second extended FS

## **RUNNING FSCK ON DOS PARTITION:**

```
# /sbin/fsck -t msdos -V -v -l /dev/nd1
```

## **CHANGING LINUX PART. TO EXT2 SO AS TO RUN FSCK:**

### **EXAMPLE:**

Control Station can see all NBS partitions, but cannot fsck /dev/nde1 (nbsnas)--error SIGSEGV

Feb 9 16:26:56 scosnas01 **kernel: EXT3-fs: invalid journal inode.**

Feb 9 16:26:56 scosnas01 mcd\_helper: : Failed to mount /nbsnas (32)

Feb 9 16:26:56 scosnas01 EMCServer: nas\_mcd: File system mount(s) failed - exiting

Feb 9 16:26:56 scosnas01 EMCServer: nas\_mcd: Slot 0 exiting

### **1. Verify File System partition info:**

```
# /sbin/dumpe2fs /dev/ndf1
```

dumpe2fs 1.26 (3-Feb-2002)

Filesystem volume name: <none>

Last mounted on: <not available>

Filesystem UUID: b19939df-f59d-4297-9b16-0764f88126aa

Filesystem magic number: 0xEF53

Filesystem revision #: 1 (dynamic)

Filesystem features: **has\_journal** filetype needs\_recovery sparse\_super

Filesystem state: clean

Errors behavior: Remount read-only

Filesystem OS type: Linux

### **2. Remove Journal File System type:**

```
# /sbin/tune2fs -O"!^has_journal" /dev/ndf1
```

tune2fs 1.26 (3-Feb-2002)

### **3. Verify:**

```
# /sbin/dumpe2fs /dev/ndf1
```

dumpe2fs 1.26 (3-Feb-2002)

Filesystem volume name: <none>

Last mounted on: <not available>

Filesystem UUID: 625050cd-4287-4cd0-8ec0-5587dcb4e0ea

Filesystem magic number: 0xEF53

Filesystem revision #: 1 (dynamic)

Filesystem features: filetype sparse\_super

### **4. FSCK File System:**

```
# /sbin/fsck /dev/ndf1
```

fsck 1.26 (3-Feb-2002)

e2fsck 1.26 (3-Feb-2002)

/dev/ndf1: clean, 2468/231360 files, 75301/461860 blocks

### **5. Recreate the Journal EXT3 File System Type:**

```
# /sbin/tune2fs -j /dev/ndf1
```

tune2fs 1.26 (3-Feb-2002)

Creating journal inode: done

This filesystem will be automatically checked every 20 mounts or 180 days, whichever comes first.

## **LINUX FILE SYSTEM DEBUG COMMAND:**

```
# /sbin/debugfs /dev/nde1
```

debugfs 1.26 (3-Feb-2002)

debugfs: stat <8>

Inode: 8 Type: regular Mode: 0600 Flags: 0x0 Generation: 0

User: 0 Group: 0 Size: 33554432

File ACL: 0 Directory ACL: 0

Links: 1 Blockcount: 65616

## **PROCEDURE TO USE WHEN FSCK'ing ROOT FILE SYSTEM ON LINUX:**

1. Boot from floppy & select “command prompt”
2. #cd /dev and type #mknod sda b 8 0 #mknod sda1 b 8 1 #mknod sda2 b 8 2 #mknod sda3 b 8 3
3. #fdisk /dev/sda & p to print out Partition Table

Ensure partitions look as follows: Then type q to quit

|           |    |     |         |    |            |
|-----------|----|-----|---------|----|------------|
| /dev/sda1 | 1  | 1   | 8001    | 83 | Linux      |
| /dev/sda2 | 2  | 18  | 136552+ | 82 | Linux swap |
| /dev/sda3 | 19 | 264 | 1975995 | 83 | Linux      |

4. Run FSCK: **#e2fsck -p /dev/sda3**

5. Mount Root: #mount /dev/sda3 /mnt

6. Verify size of “init” file: #cd /mnt/sbin;ls -la init [Size should be 25968 & file should be executable]

7. #cd / & umount: #exec umount /mnt

## **RUNNING FSCK ON NAS WHEN CS CANNOT MOUNT:**

**Note:** Stop NAS Services first

**# /sbin/fsck -t ext3 -y /dev/hda5**

fsck 1.26 (3-Feb-2002)

e2fsck 1.26 (3-Feb-2002)

/dev/hda5 has gone 189 days without being checked, check forced.

Pass 1: Checking inodes, blocks, and sizes

Error reading block 65659 (Attempt to read block from filesystem resulted in short read) while doing inode scan. Ignore error? yes

Pass 2: Checking directory structure

Pass 3: Checking directory connectivity

Pass 4: Checking reference counts

Pass 5: Checking group summary information

/dev/hda5: 6818/262144 files (1.5% non-contiguous), 154703/524112 blocks

**# /sbin/fsck -t ext3 -y /dev/hda5**

**Note:** Ran fsck again, and was clean

## **RUNNING FSCK ON CS0 FROM CS1:**

1. Power down CS0 #/nas/sbin/t2reset pwroff -s 0

2. Identify Root & Boot partitions on CS0: #cat /nas/dos/slot\_0/boot.bat

**CS0=/dev/sdc1 for boot & /dev/sdc3 for root**

**CS1=/dev/sdd1 for boot & /dev/sdd3 for root**

3. Run FSCK on boot & root partitions:

#/sbin/e2fsck -y -p /dev/sdc1

4. Mount root and boot partitions:

#mount /dev/sdc3 /mnt #mount /dev/sdc1 /mnt/boot #umount /mnt/boot

5. Powerup CS0: **#/nas/sbin/t2reset pwron -s 0**

## **FAILED 2.2 – 4.2 UPGRADE, RESULTING IN BAD JOURNAL FS:**

1. If CS1 is available, run following steps from CS1 against CS0’s Boot & Root File Systems

2. If running from CS0, unmount /nas partition first

3. Remove journal from device to FSCK:

**# /sbin/tune2fs -O ^has\_journal /dev/sdc3**

4. Run FSCK:

**# /sbin/e2fsck -f -y /dev/sdc3** [Run until fsck runs clean]

5. Mount Device:

#mount /dev/sdc3 /mnt

#mount [Run command to verify ext2/ext3]

6. Run command to create new Journal:

**# /sbin/tune2fs -j -c 60 /dev/sdc3**

7. #umount /mnt and verify that device reflects ext3 journaling

8. Repeat above steps for /boot file system if required

9. #cat /etc/fstab to verify ext3 journaling on both CS0 & CS1 side

10. Vi edit /etc/fstab to change file systems from ext2 to ext3

11. Verify internal IP address scheme

## **VERIFYING CONSISTENCY OF LINUX MOUNTPOINTS:**

**# df -h**

| Filesystem | Size | Used | Avail | Use% | Mounted on |
|------------|------|------|-------|------|------------|
| /dev/sdc3  | 1.5G | 1.1G | 400M  | 72%  | /          |

```
/dev/sdc1      30M 4.3M 24M 15% /boot
server_3:/    126M 13M 113M 11% /mnt
/dev/sde1      1.7G 683M 1003M 41% /nas
/dev/sda1      133M 126M 7.7M 95% /nas/dos
/dev/sdf1      1.7G 61M 1.5G 4% /nas/var
```

### #df -T \*

Displays partition types

```
# /sbin/dumpe2fs /dev/sdc1
```

dumpe2fs 1.26 (3-Feb-2002)

Filesystem volume name: /boot

Last mounted on: <not available>

Filesystem UUID: ed10b020-1683-11d7-85c4-a44a2b288208

Filesystem magic number: 0xEF53

Filesystem revision #: 1 (dynamic)

Filesystem features: has\_journal filetype needs\_recovery sparse\_super

Filesystem state: clean

Errors behavior: Panic

Filesystem OS type: Linux

Inode count: 8032

Block count: 32098

Reserved block count: 1604

-----output abridged-----

## Opening Up Multiple "X-Term" Sessions on Linux Control Station from Commandline:

Alt + F1 or F2 keys to toggle between sessions

## Creating Files on Linux File System:

```
# dd if=/dev/zero of=jeff1 bs=1024 count=1000
```

## USING TAR TO ARCHIVE & UNARCHIVE FILES & DIRECTORIES:

```
# tar -cvf usrmap.tar * [Creates tar file of all unhidden files from current directory]
```

```
# tar -xvf usrmap.tar [Untars file in present directory]
```

```
# tar -tvf usrmap.tar [Use this command to list the contents of a Tar file]
```

## TARRING AND GZIPPING IN ONE OPERATION:

```
# tar -zcvf usrmap.tar.gz [Tarring up directory and zipping]
```

```
# tar -zxvf usrmap.tar.gz [Unzipping and untarring directory]
```

## EXTRACTING AN ARCHIVE FILE:

```
# tar -xf archive.tar
```

## TARRING SEVERAL FILES BUT NOT ALL:

```
# tar cvf /home/nasadmin/tartest.tar a b c
```

## TARRING ALL FILES FROM A DIRECTORY BUT NOT IN SUBDIRECTORIES:

```
# tar cvf autoupload.tar --exclude="*/\/*" autoupload/
```

Note: Will tar all contents of autoupload directory but no subdirectories

## TARRING ALL FILES IN A PARTICULAR DIRECTORY, INCLUDING SUBDIRS:

```
# tar cvf support_materials.070117_1616.tar fcs01.070117_1616/*
```

Note: Above command creates a support\_materials filename and tars all log files and directories underneath the fcs directory

## TAR SWITCHES:

-t Listing contents of tar file

-f append contents to existing tar file

-u append only newer files to existing archive

-v verbose

-x Extracting

-c Create an archive file

-f File=ARCHIVE

-z Zip or Unzip a gzip file at same time as tarring or untarring

-Z compress

## **SETTING UP NFS FTP SERVICE ON DATA MOVER:**

Step 1. #server\_user server\_2 -add -passwd ftpuser

User ID: 102

Group ID: 101

Home Directory: /home/nasadmin

New Passwd: xtra10

Retype New Passwd: xtra10

**Note:** This procedure encrypts the password and places the passwd file automatically onto the datamover for “ftpuser”

Step 2. FTP to the IP address and log in to test account

## **SETTING UP PASSWD FILE--Linux DataMover, using "server\_user" command:**

For NAS Versions 2.2+: Linux uses MD5 encryption, must use #/nas/sbin/server\_user utility

### **CREATING USER CALLED NDMP AND POPULATION DATA MOVER's ./etc/passwd FILE:**

# **/nas/sbin/server\_user server\_2 -add -md5 -password ndmp**

User ID: 102

Group ID: 101

Home Directory: /home/nasadmin

New Passwd: ndmpndmp →Min. passwd length 7 characters

Retype New Passwd: ndmpndmp

**Note:** This procedure encrypts the password and appends the entry to the passwd file on the datamover

## **SETTING UP USER ACCOUNTS ON LINUX:**

1. Create Account: # **/usr/sbin/adduser -u uid -g gid -G secondarygid -d /usr/user1 user1**

2. Set Environmentals: # **cp /home/nasadmin/.bash\_profile /usr/user1**

## **CREATING NEW USER ACCOUNT ON LINUX CS WITH NASADMIN PRIVILEGES:**

1. Create User Account: # **/usr/sbin/useradd**

2. Change Passwd: # **passwd newuser**

3. Assign ‘newuser’ to NASADMIN Group 201: # **/usr/sbin/usermod -G 201 newuser**

## **SERVER USER COMMAND OPTIONS:**

\$server\_user server\_x -l [Outputs what is in the passwd file]

\$server\_user server\_x -m username [Edit an existing user's passwd entry]

\$server\_user server\_x -a -p newuser

## **MISCELLANEOUS TOPICS:**

### **Using Telnet to Linux Control Station:**

Problem: Cannot connect using Telnet or takes a long time

Fix: 1. Add Hostname and IP address of Client to CS /etc/hosts file.

2. Check /etc/resolv.conf file for errant entry for Domain [delete entry]. What happens is that an incorrect entry causes session to timeout, at which point it went to its Default Gateway, and then found a Name Resolution.  
\$netstat -r will show the route table it uses

3. Other option is to edit /etc/rc.d/rc.local file and add to bottom of file: “/bin/ping -c 1 193.1.21.200””

## **CELLERRA WATCHDOG TIMER:**

**Note:** Celerra uses a single wdog thread for each physical data mover, to monitor all cpu tasks/processes. Purpose of the wdog time is to watch for runaway processes due to lost CPU, hung threads, sluggish threads, lock contention, long structures, looping structures, or memory corruption.

→wdog sleeps for 1 second, then checks all processes on all CPUs, then goes to sleep, and repeats the cycle, setting a memory flag status on each process to complete when done. If the process cannot be checked complete, then when the NMI Handler fires (every 4 seconds), and finds an incomplete check, it will panic the Server.

→NMI is a Non-Maskable Interrupt which checks the results of the WatchDog timer, every 4 seconds, and panics the DM if the processes have not been successfully checked by Watchdog

→Memory Flags are what is used by wdog timer when checking all processes

→Watchdog Timer below 5.2 code is set at 120 seconds

→Watchdog Timer for NAS 5.2 code and higher is set at 4 seconds, with 30 second variation during bootup

## **SETTING WATCHDOG TIMEOUT VALUE ON LINUX CONTROL STATION:**

**Note:** Control Station fibre HBA is sensitive to fibre channel connectivity and may reboot if Zone change or other activity occurs that might cause a temporary loss of connectivity to the fabric.

1. Edit /etc/inittab
2. wd:3:respawn:/nasmcd/sbin/nas\_watchdog -f /dev/raw/raw1 **-t 90** [Add value to increase timeout value to 90 secs]

## **DISABLING WATCHDOG ON DM:**

**# .server\_config server\_x -v “sched watchdog stop”**

**Note:** Watchdog Timers are used to monitor threads in progress—if threads are not making any progress, i.e., are stuck in a loop, the Watchdog could panic the Server, etc.

## **FINDING WWN OF FIBRE CONTROL STATION FOR QLOGIC & EMULEX:**

### **EMULEX:**

**# ls -l /proc/scsi**

```
dr-xr-xr-x 2 root root 0 Sep 18 08:40 lpfc
```

**# cat /proc/scsi/lpfc/0**

Emulex LightPulse LPFC Driver Version: 4.21g.EMC

HBA: Emulex LightPulse LP8000 1 Gigabit PCI Fibre Channel Adapter

SerialNum: 0000c9266446

Firmware Version: 3.82A1

Hdw: 2002506d

VendorId: 0xf80010df

Portname: 10:00:00:00:c9:26:64:46 Nodename: 20:00:00:00:c9:26:64:46

Link Up - Ready:

PortID 0x21900

Fabric

Current speed 1G

lpfc0t00 DID 021300 WWPN 50:06:01:60:10:60:05:10 WWNN 50:06:01:60:90:60:05:10

**# cat /proc/scsi/lpfc/1**

Emulex LightPulse LPFC Driver Version: 4.21g.EMC

HBA: Emulex LightPulse LP8000 1 Gigabit PCI Fibre Channel Adapter

SerialNum: 0000c9266449

Firmware Version: 3.82A1

Hdw: 2002506d

VendorId: 0xf80010df

Portname: 10:00:00:00:c9:26:64:49 Nodename: 20:00:00:00:c9:26:64:49

Link Down

**Note:** Above example shows HBA0 & HBA1 ports on CS0 for Emulex driver. Rules for CNS14 & CFS14 are that only HBA0 should be zoned to the backend, and to SPA for Clariion arrays. Port 1 should NOT be attached.

### **QLOGIC:**

**# ls -l /proc/scsi**

```
dr-xr-xr-x 2 root root 0 Sep 18 08:46 qla2x00
```

**# cat /proc/scsi/qla2x00/0**

QLogic PCI to Fibre Channel Host Adapter for ISP2100/ISP2200/ISP2300:

Firmware version: 2.02.03, Driver version 4.47.18.EMC.2

Entry Address = e0817060

Host adapter: loop state= <READY>, flags= 0x20e0001

----output abridged-----

SCSI Device Information:

scsi-qla0-adapter-node=200100e08b246706;

scsi-qla0-adapter-port=210100e08b246706;

scsi-qla0-target-0=50060482bc019584;

**# cat /proc/scsi/qla2x00/1**

QLogic PCI to Fibre Channel Host Adapter for ISP2100/ISP2200/ISP2300:

Firmware version: 2.02.03, Driver version 4.47.18.EMC.2

Host adapter: loop state= <DOWN>, flags= 0x2060000

-----output abridged-----

SCSI Device Information:

scsi-qla1-adapter-node=200000e08b046706;

scsi-qla1-adapter-port=210000e08b046706;

**Note:** CNS14 & CFS14 should not use Port 1, HBA1, to connect to Symm or Clariion backends. Link should be down for Port 1 and should NOT show any targets configured—see above for proper output and configuration.

**FINDING WWN OF DATAMOVER:**

**\$server\_log server\_3 | grep -i wwn**

2005-01-27 08:50:11: CAM: 4: FCP ONLINE HBA 0: S\_ID 690813 WWN: 10000000c939ed01

2005-01-27 08:50:17: CAM: 4: FCP ONLINE HBA 1: S\_ID 610813 WWN: 10000000c939ed02

**# .server\_config server\_2 -v "fcp show"**

FCP ONLINE HBA 0: S\_ID 690813 WWN: 10000000c939ed01 LP9000 2 GHz →**S\_ID** represents Data Mover WWN

FCP scsi-0: HBA 0: D\_ID 691113 FA-02cb: 5006048accef60e1 Class 3 →**D\_ID** represents WWN for Symmetrix FA

FCP scsi-16: HBA 0: D\_ID 691413 FA-15cb: 5006048accef60ee Class 3

FCP ONLINE HBA 1: S\_ID 610813 WWN: 10000000c939ed02 LP9000 2 GHz

FCP scsi-32: HBA 1: D\_ID 610e13 FA-15ca: 5006048accef60ce Class 3

FCP scsi-48: HBA 1: D\_ID 611513 FA-02db: 5006048accef60f1 Class 3

**STOPPING/STARTING BOX MONITOR/NASMCD ON LINUX CONTROL STATION:**

**#/etc/rc.d/rc3.d/S95nas stop [start]** [Linux Red Hat 7.2]

**#/etc/rc.d/init.d/nas stop | start** [Works with Red Hat 7.2]

**#/sbin/services nas start** [Permission denied on Red Hat 7.2]

**Caution:** This will unmount /nas, /nas/dos, & /nas/var File Systems!! Remount manually if you need to run commands!

**Verify:** nas\_mcd & boxm processes are stopped

**#ps -ef |grep nas\_mcd**

root 11348 1 0 10:45 ? 00:00:00 /nasmcd/nas\_mcd -h /nasmcd /nas/

**#ps -ef |grep boxm**

root 11855 11504 0 10:45 ? 00:00:00 /nas/sbin/nas\_boxmonitor /nas -i

**MOUNTING LINUX CS TO DM:**

**#mount -a -F nfs 172.19.32.10:/fs1 /mnt**

**#df -k** #cd /mnt #cp -f -R /. [forces a copy of all files into the current working directory]

**COPYING FILES FROM (1) DATAMOVER TO ANOTHER USING CONTROL STATION:**

**#mount 192.168.1.4:/ /mnt** [Mounting to Internal IP address of Servers 4 & 5]

**#mount 192.168.1.5:/ /mnt1**

**#cd /mnt/todd/folder1** #cp /mnt1/jeff1/\* . & [Copying contents from Jeff1 to folder1]

**Linux Device Tables:** /etc/fstab

**CD-ROM Directory:** /dev/cdrom

**Floppy Directory:** /fd0

**MOUNTING LINUX CONTROL STATION TO CD-ROM:**

**# mount /dev/cdrom /mnt/cdrom or #mount -ro /dev/cdrom /mnt/cdrom**

**Unmounting CD-ROM:** # umount -a /mnt/cdrom

**Mounting CD-ROM to Linux Control Station :**

**# mount -t iso9660 /dev/cdrom /mnt/cdrom** #cd /mnt/cdrom;ls

**# mount -rt iso9660 /dev/cdrom /mnt/cdrom**

**UPGRADING CELERRA USING USB FLASH DRIVE & ISO IMAGE ON NS40**

**USING USB FLASH DRIVE TO TRANSFER NAS ISO IMAGE TO CONTROL STATION:**

1. Plug USB Flash Drive into USB port on front side of Control Station [located underneath the serial port]

2. Run Fdisk to verify the discovered device type:

**# /sbin/fdisk -l**

/dev/sda1 1 992 999813+ 6 FAT16

**# df -h**

usbdevfs /proc/bus/usb type usbdevfs (rw)

3. Create mountpoint for USB device:

# **mkdir -p /mnt/usb**

4. Mount the Flash Drive to the Control Station

# **mount /dev/sda1 /mnt/usb**

5. Copy the ISO image from the Flash Drive to a partition with enough available space (about 550MB needed)

# **cd /mnt/usb**

# **cp emcnas5.5.31.6.iso /iso**

6. Unmount Flashdrive and physically remove from the Control Station

# **umount /mnt/usb**

### **MOUNTING NAS ISO IMAGE ON LINUX CONTROL STATION:**

1. Mount the ISO image to the mountpoint defined above [/mnt/usb]

# **mount -t iso9660 -o loop /iso/emcnas5.5.31.6.iso /mnt/usb**

2. Perform NAS Upgrade from the ISO image

# **cd /mnt/usb/EMC/nas;/setup**

**Note:** After upgrading Linux kernel, system will require a reboot. Perform steps 1-2 again after the system has rebooted.

### **MOUNTING LINUX FLOPPY DRIVE:**

# **mount -t msdos /dev/fd0 /mnt/floppy #cd /mnt/floppy;ls**

### **MOUNTING CS1 FROM CS0 FOR COPYING FILES BACK & FORTH:**

**Note:** Only prerequisite is that you know what the path is for the root “/” partition of the control station that you want to mount. Sdd3 & sdc3 are normal partition names for CS1 & CS0, respectively for Red Hat 7.2

# **mount /dev/sdd3 /mnt [Mounts root partition of CS1]**

# **mount /dev/sdc3 /mnt [Mounts root partition of CS0]**

**Permanently Mounting DataMover 'rootfs' to Linux Control Station:** #mount 172.19.31.4:/ /mnt

**Note:** Linux assumes the filesystem "type" [-t nfs] is "nfs" unless otherwise specified

Remember to "# umount /mnt" after you are completed

### **COPYING FROM FLOPPY:**

**Copying Floppy to Control Station:** # dd if=/dev/fd0 of=/tmp/floppy bs=1024 [copies contents of floppy]

**Copying from Control Station to Floppy:** # dd if=/tmp/floppy of=/dev/fd0 bs=1024

**Copying Files From Floppy to Control Station:** #cp \* /scripts

# mount [Floppy Mount Listed: “/dev/fd0 on /mnt type msdos (rw)”]

**Unmounting the Floppy Drive:** # **umount -a /mnt/floppy** [not real clean, gives error, but does umount]

### **FORMATTING FLOPPY DISK ON LINUX:**

# fdformat /dev/fd0H1440

# mformat a:

# mc当地 ks.cfg a:ks.cfg

### **LOGGING IN AS ROOT USING LINUX PROFILE:**

\$su -

### **LOGGING IN AS ROOT USER VIA MODEM CONNECTION:**

Add the following line to the /etc/security file

ttys0

### **RECOVERING ROOT PASSWD IF ACCESS TO LILO PROMPT ON LINUX CS:**

Step 1. Reboot Control Station and at LILO: screen, type “linux single” without quotes

Step 2. # passwd root

Step 3. Enter new root password and reboot: # shutdown -r now

### **RECOVERING ROOT ACCESS TO LINUX CS IF MISSING PASSWORD & NO LILO SCREEN:**

Step 1. Boot from linux floppy and insert Installation CD-ROM when prompted

Step 2. Select <alt--f2> option on list and it will bring you to a "bash#" prompt

Step 3. #mknod /dev/sdc3 8 35

Step 4. #mkdir /mnt

Step 5. #mount /dev/sdc3 /mnt

Step 6. #cd /mnt/etc

Step 7. #vi shadow file and delete the encrypted portion of the root user's password line [root : :] and save

Step 8. Reboot Control Station and login normally as User "nasadmin"

Step 9. \$su

    Password: newpasswd [Just type in a new password]

### **RESETTING ROOT PASSWORD ON LINUX CS:**

1. Press F5 key at end of HBA discovery on screen [Starting MSDOS...]

2. c:>rescue [Type "rescue" to enter Linux rescue mode]

3. #mkdir /mnt/tmp   #mount /dev/sdc3 /mnt/tmp [Mount root of linux to temporary mountpoint]

4. #cd /mnt/tmp/etc

5. #vi passwd and delete password field for nasadmin and root

6. Save and reboot CS

### **USING STRACE COMMAND ON LINUX TO TRACE COMMAND OUTPUT:**

# strace -F -o trace4 server\_export server\_4 [outputs file named "trace4" for server\_export command]

# strace -f -usrmapstart.out /etc/rc.d/rc3.d/S99usrmap start &

# strace -o /tmp/aaa -f server\_export server\_2 -p -u -P cifs -n testshare

### **EXTREMELY USEFUL DEBUG OF SHELL SCRIPTS ON LINUX:**

# /bin/sh -x nas\_version

```
+ ECHO=/bin/echo -e
+ get_kit
+ /bin/echo -e /nas
+ grep emctst
+ '[' 1 = 0 ']'
+ '[' -d /nas_standby ']'
+ kit=emcnas
+ PKG_DB=/var/sadm/pkg/emcnas
+ HOME=/root
+ export HOME
+ SQUERY=/bin/rpm --qf '%{VERSION}-%{RELEASE}' --dbpath /var/sadm/pkg/emcnas -q
+ LQUERY=/bin/rpm --dbpath /var/sadm/pkg/emcnas -qi
+ parse_args
+ '[' -d /nas_standby ']'
+ /bin/rpm --qf '%{VERSION}-%{RELEASE}' --dbpath /var/sadm/pkg/emcnas -q emcnas
5.1.9-4+ /bin/echo -e
```

**Comment:** Echoes result of script as it runs

### **NEW TOOL FOR SHOWING SERVER THREADS:**

Step 1. #cd /nas/tools

Step 2. #ln -s mac\_db \_mac\_db

Step 3. #./\_mac\_db server\_2

>show thread   >show thread /SMB030 [shows partial stack trace of code]

>quit

### **USING STRACE ON LINUX TO TRACE PROCESS:**

\$strace -o ch.trace -p <PID #>

### **USING WATCH COMMAND TO MONITOR COMMAND OUTPUTS:**

# watch -n 10 server\_sysstat server\_2

Every 10s: server\_sysstat server\_2                          Mon Jul 28 20:34:26 2003

server\_2 :

```
threads runnable = 3
threads blocked = 201
threads I/J/Z = 1
memory free(kB) = 189599
cpu idle_% = 99
```

**Note:** Command will execute every 10 seconds and output to screen continuously until interrupted [default = 2secs updates]

### **ADDING USERS/GROUPS TO LINUX CS:**

#**/usr/sbin/useradd -u 102 -g 101 -d /home/nasadmin -s /bin/bash -p xtra10 iisroot**

Explanation:      UID=102  
                  GID=101  
                  -d=Home Directory  
                  -s=shell type  
                  -p=password field  
Last Entry=new user "iisroot"

#**/usr/sbin/groupadd -g 101 officemax**

**Note:** These commands will update the respective /etc/passwd; /etc/shadow; /etc/group; and /etc/shadow files on the Control Station

### **NFS “CHOWN” OR “CHGRP” COMMAND FAILS TO EXECUTE:**

For Users other than “root”, symptoms are that Client Unix system Users try changing ownership on files located on a remote mount on the Celerra and receive message indicating that Users or existing Owners of files do not have rights to change them:

#### **EXAMPLE:**

\$imisd8:/home/damnypd2>**chown admin:admin test**

**test: Not owner**

Similar operation conducted from NFS Clients to HP-UX11 remote filesystem succeeds. By default Celerra supports POSIX file systems which do not allow the above "chown" or "chgrp" operations to succeed when issued by any User other than "Root".

To disable POSIX compliance and enable "chown" to work for non-root Users, add following param to DM and Reboot:

**param filesystem rstchown=0**

**Note:** Default Celerra value is rstchown=1

### **CHANGING USER/GROUP OWNERSHIP ON LINUX:**

```
# ls -la
-rw-rw-r-- 1 root    root      0 Mar 15 11:07 um_group
-rw-rw-r-- 1 root    root      0 Mar 15 11:07 um_user
# chown nasadmin:nasadmin um_group um_user
#ls -al
-rw-rw-r-- 1 nasadmin nasadmin   0 Mar 15 11:07 um_group
-rw-rw-r-- 1 nasadmin nasadmin   0 Mar 15 11:07 um_user
```

**Note:** Chown command changes ownership

### **Changing Group Permissions on Files for Whole Directories:**

Step 1. Changing Group Permissions from Group 555 to 8 on the /mnt directory

Step 2. #find /mnt -group 555 | xargs chgrp 8

Step 3. Update /etc/group file if required

### **CHANGING OWNERSHIP RECURSIVELY FROM SPECIFIED PATHPOINT:**

# **find /mnt/32bitfs01/0/GRPF-KPC-4-15-03/n4copy -user 543 |xargs chown 683**

# **find /mnt/32bitfs01/0/GRPF-KPC-4-15-03/n4copy -group 110543 |xargs chgrp 4196**

**Note:** Command executes find of all users with UID of 110543 in the specified path, then changes ownership to 4196

### **FINDING UIDs/GIDs AND DUMPING TO FILE FROM A FILESYSTEM:**

# find /mnt/32bitfs01/0/GRPF-KPC-4-15-03/n4copy -user | -group 4196 >/home/nasadmin/tm/n4copy.users

### **Changing NFS Permissions on a File, Directory, Mountpoint using chmod:**

Examples show how to “chmod” a DataMover Mountpoint called ‘chmod’:

# **chmod 644 chmod**

```
# ls -la
drw-r--r-- 2 root    root      512 Jul 11 14:09 chmod
# chmod 777 chmod
# ls -la
drwxrwxrwx 2 root    root      512 Jul 11 14:09 chmod
```

### **NAS 5.5 COMMAND TO GRACEFULLY SHUTDOWN CONTROL STATION & DATA MOVERS:**

# **/nasmd/sbin/nas\_halt now**

\*\*\*\*\* WARNING! \*\*\*\*\*

You are about to HALT this Celerra including all of its Control Stations

and Data Movers. DATA will be UNAVAILABLE when the system is halted.

Note that this command does \*not\* halt the storage array.

ARE YOU SURE YOU WANT TO CONTINUE? [ yes or no ] : yes

Sending the halt signal to the Master Control Daemon...: Done

Jul 31 08:15:10 nx4-2 EMCServer: nas\_mcd: Check and halt other CS...: Done

-----abridged-----

Jul 31 08:16:23 nx4-2 EMCServer: nas\_mcd: Halting all servers...

**Note:** Do not use this command if there are established IP Replication sessions as it will abort the sessions. Bug prevents complete shutdown of NS600/500/700 platforms—see AR122586.

## **DATAMOVER NAMING:** (2) files that deal with server names

**\$server\_name ALL**

**\$more /nas/server/servers**

1:server\_2:1000:1:2::0:2:

2:server\_3:1000:1:3::0:3:

3:server\_4:1000:1:4::0:4:

4:server\_5:1000:4:5::0:5:

5:server\_6:1000:1:6::0:6:

6:server\_7:1000:1:7::0:7:

**\$more /nas/server/slot\_x/hostname**

hostname server\_2 selfid=2

## **CHANGING SERVER NAME:**

**\$server\_name server\_4 coyote**

**Note:** This command correctly updates the "hostname" and "servers" file as referenced above. Keep in mind that the output of this command is from DM memory and could be different from the /nas/server/servers file and /nas/server/slot\_x/hostname file. If the output of server\_name shows a discrepancy, yet both the “servers” and “hostname” files are correct, then the name is being retained in memory and will be cleaned up during next reboot.

## **CELERRA COLLECTION SCRIPT REQUIRED FOR ESCALATIONS:**

NAS 5.3.20 & 5.4.18 contain the EE collection script in [/nas/tools/collect\\_support\\_materials](#) (current version 2.9.3)

**/nas/var/emcsupport/support\_materials\_APM00081800233.090513\_1019.zip**

For NAS issues requiring Clariion backend analysis, run the [/nas/tools/.get\\_spcollect](#) script and engage Clariion Support—puts the zipped file in /nas/var/log directory

Zipping all spcollect zip files in one SPCOLLECT.zip file and putting it in the /nas/var/log directory...

adding: SL7E1081700022\_SPA\_2009-05-13\_17-24-18\_21009f\_data.zip (stored 0%)

adding: SL7E1081700022\_SPB\_2009-05-13\_17-27-00\_21004a\_data.zip (stored 0%)

**Note 1:** The /nas/tools/.get\_spcollect script is found with NAS 5.4., 5.5., 5.6

**Note 2:** With the advent of the automatic log collection & transfer utility, now running with NAS Versions 5.5.21.4 or 5.4.25.1 and higher, the collect\_support\_materials script is invoked automatically for Celerra CallHome events, as is the panic dump collection.

See emc135846. Beginning with NAS 5.5.27.5, the Log Collection, and the Transfer feature, can be enabled or disabled from Celerra Manager. When the “transfer” feature is enabled, and a CallHome event occurs that triggers an Auto Log Collection & Transfer, the following text will be written to the CSI “SR” case, which is generated by the CallHome XML header file:

<![CDATA[Check FTP server host 168.159.216.19, in remote\_directory /incoming/APM00055103645 for possible log/dump uploaded as the result of this event

## **COLLECT SUPPORT MATERIALS COLLECTION DIRECTORIES:**

/nas/var/log and /nas/var/dump for all code versions prior to NAS 5.6

/nas/var/emcsupport and /nas/var/dump for NAS 5.6 and higher

## **COLLECT SUPPORT MATERIALS CHANGE VERSION 2.8 NAS 5.5.27.x:**

**Note:** With GrandNapa and version 2.8 the /nas/tools/collect\_support\_materials script will change from a tar.gz to .zip extension, the idea being to use better compression during the log collection process. This change should not seriously impact support folks, as the Surge escalation tool and Log Parsing Tool should be able to handle collections with either extension, but it will require additional knowledge about how to use Linux zip, unzip, zipgrep, and zipinfo utilities.

### **Example of Old vs New:**

**# tar -tzvf support\_materials\_ML2809000146.070309\_1003.tar.gz |grep -v "drw" |wc -l**

1081 (files)

-rw-r--r-- 1 root root 3760171 Mar 9 10:14 support\_materials\_ML2809000146.070309\_1003.tar.gz

**# zipinfo -t support\_materials\_ML2809000146.070309\_0940.zip**

1097 files, 24894903 bytes uncompressed, 3654736 bytes compressed: 85.3%

-rw-r--r-- 1 root root 3876712 Mar 9 09:51 support\_materials\_ML2809000146.070309\_0940.zip

### **USING UNZIP TO LIST FILES:**

**# unzip -l support\_materials\_APM00025002080.070213\_1508.zip |more**

Archive: support\_materials\_APM00025002080.070213\_1508.zip

Length Date Time Name

```
-----  
0 01-31-07 10:50 nyip2.070213_1508/nas/log/webui/  
24246 02-02-07 15:34 nyip2.070213_1508/nas/log/webui/ ipc.log  
125 02-04-07 03:34 nyip2.070213_1508/nas/log/webui/alert_log  
42933 02-02-07 15:34 nyip2.070213_1508/nas/log/webui/apl_sched.log
```

### **USING UNZIP TO EXTRACT ALL FILES:**

**# unzip support\_materials\_APM00025002080.070214\_0034.zip**

### **USING UNZIP TO EXTRACT SINGLE FILES OR TO PIPE:**

**# unzip -p support\_materials\_APM00025002080.070213\_1508.zip nyip2.070213\_1508/nas/log/cmd\_log**  
**>/home/nasadmin/tm/cmd\_log.feb13**

### **USING ZIPGREP TO FIND STRINGS OR PATTERNS IN ZIPPED FILES:**

**# zipgrep reboot support\_materials\_APM00025002080.070213\_1613.zip**

nyip2.070213\_1613/cmd\_outputs/server\_2.log:2006-08-29 12:46:39: ADMIN: 4: Command succeeded: param kernel autoreboot=600

### **USING ZIPINFO TO LIST FILES OR SEE TOTALS:**

**# zipinfo -t support\_materials\_APM00025002080.070214\_0034.zip**

1027 files, 28022755 bytes uncompressed, 3254373 bytes compressed: 88.4%

**# zipinfo -1 support\_materials\_APM00025002080.070214\_0034.zip |tail** (lists out filenames only)

nyip2.070214\_0034/cmd\_outputs/server\_viruschk-audit

nyip2.070214\_0034/cmd\_outputs/webui\_use

nyip2.070214\_0034/cmd\_outputs/chkconfig-list

## **CHANGING DEFAULT LOCATION FOR COLLECT SUPPORT LOGS:**

1. Export TMPDIR first                   # export TMPDIR=/nas/var

2. Then run collection script:        # /nas/tools/collect\_support\_materials

## **UPLOAD & RUN COLLECT SUPPORT MATERIALS SCRIPT:**

1. Upload to Control Station using rz or FTP

2. Untar file:

**# tar -xvf collect\_support\_materials.tar**

collect\_support\_materials.sh

3. Run script:

**# ./collect\_support\_materials**

**Note:** Some systems may have the /nas/tools/collect\_support\_materials file without the .sh extension, and latest version also drops the sh extension. Current version is 2.7 (faster, more verbose output, added commands to gather info on various features).

Running material collection script revision 2.2.2.

Collecting /nas/log/\*, and /nas/jserver/logs

Collecting output from server\_log

-----output abridged-----

Now running material collection script for longer running commands.

Collecting complete nas dir listing

Collecting output from nas commands

Collecting output from other CS commands

Zipping archive for single file output

Material Collection File:

/tmp/support\_materials.051027\_1158.tar.gz has been generated.

4. The support\_materials collection script is generated with a current date and placed in /tmp

**support\_materials.051027\_1158.tar.gz**

5. Download and ftp to Engineering, along with appropriate escalation email and details

**Note 1:** Older versions of collect\_support\_materials script produced a couple of different output files

-rw-rw-r-- 1 nasadmin nasadmin 80584 Nov 2 07:50 support\_cmds.051102\_0744.tar.gz

-rw-rw-r-- 1 nasadmin nasadmin 3096195 Nov 2 07:45 support\_logs.051102\_0744.tar.gz

**Note 2:** To redirect the output of the collect\_support\_materials script to an alternate location on the CS, do the following:

\$ export TMPDIR=/nas/var/log

\$ echo \$TMPDIR

/nas/var/log

**Note 3:** A more current way of running the collect while working on other things

# **nohup /nas/tools/collect\_support\_materials -q -d /nas/var/dump &**

## **NAS UPGRADES & INSTALLATIONS SECTION**

### **NAS UPGRADE PRECAUTIONS & STEPS:**

Checkpoints, TimeFinder/FS, NAS Family, LUN Trespass, Flare Code version, all should be considered before doing an Upgrade  
**Note:** Completed upgrades are in /nas/log/upgrade\* log—interrupted or incomplete upgrades will have a log in /var/tmp

### **UPGRADING CELERRA USING USB FLASH DRIVE & ISO IMAGE ON NS40**

#### **USING USB FLASH DRIVE TO TRANSFER NAS ISO IMAGE TO CONTROL STATION:**

1. Plug USB Flash Drive into USB port on front side of Control Station [located underneath the serial port]

2. Run Fdisk to verify the discovered device type:

# **/sbin/fdisk -l**

/dev/sda1 1 992 999813+ 6 FAT16

# **df -h**

usbdevfs /proc/bus/usb type usbdevfs (rw)

3. Create mountpoint for USB device:

# **mkdir -p /mnt/usb**

4. Mount the Flash Drive to the Control Station

# **mount /dev/sda1 /mnt/usb**

5. Copy the ISO image from the Flash Drive to a partition with enough available space (about 550MB needed)

# **cd /mnt/usb**

# **cp emcnas5.5.31.6.iso /iso**

6. Unmount Flashdrive and physically remove from the Control Station

# **umount /mnt/usb**

#### **MOUNTING NAS ISO IMAGE ON LINUX CONTROL STATION:**

1. Mount the ISO image to the mountpoint defined above [/mnt/usb]

# **mount -t iso9660 -o loop /iso/emcnas5.5.31.6.iso /mnt/usb**

2. Perform NAS Upgrade from the ISO image

# **cd /mnt/usb/EMC/nas;/setup**

**Note:** After upgrading Linux kernel, system will require a reboot. Perform steps 1-2 again after the system has rebooted.

#### **RUNNING NAS PRE & POST-UPGRADE SCRIPT:**

1. Download & install latest Celerra Procedure Generator file from CCA website (CelerraProcGen.exe)

2. Locate upgrd-ck script.zip file in c:\Program Files\EMC\NASCPG\Application\Modules\SCRIPTS

3. Unzip to retrieve User Guide document and the tar.gz PreUpgrade check script

**upgrd-ck-document10-36.doc**

**upgrd-ckv10-36.tar.gz**

4. Upload tar.gz script to /home/nasadmin on the Control Station and unzip/untar

# **tar xvzf upgrd-ckv10-36.tar.gz**

upgrd-ckv10-36.bin

5. Execute pre-upgrade check

/home/nasadmin/upgrd-ckv10-36.bin -v 1 |-vb 1 (to backup files)

6. Log file is written to /home/nasadmin/pre-upgrade.rpt

**Note:** Check report and resolve any errors

7. Perform NAS Upgrade activity

8. Run post-upgrade check

# **/home/nasadmin/upgrd-ckv10-36.bin -v 2**

9. Check /home/nasadmin/post-upgrade.rpt

#### **RUN NAS 5.5 ENGINEERING PRE-UPGRADE CHECK SCRIPT (PUHC):**

# **/nas/tools/check\_nas\_upgrade -pre**

Check Version: 5.5.22.5028

Check Command: /nas/tools/check\_nas\_upgrade

Check Log : /nas/log/check\_nas\_upgrade.Sep-05-11:04:51.log

**Checks-----**

**Note:** Later editions of NAS 5.5 now call on the check\_nas\_upgrade script during the NAS Upgrade (./setup) and will fail the upgrade until errors are corrected.

**# ./setup**

Setting up system, please wait...

Model: NS600G

Upgrading From Version: 5.5.24-2

Mon Jan 29 12:54:45 EST 2007

Checking Upgrade Requirements

Check Version: 5.5.80.0

Check Command: /nbsnas/var/log/pkg\_5.5.80.0/check\_nas\_upgrade

Check Log : /nas/log/check\_nas\_upgrade.Jan-29-12:55:03.log

**Checks-----**

Control Station: Checking if enough free space exists ns..... **Fail**

**# cat /var/tmp/upgrade.log.Mon\_Jan\_29\_12:59:19\_EST\_2007** [Logs output from check\_nas\_upgrade]

**MORE ON UPGRADE CHECKS:****/nas/tools/upgrade\_check\_tools**

# ls -la

```
-rwxr-xr-x 1 root root 103120 Apr 15 22:09 be.pm
-rwxr-xr-x 1 root root 1443 Apr 15 22:09 be.requirements →Backend Flare vs NAS requirements
-rwxr-xr-x 1 root root 1523 Apr 15 22:09 check_full_dump.pl
-rwxr-xr-x 1 root root 870 Apr 15 22:09 check_nbs_devices.pl
-rwxr-xr-x 1 root root 19870 Apr 15 22:09 checks.pm
```

**-rwxr-xr-x 1 root root 17396 Apr 15 22:09 checks.txt →Overall checks**

```
-rwxr-xr-x 1 root root 82107 Apr 15 22:09 cmd.pm
-rwxr-xr-x 1 root root 165863 Apr 15 22:09 cs.pm
-rwxr-xr-x 1 root root 142806 Apr 15 22:09 dm.pm
-rwxr-xr-x 1 root root 176200 Apr 15 22:09 messages.txt
-rwxr-xr-x 1 root root 39469 Apr 15 22:09 msg.pm
-rwxr-xr-x 1 root root 9124 Apr 15 22:09 severities.txt
drwxrwxr-x 2 root root 4096 Apr 16 14:44 Text
-rwxr-xr-x 1 root root 42811 Apr 15 22:09 util.pm
```

**Note:** The “checks.txt” file is the master checklist used during the check\_nas\_upgrade check operation

**NAS 5.6 UPGRADES:**

CD-ROM package contains a new upgrade command:

**/EMC/nas/install\_mgr –mode upgrade**

Lab Workaround for Unsupported Hardware NAS 5.6:

**/EMC/nas/install\_mgr –mode upgrade –allow\_upgrade\_on\_obsolete\_dm** (507 & 510 DM's)

**Note:** Upgrades from 5.4 and 5.5 to 5.6 requires a CD-ROM. Additionally, the CFS platform require the use of a boot floppy

**EXAMPLE OF UPGRADE TASKS FOR 5.6.43.5 (seen after 2<sup>nd</sup> .install mgr –m upgrade):**

====Time Estimate for All Tasks=====

|    |                                  | Estimated | Status    |
|----|----------------------------------|-----------|-----------|
|    | Time(Minutes)                    |           |           |
| 1  | Copyright Information            | 0         | completed |
| 2  | Set up environment               | 0         | completed |
| 3  | Set Automatic Boot Manager       | 0         | completed |
| 4  | Run Pre Upgrade Health Check     | 0         | completed |
| 5  | Disable nas version command      | 0         | completed |
| 6  | Install or upgrade linux RPMs    | 0         | completed |
| 7  | Set up preinstall                | 0         | completed |
| 8  | Disable NAS service              | 0         | completed |
| 9  | Disable ANACRON service          | 0         | completed |
| 10 | Disable CROND service            | 0         | completed |
| 11 | Disable YPBIND service           | 0         | completed |
| 12 | Stop NAS service-                | 0         | completed |
| 13 | Reboot Control Station if needed | 0         | completed |
| 14 | Do preinstall misc               | 2         |           |

|    |   |   |
|----|---|---|
| 15 | Set up lvm device map                       | 0 |
| 16 | Set up local lvm partition                  | 0 |
| 17 | Sync cache to backend                       | 2 |
| 18 | Upgrade nas lvm                             | 0 |
| 19 | Upgrade NAS RPM                             | 8 |
| 20 | Set up etc hosts file                       | 0 |
| 21 | Do miscellaneous steps 1                    | 0 |
| 22 | Set up system files                         | 1 |
| 23 | Set up USB modem ports                      | 0 |
| 24 | Set up Control Station to reboot on panic   | 0 |
| 25 | Build administrator profile                 | 0 |
| 26 | Do miscellaneous steps 2                    | 0 |
| 27 | Install licenses                            | 0 |
| 28 | Install third party licenses                | 0 |
| 29 | Update NasStorageAPI                        | 2 |
| 30 | Install celerra srvc enabler                | 0 |
| 31 | Configure CLARiiONs                         | 1 |
| 32 | Set up DOS area                             | 0 |
| 33 | Upgrade hardware                            | 0 |
| 34 | Do miscellaneous steps 3                    | 2 |
| 35 | Set up nas checkup                          | 0 |
| 36 | Initialize banner                           | 0 |
| 37 | Set up permissions                          | 0 |
| 38 | Set up ConnectHome                          | 0 |
| 39 | Generate ssl certificates                   | 0 |
| 40 | Upgrade replication                         | 0 |
| 41 | Backup Data Mover logs                      | 0 |
| 42 | Update workpart                             | 2 |
| 43 | Set up Control Station                      | 0 |
| 44 | Upgrade Data Movers*                        | 7 |
| 45 | Translate file systems to DIR3              | 0 |
| 46 | Set partition types                         | 0 |
| 47 | Set up admin roles                          | 0 |
| 48 | Set up internationalization                 | 1 |
| 49 | Set up local storage                        | 1 |
| 50 | Enable ANACRON service                      | 0 |
| 51 | Enable CROND service                        | 0 |
| 52 | Enable YPBIND service                       | 0 |
| 53 | Set password expiration                     | 0 |
| 54 | Set password quality policy                 | 0 |
| 55 | Update user group databases                 | 0 |
| 56 | Create IDE cache                            | 3 |
| 57 | Enable NAS service                          | 0 |
| 58 | Start NAS service+                          | 4 |
| 59 | Set up backend monitor tool                 | 0 |
| 60 | Perform inventory                           | 0 |
| 61 | Run final tasks                             | 0 |
| 62 | Wait for NAS service to come up             | 1 |
| 63 | Retry upgrade for failed data movers if any | 1 |
| 64 | Stop artificial heartbeat                   | 0 |
| 65 | Restart standby Control Station             | 0 |
| 66 | Initial emailuser                           | 0 |
| 67 | Run nas checkup                             | 5 |
| 68 | Enable nas version command                  | 0 |
| 69 | Reset Automatic Boot Manager                | 0 |

- + NAS service will be unavailable between these

\* File service will be unavailable

—  
—  
—

==== Estimate

Estimated Time when Data Movers will be reset: 14:34

Estimated Time when NAS service will be restarted: 14:46

Estimated Time when upgrade will be complete: 14:57

=====

## **REMOTE UPGRADES NAS 5.6:**

RCM group performs many Celerra upgrades via remote connection

Primary utility on Linux Control Station is “screen”

# **screen** → Starts the screen terminal, after which, if you lost your connection and had to reconnect, you would use # screen -r to reconnect to the screen session

→ctrl + a, then c to create new screen window

→ctrl + a, then n to switch to next screen window (if multiple sessions)

# exit →Exits from screen program

## **NAS 5.6 INSTALL ISSUE WITH FLEET BACKEND:**

→Starting with NAS 5.6.39.5, Celerra Gateways support connectivity to Fleet Flare 28 arrays, but Celerra does not support IPv6 settings on the SP's until 5.6.43 storageAPI. AR129282 documents an issue where the Celerra cannot perform the backend discovery because the symapi cannot handle IPv6 network information--the SPs are set to IPv6 mode ‘automatic’ (DIMS292226). Changing the settings to “disabled” resolved the install issue:

**# /nas/sbin/navicli -h A\_APM00083701533 networkadmin -get -ipv6**

Storage Processor: SP A

Storage Processor IPv6 Mode: **Automatic**

Storage Processor IPv6 Address:

Storage Processor IPv6 Global Prefix:

Storage Processor IPv6 Link-local Address: fe80::260:1600:3b20:b9e

**# /nas/sbin/navicli -h 10.241.168.182 networkadmin -get -all**

Storage Processor: SP A

Storage Processor Network Name: SPA

Storage Processor IP Address: 10.241.168.182

Storage Processor Subnet Mask: 255.255.255.0

Storage Processor Gateway Address: 10.241.168.128

Storage Processor IPv6 Mode: Disabled

Management Port Settings:

Link Status: Link-Up

Current Speed: 100Mbps/full duplex

Requested Speed: Auto

Auto-Negotiate: YES

Capable Speeds: 1000Mbps half/full duplex

10Mbps half/full duplex

100Mbps half/full duplex

Auto

**# /nas/sbin/navicli -h A\_APM00083701533 networkadmin networkadmin -set -ipv6 -disable**

**IPv6 Settings on SPs changed to Disabled:**

**# /nas/symcli/bin/symcfg discover -clarion**

This operation may take up to a few minutes. Please be patient...

**# /nas/bin/server\_devconfig server\_2 -c -s -a**

server\_2 : done

**# /nas/sbin/navicli -h A\_APM00083701533 networkadmin -get -ipv6**

Storage Processor: SP A

Storage Processor IPv6 Mode: **Disabled**

## **NAS UPGRADE ISSUES RELATED TO LARGE OBJECT COUNT OR HIGH CPU ON CLARIION:**

(aka, changing CLARiiON environmental values for slow polling systems)

See emc175499. Typically, when navi commands take a long time to respond, various functions such as NAS upgrades are subject to failure because of timeouts. Generally, this problem can be caused by high CPU activity on the array, or more recently, very high object counts on the array (>than 1000 objects). Traditional workarounds have included the –np switch, along with a script developed by Engineering to make sure all commands going to the backend used the “no poll” function.

**Two-part workaround:**

1. Connect to the array's /setup program via web browser and change the Communication Timeout default from 180 to 300 secs:

### **“Set Update Parameters”>Communication Timeout: 300**

2. Restart the Management Server on the array when prompted

3. From the same Control Station session where the upgrade will be run from, issue the following:

# **export NAVI\_TIMEOUT=300**

# echo \$NAVI\_TIMEOUT

300

**Note:** exporting the navi timeout value is only valid for the current shell session

4. Make sure a copy of classic\_navicli is present in /nasmcd/sbin, with 777 permissions

## **NAS 5.5 GATEWAY INSTALL USING KSNAS.CFG FILE (auto zone & storagegroup creation):**

### **Overview:**

→Need Procedure Generator to generate tailored NAS Installation document that produces a usable ksnas.cfg file

→Need NAS CD

→Need bootable system floppy disk to start install, along with ksnas.cfg file

**Note:** Obtain a copy of the latest Celerra Procedure Generator, install on desktop system, generate checklist of required information for “NAS Installation”, produce the NAS Installation Procedure, have NAS CD\_ROM available, create system floppy disk and copy ksnas.cfg file to the floppy disk for use during installation.

### **USING CELERRA PROCEDURE GENERATOR Version 1.2.3:**

1. Download executable file and install using defaults: doubleclick “NASProcGen.exe”

2. Open “PG” shortcut on desktop>select forward ‘hand’ icon and enter customer information

3. Choose either Hardware Change, Software Change, or NAS Installation

4. Select correct Celerra Model, Type, Data Movers, Factory racked, Primary Backend Type, file name, enter

**Note:** Above produces a questionnaire that can be used as a template to answer the questions needed by ksnas.cfg script

5. Re-open the file name that was produced in Step 4, select forward button, and generate NAS Installation document

### **PERFORM NAS INSTALLATION: NAS 5.5.22.5028**

1. Insert boot floppy, CD\_ROM, connect to CS with Serial cable Hyperterminal Session using 19200 defaults, COM1 or COM2, and Hardware FC ‘None’

2. Reboot Control Station and enter following at prompt:

### **boot: serialkickstart**

Select destructive install (no other option is really available), CS reboots and goes through Preinstallation checks

Package Installation of linux occurs and reboots using Grub loader

Starting EMC NAS Factory Installation

Detecting movers in cabinet: 2

Checking for existing System RG/LUN: Shared backend detected

System Raid Group will be setup on disks 0\_2\_0 0\_2\_1 0\_2\_2 0\_2\_3 0\_2\_4 (Bus 0, Enclosure 2, Disks 0-4)

RAID Group ID is 8

ALU IDs of System LUNs: 29 30 31 33 34 35

Waiting for LUNs to finish binding...

Rebooting Data Movers...

Mover in slot\_2 →Stages: NO STATE | DART LOADED | GENERIC SLOT CONFIG | DART CONFIGURING

Setting up Fabric Zoning...

Collecting backend WWNs

Detecting Data Movers

Collecting Data Mover WWNs

Creating Zones...

Setting up Storage Groups on CLARiiON Backend...

Creating Storage Group: Celerra\_nyip2

3. NAS Installation proceeds

Starting EMC NAS installation

Creating the NAS Partition on /dev/nda1 with fdisk

Building the IDE NAS filesystem on /dev/hda5 with mke2fs

(Creates VAR & DOS Partitions next....)

## **FORCING POST UPGRADE NAS 5.6.45:**

→Generally speaking, POST & BIOS are flashed during the bootup process whenever a newer firmware version is detected.

Typically, firmware updates are provided in the NAS image.

→To force a system to update POST & BIOS

1. Edit boot.cfg to force post upgrade

```
# vi /nas/dos/slot_2/boot.cfg
flashupg post=upgrade bios=upgrade
Edit above line to read:
flashupg post=force bios=upgrade
2. Reboot server to force firmware update:
# /nasmcd/sbin/t2tty -C 2 "shutdown"
3. Verify:
# cat /nas/log/data_mover_resume.server_2.xml|grep "VERSION"
    <VERSION_INFORMATION BIOS_VERSION="03.80" POST_VERSION="Rev. 01.59" />
```

## **INSTALLING NAS WITH LUNS 0 & 1 AT 4GB vs. 11GB STANDARD: 5.5.28.1**

**Note:** Later code versions now enforce installations of Control Luns 0 & 1 at 11GB each, but for lab testing purposes, you can work around this and install at any desired size

1. Perform install as usual, with CD-ROM and floppy
2. After installation of Linux on CS, and right after configuring the CS interfaces and hostname, ssh into the CS
3. From ssh session, # cd /tftpboot/setup\_backend

**Note:** If this directory doesn't exist, extract from /nas/tools

```
# tar -xvf /nas/tools/tftpboot.tar.gz
```

4. Edit the system.pm file and change LUNs 0 & 1 from 11GB to desired LUN size:

```
# vi /tftpboot/setup_backend/system.pm
```

```
49 %systemLuns = ( $SYSTEM_LUN_ID_0, 11,
50     $SYSTEM_LUN_ID_1, 11,
51     $SYSTEM_LUN_ID_2, 2,
52     $SYSTEM_LUN_ID_3, 2,
53     $SYSTEM_LUN_ID_4, 2,
54     $SYSTEM_LUN_ID_5, 2);
```

5. Edit S95nas script in the following locations (same edit), also changing the 11GB size to 4 or 2GB:

```
# vi /etc/rc3.d/S95nas
```

```
1606 local GB_IN_BLOCKS=$[1024*1024*2]
1607 local LrgLunMinSiz=$[11*GB_IN_BLOCKS] # Minsize for Luns 0 and 1
1608 local SmlLunMinSiz=$[2*GB_IN_BLOCKS] # Minsize for Luns 2 3 4
```

```
# vi /mnt/source/EMC/nas/S95nas
```

6. Exit ssh session and continue with NAS Installation with serial session

## **NS40/NS80 INTEGRATED ADD ENCLOSURE or ADD BLADE PROCEDURES:**

### **NS40 INTEGRATED ADD BLADE PROCEDURE:**

An NS40 exists only in a Single or Dual Blade configuration. If the system is an NS41 Integrated, then a 2<sup>nd</sup> DM blade can be added. The actual steps for adding a Blade to an existing system are fairly simple. Internal Network switch cabling needs to be reconfigured for dual blades; Copper fibre channel cables need to be made between DM3 and the SPs; Celerra's Internal Network needs to be upgraded from a single blade configuration to a dual blade configuration; The new Blade contains its own Internal Management Switch (unlike the NSX/NS80 blades), which needs to be initialized so that the Celerra system and database are able to recognize the new Blade; And finally, a setup\_slot is required to setup the new Blade hardware on the Celerra.

#### **1. Internal Network Cabling Changes:**

Ethernet cable from lower management port DM3 to port labeled 2 on Control Station—eth2 (upper middle port)

Ethernet cable from upper management port DM3 to lower management port on SPB

#### **2. Fibre Channel Cabling Changes:**

DM blade 3 port BE 0 to SP A port 1 Fibre using copper SFP-to-SFP cable

DM blade 3 port BE 1 to SP B port 1 Fibre using copper SFP-to-SFP cable

#### **3. Upgrade the Internal Network:**

```
# /nasmcd/sbin/upgrade_to_dual_intnet
```

SUCCESS: Internal network upgrade succeeded. Please verify the sanity of  
the system and delete the directory /tmp/upgrade\_to\_dual\_net.23674.

**Note:** Above command stops & restarts NAS services—wait until nas\_mcd and boxmonitor processes are restored before continuing

#### **4. Verify that both internal networks are available and that Control Station has been reconfigured:**

```
# ping server_2
```

PING server\_2 (192.168.1.2) from 192.168.1.100 : 56(84) bytes of data.

64 bytes from server\_2 (192.168.1.2): icmp\_seq=0 ttl=255 time=154 usec

```
# ping server_2b
```

PING server\_2b (192.168.2.2) from 192.168.2.100 : 56(84) bytes of data.

64 bytes from server\_2b (192.168.2.2): icmp\_seq=0 ttl=255 time=184 usec

**# /sbin/ifconfig**

```
eth0    Link encap:Ethernet HWaddr 00:00:F0:9F:B3:74
        inet addr:192.168.1.100 Bcast:192.168.1.255 Mask:255.255.255.0
              UP BROADCAST NOTRAILERS RUNNING MULTICAST MTU:1500 Metric:1
eth2    Link encap:Ethernet HWaddr 00:00:F0:9F:53:07
        inet addr:192.168.2.100 Bcast:192.168.2.255 Mask:255.255.255.0
              UP BROADCAST NOTRAILERS RUNNING MULTICAST MTU:1500 Metric:1
```

**5. Remove the following directory from /tmp:**

**# rm -Rf upgrade\_to\_dual\_net.23674**

**6. Export NAS DB environmental variable:**

**# export NAS\_DB=/nas**

**7. Add new Blade's Mgmt Switch:**

**# /nas/sbin/setup\_enclosure -addMgmtswitches**

Discovered 1 Mgmt switch on Primary network and 1 Mgmt switch on Secondary network

Discover the following Mgmt switch ...

MgmtSwitch-B IP address = 192.168.2.50

MgmtSwitch-B MAC address = 00:60:16:0b:1b:a5

flashFirmware Completed

addMgmtswitches Completed

**Note:** If this fails with 0 Mgmtswitches in Secondary subnet B, run the t2net\_test command from the good switch on slot\_2 to reset the switch on slot\_3, which represents the Secondary subnet

**# /nas/sbin/t2net\_test -TyphPeercontrol**

Usage: t2net\_test -TyphPeercontrol {ipaddr operand}

where:- ipaddr = Mgmtswitch IP address

operand = 0(Peer Disable), 1(Peer Enable), 2(Peer Reset)

**# /nas/sbin/t2net\_test -TyphPeercontrol 128.221.252.50 2**

**Note:** Operand of 2 is used to reset the peer switch

**8. Verify that Secondary Switch was successfully added:**

**# /nas/sbin/setup\_enclosure -probeSystem**

Executing -probeSystem option

Discovering Primary subnet (A) ... OK

Enclosure ID = 0, BackplaneID 0 from IP 192.168.1.50

MgmtSwitch-A MAC = 00:60:16:0b:1a:ca, Peer MAC = 00:60:16:0b:1b:a5

F/W Version: Mgmtswitch=01.82, Bootblock=01.77

Discovering Secondary subnet (B) ... OK

Enclosure ID = 0, BackplaneID 1 from IP 192.168.2.50

MgmtSwitch-B MAC = 00:60:16:0b:1b:a5, Peer MAC = 00:60:16:0b:1a:ca

F/W Version: Mgmtswitch=01.82, Bootblock=01.77

**# /nas/sbin/enclosure\_status -e 0 -v**

DEVICE A DEVICE B

-----PRESENCE-----

00 60 16 0B 1A CA MAC 00 60 16 0B 1B A5

Scorpion Hardware Platform Scorpion

Present Peer Compute Blade Present

Present Compute Blade Present

Powered On Compute Blade Powered On

Present Power Supply A Present

Present Power Supply B Present (output abridged)

**9. If the switch was not properly added, try the following commands:**

**# /nas/sbin/setup\_enclosure -resetMgmtswitches or -initSystem force**

**10. Add the new Blade to the Celerra database & verify:**

**# /nas/sbin/setup\_slot -i 3**

**# nas\_server -l**

| id | type | acl  | slot | groupID | state    | name |
|----|------|------|------|---------|----------|------|
| 1  | 1    | 1000 | 2    | 0       | server_2 |      |
| 2  | 1    | 1000 | 3    | 0       | server_3 |      |

**Note:** With the later Celerra Integrated Models that use AccessLogix and Storage Groups (aka, NS20/NS40/NX4), you would use the /nas/sbin/setup\_slot -i -g 3 command to add the new host initiator records to the Storage Group

## **NS80 INTEGRATED ADD ENCLOSURE/BLADE PROCEDURE:**

An NS80 exists only in a 2, 3, or 4 Blade configuration. If the NS80 Integrated has only (2) Blades, when adding a 3<sup>rd</sup> Blade, the system would require Internal Network recabling, the addition of Fibre Optic fibre channel cables between the new Blade and the SPs, and a separate step to first add Enclosure 1 to the Celerra, followed by a step to add the 3<sup>rd</sup> or 4<sup>th</sup> blade, which aside from cabling requirements, requires a setup\_slot to recognize the new hardware on the Celerra. It is actually possible to have a 2<sup>nd</sup> Enclosure on an NSX or NS80 without a Blade, yet have the Internal Network cables cabled from the Enclosure Mgmt Switch A & B to the SPs A or B switch LAN port.

### **1.) Internal Network Cabling Changes:**

Example given is for going from a 2-blade system to a 3 to 4-blade system

--Move LAN cable from Port 4 on DM Encl 0 Switch A (the other end goes to the upper LAN port on SPA, Switch A, and is unchanged), to Port 4 on DM Encl 1, Switch A

--Move LAN cable from Port 4 on DM Encl 1, Switch B (the other end goes to the upper LAN port on SPB, Switch B, and is unchanged), to Port 4 on DM Encl 1, Switch B

--Connect new LAN cable from Port 0 on DM Encl 0, Switch A, to Port 3 on DM Encl 1, Switch A (daisychains Encl 0 & 1 together)

--Connect new LAN cable from Port 0 on DM Encl 0, Switch B, to Port 3 on DM Encl 1, Switch B (daisychains Encl 0 & 1 together)

### **2. Fibre Channel Cabling Changes:**

--Connect Port BE0 on DM 4 on Encl 1 to Port 2 on SPA using LC-to-LC fibre optic cables and SFP connectors

--Connect Port BE1 on DM 4 on Encl 1 to Port 2 on SPB using LC-to-LC fibre optic cables and SFP connectors

--Connect Port BE0 on DM 5 on Encl 1 to Port 3 on SPA using LC-to-LC fibre optic cables and SFP connectors

--Connect Port BE1 on DM 4 on Encl 1 to Port 3 on SPB using LC-to-LC fibre optic cables and SFP connectors

### **3. Add new Enclosure to System:**

/nasmcd/sbin/setup\_enclosure –addEnclosure

### **4. Add the new Blade to the Celerra database:**

/nas/sbin/setup\_slot –i 4

## ***NS80/NS40 INTEGRATED CONNECTION METHODS, INSTALLATIONS, INTERNAL NETWORK SETUP, & NDU UPGRADE PROCEDURES FOR SINGLE(NS41-only) & DUAL DM SYSTEMS***

### **NS40 NAS 5.5 INTEGRATED MODELS:**

#### **NS40 NAS Head with 4 Copper Ports and CLARiiON CX3-40 Backend:**

→Data Mover RoHS TLA Part # 100-561-998 for 1 or 2 blades

→Factory Sales Order Model # NS42C NS4xx NAS Head 100-520-605

→Field Install Sales Order Model # NS42C-A-FD NS4xx NAS Head 100-520-605

→BIOS version 3.48, Model: Sledgehammer: NAS, POST 01.30

→PROM information SAN/NAS Type (2): NAS, 1U-NAS 2DM CU CHASSIS, NS3-40 CU NAS 2

→Resume information for SPs:

# /nas/sbin/navicli -h 192.168.1.200 getresume |grep CX

Assembly Name: CX3-40 SAN/AUX

**Note:** As with the NSX models, there are no serial connections between CS & Data Mover blades (and no ctaptty logs either), meaning that the only way to observe or troubleshoot boot issues is via a local workstation connected directly to the maintenance Serial port of the Data Mover.

### **For NS40/NS80 Integrated Celerras, there are two possible configurations:**

#### **Factory-Installed systems:**

Hardware is pre-loaded with NAS software using factory-racked 40U EMC hardware. Factory-installed units are cabled into the environment, and after powering on, the Control Station can be configured with its External network address by using the InitWizard application on a Windows system that is plugged into the same network that the Control Station is cabled up to.

#### **INITWIZARD:**

The InitWizard is obtained from the Apps & Tools CD. Click InitWizard.exe and the Control Station Initialization Wizard appears. Click on Search tab and it searches the subnet via ARP broadcasts for any unconfigured Control Stations. Upon finding a CS, its MAC address is displayed. Click the “Configure” button and enter Hostname, IP Address, Netmask, Gateway, Finish. “Configuration completed successfully. Now you can access the Control Station over the network. Do you want to open a web browser and continue configuring the Control Station?” If you answer “Yes”, a browser bar opens—input the Control Station IP address. The wizard then runs through a cable check to validate the cabling—“Cable Check Passed”. Celerra Manager then launches. Configure Control Station>Call Home and then use the Wizards menu in Celerra Manager for further configuration, such as “Set Up Celerra”.

#### **CONTROL STATION INITWIZARD:** aka, iwd

When required, the IWD daemon runs on the Control Station to support the InitWizard and the new CSA wizard.

# ps -eaf|grep -i iwd

/nasmcd/sbin/iwd 040 R root 3429 3428 0 76 0 - 348 - 11:59 pts/0 00:00:00

/sbin/service nas\_ipinit stop | start | restart →Calls the /nasmcd/sbin/iwd executable

/sbin/chkconfig nas\_ipinit –list [is on by default]

**/var/log/messages** will log information regarding InitWizard starting and stopping

Jun 19 11:59:20 sludge1 iwd: InitWizard: the daemon has started.

Jun 19 11:59:20 sludge1 iwd: InitWizard: the daemon has terminated.

**Field-installed systems:**

These systems require assembly in the customer's racks, are designated with a model number, such as NS42C-A-FD, and require a completely fresh install. The Clariion SPs should already be configured with the correct SAN/AUX personality and have the correct internal IP address scheme and hostnames assigned, though the Setup Guide requests that the SP personality be verified during reboot, yet does not have User checking to ensure that SP IP Addresses are correctly set—the Celerra Procedure Generator will have the User verify each of the SP's Hostnames, IP & Gateway addresses.

**HEALTHCHECKING NS40/NS80 SYSTEM:**

**/nas/tools/auto\_checkup** → Version of PAHC tool that runs every 2 weeks per Cron entry in /nas/site/cron.d/nas\_sys

**Note:** PAHC checkup now runs once a week on Sunday

**/nas/bin/nas\_checkup** → CLI version of PAHC tool run On-demand from /nas/bin

**/nas/tools/dbchk -wvxpVs** [Switches available with latest NAS 5.5 code—not all switches may work on prior code]

**/nas/tools/dbchk -pvwx**

**/nas/tools/check\_clariion** (checks SP utilization and Disk IOPs on CLARiiON arrays)

**/nas/tools/check\_nas\_upgrade** [-pre -up -pro Upgrade check script]

**/nas/bin/nas\_storage -check -all**

**/nas/sbin/enclosure\_status -e 0 -v** → Overall system checkup of hardware components

**/nas/sbin/setup\_enclosure -checkCable**

**/nas/sbin/setup\_enclosure -checkSystem**

**/nas/sbin/setup\_enclosure -probeSystem**

**/nas/sbin/setup\_enclosure -readConfig**

**/nas/sbin/t2net\_test**

**/nas/sbin/t2net\_test -SetResumeProm**

**/nas/sbin/t2net\_test -GetResumeProm 2 38 150 30**

t2net\_test: Return values:

0000: 0x34 0x20 0x50 0x4f 0x52 0x54 0x20 0x46 0x49 0x42 4 PORT FIB

0010: 0x52 0x45 0x20 0x49 0x4f 0x20 0x4d 0x4f 0x44 0x55 RE IO MODU

0020: 0x4c 0x45 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 LE .....

**/nas/sbin/t2led**

**/nas/sbin/t2vpd -e 0**

**/nas/sbin/t2485net**

**Note:** enclosure\_status & setup\_enclosure commands are valid only with NSX, NS40, NS80 Celerras

**OUTPUT EXAMPLES:**

**# /nas/sbin/enclosure\_status -v -e 0**

DEVICE A            DEVICE B

-----PRESENCE-----

-----BROADCOM STATUS-----

-----ENCLOSURE ALARMS-----

-----BROADCOM ALARMS-----

-----RESUME CSUM ERRORS-----

-----COLDFIRE ALARMS-----

-----STATUS CONDITIONS-----

-----FRU STATUS-----

-----SYSTEM VARIABLES-----

**# /nas/sbin/setup\_enclosure -checkSystem**

Executing -checkSystem option

Checksum verification on ENCL\_DB & DHCPD\_CFG ... OK

Verify ENCL\_DB generation version ... OK

Current Enclosure database (ENCL\_DB) info ...

Enclosure ID# 0:

MgmtSwitch-A IP address = 192.168.1.50

MgmtSwitch-A MAC address = 00:60:16:0b:1a:ca

F/W Version: Mgmtswitch=01.82, Bootblock=01.77

-----abridged-----

Performing Detail System check ...

System discovery on both subnets ... OK

Detail System check OK

checkSystem Completed

### # /nas/sbin/setup\_enclosure -checkCable

Executing -checkCable option

System discovery on both subnets ... OK

Pair up all discovered Mgmtswitches ... OK

Verify no cross-connected cabling error ... OK

Collect Enclosure cabling topology .... OK

Examine Enclosure cabling ...

Cabling of Enclosure at position 0 ... OK

checkCable Completed

### # /nas/sbin/setup\_enclosure -probeSystem

Executing -probeSystem option

Discovering Primary subnet (A) ... OK

Enclosure ID = 0, BackplaneID 0 from IP 192.168.1.50

MgmtSwitch-A MAC = 00:60:16:0b:1a:ca, Peer MAC = 00:60:16:0b:1b:a5

F/W Version: Mgmtswitch=01.82, Bootblock=01.77

Discovering Secondary subnet (B) ... OK

Enclosure ID = 0, BackplaneID 1 from IP 192.168.2.50

MgmtSwitch-B MAC = 00:60:16:0b:1b:a5, Peer MAC = 00:60:16:0b:1a:ca

F/W Version: Mgmtswitch=01.82, Bootblock=01.77

### # /nas/sbin/setup\_enclosure -readConfig

Executing -readConfig option

Current Enclosure database (ENCL\_DB) info ...

Enclosure ID# 0:

MgmtSwitch-A IP address = 192.168.1.50

MgmtSwitch-A MAC address = 00:60:16:0b:1a:ca

F/W Version: Mgmtswitch=01.82, Bootblock=01.77

MgmtSwitch-B IP address = 192.168.2.50

MgmtSwitch-B MAC address = 00:60:16:0b:1b:a5

F/W Version: Mgmtswitch=01.82, Bootblock=01.77

## **INTERNAL MANAGEMENT SWITCH ISSUES:**

For dual blade systems, the Primary Management Switch A is cabled to Data Mover 2, and the Secondary Management Switch B is cabled to Data Mover 3. There are situations where the Internal Network cabling is incorrect, cables are bad, cables are removed, etc., that could lead to Mgmt Switch failures, as presented in the following example:

### **/nas/log/sys\_log**

May 2 10:16:33 2007 BoxMonitor:3:526 enclosure 0 failed to respond to ping at management switch A

May 2 10:17:41 2007 BoxMonitor:2:43 Slot 2 failed to respond to ping at server\_2

May 2 10:17:44 2007 BoxMonitor:2:43 Slot 3 failed to respond to ping at server\_3

May 2 10:20:50 2007 BoxMonitor:3:526 enclosure 0 failed to respond to ping at management switch B

May 2 10:21:55 2007 BoxMonitor:2:43 Slot 3 failed to respond to ping at server\_3b

May 2 10:21:55 2007 BoxMonitor:2:43 Slot 2 failed to respond to ping at server\_2b

**Note:** Log example shows that at one time, Mgmt Switch A & B were down, and the corresponding failure to ping the DM on the respective Primary or Secondary Network is logged

### **Manually Disabling Primary or Secondary Mgmt Switch Networks:**

**Caution:** The following commands should be used with extreme care as the Primary or Secondary Internal Network will be shutdown. Commands can be used for testing, etc.

### **# /nas/sbin/t2net\_test -TyphPeercontrol 192.168.1.50 0**

t2net\_test: Command succeed

**Note:** Disables the Secondary Management Switch & network—runs from switch in slot\_2 against its peer in slot\_3, and 0 means to disable

### **# /nas/sbin/t2net\_test -TyphPeercontrol**

Usage: t2net\_test -TyphPeercontrol {ipaddr operand}

where:- ipaddr = Mgmtswitch IP address

operand = 0(Peer Disable), 1(Peer Enable), 2(Peer Reset)

### **# /nas/sbin/t2net\_test -TyphPeercontrol 192.168.2.50 0**

t2net\_test: Command succeed

**Note:** This command disables the Primary Management Switch & network

### **# /nas/sbin/setup\_enclosure -probeSystem**

**Note:** Use the probeSystem command to see if the Primary or Secondary Mgmt Network is down, as in following example:

### **# /nas/sbin/setup\_enclosure -probeSystem**

Executing -probeSystem option

Discovering Primary subnet (A) ... OK

No Mgmtswitch is found on Primary subnet

Discovering Secondary subnet (B) ... OK

Enclosure ID = 0, BackplaneID 1 from IP 192.168.2.50

MgmtSwitch-B MAC = 00:60:16:0b:1b:a5, Peer MAC = 00:60:16:0b:1a:ca

F/W Version: Mgmtswitch=01.82, Bootblock=01.77

ZERO Mgmtswitch detection may due to one or more of the following cases:

- Miswired cabling on Mgmtswitches that causes switch looping

- Loose/bad cable between Enclosure ID# 0 and CS

- Bad Ethernet port on CS

- Bad Mgmtswitch

**Note:** The above causes of a Mgmt Switch failure should be verified prior to attempts to re-initialize the Switch. The failure can also be seen when adding a 2<sup>nd</sup> Blade to a single DM system, etc.

### **Re-initializing a failed Mgmt Switch after Blade Hardware Addition or Replacement, or Cable Replacement, etc:**

### **# /nas/sbin/setup\_enclosure -resetMgmtswitches**

**Note:** Command reboots Mgmt Switches in an attempt to recover

### **# /nas/sbin/setup\_enclosure -initSystem force**

**Note:** This command appears to rebuild the Enclosure and DHCP database, reboots both Switches, flashes the switch firmware (if needed), checks cabling, and does a discovery of the Mgmt Switches. Start with the –resetMgmtswitches first, then try –initSystem force if that fails.

Enclosure Status Commands:

### **# /nas/sbin/enclosure\_status -e 0 -v |grep -i compute**

Present Peer Compute Blade Present

Present Compute Blade Present

### **# /nas/sbin/enclosure\_status -e 0 -v**

DEVICE A DEVICE B

-----PRESENCE-----

00 60 16 0B 1A CA MAC ---

Scorpion Hardware Platform --- [output abridged]

**Note:** enclosure\_status command shows component failures in the system, but in the above example, does not show any components at all under the “DEVICE B” column, which also means that Mgmt Switch B (Secondary Network) is down or not initialized.

## **CHANGING SP HOSTNAME, IP ADDRESS, SUBNETMASK, GATEWAY ADDRESSES:**

### **# /nas/sbin/navicli -h 192.168.2.101 networkadmin -set -name spb -address 192.168.2.201 -subnetmask 255.255.255.0 -gateway 192.168.2.100**

**Note:** Running this command requires system reboot

## **UPDATING SP IP ADDRESS CHANGES IN CELERRA DATABASE:**

/nas/bin/nas\_storage –modify <APM> -network -spa | -spb xxx.xxx.xxx.xxx (Updates address in symapi db)

/nasmcd/sbin/mcd\_helper add\_SP\_info\_to\_hosts\_file (Updates /etc/hosts with new IP addresses)

/nas/site/sp\_info (File that stores SP IP addresses—update manually if necessary)

## **CHANGING SP IP ADDRESSES WITH NAS 5.5.31 AND ABOVE RUNNING PROXY ARP:**

### **# /nasmcd/sbin/clariion\_mgmt –modify –spa\_ip <new\_ip> -spb\_ip <new\_ip>**

→Changes SP IP addresses

→Updates /etc/hosts & /nas/site/sp\_info files

→Updates NASDB and SYMAPI.db with new addresses

## **OTHER COMMANDS THAT CHANGE OR IMPACT SP IP ADDRESSING:**

/nasmcd/sbin/clariion\_mgmt -stop (removes SP aliases (eth3:1, eth3:2), removes IPTables entries, reapplies old IP addresses & gateways on SPs for Napa 9 and above)

/nasmcd/sbin/clariion\_mgmt –start –spa\_ip xxx.xxx.xxx.xxx –spb\_ip xxx.xxx.xxx.xxx

## **NS40/NS80 INTEGRATED CLARIION IP ADDRESS SCHEME:**

# /nas/sbin/navicli -h 192.168.1.200 networkadmin -get

Storage Processor: SP A  
Storage Processor Network Name: SPA  
Storage Processor IP Address: 192.168.1.200  
Storage Processor Subnet Mask: 255.255.255.0  
Storage Processor Gateway Address: 192.168.1.100 (eth0 on Control Station)

# /nas/sbin/navicli -h 192.168.2.201 networkadmin -get

Storage Processor: SP B  
Storage Processor Network Name: SPB  
Storage Processor IP Address: 192.168.2.201 →Default IP for Integrated on SPB changed from 192.168.1.201  
Storage Processor Subnet Mask: 255.255.255.0  
Storage Processor Gateway Address: 192.168.2.100 (eth2 on Control Station)

### **NS42 INTERNAL MANAGEMENT SWITCH PORT CABLING—Dual DM configuration:**

- a.) Ethernet cable from lower management port DM3 to port labeled 2 on CS—eth2 192.168.2.100 Second. Mgmt (upper mid.port)
- b. ) Ethernet cable from lower management port DM2 to port labeled 10/100 on CS—eth0 192.168.1.100 Pri. Mgmt (lower left port)
- c. ) Ethernet cable from upper management port DM3 to lower management port on SPB
- d.) Ethernet cable from upper management port DM2 to lower management port on SPA

### **NS42 CS NETWORK CONFIGURATION & IMPORTANT FILES:**

# /sbin/ifconfig -a

eth0 inet addr:192.168.1.100 Bcast:192.168.1.255 Mask:255.255.255.0 →Primary Internal network (CS, DM, SPA)  
eth1 <not used>  
eth2 inet addr:192.168.2.100 Bcast:192.168.2.255 Mask:255.255.255.0 →Backup Internal network (CS, DM, SPB)  
eth3 inet addr:10.241.168.94 Bcast:10.241.168.255 Mask:255.255.255.0 →External Interface customer network

# cat /etc/sysconfig/network-scripts/ifcfg-eth0

DEVICE=eth0  
IPADDR=192.168.1.100  
NETMASK=255.255.255.0  
NETWORK=192.168.1.0  
BROADCAST=192.168.1.255  
ONBOOT=yes

# cat /etc/sysconfig/network-scripts/ifcfg-eth2

DEVICE=eth2  
IPADDR=192.168.2.100  
NETMASK=255.255.255.0  
NETWORK=192.168.2.0  
BROADCAST=192.168.2.255  
ONBOOT=yes

# cat /etc/nas\_device.map

DOSDSK=/dev/nd0  
OS1DSK=/dev/hda  
OS2DSK=/dev/hda  
VERDSK=/dev/hda  
VARDISK=/dev/ndf  
NBSDSK=/dev/nde  
ENET\_INT0=eth0  
ENET\_INT1=eth2  
ENET\_EXT=eth3  
ENET\_SP=  
ENET\_IPMI=

# cat /etc/nas\_enclosure.map

#WARNING!! DO NOT MODIFY CONTENTS OF THIS AUTO-GENERATED FILE

ENCLOSURE-0\_MGMT-A\_IP=192.168.1.50  
ENCLOSURE-0\_MGMT-A\_MAC=00:60:16:0b:1a:ca  
ENCLOSURE-0\_MGMT-B\_IP=192.168.2.50  
ENCLOSURE-0\_MGMT-B\_MAC=00:60:16:0b:1b:a5

# cat /proc/sys/net/ipv4/ip\_forward

0 →By default, IP routing is disabled on the Control Station [But not after NAS 5.5.30 and higher—we enable it]

## **NS41 INTEGRATED INTERNAL MGMT SWITCH PORT CABLING—Single DM configuration:**

- a.) Ethernet cable runs from lower management port of SPB to the port labeled 2 on the Control Station switch (upper middle port)
- b.) Ethernet cable runs from lower management port of DM2 to the 10/100 port on CS0 (lower left port)
- c.) Ethernet cable runs from upper management port of DM2 to the lower management port on SPA

**Note:** With the NS41 system, both Primary & Backup Networks are now served from the same physical interface (eth0 & an alias called eth0:0). In this arrangement, eth2 does not service the Backup network for the Celerra, but instead is reconfigured with the 192.168.2.102 address, along with a special routing entry in /etc/sysconfig/static-routes, in order for the CS to communicate back & forth to SPB. All Primary & Backup network activity, with exception of SPB & eth2, are handled via eth0 or eth0:0.

## **NS41 CS NETWORK CONFIGURATION & IMPORTANT FILES:**

### **# /sbin/ifconfig -a**

```
eth0    inet addr:192.168.1.100 Bcast:192.168.1.255 Mask:255.255.255.0      →Primary Internal network (CS, DM, SPA)
eth0:0  inet addr:192.168.2.100 Bcast:192.168.2.255 Mask:255.255.255.0      →Backup Internal network alias (CS, DM)
eth1    <not used>
eth2    inet addr:192.168.2.102 Bcast:192.168.2.255 Mask:255.255.255.254 →SPB communications only (CS, SPB)
eth3    inet addr:10.241.168.94 Bcast:10.241.168.255 Mask:255.255.255.0      →External Interface customer network
```

### **# cat /etc/sysconfig/static-routes**

```
eth2 host 192.168.2.201 dev
```

**Note:** This entry must exist in order for proper communication to occur between CS & SPB. Under normal circumstances, you should be able to ping between CS to SPs and all internal networks, navicli commands should run against either SP, etc. This file is created only with an NS41 Integrated installation.

### **# cat /etc/sysconfig/network-scripts/ifcfg-eth0:0**

```
DEVICE=eth0:0
IPADDR=192.168.2.100
NETMASK=255.255.255.0
NETWORK=192.168.2.0
BROADCAST=192.168.2.255
ONBOOT=yes
```

### **# cat /etc/sysconfig/network-scripts/ifcfg-eth0**

```
DEVICE=eth0
IPADDR=192.168.1.100
NETMASK=255.255.255.0
NETWORK=192.168.1.0
BROADCAST=192.168.1.255
ONBOOT=yes
```

### **# cat /etc/sysconfig/network-scripts/ifcfg-eth2**

```
DEVICE=eth2
IPADDR=192.168.2.102
NETMASK=255.255.255.54
NETWORK=192.168.2.201
BROADCAST=192.168.2.255
ONBOOT=yes
```

### **# cat /etc/nas\_device.map**

```
DOSDSK=/dev/nda
OS1DSK=/dev/hda
OS2DSK=/dev/hda
VERDSK=/dev/hda
VARDSK=/dev/ndf
NBSDSK=/dev/nde
ENET_INT0=eth0
ENET_INT1=eth0:0
ENET_INT2=eth2 →New entry to support NS41 communications with SPB
ENET_EXT=eth3
ENET_SP=
ENET_IPMI=
```

### **# cat /etc/nas\_enclosure.map**

#WARNING!! DO NOT MODIFY CONTENTS OF THIS AUTO-GENERATED FILE

ENCLOSURE-0\_MGMT-A\_IP=192.168.1.50

ENCLOSURE-0\_MGMT-A\_MAC=00:60:16:0b:1a:ca

ENCLOSURE-0\_MGMT-B\_IP= →Mgmt B is not used with single DM configuration

ENCLOSURE-0\_MGMT-B\_MAC= →Mgmt B is not used with single DM configuration

**Note:** Output abbreviated to show only Enclosure 0

**# cat /etc/encl\_dhcpd.conf**

#BEGIN Dynamic Primary Subnet IP range#

subnet 192.168.1.0 netmask 255.255.255.0

    allow members of "APC\_UPS";

    range 192.168.1.90 192.168.1.90;

#BEGIN Dynamic Secondary Subnet IP range#

subnet 192.168.2.0 netmask 255.255.255.0

    allow members of "APC\_UPS";

    range 192.168.2.90 192.168.2.90;

#BEGIN mgmt\_2\_3#

    fixed-address 192.168.1.50;

#BEGIN mgmt\_2\_3b#

    fixed-address 192.168.2.50;

## **NS80 INTEGRATED INTERNAL MGMTN SWITCH PORT CABLING—(4) DM configuration:**

- a.) Ethernet cable from lower left Port 3 on DM Encl 0 Mgmt Module A to port labeled 10/100 on CS—eth0 (lower left port)
- b.) Ethernet cable from lower left Port 3 on DM Encl 0 Mgmt Module B to port labeled #2 on CS—eth2 (upper middle port)
- c.) Ethernet cable from upper Port 0 on DM Encl 0 Mgmt Module A to lower left Port 3 on DM Encl 1 Mgmt Module A
- d.) Ethernet cable from upper Port 0 on DM Encl 0 Mgmt Module B to lower left Port 3 on DM Encl 1 Mgmt Module B
- e.) Ethernet cable from lower right Port 4 on DM Encl 1 Mgmt Module A to upper Mgmt port on SP Mgmt module A (right side)
- f.) Ethernet cable from lower right Port 4 on DM Encl 1 Mgmt Module B to upper Mgmt port on SP Mgmt module B (left side)

### **Internal Network Cabling for NS80 (2) Blade Configuration:**

- a.) Ethernet cable from lower left Port 3 on DM Encl 0 Mgmt Module A to port labeled 10/100 on CS—eth0 (lower left port)
- b.) Ethernet cable from lower left Port 3 on DM Encl 0 Mgmt Module B to port labeled #2 on CS—eth2 (upper middle port)
- c.) Ethernet cable from lower right Port 4 on DM Encl 0 Mgmt Module A to upper Mgmt port on SP Mgmt module A (right side)
- d.) Ethernet cable from lower right Port 4 on DM Encl 0 Mgmt Module B to upper Mgmt port on SP Mgmt module B (left side)

### **Dual Control Stations:**

- g.) Lower right Port 4 Encl 0 Mgmt Module A to CS1 10/100 port
- h.) Lower right Port 4 Encl 1 Mgmt Module B to CS1 #2 port

## **VARIOUS CONNECTION METHODS TO NS40/NS80 INTEGRATED: CS, DM, SP**

### **CONNECTING TO NS40/NS80 INTEGRATED CS**

#### **I. CONNECTING TO CS USING SERIAL CABLE & HYPERTERMINAL:**

1. Connect null modem serial cable from Workstation COM1/COM2 to front serial port on the NS40/NS80 Control Station
2. Configure Hyperterminal connection for the Control Station using the following settings:  
Name of connection; Connect using: COM1 or COM2; Bits per second: 19200 Data bits: 8 Parity: None Stop bits: 1 Flow Control: None (verify that Autodetect & ANSI are selected for terminal emulation)
3. Click on Hyperterm ‘Call’ icon to connect to Control Station & login at the prompt

#### **II. CONNECTING TO CS USING CROSSOVER ETHERNET CABLE TO EXTERNAL PORT ETH3:**

1. Temporarily remove External LAN cable from Public network and connect crossover cable from Workstation to external CS port eth3 (far left side of CS when viewed from rear) to create a private network
2. Configure an un-used IP address for the Workstation on the same network as the Control Station, & the internal network gateway:

**10.241.168.254** (example)

**255.255.255.0**

**192.168.1.100**

3. Verify connectivity from Workstation:

c:>ping 192.168.1.100 (Primary Internal Network CS0)

c:>ping 192.168.2.100 (Backup Internal Network CS0)

c:>ping 192.168.1.200 (SPA address)

c:>ping 192.168.2.201 (SPB address)

4. Open Internet Explorer Browser window for Celerra Manager or Navisphere

**192.168.1.100** or External IP Address (Celerra Manager, optionally use SSH Shell to run CLI)

**192.168.1.200** SPA (Navisphere Manager) or 192.168.2.201 SPB

**Note:** Both Celerra Manager & Navisphere Manager should show both SPs as managed when IP Forwarding is enabled

5. Enable IP Forwarding on Control Station via network connection:

# echo 1 >/proc/sys/net/ipv4/ip\_forward

# cat /proc/sys/net/ipv4/ip\_forward

1

## CONNECTING TO NS40/NS80 INTEGRATED DM

### I. CONNECTING TO NS40 BLADE USING SERIAL CABLE TO LAPTOP & HYPERTERM:

1. Connect serial cable from Workstation COM1/COM2 to DM upper serial maintenance port (indicated by wrench symbol) using special mini-DB9 cable (Part #038-003-084)

**Note:** Mini-DB9 cable is unique to NS40/NS80 Data Movers

2. Configure Hyperterminal connection for the Data Mover using the following settings:

Name of connection: Connect using: COM1 or COM2; Bits per second: 9600 Data bits: 8 Parity: None Stop bits: 1 Flow Control: None (verify that Autodetect & ANSI are selected for terminal emulation)

3. Click on ‘Call’ icon to connect to Data Mover: CONSOLE>

4. Set output to the console session and run direct commands to the Server:

**CONSOLE>logsys add output console**

**CONSOLE> version**

Product = EMC Celerra File Server

Version = T5.6.43.4 (CMB\_SLEDGEHAMMER 32-bit)

Debug = DEBUGOFF

BIOS = 3.58

POST = Rev. 01.50

1233070896: ADMIN: 6: Command succeeded: version

CONSOLE>ifconfig

5. Reset the console output to Log:

**CONSOLE>logsys delete output console**

**Note:** Reboot the Data Mover from the Control Station to watch the bootup process, or access F2 Setup/BIOS menu, or to access POST menu

### II. CONNECTING TO NS80 DM USING SERIAL CABLE & HYPERTERMINAL:

1. Connect from Workstation COM1/COM2 to Data Mover’s USB serial maintenance port (serial-to-USB adapter cable):

**Note:** Connect to righthand Management Module A’s USB port A (wrench symbol) using a Serial db9-to-USB connector, to connect to first blade in each enclosure (DM2, DM4, etc.). Connect to lefthand Management Module B’s USB port B (wrench symbol) using a Serial db9-to-USB connector, to connect to second blade in each enclosure (DM3, DM5, etc.).

2. Configure Hyperterminal connection for the Data Mover using the following settings:

Name of connection: Connect using: COM1 or COM2; Bits per second: 9600 Data bits: 8 Parity: None Stop bits: 1 Flow Control: None (verify that Autodetect & ANSI are selected for terminal emulation)

3. Click on ‘Call’ icon to connect to Data Mover: CONSOLE>

4. Set output to the console session and run direct commands to the Server:

CONSOLE>logsys add output console

CONSOLE> version

Product = EMC Celerra File Server

Version = T5.6.43.4 (HAMMERHEAD 32-bit)

Debug = DEBUGOFF

BIOS = 3.58

POST = Rev. 01.50

1233070896: ADMIN: 6: Command succeeded: version

CONSOLE>ifconfig

5. Reset the console output to Log:

CONSOLE>logsys delete output console

**Note:** Reboot the Data Mover from the Control Station to watch the bootup process, or access F2 Setup/BIOS menu, or to access POST menu

## CONNECTING TO NS40/NS80 INTEGRATED SP

### I. CONNECTING TO NS40/NS80 SPA/SPB USING SERIAL CABLE & HYPERTERMINAL:

1. Connect serial cable from Workstation COM1/COM2 to SP upper serial maintenance port (wrench symbol) using special mini-DB9 cable (Part #038-003-084)

2. Configure Hyperterminal connection for the Service Processor (SPA/SPB) using the following settings:

Name of connection: Connect using: COM1 or COM2; Bits per second: 9600 Data bits: 8 Parity: None Stop bits: 1 Flow Control: None (verify ANSIW terminal emulation)

3. Click on ‘Call’ icon to connect to SP

4. Reboot SP to observe bootup and Resume PROM settings

**Note:** Above procedure is the same for CX3-80 SP’s, just be aware that SPB’s upper maintenance serial port is located on the lefthand side Management Module B, while SPA’s serial port is located on the righthand side Management Module A.

## **II. CONNECTING TO SP MANAGEMENT LAN ON NS40i OR SERVICE LAN ON NS80i:**

1a. Connect straight-through Ethernet cable from Workstation to the upper Management LAN port on SPA or SPB for NS40 Integrated

1b. Connect straight-through cable from Workstation to lower Service LAN port on Management Module A or Management Module B for NS80 Integrated

**Note:** The NS40 SP uses the lower Service LAN port on the CX3-40 array for the Internal Celerra network, while the NS80 uses the upper Management LAN port on the CX3-80 array for the Internal Celerra network. Please note that all CX3 arrays consider the bottom LAN port as the Service port, and the upper LAN port as the Management port. A BCM switch handles the internal network communications on the SPs, and is also why a crossover cable is not required.

2. Configure the Workstation to access the hard-coded CLARiiON System Network: 128.221.1 (same network SPA, SPB)

### **System IP Address SPA or SPB:**

**128.221.1.250 (SPA) 128.221.1.251 (SPB)**

a. Configure Workstation with following IP address:

My Network Places>Properties>Local Area Connection 2>Properties>Internet Protocol (TCP/IP)>Properties:

**128.221.1.249**

**255.255.255.0** (no gateway address or DNS settings)

b. Connect to bottom Service LAN port on left side Management Module B on CX3-80 or upper Management port on SPB for CX3-40 systems in order to communicate to SPA

c. Connect to bottom Service LAN port on right side Management Module A on CX3-80 or upper Management port on SPA on CX3-40 systems in order to communicate to SPB

**Note:** When connecting to the Clariion hard-coded System IPs on the CX3 arrays, communication to SPB is done via SPA’s Service or Management port on the CX3-40 and via Management Module A on the CX3-80. Similarly, communication to SPA is done via SPB’s Service or Management port on the CX3-40 array and via Management Module B on the CX3-80 array. Also, unlike the private Celerra network, which uses two different subnets, both SP’s use the same network when using the Clariion network.

3. Alternatively, configure the Workstation to access the default Celerra Internal IP network for either SPA or SPB:

### **Celerra IP Address SPA or SPB:**

**192.168.1.200 (SPA) 192.168.2.201 (SPB)**

a. Configure Workstation with following IP address & mask:

My Network Places>Properties>Local Area Connection 2>Properties>Internet Protocol (TCP/IP)>Properties:

**192.168.1.249 (SPA) 192.168.2.249 (SPB)**

**255.255.255.0**

b. Connect to the right side Management Module A on CX3-80 or SPA on CX3-40 system in order to communicate to SPA

c. Connect to the left side Management Module B on CX3-80 or SPB on CX3-40 system in order to communicate to SPB

**Note:** Unlike the default CLARiiON system IP Addresses, when connecting to the arrays using the Celerra’s Internal IP Address scheme, use the upper Management LAN port on SPA to talk to SPA, and the upper Management LAN port on SPB to talk to SPB for the NS40 system—use the left side Management Module’s lower Service LAN port to talk to SPB and the right side Management Module’s lower Service LAN port to talk to SPA for the NS80 system.

4. Verify connectivity from Workstation using Ping:

c:>ping 192.168.1.200 (SPA) or c:>ping 128.221.1.250 (SPA)

c:>ping 192.168.2.201 (SPB) or c:>ping 128.221.1.251 (SPB)

5. Open web browser and connect to Celerra Manager via IP Address:

192.168.1.100 (Primary network)

192.168.2.100 (Backup network)

**Note:** With the Internal Network scheme properly configured for either an NS42 or NS41 Integrated Celerra, both SPs should be reachable via CLI or Celerra Manager without the need for enabling IP Forwarding. Use Celerras>Tools>SSH Shell to open CLI session on Control Station from Celerra Manager.

6. Open web browser and connect to Navisphere Manager, or to the Setup program

192.168.1.200/setup (SPA) 192.168.2.201/setup (SPB)

128.221.1.250/setup (SPA) 128.221.1.251/setup (SPB)

**Note:** At the Navisphere Security popup window saying “Global Security is not initialized”, answer “No”. Without IP Forwarding enabled on the Control Station, the SP opposite the one to which the connection is made will show up as “(Unmanaged)”, and the connected SP may also display an array Fault with “client non-data stream error: TIMEDOUT”. This issue is seen primarily when running Flare versions 19 & 22 and is not seen with Flare 24.

7. Enable IP Forwarding on Control Station to resolve any connectivity issues within Navisphere Manager

# echo 1 >/proc/sys/net/ipv4/ip\_forward

# cat /proc/sys/net/ipv4/ip\_forward

1

**Note:** Once connected to the local internal Celerra network, a Web Browser session can be used to open either Celerra Manager or Navisphere Manager.

### **III. CONNECT TO CLARIION SYSTEM IPs TO VERIFY IP ADDRESS, HOSTNAME, GATEWAY:**

**Note:** Most Integrated systems should ship with CLARiiON SP's & Hostnames already configured for the Integrated Celerra environment. Use the default CLARiiON system IP addresses and the following procedure to verify or set the correct IP address, Hostname, and Gateway addresses for an Integrated Celerra using the /setup program.

1. Connect straight-through Ethernet cable from Workstation to the upper LAN port [Management] on SPA or SPB for CX3-40 and lower Service LAN port on Management Module A or Management Module B.

2. Configure the Workstation with a static IP of 128.221.1.249 or 128.221.1.254 & the following mask:

**Note:** Reserved IP Addresses for CLARiiON SP's are 128.221.1.250, 128.221.1.251, 128.221.1.252, 128.221.1.253, 192.168.1.1, & 192.168.1.2.

My Network Places>Properties>Local Area Connection 2>Properties>Internet Protocol (TCP/IP)>Properties:

- Use the following IP address:  
**IP address:** 128.221.1.249 or .254  
**Subnet mask:** 255.255.255.0

3. Verify connectivity from Workstation using Ping:

c:>ping 128.221.1.250/.251

4. Open web browser and connect to setup program on SPA/SPB, then confirm whether each SP has the correct IP Address, Hostname, & Gateway addresses for the Celerra environment:

**SPA: 128.221.1.250/setup**

**SPB: 128.221.1.251/setup**

5. Set proper IP Address, Hostname, and Gateway address on SPA, and wait until SPA reboots before continuing.

6. Connect to SPB/setup and repeat Step 5 process

7. For array management using Navisphere Manager (if Celerra Control Station is up), enable IP Forwarding so that both SP's can be managed

### **IV. CONNECT TO CLARIION SPs USING PPP (IP over Serial line) SESSION:**

**Note:** This connection method should be considered a last resort for a number of reasons. It's slow, it's time-consuming to setup, and it can lead to subsequent NAS Installation issues if the PPP connection is not properly disconnected.

1. Connect from Workstation COM1/COM2 to SPA/SPB upper serial maintenance port (wrench symbol) using mini-db9 cable

**Note:** For CX3-80 SPs, this port is located in the right side Management module for SPA, left side module for SPB

2. Create an IP connection between computers (Windows system & SPA) using IP over the serial interface:

**W2K:**

Start>Settings>Network and Dial-up Connections>Make New Connection>Next>Connect directly to another computer>Next>Guest>Next>Select a device (e.g., Communications Port (COM1))>Next>Create this connection: For all users>Next>provide a name for the connection and click 'Finish'.

**XP:**

Start>Connect To>>Show all connections>Create a New Connection>Next>Setup an advanced connection>Connect directly to another computer>Guest>Connection Name: SPA>Select a device (e.g., COM1)>Next>Create this connection for: Anyone's use>'Finish'.

3. The Connection window remains open--select Properties>Configure>set to 115200 with hardware flow control>click o.k. & then enter following:

**Connect To dialog box:**

**User name:** clarion

**Password:** clarion!

**Note:** Once authenticated and logged in, the system automatically assigns an IP & Route address of 192.168.1.2 for the workstation, with mask 255.255.255.255 for the WAN (PPP/SLIP) Interface, and uses 192.168.1.1 for SPA or SPB, depending on whether the serial port connection is to SPA or SPB.

4. Once the session is connected, open a Browser window and connect to SPA using its default system IP Address & setup program—answer “No” if prompted to create Domain Security on the CLARiiON:

**128.221.1.250/setup (SPA)**

**Note:** Since Navisphere & java are running over a serial connection, it will take time for Navisphere to connect. If the SP's are already configured for the Celerra environment, you would not need to run the /setup program and could instead simply connect to Navisphere using SPA or SPB's IP Address

5. For Integrated Celerras, configure the Hostname for SPA as “SPA”, enter the Celerra IP Address & Gateway information, click Apply Settings and the system will reboot:

Hostname: OEM-WF3HGVFERA0→Change to SPA

192.168.1.200 (IP address SPA)

255.255.255.0 (Mask)

192.168.1.100 (Gateway)

192.168.2.201 (Peer—SPB)

**Note:** The NS40/NS80 Celerra uses a different Internal IP network for SPB than that used for SPA

6. After SPA has completed its reboot, move the Serial cable to SPB's upper maintenance serial port (wrench symbol), reconnect the PPP session from the Windows system to the SP, and then open the Browser to connect to the default IP address for SPB so as to configure its IP address and hostname—remember to answer “No” when asked to enable Domain Security:

#### **128.221.1.251/setup**

Hostname: SPB

192.168.2.201 (IP Address SPB)

255.255.255.0 (Mask)

192.168.2.100 (Gateway)

192.168.1.200 (Peer IP Address SPA)

7. After SPB reboots, reconnect using the Web Browser and check for Privileged Users:

Highlight SPB>Properties>Check the Agent tab to verify whether any “Privileged Users” have been configured on the system. If Users are already configured, add the Celerra using: root@<CS\_External\_IP\_Address>

8. Reconnect the serial cable to SPA's maintenance port, connect via PPP session, then connect to Navisphere Manager on SPA using IP Address 192.168.1.200:

a.) Click Yes at the Warning – Security screen

b.) Enter IP Address & port to use for SPA in the Navisphere Connection screen

c.) Select No at the “Confirm: Navisphere Security” screen as we do not want to setup Clariion security

## **RESUME PROM ENCLOSURE SETTINGS FOR CX3 CLARiiON SPs & DATA MOVER BLADES:**

### **NAS:**

Resume PROM for Gateway or Integrated Data Movers should always be set to NAS

DHCP client is Enabled and allows for enclosure management

VLAN restrictions are Disabled to allow for daisy chaining of enclosures

### **SAN AUX:**

Resume PROM for Clariion SP's in most Integrated Celerra systems should be set to SAN AUX

**Note:** The exception to this rule is the NS80 Integrated system, which uses optical connectors, and SP's are set to SAN

DHCP client is Disabled in this enclosure configuration

VLAN restrictions are Disabled to allow dual Control Station configurations to work for Integrated NAS systems—no VLAN security set for Service vs. Public LAN ports

POST would stop blade from booting if SFP's were found in FE ports

### **SAN:**

Resume PROM for Gateway & NS80 Integrated Clariion SPs should be set to SAN

DHCP client is Disabled

VLAN restrictions are Enabled to keep Service LAN separate from Public LAN—i.e., VLAN security is set

POST would stop blade from booting if copper cables in FE ports are detected

## **VERIFYING & CHANGING RESUME PROM ENCLOSURE SETTING FOR NS40/NS80 AUX SPs:**

1. Connect serial cable from Workstation COM1/COM2 to SP upper serial maintenance port (indicated by wrench symbol) using special mini-DB9 cable (Part #038-003-084)

2. Configure Hyperterminal connection for the Service Processor (SPA/SPB) using the following settings:

Name of connection: Connect using: COM1 or COM2; Bits per second: 9600 Data bits: 8 Parity: None Stop bits: 1 Flow Control: None (verify that terminal emulation is set to Autodetect/ANSI)

3. Click on ‘Call’ icon to connect to SP

4. Reboot SP to observe bootup and Resume PROM settings:

BIOS 3 Release 6.1 POST Version 01.27

**Model: Sledgehammer: SAN AUX** (NS40Integrated)

**Model: Hammerhead: SAN** (NS80Integrated)

**Note:** Correct setting for Clariion SP in NS40 Integrated system is SAN AUX, while correct setting for NS80 Integrated would be SAN, due to direct optical connections between DMs and SPs. Normal setting would be SAN for NS40 Gateway models.

5. Use following steps if the correct SP personality is not set

a.) During the Extended POST sequence, where characters begin scrolling across the screen [ABCab...], use the “ctrl + c” keys (or Esc key if that does not work) to let POST complete, at which point a “.... Storage System Failure - Contact your Service Representative” message will display. Enter the POST password: SHIP\_it

**Note:** Original method used to interrupt POST was the “Esc” key—use the “esc” key only if the ctrl + c sequence does not work, say for an older platform, etc.

\*\*\*\*\*

\*WARNING: ABORTING EMC EXTENDED POST VIA ESC IS UNSUPPORTED \*

\* AND MAY CAUSE SYSTEM INSTABILITY. \*

\* PLEASE REBOOT AND USE THE CTRL-C ABORT MECHANISM. \*

\*\*\*\*\*

b.) At the POST Diagnostic Menu, select menu item (15) for Resume PROM Sub-Menu

**Note:** Menu item 22) for NS80

c.) Select 2) Set Resume and 5 XP\_Chassis

d.) Enter defaults until the SAN/NAS Type (5): SAN screen appears. Enter SAN AUX for NS40 systems, which is Type (3). Enter (1) for type SAN for NS80 systems, then enter 1 – Hammerhead for Family ID, if not already set to 1. Accept all other defaults.

**Note 1:** Enter SAN/NAS Type [0-Agnostic 1-SAN 2-NAS 3-SAN AUX] (0-3) [3] for NS40

**Note 2:** Enter SAN/NAS Type [0-Agnostic 1-SAN 2-NAS 3-SAN AUX 4-SAN SYMM] (0-4) [1] for NS80

e.) Exit menus to main Diagnostic Menu, select 1) Reset Controller to reboot SP, observe personality during reboot

f.) Reboot other SP and verify personality

## **VERIFYING & CHANGING RESUME PROM ENCLOSURE SETTING FOR NS40/NS80 AUX DM:**

**Note:** It's possible that a system may not have the correct RESUME setting for the DM blade in a CX3 system.

1. Connect to upper maintenance serial port on DM using special mini-db9 connector

**Note:** Connect to righthand Management Module A's upper leftmost USB port using a Serial db9-to-USB connector, to connect to first blade in each enclosure (DM2, DM4, etc.). Connect to lefthand Management Module B's upper leftmost USB port using a Serial db9-to-USB connector, to connect to first blade in each enclosure (DM3, DM5, etc.).

2. Open HyperTerminal application and configure for DM connection:

9600; 8; None; 1; None; Autodetect/ANSI

3. Reseat Data Mover and observe bootup:

Booting Extended POST

**Model: Sledgehammer: NAS (NS40)**

**Model: Hammerhead: NAS (NS80)**

**Note:** All NS40/NS80 Integrated or Gateway Data Movers should have their Enclosure personality set to “NAS”

4. If set to anything else, access the POST menu to update the Resume Prom using the following steps:

a.) When the following characters begin scrolling across the screen, press “ctrl + c” keys (or use “Esc” key if that does not work), which lets POST complete, then login to the POST menu using “SHIP\_it” or “DB\_key”—see following output:

*ABCab << Stopping after POST >>*

*DEabcdeFGHabcdIJKLMabcdeNOPabcQabcRabcSTUabVWXabYZAABBabcCCabcDDabcEEabcFFabGGHHIIJKKLLMMabcN  
NabcOOabcPPabcQQabRRabSSabTTabUUabVVabcWWXXYYZZAAABBCCCD*

*EndTime: 09/28/2009 19:23:48*

*.... Storage System Failure - Contact your Service Representative ...*

**<Type “SHIP\_it” here>**

b.) At POST Diagnostic Menu, select menu item (15) for Resume PROM Sub-Menu on NS40, or (22) for NS80

c.) Select 2) Set Resume and then 5 XP\_Chassis

d.) Enter defaults until the SAN/NAS Type (5): SAN screen appears. Enter [2] for NAS which is Type (2).

**Note:** Choices are [0-Agnostic 1-SAN 2-NAS 3-SAN AUX 4-SAN SYMM\*] (0-4) [ ]

e.) Exit menus to main Diagnostic Menu, select 1) Reset Controller to reboot SP, observe personality during reboot

f.) Disconnect

\*SAN SYMM choice is unique to NS80

## **NS42 CABLING:**

### **Internal Network Cabling for NS42:**

a.) Ethernet cable from lower management port DM3 to port labeled 2 on Control Station—eth2 (upper middle port)

b.) Ethernet cable from lower management port DM2 to port labeled 10/100 on CS—eth0 (lower left port)

c.) Ethernet cable from upper management port DM3 to lower management port on SPB

d.) Ethernet cable from upper management port DM2 to lower management port on SPA

**Note:** For NS42F Celerra, instead of lower port on SPs, will connect to Upper LAN port on SPs for c.) and d.) above

### **Fibre Channel Cabling from DM to SP on NS42:**

a. Data Mover blade 2 port BE 0 to SP A port 0 Fibre using copper SFP-to-SFP cable

b. Data Mover blade 2 port BE 1 to SP B port 0 Fibre using copper SFP-to-SFP cable

c. For dual blade systems, Data Mover blade 3 port BE 0 to SP A port 1 Fibre using copper SFP-to-SFP cable

d. For dual blade systems, Data Mover blade 3 port BE 1 to SP B port 1 Fibre using copper SFP-to-SFP cable

### **Fibre Channel Cabling from SP to DAE3P on NS42:**

a. SPA port BE0 to LCC A port Pri on first DAE3P [Encl 0 Loop 0] using SFP-to-HSSDC2 cable

b. SPB port BE0 to LCC B port Pri on first DAE3P [Encl 0 Loop 0] using SFP-to-HSSDC2 cable

- c. SPA port BE1 to LCC A port Pri on second DAE3P [Encl 0 Loop 1] using SFP-to-HSSDC2 cable
- d. SPB port BE1 to LCC B port Pri on second DAE3P [Encl 0 Loop 1] using SFP-to-HSSDC2 cable
- e. First DAE3P LCC A port Exp [Encl 0 Loop 0] to third DAE3P LCC A port Pri [Encl 1 Loop 0] using HSSDC2-to-HSSDC2
- f. First DAE3P LCC B port Exp [Encl 0 Loop 0] to third DAE3P LCC B port Pri [Encl 1 Loop 0] using HSSDC2-to-HSSDC2
- g. Second DAE3P LCC A port Exp [Encl 0 Loop 1] to fourth DAE3P LCC A port Pri [Encl 1 Loop 1] using HSSDC2-to-HSSDC2
- h. Second DAE3P LCC B port Exp [Encl 0 Loop 1] to fourth DAE3P LCC B port Pri [Encl 1 Loop 1] using HSSDC2-to-HSSDC2

#### **SPS Mini-DB9 serial to RJ12 Data Communications Cabling for NS42:**

- a. Mini-DB9 serial connection from SPB (bottom port) to SPS-B RJ12 jack
- b. Mini-DB9 serial connection from SPA (bottom port) to SPS-A RJ12 jack

#### **HSSDC2 & SFP CONNECTORS:**

HSSDC2 stands for High Speed Serial Data Cable 2, a copper fibre cable supporting 2.5Gbits/sec and having a smaller form factor than the HSSDC cable. HSSDC2 is generally used for connecting to the DAE3P or between DAE3Ps, and uses Copper FC Arbitrated Loop. SFP stands for Small Form-Factor Pluggable connectors, and can be either Copper Fibre or LC Duplex Fibre Optical, typically used for connecting to Hosts, such as CX3 series SPs and DMs.

### **NS42F CABLING:**

#### **Internal Network Cabling for NS42F:**

- a.) Ethernet cable from lower management port DM3 to port labeled 2 on Control Station—eth2 (upper middle port)
- b.) Ethernet cable from lower management port DM2 to port labeled 10/100 on CS—eth0 (lower left port)
- c.) Ethernet cable from upper LAN port DM3 to Upper Management LAN port on SPB
- d.) Ethernet cable from upper LAN port DM2 to Upper Management LAN port on SPA

#### **Fibre Channel Cabling from DM to SP on NS42F:**

- a. Data Mover blade 2 port BE 0 to SP A port 2 Fibre using fibre optic LC-to-LC SFP cable
- b. Data Mover blade 2 port BE 1 to SP B port 2 Fibre using fibre optic LC-to-LC SFP cable
- c. For dual blade systems, Data Mover blade 3 port BE 0 to SP A port 3 Fibre using fibre optic LC-to-LC SFP cable
- d. For dual blade systems, Data Mover blade 3 port BE 1 to SP B port 3 Fibre using fibre optic LC-to-LC SFP cable

### **NS80 CABLING:**

#### **Internal Network Cabling for NS80 (4) Blade Configuration:**

- a.) Ethernet cable from lower left Port 3 on DM Encl 0 Mgmt Module A to port labeled 10/100 on CS—eth0 (lower left port)
- b.) Ethernet cable from lower left Port 3 on DM Encl 0 Mgmt Module B to port labeled #2 on CS—eth2 (upper middle port)
- c.) Ethernet cable from upper Port 0 on Encl 0 Mgmt Module A to lower left Port 3 on DM Encl 1 Mgmt Module A
- d.) Ethernet cable from upper Port 0 on Encl 0 Mgmt Module B to lower left Port 3 on DM Encl 1 Mgmt Module B
- e.) Ethernet cable from lower right Port 4 on Encl 1 Mgmt Module A to upper Management port on SP Mgmt module A (right side)
- f.) Ethernet cable from lower right Port 4 on Encl 1 Mgmt Module B to upper Management port on SP Mgmt module B (left side)

#### **Internal Network Cabling for NS80 (2) Blade Configuration:**

- a.) Ethernet cable from lower left Port 3 on DM Encl 0 Mgmt Module A to port labeled 10/100 on CS—eth0 (lower left port)
- b.) Ethernet cable from lower left Port 3 on DM Encl 0 Mgmt Module B to port labeled #2 on CS—eth2 (upper middle port)
- c.) Ethernet cable from lower right Port 4 on Encl 0 Mgmt Module A to upper Management port on SP Mgmt module A (right side)
- d.) Ethernet cable from lower right Port 4 on Encl 0 Mgmt Module B to upper Management port on SP Mgmt module B (left side)

#### **Dual Control Station Cabling:**

- g.) Lower right Port 4 Encl 0 Mgmt Module A to CS1 10/100 port
- h.) Lower right Port 4 Encl 1 Mgmt Module B to CS1 #2 port

#### **Fibre Channel Cabling from DM to SP on NS80:**

- a. Data Mover blade 2 port BE 0 to SP A port 0 Fibre using Fibre Optic multi-mode LC-to-LC cable (Optical SFP)
- b. Data Mover blade 3 port BE 0 to SP A port 1 Fibre using Fibre Optic multi-mode LC-to-LC cable
- c. Data Mover blade 4 port BE 0 to SP A port 2 Fibre using Fibre Optic multi-mode LC-to-LC cable
- d. Data Mover blade 5 port BE 0 to SP A port 3 Fibre using Fibre Optic multi-mode LC-to-LC cable
- e. Data Mover blade 2 port BE 1 to SP B port 0 Fibre using Fibre Optic multi-mode LC-to-LC cable
- f. Data Mover blade 3 port BE 1 to SP B port 1 Fibre using Fibre Optic multi-mode LC-to-LC cable
- g. Data Mover blade 4 port BE 1 to SP B port 2 Fibre using Fibre Optic multi-mode LC-to-LC cable
- h. Data Mover blade 5 port BE 1 to SP B port 3 Fibre using Fibre Optic multi-mode LC-to-LC cable

#### **Fibre Channel Cabling from SP to DAE3P on NS80:**

- a. SPA port BE0 to LCC A port Pri on first DAE3P [Encl 0 Loop 0] using SFP-to-HSSDC2 cable
- b. SPB port BE0 to LCC B port Pri on first DAE3P [Encl 0 Loop 0] using SFP-to-HSSDC2 cable
- c. SPA port BE1 to LCC A port Pri on second DAE3P [Encl 0 Loop 1] using SFP-to-HSSDC2 cable
- d. SPB port BE1 to LCC B port Pri on second DAE3P [Encl 0 Loop 1] using SFP-to-HSSDC2 cable

**Note:** Cabling schemes can vary greatly, depending on the total number of DAE2P/DAE3Ps, and on the types of drives in each enclosures. Refer to Clariion Setup Guide for more information. The above example shows Bus 0 to the 1<sup>st</sup> DAE3P and Bus 1 to the 2<sup>nd</sup> DAE3P. For an (8) DAE3P system with only two Buses, you might cable Bus 0/Loop 0 from SPA & SPB BE0 ports to LCC A & LCCB PRI, respectively, on the first DAE3P, then connect the next (3) DAE3Ps together via the EXP/PRI ports. Similarly, you

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
would then connect Bus 1/Loop 1 from SPA & SPB BE1 ports to LCC A & LCCB PRI, respectively, on the 5<sup>th</sup> DAE3P, then connect the other (3) DAE3Ps together via the EXP/PRI ports.

**SPS Mini-DB9 serial to RJ12 Data Communications Cabling for NS80:**

- a. Mini-DB9 serial connection from SPB (bottom port) to SPS-B RJ12 jack
- b. Mini-DB9 serial connection from SPA (bottom port) to SPS-A RJ12 jack

## **FRESH FIELD INSTALL--NS42 INTEGRATED SYSTEM NAS 5.5.27.5**

**Determine Factory vs. Field Install using the following verification steps:**

- I. Validate that the Clariion SP Resume PROM is set correctly (SAN AUX for NS40Integrated, SAN for NS80 Integrated)
- II. Validate that Clariion SP's are preconfigured to support the Celerra Internal Network (IP, Hostname, Mask, Gateway)
- III. Validate that Clariion array is preconfigured with Celerra Raid Groups/Luns
- IV. Perform Backend Cleanup of LUNs/Raid Groups if required
- V. Perform fresh install

**CELERRA NS42-FD FIELD INSTALLED INTEGRATED SYSTEM:**

1. After racking & cabling the NS42 array & Celerra hardware, power up the cabinet
2. Serially connect to SPA to verify the CLARiiON Resume PROM settings for the SP enclosure:
  - a. Connect serial cable from Workstation COM1/COM2 to SP upper serial maintenance port (indicated by wrench symbol) using special mini-DB9 cable (Part #038-003-084)
  - b. Configure Hyperterminal connection for the Service Processor (SPA/SPB) using the following settings:  
Name of connection: Connect using: COM1 or COM2; Bits per second: 9600 Data bits: 8 Parity: None Stop bits: 1 Flow Control: None (verify ANSIW terminal emulation)
  - c. Click on 'Call' icon to connect to SP
  - d. Reboot CLARiiON SPA, and during POST, verify that RESUME PROM values are set to SAN AUX

**Note:** An ICA's (Integrated Clariion Array) should display the following characteristic during the Extended POST

**Model: Sledgehammer: SAN AUX**

3. Use Ethernet cable to connect to SPA/SPB to verify whether the CLARiiON array has been preconfigured for NAS Installation  
**Note:** An Integrated system should ship with the array already preconfigured for the Celerra. The purpose of the following procedure is to verify that the backend SP's and Celerra Raid Groups/System LUNs have been properly configured, as well as to make the necessary changes if not.
  - a. Connect straight-through Ethernet cable from the Workstation to the upper LAN port [Management] on either SPA or SPB
  - b. Configure the Workstation with an unused IP address and mask value for SPA's network. Connect to either the Internal Celerra Network for SPA [192.168.1.0] or to the default Clariion Network for SPA [128.221.1.0], and verify/configure SPA:  
My Network Places>Properties>Local Area Connection 2>Properties>Internet Protocol (TCP/IP)>Properties:

|                              | Celerra Network (only when configured) | Clariion Network (hard-coded) |
|------------------------------|--|-------------------------------|
| <b>SPA Address:</b>          | <b>192.168.1.200</b>                   | <b>or</b>                     |
| <b>IP address (laptop):</b>  | <b>192.168.1.249 or .254</b>           | <b>128.221.1.250</b>          |
| <b>Subnet mask (laptop):</b> | <b>255.255.255.0</b>                   | <b>128.221.1.249 or .254</b>  |

c. Open web browser and connect to the setup program for SPA [192.168.1.200/setup or 128.221.1.250/setup] to confirm that IP, Hostname, Subnet Mask, & Gateway addresses are correctly setup for the Celerra environment.

**Note:** At the Navisphere Security popup window saying "Global Security is not initialized", answer "No".

d. If required, edit the following fields:

**IP Address:** **192.168.1.200**

**Hostname:** **SPA**

**Subnet Mask:** **255.255.255.0**

**Gateway:** **192.168.1.100**

e. Click the "Apply Settings" button and SPA will reboot

**Note:** Wait approximately five minutes for SPA to reboot and come back online, then proceed to next step to verify/configure SPB

f. Configure the Workstation with an unused IP address and mask value for SPB's network. Connect to either the Internal Celerra Network for SPB [192.168.2.0] or to the default Clariion Network [128.221.1.0], and verify/configure SPB:  
My Network Places>Properties>Local Area Connection 2>Properties>Internet Protocol (TCP/IP)>Properties:

|                              | Celerra Network (only when configured) | Clariion Network (hard-coded) |
|------------------------------|--|-------------------------------|
| <b>SPB Address:</b>          | <b>192.168.2.201</b>                   | <b>or</b>                     |
| <b>IP address (laptop):</b>  | <b>192.168.2.249 or .254</b>           | <b>128.221.1.251</b>          |
| <b>Subnet mask (laptop):</b> | <b>255.255.255.0</b>                   | <b>128.221.1.249 or .254</b>  |

g. Open web browser and connect to setup program for SPB [192.168.2.201/setup or 128.221.1.251/setup] to confirm that IP, Hostname, Subnet Mask, & Gateway addresses are correctly setup for the Celerra environment.

**Note:** At the Navisphere Security popup window saying "Global Security is not initialized", answer "No".

h. If required, edit the following fields:

**IP Address:** **192.168.2.201**

**Hostname:** **SPB**

**Subnet Mask:** 255.255.255.0

**Gateway:** 192.168.2.100

i. Click the “Apply Settings” button and SPB will reboot

**Note:** Wait approximately five minutes for SPB to reboot and come back online. Please note that unlike the hard-coded Clariion network scheme, which uses the 128.221.1 network for both SP’s, Celerra uses a different network for SPA (192.168.1) and SPB (192.168.2)

4. Configure the Workstation for the Internal Celerra network (192.168.1), and reconnect to SPA’s Management LAN port to verify whether the Clariion array has been preconfigured with Celerra Raid Groups/System LUNs

a. Open web browser and connect to Navisphere Manager using SPA’s IP Address: 192.168.1.200

**Note:** If a Navisphere Security popup window appears, “Global Security is not initialized”, answer “No”.

b. Note the Flare version (Array\_name>Properties>Software)

c. Check for pre-configured Celerra Raid Groups and system LUNs by expanding RAID Groups icon. Verify that LUNs 0-5 exist for RG 0 (system) and LUN 200 for RG 200 (Hot Spare)

**Note:** The presence of preconfigured Celerra Raid Groups/LUNs is a general indicator that the system is a Factory Install and should not require a fresh Field Install. However, if the preconfigured LUNs were not built with the shipped Integrated Celerra system, then the LUNs and Raid Groups would have to be destroyed before the Celerra could be installed.

d. For Fresh Installs only, delete the Celerra LUNs and Raid Groups. Unbind LUNs 0-5, then LUN 200, then rightclick RG 0 and select “Destroy”—repeat for RG200 Hot Spare

**Note:** Delete using Navisphere or Celerra CLI—if the RGs are not deleted, the fresh install process will fail because the Celerra install script will detect the Raid Groups (and Control Luns) and error out. For factory-install systems, the Raid Groups and Control Luns are pre-configured and should already exist, and you would not delete.

#### **Fresh Install Fails because System LUNs are already present:**

“Checking for existing System RG/LUN: Replacement Control Station Detected.”

“[FAILED]”

#### **e. DELETE CELERRA STORAGE GROUP, CONTROL LUNs, & RAID GROUPS USING NAVISPHERE:**

→Use straight ethernet cable & connect directly to SPB’s service LAN port (btm), configure workstation with IP address on 128.221.1 network, open web browser, enter default CLARiiON IP address for SPA for Navisphere: 128.221.1.250

- (1) Rightclick Celerra Storage Group>Connect Hosts>Select Hosts and remove
- (2) Rightclick Celerra Storage Group>Destroy
- (3) Navisphere>Expand SPA and highlight LUNs 0-5>Unbind
- (4) Navisphere>Unowned LUNS>LUN 200>Unbind
- (5) Highlight RAID Group 0>Destroy
- (6) Highlight RAID Group 200>Destroy [Continue with steps 7-10 for NS20/20FC/40/40FC complete backend cleanup]

**Note:** With the advent of CLARiiON Virtual Provisioning, supported by NAS 5.6.45, you may also need to delete ThinLuns and ThinPools from the Storagepool list

- (7) ctrl + shift + f12 messner for Engineering Mode:

Rightclick Array>Select Engineering Mode>Disable Access Logix

- (8) Navisphere (Engineering Mode)>Array>Properties>Software>select CelerraService>Uninstall

(9) Enter Clariion Setup Program (128.221.1.250/setup), change Gateway IP if necessary, on SPA, which requires reboot (then connect to SPA’s service port to access SPB and change Gateway IP address)

(10) Reconnect to SPA /setup program, choose the option “Reset all domain information and restart the Management Server” to “Destroy Security and Domain Information” on array, if required

#### **f. DELETE CELERRA CONTROL LUNs & RAID GROUPS USING CLI (Backend Cleanup Procedure):**

**Note:** If the /tftpboot/setup\_backend directories do not exist, unzip and untar the /nas/tools/tftpboot.tar.gz. Using a serial connection between Workstation and Control Station, open a HyperTerminal session to the Control Station to perform cleanup. The nas\_raid cleanup procedure is now an official procedure in all the NSxx Celerra Integrated Setup Guides, as well as the NS20 & NS40.

(1) Run the following to breakdown the array management setup:

**# /nasmcd/sbin/clariion\_mgmt -stop**

**Note:** If this fails, run the command again and add the –skip\_rules switch.

(2) Unset the NAS\_DB environmental variable prior to running cleanup:

**# unset NAS\_DB**

(3) Stop NAS Services:

**# /sbin/service nas stop**

(4) Remove Raid Groups and System LUNs—Raid Groups 0 & 200 are defaults:

**# cd /tftpboot/setup\_backend; ./nas\_raid -n ./bin/navicli -a 192.168.1.200 -b 192.168.2.201 -s cleanup**

Log will be created in the current directory →nas\_raid.log

System 192.168.1.200 is up

System 192.168.2.201 is up

Clariion Array: APM00063303725 Model: CX3-40 Memory: 4096

Lun info:

```
-----  
Lun ID: 0 RG ID: 0 State: Bound ??  
Lun ID: 1 RG ID: 0 State: Bound ??  
Lun ID: 2 RG ID: 0 State: Bound ??  
Lun ID: 3 RG ID: 0 State: Bound ??  
Lun ID: 4 RG ID: 0 State: Bound ??  
Lun ID: 5 RG ID: 0 State: Bound ??  
Lun ID: 16 RG ID: 0 State: Bound ??  
Lun ID: 17 RG ID: 0 State: Bound ??
```

Disk group info:

```
-----  
Disk Group ID: 0 r5 Disks: 0_0_0,0_1,0_0_2,0_0_3,0_0_4
```

Spare info:

```
-----  
Spare ID: 200 Disk: 0_0_5
```

!!! WARNING !!!

The CLARiiON array connected to this Celerra control station contains an existing configuration.

This could include user data.

If you continue by selecting 'y' to cleanup the system at the next prompt all existing CLARiiON configuration will be destroyed and the system will have to be reinstalled.

Proceed with caution!

!!! WARNING !!!

Do you want to clean up the system [yes or no]?: yes

Cleaning Storage Group "Celerra\_emcnas\_i0"

Removing LUN .....

Removing diskgroup ..

Removing initiators ..

Removing storage group "Celerra\_emcnas\_i0"

Removing spares

Cleanup (200) ..

Access Logix disabled

Security domain removed

Done

(5) Verify:

**# /tftpboot/bin/navicli -h 192.168.1.200 getrg**

(6) Zero out drives—only systems below Flare 24—can take a long time to complete:

**# cd /tftpboot/setup\_backend; ./nas\_raid -n ./bin/navicli -a 192.168.1.200 -b 192.168.2.201 -s zero**

5. Verify CLARiiON Software Packages:

Rightclick array APM...>Properties>Software tab to verify array software:

FLARE-Operating-Environment 03.22.040.5.005 Commit required

AccessLogix - Active

NavisphereManager - Active

6. Allocate Write & Read Memory for SPs:

Rightclick Array>Properties>Memory tab, allocate 2048MB memory for Write Cache, then 968MB Read Cache for SPA, and 968MB Read Cache for SPB--this step automatically enables the Write cache on the system

**Note:** You will need to determine what the correct values are for Read Cache vs. Write Cache—the default CX3-40 system has 4096MB memory. Under Array Properties>Memory tab, configure the maximum write cache for the system, which in this example was 2048MB. Divide the remaining cache for Read cache on SPA & SPB. In general, the Clariion Best Practices guide recommends allocating as much memory as possible to Write cache, with the balance split as Read cache between the SPs, but keep in mind that specific customer environments may dictate different cache allocation values.

7. Verify Write & Read Cache Settings:

Array>Properties>Cache tab, ensure that Write cache is selected and that Read Cache for SPA & SPB are selected

8. Check for “Privileged Users” on the SPs:

Highlight SPA>Properties and check the Agent tab to verify whether any “Privileged Users” have been configured on the system. If Users are already configured, add the Celerra using: root@<CS\_External\_IP\_Address>

9. Disconnect any PPP sessions (if used) and reboot SPA to clear any potential contention with the Internal Celerra network

192.168.1.100—see emc151721

10. Establish Serial Connection between Control Station & Workstation by connecting a null modem DB9 female cable from COM port to front serial port on the NS40 Control Station

11. Create a Hyperterminal connection between systems using the following settings:

Start>Programs>Accessories>Communications>Hyperterminal>Name of connection>Connect using: COM1 or COM2>set Bits per second: 19200 Data bits: 8 Parity: None Stop bits: 1 Flow Control: None>verify that Autodetect & ANSIW are selected for terminal emulation

12. Verify that NAS Version will be compatible with the Flare version running on the Backend—go to EMC Support Matrix, as found in eLab Navigator via the Powerlink portal. In this particular example, NAS 5.5.22.2 is the minimum code compatible with Flare 22 for the NS40 platform.

13. Create install floppy from CD-ROM for NAS 5.5.27.5 and commence the fresh NAS installation

```
# mount /dev/cdrom /mnt/cdrom  
# mount -t msdos /dev/fd0 /mnt/floppy  
# dd if=/mnt/cdrom/images/boot.img of=/dev/fd0 bs=1024
```

**Note:** Though the command comes back to the prompt immediately, keep in mind that the floppy has not yet been written to—the process will take a few minutes to complete

14. Commence Installation by inserting boot floppy and CD-ROM, then rebooting or powering on the Control Station

**Note:** Below Console outputs are informational and not comprehensive as to what takes place during the installation

a.) Type following at boot prompt:

boot: **serialinstall**

Loading initrd.img.....

Loading vmlinuz.....

Running install

Running Anaconda – please wait

b.) Destructive Linux installation: **Yes**

Pre Install

Package Installation (259 packages, takes about 10 minutes)

Complete

Installation is complete. Remove the boot media and press return to reboot.

c.) Remove boot media: **O.K.**

---

Starting EMC NAS Factory Installation

d.) Is this a Control Station Fresh Install? **Yes**

e.) At prompt, configure the Internal & External IP Addresses, Control Station hostname

Restarting the network with the new configuration...

Setting up the Enclosure...

Detecting movers in cabinet: 2

f.) Answer Yes if the number of DMs detected was correct

Checking for existing System RG/LUN:

Preparing Backend...

System 192.168.1.200 is up

System 192.168.2.201 is up

Disabling write cache...

Configuring cache...

Enabling write cache...

Start to create the system disk group and luns...

Rebooting Data Movers...

Installing NBS...

NBS on volume 1 installed successfully

NBS on volume 5 installed successfully

NBS on volume 6 installed successfully

Installing EMC NAS...

Package: EMC nfs base install

Pick a NAS administrator name

g.) Enter nasadmin with password nasadmin

emcnas Package Install

Creating the NAS Partition on /dev/nde1 with fdisk...

Building the NAS filesystem on /dev/nde1 with mke2fs...

Building the IDE NAS filesystem on /dev/hda5 with mke2fs...

Executing postinstall script

Doing Package Install...

Installing EMC NAS Version 5.5.27-5

Creating the VAR Partition on /dev/ndf1  
Creating the DOS Partition on /dev/nda1  
Building the DOS filesystem on /dev/nda1 with mkfs.msdos  
Starting NAS Storage API install...  
Starting EMC Solutions Enabler install...  
Initializing the root disk partitions  
Creating servers found in slots 2 and 3  
Initializing server in slot 2 as server 2  
Reboot server in slot 2...  
Discover disks attached to server in slot 2...  
Initializing root volumes...  
Mount root filesystem...  
Setup SymApi configuration...  
Updating the Local Datamover to Datamover Interconnect Configuration  
Renaming root\_fs\_16 to root\_fs\_common  
h.) Do you wish to enable Unicode? Yes  
Setup CLARiiON APM00063303725 storage device...  
System 192.168.1.200 is up  
System 192.168.2.201 is up  
Waiting for 2 binds to complete... (this process could take several hours, depending on the backend flare version)  
Reboot and sync servers  
Creating IDE cache of NAS file system  
The "emcnas" package install Succeeded.  
To run the new package login as <nasadmin>.  
Press Enter key to exit install :  
[ OK ]

## **NS42 INTEGRATED FLARE NDU UPGRADE PROCEDURE—LAN GUI METHOD:**

**Prerequisites:** Laptop has a copy of the Flare Upgrade package and the Navisphere Service Taskbar Version 6.24 installed.

1. Configure laptop or workstation with the following static IP address & click o.k. to save:

My Network Places>Properties>Local Area Connection 2>Properties>Internet Protocol (TCP/IP)>Properties:

- Use the following IP address:  
IP address: 192.168.1.95  
Subnet mask: 255.255.255.0  
Default gateway: 192.168.1.100

2. Connect an Ethernet cable (not a crossover cable) from the Laptop/workstation to the upper LAN port on SPA (Internal Network) & ping SPA

**Note 1:** If the connection to the SP fails, disable the DNS setting and reboot the laptop and try again

**Note 2:** The Service LAN port on SPA is actually a connection to SPB, and vice versa, while the Management LAN port on SPA is a direct connection to SPA, etc.

3. Connect to Celerra Manager via Web Browser using: <http://192.168.1.100>

4. Connect to Control Station using Celerras>Tools>SSH Shell, login, su to root, and perform the following activities:

- a.) Enable IP Forwarding on the Control Station for the NDU Upgrade procedure & verify:

**# echo 1 > /proc/sys/net/ipv4/ip\_forward**

**# cat /proc/sys/net/ipv4/ip\_forward**

1

- b.) Run Storage Check to verify backend connectivity:

**# /nas/bin/nas\_storage -check -all**

Discovering storage (may take several minutes)

done

- c.) Check for trespassed luns:

**# /nas/sbin/navicli -h 192.168.1.200 getlun -trespass**

**Note:** There are no trespassed LUNs if the output is blank

- d.) Verify SP IP & Gateway addresses to Celerra Internal Network:

**# /nas/sbin/navicli -h 192.168.1.200 networkadmin -get**

Storage Processor: SP A

Storage Processor Network Name: SPA

Storage Processor IP Address: 192.168.1.200

Storage Processor Subnet Mask: 255.255.255.0

Storage Processor Gateway Address: **192.168.1.100** (eth0 on Control Station)

**# /nas/sbin/navicli -h 192.168.2.201 networkadmin -get**

Storage Processor: SP B

Storage Processor Network Name: SPB

Storage Processor IP Address: 192.168.2.201

Storage Processor Subnet Mask: 255.255.255.0

Storage Processor Gateway Address: 192.168.2.100 (eth2 on CS)

e.) Verify currently installed and committed Clariion Software:

**# /nas/sbin/navicli -h 192.168.1.200 ndu -list**

Name of the software package: FLARE-Operating-Environment

Revision of the software package: 03.22.040.5.511

Commit Required: NO

Revert Possible: NO

Active State: NO

Is installation completed: YES

Is this System Software: NO

f.) Check for Clariion Faults:

**# /nas/sbin/navicli -h 192.168.1.200 getcrus**

SPE3 Enclosure SPE

SP A State: Present

SP B State: Present

Enclosure SPE Power A0 State: Present

Enclosure SPE Power A1 State: Present

Enclosure SPE Power B0 State: Present

Enclosure SPE Power B1 State: Present

Enclosure SPE SPS A State: Present

Enclosure SPE SPS B State: Present

Enclosure SPE SPS A Cabling State: Valid

Enclosure SPE SPS B Cabling State: Valid

DAE3P Bus 0 Enclosure 0

Bus 0 Enclosure 0 Fan A State: Present

Bus 0 Enclosure 0 Fan B State: Present

Bus 0 Enclosure 0 Power A State: Present

Bus 0 Enclosure 0 Power B State: Present

Bus 0 Enclosure 0 LCC A State: Present

Bus 0 Enclosure 0 LCC B State: Present

Bus 0 Enclosure 0 LCC A Revision: 7.66

Bus 0 Enclosure 0 LCC B Revision: 7.66

Bus 0 Enclosure 0 LCC A Serial #: FCNBD063531833

Bus 0 Enclosure 0 LCC B Serial #: FCNBD063531839

**# /nas/sbin/navicli -h spa faults -list**

The array is operating normally.

g.) Verify that Clariion Statistics Logging is enabled:

**# /nas/sbin/navicli -h 192.168.1.200 setstats**

Statistics logging is ENABLED

**# /nas/sbin/navicli -h 192.168.2.201 setstats**

Statistics logging is ENABLED

**Note:** If not Enabled, run following command on each SP to Enable

**# /nas/sbin/navicli -h 192.168.1.200 setstats -on**

**# /nas/sbin/navicli -h 192.168.2.201 setstats -on**

h.) Verify Clariion IO disk and CPU load:

**# /nas/tools/check\_clariion**

CPU UTILIZATION (MUST BE LESS THAN 50% PER STORAGE PROCESSOR)

THIS WILL TAKE APPROXIMATELY 1 MINUTE(S).

CLARIION: APM00063303725:

SP A ..... 003.50% (PASS)

SP B ..... 000.64% (PASS)

DISK I/O LOAD (MUST BE LESS THAN 100 I/O OPERATIONS PER SECOND PER DISK)

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
THIS WILL TAKE APPROXIMATELY 4 MINUTES.

CLARIION: APM00063303725:

DISK 0\_0\_0 ..... 006 OPS/SEC (PASS)  
DISK 0\_0\_1 ..... 005 OPS/SEC (PASS)  
DISK 0\_0\_2 ..... 002 OPS/SEC (PASS)  
DISK 0\_0\_3 ..... 003 OPS/SEC (PASS)v

5. Use Celerra Manager to verify that Clariion IO/s are below 15,000 per second:

Celerra Manager>Celerras>Storage System>CLARiiON: review the graph

6. Using Control Station SSH session, complete the following steps:

a.) Disable Statistics Logging on each SP:

# /nas/sbin/navicli -h 192.168.1.200 setstats -off

# /nas/sbin/navicli -h 192.168.2.201 setstats -off

b.) Disable Clariion Write Cache & Verify:

# /nas/sbin/navicli -h 192.168.1.200 setcache -wc 0

# /nas/sbin/navicli -h 192.168.1.200 getcache

|                       |          |
|-----------------------|----------|
| SP Read Cache State   | Enabled  |
| SP Write Cache State  | Disabled |
| SPA Write Cache State | Disabled |
| SPB Write Cache State | Disabled |

7. On the laptop/workstation, open the Navisphere Service Taskbar application & conduct the following actions to perform the NDU Upgrade:

a.) Connect to SPA by entering IP Address or Host Name in connection box:

Navisphere Service Taskbar>File>Connect>Select Storage System to Connect

Please enter the IP Address or Host Name of the target Storage Processor (SP)

**192.168.1.200**

b.) Software Assistant>Install Software>Welcome to the Install Software Wizard>Next

c.) On the Software Selection screen, browse to the location of the Flare Upgrade package and select>Next

d.) Software Selection screen unpacks and transfers files to the CLARiiON

e.) For the ‘Server readiness for software update’, if the High Availability check flags an issue, verify that the HA configuration is correct and then check off each box in the ‘Override HA Status’ column to continue>Next

f.) A number of Rule Checks are conducted, with results posted. If only Warnings are posted, and the details say that it is o.k. to continue after verifying each Warning or Info item, proceed with the next step.

**Note:** To enter Engineering mode with the NST tool, type Ctr + Shift + F12 and type SIR. Some failed rules is can be overridden by entering a password when running a mouse over the icon—password is “siw”.

g.) Keep the default 360 seconds setting in the ‘Non-disruptive Upgrade Delay’ screen>Next

h.) A Confirmation screen appears, select Next to continue

i.) For approximately the next hour, a ‘Software Maintenance Status’ screen will provide updates as to the progress of the Flare Upgrade, starting first with SPB, then finishing with SPA

j.) After the ‘Software Maintenance Status’ screen shows 100% completion, the ‘Post-install Tasks’ screen appears. If ready to do so, commit the newly upgraded Flare version, and click Finish.

k.) A ‘Registration Summary’ window appears, collects system statistics, and then presents the option to either Email the registration information or to Save the registration information to file—select the preferred choice in the ‘Select Transport’ screen and continue. If choosing to Save Registration, browse to a location on the laptop to save the file.

l.) A ‘Success’ screen completes the upgrade. Click ‘Finish’ to exit the NST application.

8. Complete Post-Install Steps from Celerra Manager SSH Session—log back into Celerra Manager if required--the SPA reboot does not seem to reset the network interface ports and the connection should survive.

Verify that Read & Write Cache are properly enabled—system will re-enable automatically after NDU is complete:

# /nas/sbin/navicli -h 192.168.1.200 getcache

|                      |         |
|----------------------|---------|
| SP Read Cache State  | Enabled |
| SP Write Cache State | Enabled |

# /nas/sbin/navicli -h 192.168.1.200 setcache -wc 1

**Note:** Use above command to re-enable Write Cache if required, though the system should do this automatically

9. Verify that Statistics Logging is re-enabled—enable on each SP if required:

# /nas/sbin/navicli -h 192.168.1.200 setstats

Statistics logging is ENABLED

# /nas/sbin/navicli -h 192.168.2.201 setstats

Statistics logging is ENABLED

# /nas/sbin/navicli -h 192.168.1.200 setstats –on (Use to re-enable if required, and run on each SP)

10.) Manually check & trespass back any failed over Celerra LUNs:

**Note:** Since SPA was the last SP to reboot during the NDU process, SPA luns were trespassed to SPB

**# /nas/sbin/navicli -h 192.168.1.200 getlun -tresspass**

LOGICAL UNIT NUMBER 5

Default Owner: SP A

Current owner: SP B

LOGICAL UNIT NUMBER 0

Default Owner: SP A

Current owner: SP B

LOGICAL UNIT NUMBER 16

Default Owner: SP A

Current owner: SP B

**# /nas/sbin/navicli -h 192.168.1.200 trespass lun 16**

**# /nas/sbin/navicli -h 192.168.1.200 trespass lun 5**

**# /nas/sbin/navicli -h 192.168.1.200 trespass lun 0**

11.) Turn off IP Forwarding on the Control Station & Verify:

**# echo 0 > /proc/sys/net/ipv4/ip\_forward**

**# cat /proc/sys/net/ipv4/ip\_forward**

0

12.) Verify that new Flare version is committed:

**# /nas/sbin/navicli -h 192.168.1.200 ndu -list |head**

Name of the software package: FLARE-Operating-Environment

Revision of the software package: 03.24.040.5.006

Commit Required: NO

Revert Possible: NO

Active State: YES

Is installation completed: YES

Is this System Software: NO

## **NS42 INTEGRATED FLARE NDU UPGRADE PROCEDURE—CS CLI METHOD:**

**Note:** This procedure was based on NAS 5.5.22.2 and upgrading from Flare 22 (03.22.040.5.005) to 03.22.040.5.511. Due to the differing subnets on SPA & SPB, it is recommended that the Flare NDU Upgrade be performed from the Celerra Control Station, though with IP Forwarding enabled, the Navisphere GUI can manage both SPs and the NDU Upgrade is possible using the Navisphere NST (Navisphere Service Taskbar) or Navisphere GUI when directly connected to the Internal Network.

1. Verify current CLARiiON software version and state:

**# /nas/sbin/navicli -h 192.168.1.200 ndu -list**

Name of the software package: FLARE-Operating-Environment

Revision of the software package: 03.22.040.5.005

Commit Required: NO

Revert Possible: NO

Active State: YES

Is installation completed: YES

Is this System Software: NO

2. Upload Flare bundle to Control Station directory:

**/home/nasadmin/flare/cx3-40-bundle-03.22.040.5.511.pbu**

3. Run storage check to verify backend paths, etc:

**# /nas/bin/nas\_storage -check -all**

Discovering storage (may take several minutes)

done

4. Check for trespassed luns:

**# /nas/sbin/navicli -h 192.168.1.200 getlun -tresspass**

5. Verify SP Gateway addresses to Control Station Internal Networks:

**# /nas/sbin/navicli -h 192.168.1.200 networkadmin -get**

Storage Processor: SP A

Storage Processor Network Name: SPA

Storage Processor IP Address: 192.168.1.200

Storage Processor Subnet Mask: 255.255.255.0

Storage Processor Gateway Address: **192.168.1.100 (eth0 on Control Station)**

**# /nas/sbin/navicli -h 192.168.2.201 networkadmin -get**

Storage Processor: SP B  
Storage Processor Network Name: SPB  
Storage Processor IP Address: 192.168.2.201  
Storage Processor Subnet Mask: 255.255.255.0  
Storage Processor Gateway Address: **192.168.2.100 (eth2 on Control Station)**

**# /sbin/ifconfig -a**

```
eth0    Link encap:Ethernet HWaddr 00:00:F0:9F:B3:74
        inet addr:192.168.1.100 Bcast:192.168.1.255 Mask:255.255.255.0
              UP BROADCAST NOTRAILERS RUNNING MULTICAST MTU:1500 Metric:1
              RX packets:4337789 errors:0 dropped:0 overruns:0 frame:0
              TX packets:6321336 errors:0 dropped:0 overruns:0 carrier:0
              collisions:0 txqueuelen:100
              RX bytes:1248565640 (1190.7 Mb) TX bytes:3307255372 (3154.0 Mb)

eth2    Link encap:Ethernet HWaddr 00:00:F0:9F:53:07
        inet addr:192.168.2.100 Bcast:192.168.2.255 Mask:255.255.255.0
              UP BROADCAST NOTRAILERS RUNNING MULTICAST MTU:1500 Metric:1
              RX packets:513632 errors:0 dropped:0 overruns:0 frame:0
              TX packets:513887 errors:0 dropped:0 overruns:0 carrier:0
              collisions:0 txqueuelen:100
              RX bytes:342125599 (326.2 Mb) TX bytes:43462118 (41.4 Mb)
              Base address:0xd880 Memory:fbea0000-fbec0000
```

6. Check for Clariion Faults:

**# /nas/sbin/navicli -h 192.168.1.200 getcrus**

SPE3 Enclosure SPE  
SP A State: Present  
SP B State: Present  
Enclosure SPE Power A0 State: Present  
Enclosure SPE Power A1 State: Present  
Enclosure SPE Power B0 State: Present  
Enclosure SPE Power B1 State: Present  
Enclosure SPE SPS A State: Present  
Enclosure SPE SPS B State: Present  
Enclosure SPE SPS A Cabling State: Valid  
Enclosure SPE SPS B Cabling State: Valid  
DAE3P Bus 0 Enclosure 0  
Bus 0 Enclosure 0 Fan A State: Present  
Bus 0 Enclosure 0 Fan B State: Present  
Bus 0 Enclosure 0 Power A State: Present  
Bus 0 Enclosure 0 Power B State: Present  
Bus 0 Enclosure 0 LCC A State: Present  
Bus 0 Enclosure 0 LCC B State: Present  
Bus 0 Enclosure 0 LCC A Revision: 7.64  
Bus 0 Enclosure 0 LCC B Revision: 7.64  
Bus 0 Enclosure 0 LCC A Serial #: FCNBD063531833  
Bus 0 Enclosure 0 LCC B Serial #: FCNBD063531839

**# /nas/sbin/navicli -h 192.168.2.201 getcrus** (output abridged)

7. Verify that Statistics Logging is Enabled for both SP's & enable if not on:

**# /nas/sbin/navicli -h 192.168.1.200 setstats**

Statistics logging is ENABLED

**# /nas/sbin/navicli -h 192.168.2.201 setstats**

Statistics logging is ENABLED

**# /nas/sbin/navicli -h 192.168.1.200 | 2.201 setstats -on**

**Note:** To Enable statistics

8. Verify Clariion Front-End IO/s are below 15,000 per second:

a.) Using Celerra Manager>Celerras>Storage>Systems>CLARiiON CX3-40>System: Front-End I/O Requests Graph, with I/O statistics to the right of the graph: 0.0 I/O Req/sec.

9. Disable statistics logging for both SPs during the NDU process & Verify:

**# /nas/sbin/navicli -h 192.168.1.200 setstats -off**

# /nas/sbin/navicli -h 192.168.1.200 setstats

Statistics logging is DISABLED

# /nas/sbin/navicli -h 192.168.2.201 setstats –off

# /nas/sbin/navicli -h 192.168.2.201 setstats

Statistics logging is DISABLED

10. Verify CLARiiON IO disk load and CPU load:

# /nas/tools/check\_clariion

CPU UTILIZATION (MUST BE LESS THAN 50% PER STORAGE PROCESSOR)

THIS WILL TAKE APPROXIMATELY 1 MINUTE(S).

CLARIION: APM00063303725:

SP A ..... 000.65% (PASS)

SP B ..... 000.98% (PASS)

DISK I/O LOAD (MUST BE LESS THAN 100 I/O OPERATIONS PER SECOND PER DISK)

THIS WILL TAKE APPROXIMATELY 4 MINUTES.

CLARIION: APM00063303725:

DISK 0\_0 ..... 002 OPS/SEC (PASS)

DISK 0\_0\_1 ..... 003 OPS/SEC (PASS)

DISK 0\_0\_2 ..... 004 OPS/SEC (PASS)

DISK 0\_0\_3 ..... 001 OPS/SEC (PASS)

11. Disable CLARiiON Write cache

# /nas/sbin/navicli -h 192.168.1.200 setcache -wc 0

# /nas/sbin/navicli -h 192.168.1.200 getcache

SP Read Cache State      Enabled

SP Write Cache State     Disabled

SPA Write Cache State    Disabled

SPB Write Cache State    Disabled

12. As Root User, enable IP Forwarding on Control Station for the Upgrade & Verify:

# echo 1 > /proc/sys/net/ipv4/ip\_forward

# cat /proc/sys/net/ipv4/ip\_forward

1

#### **Example of Flare Upgrade Failure if IP Forwarding is not set:**

**Note:** Following example shows results of NDU Upgrade attempt without first enabling IP Forwarding—fails because lack of route to SPB—IP Forwarding allows the Linux Control Station to serve as a router between the two internal networks

# /nas/sbin/navicli -h 192.168.1.200 ndu -install /home/nasadmin/flare/cx3-40-bundle-03.22.040.5.511.pbu -delay 360

Running install rules...

=====

Version Compatibility      : Rule passed.

Special Conditions        : Rule has warning.

Statistics Logging Disabled    : Rule failed.

Acceptable Processor Utilization    : Rule failed.

Redundant SPs            : Rule failed.

Host Connectivity        : Rule has warning.

No Trespassed LUNs      : Rule passed.

No Transitions           : Rule passed.

No System Faults        : Rule passed.

All Packages Committed    : Rule passed.

2 rule(s) have warnings.

3 rule(s) failed.

Detailed rule results:

=====

RULE NAME: Special Conditions

RULE STATUS: Rule has warning.

RULE DESCRIPTION: This rule is a warning to check for special conditions before installing software on the storage system.

RULE RESULT: Please check for the following conditions:

1. All attached servers are running failover software.

2. All attached VMWare ESX servers running pre v2.1.0 software should have I/O stopped

RULE INSTRUCTION: null

RULE NAME: Statistics Logging Disabled

RULE STATUS: Rule failed.

RULE DESCRIPTION: This rule verifies statistics logging is disabled on the storage system.

RULE RESULT: Statistics logging is enabled.

RULE INSTRUCTION: Statistics logging must be disabled before the install can proceed.

RULE NAME: Acceptable Processor Utilization

RULE STATUS: Rule failed.

RULE DESCRIPTION: This rule verifies each SP's utilization is low enough to accommodate an installation.

RULE RESULT: Cannot enable statistics logging on SP B. Verify SP B is operational.

RULE INSTRUCTION: Retry the installation when the storage system workload is expected to be low.

RULE NAME: Redundant SPs

RULE STATUS: Rule failed.

RULE DESCRIPTION: This rule checks for availability of dual SPs.

RULE RESULT: SP B: Error detected, SP is unmanaged.

RULE INSTRUCTION: All SPs must be operational for the installation to proceed.

RULE NAME: Host Connectivity

RULE STATUS: Rule has warning.

RULE DESCRIPTION: This rule checks for a path from each attached host to each SP.

RULE RESULT: During the installation, connectivity may be lost to following attached hosts:

50:06:01:60:C1:E0:1C:D0:50:06:01:68:41:E0:1C:D0 SP A Port 1

50:06:01:60:C1:E0:1C:D0:50:06:01:69:41:E0:1C:D0 SP B Port 1

50:06:01:60:C1:E0:1C:D0:50:06:01:60:41:E0:1C:D0 SP A Port 0

50:06:01:60:C1:E0:1C:D0:50:06:01:61:41:E0:1C:D0 SP B Port 0

RULE INSTRUCTION: Connectivity loss may be prevented if the storage system has 2 SPs, and each attached host has a path to each SP.

Pre installation rules have been run to ensure the success of this software upgrade. The above conditions have been detected that need to be corrected before running this command again. If this condition persists please contact your EMC Service representative.

13. Perform NDU Upgrade:

**# /nas/sbin/navicli -h 192.168.1.200 ndu -install /home/nasadmin/flare/cx3-40-bundle-03.22.040.5.511.pbu -delay 360**

Item number: 0

Name of the software package: FLARE-Operating-Environment

Revision of the software package: 03.22.040.5.511

Already Installed Revision 03.22.040.5.005

Installable YES

Disruptive upgrade: NO

Ndu Delay: 360 secs

The requested package(s) will be installed. Do you wish to proceed? : (y/n)? y

14. Tracking & Verifying Install Progress:

**# /nas/sbin/navicli -h 192.168.1.200 ndu -status**

Is Completed: NO

Status: Preparing array Settings

Operation: Install

**# /nas/sbin/navicli -h 192.168.2.201 ndu -status**

Is Completed: NO

Status: Running check scripts

Operation: Install

**# /nas/sbin/navicli -h 192.168.2.201 ndu -status**

Is Completed: NO

Status: Installing software on secondary SP

Operation: Install

**# /nas/sbin/navicli -h 192.168.1.200 ndu -status**

Is Completed: NO

Status: Deactivating software on secondary SP

Operation: Install

**# /nas/sbin/navicli -h 192.168.1.200 ndu -status**

Is Completed: NO

Status: Activating software on secondary SP

Operation: Install

**# /nas/sbin/navicli -h 192.168.1.200 ndu -status**

Is Completed: YES

Status: Operation completed successfully

Operation: Install

**# /nas/sbin/navicli -h 192.168.2.201 ndu -status**

Is Completed: YES

Status: Operation completed successfully

Operation: Install

**Note:** At this point, FLARE has been installed but not yet committed

15. Post Install Steps:

a.) Verify that Read & Write Cache has been automatically re-enabled on both SPs:

**# /nas/sbin/navicli -h 192.168.1.200 getcache**

SP Read Cache State Enabled

SP Write Cache State Enabled

**# /nas/sbin/navicli -h 192.168.1.200 setcache -wc 1**

**Note:** Use above command to re-enable Write Cache, though the system should automatically enable after NDU

b.) Verify & Enable Statistics Logging for both SPs, if required:

**# /nas/sbin/navicli -h 192.168.1.200 setstats -on**

Statistics logging is ENABLED

**# /nas/sbin/navicli -h 192.168.2.201 seststats -on**

Statistics logging is ENABLED

c.) Trespass Luns back to their Owner SPs using Celerra Manager>Celerras>Storage>Systems><clarion name>:System: Storage Processors: SP A Is Ready To Restore>Restore SP A

**Note:** Click on “Restore SP A” icon. It will restore LUNs for both SP’s, if both SP’s had trespassed luns.

d.) Alternatively, restore trespassed LUNs from the CLI:

**# /nas/sbin/navicli -h 192.168.1.200 getlun -trespass**

LOGICAL UNIT NUMBER 4

Default Owner: SP A

Current owner: SP B

LOGICAL UNIT NUMBER 5

Default Owner: SP A

Current owner: SP B

LOGICAL UNIT NUMBER 0

Default Owner: SP A

Current owner: SP B

**# /nas/sbin/navicli -h 192.168.1.200 trespass lun 5**

**# /nas/sbin/navicli -h 192.168.1.200 trespass lun 4**

**# /nas/sbin/navicli -h 192.168.1.200 trespass lun 0**

**# /nas/sbin/navicli -h 192.168.1.200 getlun -trespass**

**Note:** No output means that there are no trespassed luns on the backend

e.) Disable IP Forwarding on the Control Station & Verify:

**# echo 0 > /proc/sys/net/ipv4/ip\_forward**

**# cat /proc/sys/net/ipv4/ip\_forward**

0

f.) Commit the Flare Package to complete the NDU Upgrade & Verify:

**# /nas/sbin/navicli -h 192.168.1.200 ndu -list**

Name of the software package: FLARE-Operating-Environment

Revision of the software package: 03.22.040.5.511

Commit Required: YES

**# /nas/sbin/navicli -h 192.168.1.200 ndu -commit FLARE-Operating-Environment**

**# /nas/sbin/navicli -h 192.168.1.200 ndu -list**

Name of the software package: FLARE-Operating-Environment

Revision of the software package: 03.22.040.5.511

Commit Required: NO

**FLARE 24 PATCH 019 NS40 INTEGRATED NDU NAS 5.5.27.5:**

**# /nas/sbin/navicli -h 192.168.1.200 ndu -install /nas/var/dump/CX3-40-Bundle-03.24.040.5.019.pbu -delay**

**360**

Running install rules...

=====

Special Conditions: Rule has warning.  
Redundant SPs: Rule passed.  
All Packages Committed: Rule passed.  
No Trespassed LUNs: Rule passed.  
Statistics Logging Disabled: Rule failed.  
No Transitions: Rule passed.  
Acceptable Processor Utilization: Rule passed.

No Active Hot Spares: Rule passed.  
Version Compatibility: Rule passed.  
No Un-owned LUNs: Rule failed.  
No Active Replication I/O: Rule failed.  
No System Faults: Rule passed.  
Host Connectivity: Rule has warning.  
SP Event Log Verification: Rule failed.

**CORRECTIVE ACTIONS:**

→Disabled Statistics logging on both SPs  
**# /nas/sbin/navicli -h 192.168.1.200 setstats -off**

→Enabled IP Forwarding on Control Station

**# cat /proc/sys/net/ipv4/ip\_forward**

0

**# echo 1 > /proc/sys/net/ipv4/ip\_forward**

**# cat /proc/sys/net/ipv4/ip\_forward**

1

**# /nas/sbin/navicli -h 192.168.1.200 ndu -install /nas/var/dump/CX3-40-Bundle-03.24.040.5.019.pbu -delay 360**

Item number: 0

Name of the software package: FLARE-Operating-Environment

Revision of the software package: 03.24.040.5.019

Already Installed Revision 03.24.040.5.011

Installable YES

Disruptive upgrade: NO

Ndu Delay: 360 secs

The requested package(s) will be installed. Do you wish to proceed? : (y/n)? y

**NS41 SINGLE DM FRESH INSTALL PROCEDURE:**

**I. Install Differences between Single vs. Dual Data Mover configurations:**

→Basic NAS Installation of the single DM NS41 Integrated is different than the NS42

→Internal network cabling different than with dual Data Movers

→Fibre Channel cabling scheme is the same, minus the 2<sup>nd</sup> Data Mover

→A single Data Mover system does not populate the /nas/server/slot\_2/start file with a uniqueid, though the nas\_checkup incorrectly classifies this as a problem for NAS 5.5.27.5 and higher. See AR92218

**NAS CHECKUP FLAGS “UniqueID” PROBLEM:**

If running a NAS 5.3 system, and a check\_nas\_upgrade is run in preparation for upgrading to NAS 5.5, the checkup will flag a check\_unique\_id failure because 5.3 does not support the t2tty –C command that is being used to check the ID. With 5.5, it is also possible to have a uniqueid failure for a single Data Mover system, as the nas\_checkup command considers this a problem and flags it as an issue. In either case, there is no real problem or harm in proceeding with an actual NAS upgrade. As for a workaround for single DM systems, you could just add the following line to the bottom of the start file with the proper ID from the server:

**# .server\_config server\_2 -v "uniqueid"**

1186684020: SYSTEM: 4: uniqueid: "50060160:41e05335"

**# vi /nas/server/slot\_2/start**

**uniqueid validate=50060160:41e05335**

**Sys log:**

Apr 1 03:33:02 2007 Checkup:3:123 Scheduled nas\_checkup discovered an error.

**Nas checkup Log:**

2007-04-01T03:32:46-0400 dm::check\_unique\_id: (Fail)

→SP's use the same Internal IP addressing scheme and gateway as for the NS42, but network cabling is different

→The key difference is how the Celerra Internal Network is constructed and how it operates

→With the NS42 system, the Control Station serves the Primary Internal Network from eth0 at 192.168.1.100, between Data Movers and SPA (this remains unchanged for the NS41 system)

→With the NS42 system, the Control Station serves the Backup Internal Network from eth2 at 192.168.2.100, between Data Movers and SPB

**Note:** With the NS41 system, both Primary & Backup Networks are now served from the same physical interface (eth0 & an alias called eth0:0). In this arrangement, eth2 does not service the Backup network for the Celerra, but instead is reconfigured with the 192.168.2.102 address, along with a special routing entry in /etc/sysconfig/static-routes, in order for the Control Station to communicate back and forth to SPB. All Primary & Backup network activity, with the exception of SPB & eth2, are handled via eth0 or eth0:0.

#### **Physical Cabling of Internal Management Ports for NS41:**

- a. )Ethernet cable runs from lower management port of SPB to the port labeled 2 on the Control Station switch (upper middle port)
- b. ) Ethernet cable runs from lower management port of DM2 to the 10/100 port on CS0 (lower left port)
- c.) Ethernet cable runs from upper management port of DM2 to the lower management port on SPA

#### **Fibre Channel Cabling for NS41:**

- a. Data Mover blade 2 port BE 0 to SP A port 0 Fibre
- b. Data Mover blade 2 port BE 1 to SP B port 0 Fibre

#### **NS41 CS Interface Configuration, Routes, and other Important Files:**

# **/sbin/ifconfig -a**

```
eth0  inet addr:192.168.1.100  Bcast:192.168.1.255  Mask:255.255.255.0      →Primary Internal network (CS, DM, SPA)
eth0:0  inet addr:192.168.2.100  Bcast:192.168.2.255  Mask:255.255.255.0      →Backup Internal network alias (CS, DM)
eth1  <not used>
eth2  inet addr:192.168.2.102  Bcast:192.168.2.255  Mask:255.255.255.254 →SPB communications only (CS, SPB)
eth3  inet addr:10.241.168.94  Bcast:10.241.168.255  Mask:255.255.255.0      →External Interface customer network
```

# **cat /etc/sysconfig/static-routes**

```
eth2 host 192.168.2.201 dev
```

**Note:** This entry must exist in order for proper communication to occur between CS & SPB. Under normal circumstances, you should be able to ping between CS to SPs and all internal networks, navicli commands should run against either SP, etc.

# **cat /etc/sysconfig/network-scripts/ifcfg-eth0:0**

```
DEVICE=eth0:0
IPADDR=192.168.2.100
NETMASK=255.255.255.0
NETWORK=192.168.2.0
BROADCAST=192.168.2.255
ONBOOT=yes
```

# **cat /etc/sysconfig/network-scripts/ifcfg-eth0**

```
DEVICE=eth0
IPADDR=192.168.1.100
NETMASK=255.255.255.0
NETWORK=192.168.1.0
BROADCAST=192.168.1.255
ONBOOT=yes
```

# **cat /etc/sysconfig/network-scripts/ifcfg-eth2**

```
DEVICE=eth2
IPADDR=192.168.2.102
NETMASK=255.255.255.54
NETWORK=192.168.2.201
BROADCAST=192.168.2.255
ONBOOT=yes
```

# **cat nas\_device.map**

```
DOSDSK=/dev/nda
OS1DSK=/dev/hda
OS2DSK=/dev/hda
VERDSK=/dev/hda
VARDSK=/dev/ndf
NBSDSK=/dev/nde
ENET_INT0=eth0
ENET_INT1=eth0:0
```

**ENET\_INT2=eth2** →New entry to support NS41 communications with SPB

ENET\_EXT=eth3

ENET\_SP=

ENET\_IPMI=

# **cat nas\_enclosure.map**

#WARNING!! DO NOT MODIFY CONTENTS OF THIS AUTO-GENERATED FILE

ENCLOSURE-0\_MGMT-A\_IP=192.168.1.50

ENCLOSURE-0\_MGMT-A\_MAC=00:60:16:0b:1a:ca

ENCLOSURE-0\_MGMT-B\_IP= →Mgmt B is not used with single DM configuration

ENCLOSURE-0\_MGMT-B\_MAC= →Mgmt B is not used with single DM configuration

**Note:** Output abbreviated to show only Enclosure 0

## II. Backend Prep Work prior to Installation:

**Note:** If this is a factory install, this section can be skipped. For all field installs or re-installations, use this section to check the existing configuration on the backend, then take the appropriate action to prepare the system for a “fresh” NAS installation

1. Use a standard Ethernet cable to connect directly to the upper Clariion ethernet port on SPA from a Windows system (set an unused IP address on the Workstation (as shown below), then open a Web Browser/Navisphere Mgr session to SPA (192.168.1.200).
2. Once connected to SPA, unbind all NAS LUNs, then destroy the Raid Groups--this step should not normally be necessary if doing a Fresh NS40 Install.

### Workstation IP when connecting to SPA's Upper Management LAN Port:

192.168.1.95

255.255.255.0

192.168.1.100

DNS→De-select any entries that may be set

## III. Proceed with NAS Installation:

1. Connect to the Control Station’s front serial port from a Windows station, then use Hyperterminal to connect:

### Hyperterm Settings:

19200/8/None/1/None (use ANSI & Autodetect for connection)

2. Insert NAS CD & Boot Floppy, reboot CS—wait for prompt, then enter: boot: **serialinstall**

3. Linux install is fast & done in about 15 minutes--make sure to select Fresh Install when it detects the linux image on the IDE drive

4. NAS Installation takes an hour or so, answer prompts as needed. The Bind Operation for the Raid Groups took 4/5 hours.

**Note:** Completed upgrades result in /nas/log/upgrade log—interrupted or incomplete upgrades will have log in /tmp

## **NS41 INTEGRATED SINGLE DATA MOVER FLARE NDU UPGRADE PROCEDURE:**

1. Connect to front Control Station serial port to Workstation COM1 using Serial cable and open a Hyperterm session

2. Connect directly to SPA's Management LAN port using straight-through ethernet cable to Workstation

a.) Open Browser session to Celerra Manager

b.) Open Browser session to Navisphere Manager

3. Verify backend paths:

### **# nas\_storage -check -all**

Discovering storage (may take several minutes)

done

4. Using Celerra Manager, verify that the Celerra system is healthy

5. Using Navisphere Manager, verify that the CLARiiON array is healthy

**Note:** You may notice that SPB appears "unmanaged" when connected to SPA. Apparently this is some anomaly in certain flare versions. After upgrading to Flare 24, both SP's are correctly displayed.

6. Perform CLARiiON Healthcheck using Navicli:

### **# /nas/sbin/navicli -h 192.168.1.200 getlun –trespss**

**Note:** Check for any trespassed Celerra LUNs and restore to default Owner before proceeding

### **# /nas/sbin/navicli -h 192.168.1.200 networkadmin -get**

Storage Processor: SP A

Storage Processor Network Name: SPA

Storage Processor IP Address: 192.168.1.200

Storage Processor Subnet Mask: 255.255.255.0

Storage Processor Gateway Address: 192.168.1.100

### **# /nas/sbin/navicli -h 192.168.2.201 networkadmin -get**

Storage Processor: SP B

Storage Processor Network Name: SPB

Storage Processor IP Address: 192.168.2.201

Storage Processor Subnet Mask: 255.255.255.0

Storage Processor Gateway Address: 192.168.2.100

### **# /nas/sbin/navicli -h 192.168.2.201 getcrus** (checked for failed components)

7. Verify that CLARiiON Statistics Logging is enabled:

### **# /nas/sbin/navicli -h 192.168.1.200 setstats**

Statistics logging is ENABLED

**# /nas/sbin/navicli -h 192.168.2.201 setstats**

Statistics logging is ENABLED

**8. Verify that CLARiiON Front-end I/O throughput is < than 15,000 I/O's/sec using Celerra Manager**

Celerras>Storage>Systems>CLARiiON CX3-40 APM00063303725

**Front-End I/O Requests:**

Use I/O counter to the right of the graph: 21.5 I/O Req./sec

**9. Disable Statistics Logging on both SPs:**

**# /nas/sbin/navicli -h 192.168.1.200 setstats -off**

**# /nas/sbin/navicli -h 192.168.2.201 setstats -off**

**10. Verify that CLARiiON Back-end I/O throughput passes:**

**# /nas/tools/check\_clariion**

CPU UTILIZATION (MUST BE LESS THAN 50% PER STORAGE PROCESSOR)

THIS WILL TAKE APPROXIMATELY 1 MINUTE(S).

CLARIION: APM00063303725:

SP A ..... 000.65% (PASS)

SP B ..... 000.65% (PASS)

DISK I/O LOAD (MUST BE LESS THAN 100 I/O OPERATIONS PER SECOND PER DISK)

THIS WILL TAKE APPROXIMATELY 4 MINUTES.

CLARIION: APM00063303725:

DISK 0\_0\_0 ..... 002 OPS/SEC (PASS)

DISK 0\_0\_1 ..... 003 OPS/SEC (PASS)

DISK 0\_0\_2 ..... 002 OPS/SEC (PASS)

DISK 0\_0\_3 ..... 001 OPS/SEC (PASS)

**11. Stop NAS Services and manually remount NAS Partitions:**

**# /sbin/service nas stop**

**# ps -ef |grep nas\_m**

**# ps -ef |grep boxm**

**# df -h**

**# mount /nbsnas # mount /nas # mount /nas/dos # mount /nas/var**

**12. Put CS interface eth0:0 down & verify:**

**# /sbin/ifdown eth0:0**

**# /sbin/ifconfig -a**

**Note:** eth0:0 should not be represented in the output after running the ifdown command

**13. Enable IP Forwarding on the Control Station & Verify:**

**# echo 1 > /proc/sys/net/ipv4/ip\_forward**

**# cat /proc/sys/net/ipv4/ip\_forward**

**1**

**14. Disable CLARiiON Write Cache & Verify:**

**# /nas/sbin/navicli -h 192.168.1.200 setcache -wc 0**

**# /nas/sbin/navicli -h 192.168.1.200 getcache**

SP Read Cache State      Enabled

SP Write Cache State     Disabled

**15. Perform NDU and monitor Status from Control Station or Navisphere:**

**# /nas/sbin/navicli -h 192.168.1.200 ndu -install /home/nasadmin/flare/cx3-40-bundle-03.24.040.5.007.pbu**

**-delay 360**

Running install rules...

=====

Version Compatibility:      Rule passed.

Redundant SPs:            Rule passed.

-----abridged-----

Name of the software package:      FLARE-Operating-Environment

Revision of the software package:    03.24.040.5.007

Already Installed Revision        03.22.040.5.511

Installable                  YES

Disruptive upgrade:        NO

Ndu Delay:                360 secs

The requested package(s) will be installed. Do you wish to proceed? : (y/n)? y

**Note:** There is the possibility that the Rules Checking will prevent the NDU from proceeding. If absolutely certain that the Array is operating normally, add the "-skiprules" syntax to the end of the ndu -install command string

### # /nas/sbin/navicli -h 192.168.1.200 ndu -status

Is Completed: YES

Status: Operation completed successfully

Operation: Install

### 16. Post-Install NDU Upgrade Checks/Steps:

a.) Use Navisphere Manager or Navicli commands (getagent/getcrus/getcache/setstats, etc) to healthcheck CLARiiON backend

**Note:** Flare 24 has a fix so that both SPs will now register as “managed” when viewing the CLARiiON from Navisphere Manager

b.) Verify that Write Cache has been re-enabled on CLARiiON backend

### # /nas/sbin/navicli -h 192.168.1.200 getcache

SP Read Cache State Enabled

SP Write Cache State Enabled

**Note:** Write Cache automatically re-enables after a successful NDU

### # /nas/sbin/navicli -h 192.168.1.200 setcache -wc 1 -rca 1 -rcb 1

**Note:** Command to re-enabled Write Cache if necessary

c.) Trespass back any Celerra LUNs to their Owner SP, using either navicli trespass lun x or Celerra Manager:

### # /nas/sbin/navicli -h 192.168.1.200 getlun -tresspass

LOGICAL UNIT NUMBER 0

Default Owner: SP A

Current owner: SP B

### # /nas/sbin/navicli -h 192.168.1.200 trespass lun 0

d.) Re-enable Statistics Logging on both SPs:

### # /nas/sbin/navicli -h 192.168.1.200 setstats -on

### # /nas/sbin/navicli -h 192.168.2.201 setstats -on

### # /nas/sbin/navicli -h 192.168.2.201 setstats

Statistics logging is ENABLED

### # /nas/sbin/navicli -h 192.168.1.200 setstats

Statistics logging is ENABLED

e.) Commit flare version from navicli or Navisphere Manager & Verify:

### # /nas/sbin/navicli -h 192.168.1.200 ndu -commit FLARE-Operating-Environment

### # /nas/sbin/navicli -h 192.168.1.200 ndu -list

Name of the software package: FLARE-Operating-Environment

Revision of the software package: 03.24.040.5.007

Commit Required: NO

f.) Disable IP Forwarding on the Control Station & Verify:

### # echo 0 > /proc/sys/net/ipv4/ip\_forward

### # cat /proc/sys/net/ipv4/ip\_forward

0

g.) Bring eth0:0 back online & Verify:

### # /sbin/ifup eth0:0

### # /sbin/ifconfig -a

eth0:0 Link encap:Ethernet HWaddr 00:00:F0:9F:B3:74

inet addr:192.168.2.100 Bcast:192.168.2.255 Mask:255.255.255.0

h.) Restart NAS Services & Verify:

### # /sbin/service nas start

### # ps -ef |grep nas\_mcd

```
root 3905 1 0 13:39 ? 00:00:00 /nasmcd/nas_mcd -h /nasmcd /nas/
root 4055 3905 0 13:39 ? 00:00:00 /nasmcd/nas_mcd -h /nasmcd /nas/
```

### # ps -ef |grep boxm

```
root 5018 3905 0 13:40 ? 00:00:00 /nas/sbin/nas_boxmonitor /nas -i
root 5141 5018 0 13:40 ? 00:00:00 /nas/sbin/nas_boxmonitor /nas -i
```

-----both NAS & Box Monitor process output abridged for brevity-----

## **ADDING ARRAY ENABLERS:**

→Can be done using the NST/USM or via CLI

# /nas/sbin/navicli -h 10.241.168.179 ndu -install CompressionEnabler-01.01.5.001-fleet64\_free.ena -delay

**360**

Running install rules...

## **CHANGING PRIVATE IP ADDRESS NETWORK ON NS40 INTEGRATED CELERRA:**

**Note:** See current NAS 5.5 NS40 Integrated Setup Guide, but note that they forgot to include a section for changing SP addresses

1. Create /root/BACKUP directory
2. Copy /etc/hosts, /etc/sysconfig/network-scripts/ifcfg-eth\*, /etc/nas\_device.map, /etc/nas\_enclosure.map, /etc/encl\_dhcpd.conf, /etc/sysconfig/network & static-routes, /nas/site/nas\_param & slot\_param, and /nas/server/slot\_2/ files to /root/BACKUP
3. Stop NAS Services & manually remount NAS partitions
4. Update Secondary Network first:

**# /nasmcd/sbin/setup\_enclosure –changeSubnet 1 192.168.102.0**

**Note:** Runs CheckCable, updates ENCL\_DB & DHCPD\_CFG files, resets secondary management switches

5. Verify change:  
# /sbin/ifconfig eth2
6. Use following command to create updated files for Secondary network, review for accuracy, then copy over the original files:  
# cat /etc/hosts | sed 's/192.168.2./192.168.102./g' > /tmp/<filename>  
**Note:** Repeat process for /nas/site/nas\_param, /nas/site/slot\_param
7. Determine the list of files that require editing for each Data Mover slot, using following:  
**# fgrep -e ‘192.168.2.’ `find /nas/server/slot\_2/ -type f -maxdepth 1 |grep -v “tab\$”`**
8. Edit all files identified in Step 7 with the new subnet for the Secondary network  
[export, eof, nbs.cs.ro, nbs.cs.rw, ifconfig, etc.]
9. Export NAS DB and reboot the Server:  
# export NAS\_DB=/nas; server\_cpu server\_2 -r now  
**Note:** If system hangs, exit and reconnect to Control Station, then # ping server\_2b to verify backup network

10. Repeat sed replacement steps for the Primary network, review for accuracy, then replace original files:

**# cat /etc/hosts | sed ‘s/192.168.2./192.168.101./g’ > /tmp/<filename>**

**Note:** Repeat steps for nas\_param & slot\_param files

11. Manually unmount all NAS partitions  
**Note:** Use # /sbin/fuser -m /nbsnas and # kill -9 <pid> if needed to umount partitions
12. Stop NBS service:  
# /sbin/service nbs stop
13. Mount NAS: # mount /nas
14. Update Primary Network using setup\_enclosure command:

**# /nasmcd/sbin/setup\_enclosure –changeSubnet 1 192.168.101.0**

15. Verify using: # /sbin/ifconfig eth0
16. Umount NAS: # umount /nas
17. Restart Network Services: # /sbin/service network restart
18. Restart NBS Services: # /sbin/service nbs start
19. Remount all NAS partitions and re-export NAS DB
20. Perform Setup Slot on each Data Mover: # /nas/sbin/setup\_slot -i 2
21. Run Repair Command:

**# /nasmcd/sbin/setup\_enclosure –checkSystem autoRepair**

22. Restart NAS Services: # /sbin/service nas start
23. Change the Clariion IP Addresses using the following steps:
  - a. Using a straight-through Ethernet cable, and a valid IP on the 192.168.1 network, connect a Workstation to the upper LAN port on SPA
  - b. Open Web Browser on Workstation and access SPA’s setup program: 192.168.1.200/setup
  - c. Change SPA IP address to new subnet (Value in the last Octet must remain .200 for SPA & 201 for SPB)

**Note:** SP will automatically reboot. Wait until SPA comes back up before continuing.

  - d. Repeat above steps for SPB
  - e. From the Celerra Control Station, manually trespass luns back to their Owner SP using navicli  
# /nas/sbin/navicli -h <sp\_IPaddr> trespass lun x
  - f. Manually update the /etc/hosts file with the new Subnet addresses for SPA & SPB
  - g. Run # nas\_storage -sync <APM\_id> to update the symapi db with the revised SP IP addresses

**Note:** There was an issue where CS reboots would reset the SP addresses in the /etc/hosts file to the original values prior to the changes. The corrective action here would be to edit the /nas/site/sp\_info file and update the SP addresses to the correct values. See AR92483.

## **EXAMPLE OF sp\_info FILE:**

# cat /nas/site/sp\_info

192.1.4.214 A\_APM00071600514 SPA # CLARiiON SP

192.1.4.215 B\_APM00071600514 SPB # CLARiiON SP

### **REVERTING FLARE CODE WHEN FLARE HAS NOT YET BEEN COMMITTED:**

The process of reverting to the previous Flare version is fairly simple and straightforward, provided the Flare version that you upgraded to has not yet been committed. Rebooting an SP while in an “uncommitted” state does not revert the Flare version.

**Navisphere Manager>APM0063303725[CX3-40]>Properties>Software>FLARE-Operating-Environment**

**03.24.040.5.006 (Commit required):** Highlighted the Flare 24 uncommitted code, then selected “Revert” button and the system automatically goes back to the last running Flare version, which in this example was Flare 22 Patch 511.

### **USING CLI COMMAND TO REVERT FLARE (If not yet Committed):**

**# /nas/sbin/navicli -h 192.168.1.200 ndu -revert FLARE-Operating-Environment -delay 360**

Revert operation will revert

    FLARE-Operating-Environment

from both SPs with Ndu delay of 360 secs. Do you still want to continue. (y/n)? y

**# /nas/sbin/navicli -h 192.168.1.200 ndu -status**

Is Completed: NO

Status: Initializing

Operation: Revert

### **NAS ‘UPGRADE’ OF ENGINEERING CODE RELEASE FAILS (5.5.22.4028 to 5.5.22.5028):**

**# EMC/nas/setup**

Setting up system, please wait...

Model: NS600G

Upgrading From Version: 5.5.22-40130

Mon Jul 10 12:11:16 EDT 2006

./setup: Error: Downgrade detected. Attempt to downgrade NAS from version 5.5.22-40130 to version 5.5.22-5028 not permitted.

Run 'emcrpm --erase emcnas' to remove the previous package.

Solution: Create following file to allow for upgrade to succeed

**# touch /tmp/.allow\_downgrade**

### **NAS UPGRADE WHEN SYSTEM HARDWARE IS OUT OF REV: e.g., 507**

**# EMC/nas/setup –allow\_upgrade\_on\_obsolete\_dm**

**Note:** Purpose would be to upgrade to a code level that would support new hardware, for example

### **TROUBLESHOOTING UPGRADE ISSUES ON LINUX CONTROL STATION:**

**# rpm -qa** #rpm -qa lwc -l [Lists out Software Packages & is a good check to ensure that RPM database is o.k.]

**# rpm -a -q --last** [Lists out most recently installed packages by date]

**# rpm -qa “\*ftp”**

#### **ADDING LINUX PACKAGES:**

**# rpm -Uvh [ftp://192.168.0.254/pub/RedHat/RPMS/sendmail\\*](ftp://192.168.0.254/pub/RedHat/RPMS/sendmail*)**

#redhat-config-packages

#### **ADDING ERRATA TO ALL PACKAGES:**

**# rpm -Fhv [ftp://updates.redhat.com/current/en/os/i386/\\*.rpm](ftp://updates.redhat.com/current/en/os/i386/*.rpm)**

#### **UPGRADING KERNEL:**

1. **# rpm -ivh kernel-version.i386.rpm**

2. Reboot system and test new kernel

3. Set ‘default’ line in /boot/grub/grub.conf to 0 to load new kernel upon reboot

4. Remove old kernel if satisfied: #rpm -e old.kernel-version.i386.rpm

#### **REMOVING RPM’s:**

**# rpm -e xpdf**

#### **LOCAL RPM DATABASE:**

**/var/lib/rpm**

#### **QUERYING SPECIFICS OF A PKG:**

**# rpm -q zsh -i | -l #rpm -qa bash -i #rpm -qf /etc/imrc**

#### **EXTRACTING EXECUTABLES FROM ZIP TO LOCAL DIRECTORY:**

## **VERIFYING NASDB VERSIONS NAS 5.5+:**

**# /nas/sbin/nasdb\_tools/db\_status**

Current DB seems to be at version 3

**Note:** Command came into being to support RLL (Row Level Locking), introduced with NAS 5.5

## **DOWNLOADING & INSTALLING AN RPM PACKAGE ON LINUX CONTROL STATION:**

ftp://ftp.samba.org/pub/rsync/binaries/redhat/

ftp> mget rsync-2.5.7-1.i386.rpm

mget rsync-2.5.7-1.i386.rpm? y

**# rpm -i rsync-2.5.7-1.i386.rpm**

**# rsync**

rsync version 2.5.7 protocol version 26

## **REPAIRING EMCNAS RPM DATABASE ON CS DUE TO DESTRUCTIVE LINUX RECOVERY:**

### **Symptoms:**

**\$ nas\_version -l**

error: cannot open Packages index using db3 - No such file or directory (2)

**# ls -al /var/sadm/pkg/emcnas/**

```
drwxr-xr-x 2 root root 4096 Oct 13 19:16 install
-rw-r--r-- 1 root root 362 Oct 13 19:16 pkginfo
drwxr-xr-x 2 root root 4096 Oct 13 19:16 save
```

**Note:** 'Packages' file and others are missing

**Problem:** EMCNAS RPM database is corrupted and requires repair per the following steps—emc95561

1. Copy RPM from NAS 5.1.20.401 CD to /home/nasadmin on Control Station (or run directly from CD):

**# cp /mnt/cdrom/EMC/nas/emcnas-5.1.20-401.i386.rpm /home/nasadmin**

2. Restore emcnas RPM database using the following command on Control Station, specifying destination path and RPM source location:

**# rpm -U --justdb --force --nodeps --relocate /nas=/nas --dbpath /var/sadm/pkg/emcnas/ /home/nasadmin/emcnas-[0-9]\*.rpm**

3. Verify that database has been repaired:

**# ls -al /var/sadm/pkg/emcnas/**

```
-rw-r--r-- 1 root root 331776 Oct 14 13:01 Basenames
-rw-r--r-- 1 root root 24576 Oct 14 13:01 Dиримес
-rw-r--r-- 1 root root 12288 Oct 14 13:01 Group
drwxr-xr-x 2 root root 4096 Oct 13 19:16 install
-rw-r--r-- 1 root root 8192 Oct 14 13:01 Installtid
-rw-r--r-- 1 root root 12288 Oct 14 13:01 Name
-rw-r--r-- 1 root root 675840 Oct 14 13:01 Packages
-rw-r--r-- 1 root root 362 Oct 13 19:16 pkginfo
-rw-r--r-- 1 root root 12288 Oct 14 13:01 Providename
-rw-r--r-- 1 root root 8192 Oct 14 13:01 Provideversion
-rw-r--r-- 1 root root 12288 Oct 14 13:01 Requirename
-rw-r--r-- 1 root root 16384 Oct 14 13:01 Requireversion
drwxr-xr-x 2 root root 4096 Oct 13 19:16 save
```

**# nas\_version -l**

|                                |                                      |
|--------------------------------|--------------------------------------|
| Name : emcnas                  | Relocations: /nas                    |
| Version : 5.1.20               | Vendor: EMC                          |
| Release : 401                  | Build Date: Thu 15 Jan 2004 08:07:51 |
| AM CST                         |                                      |
| Size : 291635065               | License: EMC Copyright               |
| Packager : EMC Corporation     |                                      |
| URL : http://www.emc.com       |                                      |
| Summary : EMC nfs base install |                                      |
| Description :                  |                                      |
| EMC nfs base install           |                                      |

## **5.4 UPGRADE CHANGES:**

→ If rootfs size is not already 128MB, the upgrade will auto-extend to 128MB in size

→ root\_panic\_x partition will be created if it does not already exist (8MB)

→ Server Log has been increased from 1MB to 2MB and uses space from the work partition, root\_rdf\_channel

## **NAS UPGRADE USING NAS.EXE PATCH ONLY:**

**1. Login to Celerra, create directory for NAS Image, change to this directory:**

```
# mkdir /home/nasadmin/code;cd /home/nasadmin/code
```

**2. Connect to EMC FTP Server:**

```
# ftp 168.159.216.19 or #ftp ftp.emc.com
```

**3. At prompt, enter username of "ftp" or "anonymous" and then enter a valid email address as the password:**

```
Name (168.159.216.19:nasadmin): ftp
```

```
331 Guest login ok, send your complete e-mail address as password.
```

```
Password: joe@emc.com
```

**4. Set FTP transfer to Binary mode:**

```
ftp>bin
```

```
200 Type set to I.
```

**5. Change to the appropriate directory to obtain NAS Patch as outlined above & Verify presence of Patch:**

```
ftp> cd /outgoing/code
```

```
250 CWD command successful.
```

**Note:** Where "/outgoing/code" is an example of a patch for systems running NAS 5.1

```
ftp> ls -la
```

```
-r----- 1 g3 4920553 Jan 11 12:40 5120401.exe.gz
```

**6. Download Appropriate Gzipped NAS Version & Log off the FTP Server:**

```
ftp>get 5120401.exe.gz
```

```
ftp> bye
```

**7. Verify size of zipped NAS image and unzip:**

```
# ls -la /home/nasadmin/code/5120401.exe.gz
```

```
-rw-rw-r-- 1 nasadmin nasadmin 5423122 Jan 11 19:46 5120401.exe.gz
```

```
# gzip -d 5120401.exe.gz
```

```
# ls -la 5120401.exe
```

```
-rw-rw-r-- 1 nasadmin nasadmin 18913180 Jan 11 19:46 5120401.exe
```

**8. Determine NAS Version Running on Celerra for All Servers:**

```
# server_version ALL
```

```
server_2 : Product: EMC Celerra File Server Version: T5.1.18.804
```

```
server_3 : Product: EMC Celerra File Server Version: T5.1.18.804
```

```
server_4 : Product: EMC Celerra File Server Version: T5.1.18.804
```

```
# cat /nas/dos/slot_*/boot.bat
```

```
gload \bin\nas.exe boot.cfg
```

```
gload \bin\nas.exe boot.cfg
```

```
gload \bin\nas.exe boot.cfg
```

**9. Rename the current running "nas.exe" file located in the /nas/dos/bin directory:**

```
# ls -la *.exe
```

```
-rwxrwxr-x 1 root nasadmin 18701537 Dec 29 13:25 5118807.exe
```

```
-r-xr-xr-x 1 root nasadmin 42373 May 13 2003 gload.exe
```

```
-rwxrwxr-x 1 root nasadmin 32208 May 13 2003 loadlin.exe
```

```
-rwxrwxr-x 1 root nasadmin 18722612 Nov 20 13:12 5119504.exe
```

```
-rwxrwxr-x 1 root nasadmin 18722612 Jan 7 10:54 nas.exe
```

```
# mv /nas/dos/bin/nas.exe /nas51804.exe
```

**Note:** nas.exe is renamed to reflect the original Version Number of the patch.

**10. Run the following command to make sure that there is enough space (at least 20MB) in the /nas/dos directory to accomodate the new patch--you may need to remove older nas images:**

```
# df -h
```

| Filesystem | Size | Used | Avail | Use% | Mounted on |
|------------|------|------|-------|------|------------|
| /dev/sda3  | 1.9G | 1.2G | 607M  | 67%  | /          |
| /dev/sda1  | 7.6M | 3.5M | 3.7M  | 49%  | /boot      |
| none       | 251M | 0    | 250M  | 0%   | /dev/shm   |
| /dev/sdc1  | 1.7G | 649M | 1.0G  | 39%  | /nas       |
| /dev/sdb1  | 133M | 97M  | 37M   | 73%  | /nas/dos   |

**11. Copy the newly downloaded and unzipped patch into the /nas/dos/bin directory and rename it as "nas.exe":**

```
# cp -ip /home/nasadmin/code/5120401.exe /nas/dos/bin/nas.exe
```

```
-rwxrwxr-x 1 root nasadmin 18913180 Jan 11 20:15 nas.exe
```

**12. Verify Size and Permissions on nas.exe:**

```
# ls -la /nas/dos/bin/nas.exe
```

```
-rwxrwxr-x 1 root nasadmin 18913180 Jan 11 20:15 nas.exe
```

**13. Run server setup command to point all servers to the new "nas.exe" by updating the /nas/dos/slot \*/boot.bat gload utility:**

**# /nas/sbin/server\_setup ALL -L nas.exe**

**Note:** Where "nas.exe" is the new image.

**14. Verify that the correct image is being pointed to:**

```
# cat /nas/dos/slot_*/boot.bat  
gload \bin\nas.exe boot.cfg
```

**15. Reboot the Standby Data Mover first and Verify that the new NAS Version has Loaded:**

```
# server_cpu server_4 -r now  
# server_version server_4  
server_4 : Product: EMC Celerra File Server Version: T5.1.20.401
```

**16. To Load the New NAS Image, Failover the Production Data Mover to its Standby, then Failback:**

```
# server_standby server_2 -a mover  
# server_standby server_2 -r mover
```

**Note:** Repeat for each DM that requires the new image. Don't forget to reboot the primary data mover slot after failing the Server over to its Standby, before failing back.

**PROBLEMS WITH NAS 5.1/5.2/5.3 SOFTWARE/HARDWARE INSTALLS/UPGRADES:**

**UPGRADING TO NAS 5.3 PROBLEMS AND NAS 5.3 INSTALL ISSUES:**

**1. GENERIC NAS UPGRADE PROBLEMS:**

Problem: Missing "nas.exe" reference in boot.bat file. NAS 5.3 Upgrades needs to be able to run a script to help it identify the enclosure number for the data movers—this was necessary because we now have (4) Data Mover NS systems. In any event, the Upgrade creates a /nas/dos/slot\_99 directory, copies the boot.bat from slot\_2 to slot\_99 and creates a minimal boot.cfg to boot DM to reason code 14. Next part of the script issues resumeprom commands to verify enclosure\_id for DM\_2 only—once determined, DM reboots from its correct slot and upgrade completes (this is where the Upgrade will stall if the correct nas.exe cannot be located that contains the resumeprom commands). An enclosure\_id of 0 indicates slot\_2 or 3 and an id of 1 indicates slot\_4 or 5

**CTAPTTY Log Error:** -->/var/log/ctappty\$4.log [These logs are new to 5.3 & record DM boot process, for NS systems only]  
097961095: SYSTEM: 3: uniqueid: values do not match "LKE00040400134" vs "LKE00032400095"

**UPGRADE FAILURE SYMPTOMS—NAS 5.2.11 to 5.3.11:**

Waiting for a valid reason code on slot 2...14:done

Configuring Enclosure ID...:Failed

14:done

---

slot\_99 and Reason Code 14

**PXE BOOT & remove LKE line from /nas/server/slot\_x/start file:**

```
$ cat start  
uniqueid validate=LKE00035300071  
# .server_config server_2 -v "uniqueid"  
1175859583: SYSTEM: 4: uniqueid: "LKE00034200368"
```

**Note:** Above output from an NS700G with two Data Movers. The uniqueid is not used for single Data Mover systems, and is used for NS and NSX systems only as a way to properly identify enclosure IDs.

**Run setup\_slot to fix MAC Address & Serial Number issues:**

**Verify that SYMCLI was upgraded and is linked correctly:**

—following indicates that SYMCLI was not upgraded correctly:

```
# find / -name "symcli*" -print  
/nas/opt/emc/SYMCLI/V5.4.0/bin/symcli  
/nas/opt/emc/SYMCLI/V5.4.0/man/man1/symcli.1  
/nas/opt/emc/SYMCLI/V5.4.0/man/man1/symclient.1  
/nbsnas/opt/emc/WideSky/V5.3.0/bin/symcli  
/nbsnas/opt/emc/WideSky/V5.3.0/man/man1/symcli.1  
/nbsnas/opt/emc/WideSky/V5.3.0/man/man1/symclient.1
```

**2. DM's REMAIN IN STATE 3 LOADED NAS 4.2-to-5.2 UPGRADE:**

Data Movers remain in 'LOADED' state after NAS 4.2 to 5.2 Upgrade:

- Connect monitor to DM port and observe boot-up
- In this particular case, DM trying to locate its internal interface MAC address and cannot

**Solution:**

- Go to /nas/server/slot\_x/ifconfig & ifconfig.int files and remove following portion of the lines that reference internal interface MAC addresses: **mac=8:0:1b:43:b:63**
- Reboot data movers, should now come up to contacted state
- Conduct setup\_slot -i to complete proper backend discovery and initialization

**3. MPD CONVERSION FAILS TO COMPLETE:**

- a. Load NAS 5.2.18.2 or 5.3.11.403 or higher onto Standby Server
- b. Convert Standby server to a Primary
- c. Conduct FSCK/ALCHCK of file systems that will not convert
- d. Permanently unmount and remount affected file system to new Standby (Primary) Server
- e. Complete MPD conversion of affected file systems
- f. Mount file system back on original Server and reconfigure Standby Server

#### **4. AFTER NAS 5.3.11.4 INSTALL, SERVER COMMANDS FAIL:**

--one possible reason is that a configuration file contains an extra line at the end of the file [e.g., nas\_param file]

#### **5. NAS 5.2 GOLDEN EAGLE UPGRADE FAILS:**

If the root password on CS0 is changed using the WebUI, the password is DES encrypted. Normal passwd encryption uses MD5. If this situation has happened, when performing the NAS upgrade, upgrade will fail with “Password incorrect – please retry”. Change passwd back to MD5 encryption by using #passwd -S root

#### **6. UPGRADE 4.2-to-5.2 FAILS WITH WRONG NAS DEVICE.MAP:**

/etc/nas\_device.map.1.rev2--Upgrade script may overwrite the nas\_device.map file with incorrect information:

‘VERDSK=/dev/sde’—as in this example:

**CRITICAL FAULT:**

Inquiry failed for dev /dev/sde

**Solution:** Copy the nas\_device.map.1.rev1 file from the NAS Upgrade CD to /etc, contains ‘VERDSK=/dev/sdc’ [See emc97104]

#### **7. BIOS/POST INCOMPATIBILITY ISSUES 5.3 INSTALLS:**

##### **NS602G NAS 5.3 Install Failure:**

BIOS level of data movers was so old that the latest NAS code would not install correctly because the data movers could not boot from the backend DOS partition to complete the flash BIOS/POST upgrade. Workaround in this case was to take physical data mover and place into functioning down-rev 5.1 or 5.2 system and boot, which will flash upgrade—DM can then be used for the install on new system. Alternatively, order additional DM to replace with newer mover. In one case, the install completed on DM\_3, and DM\_3 was then used to complete the install on slot\_2—once completed, DM\_2 could be inserted in slot\_2 and setup\_slot completed.

**Note:** BIOS & POST version changes in mfg do not lead to purging of SPARES inventory—the auto-flash upgrade procedure is supposed to be able to handle this, even though we all know that it doesn’t. See AR’s 52344 & 54671.

#### **8. POST VERSION 3.37 CAUSES INSTALL FAILURES:**

Data Movers with POST version 3.37 cannot use port BE0 to boot from Backend, forcing a boot from BE1, and progressing to Reason Code 14 state, causing Install to fail. POST 3.38 or higher will resolve this problem.

#### **EXAMPLE BIOS/POST INCOMPATIBILITY FOR 5.3 INSTALLS:**

**Problem:** Down-rev BIOS & POST prevents NAS 5.3 Install from Succeeding

##### **Manual Flash Upgrade Procedure:**

→Insert problem Data Mover into functional 5.1 NS600 System

→Connect to Data Mover via HyperTerminal Session and wait for “FLASH UPGRADE UTILITY 5.7” screen to appear

##### **POST Upgrades first, followed by BIOS:**

C2\_POST.BIN <Rev. 02.30> <Rev. 03.12>

C2\_BIOS.ROM < 3.09 > < 3.26 >

→Let system reboot to complete flash upgrade, then remove hardware from system

→Installation on 5.3 system will now succeed

#### **9. NAS 5.3 FAILED INSTALL DUE TO PXE BOOT:**

##### **NAS 5.3 Install Failure on NS600G Symmetrix Backend:**

Data Mover\_2 fails to boot from the Backend using BE0, even though the POST & BIOS, Zoning, WWN’s, and Control Luns were correct. Makes sure BE0 logs into Switch and to the Clariion. In this situation, the setup\_slot starts the PXE services and the data mover will PXE boot if it cannot boot from Backend.

**Solution:** SSH to Control Station and kill the dhcpcd daemon so that PXE booting will not occur, then retry setup\_slot install.

#### **10. NAVISPHERE NOT REGISTERING DM PATHS TO SPs:**

In rare circumstances, some of the paths from the DM’s to the SP’s will get registered in Navisphere, but fail to complete the Fibre login when the first IO is sent from the DM to the CX. If this occurs, the fix would entail the following steps:

→Enter Engineering mode on Array using CTRL, SHIFT, F12 and use password “messner”

→Rightclick Storage Group for control luns, select ‘properties’>Hosts>Click Advanced

--Examine each DM for registered paths & check off the small box in front of the registered path statement on the paths not checked

--Update Navisphere and all 8 paths should register and login properly

→Navisphere Manager>rightclick SPA/SPB>properties>ALPA>verify that SCSI ID is set to 0 on all SP ports

#### **11. NAS 5.3 REASON CODE 14 ISSUES:**

- a. During DM upgrade, cannot boot from BE0—so PXE boots from IDE drive of CS0 and sits at reason code 14

**Causes:** Bad fibre switch, connection, zoning, speed settings across DM, Switch, BackEnd

- b. Unable to determine slot number (Review CTAP logs to see what slot number is being used on bootup)

c. Unable to bootup due to MAC address conflict (Do setup\_slot to correct MAC address)

d. Serial number conflict in config file (/nas/server/slot\_x/start)

**Correction:** Check start file serial # against /nas/log/data\_mover\_resume.server\_x.xml. If different, delete start file line and reboot

e. Examine CTAP logs for other indicators

f. NS700 or NS700G cannot establish enclosure ID number

**Causes:** Enclosure ID concept added for (4) DM NS Models—ensure DM's are booting from correct “nas.exe” image

g. Reason Code 14 on NAS Upgrade and Data Mover boots to c:\slot\_99 directory on Control Station:

CTAP Log shows error: Unable to determine slot number

**Note:** Upgrade script needs to be able to program the Enclosure ID number. If this cannot be determined, the t2slot.exe cannot determine its slot number and boots from slot\_99 by default.

#### **Possible Issues:**

a. Verify that data mover is pointing to nas.exe by checking boot.bat file

b. Verify that data movers 1 & 2 are pointing to correct Enclosure ID by running “resumeprom show”

**# .server\_config server\_3 -v "resumeprom show"**

NAS ENCLOSURE ID 0

**Note:** You may need to run the command via minicom session at the data mover’s CONSOLE>

c. Setting Enclosure ID’s to 0:

**# .server\_config server\_3 -v "resumeprom set xpeid=0"**

d. If all of these steps do not resolve the issue, go to FCCBOOT menu via serial session and do a Controller Reset on each DM

**“FCC Boot Sub-Menu”**

Select Option 1 to ‘Reset Controller’

## **12. NAS 5.2 to 5.3 UPGRADE FAILURE GOLDEN EAGLE CAB:**

Upgrade can fail if boot.bat is not updated properly

#### **Pre-Upgrade:**

\$ cat /nas/dos/slot\_2/boot.bat

gload -or c:\bin\nas.exe c:\slot\_2\boot.cfg

#### **After Upgrade:**

\$ cat /nas/dos/slot\_2/boot.bat

gload -or c:-m 2048 -or c:\bin\nas.exe c:\slot\_2\boot.cfg

## **13. DM FAILS TO BOOT PAST REASON CODE 3 AFTER NAS UPGRADE TO 5.3.14:**

AR56446 identifies issue with C compiler in which certain code was generated with different typecast, resulting in memory corruption problem for TOE 514 Data Movers. Using the following workaround allowed DMs to boot—fixed 5.3.18 & 5.4.17:

gload -or c:-m 2047 -or c:\bin\nas.exe c:\slot\_5\boot.cfg

## **NS SERIES ENCLOSURE ID:**

Enclosure ID of Data Movers is set during installation & is burned into eeprom, based on the serial connections to the DMs—tty4 & tty5 equate to Enclosure ID 0, while tty6 & tty7 equate to Enclosure ID 1. Furthermore, a pin on the backplane tells which data mover is in position A or B in the enclosure (0 or 1). The enclosure ID number, plus Hardware slot number from pin, are used to determine the slot number (setup\_enclosure utility in NSX). During upgrades, t2slot.exe is used to determine enclosure ID number for DM—if it cannot be determined, then data mover will boot a minimum config from slot\_99 and end up in Reason Code 14 status. In many cases, a setup\_slot can be run to correct issues. Check ctap logs for errors.

## **NAS 5.4/5.5/5.6/6.0 UPGRADE/INSTALLATION ISSUES:**

WORKPART MAGIC STRUCTURE BY CODE VERSION:

#### **NAS 5.3:**

workpart\_magic = 0xfacebeaf

#### **NAS 5.4 & 5.5:**

workpart\_magic = 0xbeaf0001

#### **NAS 5.6:**

workpart\_magic = 0xbeaf0002

**Note:** Please note that all 5.4 upgrades that do not complete have an upgrade.log in /var/tmp

#### **NAS 6.0:**

workpart\_magic = 0xbeaf0002

1. Installation with Symm backend may fail if data luns are presented. Data Movers will boot to RC4 ‘Dart Ready’ state by are unable to complete install when “camshowconfig” is run. See AR65121. Solution is to remove all data luns and use Control Luns and Gatekeeper luns only for install.

2. NAS 5.4 upgrades root\_log slices, which triggers workpart update. The workpart update fails to populate bind tables, meaning that they are then cleared and recreated. Can cause problems with systems with multiple backends or FA’s that use same target/lun for different volumes—scsi chains can become misaligned in new binding table.

3. Upgrade fails to complete (evidence is lack of upgrade log in /nas/log) and workpartition has been deleted:

**# /nas/sbin/workpart -r**

Read Work Partition layout from LBA 0x43000, 528384 bytes.

After read:

Work Partition Structure not found!

134

**Apply Primus EMC113946:**

1. Recover work partition:

**# /nas/sbin/workpart -o wpart**

134

**# /nas/sbin/workpart -r**

Read Work Partition layout from LBA 0x43000, 528384 bytes.

After read:

Dump of Work Partition Structure (size 528384 bytes):

workpart\_magic = 0xfacebeaf (workpart\_magic indicates system prior to 5.4)

dos\_slice.lba = 0x0

-----output abridged-----

2. Verify that workpartition is repaired:

**# /nas/sbin/check\_workpart\_and\_log\_slice\_consistency**

./check\_workpart\_and\_log\_slice\_consistency:

CRITICAL FAULT:

Error: Inconsistent workpart information

**Note:** If you get the above error, follow the rest of the steps outlined in Primus to complete NAS 5.4/5.5 upgrade

**# ./check\_workpart\_and\_log\_slice\_consistency**

Workpart layout is up to date, done

**# /nas/sbin/workpart -c**

Latest Version Workpart Layout Found!

134

## **UPGRADING NAS STORAGE API AFTER UPGRADE FAILURE NAS 5.4:**

**Note:** If NAS 5.4 upgrade fails with following message, download the latest nasStorageAPI and install

**Backend Storage Requirements Check Failed:**

-----  
INSTRUCTIONS: Upgrade NAS Storage API

1. Download nasStorageAPI from TS2 website (separate API required for primary CS and Standby CS)
2. Upgrade RPM: #rpm --upgrade nasStorageAPI-0.0-6.i386.rpm
3. Query to verify: **# rpm --query --info nasStorageAPI**

## **CREATING BOOT FLOPPY FROM NAS CD-ROM ON NAS 5.6 CONTROL STATION:**

**NAS 5.6 Mountpoint Changes on the Control Station:**

**/media/floppy**

**/media/cdrom** [for CD-ROM drives, ReadOnly]

**/media/cdrecorder** [for CD-RW drives, such as the NSX]

**Symptoms if using wrong mountpoints:**

**# mount -t msdos /dev/fd0 /mnt/floppy**

mount: mount point /mnt/floppy does not exist

**# mount /dev/cdrom /mnt/cdrom**

mount: mount point /mnt/cdrom does not exist

**NAS 5.6 Floppy Image Name Change on NAS CD-ROM:**

/mnt/cdrom/images/boot.img (NAS 5.5 or below)

/media/cdrom/images/floppy.img (NAS 5.6 and above)

**NAS 5.6 Boot Floppy Image:**

Since the floppy image contains no recognizable file system that Linux or Windows can detect, the only way to verify the boot floppy image is to use cksum

**Boot Floppy Creation & Verification Procedure:**

**1. Mount CD-ROM & Floppy on Control Station:**

**# mount -r /media/cdrom**

**# mount -t msdos /dev/fd0 /media/floppy**

**# mount**

```
/dev/hdd on /media/cdrom type iso9660 (ro,nosuid,nodev)  
/dev/fd0 on /media/floppy type msdos (rw)
```

**2. Create boot floppy system disk:**

```
# dd if=/media/cdrom/images/floppy.img of=/dev/fd0 bs=1024
```

1440+0 records in

1440+0 records out

**Note:** Eventhough the command prompt returns quickly, it takes a few minutes to complete the image write to the Floppy disk

**3. Use checksum to verify the image built to floppy vs. the image on the NAS CD Media:**

```
# dd if=/dev/fd0 of=/tmp/bootflop bs=1024
```

```
# ls -la /tmp/boot*
```

```
-rw-r--r-- 1 root root 1474560 Jan 31 14:36 /tmp/bootflop
```

```
# cksum /tmp/bootflop
```

```
1683387514 1474560 /tmp/bootflop
```

```
# cksum /media/cdrom/images/floppy.img
```

```
1683387514 1474560 /media/cdrom/images/floppy.img -->Checksums match on the file named "bootflop" vs. "floppy.img"
```

**4. Unmount Media from Control Station:**

```
# umount /media/floppy
```

```
# umount /media/cdrom
```

**CREATING BOOT FLOPPY FROM NAS CD-ROM (NAS 5.5 and below):**

1. # mount /dev/cdrom /mnt/cdrom
2. # mount -t msdos /dev/fd0 /mnt/floppy
3. # dd if=/mnt/cdrom/images/boot.img of=/dev/fd0 bs=1024

**Note:** Though the command comes back to the prompt immediately, keep in mind that the floppy has not yet been written to—will take a few minutes to complete

**CREATING LINUX BOOT FLOPPY FROM RUNNING LINUX SYSTEM:**

Insert diskette and issue command: **#/sbin/mkbootdisk `uname -r`**

**CREATING BOOT FLOPPY FOR NAS 4.X UPGRADE FROM CD-ROM ON WINDOWS:**

1. Requires having rawrite.exe program on the Windows system
2. Execute rawrite.exe and enter path to CD-ROM: c:>d:\images\boot.img
3. Enter diskette in a: drive, enter

**EXTRACTING FILES FROM CD-ROM:**

1. Mount CD containing NAS Version:  

```
#mount /dev/cdrom /mnt  
#cd /mnt/cdrom/EMC/nas  
#mkdir /nas/var/extract
```
2. Copy the RPM to the new directory & extract:  

```
#cp emcnas-5.1.18-3.i386.rpm /nas/var/extract  
#cd /nas/var/extract  
#rpm2cpio emcnas-5.1.18-3.i386.rpm | lcpio --extract --make-directories */nas/cifs* [*emcopy.exe* or *readme.txt*]
```
3. #cd /nas/var/extract/nas/cifs/usrmapperV3/linux [Check directories for desired files]

**EXTRACTING SINGLE FILE FROM NAS ISO IMAGE:**

1. FTP NAS ISO image to Control Station, or mount cd-rom (#mount /dev/cdrom)
2. Create mount point  

```
# mkdir /emc2
```
3. Mount ISO image using Linux loop device  

```
# mount /nas/var/temp/emcnas5.4.16.1.iso /emc2 -t iso9660 -o loop=/dev/loop0
```
4. Copy emcnas rpm to Control Station location  

```
# cd /nas/var/temp/EMC/nas  
# cp emcnas-5.4.16-1.i386.rpm /nas/var/temp
```
5. Extract the required file to /nas/var/temp using the following command.

```
#cd /nas/var/temp # rpm2cpio emcnas-5.4.16-1.i386.rpm | cpio --make-directories --extract *nas.exe*
```

```
# ls -al /nas/var/temp/nas/dosfs/bin
```

```
-rwxr-xr-x 1 root root 20667155 Aug 25 13:32 nas.exe
```

```
# umount /emc2
```

```
# rm /nas/var/temp/emcnas-5.4.16-1.i386.rpm
```

**EXTRACTING SPECIFIC PORTION OF NAS RPM TO REBUILD NAS CMD:**

**Note:** /nas/bin/nas\_cmd is invoked by numerous server commands, such as server\_df

```
#mount /dev/cdrom  
#cd /mnt/cdrom/EMC/nas  
#cp emcnas-5.4.19-4.i386.rpm /home/nasadmin  
#cd /home/nasadmin
```

**# rpm2cpio emcnas-5.4.19-4.i386.rpm | cpio -- extract --make-directories \*nas\_cmd\***

#### **EXTRACTING SINGLE FILES FROM RPM (ISO or CD-ROM):**

1. mount CD-ROM, ISO image, or place the .rpm file in a desired directory on the Control Station

**# mount /dev/cdrom # mount -t iso9660 -o loop emcnas5.4.17.5.iso /rpm**

2. Change to /nas/var/dump/pkg\_5.5.79.0 directory

3. Verify that file you want is in the RPM

**# rpm2cpio /nas/var/dump/pkg\_5.5.79.0/emcnas-5.5.79-0.i386.rpm |cpio -t |grep -i nas\_mcd.cfg**

```
./nas/sys/nas_mcd.cfg
```

4. Extract the file from RPM

```
# rpm2cpio /nas/var/dump/pkg_5.5.79.0/emcnas-5.5.79-0.i386.rpm |cpio -icduv *nas_mcd.cfg
```

```
./nas/sys/nas_mcd.cfg
```

**Note:** File will be extracted to the directory from which the command is run, and then to the path that the file exists within the RPM.  
In this example, the nas\_mcd.cfg file was extracted to the /nas/var/dump/pkg\_5.5.79.0/nas/sys directory.

#### **RESTORING NASDB FROM BACKUP:**

```
#uncompress nasdb_backup.x.tar.z [gzip -d for gz file]
```

```
#/nas/sbin/nasdb_restore /nas /nasdb_backup.x.tar
```

**Note:** Might need to conduct an FSCK on the restored file system.

#### **NASDB BACKUP & RESTORE:**

##### **BACKUP NAS DATABASE:**

**\$/nas/sbin/nasdb\_backup /nas /home/nasadmin**

##### **RESTORING NAS DATABASE:**

**# /nas/sbin/nasdb\_restore /nas /home/nasadmin/nasdb\_backup.x.tar** [Used to restore a backup copy of the NAS database]

**Example:**      uncompress nasdb\_backup.x.tar.z [or gzip -d for gz files]

```
/nas/sbin/nasdb_restore /nas /nasdb_backup.x.tar [May need to run fsck on restored file systems]
```

**Note:** NAS\_DB Backups do not currently backup SecMap or Internal Usermapper databases with 5.2 or 5.3

#### **NAS DB BACKUPS AFFECTS TIMEFINDER REFRESHES & OTHER PROCESSES:**

**Note:** There is a problem between nas\_db locking its database files and other processes not being able to complete because of this. Turns out that this is a regression issue and NAS 5.3 will contain fix to essentially not have the nas\_db backup lock the Celerra db.

##### **Workaround Fix:**

1. Stop NAS\_DB Cron Job before editing & comment out the Job:

```
# vi /nas/site/cron.d/nas_sys
```

```
#1 * * * * nasadmin /nas/sbin/nasdb_backup /nas /home/nasadmin 1 &>/dev/null
```

2. Edit nasdb\_backup Script to add “return 0” to following two areas in script:

```
# vi /nas/sbin/nasdb_backup
```

```
return 0
```

```
DIR=$1
```

```
DIR_LOCKED=$1$DIR_LOCK_EXT
```

3. Uncomment nasdb\_backup Job in cron.d directory

4. Reset Cron by doing: #touch /etc/crontab

**Note:** See AR59118 for more details about the fix. Fixed in NAS 5.3.17.0/5.4.15.0→nasdb\_backup script used to grab a lock while creating the SCCS archives. While holding this lock, all other administrative operations would not be permitted—was done to ensure consistency in the SCCS archive—this has found not to be required and is no longer done. One problem caused was that a SavVol extend might not be executed if the NAS\_DB had a lock, leading to SavVol full and inactive condition.

#### **MONITORING PROGRAM FOR NS SERIES SERIAL PORTS:**

**Note:** CTAP logs record the serial boot history of a data mover and is new to NAS 5.3—up to 11 logs per DM are kept

##### **Location:**

**/var/log/ctapptyS4.log** (Data Mover 2)

**/nas/sbin/ctap/ctapmon** →Program to monitors serial ports listed in /etc/ctapmon.conf

**# ps -ef |grep -i ctap**

```
root 30531 1 0 08:51 pts/7 00:00:00 /nasmcd/sbin/ctap_mon /dev/ttyS4
```

```
root 30539 1 0 08:51 pts/7 00:00:00 /nasmcd/sbin/ctap_mon /dev/ttyS5
root 30547 1 0 08:51 pts/7 00:00:00 /nasmcd/sbin/ctap_mon /dev/ttyS6
root 30555 1 0 08:51 pts/7 00:00:00 /nasmcd/sbin/ctap_mon /dev/ttyS7
```

# /nas/sbin/ctap/ctapmon -restart

# cat /etc/ctapmon.conf

```
/dev/ttyS4:/var/log/ctapttyS4.log →Data Mover_2
/dev/ttyS5:/var/log/ctapttyS5.log →Data Mover_3
/dev/ttyS6:/var/log/ctapttyS6.log →Data Mover_4
/dev/ttyS7:/var/log/ctapttyS7.log →Data Mover_5
```

**Note:** Logs contain information that may be useful in troubleshooting install/upgrade failures

#### **SAMPLE LOG OUTPUT:**

```
CONSOLE> rcq<0D>sib<0D>s reason_code=05 <0D><0A>
Fri Oct 15 17:05:56 2004
```

## **CELLERRA PRIMARY, SECONDARY, & IPMI INTERNAL NETWORKS**

Celerra uses a private IP communications network based on a Primary & Secondary network, for monitoring and control functions on the Data Movers, Control Station, and for Clariion SPs for Integrated systems. Additionally, there is a 3<sup>rd</sup> private network called IPMI, used only for those systems that use dual Control Stations (NSX, NS80, old CNS, CFS-14, etc)

### **CELLERRA INTERNAL IP NETWORKS ON CS:**

#### **SCO & LINUX Upgrades From SCO:**

192.1.1.100/192.1.2.100

#### **LINUX INSTALLS ABOVE 2.2.35.4:**

192.168.1.100/192.168.2.100

#### **NAPA 2 5.5.22.x SYSTEMS:**

Integrated SP addresses on SPB change from 192.168.1.201 to 192.168.2.201

#### **NOVEMBER 2007 NAS 5.5.31.x PRIVATE ADDRESS CHANGE—Current vs. New:**

#### **CURRENT CELERRA PRIVATE NETWORK SCHEME:**

Primary internal network IP: 192.168.1.100

Backup internal network IP: 192.168.2.100

Netmask: 255.255.255.0

IP address of storage processor A: 192.168.1.200

IP address of storage processor B: 192.168.2.201 (CX3-series Integrated NS)

IP address of storage processor B (NS350 Only & CX3-series NS): 192.168.1.201

Gateway IP address of storage processor A: 192.168.1.104

Gateway IP address of storage processor B: 192.168.2.104

IMPI network (NS80 only): 192.168.3.100; 192.168.3.101

#### **CELLERRA PRIVATE NETWORK SCHEME (NAS 5.5.32+):**

Primary internal network IP: 128.221.252.100

Backup internal network IP: 128.221.253.100

Netmask: 255.255.255.0

IP address of storage processor A: 128.221.252.200

IP address of storage processor B: 128.221.253.201

IP address of storage processor B (NS350 Only): 128.221.252.201

Gateway IP address of storage processor A: 128.221.252.104

Gateway IP address of storage processor B: 128.221.253.104

IMPI networks (NS80 only): 128.221.254.100 emcnasipmi; 128.221.254.101 emcnasotherIPMICS\_i3

**Note:** With the release of the NS20 for Napa 8, it was found that many customers were already using the 192.168.1/2 networks for their internal LAN, which conflicts with the default Celerra networks, resulting in the need to perform fresh reinstalls for systems that are being touted as factory-installed. As a result, the 128.221.252/253/254 networks will become the new default private networks for the Celerra. All systems shipped from Manufacturing as of 11/05/2007 will ship with the new defaults. As of NAS 5.5.32.x, the DART code will be updated to reflect the new Private Network scheme.

## **INTEGRATED SP ADDRESSES & GATEWAY FOR CONTROL STATION COMMUNICATIONS:**

#### **Pre-5.5.30 NS SYSTEMS USING CX3-ARRAYS:**

SPA 192.168.1.200

gw 192.168.1.100 → eth0 192.168.1.100

SPB 192.168.2.201

gw 192.168.2.100 → eth2 192.168.2.100

#### **NAS 5.5.30+ OLD INTERNAL IPs:**

SPA 192.168.1.200

gw 192.168.1.104 → eth0:1 192.168.1.104

SPB 192.168.2.201

gw 192.168.2.104 → eth2:1 192.168.2.104

**NAS 5.5.30+ NEW INTERNAL IPs:**

SPA 128.221.252.200

gw 128.221.252.104 → eth0:1 128.221.252.104

SPB 128.221.253.201

gw 128.221.253.104 → eth2:1 128.221.253.104

**NAS 5.5.30+ Proxy ARP:**

SPA 10.250.169.50

gw 10.250.169.1 → eth0:1 128.221.252.104 (remains configured on CS but is no longer applicable)

SPB 10.250.169.51

gw 10.250.169.1 → eth2:1 128.221.253.104 (remains configured on CS but is no longer applicable)

**CHANGING GET REASON CODE FROM CONFIGURED (4) TO CONTACTED (5):**

# /nas/sbin/setreason /nas 4 5 [Setting slot\_4 to ‘Sib’ Contacted State, 5--/nas required to define NAS\_DB path]

**CELLERRA FILE SERVER OPERATION:**

**DETERMINING DATAMOVER STATUS:**

# /nas/sbin/getboxmask -r [Shows DM’s powered on or not. Empty slot may mean DataMover removed]

# /nas/sbin/getreason [If both commands show the slot empty, then the DataMover is not plugged into backplane]

# /nasmcd/getreason [alternate command to obtain slot status]

# /nas/bin/nas\_server -l [Slot populated if a setup\_slot was performed—added to database & will still show if DM removed]

**Note:** Above command reads information from the /nas/server/servers directory

# /nas/sbin/t2adc

**Purpose:** pwr & voltage readings for DM; Will also show if a slot is physically empty

# nas\_server -a -i [Shows which DataMovers are Out\_of\_Service and which Slots are empty]

**Key Directories:** /nas/sbin and /nas/bin and /nas/server /nas/dos /nas/log /nas/sys etc Hosts Inet

**OBTAINING SLOT STATUS IF BOX MONITOR SHUT OFF OR FROM SECONDARY CS:**

# /nasmcd/getreason

**Note:** This command will return slot status even if Box Monitor is disabled or if you are telnetted into CS1 running as Secondary

**NAME RESOLUTION: Control Station and DataMovers**

**/etc/hosts File:** Normally configured on Control Stations and Sun Solaris servers, etc, but could have Hosts file for DM too.

Step 1. Mount the DataMover directly to the Control Station [#mount -F nfs server\_2:/ /mntpoint

Step 2. #ls -la to locate the .etc file

**Copying Hosts file from Control Station to DM:** \$cp /etc/hosts /tmp \$cd /tmp; #server\_file server\_2 -put hosts hosts

**Note:** use \$server\_file server\_x -put to copy files from your current directory to the default ./etc directory of the DataMover

**ORDER OF HOST NAME RESOLUTION ON DATAMOVER:** ./etc/hosts file; NIS Server; DNS Server

**HARDWARE OPERATIONS:**

**Removing a Production Data Mover from the database:**

**Example:** Removing slot\_5 on a CFS-14, with updates for NAS 5.6

**Caution:** There may be additional steps required for updating enclosure database information

Step 1. Verify that file systems are not in use before unmounting (nas\_fs –list):

# nas\_fs -list

Step 2. Permanently Unexport and Unmount all FileSystems

# server\_export server\_5 -u -p -a

# server\_umount server\_5 -perm -a

Step 3. If removing a primary Data Mover where a Standby has been configured, delete the standby relationship:

# server\_standby server\_5 -delete mover=server\_3 (where server\_3 is the Standby in this example)

Step 4. Permanently unmount the Root FileSystem:

# server\_umount server\_5 -p root\_fs\_5

Step 5. Leave the Data Mover running, in its slot, and run the following command:

# /nas/sbin/setup\_slot -d 5 [Will prompt that database will be deleted—select ‘Yes’]

Step 6. Rebuild SYMAPI Database using the following command:

**Prior to NAS 5.6:**

```
# mv /nas/symapi/db/symapi_db.bin /nas/symapi/db/symapi_db.old  
# /nas/sbin/nas_rdf –localinit  
# chown nasadmin:nasadmin symapi_db.bin
```

**NAS 5.6 or Higher:**

```
# server_devconfig ALL –c –s –a
```

Step 7. Conduct additional database cleanup:

```
# vi /nas/server/servers and remove the line for server_5  
# cd /nas/server;ls -la and remove link that refers to "slot_5 -->server_4"  
# rm slot_5  
# cd /nas/server/server_4, verify your location using ‘pwd’, then delete contents of /nas/server/server_4  
# rm -Rf *
```

**Note:** Due to an old Celerra system numbering anomaly, note that the /nas/server/server\_4 directory is actually for slot\_5. Exercise care when selecting the right slot vs. the right server directory!

Step 8. Verify that Root filesystem for Slot\_5 is unmounted:

```
#cat /nas/volume/filesys and verify that "n" is set, as shown below:
```

```
5:root_fs_5::0:10::n:1:18:::0:::0:0::
```

Step 9. Power down slot\_5 and remove the hardware

Step 10. If running NAS 5.6, and intending to move the Hardware to a new physical slot, make sure that nas\_cel –update id=0 is run to ensure that the Data Mover loopback interconnects are created

Step 11. Do system healthchecks:

```
# nas_storage –check -all  
# nas_checkup  
# /nas/sbin/setup_enclosure -checkSystem  
# /nas/sbin/enclosure_status
```

**REBOOTING, HALTING, USING T2RESET:**

```
#/nas/sbin/t2reset reboot -s 2 {slot # of DM}  
#/nas/sbin/t2reset pwoff {pwron} -s slot  
#server_cpu server_2 -r -monitor now  
#server_cpu server_2 -halt now  
#/nas/sbin/t2reset reboot -s 2  
# /sbin/halt  
# /nas/sbin/nas_halt now  
Perform a controlled halt of the Control Station(s) and Data Mover(s)
```

**RESETTING DM USING MOTHERBOARD CMD: #/nas/sbin/t2reset mother -s <slot\_#>**

**Note:** This command may be useful if DM is stuck in reason code 0 or 7, or if the comm. board is bad, for older CNS/CFS frames. Command issues a Tier 2 reset to the motherboard.

**WHY "T2RESET" vs. "SERVER CPU" COMMAND?**

1. Use the “t2reset” whenever you’ve edited the “boot.cfg” file, as in preparing to run an FSCK. The Server will reboot from the /nas/dos/slot\_x/boot.cfg and NOT rebuild the “boot.cfg” from /nas/server/slot\_x files as it normally does on reboot. The server\_cpu command invokes a build\_config which recompiles the boot.cfg based on the NASDB configuration files of the Data Mover slot.
2. Similarly, use the “server\_cpu” reboot command when you do want to rebuild the server’s configuration files [aka, “boot.cfg”], such as when the server won’t progress from “Status 4 configured” to “Status 5 contacted”. The Server rebuilds the “boot.cfg” from the files located in the /nas/server/slot\_x using build\_config.

**STOPPING NAS SERVICES:**

**#/sbin/service nas stop**

```
#ps -ef |grep boxm  
#ps -ef |grep nasmcd
```

**Note:** Do not try to stop NAS services while in any of the NAS partition directories

**SHUTTING DOWN/REBOOTING/CHANGING RUNLEVEL OF LINUX CONTROL STATION:**

**# sync;reboot | # reboot | # reboot -n -f** [Force, no sync reboot] | **# sync; sync;sync;init 6 | init 0 | # shutdown -r now**

## # /nas/sbin/nas\_halt now

Perform a controlled halt of the Control Station(s) and Data Mover(s)

## # /sbin/halt

### **Data Mover Reason Codes--BootUp:**

- 0 Data Mover performing BIOS check, then begins boot sequence
- 1 Data Mover passes SIB POST
- 2 Data Mover fails SIB POST
- 3 Data Mover booting EMC NAS
- 4 Data Mover in running state
- 5 Box monitor operational and Data Mover in normal running state
- 7 Data Mover Crash
- 9 Reboot

**NAS.EXE File:** \$/nas/dos/bin/nas.exe [/nas/dos/slot\_x contains gload.exe, boot.cfg; /nas/dos/bin contains nas.exe and misc. files]

**BOOT.BAT & BOOT.CFG Files:** \$/nas/dos/slot\_x/boot.bat & boot.cfg [DataMover configuration files]

## **CELLERRA NASDB**

### **NAS DATABASE:**

- Located in /nas/server directory and contains configuration information for each Data Mover, as well as general information for all Servers in the cabinet (servers & server\_setup files)
- Contains all files that combine to create the /nas/dos/slot\_x/boot.cfg file and complete Data Mover configuration on bootup
- Files contained in /nas/server/server directory allow server to mount file systems (camdisk; mount; mountpoint;ufs;volume)
- See /nas/sbin/build\_config script to see how NAS DB files are used to build the boot.cfg

### **CONTROL STATION DATABASE:**

- Located in /nas/volume directory, contains files common to all data movers regarding Storage volumes & file systems

## **CELLERRA ROWLEVEL LOCKING NAS 5.5 (Row Locking Files):**

- NAS 5.5 introduced RLL in order to support the DBMS database that will be put into service with NAS 5.6

- Row Locking done on /nas/volume/disks, volumes, filesystems, and slices files

## **/nas/sbin/nasdb\_tools**

```
# ls -la
-rwxrwxr-x 1 nasadmin nasadmin 1295 Jan 31 10:48 db_status
-rwxrwxr-x 1 nasadmin nasadmin 4583 Jan 31 10:48 downgrade_db
-rwxrwxr-x 1 nasadmin nasadmin 1685 Jan 31 10:48 downgrade_table
-rwxrwxr-x 1 nasadmin nasadmin 56223 Jan 31 10:48 fixdb3
-rwxrwxr-x 1 nasadmin nasadmin 2812 Jan 31 10:48 lock_cleanup
-rwxrwxr-x 1 nasadmin nasadmin 210143 Jan 31 10:48 lock_wrapper
-rwxrwxr-x 1 nasadmin nasadmin 5425 Jan 31 10:48 remove_lock_owners
-rwxrwxr-x 1 nasadmin nasadmin 5837 Jan 31 10:48 upgrade_db
-rwxrwxr-x 1 nasadmin nasadmin 1784 Jan 31 10:48 upgrade_table
```

## **# /nbsnas/sbin/nasdb\_tools/db\_status**

Current DB seems to be at version 3

**Note:** Above output reflects RLL with NAS 5.5.27.0

## **# /nas/sbin/nasdb\_tools/db\_status**

Current DB seems to be at version NO\_ROW\_LOCKING

**Note:** Output as seen from NAS 5.4 or earlier. An issue like this could be seen if upgrading from 5.4 to 5.5 and the upgrade did not complete the update of the NASDB.

→ NAS 5.5 contains Row Level Locking (RLL) in each db file to support auto fs extension and Thin Provisioning

### **ROWLOCK EXAMPLE FROM FILESYS, DISKS, & VOLUMES FILE:**

## **\$ head filesystem**

### **ROWLOCK:FILESYS:1:24:4::**

**Note:** The :24: field represents the last index number used in the last file entry. The last empty :: field would show a <pid> if the db was locked by another process

→ disks, volumes, slices, & filesystems files are all directly related and must match info posted in camdisk files

→ pools & profiles files are used by AVM to select disks during file system creation

## **\$ cat disks**

### **ROWLOCK:DISKS:1:8:8::**

```
1:root_disk::0:9::y:11263:APM00073801833:1:1:0000:7:  
2:root_ldisk::0:9::y:11263:APM00073801833:1:2:0001:7:  
3:d3::0:8::y:2047:APM00073801833:1:3:0002:7:  
4:d4::0:8::y:2047:APM00073801833:1:4:0003:7:  
5:d5::0:8::y:2047:APM00073801833:1:5:0004:7:  
6:d6::0:8::y:32767:APM00073801833:1:6:0005:7:  
7:d7::0:5::n:175473:APM00073801833:1:83:0010:7:  
8:d8::0:5::n:175473:APM00073801833:1:84:0011:7:
```

### \$ cat volumes

#### ROWLOCK:VOLUMES:1:84:6::

```
1:root_disk::0:41::y:4:0:1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,52:1:  
2:root_ldisk::0:23::y:4:0:35,36,37,38,39,40,41,42,43,44,45,46,47,48,49,50,51:2:  
3:d3::0:6::y:4:1:77:3:  
4:d4::0:6::y:4:1:78:4:  
5:d5::0:6::y:4:1:79:5:  
6:d6::0:6::y:4:1:80:6:  
---removed----  
83:d7::0:3::n:4:0::7:  
84:d8::0:3::n:4:0::8:
```

### **DB Status if Row Locking File missing with NAS 5.5:**

```
# ls -la /nas/volume/.row_locking_db_version*
```

**Note:** File not found

```
# /nas/sbin/nasdb_tools/db_status
```

Current DB seems to be at version NO\_ROW\_LOCKING

#### **Recreating & Setting Permissions on the Row Locking DB File:**

```
/nas/volume/
```

1. \$ echo -n "3" >/nas/volume/.row\_locking\_db\_version
2. \$ chmod 775 /nas/volume/.row\_locking\_db\_version
3. \$ ls -al /nas/volume/.row\_locking\_db\_version  
-rwxrwxr-x- 1 nasadmin nasadmin 1 < date time > /nas/volume/.row\_locking\_db\_version

#### 4. \$ cat .row\_locking\_db\_version

3

5. Verify:

```
# /nas/sbin/nasdb_tools/db_status
```

Current DB seems to be at version 3

### **DBCHK TOOL:**

```
/nas/tools/dbchk -wvxpVs [Not all switches are valid prior to NAS 5.5 code]
```

```
/nas/tools/dbchk.jar (Actual code for dbchk)
```

**Note:** dbchk was introduced in NAS 5.3 as a way to verify the relationship of db entries. Switches: w=warnings; v=verbose; x=extended checks; p=compare NAS\_DB to probe; V=volume consistency check; s=suggestion mode

### **NAS DB SIGNATURE FILES:**

→Each NAS\_DB file has a corresponding hidden signature file used to verify consistency and integrity of db files

Signature Files Found in Following Directories:

```
/nas/server  
/nas/server/slot_*  
/nas/server/vdm/  
/nas/site/  
/nas/sys/  
/nas/volume/
```

#### **EXAMPLE:** /nas/server/slot\_2/camdisk

```
-rw-r--r-- 1 nasadmin nasadmin 464 Jul 8 15:15 camdisk  
-rw-r--r-- 1 nasadmin nasadmin 13 Jul 8 15:15 .camdisk.sig  
# cat .camdisk.sig
```

1263787786:8: →Hash value, compared to bit value of original file, used to validate if file contents are valid

**Note:** If hash process detects problem, an error is logged in sys\_log. In above Hash value example, last colon-separated value “:8:” represents the last dvolume added to the database.

**GLOBAL PARAM DATABASE--Boot.Cfg File:**

Also known as the “Global Param Database” file. Location: */nas/dos/slot\_x*

**Note:** The boot.cfg is compiled from DART memory as it goes down for reboot. It contains information from the various configuration files located in the */nas/server\_slot\_x* directory

---

```
#verify_start
logsys add output disk=root_log_10 bufsz=256
# This is the global param file
param nfs v3xfersize 32768
# buff
bufcache
device isa isa-0
device pci pci-0
dskdumpconfig full slot=10
pciautoconfig
el3config tier2-aui
sysconfig net name=ace0 option="linkneg=disable"
ifconfig el30 protocol=IP device=el30 local=192.1.1.10 netmask=255.255.255.0 broadcast=192.1.1.255 mtu=1500 mac=0:1:2:c0:69:8b
ifconfig el31 protocol=IP device=el31 local=192.1.2.10 netmask=255.255.255.0 broadcast=192.1.2.255 mtu=1500 mac=0:1:2:c0:7d:58
ifconfig ace0 protocol=IP device=ace0 local=172.20.30.21 netmask=255.255.0.0 broadcast=172.20.255.255 mtu=1500 mac=0:60:cf:20:4f:1f
hostname abv-nfs21 selfid=10
file initialize nodes=10240 dnlc=8192
volume disk 1 c0t0l0
volume disk 1 c1t0l0
--abridged-----
volume hyper 224 2 214 215
volume disk 2 c0t0l1
volume disk 2 c1t0l1
volume slice 123 1179648 131072 2
file recover ufs 123=131072 69=10 220=59 221=60 222=61 223=62 224=63
log sectors=131072 123
file mount uxfss rw / 69=10 rw
file mount uxfss rw /vol01 220=59 rw
file mount uxfss rw /vol02 221=60 rw
file mount uxfss rw /vol03 222=61 rw
file mount uxfss rw /vol04 223=62 rw
file mount uxfss rw /vol05 224=63 rw
export "/" anon=0 access=192.1.1.100:192.1.2.100:192.1.1.101:192.1.2.101
export "/vol01" access=prod root=rootgp
export "/vol02" access=prod root=rootgp
export "/vol03" access=prod root=rootgp
export "/vol04" access=prod root=rootgp
export "/vol05" access=prod root=rootgp
transport start
udp checksum on
tcp
routed
rpc
communityname public
syscontact nasadmin
syslocation here
#
dns server=172.20.40.21 udp domain=prod.myplay.com
#
#
nfs start openfiles=4096 nfsd=96
statd
lockd
pax
rquotad action=start
ftpd
sysman
ndmp port=10000
mac
logsys set output disk=root_log_10 bufsz=256
#verify_eof
```

**REBUILDING THE BOOT.CFG FILE WITHOUT REBOOTING:**

**Comment:** Use this feature with extreme care; backup NAS\_DB first; save extra copy of boot.cfg file!!

In certain situations it becomes necessary to edit param files contained in the */nas/server/slot\_x* directory. Once the edits are done, you can execute the following command to rebuild the Server's "boot.cfg" file from the modified files:

**\$/nas/sbin/build\_config /nas/server/slot\_5 /nas/dos/slot\_5**

**\$/nas/sbin/build\_config -min /nas/server/slot\_7 /nas/dos/slot\_7**

**Note:** Use this second command to build a ‘minimum’ configuration for a DataMover that you are having problems trying to bootup

## **SERVER LOG EXAMPLE: NORMAL CPU BOOT-UP:**

```
2001-05-23 05:34:13: ADMIN: 4: Command succeeded: logsys add output disk=root_log_10 bufsz=256 [Beginning of reboot]
2001-05-23 05:34:13: ADMIN: 3: Command failed: param nfs v3xfersize 32768
2001-05-23 05:34:13: ADMIN: 4: Command succeeded: bufcache
2001-05-23 05:34:13: ADMIN: 4: Command succeeded: device isa isa-0
2001-05-23 05:34:13: KERNEL: 3: PCI BIOS Rev 02.10 [bios and POST check]
2001-05-23 05:34:13: KERNEL: 4: CMB-100 Motherboard
2001-05-23 05:34:13: ADMIN: 4: Command succeeded: device pci pci-0
2001-05-23 05:34:13: ADMIN: 4: Command succeeded: dskdumpconfig full slot=10
2001-05-23 05:34:13: DRIVERS: 4: scsi-0 (AHA3944AUW Ch: A) @ 1400, irq 7, bus 0, func 0 [scsi drivers loading]
2001-05-23 05:34:14: DRIVERS: 4: scsi-1 (AHA3944AUW Ch: B) @ 1800, irq f, bus 0, func 1
2001-05-23 05:34:16: DRIVERS: 4: ace0: Serial Number is 0:60:cf:20:4f:1f [network interface drivers loading]
2001-05-23 05:34:16: DRIVERS: 4: ace0: Part number is 200007B02
2001-05-23 05:34:16: DRIVERS: 4: ace0: Board Revision is C
2001-05-23 05:34:16: DRIVERS: 3: ana0 Link setting change: Autonegotiate enable
2001-05-23 05:34:16: DRIVERS: 4: ana0 : csr port = 2000, bus 1, irq = 15
2001-05-23 05:34:16: DRIVERS: 3: ana1 Link setting change: Autonegotiate enable
2001-05-23 05:34:16: DRIVERS: 4: ana1 : csr port = 2080, bus 1, irq = 15
2001-05-23 05:34:16: DRIVERS: 3: ana2 Link setting change: Autonegotiate enable
2001-05-23 05:34:16: DRIVERS: 4: ana2 : csr port = 2400, bus 1, irq = 15
2001-05-23 05:34:16: DRIVERS: 3: ana3 Link setting change: Autonegotiate enable
2001-05-23 05:34:16: DRIVERS: 4: ana3 : csr port = 2480, bus 1, irq = 15
2001-05-23 05:34:16: ADMIN: 4: Command succeeded: pciautoconfig
2001-05-23 05:34:17: ADMIN: 4: Command succeeded: volume disk logdisk c0t0l0 [volume for logging]
2001-05-23 05:34:17: ADMIN: 4: Command succeeded: volume disk logdisk c1t0l0
2001-05-23 05:34:17: ADMIN: 4: Command succeeded: volume slice root_log_10 821 248 2047 logdisk
2001-05-23 05:34:17: DRIVERS: 4: el30 : irq = a, csr = 300, type = 1, address = 0:1:2:c0:69:8b [more interface drivers]
2001-05-23 05:34:17: DRIVERS: 4: el3Device::Initialize : csr =300, name = el30, unit =0
2001-05-23 05:34:17: DRIVERS: 4: link ip el30
2001-05-23 05:34:17: DRIVERS: 4: link amgr el30
2001-05-23 05:34:18: DRIVERS: 4: el31 : irq = b, csr = 310, type = 1, address = 0:1:2:c0:7d:58
2001-05-23 05:34:18: DRIVERS: 4: el3Device::Initialize : csr =310, name = el31, unit =1
2001-05-23 05:34:18: DRIVERS: 4: link ip el31
2001-05-23 05:34:18: DRIVERS: 4: link amgr el31
2001-05-23 05:34:18: ADMIN: 4: Command succeeded: el3config tier2-aui
2001-05-23 05:34:18: ADMIN: 4: Command succeeded: sysconfig net name=ace0 option="linkneg=disable"
2001-05-23 05:34:18: ADMIN: 4: Command succeeded: ifconfig el30 protocol=IP device=el30 local=192.1.1.10 netmask=255.255.255.0 broadcast=192.1.1.255
mtu=1500
mac=0:1:2:c0:69:8b
2001-05-23 05:34:18: ADMIN: 4: Command succeeded: ifconfig el31 protocol=IP device=el31 local=192.1.2.10 netmask=255.255.255.0 broadcast=192.1.2.255
mtu=1500
mac=0:1:2:c0:7d:58
2001-05-23 05:34:18: ADMIN: 4: Command succeeded: ifconfig ace0 protocol=IP device=ace0 local=172.20.30.21 netmask=255.255.0.0 broadcast=172.20.255.255
mtu=15
00 mac=0:60:cf:20:4f:1f
2001-05-23 05:34:18: ADMIN: 4: Command succeeded: hostname abv-nfs21 selfid=10
2001-05-23 05:34:18: ADMIN: 4: Command succeeded: file initialize nodes=10240 dnlc=8192
2001-05-23 05:34:18: ADMIN: 4: Command succeeded: volume disk 1 c0t0l0 [volumes discovered]
2001-05-23 05:34:18: ADMIN: 4: Command succeeded: volume disk 1 c1t0l0
2001-05-23 05:34:18: ADMIN: 4: Command succeeded: volume slice 68 571392 32768 1
2001-05-23 05:34:18: ADMIN: 4: Command succeeded: volume hyper 69 1 68
2001-05-23 05:34:18: DRIVERS: 4: ace0: Gigabit Ethernet link is up [GB NIC Up and ready]
2001-05-23 05:34:18: DRIVERS: 4: ace0: rx flow control is disabled, tx flow control is disabled
2001-05-23 05:34:18: ADMIN: 4: Command succeeded: volume disk 206 c0t1l8
2001-05-23 05:34:19: ADMIN: 4: Command succeeded: volume disk 206 c1t1l8
2001-05-23 05:34:19: ADMIN: 4: Command succeeded: volume disk 207 c0t1l9
2001-05-23 05:34:19: ADMIN: 4: Command succeeded: volume disk 207 c1t1l9
2001-05-23 05:34:19: ADMIN: 4: Command succeeded: volume hyper 220 2 206 207
2001-05-23 05:34:20: ADMIN: 4: Command succeeded: volume disk 208 c0t1l10
2001-05-23 05:34:20: ADMIN: 4: Command succeeded: volume disk 208 c1t1l10
2001-05-23 05:34:21: ADMIN: 4: Command succeeded: volume disk 209 c0t1l11
2001-05-23 05:34:21: ADMIN: 4: Command succeeded: volume disk 209 c1t1l11
2001-05-23 05:34:21: ADMIN: 4: Command succeeded: volume hyper 221 2 208 209
2001-05-23 05:34:21: ADMIN: 4: Command succeeded: volume disk 210 c0t1l12
2001-05-23 05:34:21: ADMIN: 4: Command succeeded: volume disk 210 c1t1l12
2001-05-23 05:34:21: ADMIN: 4: Command succeeded: volume disk 211 c0t1l13
2001-05-23 05:34:21: ADMIN: 4: Command succeeded: volume disk 211 c1t1l13
2001-05-23 05:34:21: ADMIN: 4: Command succeeded: volume hyper 222 2 210 211
2001-05-23 05:34:21: ADMIN: 4: Command succeeded: volume disk 212 c0t1l14
```

2001-05-23 05:34:21: ADMIN: 4: Command succeeded: volume disk 212 c1t1114  
2001-05-23 05:34:21: ADMIN: 4: Command succeeded: volume disk 213 c0t1115  
2001-05-23 05:34:21: ADMIN: 4: Command succeeded: volume disk 213 c1t1115  
2001-05-23 05:34:21: ADMIN: 4: Command succeeded: volume hyper 223 2 212 213  
2001-05-23 05:34:21: ADMIN: 4: Command succeeded: volume disk 214 c0t210  
2001-05-23 05:34:21: ADMIN: 4: Command succeeded: volume disk 214 c1t210  
2001-05-23 05:34:21: ADMIN: 4: Command succeeded: volume disk 215 c0t211  
2001-05-23 05:34:21: ADMIN: 4: Command succeeded: volume disk 215 c1t211  
2001-05-23 05:34:21: ADMIN: 4: Command succeeded: volume hyper 224 2 214 215  
2001-05-23 05:34:21: ADMIN: 4: Command succeeded: volume disk 2 c0t011  
2001-05-23 05:34:21: ADMIN: 4: Command succeeded: volume disk 2 c1t011  
2001-05-23 05:34:21: ADMIN: 4: Command succeeded: volume slice 123 1179648 131072 2  
2001-05-23 05:34:21: UFS: 4: Recovering filesystems *[beginning of File System discovery & mounting]*  
2001-05-23 05:34:21: UFS: 4: successful recovery took 1 ticks  
2001-05-23 05:34:21: UFS: 4: processed records 0 - 0, 14 reads, 2 writes  
2001-05-23 05:34:21: ADMIN: 4: Command succeeded: file recover ufs 123=131072 69=10 220=59 221=60 222=61 223=62 224=63  
2001-05-23 05:34:23: ADMIN: 4: Command succeeded: log sectors=131072 123  
2001-05-23 05:34:23: UFS: 4: Cleaners into 5c3b8, cg 5c454, dirty 5e628  
2001-05-23 05:34:23: ADMIN: 4: Command succeeded: file mount ufs rw / 69=10 rw *[root file system is mounted]*  
2001-05-23 05:34:26: UFS: 4: Cleaners into 5c7fc, cg 5c760, dirty 5c898  
2001-05-23 05:34:26: ADMIN: 4: Command succeeded: file mount ufs rw /vol01 220=59 rw  
2001-05-23 05:34:28: UFS: 4: Cleaners into 5ca6c, cg 5c9d0, dirty 5cb08  
2001-05-23 05:34:28: ADMIN: 4: Command succeeded: file mount ufs rw /vol02 221=60 rw  
2001-05-23 05:34:30: UFS: 4: Cleaners into 5ccdc, cg 5cc40, dirty 5cd78  
2001-05-23 05:34:30: ADMIN: 4: Command succeeded: file mount ufs rw /vol03 222=61 rw  
2001-05-23 05:34:33: UFS: 4: Cleaners into 5cf4c, cg 5ceb0, dirty 5cf8  
2001-05-23 05:34:33: ADMIN: 4: Command succeeded: file mount ufs rw /vol04 223=62 rw  
2001-05-23 05:34:35: UFS: 4: Cleaners into 5d1bc, cg 5d120, dirty 5d258  
2001-05-23 05:34:35: ADMIN: 4: Command succeeded: file mount ufs rw /vol05 224=63 rw  
2001-05-23 05:34:35: ADMIN: 4: Command succeeded: export "/" anon=0 access=192.1.1.100:192.1.2.100:192.1.1.101:192.1.2.101  
2001-05-23 05:34:35: ADMIN: 4: Command succeeded: export "/vol01" access=prod *[file systems exported]*  
root=rootgp  
2001-05-23 05:34:35: ADMIN: 4: Command succeeded: export "/vol02" access=prod root=rootgp  
2001-05-23 05:34:35: ADMIN: 4: Command succeeded: export "/vol03" access=prod root=rootgp  
2001-05-23 05:34:35: ADMIN: 4: Command succeeded: export "/vol04" access=prod root=rootgp  
2001-05-23 05:34:35: ADMIN: 4: Command succeeded: export "/vol05" access=prod root=rootgp  
2001-05-23 05:34:35: IP: 4: link ip loop  
2001-05-23 05:34:35: IP: 4: route add host 127.0.0.1 127.0.0.1 loop *[network & other services started]*  
2001-05-23 05:34:35: ADMIN: 4: Command succeeded: transport start  
2001-05-23 05:34:35: ADMIN: 4: Command succeeded: udp checksum on  
2001-05-23 05:34:35: ADMIN: 4: Command succeeded: tcp  
2001-05-23 05:34:35: ADMIN: 4: Command succeeded: routed  
2001-05-23 05:34:35: ADMIN: 4: Command succeeded: rpc  
2001-05-23 05:34:35: ADMIN: 4: Command succeeded: communityname public  
2001-05-23 05:34:35: ADMIN: 4: Command succeeded: syscontact nasadmin  
2001-05-23 05:34:35: ADMIN: 4: Command succeeded: syslocation here  
2001-05-23 05:34:35: ADMIN: 4: Command succeeded: dns server=172.20.40.21 udp domain=prod.myplay.com  
2001-05-23 05:34:35: ADMIN: 4: Command succeeded: nfs start openfiles=4096 nfsd=96  
2001-05-23 05:34:35: ADMIN: 4: Command succeeded: statd  
2001-05-23 05:34:35: ADMIN: 4: Command succeeded: lockd  
2001-05-23 05:34:35: ADMIN: 4: Command succeeded: pax  
2001-05-23 05:34:35: ADMIN: 4: Command succeeded: rquotad action=start  
2001-05-23 05:34:35: ADMIN: 4: Command succeeded: ftpd  
2001-05-23 05:34:35: ADMIN: 4: Command succeeded: LOG (6, defaultlog): System management component initializing...  
LOG (6, defaultlog): E003 - evd (V1.10) initialization complete  
LOG (6, defaultlog): sysman  
2001-05-23 05:34:35: NDMP: 3: RecBuf in NDMP Pool (count:0) (./ndmp.cxx:181)  
2001-05-23 05:34:35: ADMIN: 4: Command succeeded: ndmp port=10000  
2001-05-23 05:34:35: ADMIN: 4: Warning: Any client will be provided with the Management and Administration Control(MAC) services.  
2001-05-23 05:34:35: ADMIN: 4: Command succeeded: mac  
2001-05-23 05:34:35: ADMIN: 4: Command succeeded: logsys set output disk=root\_log\_10 bufsz=256  
CONSOLE>  
LOG (6, defaultlog): M004 - snmp\_pe (V1.10) initialization complete  
2001-05-23 05:34:35: NDMP: 3: BlockBuf in PAX Pool (count:0) (pax\_rbuf:4) (./buf\_subs.cxx:117)  
2001-05-23 05:34:42: SYSTEM: 4: Setting SIB attention status to 5, contact established  
2001-05-23 05:34:42: ADMIN: 4: Command succeeded: sib contacted *[Box Monitor running & recognizing server as Up]*

## NEW CELERRA DATABASE QUERYING SUBSYSTEM:

# nas\_disk -query:tags  
# nas\_fs -query:tags  
# nas\_volume -query:tags  
# nas\_slice -query:tags

**Note:** Run any of the above commands to see what type of Queries are available—following topics are available for FS queries:

## **AVAILABLE QUERIES FOR FILE SYSTEM TOPICS:**

Supported Query Tags for nas\_fs:

Supported Query Tags for CDMSConnections:

Supported Query Tags for CDMSThreads:

Supported Query Tags for Checkpoints:

Supported Query Tags for DefaultQuotas:

Supported Query Tags for Disks:

Supported Query Tags for ServerTable:

Supported Query Tags for GroupQuotas:

Supported Query Tags for ReplicationSessions:

Supported Query Tags for TreeQuotas:

Supported Query Tags for UserQuotas:

## **QUERYING SERVERS BY NAME & ID:**

# **nas\_server -query:\*** -format:'%s\n' -Fields:Name

server\_2

server\_3

**Note:** Example of query for list of all server names

# **nas\_server -query:\*** -format:'%s\n' -Fields:Id

## **QUERY EXAMPLES:**

\$ **nas\_cel -exec id=1 "nas\_fs -s fs04"**

total = 3223813 avail = 1755380 used = 1468433 ( 45% ) (sizes in MB) ( blockcount = 6704695296 )

volume: total = 3273777 (sizes in MB) ( blockcount = 6704695296 )

# **nas\_disk -query:\*** -Fields:'id,DiskType,StorageProfile,Protection' -format:'Disk id %s has %s, %s, type %s\n'

Disk id 1 has CLSAS, clarsas\_archive, type RAID5(4+1)

Disk id 2 has CLSAS, clarsas\_archive, type RAID5(4+1)

# **nas\_disk -query:\*** -Fields:'id,DiskType,StorageProfile,Protection' -format:'Disk id %s has %s, %s, %s type\n' |head

Disk id 1 has STD, symm\_std, 2-Way Mir type

Disk id 2 has STD, symm\_std, 2-Way Mir type

# /nas/bin/nas\_fs -query:\*

-format:"\n%-40s%q" -fields:name,disks -query:\*

-format:"%s" -fields:name

root\_fs\_1 root\_disk

root\_fs\_2 root\_disk root\_ldisk

## **QUERYING FILE SYSTEM ID & NAME:**

# **nas\_fs -query:Name==fs9** -format:'%s\n' -Fields:Id

# **nas\_fs -query:id==39** -format:'%s\n' -Fields:Name

## **OPERATOR DEFINITIONS:**

= having the pattern

== exact string match

=- integer minimum

=+ integer maximum

=\* any

=^ not having the pattern

=^= not an exact match

=^- not integer minimum, less than

=^+ not integer maximum, more than

=^\* not any

## **Example of Disks used by a Server:**

# **nas\_disk -query:servers=3 -fields:name,size** -format:'name=%L size=%L \n'

## **QUERYING FILE SYSTEMS USING NAS\_FS QUERY:**

# /nas/bin/nas\_fs -query:\*

-format:"\n%-40s%q" -fields:name,disks -query:\*

-format:"%s" -fields:name

>/tmp/filesystems.log

# tail /tmp/filesystems.log

-----abridged-----

cifs1

d7

root\_avm\_fs\_group\_3 d7  
event\_log d38 d39 d40  
root\_avm\_fs\_group\_36 d38 d39 d40

## **CELERRA MANAGER/CELERRA MONITOR:** [http://172.24.80.64:8000/top\\_level.htm](http://172.24.80.64:8000/top_level.htm)

**Prerequisites:** Celerra Monitor is a Java client/server application consisting of “Java Server” on Control Station [/nas/sys/nas\_mcd.cfg] and Java applets that run on Client’s Web Browser. Cookies must be enabled since SSL is used between Browser and Control Station connections. Celerra Monitor is only available with the Advanced Edition of Celerra Manager.

## **CELERRA MANAGER BASIC vs. ADVANCED MANAGER:**

Celerra Manager Advanced Edition adds ability to manage multiple Celerra systems, use Manual Volume Management, and use the CDMS (Celerra Data Migration Services) feature.

### **CELERRA MANAGER 5.2 (WebUI):**

- Basic Edition=free & Advanced Edition requires license fees [NFS and/or CIFS and CDMS Data Migration with NAS 5.3]
- Interface emulates ECC “Highland” product & combines WebUI, Native Manager, & Celerra Monitor

### **NAVIAGENT ON LINUX:**

# /nas/sbin/naviagent -help

@(#)Navisphere Agent Revision 6.6.0.3.8 for Linux on 12/12/2003 01:21:32

# /nas/opt/Navisphere/bin/naviagent -help

@(#)Navisphere Agent Revision 6.6.0.3.8 for Linux on 12/12/2003 01:21:32

**Note:** NAS Upgrades can sometimes fail to upgrade these directories with the same files and Celerra Manager may stop working.

## **CHANGING CONTROL STATION HOSTNAME:**

### **Use Celerra Manager:**

Celerras>Control Station Properties>Hostname: new\_csname

**Note:** Updates the /etc/hosts and /etc/sysconfig/network files with new hostname. Logoff and log back on to see new hostname.

### **Changing Name Manually:**

1. Edit /etc/hosts file to new name
2. Edit /etc/sysconfig/network file to new name
3. Reboot or Logoff and log back onto CS
4. #/bin/hostname [Verify name changed—if not, use #hostname newname]

**Note:** Or, use commandline GUI programs called “linuxconf” or “/sbin/netconf” [Must change Hostname in Hostname section and in the Adapter section for eth3, then reboot CS]

## **TROUBLESHOOTING CELERRA MANAGER(WEBUI)/CELERRA MONITOR:**

### **CELERRA MANAGER:**

Celerra Manager is a web-based gui for remote administration (aka WebUI) & operates as a secure SSL Apache Web Server that listens for requests from Browsers—its purpose is to configure and administer the Celerra from a GUI interface. Users log into WebUI with an authentication token that requires ability of browser to use HTTP cookies (set Browser to accept cookies). Jakarta Tomcat consists of .jsp pages/servlets and forwards Browser requests received from Apache, to the WebUI application. NAS 5.2 and higher uses XML requests from WebUI to the JServer daemon for graph, file system, and data mover notification inquiries initiated from the WebUI. The APL Task Manager listens to emcapl socket to receive and process XML tasks created by the WebUI, and returns output to socket for Tomcat to retrieve and send to Web Browser. JServer is allowed to consume up to 230MB of CS RAM, while the WebUI is allowed up to 96MB memory. Control Station will consume physical memory, then all the page memory that is allocated—if it exceeds the total allocated memory, then processes that are using the greatest amount of memory are killed.

#### **Basic Edition:**

Basic management tasks and can use with only a single Celerra.

#### **Advanced Edition:**

Manage multiple Celerras; create & manage volumes on the backend; CDMS Data Migration feature; Celerra Monitor; MVM wizard to create sub-structure to build file system on top of; Notifications tab to set file system & DM resource notifications

## **ACCESSING CELERRA MANAGER PAGES DIRECTLY VIA BROWSER:**

**Note:** You can attempt to display JavaServer pages directly by entering the proper .jsp page in a Web Browser /nas/tomcat/webapps/ROOT [All Control Station .jsp pages are stored here]

<https://10.241.169.53/action/shareDisplay> [After logging in, CIFS Share page will be displayed]

## [https://<system\\_name>/action/filesystemDisplay](https://<system_name>/action/filesystemDisplay)

filesystemCheckPtProperties  
filesystemExtend  
filesystemNew  
filesystemNotificationProperties  
filesystemPredict  
filesystemProperties  
filesystemQuotaProperties  
filesystemStats

### **USING COMMAND LINE TO QUERY:**

\$ /nas/bin/nas\_server -query:name=server\_3 -fields:MemoryUsage -format:"%s\n"

54

# nas\_fs -query:id==37 -fields:Name,type -format:"%s %s\n"

EPWxtra2\_DR uxf5

**Note:** FSID is 37 for the above query--Queries can be run against different nas components, such as nas\_server, nas\_fs, nas\_xml, etc.

# nas\_storage -query:\* -fields:StorageProcessorTable -format:"%q" -query:\* -fields:Address -format:"%s," 192.168.1.200,192.168.1.201

192.168.1.200,192.168.1.201

**Note:** Command queries for SP IP Address confirmation

### **CELLERRA MONITOR V2.3:**

Celerra Monitor is a java-based client/server application used to monitor performance of Celerra Servers & Storage subsystems and is generally launched from within the WebUI window (NAS 5.2 +). Celerra Monitor runs as a Java Client/Server application launched from a Browser. The **Java Server** [poller] runs on the Control Station and a **Java applet** [application] runs on the Client's Web Browser. JServer API is known as Celerra XML API (aka xhmp)—XML is a session-based protocol used to encapsulate data carried to and from Control Station and Data Movers. Java archive files (.jar) are downloaded to Client browser from Control Station when connecting.

**Default Configuration used Celerra Manager>Tools>Celerra Monitor:**

/nas/http/htdocs/cmv2/cmv2.jnlp

### **CELLERRA XHMP PROTOCOL:** New feature with CFS 4.0 release—“XML Packet Messaging”

EMC W3C-compliant messaging protocol used for managing a Celerra using XML V2 to carry data messages and HTTP POST to communicate between Control Station & DM.

<http://10.241.168.51/schemas/xhmp/xhmp.htm>

Essentially, the XML API resides on the front-end of the Apache HTTP Server, receives Client Browser Request Packets using HTTP POST, passes to Apache JServer Servlet layer for parsing. All requests are sequential and Session-based. Replies are sent in Reply Packets & sent back to the Client.

### **XHMP OPERATION:**

**XML API V2:** Possible future replacement of JServer XML API V1

#### **Start & Stop:**

XML API V2 is started by configuration found in /nas/sys/nas\_mcd.cfg--start script is [/nas/sbin/start\\_xml\\_api\\_server](/nas/sbin/start_xml_api_server)

```
daemon "XML API Server"
executable  "/nas/sbin/start_xml_api_server"
optional    yes
canexit     yes
autorestart yes
ioaccess    no
```

In order to stop the XML API, comment out entries in nas\_mcd.cfg & shutdown using [/nas/sbin/hup\\_api](/nas/sbin/hup_api)

#### **Configuration & Ports Used:**

Configuration file is /nas/sys/xml\_api.conf, and contains port information for JServer, APL, & Indications Manager.

#### **Logs:**

→/nas/log/cel\_api.log XML API V2 Server log for User requests, APL requests & replies, & APL indications  
→/nas/log/webui/cel\_api.log XML API V2 servlet debug log  
→/nas/log/webui/cel\_api\_error.log XML API V2 servlet error log

#### **Processes:**

# ps fax |grep -i api

17476 pts/0 S 0:00 \\_ /bin/sh ./start\_xml\_api\_serve

# ps fax |grep java |grep -i xmxa180

```
2240 ? S 0:07 | \_ /usr/java/bin/java -server -Xmx180m -Xss10485
2337 ? S 0:00 | \_ /usr/java/bin/java -server -Xmx180m -Xss1
2339 ? S 37:10 | \_ /usr/java/bin/java -server -Xmx180m -
2342 ? S 0:00 | \_ /usr/java/bin/java -server -Xmx180m - [Multiple processes are running for java server]
```

### **ACCESSING CELERRA MONITOR DIRECTLY WITHOUT CELERRA MANAGER:**

1. Have JRE1.4.2 installed on Client computer
2. Use following URL to access Monitor & enter username & password at prompt:

**<http://172.168.2.20/cmv2/cmv2.jnlp>**

**Note:** This view shows Data Movers, Storage, File Systems, Volume Config, Alerts, System Log, OSM Log, & Graphs

### **BASIC THINGS TO CHECK OR DO WHEN TROUBLESHOOTING:**

- In all cases, you must first determine if this a WebUI or Celerra Monitor issue, or both. Always determine if the WebUI itself is operational and whether it is only Celerra Monitor that is not operational.
- If it is a WebUI issue, you must determine if a NAS Upgrade or other change has been done recently. If not, then the most likely problem is that the Apache & Tomcat processes are not recoverable except for a CS reboot—try to get customer to allow CS reboot.
- But more often than not it will be a Celerra Monitor. Again, determine if Monitor has ever been working correctly, and if so, whether an upgrade has just occurred.
- Obtain screen capture of problem from customer, if applicable
- Verify Browser and JRE versions in use and try access from a different Client and O/S [if possible]
- Verify that customer is not use a Proxy Server for the Web Browser [Uncheck ‘Use Proxy Server in Celerra Monitor’]

**Note:** If proxy does not consistently translate Client’s IP address, then Monitor will not operate correctly, especially if traffic is returning to Monitor with Proxy IP address but requests are being presented with Client’s real IP address.

→ Verify that Celerra Monitor is properly enabled via the #nas\_license –l output

```
advancedmanager online
```

### **MINIMUM SOFTWARE, JRE, & BROWSER REQUIREMENTS:**

→ Java Runtime Environment (JRE) 1.4.2 recommended, or higher on Client system [download from [java.sun.com/getjava](http://java.sun.com/getjava) to update Client system]. JRE 5.0 is available, but there are issues with this latest version.

→ Internet Explorer 6.0 SP2 is preferred

**Note:** Please note that Celerra Manager is not currently compatible with Internet Explorer 7.0—see Primus emc147371

→ Netscape 6.2.3 or higher, with 7.1 the preferred

→ Firefox 1.5 supported

→ Java Web Start package is required for Celerra Monitor & is included in the JRE package

→ Don’t used Proxy connections, may be unable to resolve hostname & IP address and login will fail

→ NAS 5.4 requires that Browsers accept all cookies, or else have CS in trusted site list

**Note:** If you are using a locale other than English, you must use JRE 1.3.1\_02-I to use Celerra Monitor!

### **BASIC CELMGR CELMONITOR TROUBLESHOOTING STEPS:**

#### **1. GET EXACT DESCRIPTION OF PROBLEM/FAILURE:** Check Primus & Logs

#### **2. VERIFY CONTROL STATION PROCESSES:** (**Apache/httpd, Tomcat, JServer, APL\_Task\_Mgr**)

**Note:** WebUI does NOT need JServer process to run, but does use JServer when polling the data movers for system information! JServer is a Server side process used for Celerra Monitor and also for graphs/resource notifications in WebUI.

**\$ ps -ef |grep -i jserv**

```
root 1816 1165 0 Jan06 ? 00:00:00 /bin/sh /nas/sbin/run_jserver
root 2867 2866 0 Jan06 ? 00:00:04 /bin/sh /nas/jserver/jexec
root 10408 10407 0 Jan06 ? 00:00:00 /nas/jserver/jserver_tail -f /na
root 10517 10515 0 01:14 ? 00:00:00 /nas/jserver/jserver_tail -f -c
```

**Note:** Above output reflects a correctly running JServer

**\$ ps -axlww |grep -i tomcat**

```
100 0 3134 1 9 0 284848 41084 wait_f S ? 0:15 /nas/http/webui/tools/j2sdk1.4.2_01/bin/java -server -
Djava.endorsed.dirs=/nas/tomcat/common/endorsed -classpath /nas/http/webui/tools/j2sdk1.4.2_01/lib/tools.jar:/nas/tom
cat/bin/bootstrap.jar -Dcatalina.base=/nas/tomcat -Dcatalina.home=/nas/tomcat -D java.io.tmpdir=/nas/tomcat/temp
org.apache.catalina.startup.Bootstrap start
```

**Note:** It is not unusual to see dozens of Tomcat processes running

**\$ ps -eafl |grep -i apache**

```
140 S apache 6897 2278 0 69 0 - 3297 semop Jan06 ? 00:00:01
/nas/sbin/httpd -D HAVE_PERL -D HAVE_SSL -f /nas/http/conf/httpd.conf
140 S apache 13167 2278 0 69 0 - 3287 semop Jan06 ? 00:00:01
```

**Note:** There may be dozens of httpd/apache processes running. Verify that the processes are running with a normal subroutine called “semop” or “semaphore operation”. If the processes are in a “wait\_if” state, could indicate a problem. Also check to ensure that the httpd daemon is running with proper configuration: -D HAVE\_PERL -D HAVE\_SSL You can verify its proper configuration by grepping for PERL from /nas/sbin/mcd\_helper file:

**\$ cat mcd\_helper |grep -i have\_perl**

\$SBIN/\$HTTPD -D HAVE\_PERL -D HAVE\_SSL \

**# ps fax |egrep httpd**

```
18504 pts/3 S 0:00 \ egrep httpd
17186 ? S 0:00 /nas/sbin/httpd -D HAVE_PERL -D HAVE_SSL -f /nas/http (Parent HTTPD process)
17187 ? S 0:00 \_ /nas/sbin/httpd -D HAVE_PERL -D HAVE_SSL -f /nas/
17188 ? S 0:00 \_ /nas/sbin/httpd -D HAVE_PERL -D HAVE_SSL -f /nas/
```

**# ps -ef |grep task**

```
nasadmin 16275 1 0 11:40 pts/2 00:00:00 /nas/sbin/apl_task_mgr
root 29181 1165 0 13:42 ? 00:00:00 /bin/sh /nas/sbin/start_apl_task
root 29235 29181 0 13:42 ? 00:00:00 /bin/sh /nas/sbin/start_apl_task
nasadmin 29236 29235 0 13:42 ? 00:00:00 /nas/sbin/apl_task_mgr
nasadmin 29239 29236 0 13:42 ? 00:00:00 /nas/sbin/apl_task_mgr
nasadmin 29240 29239 0 13:42 ? 00:00:00 /nas/sbin/apl_task_mgr
nasadmin 29242 29239 0 13:42 ? 00:00:00 /nas/sbin/apl_task_mgr
```

**Note:** Above represents normal processes for APL Task Manager with NAS 5.3 and higher (4 apl\_task\_mgr processes)

**# ps axfw |less |grep -i jserv**

```
8446 pts/0 S 0:00 \ grep -i jserv
10721 ? S 0:00 \_ /bin/sh /nas/sbin/run_jserver
11852 ? S 0:12 | | \_ /bin/sh /nas/jserver/jexec
1021 ? S 0:00 | | \_ /nas/jserver/jserver_tail -f /nas/log/sys_log
20322 ? S 0:00 | | \_ /nas/jserver/jserver_tail -f -c +0 /nas/log/cmd_log
```

**# netstat -ant |grep 9824**

```
tcp 0 0 127.0.0.1:9824 0.0.0.0:* LISTEN
tcp 0 0 127.0.0.1:9824 127.0.0.1:37822 TIME_WAIT
tcp 0 0 127.0.0.1:9824 127.0.0.1:37827 TIME_WAIT
tcp 0 0 127.0.0.1:9824 127.0.0.1:37832 ESTABLISHED
tcp 0 0 127.0.0.1:37832 127.0.0.1:9824 ESTABLISHED
```

**Note:** Above output indicates that a Client has an established socket connection to the WebUI at port 9824

**# ps -ef |grep mac** [Verify presence of running mac threads]

```
root 3975 1928 0 2004 ? 00:01:43 [macstat <defunct>]
root 3988 1928 0 2004 ? 00:00:49 /nas/sbin/macstat 3 8033
root 446 1928 0 Mar11 ? 00:00:01 /nas/sbin/macstat 2 8032
```

**# top -n 3 -b ><filename>**

### **3. RUN NAS XML COMMAND TO EACH DM TO VERIFY XML FUNCTION:**

**# nas\_xml -info:server -level:3 2>&1 |more**

```
<CELERRA SRC='controlstation'>
<CELERRA_MANAGEMENT_UNIT NAME='0001877005320002' TYPE='Anaconda'
PRODUCT_NAME='Celerra CNS' SERIAL_NO='ml2805001458'
```

**Note:** If you do not get a screen stack failure trace, XML is working properly. Review errors if command does dump.

### **4. INVESTIGATE LOGS:**

#### **TOMCAT (/nas/tomcat/logs):**

/nas/tomcat/logs/mod\_jk.log [look here for Tomcat hangs and errors]

/nas/tomcat/logs/catalina\_log.2004-12-29.txt

# cat /nas/tomcat/logs/catalina.out [records stops and starts of Tomcat]

Mon Jan 10 11:28:31 EST 2005 Starting tomcat web server.

Starting service Tomcat-Standalone

#### **APACHE/HTTP (/nas/http/logs):**

/nas/http/logs/error\_log [Apache errors from browser transactions]

/nas/http/logs/access\_log [Browser transactions]

#### **JSEVER (/nas/jserver/logs):**

/nas/jserver/logs/error\_log.0 [Java exceptions]

/nas/jserver/logs/system\_log | system\_log.1 [main JServer log]

/nas/jserver/logs/jvm.out [timestamp on this file will show last successful start time of JServer]

/nas/jserver/logs/problem\_report.0 [Reports on issues with commands being run—new with NAS 5.4]

**Note:** The error\_log & system\_logs will not show a current timestamp if the JServer process itself cannot be started—this would be an indication of a major failure, perhaps related to a disk database consistency issue on the Celerra—tie the last timestamp date to a possible event that crippled the JServer. JServer will retry to start itself 3 times before going to sleep, and retrying again. JServer process itself is only allowed to consume up to 230MB RAM. Important to point out that JServer events are generally logged in the sys\_log or alert\_log, and NAS 5.5 has added more specific JServer event codes:

### # nas\_event -l -f JServer

```
id    description
0    Log Errors/Warnings into /nas/jserver/logs/system_log.x
101   Call Home - JServer failed to come up after 3 retries
102   Call Home - Generated debug file xxx from the core dump
103   Call Home - JServer reports nas_cmd errors
104   Call Home - JServer generated problem report zip file /nas/jserver/x.zip
105   Call Home - API Server failed to come up after 3 retries
1001  Crossed the Jserver file system usage threshold
1002  Dropped below the Jserver file system usage threshold
```

### SYSTEM LOG (/nas/log/sys\_log):

```
Jan 9 13:31:31 2005 JServer:4:0 "JServer stopped"
Jan 9 13:32:21 2005 JServer:4:0 "Polling of data mover in slot 2 has been restarted"
Jan 9 13:32:31 2005 JServer:4:0 "Polling of data mover in slot 3 has been restarted"
Jan 9 13:32:33 2005 JServer:4:0 "CLARiiON APM00042103183 poller has been restarted."
```

**Note:** Above output shows that JServer was stopped and then restarted normally.

Jan 4 11:03:46 2005 JServer:4:100 Failed to come up after 3 retries, shutting down

**Note:** Above message shows a JServer that is unable to start, as opposed to a JServer process that starts, then shuts down later

### API TASK MANAGER/WEBUI LOGS (/nas/log/webui):

#### **/nas/log/webui/apl\_tm.log** [Logs Starts & Stops of APL Task Mgr, also stack trace errors, Internal Errors]

2005-12-07 18:19:22.609 ATM> Starting APL indication manager with options:

```
port=9824
start_periodic_cleanup=True
save_task_files=False
run_as_daemon=False
default_expire_age=3:0:0:0
task_log_file=/nas/log/webui/apl_tm.log
scheduler_log_file=/nas/log/webui/apl_sched.log
indication_log_file=/nas/log/webui/apl_ind.log
iport=8886
jport=8887
```

**Note:** Above output from apl\_tm.log shows restart of Indications Manager after Control Station reboot. Indications Manager is new with NAS 5.4 and runs on Port 9824

/nas/log/webui/webui.log [Celerra Mgr internal errors can show up here, Java exceptions with source code messages]

/nas/log/webui/alert\_log [Event Notifications from WebUI or generated by the System]

**Note:** See /nas/http/webui/etc/web\_client\_eventlog.cfg for manually created WebUI event notifications

/nas/log/webui/apl\_sched.log [Log for Snapsure scheduling or WebUI initiated events]

/nas/tasks [Jobs in progress or incomplete]

/nas/tasks/schedule [Directory contains schedules created by WebUI, such as Checkpoints, etc.]

/nas/log/webui/webui\_syr.log [New log NAS 5.4 tracking actions taken with WebUI]

### **#cat /nas/site/task\_mngr.cfg**

port=9824 →Check to make sure that this emcapl socket port is defined correctly in /etc/services

start\_periodic\_cleanup=True

save\_task\_files=False

run\_as\_daemon=False

default\_expire\_age=3:0:0:0

task\_log\_file=/nas/log/webui/apl\_tm.log

scheduler\_log\_file=/nas/log/webui/apl\_sched.log

indication\_log\_file=/nas/log/webui/apl\_ind.log

iport=8886

jport=8887

### **# cat services |grep -i 9824**

emcapl 9824/tcp

### NAS 5.6:

```
# cat /nas/site/task_mgr.cfg
```

```
run_as_daemon=False
log_file=/nas/log/webui/apl_tm.log
port=9827
mgmtd_port=9824
save_task_files=False
default_expire_age=3:0:0:0
```

#### **INDICATIONS MANAGER LOGS:** (NAS 5.4)

The Indications Manager manages cache coherency for Celerra Manager for objects such as fs, shares, exports, tree quotas, checkpoints. When configuration changes occur, the APL Indications Manager will update its cache. NAS 5.5 adds more objects to cache.

```
/nas or /nbsnas/log/nas_log.al.indication
/nas or /nbsnas/sbin/indication_mngr
/nas or /nbsnas/sbin/start_indication_mngr
/nas or /nbsnas/site/indication_mngr.cfg
```

```
# cat indication_mngr.cfg
```

```
dport=9825
bport=9826
run_as_daemon=False
indication_log_file=/nas/log/nas_log.al.indication
```

#### **SYMAPI LOG:**

```
/nas/log/symapi.log [For issues related to backend, check this log to see what errors are occurring]
```

#### **CLIENT LOGS FOR CELERRA MANAGER:**

```
C:\Documents and Settings\<username>\Application\Java\Deployment\log→.plugin & .trace files
```

#### **XHMP API LOGS: (new with NAS 5.5)**

```
/nas/log/cel_api.log → Logs requests, replies, indications
```

```
/nas/log/webui/cel_api.log | cel_api_error.log → Servlet debug & error logs, respectively
```

### **5. INVESTIGATE FILES IN JSERVER DIRECTORY:**

```
# ls -la /nas/jserver
```

|                   |          |             |             |                       |   |
|-------------------|----------|-------------|-------------|-----------------------|---|
| -rw-r--r--        | 1        | root        | root        | 2 Jan 9 13:31         | CONTROL_STATION_SLOT  |
| -rw-rw-r-         | 1        | root        | sys         | 6756490 Sep 13 19:04  | emcaux.jar  |
| -rw-rw-r-         | 1        | root        | sys         | 2921 Sep 13 19:04     | esmif.jar   |
| <b>-rw-r--r--</b> | <b>1</b> | <b>root</b> | <b>root</b> | <b>13 Jan 5 15:26</b> | <b>fail_setup_logged</b> [Indicator of a JServer problem—delete file before restarting] |
| drwxr-xr-x        | 2        | root        | root        | 4096 Jan 6 17:07      | event_config  |
| drwxrwxr-x        | 2        | root        | sys         | 4096 Sep 13 19:04     | icons   |
| -rwxr--r--        | 1        | root        | root        | 300 Jan 9 13:31       | jexec   |
| -rwxr--r--        | 1        | root        | root        | 38 Jan 9 13:31        | jexec_any   |
| -rwxr--r--        | 1        | root        | root        | 43 Jan 9 13:31        | jexec_server_conf_fc  |
| -rwxr--r--        | 1        | root        | root        | 47 Jan 9 13:31        | jexec_server_conf_fct   |
| <b>-rw-r--r--</b> | <b>1</b> | <b>root</b> | <b>root</b> | <b>2 Jan 5 15:30</b>  | <b>jserver_retry</b> [Indicator of a JServer problem—delete file before restarting]     |
| lrwxrwxrwx        | 1        | root        | root        | 13 Jan 6 13:12        | jserver_tail -> /usr/bin/tail   |
| -rwxrwxr-x        | 1        | root        | sys         | 20233 Sep 13 19:04    | libj2c.so   |
| -rwxrwxr-x        | 1        | root        | sys         | 39589 Sep 13 19:04    | libj2sapi.so  |
| -rw-r--r--        | 1        | root        | root        | 0 Jan 9 13:31         | Linux   |
| drwxr-xr-x        | 2        | root        | root        | 4096 Oct 6 14:31      | logs  |
| drwxr-xr-x        | 2        | root        | root        | 4096 Nov 3 13:27      | odb   |
| -rw-r--r--        | 1        | root        | root        | 4 Jan 9 13:31         | SCHEMA_VERSION  |
| drwxr-xr-x        | 4        | root        | root        | 4096 Nov 3 13:27      | sdb   |
| drwxrwxrwx        | 5        | root        | sys         | 4096 Dec 6 18:22      | servlet   |
| -rw-r--r--        | 1        | root        | root        | 1 Jan 9 13:49         | timer_stamp   |
| -rw-rw-r-         | 1        | root        | sys         | 510088 Sep 13 19:04   | voyager.jar   |
| -rw-rw-r-         | 1        | root        | sys         | 1130741 Sep 13 19:04  | xerces.jar  |

**Note:** If possible, compare .jar (Java Archive) files to known good system

```
-rw----- 1 root root 8994816 May 29 02:12 debug_of_core.6925
```

#### **Debug of core Files:**

**Note:** NAS 5.4 introduces a feature that looks for JServer core dump files every 10 minutes. If it finds a core file, it parses it through the “gdb” utility to produce a “debug\_of\_core” file. Regular ‘core’ dumps are removed every 10 days and ‘debug\_of\_core’ files are removed every 28 days.

### **6. CORRECTIVE ACTIONS:**

→Reboot CS0 for WebUI issues, if allowed

→For JServer issues, cleanup any jserver\_retry files, kill off any sleep 1000000000 processes, kill off any residual JServer processes, clean JServer database, conduct fresh restart:

**# rm jserver\_retry**

**# kill -9 22415** [example of sleep 1000000000 process]

**# /nas/sbin/js\_kill -f**

**# /nas/sbin/js\_fresh\_restart**

→Kill Apache/httpd, Tomcat, and APL\_task\_mgr processes—conduct manual restart of httpd/apache & tomcat

## **GRACEFUL WAY TO STOP & RESTART TOMCAT & APL TASK MANAGER PROCESSES:**

**# /nas/http/nas\_ezadm/etc/script restart**

**Note:** See AR89253. Script is currently calling out wrong path to the /nas/http/webui/etc/config file. Edit the /nas/http/webui/etc/script file to correct the path:

**. /nas/http/webui/etc/config**

**# /nas/http/nas\_ezadm/etc/script restart &** [NAS 5.6.43]

[1] 2892

Stopping tomcat web server.

Stopping apl\_task\_mgr.

Starting apl\_task\_mgr.

Starting tomcat web server.

EventLog : will load /nas/http/webui/etc/web\_client\_eventlog.cfg... done

## **STOPPING & RESTARTING APACHE HTTPD DAEMON:**

1. Identify and Kill the master Apache httpd process:

**# ps fax |egrep httpd**

```
18504 pts/3 S 0:00 \_ egrep httpd
17186 ? S 0:00 /nas/sbin/httpd -D HAVE_PERL -D HAVE_SSL -f /nas/http
17187 ? S 0:00 \_ /nas/sbin/httpd -D HAVE_PERL -D HAVE_SSL -f /nas/
17188 ? S 0:00 \_ /nas/sbin/httpd -D HAVE_PERL -D HAVE_SSL -f /nas/
```

**# kill 17186**

**Note:** Do not use kill -9 to stop the Apache httpd daemon

**# ps fax |egrep httpd**

```
18824 pts/3 S 0:00 \_ egrep httpd
```

2. Restart Apache httpd process using following command:

**# /nas/sbin/httpd -D HAVE\_PERL -D HAVE\_SSL -f /nas/http/conf/httpd.conf** [Restarts HTTPD/Apache Server]

3. Verify Apache process:

**# ps fax |egrep httpd**

```
19571 pts/3 S 0:00 \_ egrep httpd
19519 ? S 0:00 /nas/sbin/httpd -D HAVE_PERL -D HAVE_SSL -f /nas/http
19521 ? S 0:00 \_ /nas/sbin/httpd -D HAVE_PERL -D HAVE_SSL -f /nas/
```

**# /nas/http/webui/etc/tomcat restart** [Restarts Tomcat]

**# killall apl\_task\_mgr** [Stops apl\_task\_mgr—NAS\_MCD automatically restarts]

**Note:** Allow MCD to restart the apl\_task\_mgr processes

**#/nas/sbin/js\_cleandb**

**Caution:** Do not run the js\_cleandb without customer permission, as it deletes any historical performance statistics from the /nas/jserver/sdb directory. Older NAS versions also used to delete any customer alerts that have been setup, from the /nas/jserver/odb directory. AR67281 has been opened to see if current behavior of cleandb script can be changed. When running js\_cleandb with any GA version of NAS 5.5, it will prompt you whether you want to delete the statistical database information, and the alert configuration database. It does not remove Checkpoint schedules that have been created. More recently, AR85690 introduces two separate scripts, one to delete the statistical data, and one to delete customer notifications.

## **PRE-NAS 5.5 JS CLEANDB BEHAVIOR:**

**#/nas/sbin/js\_cleandb**

This script will clean Jserver database.

You will have to login as root to run this script.

All statistic history will be lost.

Do you wish to continue? [yes or no]: yes

Done.

**Note:** Prior to NAS 5.5, the js\_cleandb script would delete both statistical data and customer alerts!

### **CURRENT JS CLEANDB BEHAVIOR NAS 5.5.22.2:**

#### **# /nas/sbin/js\_cleandb**

This script will clean Jserver database.

You will have to login as root to run this script.

All statistic history will be lost.

Do you wish to continue? [yes or no]: yes

Warning - Deleting the JServer configuration data will remove also remove any user defined alert settings.

Do you wish to remove existing JServer configuration data? [yes or no]: no

Done.

**Note:** The wording is pretty confusing. If you select ‘Yes’ to the first question, it will delete the statistical data stored in the /nas/jserver/sdb directory. If you answer ‘No’ to the second question, it will KEEP the customer notifications/alerts that have been created in the /nas/jserver/odb directory—if you answer ‘Yes’ to both, it deletes both.

### **NAS 5.5.27.x JS CLEANDB CHANGES:** AR85690, fixed in GrandNapa, which will be NAS 5.5.27.x

**/nas/sbin/js\_cleandb** will delete only the historical statistical information data collected in the /nas/jserver/sdb directory

**/nas/sbin/js\_cleanconfig** will delete only the customer alerts/notifications that are stored in the /nas/jserver/odb directory

### **7. CONTROL STATION CHECKS:**

→ Verify Control Station hostname, interface configuration, and default routing are correct

[#uname –a #hostname #cat /etc/hosts #cat /etc/sysconfig/network #netstat -r]

#### **# uname -r**

2.4.9-34.5405.EMC

#### **# netstat -r**

Kernel IP routing table

| Destination | Gateway      | Genmask       | Flags | MSS  | Window | irtt | Iface |
|-------------|--------------|---------------|-------|------|--------|------|-------|
| 192.168.2.0 | *            | 255.255.255.0 | U     | 40 0 | 0      |      | eth0  |
| 192.168.1.0 | *            | 255.255.255.0 | U     | 40 0 | 0      |      | eth0  |
| 10.241.0.0  | *            | 255.255.0.0   | U     | 40 0 | 0      |      | eth1  |
| 127.0.0.0   | *            | 255.0.0.0     | U     | 40 0 | 0      |      | lo    |
| default     | 10.241.169.1 | 0.0.0.0       | UG    | 40 0 | 0      |      | eth1  |

### **8. WEBUI/CELERRA MONITOR ESCALATIONS:**

→ Run Eng. support script to gather necessary logs:

**#/nas/tools/collect\_support\_materials** [Bundles up log files and puts in /tmp; with NAS 5.6, logs are placed in /nas/var/emcsupport directory]

→ Provide NAS Version

→ Provide screenshot example of Client error seen by customer

→ Provide synopsis of issue and output of JServer processes and any errors noted in sys\_log

→ Determine if this is a new issue or never worked, and whether a code upgrade occurred

→ Verify Client Browser [i.e., IE 6.0, Netscape 6.2.3] and JRE versions [JRE 1.4.2 minimum]

### **9. VERIFY CONTROL STATION PARTITION USAGE:**

→ If the Control Station partitions are at 100%, Celerra Manager may not run without errors, as in following example:

#### **Key Symptom:**

#### **# df -h**

| Filesystem | Size | Used | Avail | Use% | Mounted on |
|------------|------|------|-------|------|------------|
| /dev/hda3  | 2.0G | 779M | 1.1G  | 41%  | /          |
| /dev/hda1  | 30M  | 2.7M | 26M   | 10%  | /boot      |
| none       | 251M | 0    | 250M  | 0%   | /dev/shm   |
| /dev/nde1  | 1.7G | 479M | 1.1G  | 100% | /nbsnas    |
| /dev/hda5  | 2.0G | 393M | 1.4G  | 21%  | /nas       |
| /dev/nda1  | 133M | 37M  | 96M   | 28%  | /nas/dos   |
| /dev/ndf1  | 1.7G | 77M  | 1.5G  | 5%   | /nas/var   |

**Note:** With /nbsnas at 100%, Celerra Manager returned following error: “Query Celerras All. Internal Error: Missing response from Celerra Manager agent” or “Internal Server Error....Apache 1.3.27 Server at 192.168.32.123 Port 443.”

**Key Message:** Feb 9 11:18:30 2005 JServer:4:0 "FS Usage exceeds 80%. DbMgr has been suspended [See Primus emc101854]

**Resolution:** Unfortunately, it can be difficult cleaning up files from /nbsnas because of the “dirstync” process that is tasked with keeping /nas and /nbsnas in sync. Stop NAS services before cleaning up /nbsnas & /nas directories.

### **NBSNAS 100% FULL:**

### **DELETING FILES FROM BOTH /NAS & /NBSNAS:**

1. Run #df -h
2. Stop NAS Services
3. Remount /nbsnas /nas /nas/dos /nas/var
4. Delete necessary files from /nas first, then same files from /nbsnas
5. Start NAS Services
6. #df -h to verify

**Note:** For example, if you need to cleanup JServer files, you would have to delete from both locations, /nas & /nbsnas

## **PROBLEM WITH WEB MANAGER UNABLE TO COMMUNICATE WITH APL MANAGER:**

### **Symptoms:**

/nas/log/webui/webui.log

Catching an exception in task [[15649](#)]

Class / Object Routine Nature of exception Effect

CLARIION\_DISK\_GROUP is\_spindle\_type\_ata have\_spindles:

<41252DE8> Precondition violated. Fail

/nas/log/webui/apl\_tm.log

ERROR 2005-05-04 13:24:55,272 [COMMON]: Empty response XML from APL for task Query file systems All:All

### **Celerra Manager File System Query:**

Internal Error: Missing response from Celerra Manager agent

**Workaround NAS 5.3:** (NAS 5.4 uses nas\_storage -check or server\_devconfig -create to force discovery)

1. Determine path to use for creating a temp gatekeeper file:

# nas\_disk -i id=1  
server = server\_2      addr=c2t0l0

2. Take the first path (c2t0l0), and make a temp gatekeeper file in /nas/dev

# touch /nas/dev/c0t0l0s2 (s2 specifies server\_2)

3. Run discover command:

# /nas/symcli/bin/symcfg discover –clariion

4. Remove the temp gatekeepr file:

# rm /nas/dev/c0t0l0s2

5. Verify results using:

# nas\_storage -i -a

**Cause of Problem:** Inconsistency in symapi database causes Celerra Manager error outlined above

## **UNABLE TO LAUNCH CELERRA MANAGER – Internal Server Error:**

Generally Apache or Tomcat issue after logging into Celerra Manager. Restart Apache and Tomcat. Check webui.log & mod\_jk.log.

### **INTERNAL ERROR – Celerra Manager Agent not Responding:**

Generally APL Task Mgr issue. Check for task processes, sys\_log, netstat –ant lgrep 9824

## **UNABLE TO LAUNCH CELERRA MONITOR – Connection refused HTTP Status 503:**

JServer processes not running.

### **NAS AGENT FOR ECC (CNN):**

ECC uses a NAS agent for discovering & polling (every 15 minutes) various information about the Celerra and its data movers, which uses xhmp queries to obtain Control Station-generated nas\_xml output. Information is then displayed in the ECC StorageScope GUI.

### **NAS LICENSE:**

# nas\_license -l

|                 |        |             |
|-----------------|--------|-------------|
| key             | status | value       |
| site_key        | online | 42 10 f4 09 |
| advancedmanager | online |             |
| nfs             | online |             |
| cifs            | online |             |
| enterprise      | online |             |
| snapsure        | online |             |
| replicator      | online |             |

## **REPLACING JAR FILES FOR JSERVER ISSUES Pre-5.5 Code:**

1. Copy replacement jar file to CS /nas/jserver
2. #chmod 664 emcaux.jar

3. Stop JServer: #/nas/sbin/js\_kill -f
4. Remove any /nas/jserver or /nbsnas/jserver\_retry files, and sleep 1000000 processes
5. Start JServer: #/nas/sbin/js\_fresh\_restart

### **REPLACING JAR FILE FOR JSERVER ON NAS 5.5.28 CODE:**

1. # cd /nas/j\_lib
2. # mv emcaux.jar emcaux\_ori
3. Copy new jar file into this directory
4. # chown root:sys emcaux.jar
5. # chmod 664 emcaux.jar
6. Restart JServer  
/nas/sbin/js\_shutdown  
/nas/sbin/js\_kill  
/nas/sbin/js\_fresh\_restart

### **OVERVIEW OF CELERRA MANAGER (WEBUI) ARCHITECTURE:**

#### **MASTER CONTROL DAEMON (MCD) & MCD HELPER SCRIPT:**

Overall management and control of Celerra Manager. MCD starts and maintains APL Task Manager, JServer, & Trimmer Logs while MCD Helper script starts Apache and Tomcat applications.

#### **WEBUI:**

WebUI creates XML documents that describes ‘tasks’ to be performed, and posts the document to the socket used by APL Task Manager, defined by emcapl in /etc/services. Default port for emcapl is 9824. Requests for JServer are sent as XML documents over HTTP to establish session with JServer.

**webui\_syr.log** → Log used for System Reporting (SYR) of statistics for Celerra Manager

#### **APL TASK MANAGER (APpliance Layer):**

Runs as daemon started by MCD using /nas/sbin/start\_apl\_task\_manager. Listens on port 9824 (emcapl socket defined in /etc/services) for XML tasks generated by WebUI. Processes XML tasks and returns results to socket in XML format for Tomcat to retrieve and forward to Browser. Most tasks logged to /nas/log/webui/apl\_tm.log & /nas/tasks/JID number (Job ID number).

#### **Stopping & Restarting APL Task Manager:**

```
# killall apl_task_mgr [NAS MCD will automatically restart]  
# ps -ef |grep task  
nasadmin 16275 1 0 11:40 pts/2 00:00:00 /nas/sbin/apl_task_mgr  
root 29181 1165 0 13:42 ? 00:00:00 /bin/sh /nas/sbin/start_apl_task  
root 29235 29181 0 13:42 ? 00:00:00 /bin/sh /nas/sbin/start_apl_task  
nasadmin 29236 29235 1 13:42 ? 00:00:00 /nas/sbin/apl_task_mgr  
nasadmin 29239 29236 0 13:42 ? 00:00:00 /nas/sbin/apl_task_mgr  
nasadmin 29240 29239 0 13:42 ? 00:00:00 /nas/sbin/apl_task_mgr  
nasadmin 29242 29239 0 13:42 ? 00:00:00 /nas/sbin/apl_task_mgr
```

**Note:** Above reflects normal apl\_task\_mgr processes

#### **APL INDICATION MANAGER:**

New with NAS 5.4, the Indication Manager manages Browser cache coherency for Celerra Manager for objects such as volumes, file systems, shares, exports, tree quotas, checkpoints.

**apl\_ind.log** → Log used by APL Indications Manager for object caching

**ipc.log** → Internal log for Indication communication

#### **JSERVER:**

Runs as daemon on CS under nas\_mcd control and serves as server side process for Celerra Monitor, started by /nas/sbin/run\_jserver. Client side application is launched using .jar files which are downloaded to the client when launching Celerra Monitor (/nas/http/htdocs/cmv2). The JServer API is known as Celerra XML or XMHP. When invoked, Jserver polls data movers, file systems, and storage systems for status, configuration, and performance statistics, sending information to WebUI via XML documents using HTTP. Although the Polling interval defaults to 5 minutes, you can set to a different value from Celerra Manager, though the historical data only captures a minimum of 5 minutes. Most logs found in /nas/jserver/logs directory. Live graph displays, file system info, and Data Mover notifications are all processed via XML requests to the JServer daemon. JServer polling intervals are set via Celerra Monitor, and now Celerra Mgr with NAS 5.4. Default data retention for polling information on components is 1 week, configurable up to 26 weeks, and file system usage & capacity is hard-coded at 12 weeks. Configuration changes are updated every 20 minutes and recorded in the /nas/jserver/sdb/dbmgr\_persist file. JServer also monitors and acts on messages given by the Indications Manager.

#### **JSERVER STARTUP:**

→ nas\_cmd queries are conducted in DART to gather configuration info

→ Storage API queries are made to update backend information

## **WHICH NAS COMPONENTS USE JSERVER?**

- Celerra Monitor
- Celerra Manager
- ECC Solutions Enabler
- XML API v1 Server [JServer & APL APIs for 3<sup>rd</sup> Party applications]

## **STATISTICS & NOTIFICATION INFORMATION:**

### **/nas/jserver/odb**

File System and Data Mover usage notification settings are stored in this directory as binary information

### **Stopping & Restarting JServer:**

**#/nas/sbin/js\_kill -f**

**#/nas/sbin/js\_fresh\_restart**

**# ps -ef |grep -i jserv**

```
root  21034 1165 0 Jan09 ?    00:00:00 /bin/sh /nas/sbin/run_jserv
root  21305 21304 0 Jan09 ?    00:00:01 /bin/sh /nas/jserver/jexec
root  21810 21809 0 Jan09 ?    00:00:00 /nas/jserver/jserver_tail -f -c
root  21845 21844 0 Jan09 ?    00:00:00 /nas/jserver/jserver_tail -f /na
```

**Note:** Above output reflects normal JServer processes

### **DUMPING DEBUG OUTPUT OF JSERVER:**

**#/nas/sbin/js\_kill -debug**

**#cat /nas/jserver/logs/jvm.err**

Full thread dump Java HotSpot(TM) Server VM (1.4.2\_04-b05 mixed mode):

```
"com.emc.cs.Timer" prio=1 tid=0x08197fb8 nid=0x784b in Object.wait() [bcbff000..bcbff908]
```

```
at java.lang.Object.wait(Native Method)
- waiting on <0x45c963c8> (a java.lang.Object)
at com.emc.cs.Timer.waitInternal(Timer.java:251)
- locked <0x45c963c8> (a java.lang.Object)
at com.emc.cs.Timer.run(Timer.java:179)
```

**Note:** Output may give indication of what process is hanging

## **SYSLOG TRIMMER:**

Part of nas\_mcd.cfg script, checked every 24 hours and backed up when exceeding 1000 lines—4 backup versions kept.

## **APACHE WEB SERVER:**

Serves as secure HTTP Web Server when using SSL, listens to browser requests, and requires session token (HTTP cookie) for user authentication, generated by CS after user logs on. Browser requests are sent and retrieved to WebUI via Apache using Jakarta Tomcat. Apache version 1.3.27.

### **Stopping & Restarting Apache/HTTP:**

#### **/nas/sbin/mcd\_helper**

**#kill <master http pid>** [Kill all httpd processes & verify that Apache & HTTP processes are stopped]

**#/nas/sbin/httpd -D HAVE\_PERL -D HAVE\_SSL -f /nas/http/conf/httpd.conf** [Use this to restart Apache/http]

**# ps -eafl |grep -i apache**

```
140 S apache 13258 27562 0 69 0 - 3238 do_sel 12:58 ? 00:00:00
/nas/sbin/httpd -D HAVE_PERL -D HAVE_SSL -f /nas/http/conf/httpd.conf
140 S apache 13331 27562 0 69 0 - 3247 semop 12:59 ? 00:00:00
/nas/sbin/httpd -D HAVE_PERL -D HAVE_SSL -f /nas/http/conf/httpd.conf
```

**Note:** There may be dozens of Apache and HTTP processes running on a system

**# ps -eafl |grep -i http**

```
/nas/http/webui/tools/j2sdk1.4.2_01/bin/java -server -Djava.endorsed.dirs=/nas/t
040 S root 22686 22681 0 69 0 - 69725 rt_sig 14:53 pts/2 00:00:00
/nas/http/webui/tools/j2sdk1.4.2_01/bin/java -server -Djava.endorsed.dirs=/nas/t
040 S root 22687 22681 0 69 0 - 69725 rt_sig 14:53 pts/2 00:00:00
```

## **JAKARTA TOMCAT:**

Tomcat supports web browser requests using the Apache Web Server (WebUI jsp pages and applets/servlets). Version 4.0.4 NAS 5.2, and 4.1.29 for NAS 5.3. Tomcat-Apache mod\_jk module version 1.2.5 is used as the communication pipe between Apache & Tomcat

### **Stopping & Restarting Tomcat:**

**/nas/sbin/mcd\_helper**

**# /nas/http/webui/etc/tomcat restart**

**# ps -axlww |grep -i tomcat**

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
100 0 22679 1 9 0 278900 85816 wait\_f S pts/2 0:15 /nas/http/webui/tools/j2sdk1.4.2\_01/bin/java -server -Djava.endorsed.dirs=/nas/tomcat/common/endorsed -classpath /nas/http/webui/tools/j2sdk1.4.2\_01/lib/tools.jar:/nas/tomcat/bin/bootstrap.jar -Dcatalina.base=/nas/tomcat -Dcatalina.home=/nas/tomcat -D java.io.tmpdir=/nas/tomcat/temp org.apache.catalina.startup.Bootstrap start

**Note:** There may be dozens of Tomcat processes running

#tail -20 /nas/tomcat/logs/catalina.out [Example of normal Tomcat startup in log]

Mon Jan 10 14:53:29 EST 2005 Starting tomcat web server.

tomcat start: http ready to go

Starting service Tomcat-Standalone

Apache Tomcat/4.1.29

Jan 10, 2005 2:54:01 PM org.apache.struts.util.PropertyMessageResources <init>

INFO: Initializing, config='org.apache.struts.util.LocalStrings', returnNull=true

Jan 10, 2005 2:54:01 PM org.apache.struts.util.PropertyMessageResources <init>

INFO: Initializing, config='org.apache.struts.action.ActionResources', returnNull=true

Jan 10, 2005 2:54:04 PM org.apache.struts.util.PropertyMessageResources <init>

INFO: Initializing, config='org.apache.webapp.admin.ApplicationResources', returnNull=true

Jan 10, 2005 2:54:34 PM org.apache.jk.common.ChannelSocket init

INFO: JK2: ajp13 listening on /127.0.0.1:8009

Jan 10, 2005 2:54:34 PM org.apache.jk.server.JkMain start

INFO: Jk running ID=0 time=16/268 config=/nas/tomcat/conf/jk2.properties

### **How to Set Linux Hostname to Make Celerra Manager Work:**

Updating Hostname on Box: vi \$/etc/HOSTNAME [Linux] or vi /etc/nodename [SCO] vi /etc/hosts file #set name celctrl01

**Caution:** Must be logged in as "root" and use 'linuxconf' or 'netconf' GUI program to change Hostname permanently!

Updating Hostname on External Interface Card: #linuxconf or #netconf or edit directly; #/etc/sysconfig/network "Hostname"

### **CELERRA MONITOR:**

Celerra Monitor is a java-based client/server application used to monitor performance of Celerra Servers & Storage subsystems.

Celerra Monitor runs as a Java Client/Server application launched from a Browser or Desktop Windows or Unix machine. The **Java Server** [poller] runs on the Control Station and a **Java applet** [application] runs on the Client's Web Browser.

#### **ACCESSING CELERRA JSERVER API HELP FILES:**

<http://10.241.169.43/schemas/xhmp/xhmp.htm> [Celerra Control Station]

#### **PURPOSE:**

**File System Statistics:** File System I/O performance history, HWMarks, Alerts, Inode Usage, predictive monitoring of fs capacity

**DataMover Statistics:** DM Configuration, NFS, CIFS, TCP Stats, CPU & Memory Usage, alerts.

**Storage Backend Stats:** Director Performance Stats\*

\***Note:** This statistic is not yet available with CLARiiON platforms [NAS 5.1 or lower]

→Used to view Run Time, Logs, configure Polling times for data collection

→Status Monitor provides flashing icon for system problems

→Data export capability to CSV export to files

### **USING CELERRA MONITOR WITH MULTIPLE SYMMS ATTACHED TO A SINGLE CELERRA:**

Step 1. Run \$/nas/bin/nas\_symm -l [Outputs Serial Numbers of Symmetrixes]

Step 2. Run \$/nas/bin/nas\_symm -i serialnumber [Verifies SymAPI is parsing through the GateKeeper to the respective Symm]

Step 3. Run \$server\_devconfig server\_x -p -s -a

#### **Check to see if Gatekeepers are properly recognized:**

tid/lun= 0/15 type= disk val= -99 info= 5266470B4050 diskerr= -1 [Correct entry for a GateKeeper]

tid/lun= 0/15 type= disk sz= 3 val= -1 info= 52667202E010 [Incorrect entry]

**Caution:** Celerra Monitor will not see 4-Port FA's at this level of code [fixed in 4.0 +]

### **TROUBLESHOOTING CELERRA MONITOR:**

**Browser Error:** 404 Not Found. The requested URL...was not found on this server. [Java Server is not starting]

Resolution? Ensure /etc/hosts and /etc/nodename are identical!

**Processes on CS:** \$ps -ef | grep java lagent lmacstat

**/nas/jserver/sdb/control\_station/data\_movers/1/cifs\_stats:**

Check for evidence that Stats are populating the Data Mover directories—if not, Jserver is not working

#### **VERIFY HTTP.CONF FILE: #ps -ef |grep http**

a. Telnet to port 8000; Type GET and check default HTM page

b. If connection refused may indicate bad "httpd.conf"

c. Copy good "httpd.conf" over: \$cp /nas/http/conf/httpd.conf /etc/httpd/conf/httpd.conf

## **CONTROL STATION MEMORY UTILIZATION HIGH DUE TO JAVA PROCESSES:**

```
$ ps -axuw grep java
root 10306 0.2 4.5 170348 23500 ? S 2002 30:50 /usr/java/bin/i38 [4.5=% Memory Utilization for PID 10306]
6/green_threads/java -classic -Xmx128M com.objectspace.voyager.system.Main 8010
-r -q -b com.emc.cs.Initiator 8010 8000 /nas 1 2
root 11436 0.0 3.2 180228 16852 ? S 2002 9:23 /usr/java/bin/i38
6/native_threads/java -Xss1048576 com.emc.cs.symmpoller.SymmAgentSapi
```

## **STOPPING & RESTARTING JSERVER:**

1. Remove any `/nas/jserver/jserver_retry` if present
2. `#kill -9 sleep 1000000` if required
3. **`#/nas/sbin/js_kill -f`**
4. Startup using **`#/nas/sbin/js_fresh_restart`**
5. Grep for jserver processes:

```
# ps -ef |grep -i jserver
root 24027 869 0 20:16 ? 00:00:00 /bin/sh /nas/sbin/run_jserver
root 24167 24103 0 20:16 ? 00:00:00 /bin/sh /nas/jserver/jexec
root 25294 24103 0 20:19 ? 00:00:00 /nas/jserver/jserver_tail -f -c
root 25350 24103 0 20:19 ? 00:00:00 /nas/jserver/jserver_tail -f /na
```

## **ADDITIONAL JS START & CLEANUP OPTIONS:**

### **`#/nas/sbin/js_cleandb`**

This script will clean Jserver database.

You will have to login as root to run this script.

All statistic history will be lost.

Do you wish to continue? [yes or no]: yes

Done.

### **`#/nas/sbin/js_fresh_restart`**

### **`# ps -ef |grep js`**

```
root 11258 1 0 09:06 ? 00:00:00 /bin/sh /nas/sbin/js_start
root 11457 28660 0 09:06 ? 00:00:00 /bin/sh /nas/sbin/run_jserver
root 11588 11392 0 09:06 ? 00:00:00 /bin/sh /nas/jserver/jexec
```

**Creating Large Celerra Monitor Database for Analysis:** `#/nas/jserver/sdb` [move to another location & softlink to it]

**Changing Default Passwd "user" & "root":** `/nas/http/esteem`

## **Three Key Measurements that Celerra Monitor Can Provide:**

- 1.) Assessing activity on a FileSystem basis—not read via any CLI commands
- 2.) View Input/Output activity for a specific DataMover interface
- 3.) CPU & RAM usage for system

## **USING CONSOLE UTILITY:**

→Special utility used by Support & EE to view & manage different NAS processes, such as Tasks, Databases, Events, Logs, etc.

### **`/nas/tools/console`**

1) Create link to the console utility

```
# cd /nas/tools
```

```
# ln -s console _console
```

2) Run the console utility

```
# ./_console
```

console 127.0.0.1 6666

Commands:

```
cic.dump
cic.help
cic.setlog
cic.settimeout
db.bdb
db.list
```

```
dummy.create
dummy.help
evmonitor.dump
evmonitor.help
evmonitor.restart
evmonitor.setinterval
legacy.dump
log.help
log.info
log.restore
log.set
shlib.help
shlib.list
shlib.load
shlib.unload
shlib.version
sighndl.r.dump
sighndl.r.help
stats.create
stats.dump
stats.help
taskmgr.destroy
taskmgr.dump
taskmgr.setmaxtasks
watchdog.dump
watchdog.eventlog
watchdog.help
watchdog.monitor
watchdog.setinv
watchdog.setparam
watchdog.thread
```

Command {Enter "exit" to quit} >> exit

### **EXAMPLE OUTPUTS:**

**Command {Enter "exit" to quit} >> db.list**

List of all registered databases:

---

| DbId | Access | Key/pg | PgSize | DbName |
|------|--------|--------|--------|--------|
|------|--------|--------|--------|--------|

---

|   |       |   |      |                           |
|---|-------|---|------|---------------------------|
| 1 | BTree | 2 | 4096 | /db.taskid                |
| 2 | BTree | 2 | 4096 | /db.command.context       |
| 3 | BTree | 2 | 4096 | /db.task.context          |
| 4 | BTree | 2 | 4096 | /db.communication.context |
| 5 | BTree | 2 | 4096 | /db.application.context   |
| 6 | BTree | 2 | 4096 | /db.service.context       |

---

List of all registered FlatFile databases:

---

| DbId | DbName |
|------|--------|
|------|--------|

---

|    |                         |
|----|-------------------------|
| 7  | NAS_DB//volume/filesys  |
| 8  | NAS_DB//server/servers  |
| 9  | NAS_DB//volume/pools    |
| 10 | NAS_DB//server/vdm/vdms |
| 11 | NAS_DB//site/chosts     |
| 12 | NAS_DB//tasks/          |
| 13 | NAS_DB//volume/volumes  |
| 14 | NAS_DB//tasks/          |
| 15 | NAS_DB//sys/            |

---

**Command {Enter "exit" to quit} >> watchdog.dump**

```

Monitoring interval: 30
Event mode: 0
Monitor mode: 1
Registration list
Caller      Interval
=====
Watchdog      60
EventMgr      30
TaskMgr       30
AdminServer   30
CLTimer       30
HouseKeeping  120
CicSender     240

```

**Command {Enter "exit" to quit} >> cic.dump**

```

Default connection timeout: 20
Default response timeout: 100
Default async timeout: 900

```

### **CELERRA PARAMETERS:**

→ param tuning is done using /nas/site/slot\_param or /nas/server/slot\_x/param file to make changes outside the default values  
 → See server\_param command for outputting & setting param values NAS 5.3 +  
 → Note that NAS 5.4 + has begun hiding many parameters from view, though they may still exist  
 → NAS 5.3 + are now beginning to port commands to xml output, removing the output limitation—not all commands ported yet

### **DATA MOVER PARAM FILES:**

**/nas/server/slot\_x/param    /nas/server/slot\_x/paramtab    /nas/site/slot\_param** (Global param file)

**Note:** During Server reboot, the last param settings applied are from the local /nas/server/slot\_x and if different than the /nas/site/slot\_param, will override the global param value

### **EXAMPLE OF HIDDEN PARAM & FULLDESCRIPTION OPTION NAS 5.5:**

```

# .server_config server_2 -v "param cifs wins.Refresh" -->Hidden cmd, won't output current value
1143835291: SMB: 3: cifs.wins.Refresh must be <= 3600 sec and >= 30 sec
1143835291: ADMIN: 3: Param-attached function failed.

# .server_config server_2 -v "param fulldescription cifs wins.Refresh" -->Outputs current & default values in hex
cifs.wins.Refresh 0x01fd8798 0x00000384 0x00000384
1143835301: ADMIN: 4: Command succeeded: param fulldescription cifs wins.Refresh

```

**\$ server\_param server\_2 -facility usrmap -list**

```

server_2:
name          facility default  current  configured
maxgid        usrmap 2147483647 2147483647
minuid        usrmap    16      16
maxuid        usrmap 2147483647 2147483647
mingid        usrmap    16      16

```

### **USEFUL COMMANDS TO DUMP PARAMETERS AND VALUES:**

**.server\_config server\_6 -v 32768 “param”**

**Note:** NAS 5.2 and higher has increased the buffer output for .server\_config commands to 32k from 16k

**# .server\_config server\_2 -v "param cifs verbose" | grep mac** → New with Cognac release, shows more than 5.5  
 cifs.macOS INT 0x09efd718 0 0 (0,4294967295) FALSE REBOOT 'NA'

**.server\_config server\_6 “help” “fcp bind help” “wins help” “ds help” “secmap help” “help /s”**

**.server\_config server\_3 -v "help /s"**

**.server\_config server\_2 "printstats all"**

### **SERVER CONFIG COMMANDS & PARAM SETTINGS: NAS 5.5**

| Name             | Location   | Current    | Default    |
|------------------|------------|------------|------------|
| IPCOPY.snaplimit | 0x032df8b8 | 0x00000060 | 0x00000060 |
| NDMP.bufsz       | 0x032dcc90 | 0x00000080 | 0x00000080 |

```

NDMP.convDialect      0x032dce64 '8859-1' '8859-1'
NDMP.dialect          0x032dce60 " "
NDMP.maxProtocolVersion 0x01585e94 0x00000004 0x00000004
NDMP.md5              0x032dce68 0x00000000 0x00000000
NDMP.scsiReserve      0x01585e98 0x00000001 0x00000001
NDMP.v4OldTapeCompatible 0x01585e9c 0x00000001 0x00000001
NTsec.neverAddDC       0x01571470 0x00000000 0x00000000
PAX.allowVLCRestoreToUFS 0x0157080c 0x00000000 0x00000000
PAX.checkUtf8Filenames 0x015707d8 0x00000001 0x00000001
PAX.dump               0x015707dc 0x00000000 0x00000000
PAX.filter.caseSensitive 0x0156ea84 0x00000001 0x00000001
PAX.filter.numDirFilter 0x0156ea88 0x00000005 0x00000005
PAX.filter.numFileFilter 0x0156ea90 0x00000005 0x00000005
PAX.nFTSThreads        0x0156e568 0x00000008 0x00000008
PAX.nPrefetch           0x015707ec 0x00000008 0x00000008
PAX.nRestore            0x015707e8 0x00000010 0x00000010
PAX.nThread             0x015707e4 0x00000040 0x00000040
PAX.noFileStreams       0x01570814 0x00000000 0x00000000
PAX.paxReadBuff         0x015707fc 0x00000040 0x00000040
PAX.paxStatBuff         0x01570804 0x00000080 0x00000080
PAX.paxWriteBuff        0x015707f8 0x00000040 0x00000040
PAX.readWriteBlockSizeInKB 0x01570800 0x00000040 0x00000040
PAX.scanOnRestore        0x01570810 0x00000001 0x00000001
PAX.writeToArch          0x015707f0 0x00000001 0x00000001
PAX.writeToTape           0x015707f4 0x00000001 0x00000001
RCP.tcpwindow           0x032deb78 0x00000000 0x00000000
RCP.tcpwindowlowat      0x032deb88 0x00000000 0x00000000
VRPL.freeze              0x032db3d8 0x00000000 0x00000000
VRPL.nchunkreserved     0x02f88460 0x00000002 0x00000002
VRPL.ncopythreads       0x02f885b8 0x0000000a 0x0000000a
VRPL.readonly            0x032db418 0x00000000 0x00000000
ana.rxburst              0x02014b98 0x00000100 0x00000100
arp.holddown             0x0153e700 0x00000258 0x00000258
cfs.deleteDelay           0x0154b4c0 0x00000001 0x00000001
cfs.showChildFsRoot      0x0154b034 0x00000000 0x00000000
cfs.showHiddenCkpt       0x0154b038 0x00000001 0x00000001
cifs.DC.useFastest        0x01571458 0x00000001 0x00000001
cifs.LanmanServer.IdleUserAutoL 0x015780a4 0xfffffff 0xfffffff
cifs.LanmanServer.MaxMpxCount 0x0324f198 0x0000007f 0x0000007f
cifs.LanmanServer.SessUsers 0x0157392c 0x00000800 0x00000800
cifs.NTBufSz              0x01573934 0x0000ffff 0x0000ffff
cifs.NTsec.DCConnectTimeout 0x0324f9b8 0x00001388 0x00001388
cifs.NTsec.DCTimeout       0x0324f978 0x00004e20 0x00004e20
cifs.NTsec.getDCfromADServices 0x03249b58 0x00000001 0x00000001
cifs.ReadOnly.Comp         0x0232c8f8 0x00000000 0x00000000
cifs.ReadOnly.Delete       0x0232c938 0x00000000 0x00000000
cifs.W2KBufSz              0x01573930 0x0000ffff 0x0000ffff
cifs.W95BufSz              0x01573938 0x0000ffff 0x0000ffff
cifs.acl.FailOnSDRestoreError 0x01575834 0x00000001 0x00000001
cifs.acl.checkAclConsistency 0x0154b438 0x00000001 0x00000001
cifs.acl.checkUnixXForCifsOpen 0x0244d678 0x00000000 0x00000000
cifs.acl.checkacl           0x0154b42c 0x00000001 0x00000001
cifs.acl.extacl            0x0244d638 0x00000000 0x00000000
cifs.acl.extendExtraGid     0x0324f118 0x00000000 0x00000000
cifs.acl.mappingErrorAction 0x0244d6b8 0x00000000 0x00000000
cifs.acl.restrictedTakeOwnershi 0x0244d6f8 0x00000000 0x00000000
cifs.acl.retryAuthSid       0x032564b8 0x0000000a 0x0000000a
cifs.acl.sortAces           0x0154b430 0x00000000 0x00000000
cifs.acl.takegroupship     0x01575830 0x00000000 0x00000000
cifs.acl.useUnixGid         0x0244d538 0x00000000 0x00000000
cifs.admin.adminsAreRoot     0x01574b90 0x00000001 0x00000001
cifs.admin.shareC_RO         0x01571cc4 0x00000000 0x00000000
cifs.allowSnapSureVss        0x015769fc 0x00000001 0x00000001
cifs.audit.eventsCountFlush 0x0154b650 0x00000064 0x00000064
cifs.capabilities           0x0324f258 0x0000a3f9 0x0000a3f9
cifs.cifsclient.timeout      0x0157c438 0x00004e20 0x00004e20
cifs.comment_filtered        0x0157145c 0x00000000 0x00000000
cifs.djAddAdminToLg          0x0324a298 0x00000001 0x00000001
cifs.djEnforceDhn            0x0324a318 0x00000001 0x00000001
cifs.djUseKpasswd            0x0324a2d8 0x00000000 0x00000000
cifs.enableFileFiltering      0x032555c8 0x00000007 0x00000007
cifs.gpo                     0x0157c36c 0x00000001 0x00000001
cifs.gpocache                 0x0157c3f0 0x00000001 0x00000001
cifs.listBlocks                0x0325b5b8 0x00000040 0x00000040
cifs.lookup.secmapOnly        0x01576aa4 0x00000001 0x00000001

```

cifs.lookup.traceLevel 0x01576aa0 0x00000003 0x00000003  
cifs.lsarpc.maxDomains 0x03258eb8 0x00000020 0x00000020  
cifs.majorVersion 0x03249b9c 0x00000000 0x00000000  
cifs.maxForceDCInfoRefresh 0x01571454 0x0000000a 0x0000000a  
cifs.maxLockXPending 0x01575488 0x00000000 0x00000000  
cifs.maxMpxCount 0x0324f198 0x0000007f 0x0000007f  
cifs.maxVCThreads 0x0324a9f8 0x00000003 0x00000003  
cifs.minFreeFS 0x0324f0d8 0x0000000a 0x0000000a  
cifs.minorVersion 0x03249b98 0x00000000 0x00000000  
cifs.nanoroundoff 0x0325a3d8 0x00000000 0x00000000  
cifs.nullSession 0x0157393c 0x00000001 0x00000001  
cifs.nullSessionNotOnFS 0x01573940 0x00000001 0x00000001  
cifs.popupToConnectedUser 0x0324efa8 0x00000000 0x00000000  
cifs.prealloc 0x0325a290 0x00000004 0x00000004  
cifs.resolver 0x0157b500 0x00000000 0x00000000  
cifs.secmap.enable 0x0157c324 0x00000000 0x00000001  
cifs.sendMessage 0x03255678 0x00000001 0x00000001  
cifs.set\_eas\_ok 0x03255c78 0x00000000 0x00000000  
cifs.share.default.umask 0x01571dd0 0x00000012 0x00000012  
cifs.sidcache.globalSidCacheSiz 0x01576744 0x00000191 0x00000191  
cifs.sidcache.size 0x0157673c 0x00000035 0x00000035  
cifs.simulateNTFS 0x0325b9fc 0x00000001 0x00000001  
cifs.smbsigning 0x0157436c 0x00000001 0x00000001  
cifs.srvmgr.globalShares 0x03259538 0x00000000 0x00000000  
cifs.srvpwd.encryptAccountFile 0x01571740 0x00000000 0x00000000  
cifs.srvpwd.maxHistory 0x0157173c 0x00000002 0x00000002  
cifs.srvpwd.updtMinutes 0x01571464 0x000002d0 0x000002d0  
cifs.sync.stat 0x03249d8c 0x00001388 0x00001388  
cifs.tcpkeepalive 0x0324a9b8 0xff01030a 0xff01030a  
cifs.useADMap 0x0157b504 0x00000000 0x00000000  
cifs.vnodepercent 0x0325a418 0x00000050 0x00000050  
cifs.wins.UseDCs 0x03254fd8 0x00000001 0x00000001  
config.userMapper.version 0x0325b7b8 0x00000003 0x00000003  
cvfs.virtualDirName 0x01584cf0 'ckpt' 'ckpt'  
dic.httpPort 0x0157dc44 0x000013d9 0x000013d9  
dns.updateMode 0x0149ac00 0x00000002 0x00000002  
dns.updatePTRrecord 0x0149ac04 0x00000000 0x00000000  
ds.RefreshDC 0x03255258 0x00000384 0x00000384  
ds.RetryRefreshDC 0x032552d8 0x00000028 0x00000028  
ds.useADSite 0x03255358 0x00000001 0x00000001  
ds.useDCLdapPing 0x03255318 0x00000001 0x00000001  
ds.useDSFile 0x01fbe838 0x00000000 0x00000000  
fcTach.linx\_speed\_aux0 0x03196210 0x00008000 0x00008000  
fcTach.linx\_speed\_aux1 0x03196214 0x00008000 0x00008000  
file.asyncThreshold 0x0232ca38 0x00000020 0x00000020  
file.fsInodeThreshold 0x02509058 0x0000005a 0x0000005a  
file.fsSizeThreshold 0x0244f2d8 0x0000005a 0x0000005a  
filesystem.rstchown 0x0232c878 0x00000001 0x00000001  
ftpd.bounceAttackChk 0x031a3a38 0x00000001 0x00000001  
ftpd.defaultdir 0x031a3978 '/' '/'  
ftpd.maxCnx 0x031a3a78 0x0000ffff 0x0000ffff  
ftpd.wildcharsInDir 0x031a39f8 0x00000000 0x00000000  
http.quarantineTime 0x0157d7bc 0x0000003c 0x0000003c  
ifpcache.shared 0x02048b98 'true' 'true'  
ifpcache.size 0x0153ebb4 0x00000100 0x00000100  
ike.family 0x022e5230 'AF\_INET' 'AF\_INET'  
ip.reflect 0x02038078 0x00000001 0x00000001  
ipsec.avoidfrag 0x01545d70 0x00000001 0x00000001  
ipsec.sadsize 0x01545d78 0x00000400 0x00000400  
ipsec.spdsize 0x01545d74 0x00000080 0x00000080  
iscsi.AsyncEvent 0x0158b544 0x00000001 0x00000001  
iscsi.CollPerfStats 0x0158b540 0x00000000 0x00000000  
iscsi.ImmedData 0x0158b528 0x00000001 0x00000001  
iscsi.InitialR2t 0x0158b524 0x00000000 0x00000000  
iscsi.MaxConnections 0x0158b52c 0x00000004 0x00000004  
iscsi.PreferDataDigest 0x0158b518 0x00000000 0x00000000  
iscsi.PreferHeaderDigest 0x0158b514 0x00000001 0x00000001  
iscsi.RequireChap 0x0158b51c 0x00000000 0x00000000  
iscsi.RequireDiscoveryChap 0x0158b520 0x00000000 0x00000000  
iscsi.SendTargetsMode 0x0158b538 0x00000000 0x00000000  
iscsi.WinCompat 0x0158b530 0x00000001 0x00000001  
kernel.threads.alertOptions 0x014772b0 0x00000001 0x00000001  
kernel.threads.maxBlockedTime 0x014772a4 0x00000168 0x00000168  
kernel.threads.minimumAlertBloc 0x014772ac 0x00000000 0x00000000  
kernel.threads.panicIfHung 0x014772a8 0x00000000 0x00000000

|                                 |            |            |            |
|---------------------------------|------------|------------|------------|
| ldap.SecurityLayer              | 0x01543ef0 | 0x00000002 | 0x00000002 |
| lockd.OpLockTO1                 | 0x02458718 | 0x00001388 | 0x00001388 |
| lockd.OpLockTO2                 | 0x02458758 | 0x00001388 | 0x00001388 |
| lockd.OpLockTO3                 | 0x02458798 | 0x00000009 | 0x00000009 |
| lockd.asyncTO                   | 0x02457988 | 0x0000000a | 0x0000000a |
| lockd.gpDuration                | 0x02457ad8 | 0x0000002d | 0x0000002d |
| lockd.grantCS                   | 0x02457a98 | 0x00000400 | 0x00000400 |
| lockd.grantRetry                | 0x02457a18 | 0x00000000 | 0x00000000 |
| lockd.grantTO                   | 0x02457a58 | 0x00000fa0 | 0x00000fa0 |
| mgfs.listAllOfflineInodes       | 0x0157fee0 | 0x00000000 | 0x00000000 |
| mgfs.offlineCheckThreadsNum     | 0x0157fee4 | 0x00000001 | 0x00000001 |
| mount.allowNullCred             | 0x01548138 | 0x00000000 | 0x00000000 |
| nbs.sparseTws                   | 0x0158a03c | 0x00000000 | 0x00000000 |
| nfs.NTcred.TTL                  | 0x01576da4 | 0x00000014 | 0x00000014 |
| nfs.NTcred.size                 | 0x01576da0 | 0x000003f1 | 0x000003f1 |
| nfs.NTcred.trace                | 0x01576dac | 0x00000003 | 0x00000003 |
| nfs.NTcred.winDomain            | 0x03258a78 | " "        | " "        |
| nfs.rpcgss.discardReplay        | 0x01541de8 | 0x00000001 | 0x00000001 |
| nfs.secureExportMode            | 0x0154804c | 0x00000000 | 0x00000000 |
| nfs.v3xferSize                  | 0x0230a764 | 0x00008000 | 0x00008000 |
| nfsv4.32bitClient               | 0x01548b40 | 0x00000001 | 0x00000001 |
| nfsv4.delegLeaseDuration        | 0x0230ad34 | 0x000000b4 | 0x000000b4 |
| nfsv4.domain                    | 0x01549200 | " "        | " "        |
| nfsv4.leaseDuration             | 0x0230abdc | 0x00000028 | 0x00000028 |
| nfsv4.recallTimeout             | 0x0230ad2c | 0x0000000a | 0x0000000a |
| quota.countRootUsageInQuotaTree | 0x0284dcf8 | 0x00000001 | 0x00000001 |
| quota.maxuid                    | 0x0244f458 | 0x00000000 | 0x00000000 |
| quota.policy                    | 0x0244f498 | 'blocks'   | 'blocks'   |
| quota.useQuotasInFsStat         | 0x0284dd98 | 0x00000000 | 0x00000000 |
| replication.highcapacity        | 0x02f88578 | 0x00000001 | 0x00000001 |
| shadow.asciiifilter             | 0x02b53ab8 | 0x00000000 | 0x00000000 |
| shadow.flushCount               | 0x02b53678 | 0x00000400 | 0x00000400 |
| shadow.followabsolutpath        | 0x02b534f8 | 0x00000000 | 0x00000000 |
| shadow.followdotdot             | 0x02b53638 | 0x00000000 | 0x00000000 |
| shadow.stream                   | 0x02b53b38 | 0x00000001 | 0x00000001 |
| statd.hostname                  | 0x02458238 | ' / '      | ' / '      |
| stream.mblkInitial              | 0x01bd0ba4 | 0x00004000 | 0x00004000 |
| stream.mblkMax                  | 0x01bd0bf8 | 0x001fffff | 0x001fffff |
| streamio.timeout                | 0x01552fac | 0x00000708 | 0x00000708 |
| svtl.dbLocation                 | 0x032dd8c8 | 'svtl'     | 'svtl'     |
| svtl.discardTapeData            | 0x015871d8 | 0x00000000 | 0x00000000 |
| tcp.ackpush                     | 0x0218b868 | 0x00000000 | 0x00000000 |
| tcp.backlog                     | 0x0218ba50 | 0x00000064 | 0x00000064 |
| tcp.do_newreno                  | 0x0218b8f8 | 0x00000001 | 0x00000001 |
| tcp.do_sack                     | 0x0218b8b8 | 0x00000001 | 0x00000001 |
| tcp.fastRTO                     | 0x0218c108 | 0x00000000 | 0x00000000 |
| tcp.maxStreams                  | 0x0218bd78 | 0x0000ffff | 0x0000ffff |
| tcp.maxburst                    | 0x0218bab0 | 0x00000000 | 0x00000000 |
| tcp.sndcwnd                     | 0x022d6dbo | 0x00000000 | 0x00000000 |
| tftp.maxthreads                 | 0x031a42d8 | 0x00000020 | 0x00000020 |
| tftp.thrdtimeout                | 0x031a42e4 | 0x00000e10 | 0x00000e10 |
| trunk.LoadBalance               | 0x02032260 | 'ip'       | 'ip'       |
| ufs.dirOffsetHashSize           | 0x0154d554 | 0x000553ff | 0x000553ff |
| ufs.gid32                       | 0x0154e09c | 0x00000001 | 0x00000001 |
| ufs.indBlkHashSize              | 0x0284dc38 | 0x00000004 | 0x00000004 |
| ufs.inodelimit                  | 0x0154d8f0 | 0x0f600000 | 0x0f600000 |
| ufs.skipFsck                    | 0x0154d8ec | 0x00000000 | 0x00000000 |
| ufs.skipMapBlock                | 0x0154f4cc | 0x00000000 | 0x00000000 |
| ufs.syncwatchdog                | 0x0154d8f8 | 0x00000006 | 0x00000006 |
| ufs.verifySummary               | 0x0154d078 | 0x00000000 | 0x00000000 |
| ufs.xlateMaxThreads             | 0x0154e22c | 0x0000000a | 0x0000000a |
| ufs.xlateMinThreads             | 0x0154e230 | 0x0000000a | 0x0000000a |
| ufs.xlateSpaceHighCeiling       | 0x0154e248 | 0x00000064 | 0x00000064 |
| ufs.xlateSpaceLowCeiling        | 0x0154e24c | 0x000003e8 | 0x000003e8 |
| ufs.xlateToInline               | 0x0154e240 | 0x00000001 | 0x00000001 |
| usrmap.maxgid                   | 0x0157eb38 | 0x7fffffff | 0x7fffffff |
| usrmap.maxuid                   | 0x0157eb2c | 0x7fffffff | 0x7fffffff |
| usrmap.mingid                   | 0x0157eb34 | 0x00000010 | 0x00000010 |
| usrmap.minuid                   | 0x0157eb28 | 0x00000010 | 0x00000010 |
| vbb.tempDir                     | 0x015863c0 | 'vbtemp'   | 'vbtemp'   |
| viruschk.Notify                 | 0x0244e558 | 0x00000007 | 0x00000007 |
| viruschk.Traces                 | 0x0244e518 | 0x00000000 | 0x00000000 |
| viruschk.chunkQuota             | 0x0325c198 | 0x00000041 | 0x00000041 |
| viruschk.noRetry                | 0x0325bf98 | 0x00000018 | 0x00000018 |
| viruschk.reconnectTime          | 0x0325c0b8 | 0x0000012c | 0x0000012c |

```
viruschk.vnodeHWM      0x0325bfdc 0x0000005a 0x0000005a  
viruschk.vnodeLWM      0x0325bfd8 0x0000003c 0x0000003c  
viruschk.vnodeMax      0x0325bfe0 0x000007d0 0x000007d0
```

#### **CONTROL STATION HISTORY LOG:**

\$more /home/nasadmin/.sh\_history [command history log from root shell]

#[/home/nasadmin/.bash\\_history](#) → Current Linux command history file for root user #history **#history -c** [Clear History log]

## **CELERRA TROUBLESHOOTING LOGS/CELERRA LOGS:**

### **I. CONTROL STATION LOGS:**

Most of the important Control Station logging for Celerra activities is located in /nas/log directory, though Celerra Manager, and its associated components, are logged in various directories. Additionally, some important tasks, such as NAS Upgrades, or SRDF failover activities, are logged in /tmp directory until the operation is completed. Linux-related logging is located in the traditional /nas/log directory. The Control Station runs a nas\_eventcollector daemon that retrieves events from a memory buffer and uses ReportEvent and PostEvent to post to the nas\_eventlog, which is turn gets recorded to the sys\_log.

#### **/nas/log directory:**

install.log: NAS install log **upgrade.5.3.15-3.Apr-15-17:54.log:** NAS Upgrade log

**nas\_log.al** Successful commands **nas\_log.al.err** Cmds that err **nas\_log.al.trace** (Eiffel dump location)

**sys\_log:** Control Station log of Hardware Events—this is what Call\_Home utility uses **sympapi.log** (sympapi events)

**osmlog:** /var/adm/log/osmlog **setup\_clariion2.log** (clariion setup log) **nas\_rdf.log** (rdf issues) **cmd\_log cmd\_log.err**

**nas\_log.al.rll**

#### **Callhome Logs:**

\$ /nas/log/callhome.phone\_calls [Prior to NAS 5.5.31--Callhome information, times & numbers of calls made]

\$ /nas/opt/connectemc [NAS 5.5.31 and higher]

#### **/var/log directory:**

messages file

dmesg file

boot.log file

### **II. CELERRA MANAGER/JSERVER LOGS:**

→ see Celerra Manager section for complete list of logs for APL Task Mgr, Tomcat, Apache, JServer, etc.

JServer events are posted in the sys\_log and the alert\_log, not the Server Log

### **III. SERVER LOGS—DART:**

DART uses a logmsg facility to post events into the Server Log, and a logEvents facility to post events into a memory buffer that gets checked by the NAS Event Collector daemon via MAC/XML interface (?)

## **TROUBLESHOOTING USING THE SERVER LOG:**

### **REAL SERVER LOG vs. HISTORICAL LOG:**

**RealTime Server log:** *format* is always \$server\_log server\_2 and shows information since last reboot, or in case of a failover or panic, information from point of failure to present time on the Standby Server if running as the Primary Server

**Historical Server log:** *format* is always \$server\_log server\_2.faulted.server\_5 -a -s and shows all historical information up to the point of panic or failover.

## **USING SERVER LOGS TO TROUBLESHOOT PANICS OR FAILOVERS:**

**Scenario:** Primary Server\_2 panics & fails over to Standby Server\_5

- Step 1. Locate Hostname in **historical server\_log**: \$server\_log server\_2.faulted.server\_5 | grep hostname  
[Locate section where server\_5 renamed itself server\_2]
- Step 2. Locate Panic Info in sys\_log: \$grep panic /nas/log/sys\_log  
[Obtain TimeStamp{CS0} for Panic in question; 2001-04-26 07:30:05]
- Step 3. Note diff. in CS0 & Server Dates: \$date \$server\_date ALL
- Step 4. **For Historical Events for Server 2:** \$server\_log server\_2.faulted.server\_5 -a -s |more /2001-04-26 07:30:05
- Step 5. **For RealTime Events for Server 2:** \$server\_log server\_2  
[Observe that Standby Server\_5 is running as Server\_2 in Slot\_5]

#### **Using Logs to Focus in on Data Mover Problem:**

Step 1. Use Control Station “Sys\_log” to ID problems and obtain timestamps of events, etc

Step 2. Go to appropriate Server\_log \$server\_log server\_2 -a -s | more

Step 3. Run this Server Log Command to view Log “live”: \$server\_log server\_2 -s -f

**Note:** DataMover reboots are logged in the “sys\_log”

**TAILING SERVER LOG TO OBSERVE ‘REALTIME’:** [Contains Log info since last server reboot]**\$server\_log server\_2 -s -f**

-f switch keeps log open -s switch applies date/timestamp

**Using Server Log Command to Open the Archival Logs:**

\$server\_log server\_2 -a -s

**CREATING SCRIPT TO SAVE SERVER LOG CONTENTS USING CRON JOB:**

1. Create following script on Control Station and chmod 777—make sure you create folder in which the logs will be collected, and insert the folder name in the script—in this example, the script will create log files in /nas/var/slog:

```
#!/bin/sh
DMname=server_2
NAS_DB=/nas
export NAS_DB
$NAS_DB/bin/server_log $DMname -s -a > /nas/var/slog/log.$DMname.`date '+%y%m%d%H%M'`
```

2. Create Cron Job to run every 5 minutes, or set the interval as desired

**0-59/5 \* \* \* \* /nas/var/slog/logscript.log &****Note:** Cronjob example creates a new server\_log with timestamp every 5 minutes**CFS ENTRIES IN SERVER LOG:**

CFS=Common File System, the software layer that resides between the NFS &amp; CIFS protocol and the filesystem.

**OPENING SERVER LOG AND OUTPUTTING CONTENTS INTO FILE:**

**Purpose:** Since the total size of Server Log is limited to 1 MB it is useful to pipe the Server Log into an open file so as to capture debugging information:

**\$server\_log server\_3 -f -s > s3log.log & [\$jobs #fg %1 ctrl + c]****SERVER LOG SIZE INCREASED TO 2MB WITH NAS 5.4:****Note:** NAS 5.4.5.0 increased Server Log to 2MB from 1MB—see AR47260--root\_rdf\_channel volume.**NAS 5.3 WORKPART MAGIC & SERVER LOG SLICE SIZE:****# /nas/sbin/workpart -r**

Read Work Partition layout from LBA 0x43000, 528384 bytes.

After read:

Dump of Work Partition Structure (size 528384 bytes):

workpart\_magic = 0xfacebeaf      --&gt;0xfacebeaf = pre-5.4 workpart\_magic

-----  
log\_slices[0].size = 0x7ff      -->0x7ff = 2047 decimal \* 512 = 1048064 (1MB)**# /nas/sbin/rootnas\_volume -i root\_rdf\_channel**

id = 41

name = root\_rdf\_channel

acl = 0

in\_use = False

type = slice

slice\_name = root\_log\_1

slice\_of = root\_disk

offset(MB) = 391

**size (MB) = 1**

disks = root\_disk

**NAS 5.4/5.5 WORKPART MAGIC & SERVER LOG SLICE SIZE:****# /nas/sbin/workpart -r |head**

Read Work Partition layout from LBA 0x43000, 528384 bytes.

After read:

Dump of Work Partition Structure (size 528384 bytes):

workpart\_magic = 0xbeaf0001      --&gt;0xbeaf0001 = 5.4/5.5 workpart\_magic

-----  
log\_slices[0].size = 0xffff      -->0xffff = 4095 \* 512 = 2096640 (2MB) Server Log**-rw-r--r-- 1 root root 2328452 Dec 28 15:22 slog2.2mb**

# /nas/sbin/rootnas\_volume -i root\_rdf\_channel

```
id      = 165
name    = root_rdf_channel
acl     = 0
in_use   = False
type    = slice
slice_name = root_rdf_channel
slice_of  = root_disk
offset(MB) = 391
```

**size (MB) = 2**

```
disks   = root_disk
```

# /nas/sbin/rootnas\_volume -s root\_rdf\_channel

```
total = 2 avail = 2 used = 0 ( 0% ) (sizes in MB)
```

## NAS 5.6 WORKPART MAGIC & SERVER LOG SLICE SIZE:

# /nas/sbin/workpart -r | head

Read Work Partition layout from LBA 0x43000, 528384 bytes.

After read:

Dump of Work Partition Structure (size 528384 bytes):

```
workpart_magic      = 0xbeaf0002
log_slices[0].size  = 0x1fff → 0x1fff = 8191 * 512 = 4193792 (4MB) Server Log
```

## OUTPUTTING RAW SERVER LOG WITH LOG SLOT:

\$ /nas/sbin/log\_slot -r slot\_2.logfile 2

Creating raw log dump file "slot\_2.logfile".

Dumping slot 2 raw log to "slot\_2.logfile".

## Looking at Correct Log for the Primary and Faulted-Over DataMover:

1. For events prior to the Panic on Primary Server\_2:

\$server\_log server\_2 -s -f

2. For events after failover, while faulted over to the Standby Server\_4:

\$server\_log server\_2.faulted.server\_4

## SETTING SMB DEBUG LOGGING ON DM:

1. Turning on Debug Logging: #.server\_config server\_x "logsys set severity SMB=LOG\_DEBUG"

2. Turn off Debug Logging: #.server\_config server\_x "logsys set severity SMB=LOG\_PRINTF"

## VERIFYING THE CURRENT LEVEL OF DEBUG LOGGING FOR CATEGORIES:

# .server\_config server\_2 -v "logsys get severity SMB"

1158591814: LIB: 4: Server log severity for facility SMB is 4

## SERVER LOG CATEGORY PRECEDENCE NUMBERS/CLASSIFICATION CODES:

### Existing DART Logging Severity Levels:

0 = LOG\_EMERG or LOG\_EVENT

0=Emergency, system unusable

1 = LOG\_ALERT

1=Alert, take action

2 = LOG\_CRIT

2=Critical, critical condition

3 = LOG\_ERR

3=Error, error condition (old default logging level)

4 = LOG\_PRINTF or LOG\_WARNING or LOG\_STAT

4=Warnings (this appears to be the new default logging level)

5 = LOG\_NOTICE or LOG\_DEBUG or LOG\_DBG1

5=Notice, normal, but significant messages

6 = LOG\_DBG2 or LOG\_INFO

6=Info, informational messages

7 = LOG\_DBG3

7=Highest debug level messages

### Proposed CCMD Logging Levels:

0 = Emergency → LOG\_EMERG=0

0=Emergency, system unusable

1 = Alert → LOG\_ALERT=1

1=Alert, take action

2 = Critical → LOG\_CRIT=2

2=Critical, critical condition

3 = Error → LOG\_ERR=3

3=Error, error condition (old default logging level)

4 = Warning → LOG\_WARNING=4

4=Warnings (this appears to be the new default logging level)

5 = Notify → LOG\_NOTICE=5

5=Notice, normal, but significant messages

6 = Info → LOG\_INFO=6 [or LOG\_PRINTF or LOG\_STAT]

6=Info, informational messages

7 = Debug → LOG\_DEBUG=7

7=Highest debug level messages

## BACK-END STATISTICS:

`$server_config server_2 -v "printstats scsi full" or "printstats scsi full reset" [resets stats buffer to zero]`

### **PRINTSTAT SCSI OUTPUT DEFINED:**

**`$ .server_config server_2 -v "printstats scsi"`**

io count 82344, time 221553637 usecs, ave 2690

tot time on nexus queue 1741478392 usecs, ave 21148

ctr0: maxqueued 360, maxpend 33, byteCount 603410320, busy 29286206695, idle 4611494911468 = 99%

io count 82344 -->Number I/O's since last reboot

time 221553637 usecs -->Time used to perform 'io count' totals

ave 2690 -->Average time to complete one i/o

maxqueued -->Max number of i/o's queued--queue depth is 8/lun (Clariion) and 1/hyper (Symmetrix)--number indicates chain

maxpend -->Number of pending I/O's to luns on indicated chain--pending, not queued I/O

byteCount -->Ticker value for bytes processed

busy -->Ticker value for the time of actual work down chain

Idle -->Idle value is usually near 99% but can be lower as I/O increases and decreases--represents amount of time chain fulfills request.

**Assessing Celerra-to-Symmetrix Performance:** Writes to the Symm from the Celerra

**`$server_config server_2 -v "printstats scsi"`**

**`$server_config server_2 -v "printstats filewrite"`**

### **CELERRA SERIAL NUMBER:**

**`# /nas/sbin/serial`**

APM00024300875

**`# cat /nas/sys/callhome.config`**

modem phone number:9999999

site name:35021551

serial number:APM00024300875

**`# expect -f /nas/sbin/encl_mon.exp resume server_2 5080 |grep RESUME_INFORMATION_MIDPLANE`**

IDS exp7 1 RESUME\_INFORMATION\_MIDPLANE

<RESUME\_INFORMATION\_MIDPLANE EMC\_PART\_NUMBER="005047766" EMC\_ARTWORK\_REVISION=""  
EMC\_ASSEMBLY\_REVISION="A13" EMC\_SERIAL\_NUMBER="APM00024300875" VENDOR\_NAME="" L  
OCATION\_OF\_MANUFACTURE="Apex, NC USA" YEAR\_OF\_MANUFACTURE="2002" MONTH\_OF\_MANUFA  
CTURE="10" DAY\_OF\_MONTH\_OF\_MANUFACTURE="28" ASSEMBLY\_NAME="XPE NAS CHASSIS" WORL  
D\_WIDE\_NAME\_SEED="006006aa" />

**`# nas_xml -info:server |grep SERIAL_NO`**

PRODUCT\_NAME=Celerra NS600G' SERIAL\_NO='APM00024300875'

**`$ /nas/sbin/get_data_mover_status -resume server_2 5080 /tmp/dm2_resume.xml`**

### **HOW /nas/sbin/serial COMMAND DERIVES CELERRA SERIAL NUMBER:**

#### **I. CELERRA IP SYSTEMS (e.g., NS500, 700, etc.):**

**1. /nas/sbin/serial command determines "Celerra IP" system**

a) Gets Data Mover Resume information from Server\_2

→From serial script: \$NAS\_DB/sbin/get\_data\_mover\_status -resume server\_2 5082 \$dm\_file

b) Sed & Greps for EMC\_SERIAL\_NUMBER from the RESUME\_INFORMATION\_MIPLANE xml field

**`# cat /nas/log/data_mover_resume.server_2.xml`**

<RESUME\_INFORMATION\_MIDPLANE EMC\_PART\_NUMBER="005048417" EMC\_ARTWORK\_REVISION=""  
EMC\_ASSEMBLY\_REVISION="A03" EMC\_SERIAL\_NUMBER="APM00042403637"

c) Outputs Celerra Serial number

**`# /nas/sbin/serial`**

APM00042403637

#### **II. HAMMERHEAD/SLEDGEHAMMER SYSTEMS (NS20, NS40, NX4, NSX, NS-120, etc.):**

**1. /nas/sbin/serial command determines "Celerra Hammerhead" | "Celerra SledgeHammer" system**

a) Gets Enclosure Resume information from Enclosure\_0

→From serial script: \$NAS\_DB/sbin/get\_enclosure\_status -resume 0 \$encl\_file

b) Sed & Greps for EMC\_SERIAL\_NUMBER from the RESUME\_INFORMATION\_MIPLANE xml field

**`# cat enclosure_resume.enclosure_0.xml`**

<?xml version="1.0"?>

<EnclosureResume EnclosureID="0">

```
<RESUME_INFORMATION_MIDPLANE
  EMC_PART_NUMBER="100-520-327    "
  EMC_ARTWORK_REVISION="   "
  EMC_ASSEMBLY_REVISION="A05"
  EMC_SERIAL_NUMBER="APM00071601776 "
```

c) Outputs Celerra Serial number

```
# /nas/sbin/serial
```

```
APM00071601776
```

### III. FOXGLOVE SYSTEMS (NS-960/NS-G8):

**Note:** Beginning with Foxglove, the EMC\_PRODUCT\_SERIAL\_NUMBER becomes the system serial number, as found in the enclosure\_resume\_enclosure\_0.xml 'RESUME\_INFORMATION\_MIDPLANE' field, and also on the PSNT tag hanging on the back of DME0. This is the serial number used to install the system in CSI and is the serial number used by the Celerra when it calls home. The serial and nas\_connecthome commands should list this serial number.

#### 1. /nas/sbin/serial command determines "Celerra Foxglove" system

→From serial script: serial=`get\_db\_psn`

a) Gets PSN (Product Serial Number) from /etc/product\_numbers.db file

```
# cat product_numbers.db
```

#Product Numbers Database file

#This file contains the Product Serial Number(PSN) and Product Part Number(PPN).

Version=1.0

**PSN=FNM00083800203**

PPN=900-525-001

b) If the PSN is not available from the product\_numbers.db file, defaults to using Enclosure\_0's default serial number (NOT the PSN number):

→From serial script:

```
$NAS_DB/sbin/get_enclosure_status -resume 0 $encl_file
```

```
serial=`sed -n '/<RESUME_INFORMATION_MIDPLANE/,/\/>/p' $encl_file | grep EMC_SERIAL_NUMBER
```

**Note:** It's important to repeat that the EMC\_SERIAL\_NUMBER in the RESUME\_INFORMATION\_MIDPLANE represents the Serial number of the actual Enclosure\_0 midplane, and NOT the Product Serial Number that is found on the PSNT tag and inputted by the factory into the Enclosure\_0 RESUME PROM

#### # cat enclosure\_resume.enclosure\_0.xml

```
<?xml version="1.0"?>
```

```
<EnclosureResume EnclosureID="0">
```

```
<RESUME_INFORMATION_MIDPLANE
```

```
  EMC_PART_NUMBER="100-520-839    "
```

```
  EMC_ARTWORK_REVISION="   "
```

```
  EMC_ASSEMBLY_REVISION="A01"
```

**EMC\_SERIAL\_NUMBER="FNM00083800197 "** →Enclosure serial number, not the PSN

```
  EMC_PRODUCT_PART_NUMBER="900-525-001    "
```

**EMC\_PRODUCT\_SERIAL\_NUMBER="FNM00083800203 "** →Real PSN number

c) Outputs Celerra PSN Serial number (Same as written on the PSNT Tag)

```
# /nas/sbin/serial
```

FNM00083800203 →Actual PSN number as derived from PSNT tag and factory installation

d) Output from Serial command when product\_numbers.db file is missing or PSN field blank

```
# /nas/sbin/serial
```

FNM00083800197 →Enclosure serial number

**Note:** It's worth commenting that with the release of the Foxglove system, the traditional "EMC\_SERIAL\_NUMBER", as read in the Data Mover or Enclosure\_0 RESUME information, is no longer the system Celerra Serial number, and instead represents the physical serial number used for enclosure\_0. Instead, the EMC\_PRODUCT\_SERIAL\_NUMBER (PSN) is now considered the top-level system serial number and is the "Celerra Serial number".

#### CELERRA SERIAL NUMBER & CONNECTHOME ISSUE:

**Note:** See emc211169 for an issue where the system serial number reported has changed to display the serial number for Data Mover enclosure 0, both for the /nas/sbin/serial output and in the connecthome.config file. This occurs only for older 'fish' NS Series (NS500, NS350, NS600, NS700) that upgrade from NAS 5.5 or earlier 5.6 codes to NAS 5.6.42.5 or above. See AR140547. This issue was fixed in NAS 5.6.45.

The Celerra Serial number is referenced in the /nas/sys/connecthome.config file on all systems using the ConnectHome feature:

```
# cat /nas/sys/connecthome.config |grep Serial
```

```
celerraSerial:FNM00083800203
```

```
# /nas/sbin/nas_connecthome -modify -serial_number
```

Serial number automatically detected

Ok

**Note:** The system will automatically detect the Celerra Serial number from the product\_numbers.db or enclosure resume file and write the serial number to the connecthome.config file

# /nas/sbin/nas\_connecthome -info|head

ConnectHome Configuration:

Encryption Enabled = no

Dial In :

Enabled = yes

Modem phone number =

Site ID = 117900

Serial number = FNM00083800203

### **REPAIRING SERIAL NUMBER ON FOXGLOVE SYSTEMS:**

# /nas/sbin/serial -get

Error 255: serial

```
-get { -psn | -ppn } [ <enclosure_id> ]
| -db_check
| -repair
| -c <Serial Number>
```

/var/log/serial.log

### **REPAIRING PPN ON GALILEO SYSTEMS:**

#### **Scenario:**

During the install, the PPN for VG2 was inadvertently used for the VG8 hardware. Although the model correctly reported itself, the PSN number was wrong.

# /nas/sbin/serial -db\_check

Error 28: PPN is inconsistent between the db file and the enclosures

#### **Corrective Actions:**

- 1) First step is to connect to the serial port of either Blade 2 or Blade 3 in Enclosure 0
- 2) Boot the blade, access the POST program [passwd=SHIP\_it], select 2) Set Resume, then edit the following field with the correct PPN number for a VG8:

Enter Product PN (16 characters)[900-567-004]: 900-567-005

3) Reboot blade

4) Update the Enclosure Resume status file:

# /nas/sbin/get\_enclosure\_status -resume 0 /nas/log/enclosure\_resume.enclosure\_0.xml >/dev/null 2>&1

5) Vi edit the /etc/product\_numbers.db file and enter the correct PPN

6) Reboot CS and verify that the product\_numbers.db retains the edited values

### **CELLERRA CALLHOME (Pre-NAS 5.5.31):** Modem Number & Site info in "callhome.config" file

#### **File Path= /nas/sys/callhome.config:**

- |  |  |
|--|--|
| Step 1. Verify CallHome Configuration: | #more /nas/sys/callhome.config   |
| Step 2. Stop CallHome Service:         | #/nas/sbin/ch_stop [Stops any current CallHomes & clears callhome buffers]   |
| Step 3. Modify CallHome.Config:        | #/nas/sbin/ch_cfg  |
| Step 4. Modify Fieldwatch Numbers:     | #/nas/sbin/ch_dd [/nas/sys/callhome.dial_dir]  |
| Step 5. Test CallHome:                 | #/nas/sbin/ch_test [Tests callhome]  |
| Step 6. Running Modem Test:            | #/nas/sbin/modemtst  |
| Step 7. Enabling or Disabling Dial-In: | #/nas/sbin/ch_pm [Must run this after running "modemtst", which disables Dial-In!] >on [for Enabling Dial-In for "ttyS0"] >off [for Disabling Dial-In for port monitor "ttyS0"]] |

### **XML CALLHOME PAYLOAD DIRECTORY NAS 5.2:**

/nas/log/ConnectHome

# ls -la

```
-rw-r--r-- 1 root root 2468 Jun 22 09:26 RSC_ML2805001493_062204_092650000.xml
```

### **CALLHOME PROCESS FOR CLARIION BACKEND EVENTS:** see primus emc125099

The naviagent is running on the control station to monitor backend events. For any informational/warning/error/critical events generated from the Backend, the naviagent will call syslog() function with a specific event message [Note that this new naviagent has changed message format to syslog() function]. Once an event is generated, the syslog() function will direct the message to the Control Station's syslogd daemon, which will then simultaneously redirect the messages to /var/log/messages & /nasmcd/event\_fifo, based on the contents of the configuration file located in /etc/syslog.conf. Another Control Station process, called /nas/sys/navilog\_mon, monitors for any backend events posted to /nasmcd/event\_fifo, and upon finding events, will then post the

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
message to /nas/log/sys\_log. The problem here is that /nas/sys/navilog\_mon could read the redirected message from  
/nasmcd/event\_fifo but could not parse it correctly because the new naviagent has added four additional fields in the event message.

#### **ORIGINAL NAVIAGENT FORMAT (6.16.2.0.0):**

Jan 2 14:32:48 ns600-1 EV\_AGENT[5161]: Time Stamp 01/03/06 00:32:09 Event Number a07 Severity Critical Error Host ns600spa Storage Array APM00023700164 SPA Device Bus 1 Enclosure 0 Disk 2 Description CRU Powered Down

#### **NEW NAVIAGENT FORMAT (6.19.0.4.14):**

Dec 21 03:26:15 cs\_kama EV\_AGENT[5073]: Time Stamp 12/21/05 03:25:06 Event Number **0xa07 ExtCode1 0xa07 ExtCode2 0xa07** Severity Critical Error Host OEM-2MYA58ZGAQT Storage Array CK200054201137 SPA Device Enclosure 0 Disk 9 SoftwareRev 6.19.0 (4.14) Unknown Error 2.19.0.500.5.016 Description CRU Powered Down

**Note:** Issue fixed with NAS 5.3.24.1 & 5.4.21.2

**# /nas/sbin/naviagent -help**

@(#)Navisphere Agent Revision 6.19.0.4.14

### **TEMPORARILY DISABLING CALLHOME EVENTS WHILE CONDUCTING MAINTENANCE:**

1) Create file to prevent ch\_monitor process from performing CallHome:

**# touch /nas/tmp/CallHomeTESTING**

**Note:** Above applies to NAS 5.3.19.0, 5.4.17.0, 5.5.3.0 only. Use # touch /nas/log/ConnectHome/TESTING for other versions.

2) Perform the maintenance activity: maintenance, upgrade, hardware replace, etc.

3) Clear CallHome queue and remove the CallHomeTESTING file:

**# /nas/sbin/ch\_stop**

**Note:** For NS systems, go to /nbsnas/log/ConnectHome directory & manually delete XML files

4) Verify that the file (CallHomeTESTING or TESTING) was deleted—remove manually if required

### **CALLHOME EVENTS NAS 5.4:**

CPU Blade faults, Management Switch faults, Power Supply, Fan faults, slot boot errors, Voltage, Fan, Temp alarms, APC UPS faults

#### **CALLHOME CONFIG FILES:**

**Configuring:** #/nas/sbin/ch\_cfg

**Stop Service:** #/nas/sbin/ch\_stop

**Start Editor:** #/nas/sbin/ch\_dd [R to revise the table. Modify Fieldwatch numbers]

**Conduct Test:** #/nas/sbin/ch\_test {if successful, then test from Remote Dial-In}

**Useful Log:** Call Home event.dir

### **CALL-IN UTILITY:**

**Default Settings:** No parity; 8 data bits; 1 stop bit; speed 9600bps Default Terminal is ttyS0 enabled, ttyS1 disabled.

**Note:** To set correct emcnas command set path; #set path =\$path: /nas/bin: /nas/sbin export PATH

#### **DISABLING CALL-IN UTILITY:**

**#/nas/sbin/ch\_dialin\_disable ttyS0 | ttyS1** ['enable' to reactivate]

**Modem Daemon:**

#top esc + M → Look for modem process: /sbin/mgetty -x 0

### **CALL HOME EVENTS FOR NAS 5.5:**

**# nas\_event -list -a callhome**

| facility      | id | description                                |
|---------------|----|--|
| CAM           | 7  | CAM I/O error                              |
| CFS           | 4  | Crossed the root filesystem size threshold |
| DRIVERS       | 6  | NIC device panic                           |
| FSTOOLS       | 0  | FSCK Started                               |
| MasterControl | 6  | Unexpected daemon exit                     |
| MasterControl | 7  | Control Station heartbeat missing          |
| MasterControl | 8  | Sibpost failed                             |
| MasterControl | 9  | Unknown slot ID                            |
| MasterControl | 10 | File system error                          |
| MasterControl | 12 | NBS Device Error                           |
| MasterControl | 13 | IPMI connection failure                    |
| BoxMonitor    | 1  | EPP did not initialize                     |
| BoxMonitor    | 3  | Mandatory thread create failed             |
| BoxMonitor    | 4  | SIB read failure                           |
| BoxMonitor    | 5  | SIB write failure                          |
| BoxMonitor    | 6  | AC power failure                           |
| BoxMonitor    | 7  | Fan 0 has failed                           |
| BoxMonitor    | 8  | Fan 1 has failed                           |

|            |     |  |
|------------|-----|--|
| BoxMonitor | 9   | Both fans have failed                          |
| BoxMonitor | 10  | High temperature warning                       |
| BoxMonitor | 11  | High temperature failure                       |
| BoxMonitor | 12  | 12V DC supply high                             |
| BoxMonitor | 13  | 12V DC supply low                              |
| BoxMonitor | 14  | 5V DC supply high                              |
| BoxMonitor | 15  | 5V DC supply low                               |
| BoxMonitor | 18  | -12V DC supply high                            |
| BoxMonitor | 19  | -12V DC supply low                             |
| BoxMonitor | 20  | -5V DC supply high                             |
| BoxMonitor | 21  | -5V DC supply low                              |
| BoxMonitor | 38  | Invalid hardware configuration                 |
| BoxMonitor | 39  | ShoeBox boot failed                            |
| BoxMonitor | 43  | Shoebox ping failed to respond                 |
| BoxMonitor | 57  | Battery power is low                           |
| BoxMonitor | 58  | Lost AC Power                                  |
| BoxMonitor | 62  | EPO failure alarms                             |
| BoxMonitor | 63  | Battery related alarms                         |
| BoxMonitor | 64  | AC input power alarms                          |
| BoxMonitor | 65  | DC internal power alarms                       |
| BoxMonitor | 66  | Power supply failure alarms                    |
| BoxMonitor | 67  | Fan failure alarms                             |
| BoxMonitor | 70  | /etc/hosts file error                          |
| BoxMonitor | 72  | Shoebox sibpost failure                        |
| BoxMonitor | 73  | Serial connection to the Data Mover has failed |
| BoxMonitor | 100 | Slot 00 reason code is stale                   |
| BoxMonitor | 101 | Slot 01 reason code is stale                   |
| BoxMonitor | 102 | Slot 02 reason code is stale                   |
| BoxMonitor | 103 | Slot 03 reason code is stale                   |
| BoxMonitor | 104 | Slot 04 reason code is stale                   |
| BoxMonitor | 105 | Slot 05 reason code is stale                   |
| BoxMonitor | 106 | Slot 06 reason code is stale                   |
| BoxMonitor | 107 | Slot 07 reason code is stale                   |
| BoxMonitor | 108 | Slot 08 reason code is stale                   |
| BoxMonitor | 109 | Slot 09 reason code is stale                   |
| BoxMonitor | 110 | Slot 10 reason code is stale                   |
| BoxMonitor | 111 | Slot 11 reason code is stale                   |
| BoxMonitor | 112 | Slot 12 reason code is stale                   |
| BoxMonitor | 113 | Slot 13 reason code is stale                   |
| BoxMonitor | 114 | Slot 14 reason code is stale                   |
| BoxMonitor | 115 | Slot 15 reason code is stale                   |
| BoxMonitor | 200 | Slot 00 ping failed on both ifaces             |
| BoxMonitor | 201 | Slot 01 ping failed on both ifaces             |
| BoxMonitor | 202 | Slot 02 ping failed on both ifaces             |
| BoxMonitor | 203 | Slot 03 ping failed on both ifaces             |
| BoxMonitor | 204 | Slot 04 ping failed on both ifaces             |
| BoxMonitor | 205 | Slot 05 ping failed on both ifaces             |
| BoxMonitor | 206 | Slot 06 ping failed on both ifaces             |
| BoxMonitor | 207 | Slot 07 ping failed on both ifaces             |
| BoxMonitor | 208 | Slot 08 ping failed on both ifaces             |
| BoxMonitor | 209 | Slot 09 ping failed on both ifaces             |
| BoxMonitor | 210 | Slot 10 ping failed on both ifaces             |
| BoxMonitor | 211 | Slot 11 ping failed on both ifaces             |
| BoxMonitor | 212 | Slot 12 ping failed on both ifaces             |
| BoxMonitor | 213 | Slot 13 ping failed on both ifaces             |
| BoxMonitor | 214 | Slot 14 ping failed on both ifaces             |
| BoxMonitor | 215 | Slot 15 ping failed on both ifaces             |
| BoxMonitor | 300 | Slot 00 panicked                               |
| BoxMonitor | 301 | Slot 01 panicked                               |
| BoxMonitor | 302 | Slot 02 panicked                               |
| BoxMonitor | 303 | Slot 03 panicked                               |
| BoxMonitor | 304 | Slot 04 panicked                               |
| BoxMonitor | 305 | Slot 05 panicked                               |
| BoxMonitor | 306 | Slot 06 panicked                               |
| BoxMonitor | 307 | Slot 07 panicked                               |
| BoxMonitor | 308 | Slot 08 panicked                               |
| BoxMonitor | 309 | Slot 09 panicked                               |
| BoxMonitor | 310 | Slot 10 panicked                               |
| BoxMonitor | 311 | Slot 11 panicked                               |
| BoxMonitor | 312 | Slot 12 panicked                               |
| BoxMonitor | 313 | Slot 13 panicked                               |
| BoxMonitor | 314 | Slot 14 panicked                               |
| BoxMonitor | 315 | Slot 15 panicked                               |
| BoxMonitor | 500 | CPU blade is removed                           |
| BoxMonitor | 502 | CPU blade fault                                |

BoxMonitor 504 CPU IO Module 0 fault  
 BoxMonitor 506 CPU IO Module 1 fault  
 BoxMonitor 508 I/O Annex is removed  
 BoxMonitor 510 I/O Annex fault  
 BoxMonitor 512 Management Switch is removed  
 BoxMonitor 514 Management Switch fault  
 BoxMonitor 515 Management Switch internal error  
 BoxMonitor 517 Power Supply is removed  
 BoxMonitor 519 Power Supply fault  
 BoxMonitor 521 Fan fault  
 BoxMonitor 523 Multiple Fan error  
 BoxMonitor 524 Slot boot error  
 BoxMonitor 527 Both Management Switches Ping Failure  
 JServer 101 Call Home - JServer failed to come up after 3 retries  
 JServer 102 Call Home - Generated debug file xxx from the core dump  
 JServer 103 Call Home - JServer reports nas\_cmd errors  
 JServer 104 Call Home - JServer generated problem report zip file /nas/jserver/x.zip  
 JServer 105 Call Home - API Server failed to come up after 3 retries  
 VRPL 1 Replication on Source Filesystem Inactive  
 VRPL 2 Resync asked by (previous) Destination Filesystem  
 VRPL 4 Destination Filesystem in Error  
 NaviEventMonitor 4 Navi Event with severity CRITICAL was received  
 NaviEventMonitor 100 A control lun has been cache compromised  
 LocalHardwareMonitor 100 VCCP Voltage Alarm  
 LocalHardwareMonitor 101 FSB Voltage Alarm  
 LocalHardwareMonitor 102 MCH CORE 1.5V Alarm  
 LocalHardwareMonitor 103 MEM VTT 0.98V Alarm  
 LocalHardwareMonitor 104 MEM CORE 1.8V Alarm  
 LocalHardwareMonitor 105 NIC CORE 1.5V Alarm  
 LocalHardwareMonitor 106 BB 3.3VSB Alarm  
 LocalHardwareMonitor 107 BB 5VSB Alarm  
 LocalHardwareMonitor 108 BB 3.3V Alarm  
 LocalHardwareMonitor 109 BB 5V Alarm  
 LocalHardwareMonitor 110 BB 12V Alarm  
 LocalHardwareMonitor 111 BB -12V Alarm  
 LocalHardwareMonitor 112 Fan 1H Alarm  
 LocalHardwareMonitor 113 Fan 2H Alarm  
 LocalHardwareMonitor 114 Fan 3H Alarm  
 LocalHardwareMonitor 115 Fan 4H Alarm  
 LocalHardwareMonitor 116 Fan 5H Alarm  
 LocalHardwareMonitor 117 Fan 1L Alarm  
 LocalHardwareMonitor 118 Fan 2L Alarm  
 LocalHardwareMonitor 119 Fan 3L Alarm  
 LocalHardwareMonitor 120 Fan 4L Alarm  
 LocalHardwareMonitor 121 Fan 5L Alarm  
 LocalHardwareMonitor 122 BB Temp Alarm  
 LocalHardwareMonitor 123 Amb Temp Alarm  
 LocalHardwareMonitor 124 CPU Temp Alarm  
 CHAMIIENCMON 6 Power Supply A failed  
 CHAMIIENCMON 16 Power Supply B failed  
 CHAMIIENCMON 21 One or more fans in Fan Module 1 failed  
 CHAMIIENCMON 23 One or more fans in Fan Module 2 failed  
 CHAMIIENCMON 25 One or more fans in Fan Module 3 failed  
 CHAMIIENCMON 26 Multiple fans failed. Critical  
 CHAMIIENCMON 31 Power Supply A is going to shutdown  
 CHAMIIENCMON 32 Power Supply B is going to shutdown  
 CHAMIIENCMON 33 Both Power Supplies are going to shutdown  
 UFS 7 Skip auto fsck for corrupted filesystem at mount  
 UFS 11 Crossed the root filesystem inode threshold  
 UFS 15 Filesystem size incorrect in superblock  
 STORAGE 1 Disk mark does not match  
 STORAGE 6 Incorrect reference/dereference on a volume  
 UPSMonitor 3 The UPS has failed its internal self-test  
 UPSMonitor 12 The UPS has been switched off by a management station  
 UPSMonitor 17 The UPS batteries require immediate replacement  
 UPSMonitor 18 An Environment contact closure has faulted  
 UPSMonitor 20 The UPS is on bypass due to an internal fault  
 UPSMonitor 24 The base module bypass power supply needs repair  
 UPSMonitor 25 The base module fan needs repair  
 UPSMonitor 51 The battery charger has failed  
 UPSMonitor 53 The battery temperature threshold has been violated  
 UPSMonitor 77 An abnormal condition has been detected  
 NFS 0 Failed to start NFS

## **SYS LOG RECORDS EVENTS BY SEVERITY AND EVENT ID:**

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
May 12 21:00:00 2006 NaviEventMonitor:**4:2** Backend Event Number 0x840 Host OEM-XOO25IL9VL9 Storage Array  
APM00041700339 SPA Device

Bus 1 Enclosure 1 Disk 11 SoftwareRev 6.19.0 (4.14) Unknown Error 2.19.0.701.5.027 Description Data Sector Invalidated  
May 12 21:00:00 2006 NaviEventMonitor:**3:3** Backend Event Number 0x957 Host OEM-XOO25IL9VL9 Storage Array  
APM00041700339 SPA Device

Bus 1 Enclosure 1 Disk 12 SoftwareRev 6.19.0 (4.14) Unknown Error 2.19.0.701.5.027 Description Uncorrectable Sector

**Note:** First number represents Severity, second number is the EventID for the facility called “NaviEventMonitor”. Use # nas\_event -list -f -i to obtain the list of facilities.

**USING SERVER SYSSTAT:** Operating System statistics by kernel thread; free memory; free CPU time

\$server\_sysstat server\_2 server\_3 :

|                          |  |
|--------------------------|--|
| threads runnable = 2     | [number processes running and not waiting for I/O] |
| threads blocked = 170    | [number processes running and waiting due to I/O]  |
| threads I/J/Z = 1        | [threads in zombie status, not an error!]          |
| memory free(kB) = 419998 | [free memory]                                      |
| cpu idle_% = 99          | [processor idle time in %]                         |

**SCO Control Station Statistics:** \$sar -u 5 5 or \$sar -u 1 5 [Provides STATS on CS0 CPU activity, et al for SCO 2.1.3]  
\$sar -d 15 10 [Provides disk activity report every 15 seconds for 10 times]

**Linux CS Statistics:** \$top and \$vmstat commands

**Quick Way to See if DataMovers Have Panic Information Logged:**

\$egrep -i panic /nas/log/sys\_log [/nas/log/eventhome]

## **CELERRA DUMP FILES/PANICS:**

**Note:** DataMover “panics” create memory ‘dump’ files that can be extracted and used for analysis by Engineering. Default location for writing dumps to CS is /nas/var/dump [5.3 & 5.4 only]. When dumpfiles are not automatically written to the CS, there are still retained in slots on LUN 0 backend. These dumps can be forced from the slot by using the dump\_slot utility. Dump files are actually written to LUN 0 on disk. For the Celerra backend, there are three full, dedicated dump\_slots that are used for full panic dumps, with a total capacity of 3.1GB each, though with current 5.4 code, a typical dump is 1.6GB. There are also (2) partial dump slots that can capture up to 4MB of data, per Data Mover. Please note that with AR44489, NAS Versions 5.2.22.0, 5.3.16.0, & 5.4.12.0, panic dumps are no longer automatically extracted and written to the Control Station, /nas/var/dump, but must be deliberately and manually extracted using the dump\_slot utility. Dumps are generated directly by DART when the panic() function is called, and written directly to one of the available dump\_slots on LUN 0 of the backend. The Control Station is not involved in this process. If the Server loses connectivity to backend, then the panic dump cannot be written to LUN 0, so panic handler will not be updated, but a sys\_log entry should still be generated.

## **NAS 5.5 DUMP SLOT SIZES:**

# /nas/sbin/workpart -r

|                        |                                 |
|------------------------|---------------------------------|
| workpart_magic         | = 0xbeaf0001                    |
| full.dumps[0].size     | = 0x300004 = 3,145,732 or 3.1GB |
| full.dumps[1].size     | = 0x300004                      |
| full.dumps[2].size     | = 0x300004                      |
| dump_slices[0][0].size | = 0xffff = 4096 or 4MB          |
| dump_slices[0][1].size | = 0xffff                        |

**Note:** Celerra can write a total of (3) full 3.1GB dumps to the backend LUN 0 and (2) Partial 4MB dumps for each Data Mover slot. Cognac will remove one of the partial dump slots on each DM. Compressed dumps will be introduced with GrandNapa release. Crash header will show if dump is compressed or not and normal dump\_slot commands will still apply. Eng. recommends gzipping any dumps prior to FTP. Can transfer nas\_crash and dump\_slot binaries from newer code to older systems.

→Use #/nas/sbin/& -v -F utility to mark the 'dump' partition for cleanup of Full Dumps

→Use #/nas/sbin/dump\_slot -d /tmp/slot2.dump -v -F 2 [use to extract dump from 2<sup>nd</sup> full dump slot]

→Use #/nas/sbin/dump\_slot -v to cleanup partial dumps

→Use #/nas/sbin/nas\_crash -F 0 | 1 | 2 to see the panic headers of any existing dumps in the full slots

## **ZEROING OUT DUMP SLOTS SO THAT NEW PANICS CAN BE WRITTEN TO BACKEND:**

### **1. Verifying current dump\_slot Status:**

# /nas/sbin/dump\_slot -v -F -x

Partition 0 starts at 0x1, size 273042 (0x42a92) blocks.

DART Work Partition Layout found @ LBA 0x43000 (134MB boundary)

Checking LBA 0xf3800 for Full Dump...

Tue Feb 14 08:19:31 2006 UTC

Product: EMC Celerra File Server

Uptime: 006 days, 21:03:52

Version: T5.5.18.2

IP Address: 192.168.2.2

Host Name: server\_2

Checking LBA 0x3f380a for Full Dump...

Checking LBA 0x6f3814 for Full Dump...

**Note:** Output indicates that a dumpfile exists in the first dump\_slot position but that no dumps have been written to the other (2) dump\_slots

**2. Verify whether the dumps are valid or invalid, or use to dump to CS partition:**

**# /nas/sbin/dump\_slot -v -F**

Partition 0 starts at 0x1, size 273042 (0x42a92) blocks.

DART Work Partition Layout found @ LBA 0x43000 (134MB boundary)

Checking LBA 0xf3800 for Full Dump...

\*\*\*\*\*

DART Slot 2, Full Dump 0 Found @ LBA 0xf3800 (997376)

Dump info:

DART Slot 2

magic = 0xfacedead

dump\_lba = 0xf3805

dump\_type = 0x3

dump\_chunks = 0x30000

dump\_chunksz = 0x2000

DART panic message: >>PANIC in file: ../../dskdump.cxx at line: 1450 :

to see if ip rep survives

DART stack trace:

0xd6617a80: 0x13d047 waitForReboot+0x8b

0xd6617aa0: 0x13d24f fault\_dump+0x67

0xd6617ac0: 0x13d14f PANIC+0x2b

0xd6617ad0: 0x830c04 \_Z11paniccfgcmdR6Option+0x178

0xd6617c10: 0x161765 \_ZN11cfgcmd\_desc6invokeEPKcPP11CCMDMessageii+0

0xd6617c80: 0x1fb25e \_Z10xml\_cfgcmdP7macDataRP8CCMD\_Msg+0x30e

0xd6617cd0: 0x1f7652 \_ZN17macCCMDCommand\_DB6invokeEPKcP7macDataRP8C

0xd6617d00: 0x1f7b76 \_Z17mac\_CCMDDocParserP7macDataRP8CCMD\_Msg+0x20

**Note:** Run dump\_slot -v -F to view panic header and also to verify the magic value of the dump\_slot. When the magic value is something other than 0, a dump\_slot exists and has not yet been marked for overwrite.

**3. Use the following command to zero out the magic values for each dump\_slot, thereby enabling the Celerra to overwrite the dump\_slots with new panic dumps:**

**# /nas/sbin/dump\_slot -c 0**

**# /nas/sbin/dump\_slot -c 1**

**# /nas/sbin/dump\_slot -c 2**

**Note:** The above command does not return a success or failure message, though each dump\_slot becomes marked for overwrite (deletion).

**4. Verify Dump\_slot -c Results:**

**# /nas/sbin/dump\_slot -v -F**

Partition 0 starts at 0x1, size 273042 (0x42a92) blocks.

DART Work Partition Layout found @ LBA 0x43000 (134MB boundary)

Checking LBA 0xf3800 for Full Dump...

Dump info FOUND:

Dart Slot = 2

magic = 0

dump\_lba = 0xf3805

dump\_type = 0x3

dump\_chunks = 0x30000

dump\_chunksz = 0x2000

No Full Dump Found

**Note:** Magic value is set to zero and "No Full Dump Found" is indicated this time--system is now able to overwrite any existing dumps if a new panic occurs. Please note that as in past behavior, until the slot is actually overwritten by a new panic dump, the contents of the existing dump still exist on disk and can be extracted to the Control Station.

## **GZIPPING DUMP FILES ON PARTITIONS WITH LITTLE FREE SPACE:**

```
# cat slot3.dump0 | gzip -9 -c >/nas/var/dump/slot3.dump.gz &
```

```
/dev/ndf1      1.7G  1.4G  354M  79% /nas/var  
# ls -la  
-rwxr-xr-x  1 root  root  1154796032 Feb  9 17:30 slot3.dump0  
-rw-r--r--  1 root  root   29032448 Feb  9 17:41 slot3.dump.gz
```

## **CHECKING FOR FULL DUMP SLOT PANICS:**

```
# /nas/sbin/dump_slot -v -F -x
```

Partition 0 starts at 0x1, size 273042 (0x42a92) blocks.  
DART Work Partition Layout found @ LBA 0x43000 (134MB boundary)

**Checking LBA 0xf3800 for Full Dump...**

Wed Oct 19 13:44:31 2005 UTC

Product: EMC Celerra File Server

Uptime: 009 days, 04:12:34

Version: T5.4.14.3030

IP Address: 192.168.1.4

Host Name: server\_4

**Checking LBA 0x3f3805 for Full Dump...**

Tue Jun 14 21:06:30 2005 UTC

Product: EMC Celerra File Server

Uptime: 001 days, 00:37:41

Version: T5.4.12.0

IP Address: 192.168.1.2

Host Name: server\_2

**Checking LBA 0x6f380a for Full Dump...**

Tue Jun 14 21:11:50 2005 UTC

Product: EMC Celerra File Server

Uptime: 000 days, 00:03:17

Version: T5.4.12.0

IP Address: 192.168.1.2

Host Name: server\_2

## **NAS 5.4 CONTAINS (3) FULL DUMPSLOTS FOR MEMORY DUMPS:**

**Note:** As seen from /nas/sbin/workpart -r output

|                    |            |
|--------------------|------------|
| full.dumps[0].lba  | = 0xf3800  |
| full.dumps[0].size | = 0x300009 |
| full.dumps[1].lba  | = 0x3f380a |
| full.dumps[1].size | = 0x300009 |
| full.dumps[2].lba  | = 0x6f3814 |
| full.dumps[2].size | = 0x300009 |

## **AUTOREBOOT AFTER PANIC:**

```
param kernel autoreboot=600
```

**Note:** Amount of time in seconds that DART will wait after panicking, before rebooting

## **DEFAULT LOCATION OF DART AUTOMATIC & MANUAL DUMPS:**

```
/nas/var/dump
```

**Note:** NAS 5.3.6.0 & 5.4.1.0 changed manual dump\_slot behavior from putting dump in /nas/log to /nas/var/dump, but later versions write dumps to LUN 0 dump\_slots on backend only and do not extract to /nas/var.

**DEPOSIT DUMP TO ALTERNATE LOCATION ON CS: /nas/var/dump /dev/sdf1**

```
$/nas/sbin/dump_slot -v -F -d /nas/var/dump/dump_slot.2
```

**Note:** Later NAS versions write panic dumps automatically to /nas/var/dump, but the default location when using /nas/sbin/dump\_slot would still output dump to /nas/log, creating potential problems with the /nas partition filling up. NAS 5.3.6.0 & 5.4.1.0 now use

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
/nas/var/dump as the default location when using the dump\_slot command manually. NAS 5.2.22.0, 5.3.16.0, & 5.4.12.0 no longer writes the dump to /nas/var/dump by default, just to the backend dump\_slot.

### **WATCHDOG PANICS:**

Watchdog panics mean that something in the system was in a deadlock or a never-ending state

### **NMI EXCEPTION PANICS:**

NMI is a “Non-Maskable Interrupt” that exists so that the processor can be interrupted no matter what. This is only necessary when there are certain hardware errors that require immediate shutdown so as to stop the system from further damaging itself. An uncorrectable memory error would be an example.

### **COMPRESSING DUMPS:**

```
$compress -v slot3.dump.0203121420      [536873472]      [SCO]  
slot3.dump.0203121420.Z                  [200161175]  
$ gzip slot3.dump.0211121824            [LINUX]  
-rwxr-xr-x 1 nasadmin nasadmin 145053685 Nov 12 18:26 slot3.dump.0211121824.gz
```

### **MANUALLY COLLECTING PANIC DUMPS:** [/nas/var/dump    /nas/log]

#### **Step 1. Locate the Dump Slot Files:**

```
# ls -l /nas/log/slot*          [SCO]  
-rwxrwxr-x 1 nasadmin nasadmin 109731331 Nov 17 09:31 /nas/log/slot4.dump0.1171.Z  
#ls -la /nas/var/dump*        [LINUX SYSTEMS]  
-rwxr-xr-x 1 root  root  529574400 Nov 12 18:26 slot3.dump.0211121824
```

#### **Complete 536MB Dump File:**

```
-rwxrwxrwx 1 nasadmin root  536873472 Mar 12 14:22 slot3.dump.0203121420
```

#### **Step 2. Ensure there is enough room to Uncompress Dumpfile:** #df -k

```
/dev/sdd1    1818352 123640 1602340 7% /nas/var
```

#### **Step 3. Uncompress the Dump File:**#uncompress /nas/log/slot4.dump0.1171.Z & [Run as bg process--takes about 15 min]

```
# jobs  
[1]+ Running      uncompress /nas/log/slot4.dump0.1171.Z &
```

**Note:** #fg %1 to bring background job to foreground; use ctrl + c to end job

#### **Step 4. # wait** [Displays prompt when background job is completed]

```
# ls -al slot*  
-rw----- 1 root  root  209350656 Nov 19 16:02 slot4.dump0.1171  
-rwxrwxr-x 1 nasadmin nasadmin 109731331 Nov 17 09:31 slot4.dump0.1171.Z
```

#### **Step 5. View Dump Header using Nas Crash:**

```
# /nas/sbin/nas_crash slot4.dump0.1171 -i [Always use -i switch!]
```

-----  
DART time of dump: Sat Nov 17 12:37:00 2001

Product: Symmetrix Network File Storage

Uptime: 003 days, 06:17:26

Version: T2.2.35.4

IP Address: 192.168.1.4

Host Name: server\_4

-----  
DART panic/fault message:

```
>>PANIC in file: ./addrspac.cxx at line: 410 :
```

***Out of memory***

DART stack trace:

```
0x37eb4: 0x139a64 _PANIC+0x28  
0x37ed4: 0x13678d _allocPages_19Memory_AddressSpaceiG12Memory_Owne  
0x37f28: 0x1678f5 _more_memory_FiP9free_list+0x51  
0x37f4c: 0x167a05 _malloc+0x8d  
0x37f68: 0x168252 __builtin_new+0x16  
0x37f80: 0x1682ca __builtin_vec_new+0xa  
0x37f98: 0x19f51c _mactrh_v_cfgcmd_FP4msgb+0x20  
0x3800c: 0x19171e _start_11Mac_Initial+0x26
```

DART panic/fault current cpu registers:

```
edi   =     0    esi   = 1362db  
ebp   = 37ed0    oldEsp = 37ebc  
ebx   = fffe    edx   =     1  
ecx   =     4    eax   =     0  
eflags = 202
```

**Step 6. Take snapshot of Server Log 20 lines before and 10 lines after Panic**

```
#server_log server_4 -a -s > slog4.log
```

**Step 7. Compress the dump file: #compress -v slot3.dump0 [slot3.dump0.Z]**

[or use gzip to compress]

**Step 8. FTP the “slot dump” and ProComm “capture file” to Engineering using TFTP process**

**Step 9. Clean up DUMP Files from Celerra:**

```
#/nas/sbin/dump_slot -v -F [full dumps] | -v [partial dumps]
```

Creating dump file "/nas/log/slot4.dump0"

```
#/nas/var/dump      #/nas/log
```

**Note:** Answer Yes to all questions except if you want to "Continue Y/N" make selection

**(Old Method—see new dump\_slot -c procedure) CLEANING UP DUMPFILE SLOTS ON CELERRA:**

**1. Log into system and verify current partition usage on the Control Station:**

```
$ df -h
Filesystem      Size Used Avail Use% Mounted on
/dev/sdc3        1.4G 710M 697M 51% /
/dev/sdc1        30M  4.0M 24M 14% /boot
none            251M  0 250M 0% /dev/shm
/dev/sde1        1.7G 1.1G 581M 66% /nas
/dev/sda1        133M 78M 55M 59% /nas/dos
/dev/sdf1        1.7G 886M 799M 53% /nas/var
```

**Note:** Default directory for writing dumps is /nas/var/dump—newer 5.4 dumps require 1.6GB space

**2. Go to /nas/var/dump directory and delete any old dump files that might be taking up the space:**

```
$ ls -la
-rwxrwxr-x  1 nasadmin nasadmin 111550995 Sep 20 10:48 slot2.dump0.gz
$ rm *
```

**3. Verify space on Control Station:**

```
# df -h
Filesystem      Size Used Avail Use% Mounted on
/dev/sdc3        1.4G 710M 697M 51% /
/dev/sdc1        30M  4.0M 24M 14% /boot
none            251M  0 250M 0% /dev/shm
/dev/sde1        1.7G 1.1G 585M 66% /nas
/dev/sda1        133M 78M 55M 59% /nas/dos
/dev/sdf1        1.7G 48M 1.5G 3% /nas/var -->Now have enough space to write a full panic dumpfile
```

**4. Now we can begin the tedious process of cleaning up each memory slot:**

**Option 1:** For systems below NAS Version 5.3, use following syntax to redirect the dumpfile to /nas/var/dump. Failure to do so can fill /nas to 100% and cause file system corruption:

```
#/nas/sbin/dump_slot -v -F -d /nas/var/dump/dump_slot.2 (directing dumpfile to specific location and assigning name to the dump)
```

**Option 2:** For NAS 5.3 and higher, use following command to extract dumps:

```
# /nas/sbin/dump_slot -v -F
```

Partition 0 starts at 0x1, size 273042 (0x42a92) blocks.

DART Work Partition Layout found @ LBA 0x43000 (134MB boundary)

Checking LBA 0xf3800 for Full Dump...

```
*****
```

DART Slot 2, Full Dump 0 Found @ LBA 0xf3800 (997376)

Dump info:

DART Slot 2

magic = 0xfacedead

dump\_lba = 0xf3805

dump\_type = 0x3

dump\_chunks = 0x20022

dump\_chunksz = 0x2000

DART panic message: finishTake: bad magic in CG

DART stack trace:

```
0x8cf0eca0: 0x13b74b waitForReboot+0x8b
```

```
0x8cf0ecc0: 0x13b94a fault_dump+0x5e
```

```
0x8cf0ece0: 0x13b853 PANIC+0x2b
```

```
0x8cf0ecf0: 0x58f561 _ZN14UFS_FileSystem28markCorruptedFsForUnmount
```

```
0x8cf0ed80: 0x5a62fd _ZN14UFS_CgBlkEntry10finishTakeEv+0x157  
0x8cf0ede0: 0x5a7921 _ZN14UFS_CgBlkEntry11preAllocInoEP20UFS_Update  
0x8cf0ee10: 0x597c3e _ZN14UFS_FileSystem9hashallocEP20UFS_UpdateDes  
0x8cf0ee50: 0x597998 _ZN14UFS_FileSystem8allocInoEP20UFS_UpdateDesc
```

Do you wish to save this dump? [y or n] y

Dump size is 1074022912 bytes. A dump file of this size will be created.

Do you wish to continue? [y or n] y

Creating dump file "/nas/var/dump/slot2.dump0".

**Note:** At this point, it will take the system 10-15 minutes to completely write the full dump to the /nas/var/dump directory. Answer yes to all questions that are asked.

Dump to file "/nas/var/dump/slot2.dump0" complete.

Do you wish to erase this dump from the DART disk? [y or n] y

Checking LBA 0x3f3805 for Full Dump...

```
*****
```

DART Slot 2, Full Dump 1 Found @ LBA 0x3f3805 (4143109)

Dump info:

DART Slot 2

magic = 0xfacedead

dump\_lba = 0x3f380a

dump\_type = 0x3

dump\_chunks = 0x1afb4

dump\_chunksz = 0x2000

DART panic message: >>PANIC in file: ../../dskdump.cxx at line: 1443 :

Force panic for hung FS investigation

DART stack trace:

```
0x8f665ae0: 0x13b74b waitForReboot+0x8b
```

```
0x8f665b00: 0x13b94a fault_dump+0x5e
```

```
0x8f665b20: 0x13b853 PANIC+0x2b
```

```
0x8f665b30: 0x7db076 _Z11paniccfgcmdR6Option+0x178
```

```
0x8f665c70: 0x15b0f8 _ZN11cfgcmd_desc6invokeEPKcii+0x124
```

```
0x8f665ce0: 0x1ee302 _Z10xml_cfgcmdP7macData+0x282
```

```
0x8f665d20: 0x1eb52c _ZN13macCommand_DB6invokeEPKcP7macData+0x74
```

```
0x8f665d50: 0x1eb7e3 _Z13mac_DocParserP7macData+0x19
```

**5. From the crash header above, determine if dumpfile is needed or not (i.e., check the date and size to see if it's a keeper)**

**6. Delete dumpfile from /nas/var/dump and repeat step 4. to extract any other full dumps that might be written to memory.**

**7. After full dumps have been extracted, erased from the memory slot, and deleted, run the following command to extract any "partial" dumpfiles that might exist.** Follow same procedure as in steps 4-5 and repeat until all slots are empty:

#/nas/sbin/dump\_slot -v -d /nas/var/dump/dump\_slot\_partial.2 (NAS 5.2 or below)

#/nas/sbin/dump\_slot -v (NAS 5.3 or higher)

## **FORCING DUMP & REDIRECTING TO ALTERNATE DIRECTORY:**

#/nas/sbin/dump\_slot -d /home/nasadmin/slot3.dump0 -v -F [ -d switch redirects to specified path]

**Note:** Make sure you give the dump a name when specifying the new location or the script won't work!

## **GZIPPING DUMPFFILE:**

#cat slot3.dump0 | gzip -9 -c > /nas/log/slot3.dump0.gz &

**Note:** This procedure will allow you to gzip a large dump in a partition that would not normally be able to conduct a gzip due to the partition being nearly filled.

## **COLLECTING DUMP FILES ON DATAMOVERS:**

**Checking for DumpFiles:** #ls -l /nas/log/slot\* [this is where slot\_x.dump0 would be if dumps had occurred]  
#/nas/log /slot2.dump0

**Dump Summary Info:** \$/nas/sbin/nas\_crash slot2.dump.9906231630 -i >slot\_2.dump1 #more slot\_2.dump1

**Collecting Dump:** #/nas/sbin/dump\_slot -v -F [Polls all slots on DM's & finds dumps. Prompts for extraction or not]  
a.) Do You Wish to Save Dump? Yes  
b.) Do you Wish to erase this Dump from the DART Disk? Yes {**Note:** Doesn't actually remove dump, frees mem slot}

## **FORCING PANIC DUMPS FROM LUN 0 DUMP SLOTS:**

#/nas/sbin/dump\_slot -v -F

## **PANIC DUMPS & THE CFS 4.0 SERVER:**

Due to the increase in Memory for the 510 DataMovers, dump files have changed to a new "sparse" format  
If using a 2GB system disk config, the default is for (1) 1GB dumpslot  
If using a 4GB system disk config, the default is for (3) 1GB dumpslots  
**Note:** System can be reconfigured for (1) 3GB dumpslot if necessary

## **EXTRACTING DUMPS DIRECTLY FROM DUMP SLOT & UPLOADING TO EMC FTP SERVER:**

**Note 1:** With NAS 5.4 there are (3) full 1.6GB dump\_slots that are shared between the Data Movers. When a Data Mover panick occurs, the memory from the DM is written to one of the full dump\_slots, if available, or to (2) partial dump\_slots for each Server, to backend Control LUN 0. Because of the recent NAS 5.4 increase in dump size, there is generally not enough native free space available to extract the 1.6GB dump to /nas/var/dump. The following procedure and /nas/tools/ftp\_dump tool can be used to transfer dumps directly from dump\_slots to an FTP Server directory, precluding the need to output dumps to the Control Station first.

**Note 2:** AR44489 stopped the autosaving of panic dumps to the Control Station partition—all dumps need to be manually extracted with NAS Versions 5.2.22.0, 5.3.16.0, 5.4.12.0.

Note 3: The FTP Dump utility will no longer be supported in NAS 5.6

### **1. Configure an ftp.cfg transfer file in /nas/tools directory with the following information:**

# cat /nas/tools/ftp.cfg

host ftp.emc.com

user anonymous

pass ftp\_ftp@emc.com

**Note:** Host can be in form of hostname or IP Address. Email address needs to be in format "user\_name@emc.com" to be a valid password entry

### **2. Execute the following command to list out the various dates associated with Full Dump Slots:**

# /nas/sbin/dump\_slot -v -F -x

Partition 0 starts at 0x1, size 273042 (0x42a92) blocks.

DART Work Partition Layout found @ LBA 0x43000 (134MB boundary)

Checking LBA 0xf3800 for Full Dump...

Sun Dec 18 18:49:17 2005 UTC

Checking LBA 0x3f3805 for Full Dump...

Sat Dec 17 18:51:41 2005 UTC

Checking LBA 0x6f380a for Full Dump...

Sat Dec 17 19:33:52 2005 UTC

**Note:** Identify the dump date that is relevant and note the LBA partition information for use when extracting dump

### **3. Execute the script to find and transfer the desired dump to the EMC FTP Site:**

**Note:** Since there may be a number of dumps available, ranging in age from current to old dumps, a decision will need to be made on which ones to extract and transfer, and which ones not to extract.

#/nas/tools/ftp\_dump -f ftp.cfg

Please enter the remote-directory: /incoming/ftp\_dmp

Partition 0 starts at 0x1, size 273042 (0x42a92) blocks.

DART Work Partition Layout found @ LBA 0x43000 (134MB boundary)

Checking LBA 0xf3800 for Full Dump...

No Full Dump Found

**Do you wish to save this dump anyway? [y or n] y**

\*\*\*\*\*

DART Slot 2, Full Dump 0 Found @ LBA 0xf3800 (997376)

DART panic message: >>PANIC in file: ../dskdump.cxx at line: 1437 :

Server\_2 PANICKED

DART stack trace: -----output abbreviated-----

**Do you wish to save this dump? [y or n] y**

Dump size is 660834816 bytes. A dump file of this size will be created.

**Do you wish to continue? [y or n] y**

Creating dump file "/nas/var/dump/29797.dump".

**Dump file "/nas/var/dump/29797.dump" already exists. Overwrite it? [y or n]: y**

**Note:** A process file 29797.dump is created in /nas/var/dump during the dump transfer from dump\_slot to FTP Server directory, and then erased after the operation is completed. This process file works in 4k chunks—a minimum of 4k of freespace is required in the /nas/var/dump directory for this procedure to work:

prw-r--r-- 1 root root 0 Jan 9 13:41 29797.dump

prw-r--r-- 1 root root 4096 Jan 9 13:42 29797.dump

**Dump to file "/nas/var/dump/29797.dump" complete.**

**Do you wish to erase this dump from the DART disk? [y or n] y**

**4. At this point, the dumpfile has been transferred to FTP Server and the dump\_slot marked for deletion of the old dump**

**5. Answer following questions below, to either transfer other dump files, or to complete the current activity:**

Checking LBA 0x3f3805 for Full Dump...

DART Slot 2, Full Dump 1 Found @ LBA 0x3f3805 (4143109)

Dump info:

DART Slot 2

magic = 0xfacedead -----output abbreviated-----

DART panic message: >>PANIC in file: ./dskdump.cxx at line: 1437 :

Forced panic on DM2

DART stack trace: -----output abbreviated-----

**Do you wish to save this dump? [y or n] n**

Checking LBA 0x6f380a for Full Dump...

\*\*\*\*\*

DART Slot 2, Full Dump 2 Found @ LBA 0x6f380a (7288842)

Dump info:

DART Slot 2

magic = 0xfacedead

DART panic message: >>PANIC in file: ./dskdump.cxx at line: 1437 :

This is Test Panic for Server 2

DART stack trace:

**Do you wish to save this dump? [y or n] n**

**dump saved as 29797.dump @ /incoming/ftp\_dmp**

done.

**6. Dump file has been transferred to FTP site /incoming/ftp\_dmp:**

-r----- 1 u3 g3 660834816 Jan 17 19:02 29797.dump

## **SPLITTING BINARY DUMP FILES INTO SEGMENTS IN ORDER TO FTP TO EMC:**

Step 1. Prior to splitting a Binary Dump File, run the following command to obtain a CheckSum of the File

(Checksum number will be used after the files are concatenated back together to verify the integrity of the final Dump File):

**#/usr/bin/cksum slot6.dump.0208261933 [or #sum /var/dump/]**

**3661864574 268184064 slot6.dump.0208261933**

[3661864574 is the CheckSum of the Original Dump File]

Step 2. From a directory on Control Station that has adequate space, run the split command:

**\$/nas/var/split -b 250000000 /nas/log/slot6.dump.0208261933**

**Note:** This command will split the binary dump file into 250MB segments that are named “xaa, xab, xac,” etc. This eliminates the original slot6.dump filename and deposits the segments in the directory from which the command was run.

Step 3. Gzip the segments #gzip \* [xaa.gz, xab.gz, xac.gz]

Step 4. FTP the zipped files to EMC, unzip, then concatenate the binary files together again [\$gzip -d xaa ]

**\$cat xaa xab xac >slot6.dump.0208261933**

Step 5. Verify the integrity of the binary file:

**#/usr/bin/cksum slot6.dump.0208261933**

**3661864574 268184064 slot6.dump.0208261933**

## **GZIPPING DUMP USING CAT COMMAND TO REDUCE SIZE OF DUMPFLE:**

1. #cd /nas/var/dump

2. Zip dump to location with adequate room: **#cat slot2.dump0 | gzip -9 -c > /home/nasadmin/slot2.dump0.gz &**

3. Upload file to FTP Site & send to Engineering

## **ANOTHER METHOD FOR GZIPPING DUMP TO SPECIFIC DIRECTORY:**

1. #mkfifo /tmp/fifo

2. #cat /tmp/fifo lzip > /nas/var/dump/slot2.dump0.gz &

3. #/nas/sbin/dump\_slot -s -d /tmp/fifo -v -F

4. #rm -Rf /tmp/fifo

## **STABILIZING DM's IN "Rolling Panic" SITUATIONS:**

1. Stop NAS Services, if time permits, to stop the annoying CallHome messages on screen:

**#/sbin/service nas stop**

2. Verify that Box Monitor and nas\_mcd are no longer running:

**#ps -ef |grep boxm**

**#ps -ef |grep -i nas\_mcd**

3. Manually remount file systems on Control Station:

**# mount /nbsnas # mount /nas # mount/nas/dos #mount/nas/var**

4. Identify panicked slot and obtain Server Log:

**# nas\_server -l**

**# /nas/sbin/getreason**

**# /nas/sbin/log\_slot -s -a 2 >s2log**

5. Look for cause of panic in the log. Rolling panics are often caused by a corrupted file system that cannot be remounted by the Data Mover during bootup. Look for evidence of file system corruption, FSTOOLS activity for fsck's, in the log.

**Note:** NAS 5.4 should be able to handle corrupted file systems better and allow for the Server to bootup without the corrupted file system mounted, conducting autofsck, etc.

6. Once identified, make a copy of the /nas/server/slot\_x/mount file, then vi edit and delete the offending file system line entirely from the mount file:

**#cp -ip mount mount\_ori**

**# vi mount** (edit the file & remove following line using dd to delete the line)

**uxfs rw /fs31 359=1283 rw**

**Note:** If DM continues to panic, find the corresponding “359=1283” entry in the ufs file and remove this entry

**# cat ufs**

file recover ufs 73=131072 14=5 122=18 36=16 357=1282 **359=1283**

**Note:** Remove the highlighted entry, which is for file system /fs31 or volume # 359

7. Use **# server\_cpu server\_x -r now** to reboot server. This will rebuild the Server configuration based on the current state of its database files, and therefore not attempt to mount the offending file systems.

#### **ADDITIONAL COMMENTS:**

→ In some situations, Checkpoints can cause the Rolling Panic scenario to occur, and may require a permanent unmount and deletion of the Checkpoints associated with a particular file system. Again, the Server Log should be the first place to check to identify the source of the rolling panics

→ In other situations, IP Replication file systems can also cause rolling panics. Determine the offending file system, then comment out the respective ipfsclone line from the panicking server's netd file

→ In rare cases, a newly implemented Server Param entry may be the cause of the rolling panic and would need to be removed

→ An alternative method of commenting out file systems would involve making a copy of the appropriate /nas/dos/slot\_x/boot.cfg file and editing the boot.cfg as in following example, followed by **#/nas/sbin/t2reset reboot -s 2**. The T2reset command will boot the server based on the current state of the boot.cfg file:

file recover ufs 73=131072 14=5 122=18 36=16 357=1282 **359=1283 (remove this entry)**

#file mount ufs rw /fs31 359=1283 rw

### **EXAMPLES OF FILE SYSTEM RELATED PANICS:**

#### **IO FAILURE PANIC:**

>>PANIC in file: ../../volume.cxx at line: 214 :

IO failure despite all retries/failovers

---

#### **READ BLOCK FAILURE PANIC:**

>>PANIC in file: ../../ufsio.cxx at line: 340 :

readBlock failed

---

#### **BAD MAGIC PANIC:**

finishTake: bad magic in CG

#### **Corresponding Server Log Entry specifies problem file system:**

2004-08-23 15:58:38: UFS: 3: bad magic num 8c3a5178 in cylinder group at blk 6ac0030

2004-08-23 15:58:38: CFS: 3: **Unmounting fs 56**: Reason: finishTake: bad magic in CG

---

#### **MANGLED DIRECTORY ENTRY PANICS:**

DART panic/fault message: mangled directory entry

#### **Corresponding Server Log Entry:**

UFS: 3: fs /mnt02: bad dir ino 615176 at offset 7658: mangled entry

>>PANIC in file: ./ufsdir.cxx at line: 184 :

UFS\_DirHashIter::getEntry: mangled entry

### **VARIOUS SERVER LOG ENTRIES FOR FILE SYSTEM PANICS:**

UFS: 3: Partially truncated inode 292588 in fs 77

You can recover the storage for this file by running fsck at a convenient time

UFS: 3: bad magic num b1a3ca0f in cylinder group at blk c9a0030

UFS: 3: Corrupted file: ino 522768, fsid 75, size 100000000001

UFS: 3: fs /fs2: bad dir ino 1515984 at offset 192: mangled entry

CFS: 3: Unmounting fs 19: Reason: mangled directory entry

STORAGE: 3: readBlock() : Volume 156 Bad Block Index 233293f01067406870: UFS: 3: readBlock failed, err 16

>>PANIC in file: ./ufsinode.cxx at line: 894 : Oversized file: need to run fsck

DART panic/fault message: mangled directory entry

>>PANIC in file: ./ufsio.cxx at line: 340 : readBlock failed

### **CASE STUDY TO FIND FILE SYSTEM BASED ON TARGET & LUN INFO FROM LOG ONLY:**

1. In certain cases, you may see only the following entries in the Log just prior to a panic, and may need to do some detailed work to find out what the offending file system name is. In the following example, all that is recorded is the Target 11 & Lun 14.

#### **Server Log:**

2005-09-27 12:20:03: CAM: 3: FCP I/O Error: c16t11114: CamStatus 04 ScsiStatus18

2005-09-27 12:20:03: CAM: 3: FCP I/O Error: c0t11114: CamStatus 04 ScsiStatus 18

2. Grep the camdisk file to find the “d” volume & device that matches up with Target 11, Lun 14.

# egrep t11114 /nas/server/slot\_7/camdisk

**68:c0t11114+556812~~2E2~~031+,c16t11114+5568122E2201+:** (Line 68 is d68 & device is 2E2)

3. Use following commands to determine file system name:

# nas\_disk -i d68

clnt\_volume = stv9

#nas\_volume -i stv9

clnt\_volume = mtv3

# /nas/sbin/rootnas\_volume -i mtv3

**clnt\_filesys= prodFO\_fs**

# nas\_fs -l

20 y 1 0 231 prodFO\_fs 1 (FSID is 20 and Volume # is 231)

### **HOW TO VERIFY IF A PANIC IS IN PROGRESS:**

**\$ps -ef |grep recover\_slot** [If this process is running, a panic is in progress & you must let it play out]

### **GENERATING SERVER PANIC:**

**\$server\_config server\_x “panic for a 0 percent cpu condition”**

### **FORCING SERVER PANIC USING MAC\_DB TOOL:**

1. Create link to MAC\_DB Utility in /nas/tools directory:

**# ln -s /nas/tools/mac\_db \_mac\_db**

2. Open MAC\_DB on Server\_3:

**# /nas/tools/\_mac\_db server\_3**

**macdb>**

3. Obtain List of Memory Registers by running "Show Thread" command:

**macdb> sh thr**

Dummy at 35f6010: State = IDLE

---output abridged-----

logXmlS at dd8cc448 0: Waiting for Condition at 211a910

logXmlS at dd8cc4fc 0: Waiting for Condition at 211a910

logXmlS at dd8cc5b0 0: Waiting for Condition at **211a910**

4. Select a memory register from something that appears innocuous—consult with TS2!:

**macdb> sh thr 211a910**

**Note:** System hangs at this point, indicating a reboot

5. Use ctrl + C to exit from the macdb> prompt

6. Grep for Panic Dump Recovery Process, check /nas/var/dump, or run dump\_slot utility to see if dump was written:

### # ps -ef |grep recover

```
root      9  1 0 Apr19 ?    00:00:00 [mdrecoveryd]
root   24152 1583 0 07:02 ?    00:00:00 /bin/sh /nas/sbin/recover_slot /
nasadmin 24410 24152 3 07:02 ?    00:00:00 /nas/sbin/slot_recover 3
```

### # /nas/sbin/dump\_slot -v -F -x

**Note:** If first attempt does not panic Server, continue using 'sh thr' command with different Memory addresses until successful  
7. Remove \_mac\_db link from /nas/tools directory

### **ANOTHER METHOD TO PANIC A TOTALLY HUNG DM:**

1. Obtain the server\_xml.tar.gz file from Kore: /pub2/utils
2. Push file to Celerra Control Station & copy to newly created directory
3. #tar -zxpvf server\_xml.tar.gz
4. #chmod 700 \*

### 5. #./server\_xml server\_x cfgcmd

**Note:** This will create a server panic and can be used with NAS 5.2, 5.3, 5.4

### **YET ANOTHER METHOD TO PANIC HUNG DATA MOVER:**

### # /nas/sbin/t2tty -c 2 "panic hung mac thread issue"

#### **How to Capture Screen Inputs and Displays on the Linux Control Station:**

**#cd; script** [Begins file capture of screen session to file ‘typescript’} –Use ctrl + d to end script session

**STARTING LINUX SCREEN SCRIPT CAPTURE:** \$script /tmp/aaa [starts script session in /tmp directory with filename “aaa”]

### **CELERRA EVENTS:**

\$nas\_event -l -f MasterControl | BoxMonitor [This command lists events for each ‘facility’]

**\$ nas\_event -l -f -i** [Lists the facilities on the Celerra: AAF; VideoService;BoxMonitor;MasterControl]

\$/nas/bin/nas\_event -list -f UFS [Lists all events for UFS facility such as Quota warnings]

**\$ /nas/bin/nas\_event -list -f CFS** [Lists all events for CFS facility]

\$/nas/bin/nas\_event -list -f SVFS [lists Checkpoint Save Volume Events]

### # nas\_event -list -k panick

facility id description

BoxMonitor 300 Slot 00 panicked

BoxMonitor 301 Slot 01 panicked →output abridged---list of all panic events

\$/nas/bin/nas\_event -list -k quota [Lists all quota-related events]

\$nas\_event -l -a callhome [Lists all events that generate CallHomes]

\$/nas/bin/nas\_event -L /nas/site/event.cfg

**\$ /nas/bin/nas\_event -L -info** NAS Events that are loaded on the Control Station

\$/nas/bin/nas\_event -list -a mail [Lists all events for mail action]

**\$ nas\_event -l -a trap** [Viewing SNMP Trap Events]

**# nas\_event -list -a -info** [Lists all possible actions for Celerra]

action

udprpc

exec

terminate

trap

callhome

logfile

**# nas\_event -list -a <list above possible actions here to get output>**

### **OUTPUTTING ALL NAS 5.6\6.0 EVENTS TO FILE:**

```
$ export NAS_DB=/nas ; /nas/bin/nas_event -l -c -i \| awk '/^[*][0-9]+/{print $1}' \| xargs -n1 -i bash -c \ "export COMP={} ; /nas/bin/nas_event -l -c {} -i \| awk '/^[*][0-9]+/{print $1}' \| xargs -n1 -i /nas/bin/nas_event -l -c \$COMP -f \{\} -id" \| fgrep -v 'DEBUG(7)' >events.out
```

**Note:** Event categories for 5.6 are DART; CS\_CORE; XML\_API; & CS\_PLATFORM

### **NAS EVENT SEVERITY CODES:**

1 = Alert  
2 = Critical  
3 = Error  
4 = Warning  
5 = Notice  
6 = Info  
7 = Debug

### **SYS LOG EXAMPLES:**

Aug 27 11:42:51 2007 NaviEventMonitor:**3:3** Backend Event Number 0x904 Host OEM-2MYA58ZGAQT Storage Array APM00062101038 SPA Device Bus 1 Enclosure 0 Power A SoftwareRev 6.24.1 (5.0) BaseRev 2.19.0.500.5.040 Description VSC Shutdown/Removed

Aug 27 11:42:51 2007 NaviEventMonitor:**2:4** Backend Event Number 0xa07 Error Host OEM-2MYA58ZGAQT Storage Array APM00062101038 SPA Device Bus 1 Enclosure 0 Disk 1 SoftwareRev 6.24.1 (5.0) BaseRev 2.19.0.500.5.040 Description CRU Powered Down

**3:3 →Severity 3 Error, Event ID 3**

**2:4 →Severity 2 Critical, Event ID 4**

## **CELLERRA EVENTS WITH NAS 5.6:**

### **LIST OF COMPONENTS:**

**# nas\_event -list -c -info**

```
Id Component
1 DART
2 CS_CORE
5 XML_API
6 CS_PLATFORM
```

### **LIST OF FACILITIES WITHIN A COMPONENT:**

**# nas\_event -l -c DART**

```
DART(1)
|->Id Facility
 24 ADMIN
 26 CAM
 27 CFS
 36 DRIVERS
 40 FSTOOLS
 43 IP
 45 KERNEL
 51 NDMP
 52 NFS
 54 SECURITY
 56 SMB
 58 STORAGE
 64 UFS
 68 LOCK
 70 SVFS
 72 XLT
 73 NETLIB
 75 MGFS
 77 VRPL
 81 VC
 83 RCPD
 84 VMCAST
 86 CHAMII
 93 USRMAP
101 ACLUPD
102 FCP
108 REP
111 DPSVC
115 SECMAP
117 WINS
118 DNS
122 DBMS
144 PERFSTATS
146 CEPP
```

### **LIST OF EVENTS FOR A FACILITY WITHIN A COMPONENT:**

**# nas\_event -l -c DART -f DBMS**

```
DART(1)
```

|--&gt; DBMS(122)

BaseID Severity Brief\_Description

1 ALERT(1) The environment database of the VDM might be corrupted and a recovery procedure must take place  
 2 CRITICAL(2) Only \${freeblocks,3,%llu} free blocks in the root file system (fsid \${fsid,2,%u}) of the VDM \${vdm,8,%s}.  
 3 ALERT(1) The root file system (fsid \${fsid,2,%u}) of the VDM \${vdm,8,%s} is full. There are only \${freeblocks,3,%llu} free blocks.

**# nas\_event -l -c CS\_PLATFORM -f NaviEventMonitor**

CS\_PLATFORM(6)

|--&gt; NaviEventMonitor(138)

BaseID Severity Brief\_Description

1 INFO(6) CLARiiON event number 0x\${navi\_event\_str,8,%s} \${desc,8,%s}.  
 2 WARNING(4) CLARiiON event number 0x\${navi\_event\_str,8,%s} \${desc,8,%s}.  
 3 ERROR(3) CLARiiON event number 0x\${navi\_event\_str,8,%s} \${desc,8,%s}.  
 4 CRITICAL(2) CLARiiON event number 0x\${navi\_event\_str,8,%s} \${desc,8,%s}.  
 5 DEBUG(7) CLARiiON event number 0x\${navi\_event\_str,8,%s} \${desc,8,%s}.  
 10 ALERT(1) CLARiiON Event Monitor daemon terminated unexpectedly.  
 100 ALERT(1) Disk \${dname,8,%s} has been cache-compromised.  
 101 CRITICAL(2) Disk \${dname,8,%s} has been trespassed.  
 200 INFO(6) Device group \${name,8,%s} on \${BACKEND\_NAME,8,%s} is \${condition,8,%s}.  
 201 ERROR(3) \${BACKEND\_NAME,8,%s} \${name,8,%s} is \${condition,8,%s}.  
 202 ERROR(3) Device group \${name,8,%s} on \${BACKEND\_NAME,8,%s} is \${condition,8,%s}.  
 203 WARNING(4) Device group \${name,8,%s} on \${BACKEND\_NAME,8,%s} is \${condition,8,%s}.

**LIST OF POSSIBLE ACTIONS FOR EVENTS:****# nas\_event -list -a -info**

action

terminate

trap

exec

callhome

logfile

**LIST OF EVENTS FOR A PARTICULAR ACTION:****# nas\_event -list -a callhome|head**

DART(1)

|--&gt; CFS(27)

BaseID Severity Brief\_Description

4 WARNING(4) File system size threshold (\${size,5,%u}%) was crossed for (fs \${fsname,68,%s})

DART(1)

|--&gt; DRIVERS(36)

BaseID Severity Brief\_Description

6 ERROR(3) SlicNicDevice: TOE \${deviceName,45,%s} panic detected. The TOE core image is being saved to \${filename,8,%s}'.

**# nas\_event -list -a callhome →CS\_PLATFORM for facility NaviEventMonitor BaseIDs**

CS\_PLATFORM(6)

|--&gt; NaviEventMonitor(138)

BaseID Severity Brief\_Description

4 CRITICAL(2) CLARiiON event number 0x\${navi\_event\_str,8,%s} \${desc,8,%s}.

100 ALERT(1) Disk \${dname,8,%s} has been cache-compromised.

**# nas\_event -list -a trap|head**

DART(1)

|--&gt; FSTOOLS(40)

BaseID Severity Brief\_Description

1 CRITICAL(2) The file system utility, fsck, started on file system id: \${id,22,%u} for type: (\${type,8,%s}).

2 INFO(6) FsId: \${id,5,%u} Fsck Started (\${type,8,%s}).

3 CRITICAL(2) FsId: \${id,5,%u} Fsck Succeeded

4 INFO(6) FsId: \${id,5,%u} Fsck Succeeded

5 CRITICAL(2) FsId: \${fsid,5,%u} Fsck Failed

6 INFO(6) FsId: \${fsid,5,%u} Fsck Failed

**NOTIFICATION ACTIONS CS CAN MAKE FOR EVENTS:**

Logging event to file

Sending email message  
Generating SNMP trap  
CallHome  
Running a script  
Sending RPC message to Host  
Terminating processing of an event

## **FACILITIES WITHIN NAS 5.5:**

# **nas\_event -l -f -i**

```
id facility
140 UPSMonitor
139 LocalHardwareMonitor
138 NaviEventMonitor
137 NASDB
135 JServer
131 BoxMonitor
129 MasterControl
111 DPSVC
108 REP
102 FCP
101 ACLUPD
96 DHSM
93 USRMAP
91 SNAPSURE_SCHED
86 CHAMIIENCMON
84 VMCAST
83 RCPD
81 VC
77 VRPL
75 MGFS
70 SVFS
64 UFS
58 STORAGE
54 SECURITY
52 NFS
51 NDMP
46 LIB
40 FSTOOLS
36 DRIVERS
27 CFS
26 CAM
24 ADMIN
```

## **NAS EVENTS WITHIN EACH FACILITY 5.5:**

# **nas\_event -list -f VC**

```
id description
0 Normal state
1 High water mark reached
2 Low water mark reached
3 No Virus Checker server available
4 No Virus Checker server; stop CIFS
5 No Virus Checker server; stop virus checking
6 File not checked
7 Virus Checker server online
8 Virus Checker server offline
9 Filesystem scan started
10 Filesystem scan terminated
11 Filesystem scan aborted
12 File has been deleted by the Virus Checker engine
```

13 File has been renamed by the Virus Checker engine  
14 File has been modified (cleaned) by the Virus Checker engine  
15 Virus Checker stopped  
16 Virus Checker started

**# nas\_event -list -f UFS**

id description  
0 I18N OK  
1 I18N - Out of Inodes  
2 I18N - Cannot balloc  
3 I18N - Recovery failed, filesystem not mounted  
4 Soft quota (warning limit) crossed  
5 Hard quota limit reached/exceeded  
6 Directory translate status message  
7 Skip auto fsck for corrupted filesystem at mount  
8 Gid map file is not present  
9 Gid map file is full  
10 Directory translate major error  
11 Crossed the root filesystem inode threshold  
12 Crossed the filesystem inode threshold  
13 Dropped below the root filesystem inode threshold  
14 Dropped below the filesystem inode threshold  
15 Filesystem size incorrect in superblock

**# /nas/bin/nas\_event -l -f VMCAST**

id description  
1 Filesystem Copy over IP done  
2 Filesystem Copy over IP failed  
3 Volume Copy over IP done  
4 Volume Copy over IP failed  
5 Filesystem Copy over IP interrupted, unmount  
6 Filesystem Copy clean up done  
7 Filesystem Copy convert failed  
8 Filesystem Copy clean up failed  
9 Resync Copy failed. Full Copy started

**INTERPRETING NAS EVENT FACILITY, SEVERITY, & EVENT ID FROM SYS LOG:**

**# nas\_event -l -f CAM** →All Event IDs for the Facility CAM

id description  
0 CAM OK  
1 CAM I/O path has failed  
2 CAM I/O path is inaccessible  
3 CAM command timeout  
4 CAM hung I/O  
5 CAM internal error or bad request  
6 CAM I/O path is again normal  
7 CAM I/O error

**Event Severity Levels for Celerra:** Lower the number, the greater the Severity

0 Emergency  
1 Alert  
2 Critical  
3 Error  
4 Warning  
5 Notice  
6 Info  
7 Debug

**Sys log examples:**

Nov 3 15:38:14 2006 **CAM:3:1** Slot 2: 1162586223: WARNING: Scsi-48 hba FAILED! →Facility CAM, Severity 3, EventID 1  
**(IO path failure)**

Nov 3 15:53:14 2006 **CAM:3:7** Slot 2: 1162587155: I/O Error: c16t115 Irp 0xde185904 CamStatus 0x84 SessStatus 0x02 Sense 0x03/0x11/0x00 →Facility CAM, Severity 3, EventID 7 (IO Error)

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
Nov 1 12:31:36 2006 **CAM:3:2** Slot 2: 1162402259: Warning: Scsi-16 hba is INACCESSIBLE. →Facility CAM, Severity 3,

**EventID 2 (Inaccessible path)**

**3:1 →Severity 3 Error, Event ID 1**

**3:7 →Severity 3 Error, Event ID 7, etc.**

## **DECIPHERING NAS EVENTS FROM THE MASTER EVENT FILE:**

**Note:** Typically, each NAS Event generation requires a facilitypolicy line and a disposition line

**# cat /nas/sys/nas\_eventlog.cfg**

# CAM

#

**facilitypolicy 26, 4 →Facility 26 is the ID assigned to CAM, as seen in nas\_event -l -f -i, 4 is the Event Severity**

disposition range=0-7 severity=0-4, logfile "/nas/log/sys\_log"

disposition **range=7-7 severity=0-3, callhome materials** →range=list of CAM events, 0-3=severity level, and ‘callhome materials’ is the action to take—in this case, only (1) CAM Event generates CallHome events for Event Severity 0-3, which is a CAM I/O Error

## **POSTEVENTS:**

**\$/nas/sbin/postevent**

postevent v0.1

usage: postevent <-f ev\_facility> <-s ev\_severity> <-i ev\_ID> [ev\_text]

Posts an ev\_ID with ev\_text to the nas\_eventlog facility as if it came from ev\_facility with ev\_severity. The nas\_eventlog will dispose of the event based on instructions in the file nas\_eventlog.cfg

-f Facility under which to log the event (e.g. Box Monitor = 131).

-i Event ID.

-s Severity to assign this event.

- 1 - Alert
- 2 - Critical
- 3 - Error
- 4 - Warning
- 5 - Notice
- 6 - Info
- 7 - Debug

## **EXAMPLE OF POSTING AN EVENT TO SYS LOG:**

**\$/nas/sbin/postevent -f 84 -s 4 -i 1 "Slot 2: 1099502624: Group:77\_CK2000433012710000 FSID:76 noconvert Copy done (t:11120531190117096)"**

**Note:** Obtain “1099502624” from date command: \$date +%

## **TESTING DEFINED CELERRA EVENTS & CALLHOMES:**

**# .server\_config server\_2 "logtest event facility=CAM severity=LOG\_EMERG eventid=7 message=toast"**

**Note:** If successful, will output the event in both sys\_log and server\_log, and if the event qualifies as a CallHome event [# nas\_event -list -a callhome], then the system will generate a CallHome.

Sys\_log → Nov 15 16:00:36 2006 CAM:0:1 Slot 2: 1163624407: toast

**# .server\_config server\_2 "logtest server facility=CAM severity=LOG\_EMERG eventid=7 message=toast"**

**Note:** If successful, this outputs only to the Server Log

2006-11-15 16:00:33: ADMIN: 4: Command succeeded: logtest server facility=CAM severity=LOG\_EMERG eventid=7 message=toast

## **POSSIBLE ACTIONS THAT CAN BE TAKEN:**

**# nas\_event -l -a -i**

action

udprpc

exec

terminate

trap

callhome

mail  
logfile

## **CELERRA NAS EVENTS: /nas/sys/nas\_eventlog.cfg [Caution: Do not edit this file!]**

DataMover & Control Station Internal operating events are logged in this file. Internal events contain the name of the Celerra facility, the "High Water mark" severity of the event [1-7], the Event ID number, short description of event, and System action being taken.

## **CONFIGURING ADDITIONAL CELERRA EVENTS:**

Celerra supports additional custom-defined Events. Events are added by creating a new event.cfg file and loading it to the existing event files. For example, an event.cfg file contains a line for "facility policy" that describes facility ID & severity level. Also contains lines for "disposition" describing Event ID ranges & actions to be taken.

### **EXAMPLE:**

Step 1. Create "event.cfg" file using appropriate format [see other examples below]

Step 2. Place this file in the /nas/site directory

**\$/nas/bin/nas\_event -L /nas/site/event.cfg**

Step 4. Verify that new Event file has been added: **\$/nas/bin/nas\_event -L -info**

### **Loaded config. files:**

1: /nas/sys/nas\_eventlog.cfg

2: /nas/http/webui/etc/web\_client\_eventlog.cfg

**Note:** #nas\_event -L -info reads file /nas/site/nas\_eventlog.cfg

Step 5. \$cat /nas/sys/nas\_eventlog.cfg [should see reference to new 'event.cfg' file]

Step 6. To Unload an Event File: **\$/nas/bin/nas\_event -Unload /nas/site/event.cfg**

## **CELERRA FAILS TO CALL HOME OR LOG ENTRIES IN SYS LOG:**

**PROBLEM:** nas\_eventlog.cfg file is no longer "loaded" or is corrupted. Without a properly loaded nas\_eventlog.cfg file, Celerra cannot post events to the sys\_log, which is used to trigger CallHome events.

**sys log:** no current entries being logged at all

### **/var/log/messages**

Feb 27 00:01:53 emc\_cs02 EMCServer: ReportEvent: daemon "Event Log" unexpectedly exited (status = 256)

Feb 27 00:02:26 emc\_cs02 EMCServer: ReportEvent: "Event Log" respawning too fast: disabled for 5 minutes

Feb 27 00:05:20 emc\_cs02 get\_datamover\_status: Data Mover server\_3 is not responding "messages" [readonly] 8216L, 855559C

### **Nas Event Log shows as not loaded:**

**\$ nas\_event -L -i**

Loaded config. files:

1: /nas/http/webui/etc/web\_client\_eventlog.cfg

### **Nas Event Log with corrupted entries:**

# cat nas\_eventlog.cfg (Shows corrupted entries)

1:export "/" anon=0 access=192.168.1.100:192.168.2.100:192.168.1.101:192.168.22:

/nas/http/webui/etc/web\_client\_eventlog.cfg::

## **RELOADING NAS EVENTLOG.CFG FILE:**

### **1. Unload all NAS Events that are listed in output of #nas\_event -L -i:**

**# nas\_event -U /nas/sys/nas\_eventlog.cfg**

Config file (/nas/sys/nas\_eventlog.cfg) not loaded

# nas\_event -U /nas/site/cwm\_notify.cfg

Config file (/nas/site/cwm\_notify.cfg) not loaded

# nas\_event -U /nas/http/webui/etc/web\_client\_eventlog.cfg

Config file (/nas/http/webui/etc/web\_client\_eventlog.cfg) not loaded

### **2. Verify that all files are unloaded:**

**# nas\_event -L -i**

Loaded config. files: <none>

### **3. Stop NAS Services:**

**4. Manually remount: /nas /nbsnas /nas/dos /nas/var**

**5. Edit the /nas/site/nas\_eventlog.cfg to add the correct entries back to the file:**

**# vi /nas/site/nas\_eventlog.cfg**

1:/nas/sys/nas\_eventlog.cfg:

2:/nas/http/webui/etc/web\_client\_eventlog.cfg:

3:/nas/site/cwm\_notify.cfg:

**6. Restart NAS Services & verify that nas\_mcd & box monitor restart**

**7. Verify NAS Event processes:**

**# nas\_event -L -i**

Loaded config. files:

- 1: /nas/sys/nas\_eventlog.cfg
- 2: /nas/http/webui/etc/web\_client\_eventlog.cfg
- 3: /nas/site/cwm\_notify.cfg

**\$ ps -eafl |grep -i nas\_event**

```
/nas/sbin/nas_eventcollector -I 10
040 S root 11054 10991 0 69 0 - 7728 nanosl Mar02 ? 00:00:00 -->Number of nas_eventcollector processes running
/nas/sbin/nas_eventcollector -I 10
100 S root 14136 10460 0 69 0 - 1453 do_pol Mar02 ? 00:00:00
/nas/sbin/nas_eventlog /nas nas_eventlog.cfg
040 S root 14138 14136 0 69 0 - 1453 do_pol Mar02 ? 00:00:00 -->Number of nas_eventlog.cfg processes running
/nas/sbin/nas_eventlog /nas nas_eventlog.cfg
140 S root 14141 14138 0 69 0 - 1453 rt_sig Mar02 ? 00:00:0
```

**CELERRA SNMP: /nas/sys/emccelerra.mib**

An "SNMP Community" is configured with the DataMover or Control Station as the "SNMP Agent" and "SNMP Managers" running on PC or Unix Workstations. SNMP Agent recognizes users by community name. Because SNMP communications are 'plain text', it's recommended that SNMP implementations be conducted on trusted networks. SNMP uses UDP ports 161/162.

**THREE TYPES OF SNMP MESSAGES:**

Get requests from SNMP manager for information from remote devices

Set request from SNMP manager to modify configuration of remote devices

Trap messages from SNMP agent located on remote devices for notification and monitoring of events

**TROUBLESHOOTING FILES & SNMPD SERVICE:****/nas/site/trap.config****/sbin/service --status-all |grep snmpd****# snmpwalk -Cc -c public server\_x****TRANSLATING EMCCELERRA MIB:****# snmptranslate -Tp -m ./emccelerra.mib**

```
+--emcCelerra#(0)
  |
  | +-+celReboot(1)
  | +-+celMasterCtlFault(2)
  | +-+celHWFFailure(3)
  | +-+celSlotStale(4)
  | +-+celSlotPanicked(5)
  | +-+celIntfFailure(6)
  | +-+celAAF(7)
  | +-+celCSStart(8)
  | +-+celJServer(9)
  | +-+celWebGUI(10)
  |
  +-+celEventTable(1)
    |
    +-+celEvent(1)
      | Index: celEventFacility, celEventID
      |
      +-+R-- INTEGER celEventFacility(1)
      +-+R-- INTEGER celEventID(2)
      +-+R-- INTEGER celEventSeverity(3)
      +-+R-- String celEventDescr(4)
      Textual Convention: DisplayString
```

**CONTROL STATION EVENTS:** The Control Station SNMP Agent sends "traps" back to Management Station when predefined "events" occur. These events are logged in an "event log" in /nas/log/sys\_log. Additional "traps" can be defined in a customized version of the event log via a "trap.cfg" file, which supplements, but does not replace, the existing "nas\_eventlog.cfg" file.

**CONFIGURING SNMP EVENTS ON CONTROL STATION:****CreatingTraps and Eventlog Configuration Files**

Step 1. Set SNMP Manager: #vi /nas/site/trap.cfg and add following:

“snmpmanager 193.1.21.210 ; communityname comics”

Step 2. Set DNS configuration for Email: #vi /etc/resolv.conf

**nameserver 192.168.1.200**

**search example.com**

Step 3. Create Event Log & Assign Policy Statements: #vi /nas/site/new\_eventlog.cfg

**Caution:** Do not modify the original /nas/site/nas\_eventlog.cfg file

Step 4. Load Eventlog File : #nas\_event -L /nas/site/new\_eventlog.cfg

## **CONFIGURING CELERRA FOR SNMP TRAPS FOR FILESYSTEM ISSUES:**

1. Set the FileSystem threshold on the Data Mover by adding the following to /nas/server/slot\_x/param and rebooting the Server:

**param file fsSizeThreshold=80** [notify at 80% full]

2. Create a new config file called /nas/site/fs\_eventlog.cfg that contains the following:

**# vi /nas/site/fs\_eventlog.cfg**

#CFS High Water Mark Event Control

facilitypolicy 27, 7

disposition range=1-1, mail "root@Celerra\_CS0"

disposition range=1-1, trap "/nas/site/trap.cfg 11"

**Note :** 27 = CFS facility, 7 = Event severity code, 1-1 means events 1 thru 1, equating to ‘crossed fs threshold’, 11 is the newly assigned trap number from the emccelerra.mib file, and “root@Celerra\_CS0” is an example of an email address for notification.

3. Edit the /etc/resolv.conf file so as to setup the CS DNS Client and allow for email notification:

domain mydomain.com

search localdomain

nameserver 192.10.20.10

4. Edit /nas/site/trap.cfg so that it contains the following configuration setup:

snmpmanager 192.168.1.100; communityname public

**Note:** IP address of SNMP manager that will receive traps; “public” = community name used for authenticating with SNMP Mgr

5. Modify the Celerra MIB by editing the /nas/sys/emccelerra.mib file and adding to the bottom of the file with the following syntax, first checking to see what the next available trap number is--in this case, 11 is the next trap number to be assigned:

celCFS TRAP-TYPE

ENTERPRISE emcCelerra

VARIABLES { celEvent }

DESCRIPTION

“Trap message will be sent when a file system exceeds the threshold.”

::= 11

6. Using HP OpenView, or some other SNMP Management Program, recompile the Celerra MIB by choosing “Options -> Load/Unload MIBs:SNMP”

7. Ensure Control Station Hostname and email Domain name are correctly registered in the /etc/hosts file:

# Do not remove the following line, or various programs

# that require network functionality will fail.

127.0.0.1 localhost.localdomain localhost

172.24.96.82 Celerra\_CS0 Celerra\_CS0.mydomain.com

**Note:** mydomain.com is the domain that has the mail server--substitute with your domain information

8. Load the new config file:

**# /nas/bin/nas\_event -L /nas/site/fs\_eventlog.cfg**

9. Verify:

**# nas\_event -L -i**

**# nas\_event -l -a mail**

**# nas\_event -l -a trap**

10. Test the new Trap:

**# /nas/sbin/nas\_snmptrap /nas/site/trap.cfg -m /nas/sys/emccelerra.mib -f 10 -f 2 -i 1 -s 4 -d**

**# /nas/sbin/nas\_snmptrap /nas/site/public.cfg -m /nas/sys/emccelerra.mib -r 10 -f 131 -i 8 -s "test trap fan 1 error"**

11.Add cron tab for snmp\_trap\_full\_fs

#crontab -l >nas.cron

#vi nas.cron

5,10,15,20,25\*\*\*\* /home/nasadmin/SNMP/snmp\_trap\_full\_fs 1>/dev/null 2>&1

30,35,40,45,50,55\*\*\*\* /home/nasadmin/SNMP/snmp\_trap\_full\_fs 1>/dev/null 2>&1

## CONFIGURING TRAP FOR FILESYSTEM USAGE THRESHOLD:

- ```
1. Specifying recipient of snmp Trap: #cat /nas/site/trap.cfg  
snmpmanager localhost; communityname public  
2. Start Trap Daemon on CS0: #!/usr/sbin/snmptrapd -P &  
3. Shell Script for >than 90% Full FS:  
#!/bin/sh  
SERVER=server_2  
FILESYSTEM=fs02  
HWM=90  
CAP1=`/nas/bin/server_df $SERVER | egrep $FILESYSTEM | awk '{print $5}'`  
CAP2=`echo $CAP1 | sed 's/%//'`  
MESSAGE="Filesystem $FILESYSTEM on $SERVER exceeds $HWM, totals $CAP1"  
if [ $CAP2 -ge $HWM ]; then  
/nas/sbin/nas_snmptrap /nas/site/trap.cfg -m /nas/sys/emapcelerra.mib -r 4444 -i 5555 -s 6666 -d capacity_of_fs02_is_over_$CAP1  
4. When triggered, Filesystem event will be sent to snmptrapd daemon
```

## **VIEWING SNMP TRAP EVENTS:**

**\$ nas event -l -a trap**

Trap messages consist of the Name of the Celerra facility for the 'event'; the High Water Mark event severity, from 1-7; the Event ID number; Description of Event; and System-defined action to be taken [defined in \$NAS\_DB/site/trap.cfg]

**SNMP DIRECTORIES:** /nas/site      /nas/sys

/nas/site [trap.cfg and eventlog.cfg files] /nas/sys [trap.cfg; nas eventlist; nas eventlog.cfg--events; emccelerra.mib--MIBs]

## **ISSUING TEST SNMP TRAPS:**

```
#/nas/sbin/nas_snmptrap /nas/site/2031806010public.cfg -m /nas/sys/emccelerra.mib -r 10 -f 131 -i 8  
-s "emc test trap for fan 1 error"
```

#### **INFORMATION FOR 5.5 and 5.6 WHEN TESTING SNMP TRAPS AB109619:**

1. Kill the default UPS snmptrap service  
**# killall snmptrapd**
  2. Perform SNMP Trap testing
  3. Reboot CS after testing to restore UPS snmptrap service, or issue following:  
**# /usr/sbin/snmptrapd -c /nas/sys/snmptrapd.conf -p 162 -u /var/run/snmptrapd.pid >/dev/null 2>&1 &**

#### **CELLERRA SNMP: Only MIB II supported**

## SNMP CONFIGURATION:

**\$ server snmp server 4 -community**

server 2 : public [Default com

**SNMP CONTACT NAME:**

---

```
$server_snmp server_4 -syscom
```

## **SNMP AGENT LOCATION:**

\$Server\_snmp server\_4 -location location name

## **SNMP EVENTS & TRAPS:** /nas/site/trap.cfg

Two ‘facilities’ issues traps: **Box Monitor 131** & **Master Control Daemon 1**

Other facilities also exist, such as Call Home, D

- SEVEN PREDEFINED CELERRA TRAPS**

  - 1. DataMover Reboot
  - 2. Missing Control Station heartbeat/serious fault
  - 3. Celerra Hardware failure
  - 4. Stale Reason Code
  - 5. DataMover panic
  - 6. Failure to ping both DM internal interfaces
  - 7. Automated Archive Facility Error

**Note:** In the case that one of these (7) predefined "events" occur, a trap message will be sent that includes the Facility Name, Event ID, Severity, and Text description of event.

## **CELERRA EVENTLOG CONFIG FILE:**

**/nas/sys/nas\_eventlog.cfg** [Celerra "binding" file that correlates internal events to the predefined system traps]

**Warning:** Never Edit this file if adding new Events to Trap as this file controls Celerra CallHome Events!!!]

## **CHECKING FACILITIES AND TRAPS:**

### **LISTING SNMP FACILITIES:**

**\$nas\_event -l -f -i**

**\$nas\_event -l -f UFS**

### **USE FOLLOWING KEYWORD TO FIND EVENTS FOR QUOTA, FILESYSTEM, etc.:**

**\$nas\_event -l -k quota**

```
facility id description
USRMAP 7 Usermapper filesystem quota exceeded
UFS 4 Soft quota (warning limit) crossed
UFS 5 Hard quota limit reached/exceeded
```

**# nas\_event -list -k filesystem**

```
facility id description
USRMAP 7 Usermapper filesystem quota exceeded
SVFS 3 Source filesystem restore done
UFS 3 I18N - Recovery failed, filesystem not mounted
UFS 7 Skip auto fsck for corrupted filesystem at mount
UFS 11 Crossed the root filesystem inode threshold
UFS 12 Crossed the filesystem inode threshold
UFS 13 Dropped below the root filesystem inode threshold
UFS 14 Dropped below the filesystem inode threshold
NDMP 1 Create a checkpoint for the backup filesystem and mount it
NDMP 2 Unmount the checkpoint for the backup filesystem and delete it
CFS 1 Crossed the filesystem size threshold
CFS 3 Dropped below the filesystem size threshold
CFS 4 Crossed the root filesystem size threshold
CFS 5 Dropped below the root filesystem size threshold
```

### **LISTING SNMP TRAPS:**

**\$nas\_event -l -a trap**

### **DETERMINING WHAT EVENT.CFG FILES ARE LOADED ON CS0:**

**#nas\_event -L -info**

Loaded config. files:

- 1: /nas/sys/nas\_eventlog.cfg
- 2: /nas/http/webui/etc/web\_client\_eventlog.cfg
- 3: /nas/site/mail\_events.cfg

### **Seven Default System Actions:**

|           |                               |                        |
|-----------|-------------------------------|------------------------|
| SNMP Trap | Sent to SNMP Manager          | \$NAS_DB/site/trap.cfg |
| UDPRPC    | RPC message to specified host |                        |
| Exec      | Execute a procedure           |                        |
| Terminate | Shutdown process              |                        |
| Callhome  | Call Home to EMC              |                        |
| Mail      | Send email                    |                        |
| LogFile   | Log an entry                  |                        |

**Action Listings:** #nas\_event -l -a -i      #nas\_event -l -a callhome      #nas\_event -l -f BoxMonitor [MasterControl]

### **SAMPLE TRAP.CFG FILE:**

```
#snmp trap configuration file
#snmpmanager 128.154.11.20 ; communityname public
#snmpmanager host1 ; communityname public
```

### **SAMPLE NASEVENTLOG.CFG FILE:**

```
# Test configuration file
#
defaultpolicy 7
```

```
disposition , logfile "/nas/log/sys_log"
facilitypolicy 0x10, 3
    disposition range=1-100 severity=2-2, mail "user@host"
    disposition range=50-100 threshold=6 resetafter=20, logfile "/nas/lo
g/sys_log"
        disposition range=5-5, callhome immediate
        disposition range=5-5, trap "/nas/site/trap.cfg 2"
facilitypolicy 17, 4
    disposition range=1-5, terminate
    disposition range=6-100 threshold=4 rearm=2, exec "/usr/bin/xxx"
    disposition range=80-99, udprpc "1,706001,1@remote.domain.com"
#
# For now all Dart log messages use facility 128
#
facilitypolicy 128, 4
    disposition , logfile "/nas/log/sys_log"
        disposition , mail "user@remote"
        disposition , udprpc "1,706001,1@remote"
#
# Master Control
#
facilitypolicy 129, 4
    disposition range=0-1000, logfile "/nas/log/sys_log"
    disposition range=6-10, callhome immediate
    disposition range=6-7, trap "/nas/site/trap.cfg 2"
#
# BoxMonitor
#
facilitypolicy 131, 7
    disposition range=0-1000 , logfile "/nas/log/sys_log"
    disposition range=1-1, callhome immediate
    disposition range=1-1, trap "/nas/site/trap.cfg 2"
    disposition range=3-15, callhome immediate
-----abridged-----
```

## **CREATING ADDITIONAL CELERRA EVENTS:**

Step 1. Copy the /nas/sys/nas\_eventlog.cfg to /nas/site/new\_eventlog.cfg

Step 2. Edit file to add new traps but leave facilitypolicy and disposition

Step 3. Load new file: **\$/nas\_event -L /nas/site/new\_eventlog.cfg**

## **LINUX SNMPWALK:** [Other Unix hosts use 'getmib']

/usr/bin/snmpwalk -V UCD-snmp version: 4.1.1 -H config files -h help files  
SNMP GET NEXTS requests are used to query community tree for information.

## **OUTPUTTING DATA MOVER OIDs:**

```
$ /usr/bin/snmpwalk -Cc server_2 public
$ /usr/bin/snmpwalk -Cc server_2 public at
$ /usr/bin/snmpwalk -Cc server_2 public tcp
$ /usr/bin/snmpwalk -Cc server_2 public ip
# snmpwalk -v 1 -c public server_2 system
```

SNMPv2-MIB::sysDescr.0 = STRING: Product: EMC Celerra File Server Project: SNAS Version: T5.6.37.5

SNMPv2-MIB::sysObjectID.0 = OID: SNMPv2-SMI::enterprises.1139

SNMPv2-MIB::sysUpTime.0 = Timeticks: (172213913) 19 days, 22:22:19.13

SNMPv2-MIB::sysContact.0 = STRING: nasadmin

SNMPv2-MIB::sysName.0 = STRING: server\_2

SNMPv2-MIB::sysLocation.0 = STRING: here

SNMPv2-MIB::sysServices.0 = INTEGER: 72

## **USING SNMPWALK ON LINUX:**

```
# /usr/bin/snmpwalk -Cc -c public server_5 |grep ifMtu
```

interfaces.ifTable.ifEntry.ifMtu.1 = 9000

```
# /usr/bin/snmpwalk -t 10 192.168.10.10 public [Where 10 is timeout value of command]
```

**OTHER DIRECTORIES:** /usr/share/snmp/snmpd.conf lsnmpd.local.conf | mibs [list of all mibs]

## **SNMPD DAEMON:**

The snmpd daemon binds to a port, receives SNMP Management request, collects requested information, performs requested operation, and returns to requestor.

## **SNMP TRAP DAEMON:**

**Note:** Works as an SNMP application that receives and logs snmp trap messages that are sent to ‘SNMP-TRAP’ Port 162

### **STARTING SNMPTRAPD:**

**#/usr/sbin/snmptrapd -P &**

```
#ps -ef |grep -i snmp
root 1396 1 0 Dec24 ? 00:00:00 /usr/sbin/snmpd -s -l /dev/null
root 15187 15149 0 18:26 pts/0 00:00:00 /usr/sbin/snmptrapd -P
```

### **TRAPPING CONTROL STATION SNMP EVENTS TO A SPECIFIC LOG or MESSAGES FILE:**

**# /usr/sbin/snmptrapd -P >/home/nasadmin/snmp.log 2>&1 &**

**# /usr/sbin/snmptrapd -P -o custom\_events\_messages &**

### **DEFAULT SNMP PORT FOR AGENT:** Port=161 for Agent

**Verify Port: # netstat -a -p |grep snmp**

**Comment:** SNMP is one of the basis of many Celerra Commands, such as "nfsstat", "netstat", etc

### **TURNING ON SNMP:**

**#/nas/sbin/chkconfig snmpd on** [Configures SNMP daemon to start on bootup]

**#/etc/rc.d/init.d/snmpd start** [or #/sbin/service snmpd start]

### **CONFIGURING NAS EVENTS ON LINUX:**

1. Edit the /nas/site/trap.cfg file create an entry similar to the one below:

**snmpmanager 192.168.4.253 ; communityname public**

2. Turn on the SNMP daemon on the control station (as root), the daemon will start by default at boot up.

**/sbin/chkconfig snmp on**

3. Verify which snmp traps are set for nas\_events, (most events that would be desirable are set by default)

**nas\_event -l -a trap**

4. Use the following command (as root) to start the snmpd:

**/etc/rc.d/init.d/snmpd start**

**NOTE:** It is possible to use the control station to verify operation by setting the control station to be the snmp manager in the trap.cfg file, and then starting another process to catch the traps. By issuing the following command it is possible to log all snmp activity to a log file, provided the control station is identified in the trap.cfg file as the SNMP manager. This should only be used to verify functionality.    /usr/sbin/snmptrapd -P >/home/nasadmin/snmp.log 2>&1 &

### **NAS EVENT LOGGING PROCESSES:** NAS 5.2

NAS\_MCD controls the NAS Event Collector and Logging mechanism that generates messages to sys\_log and nas\_eventlog:

**/nas/sys/nas\_mcd.cfg**

**daemon "Event Log"**

```
executable  "/nas/sbin/nas_eventlog"
optional    no
autorestart yes
ioaccess   no
cmdline    "/nas/nas_eventlog.cfg"
eventlogger yes
```

**daemon "Event Collector"**

```
executable  "/nas/sbin/nas_eventcollector"
optional    no
autorestart yes
ioaccess   no
cmdline    "-I 10"
```

**#/nas/sbin/nas\_eventcollector** runs as a daemon that catches events and posts to /nas/log/sys\_log. There is a problem where a cloned eventcollector process may die without the parent process recognizing a problem, and leads to zombie status of the eventcollector. Proper operation is that parent nas\_eventcollector should recognize a problem and then kill off all processes, at which point the NAS\_MCD will restart the processes cleanly. See AR47104. End result of this problem is that events may not get logged, and processes such as Checkpoint SavVol extensions, may fail as well.

**# ps -axuww |grep nas\_event**

```
root 1833 0.0 0.1 5924 892 ? S Dec22 0:00 /nas/sbin/nas_eventlog /nas nas_eventlog.cfg
root 1834 0.0 0.1 5924 892 ? S Dec22 0:00 /nas/sbin/nas_eventlog /nas nas_eventlog.cfg
root 1835 0.0 0.1 5924 892 ? S Dec22 0:39 /nas/sbin/nas_eventlog /nas nas_eventlog.cfg
root 1836 0.0 0.1 5924 892 ? S Dec22 0:00 /nas/sbin/nas_eventlog /nas nas_eventlog.cfg
root 1838 0.0 0.1 31948 936 ? S Dec22 0:00 /nas/sbin/nas_eventcollector -I 10
root 1841 0.0 0.1 31948 936 ? S Dec22 0:00 /nas/sbin/nas_eventcollector -I 10
# ls -la /proc/1833
lrwxrwxrwx 1 root root 0 Dec 30 11:07 exe -> /nas/sbin/nas_eventlog
```

## **CONFIGURING SNMP MAIL EVENTS ON CELERRA:**

Step 1. Create Map file called “mail\_events.cfg” and place in /nas/site directory

### **Contents of Mail Events Map File:**

```
#  
facilitypolicy 129, 4  
    disposition range=6-7, mail "nasadmin@sun1"  
# BoxMonitor  
#  
facilitypolicy 131, 7  
    disposition range=1-1, mail "nasadmin@sun1"  
    disposition range=3-15, mail "nasadmin@sun1"  
    disposition range=18-21, mail "nasadmin@sun1"  
    disposition range=38-39, mail "nasadmin@sun1"  
    disposition range=43-43, mail "nasadmin@sun1"  
    disposition range=57-58, mail "nasadmin@sun1"  
    disposition range=60-60, mail "nasadmin@sun1"  
    disposition range=100-315, mail "nasadmin@sun1"
```

Step 2. Load Map file to control station:

**\$/nas/bin/nas\_event -L /nas/site/mail\_events.cfg**

EventLog: will load /nas/site/mail\_events.cfg...done

Step 3. The following mail message was received when a test datamover panic was done:

**\$ mail**

```
From root@jessica Mon Mar 20 16:57:22 2000
From: root@jessica
To: nasadmin@sun1
Date: Mon, 20 Mar 2000 21:58 GMT
Subject: EMCServer Event
Message-ID: <38d69f220.e64@jessica>
Content-Length: 58
Mar 20 16:58:57 2000 BoxMonitor:2:302 Slot 2 has panicked
```

Step 4. Issue the following command to determine what events emails will be sent for:

**#nas\_event -l -a mail**

```
facility id description
AAF 4 Not much tapes left available
MasterControl 6 unexpected daemon exit
MasterControl 7 Control Station heartbeat missing
BoxMonitor 1 EPP did not initialize
```

Step 5. To unload the current email events configuration so that changes can be made:

**#/nas/bin/nas\_event -U /nas/site/mail\_events.cfg**

## **MODIFYING CELERRA MIBs FOR NEW TRAP EVENTS:**

1. #vi /nas/sys/emccelerra.mib & add new Trap Event at bottom of file, using the next available Trap number:

```
celCFS TRAP-TYPE
ENTERPRISE emcCelerra
VARIABLES { celEvent }
DESCRIPTION
    "Trap message will be sent when a file system exceeds the threshold."
 ::= 11
```

**Note:** Once modified, the “emccelerra.mib” file needs to be recompiled into the SNMP Manager Program

## **TROUBLESHOOTING SNMP/CELERRA NAS EVENTS:**

1. Is Event File loaded? \$nas\_event -L -info
  2. View Server Log. If Event not recorded, event is not being issued.
  3. View Eventlog.cfg to ensure that Facility, Severity, Disposition Range, Severity Range, Trap numbers are setup correctly
  4. Verify IP Address & Community name of trap.cfg and ping IP Address
  5. Conduct Mail Test on CS0: #mail -s test "joeblow@emc.com"
  6. Manually initiate trap: #/nas/sbin/nas\_snmptrap /nas/site/trap.cfg -m /nas/sys/emccelerra.mib -r 8 -d test
- Note:** Would need to check on SNMP Manager (HP OpenView, etc) to see if Trap is received
7. Manually initiate Trap to CS0: #/usr/sbin/snmptrapd -P
- Note:** Requires IP address of CS0 in Trap.cfg file and should output to screen
8. I18N Events are not yet supported
  9. SNMP Events are usually a 1-time notification, but with NAS 4.x and above, the following Events are now routinely duplicated:  
Checkpoint; File System Full; Soft/Hard Quota
  10. Following is an Example of Server Log Duplicate Event Entries being Logged:  
2002-11-09 23:11:29: LIB: 6: last message repeated 1 times

## **MAX SIZE OF SNMP QUERY FOR DART IS 494 BYTES:**

If data mover receives SNMP queries > than 494 bytes in size, following message is logged:

LOG\_SYSLOG: 4: SNMP AGENT: Too big message – discarded

**Note:** NAS 5.0.9.1

## **SNMP INQUIRIES CREATE ANNOYING SYS LOG/SERVER LOG ENTRIES:**

Modprobe: Can't locate module block-major-22

**Solution:** Vi edit /etc/modules.conf and insert following entry for each offending module

**alias block-major-22 off**

## **DATA MOVER FAILOVER:**

### **DATA MOVER HARDWARE FAILOVER RULES:**

- Standby Server should contain same amount of memory and CPU as Primary [SnapSure, Celerra Replicator are memory hogs]
- 507 DM can only failover to another 507, 510, or 514, provided network interface rules are met
- 510 DM can only failover to another 510 or 514, provided network interface rules are met
- 514 DM can only failover to another 514, provided network interface rules are met

#### **Interface Rules:**

- Primary & Standby Servers must share a similar “ana”, “ace”, or “cge” interface in order to failover
- Can failover from copper ana interface to Standby with copper cge interface if cge is set to auto
- Copper cannot be used to failover to Fibre and vice versa

#### **Storage Rules:**

- Fibre SW can be replaced by Fibre SW or SCSI and vice versa
- Fibre AL can only failover to another Fibre AL system [NS600G]
- Standby Server must be able to probe and see the same devices that the Primary Server sees

#### **Tape Rules:**

- NDMP backups from a Standby Server are not supported
- 507 DM uses one FC-SW HBA for disks and a SCSI HBA for tape backup [not to be used for Disk storage]
- 507 DM does not support dual FC-SW HBA's
- 510 & 514 with dual FC-SW HBA's, the 2<sup>nd</sup> HBA can only be used for connecting to tape devices via Switch

#### **Common Error for this: "hardware components mismatch"**

- DataMover cannot Failover or Assume duties of a ‘failed’ server if Control Station is not operational
- Prior to “Failing Back”, ensure that the original server is up & running at Status 5 & not in a reset 0 state--reboot original server & check for Status 5

## **DATA MOVER STATUS & STANDBY SERVERS:**

### **CREATING STANDBY SERVER 3 for PRIMARY SERVER 2:**

**\$ server\_standby server\_2 -c mover=server\_3 -policy auto**

server\_2 : server\_3 is rebooting as standby

**# nas\_server -l**

|    |      |      |      |         |          |      |
|----|------|------|------|---------|----------|------|
| id | type | acl  | slot | groupID | state    | name |
| 1  | 1    | 1000 | 2    | 0       | server_2 |      |

2 4 1000 3 0 server\_3

**Note:** Type 1 indicates a Primary NAS Server, Type 4 indicates a Standby Server

**EXAMPLE: FAILING OVER PRIMARY (server 2) TO ITS STANDBY SERVER (server 3):**

**# server\_standby server\_2 -a mover**

server\_2 : operation in progress (not interruptible).....

server\_2 : going offline

server\_3 : going active

replace in progress ...done

failover activity complete

commit in progress (not interruptible)...done

server\_2 : renamed as server\_2.faulted.server\_3

server\_3 : renamed as server\_2

**# nas\_server -l**

id type acl slot groupID state name

1 4 1000 2 2 server\_2.faulted.server\_3

2 1 1000 3 0 server\_2

**Note:** nas\_server output shows if Data Movers are failed over to Standby server. In this example, the Primary Server\_2, who's normal slot is slot\_2, is “faulted” and failed over to its Standby server in slot\_3. Keep in mind that regular commands issued to the server will still produce the correct output, but that the Data Mover’s actual database is located in slot\_3 while failed over.

**DELETING STANDBY RELATIONSHIP & CREATING PRIMARY NAS SERVER:**

**# nas\_server -l**

id type acl slot groupID state name

1 1 1000 2 0 server\_2

2 4 1000 3 0 server\_3

**# server\_standby server\_2 -d mover** [removes server\_3’s Standby relationship from server\_2]

server\_2 : done

**# server\_setup server\_3 -type nas**

server\_3 : is rebooting as type nas

**Note:** You cannot use setup\_slot -init to convert a Standby Server to a Primary. Only server\_setup can do this.

**# nas\_server -l**

id type acl slot groupID state name

1 1 1000 2 0 server\_2

2 1 1000 3 0 server\_3

**\$ nas\_server -a -i** (shows data mover type, nas or standby; shows standby relationships; shows actual server status)

id = 1

name = server\_2

acl = 1000, owner=nasadmin, ID=201

type = nas

slot = 2

member\_of =

standby = server\_3, policy=auto

status :

defined = enabled

actual = online, active

id = 2

name = server\_3

acl = 1000, owner=nasadmin, ID=201

type = standby

slot = 3

member\_of =

standbyfor= server\_2

status :

defined = enabled

actual = online, ready

**\$ server\_standby ALL -v mover** (shows if standby relationship has been defined)

server\_2 : ok

server\_3 : ok

**# server\_standby ALL -v mover** (No standby relationships defined)

server\_2 :

Error 4003: server\_2 : standby is not configured

server\_3 :

Error 4003: server\_3 : standby is not configured

## **DATA MOVER FAILOVER MECHANISM:**

When data mover fails over to standby, the Standby Server inherits the MAC addresses, IP addresses, and any physical interface settings that are set, such as speed, duplex, link negotiation, etc. [Configure Switch ports of Standby with same settings as Primary]. After failing over to Standby, Standby sends one gratuitous ARP broadcast so that the Switch can learn the new location of the MAC address, thereby updating its CAM Tables with the new source MAC address, mod/port number, and VLAN number. This also serves to update client ARP tables on local VLANs.

## **CHANGING DATA MOVER FAILOVER TIMEOUT VALUE:**

# cat /nas/server/slot\_2/bmparms

ScanInterval:5: →Scan Interval is number of secs between checks by Box Monitor of DM progress during failover

StaleReasonThresh:49: →Number of times Box Monitor will use the ScanInterval to check failover progress before aborting

#FailedPingThresh:10:

#wait\_ready:162:

**Note:** Change “StaleReasonThresh” value to a max. of 120, which translates as  $120 * 5 = 600\text{secs} / 60 = 10\text{ minutes}$ . This value dictates on how long Control Station will wait for Data Mover to failover before aborting the failover. Please note that after any changes are made to the bmparms file, stop and restart Box Monitor to read-in the new changes. Also, setup\_slot and NAS Upgrades may change the file settings back to system defaults!

## **FAILOVER ACTIONS DURING SERVER PANICS:**

→Current code requires that panic dump completes before restarting services on failing over to Standby Server

→Cognac will allow failover to begin after panic handlers have written recovery data to disk, & compress dumps to disk

1. DART panic, panic handlers run, dump written

2. CS Box Monitor kicks off recover\_slot after panic, waits for handlers or dump complete, forks off to slot\_recover

3. Slot\_recover changes faulted DM boot.cfg to minconfig, feeds DM config to standby over internal network, changes target Standby boot.cfg to primary, reboots faulted slot after dump reported complete

## **NAS 5.0 & STANDBY DATAMOVERS:**

A Primary Datamover can now have more than (1) Standby Datamover to fail over to

\$server\_standby server\_2 -c mover=server\_4 -policy auto

\$server\_standby server\_2 -c mover=server\_8 -policy auto

## **Data Mover Hardware Prerequisites for Standby Configuration:**

Must have same Network configuration/setup as Primary

Must use same NIC & SCSI Hardware

## **ESTABLISHING STANDBY SERVER & FAILOVER RELATIONSHIP:**

**Rules:** Recommended 3:1 Data Mover relationship for failover [Failover Policies: auto/manual/retry]

*Failing back to original Primary Data Mover slot requires manual step by design!*

Step 1. Unexport & Unmount All File Systems from DataMover to be used as Standby

Step 2. Create Standby Server:

\$server\_standby server\_2 -create mover=server\_3 -policy auto [Standby server will reboot]

**Note:** First server in command is the “Primary”, second server is the “Standby” unit

**Verify Status:** \$nas\_server -l [Server\_3 now listed as type “4”]      \$nas\_server -i server\_2

## **Data Mover Failover Policies in /nas/server/server\_setup file:**

1=Manual failover (Default Celerra policy; all RDF setups are manual failover policy)

2=Retry (Panic will cause reboot to clear condition, if unsuccessful, failover occurs)

3=Auto (Data Mover will failover automatically after panicking)

## **DATAMOVER FAILOVER POLICIES:**

Auto →Primary Server fails over to designated Standby immediately & automatically

Manual (default) →Control Station reboots faulted Datamover to attempt recovery, but does not initiate failover to Standby

Retry →Control Station reboots primary DM once, then fails it over if still in faulted condition

**Note:** When setting up Failover Policies, keep in mind that “manual” is the default--make sure to specify “auto” instead

### **Local Celerra with (1) Primary & (1) Standby Server defined (NO RDF relationships):**

```
# nas_server -l
id type acl slot groupID state name
1 1 1000 2 0 server_2
2 4 1000 3 0 server_3
```

**# cat /nas/server/server\_setup**

1:2:1:3:: -->Fields defined, from Left-to-Right

**Primary\_id: Backup\_id: Component: Policy: Slot#, ACL, State:**

Server\_2 : Server\_3 : constant value : Failover Policy : RDF relationship as slot\_x, ACL, State

**Note:** This file is empty if there are no Standby Server relationships, either local or RDF, defined

### **Local Celerra with (1) Primary and (1) Standby Server, but NO Standby relationships defined:**

```
# nas_server -l
```

```
id type acl slot groupID state name
1 1 1000 2 0 server_2
2 4 1000 3 0 server_3
```

**# cat /nas/server/server\_setup** -->No entries since no Standby relationships have been defined

### **Local Celerra with (1) Primary and (1) Standby Server defined with Manual Failover Policy:**

```
# nas_server -l
```

```
id type acl slot groupID state name
1 1 1000 2 0 server_2
2 4 1000 3 0 server_3
```

**# cat /nas/server/server\_setup**

1:2:1:1::

### **Local Celerra with both Data Movers defined as Primaries, NO Standbys or RDF Standbys:**

```
# nas_server -l
```

```
id type acl slot groupID state name
1 1 1000 2 0 server_2
2 1 1000 3 0 server_3
```

**# cat /nas/server/server\_setup** -->File is empty because there are no local or RDF standbys defined

**Active/Active MirrorView Disaster Recovery:** No local Standby servers. Server\_2 Local Primary, with Server\_2 Remote RDF Standby. Server\_3 Remote Primary, with Server\_3 Local RDF Standby:

Local Celerra-->**# nas\_server -l**

```
id type acl slot groupID state name
1 1 1000 2 0 server_2
2 4 2000 3 0 server_3
```

**# cat /nas/server/server\_setup**

1:0:1:1:2,0,0,:;

Remote Celerra-->**# nas\_server -l**

```
id type acl slot groupID state name
1 4 2000 2 0 server_2
2 1 1000 3 0 server_3
```

**# cat /nas/server/server\_setup**

2:0:1:1:3,0,0,:;

### **Active/Passive MirrorView Disaster Recovery:**

Local Active side with Server\_2 Primary & local Standby defined as Server\_3. Local Standby Server\_3 has RDF failover relationship on slot\_3 on the Remote Celerra & Local Primary Server\_2 has RDF Standby (failover) relationship on slot\_2 on Remote Celerra

**# nas\_server -l (Active Source side)**

```
id type acl slot groupID state name
1 1 1000 2 0 server_2
2 4 1000 3 0 server_3
```

**# cat /nas/server/server\_setup**

2:0:1:1:3,1000,0,: -->Server\_3:no local backup:1=n/a:1=manual failover:3,1000,0=RDF Stdby slot\_3, ACL 1000, State 0

1:2:1:3:2,1000,0,: -->Server\_2:local backup server\_3:1=n/a:3=auto failover:2,1000,0=RDF Stdby slot\_2, ACL 1000 State 0

### **Fields Defined:**

Primary\_id: Backup\_id: Component n/a: Failover Policy: RDF Standby Slot#, ACL, State:

**# nas\_server -l (Passive Target side)**

```
id type acl slot groupID state name
1 4 1000 2 0 server_2
```

```
2   4  1000 3      0  server_3
# cat server_setup
1:2:1:3:: -->Server_2:local backup server_3:1=n/a:3=auto failover:no RDF relationships defined on this side
-->Server_3 not referenced since it has no local Standby or RDF relationship defined
```

## **CAUSES OF DATA MOVER FAILOVER:**

1. Failure of both internal NIC's [results in failed to ping error & failover]
2. Failure of Power Supply
3. SW, Panic or GP Exception
4. Memory Error
5. Loss of all storage connectivity
6. HW failure

**Note:** Current NAS codes [4.2.x & 5.1.x] behave as follows: If either path is lost to Symm, HBA will failover to alternate controller & port. If both cables are detached [both paths lost], then data mover will failover to Standby server.

## **TESTING DATA MOVER FAILOVER:**

Step 1: Failover Test: **\$server\_standby server\_2 -activate mover** [\$nas\_server -i]

Step 2: Failback Test: **\$server\_standby server\_2 -restore mover**

**Note:** Standby server assumes Primary's IP Address; MAC; Hostname; FS Configuration and Mounts. Can failback using CLI command or using Celerra Manager "Restore" button

## **DELETING A STANDBY SERVER & CREATING AS A NEW PRIMARY SERVER:**

Step 1. Delete Standby Status of Server\_5 for Server\_2:

**\$server\_standby server\_2 -delete mover**

Step 2. Make Server\_5 a Primary:

**\$server\_setup server\_5 -type nas**

**Note:** Data Mover should reboot automatically to build its Primary database, but reboot manually if need be.

## **EXAMPLE OF CHANGING SERVER FROM PRIMARY TO STANDBY:**

**# /nas/bin/server\_setup server\_5 -type standby**

server\_5 : is rebooting as type standby

## **SECONDARY CONTROL STATIONS:**

**\$/nas/sbin/getreason** Primary = slot\_0, status 10 Backup = slot\_1, status 11

**\$/nasmcd/getreason** [Running getreason from a Secondary Control Station]

**Note:** Each CS has its own O/S, but only one CS can mount the /nas & /nas/dos FileSystems at a time—first CS that boots will take over as Primary & is determined by the MCD daemon, which runs on both Control Stations. MCD (Master Control Daemon) runs on both CS but Box Monitor runs only on the Primary CS, as does nas\_eventlog, emcnas, etc. CS1 requires its own Modem, Phoneline, and Cable. Hourly NAS Database backups are written to /home/nasadmin.

**Scenario:** nas\_mcd listens for heartbeat of CS0. If no heartbeat registered, fails over to CS1 and initiates Call Home.

## **CONDUCTING MANUAL CONTROL STATION FAILOVER FOR LINUX:**

### **FAILING OVER FROM PRIMARY CS TO STANDY CS:**

**# /nasmcd/sbin/cs\_standby -failover**

**Note:** Issue command on whichever CS is the Primary to failover to the alternate CS—will reboot system you are on, to failover

### **FAILING OVER FROM PRIMARY CS0 to CS1:**

**# /nasmcd/sbin/cs\_standby -takeover**

**Note:** Issue this command on the Standby Control Station—unexports filesystems from Primary CS, mounts, and reboots the other Control Station as a Secondary. i.e., “takes over” from Primary to become the new Primary CS.

## **RESETTING CONTROL STATION:**

**# /nasmcd/sbin/t2reset pwron -s 1 | 0**

## **CONFIGURING CONTROL STATION TIME ZONE, DATE, & HW CLOCK NAS 5.6**

### **CHANGING CONTROL STATION TIME ZONE 5.6 (see emc214902):**

**Note:** The /usr/sbin/timeconfig utility does not exist in 5.6

**1. Verify current settings:**

**# date**

Tue May 5 08:14:20 EDT 2009

**# ls -la /etc/localtime**

lrwxrwxrwx 1 root root 36 Mar 20 09:57 /etc/localtime -> /usr/share/zoneinfo/America/New\_York

**# cat /etc/sysconfig/clock**

ZONE="America/New\_York"

UTC=false

ARC=false

**2. Use the following Perl Script to make the Time zone change, then verify:**

**# /usr/bin/perl /nas/http/webui/bin/timezone.pl -s America/Denver**

**# date**

Tue May 5 06:27:03 MDT 2009

**# ls -la /etc/localtime**

lrwxrwxrwx 1 root root 34 May 5 06:21 /etc/localtime -> /usr/share/zoneinfo/America/Denver

**# cat /etc/sysconfig/clock**

ZONE="America/Denver"

UTC=false

ARC=false

**Note:** The time zone string is based on "America/<city\_name>"--the city names can be found in the /usr/share/zoneinfo/America directory.

**3. Reboot the Celerra Control Station in order to update persistent applications, such as Apache, Tomcat, or XML API, with the new Time zone information.**

**ALTERNATIVE METHOD FOR CHANGING TIME ZONE 5.6:**

**1. Remove old time zone link file:**

**# rm -f /etc/localtime**

**2. Create new time zone link file based on city of choice from /usr/share/zoneinfo/America directory:**

**# ln -sf /usr/share/zoneinfo/America/Chicago /etc/localtime**

**3. Manually edit /etc/sysconfig/clock with new time zone string:**

ZONE="America/New\_York" --> change to --> **ZONE="America/Chicago"**

**4. Reboot Control Station and verify**

**# ls -l /etc/localtime**

lrwxrwxrwx 1 root root 35 May 5 10:09 /etc/localtime -> /usr/share/zoneinfo/America/Chicago

**# cat /etc/sysconfig/clock**

ZONE="America/Chicago"

UTC=false

ARC=false

**# date**

Tue May 5 10:10:42 CDT 2009

**VIEWING DAYLIGHT SAVINGS TIME INFORMATION 5.6:**

**# /usr/sbin/zdump -v /etc/localtime |grep 2009**

/etc/localtime Sun Mar 8 06:59:59 2009 UTC = Sun Mar 8 01:59:59 2009 EST isdst=0 gmtoff=-18000

/etc/localtime Sun Mar 8 07:00:00 2009 UTC = Sun Mar 8 03:00:00 2009 EDT isdst=1 gmtoff=-14400

/etc/localtime Sun Nov 1 05:59:59 2009 UTC = Sun Nov 1 01:59:59 2009 EDT isdst=1 gmtoff=-14400

/etc/localtime Sun Nov 1 06:00:00 2009 UTC = Sun Nov 1 01:00:00 2009 EST isdst=0 gmtoff=-18000

**CHANGING CONTROL STATION DATE & TIME 5.6:**

**# date 0504133709** →Where 05 is mnth, 04 is day, 1337 is time, 09 is year

Mon May 4 13:37:00 EDT 2009

**# date;date -u** (latter switch displays UTC time)

Mon May 4 15:47:19 EDT 2009

Mon May 4 19:47:19 UTC 2009

**CHANGING CONTROL STATION HW CLOCK TO MATCH LINUX SYSTEM CLOCK 5.6:**

**# /usr/sbin/hwclock**

Mon 04 May 2009 01:37:45 PM EDT -0.133015 seconds

**# /usr/sbin/hwclock -systohc (-w)**

**Note:** Sets hwclock based on current System Time (Linux kernel clock)

# /usr/sbin/hwclock

Mon 04 May 2009 01:37:24 PM EDT -0.072162 seconds

**HWClock Drift:**

/etc/adjtime

27.391

## ***CONFIGURING CONTROL STATION NTP CLIENT SERVICE NAS 5.6***

**GUI Method for Setting up NTP on CS:**

→Open Celerra Manager, go to Control Station properties, find NTP section, input IP address for NTP server to be used by the Control Station client service

**CLI Method:**

**1. Verify whether NTPD daemon is running:**

# ps -ef |grep ntpd

root 17682 11902 0 11:26 pts/0 00:00:00 grep ntpd

# /sbin/service ntpd status

ntpd is stopped

# /sbin/chkconfig ntpd --list

ntpd 0:off 1:off 2:off 3:off 4:off 5:off 6:off

**2. Add NTP Server IP to the bottom of the /etc/ntp.conf file per the following example:**

server 10.246.18.40

**3. Add IP Address of NTP Server (s) to /etc/ntp/step-tickers file:**

10.246.18.40

**4. If not already configured, set the NTP daemon for run-levels 3, 4, & 5, then verify:**

# /sbin/chkconfig --level 345 ntpd on | off [Enables or disables daemon for auto restart]

# /sbin/chkconfig ntpd --list

ntpd 0:off 1:off 2:off 3:on 4:on 5:on 6:off

**5. Start or restart the NTP Daemon:**

# /sbin/service ntpd start

ntpd: Synchronizing with time server: [ OK ]

Starting ntpd: [ OK ]

# /sbin/service ntpd restart

Shutting down ntpd: [ OK ]

ntpd: Synchronizing with time server: [ OK ]

Starting ntpd: [ OK ]

**Note:** If 'Synchronizing with time server' comes back as "OK", the NTP client was able to communicate with the NTP Server.

**6. Grep for the process:**

# ps -ef |grep ntp

ntp 25048 1 0 13:09 ? 00:00:00 ntpd -u ntp:ntp -p /var/run/ntp.pid

**7. Check Status of NTP:**

# /sbin/service ntpd status

ntpd (pid 25048) is running...

# /usr/sbin/ntpq -p

remote refid st t when poll reach delay offset jitter

=====

10.246.18.40 192.1.4.236 2 u 91 256 1 0.809 0.492 0.001

**8. Examine /var/log/messages file for ntpd messages. If you get errors, you may need to reboot the Control Station.**

May 4 13:09:45 fox1 ntpd[25048]: precision = 1.000 usec

May 4 13:09:45 fox1 ntpd[25048]: Listening on interface eth3, 10.241.168.30#123

May 4 13:09:45 fox1 ntpd: ntpd startup succeeded

May 4 13:09:45 fox1 ntpd[25048]: kernel time sync status 0040

May 4 13:09:45 fox1 ntpd[25048]: frequency initialized 0.000 PPM from /var/lib/ntp/drift

**9. Optionally, setup a drift value in the following file**

# vi /var/lib/ntp/drift

27.391

## ***EXAMPLE NTP SERVICE FAILING INVALID IP:***

```
# cat /etc/ntp/step-tickers  
10.241.168.172  
# /sbin/service ntpd restart  
Shutting down ntpd: [ OK ]  
ntpd: Synchronizing with time server: [FAILED]  
Starting ntpd: [ OK ]  
# /usr/sbin/ntpq -p  
remote refid st t when poll reach delay offset jitter  
=====
```

LOCAL(0) LOCAL(0) 101 8 64 1 0.000 0.000 0.001

**Note:** Above output indicates NTP not working & has an invalid NTP server IP address

### **EXAMPLE NTP SERVICE WORKING WHEN USING VALID IP:**

```
# cat /etc/ntp/step-tickers  
10.246.18.40  
# /sbin/service ntpd restart  
Shutting down ntpd: [ OK ]  
ntpd: Synchronizing with time server: [ OK ]  
Starting ntpd: [ OK ]  
# /usr/sbin/ntpq -p  
remote refid st t when poll reach delay offset jitter  
=====
```

LOCAL(0) LOCAL(0) 101 11 64 1 0.000 0.000 0.001  
10.246.18.40 192.1.4.236 2 u 10 64 1 0.682 -1.957 0.001

### **SETTING UP CS AS AN NTP SERVER:**

#### **a.) Configure NTP Service on CS0:**

```
# vi /etc/ntp.conf  
server 172.24.80.11 [Add this entry for the local Control Station External IP Address]  
server 127.127.1.0 # local clock [Leave this default entry]  
fudge 127.127.1.0 stratum 10 [Leave this default entry]
```

#### **b.) Enable NTP Daemon for NAS 5.0:**

```
# /sbin/chkconfig ntpd on
```

#### **c.) Start the Time Service on CS0:**

```
# /etc/rc.d/init.d/xntpd start | stop [NAS 4.x]  
# /sbin/service ntpd start [NAS 5.0]
```

#### **d.) Verify Time Service is Running:**

```
# ps -ef |grep xntpd "...root 16797 xntpd -A" ---or---  
# ps -ef |grep ntp  
ntp 19772 1 0 17:18 ? 00:00:00 ntpd -U ntp  
# /sbin/service ntpd status  
ntp is stopped  
# /sbin/chkconfig --list |grep ntp  
ntpda 0:off 1:off 2:off 3:off 4:off 5:off 6:off
```

#### **e.) Test by starting DataMover's NTP Service:**

```
# server_date server_2 timesvc start ntp 192.168.1.100
```

#### **f.) Synchronize DataMover to CS0 using “update” command or manually if required:**

```
# server_date server_2 timesvc update ntp  
# server_date server_2 0209201130
```

**Note:** Date is in Year, Month, Day, Time format

#### **g.) Verify by comparing Control Station Time to DataMover Time:**

```
# date # server_date server_2
```

#### **h.) Create Symbolic Link to restart Time Service upon rebooting Control Station:**

```
# /sbin/chkconfig xntpd on [Creates symbolic link /etc/rc.d/rc3.d/S55xntpd → /init.d/xntpd]
```

#### **i.) Reboot Control Station & verify**

**Notes:** Default polling to the Time Server is hourly, but this can be modified. Normal ‘fudge’ configuration is to use “stratum 10” as local ‘unreliable clock source’—but since it’s the primary Time Server here, change value to (8) to preclude client sync problems. Or,

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
if you would like to setup the Control Station to point to an external Time Server source, yet have the DataMovers point to the Control Station, and then use the CS's "local clock" as a backup if the External NTP Server is unavailable, setup ntp.conf file like this:

```
server 200.1.1.1      prefer # IP of external time server
server 127.127.1.0
```

## **SETTING UP NTP CLIENT ON LINUX 6.2/7.2 CONTROL STATIONS (pre 5.6):**

- Step 1. #vi /etc/ntp/step-tickers  
192.10.4.5 [IP Address of NTP Server]
- Step 2. Set Control Station to Autostart Time Service on Reboots:  
#/sbin/chkconfig xntpd on [RedHat 6.2]  
#/sbin/chkconfig ntpd on [RedHat 7.2]
- Step 3. Verifying Configuration & Startup Level:  
#/sbin/chkconfig --list xntpd [RedHat 6.2]  
#/sbin/chkconfig --list ntpd [RedHat 7.2]
- Step 4. Manually Starting the NTP Daemon:  
#/etc/rc.d/init.d/xntpd start [RedHat 6.2]  
#/etc/rc.d/init.d/ntpd start [RedHat 7.2]  
Synchronizing with time server: OK ]  
Starting ntpd: OK ]
- Step 5. Verify NTP Service:  
#ps -ef |grep ntp  
ntp 26728 1 0 16:38 ? 00:00:00 ntpd -U ntp

## **TROUBLESHOOTING NTP ON LINUX:**

- check to ensure that service is running  
--# /usr/sbin/ntpq -p  
--#ps -ef |grep xntpd "...root 16797 xntpd -A"  
--# ps -ef |grep ntp  
ntp 19772 1 0 17:18 ? 00:00:00 ntpd -U ntp  
--#/sbin/service ntpd status  
ntpd is stopped  
--# /sbin/chkconfig --list |grep ntp  
ntpd 0:off 1:off 2:off 3:off 4:off 5:off 6:off  
--ping the NTP Server  
--# /usr/sbin/rpcinfo -p 192.10.3.2 [Verify that Time Service port is registered]  
--#/usr/sbin/ntptrace

## **UTILITIES TO CHECK NTP:**

```
#/usr/sbin/ntptrace
#/usr/sbin/ntpdate
#/usr/sbin/ntpq
>peers
#/usr/sbin/ntpdc
>sysinfo
>listpeers
```

## **FILES USED WITH NTP ON CONTROL STATION:**

/etc/ntp.conf  
/etc/rc3.d/K74ntpd  
/sbin/chkconfig -list |grep -i ntpd [Verify if NTP daemon is running]  
/usr/sbin/clockdiff -o IPaddress [IP Address of NTP Server]  
/usr/sbin/timeconfig [Useful GUI for configuring TimeZones on CS—NAS 5.5 and lower]  
/etc/sysconfig/clock [Shows TimeZone currently configured]  
/bin/date [Current date & time—also use this to change date & time]  
/etc/ntp/drift or /var/lib/ntp/drift

## **SETTING DATE & TIME ON LINUX CONTROL STATION (NAS 5.5 and before):**

1. # date -s "4/09/03 16:15:15"
2. # /usr/sbin/setclock

### 3. # date → Wed Apr 9 16:15:30 MDT 2003

#### OBTAINING UNIX DATE ON LINUX CS:

# /bin/date '+%s'

1074016065

#### Convert to Number of Bits:

Open scientific calculator, input above number as decimal, then convert to binary:

100000000000100001011101000001

Note: Above example shows 31 bits in use to display Unix date since Jan 1, 1970

#### RESTARTING TIME SERVICE FOR LINUX 6.2 CS:

#/sbin/chkconfig xntpd on [This command creates the link → /etc/rc.d/rc3.d/S55xntp]

Note: Create symbolic link to auto-restart the service after Control Station reboots

#### CHANGING CONTROL STATION TIMEZONE ON LINUX:

The default system TimeZone for Linux generally resides in /usr/share/zoneinfo, with /etc/localtime as a symbolic link to the actual Timezone file that is desired as the reference. See man pages for tzset for more information.

#### USING CLI SYMLINK TO CHANGE FROM DEFAULT EST TO PST:

1. Create the link, or, find the proper timezone file (/usr/share/zoneinfo/PST8PDT) and copy to /etc/localtime

# cp -ip /usr/share/zoneinfo/PST8PDT /etc/localtime

Or # ln -s /usr/share/zoneinfo/PST8PDT /etc/localtime

# ls -la /etc/localtime

lrwxrwxrwx 1 root root 27 Nov 29 07:43 localtime -> /usr/share/zoneinfo/PST8PDT

2. Edit /etc/sysconfig/clock and change to reflect the correct timezone

# cat /etc/sysconfig/clock

ZONE="America/New\_York"

UTC=false

ARC=false

# cat /etc/sysconfig/clock

ZONE="PST8PDT"

UTC=false

ARC=false

#### USING TIMECONFIG TO CHANGE TIMEZONE (NAS 5.5 and below):

Step 1. # /usr/sbin/timeconfig [Select timezone of choice for Control Station]

Step 2. Manually set date and time on all Datamovers

Note: Using the GUI is probably the best way to change timezones on the Control Station, since the changes are written to /etc/sysconfig/clock and the correct timezone file is copied to /etc/localtime.

# cat /etc/sysconfig/clock

ZONE="EST5EDT"

UTC=false

ARC=false

#### VERIFYING CONTROL STATION DST TIMEZONE CHANGES:

NAS 5.4.27.2 & 5.5.24.2 contain code fixes to update the Linux glibc files so as to be able to honor the new 2007 Timezone. Run the following command to determine if the system has been updated to recognize the new start & end dates of DST, which is the 2<sup>nd</sup> Sunday in March to the first Sunday in November.

# /usr/sbin/zdump -v /etc/localtime |grep 2007

/etc/localtime Sun Mar 11 06:59:59 2007 UTC = Sun Mar 11 01:59:59 2007 EST isdst=0 gmtoff=-18000

/etc/localtime Sun Mar 11 07:00:00 2007 UTC = Sun Mar 11 03:00:00 2007 EDT isdst=1 gmtoff=-14400

/etc/localtime Sun Nov 4 05:59:59 2007 UTC = Sun Nov 4 01:59:59 2007 EDT isdst=1 gmtoff=-14400

/etc/localtime Sun Nov 4 06:00:00 2007 UTC = Sun Nov 4 01:00:00 2007 EST isdst=0 gmtoff=-18000

Note: See emc150583 for more details.

#### DATA MOVER TIME SERVICE, TIMEZONE, DATE

#### VERIFYING DATAMOVER TIME SERVICE & SETTING NTP SERVERS:

#server\_date server\_6 timesvc start ntp 162.69.63.8

#server\_date server\_6 timesvc

Timeservice State

time: Wed Apr 24 13:43:42 EDT 2002  
type: ntp  
sync delay: off  
interval: 60  
hosts: 162.69.63.8

### **DATA MOVER DISPLAY TIME:**

Data Mover display and log times are calculated from the 1970 UNIX clock in seconds, then converted to show the local timezone that the Control Station is running. In otherwords, the following command will always output the localtime as set on the CS—to make matters more confusing, the default Timezone of the Data Mover is always GMT unless specifically changed:

# server\_date server\_2

server\_2 : Wed Dec 13 12:54:27 EST 2006

# server\_date server\_2 timezone

server\_2 : Local timezone: PST

### **SETTING TIMEZONE ON DATA MOVER(old):**

# .server\_config server\_2 -v "time PST+8" (sets Timezone but without Savings time set)

**Note:** Supported 5.2.14.0 and 5.3.6.0

# server\_date server\_2 timezone CST6CDT5,M4.1.0,M10.5.0 [Stnd time +offset, Savings time + offset, then Month, Week, Day of Week—default time is /2:00 for 2:00 a.m.]

# server\_date server\_2 timezone EDT+4

# server\_date server\_2 timezone BST-1

**Note:** If desired, plug in daylight savings time information—First Sunday April, 5<sup>th</sup> Sunday October. Also, must point to an NTP server in the same timezone in order for the proper time to be returned with server\_date. See also extra examples of different timezones and syntax required.

### **VERIFYING DATA MOVER TIMEZONE DATE INFO:**

1. \$ .server\_config server\_2 -v "time"

1089329961: KERNEL: 4: Local time: Thu Jul 8 23:39:21 2004

1089329961: KERNEL: 4: GMT time: Thu Jul 8 23:39:21 2004

**Note:** Output shows default GMT time, as well as local time, which is really keyed off the local CS clock? If you specifically change DM timezone, the Local time will then change to reflect the offset from GMT time for the timezone specified.

2. Following entry will be written to boot.cfg file only when Data Mover is set to a timezone other than its default for GMT

# cat /nas/dos/slot\_2/boot.cfglhead

time timezone=PST8PDT7,M3.2.0,M11.1.0

**Note:** There is no “time timezone” entry for the default Data Mover GMT timezone

3. Following Server system file will also contain the timezone entry:

# cat /nas/server/slot\_2/system

time timezone=PST8PDT7,M3.2.0,M11.1.0

4. Use server\_date command:

# server\_date server\_2 timezone

server\_2 : Local timezone: PST

### **CHANGING DATA MOVER TIMEZONES(new):**

**Note:** Default DM timezone is GMT

# server\_date server\_2 timezone

server\_2 : Local timezone: GMT

1. To change Timezone of Data Mover to Pacific Standard Time with the new Pacific Daylight Time for 2007:

# server\_date server\_2 timezone PST8PDT7,M3.2.0/2:00,M11.1.0/2:00

**Note:** PST8 represents Pacific Stnd Time and its – 8 hour offset from GMT; PDT7 is Pacific Daylight Time with an offset of -7; M3.2.0/2:00 is the start of Savings Time; M11.1.0/2:00 is the end of Savings Time

M3 = Month (1-12), in this case March

2 = Week of the month (1-5), 2<sup>nd</sup> week of March

0 = Day of the week (0-6), with 0 always representing Sunday

/2:00 = time, where 2:00 is a.m.

### **BEFORE TIMEZONE CHANGE:**

```
# server_date server_2 timezone  
server_2 : Local timezone: GMT  
# .server_config server_2 -v "time"  
1164813455: KERNEL: 4: Local time: Wed Nov 29 15:17:35 2006  
1164813455: KERNEL: 4: GMT time: Wed Nov 29 15:17:35 2006
```

```
# server_date server_2  
server_2 : Wed Nov 29 10:17:41 EST 2006
```

#### AFTER TIMEZONE CHANGE:

```
# server_date server_2 timezone  
server_2 : Local timezone: PST  
# .server_config server_2 -v "time"  
1164813498: KERNEL: 4: Local time: Wed Nov 29 07:18:18 2006  
1164813498: KERNEL: 4: GMT time: Wed Nov 29 15:18:18 2006
```

```
# server_date server_2
```

server\_2 : Wed Nov 29 10:22:49 EST 2006 →This is actually correct since server\_date uses same time as Control Station, which is set to EST in this example—so, server\_date has nothing to do with actual DM timezone info:

```
# date
```

Wed Nov 29 10:33:02 EST 2006

```
# cat /nas/server/slot_2/system
```

time timezone=EST5EDT4,M3.2.0,M11.1.0 →File where actual timezone changes are written if anything other than default GMT

#### SETTING TIMEZONE BACK TO GMT DEFAULT:

```
# server_date server_3 timezone GMT+0
```

```
server_3 : done  
# cat /nas/server/slot_2/system  
time timezone=GMT+0
```

#### SETTING TIMEZONE INFO ON DATA MOVERS WITH NAS 5.6:

```
$ server_date server_2 timezone -name America/New_York
```

Note: Can use above format to set timezone vs. posix entry

#### DELETING TIME SERVICE ON DATAMOVER:

```
$server_date server_2 timesvc delete ntp
```

#### CELLERRA TIME SERVICES: NAS 5.1

##### Terminology:

Clock Offset—difference between local system time and Time Server time, expressed in UTC (Coordinated Universal Time), which is independent of Time Zones. Clock Offset can be negative or positive. Datamover typically slews at a rate of 10%, meaning that it would take (6) minutes to recover from an offset of (1) hour

NTP (Network Time Protocol)—Time service specifying accuracy within one millisecond over large networks

SNTP (Simple Network Time Protocol)—Subset of NTP, does not demand as much accuracy

Slewing—Gradual self-correction process to slew back to NTP Time

#### TIME SERVICE COMMANDS:

```
$server_date server_x timesvc set ntp [Polls NTP Server & syncs system time immediately—no slewing--Default]
```

```
$server_date server_x timesvc update ntp [Polls NTP Server & begins slewing to sync back up to Time Server]
```

```
$server_date server_x timesvc start ntp -synch_delay -interval 01:00 host 192.10.3.2 [This option will begin slewing system time to NTP Time after reboots or startup, per the default interval of (1) hour between polls]
```

#### SETTING UP MULTIPLE NTP SERVERS AT ONCE:

```
# server_date server_2 timesvc start ntp 10.92.5.189 10.70.154.205
```

Note: Normally, system reboots or startups will resynchronize System time to NTP Time immediately—there might be occasions where you would not want to do this.

```
$server_date server_x timesvc stats ntp [Statistics about time services]
```

```
$server_date server_x timesvc stop ntp
```

```
$server_date server_x timesvc delete ntp
```

#### USING SCRIPT TO SYNC DATA MOVER TO CONTROL STATION, RUNNING AS NTP PEER:

```
# cat time  
#!/bin/sh  
NAS_DB=/nas
```

```
export NAS_DB
now=`/bin/date '+%y%m%d%H%M%S'
sleep 40
$NAS_DB/bin/server_date server_3 $now
FOR DUAL CS ENVIRONMENT:
$ cat time
#!/bin/sh
NAS_DB=/nas
export NAS_DB
if [ ! -d /nas/sbin ]; then
exit
else
/nas/bin/server_date ALL $(/bin/date +%y%m%d%H%M%S)
fi
```

## **CELERRA SCSI SUPPORT [Wide Ultra2 SCSI--80MBps—UFWD HVD]:**

**SCSI:** Parallel interface with speed bursts. Max cable lengths 3 – 25m. Total of 16 interconnected devices. SCSI III Protocol.

**Note:** Celerra uses Wide Ultra2 SCSI, with bus speeds up to 80 MB/second. Data Mover SCSI HBA's are single-ended HVD connections, but can use an adapter to LVD

Older configurations may still exist, which use Ultra2 SCSI, 40MB/second

### **Rules:**

→Two SCSI ports must be mapped to same SYMM devices—Why?

1. For High Availability

2. For Performance reasons [507DM can overdrive a single SCSI port]

→Up to 240 SCSI id's per DM

→16 Targets per SCSI Bus Chain—one reserved for HBA

→16 LUNS typical per Target. SCSI Port A = Data & Boot Port on Bottom Port of card. Lose this and must Failover to Standby. Port B=Data Only and is Top Port on SCSI Card.

**Note:** 2.1.24.4 Code uses both ports as “Boot” and “Data” ports to eliminate this earlier failover issue!! [Still 20% occurrence?]

CTD Terminology: Controller, Target, Device{LUN}

**Example:** **/dev/rdsck/c1t1d2 c1=controller t1=target d2=LUN**

Adaptec AHA-3944UWD Ultra Fast Wide Differential Dual Port PCI SCSI Cards on 505+ using C6mini68s or c12mini68s cables

**Cables:** Usually mini-to-mini for Celerra to Symm; But, mini-to-large connector for PC to Celerra, tape unit, etc.

SCSI communicates to Symmetrix using USD4; WD4; or OSD4 electronic negotiation.

**Note:** Minichamp = 6 or 12 meter cables

**SCSI :** Small Computer System Interface—parallel interface standard for attaching peripherals to a computer

**CDB :** Command Descriptor Block

**CAM:** Common Access Method, segment of code dealing with raw I/O communications to backend storage devices

## **SCSI DISTANCE LIMITATIONS:**

Single Ended Cables = 3 meters

Differential SCSI = 25 meters

## **CELERRA/CLARIION SCSI TIMEOUT VALUES:**

**Note:** NAS codes from 4.2.8.1 - 4.2.11.x & 5.0.9.1 – 5.0.11.x required setting the io\_time parameter. NAS 4.2.12.0 & 5.1.9.4 and higher do not require the param setting as the backend type is detected automatically and the setting made.

--Timeout Values for Celerra Servers attached to SYMMETRIX must be 90 secs

--Timeout Values for Celerra Servers attached to CLARiiON must be 300 secs [5 minutes]

## **CELERRA CONFIGURATION CHANGE REQUIRED:**

**Note:** As a result of this new behavior for SCSI timeouts, Celerras attached to SYMMETRIX must have following parameter added for all NAS Codes 4.2.8.1 – 4.2.12.x. The effect of not doing this means that failover, for loss of connectivity, will not occur for 5 minutes. See Primus emc60921.

1. #vi /nas/site/slot\_param

**param storage io\_time=90**

2. Reboot all Servers

## **SCSI INTERFACE PROTOCOL:**

- SCSI is a device independent I/O bus used to connect a variety of peripheral devices to a single computer bus
- Data is addressed logically, not physically
- SCSI buses can host (8) devices on a single Bus, but only (2) devices can converse with each other at a time
- SCSI Roles are as Initiator or Target
- Initiator [SCSI Host Adapter] begins a transaction by giving another device a task
- Target [Disk Drive] carries out the task
- Typical conversation involves Initiator sending command to Target, Target carrying out Task, & then reporting outcome [Target controls Bus until released]

**Example:** Initiator instructs target to fetch a certain number of blocks starting at block number. Target calculates physical address of cylinder, head, and sector number from logical block number and sends data to Initiator.

**Note:** Computer Systems connect to SCSI bus via Host Adapter, generally built into motherboard  
Peripherals connect to SCSI bus via Controller, embedded in drive electronics

## **SCSI CONTROLLER:** Used to connect peripherals to Bus; Controller has SCSI ID and Targets have LUNs

- SCSI ID & Address are unique on the Bus
- SCSI Reset is a special bus signal that affects all devices on the Bus
- Odd Parity Bit is set to 0 if total number of 1's are odd (1 parity bit for every 8 data bits)

## **PHASES IN SCSI COMMUNICATION:**

**Bus Free Phase:** SCSI bus is free, set by Power On, SCSI reset, execution of command completion by Target, or when Target relinquishes Bus with Disconnect message

**Arbitration Phase:** Devices must arbitrate for control of Bus (BSY)

Devices with highest SCSI ID (7) wins arbitration and uses SEL signal to allow other devices to release any BSY signals from wire and to remove their ID bit from data bus

**Selection Phase:** BSY, SEL, Initiator ID

Connection established by Initiator to desired Target, except in ‘reselection’ where Target asserts I/O ATN data signal issued to Target ID; Target asserts BSY signal & Initiator releases SEL signal

**Reselection Phase:** Allows Target to reconnect to Initiator after disconnecting to complete a command

Also allows Target to Send I/O reselection signal to Initiator after successful Arbitration

**Message Phase:** Message Out always follows successful selection & is used by Target to send or receive message byte Initiator sends ATN to Target to send message bytes, and then releases

**Command Phase:** Command opcodes used by Target to receive commands from Initiator

Target reads first byte to determine number of bytes to follow

**Data Phase:** Target sends data using I/O DATA IN signal, or is ready to receive data by de-asserting I/O in DATA OUT signal. Data is transferred asynchronously or synchronously

**Status Phase:** Target sends 1 byte Success or Failure message to Initiator after command completed, or if command interrupted or refused by Target. Target releases the Bus for Bus Free Phase.

## **SCSI BUS:**

8-32 bits wide, 50-pin ribbon cable, with built-in terminating resistors

Commands, messages, Status are sent over bus as asynchronous transfers [SCSI-2 can transfer data synchronously]

Single-ended SCSI is most common, but where cable lengths need to be > than 6 meters, Differential SCSI is used

## **SCSI/FIBRE TARGET & LUN CONCEPTS:**

Scsi Bus is from 8 – 32 bits wide. The 8-bit bus uses a 50pin cable. The 16 and 32-bit buses are called “Wide SCSI”.

|                    |        |                         |                      |
|--------------------|--------|-------------------------|----------------------|
| Wide SCSI-2:       | 20MBps | 16-bit bus = 16 Devices | 3 Meter Cable Length |
| Fast Wide SCSI-2:  | 20MBps | 16-bit bus = 16 devices | 3 Meter Cable Length |
| Ultra SCSI-3:      | 40MBps | 16-bit bus = 16 devices | 1.5 Meters           |
| Ultra-2 SCSI:      | 40MBps | 8-bit bus = 8 devices   | 12 Meters            |
| Wide Ultra-2 SCSI: | 80MBps | 16-bit bus = 16 devices | 12 Meters            |

## **SCSI STATUS CODES:**

|     |                 |                                                               |
|-----|-----------------|---------------------------------------------------------------|
| 00h | GOOD            | Device server has successfully completed task                 |
| 02h | CHECK CONDITION | ACA or CA condition occurred; autosense data may be delivered |
| 04h | CONDITION MET   | Status returned after requested operation satisfied           |
| 08h | BUSY            | Logical Unit is busy—reissue command at a later time          |
| 40h | TASK ABORTED    |                                                               |

## **DEVICE CLASSES:**

00h—Disk Drives  
01h—Tape Drives  
02h—Printers  
03h—Processor devices  
04h—WORM devices  
05h—CD-ROM drives  
06h—Scanners  
07h—Optical Disks  
08h—Media Changer  
09h—Communication devices

## **TARGET & LUN SUPPORT ON CELERRA:**

SCSI=240LUNS

FIBRE=256LUNS

**Note:** This is a Symmetrix limitation. Celerra could otherwise support up to 4096 LUNs

## **TARGET & LUN LIMITATIONS ON CELERRA:**

For NAS 2.x and above, use only Target and Lun assignments for Data Volumes that start at Target 1 Lun 0 [10]

Target/LUNs 006; 007; 008; 009; 00A; 00B; 00C; 00D are reserved for Control Volumes only! Errors will be seen on Control Station if these LUNs are used for Data Volumes

**SCSI CHAINS or PORTS:** Represents the whole physical SCSI port [aka chain] to which a cable is connected to!

Chains C0, C1, C2, C3 are available for SCSI

Chains C0 -C100+ available for Fibre

Each SCSI chain is simply a SCSI connector port for a given Server.

**For Example:** A dual-scsi connector would give you Chains 0 and 1

Two Dual-SCSI cards would give you Chains 0, 1, 2, 3

**Limitations:** With our SYMMETRIX and SCSI, we are limited to a total of 240 LUNS per chain!

**Note:** In Theory, DataMover could support up to 4096 devices. Fiber supports up to 128 LUNS [256with RPQ]

**SCSI PINS or TARGETS:** There are a total of (15) pins per connector, which represent “Targets” !

8-bit SCSI bus has Targets 0-7

16-bit SCSI bus has Targets 0-15 [lose one target for termination at target 7]

Each pin is therefore equivalent to a “Target” that transmits data down a wire

Each Target or Pin can have up to 15 LUNS mapped to it as well

**Limitations:** Pin #7, or Target 7, is always reserved for SCSI termination and cannot be used!!!!!!

Therefore, maximum LUNS available are 240 for a 16-bit SCSI Target

**SCSI LUNS:** Virtual addresses—can have a total of (15) virtual addresses on each SCSI Target or Pin 16-bit buses

Listed as LUNS 0—15 [lose one LUN for termination for ‘target’ 7]

## **SCSI3 RESERVATIONS:**

SCSI3 Reservations are used in some Tape Backup implementations where Shared drives are used. Hosts can take out a persistent reservation against a drive, but CommVault does not support SCSI3 Reservations for NAS-attached drives. Celerra ends up not being able to access the drive because of a reservation conflict error. Work is being conducted by both CommVault and EMC to review this limitation.

## **HOT PLUGGING/UNPLUGGING SCSI CABLES:**

Celerra SCSI connections are not like “hot swappable” disk drives that can be removed or inserted on live systems [Hot Swappable devices have built-in circuitry & longer ground pins to allow this] without potential impact.

### **Potential Problems Due to Hot Plugging or Unplugging SCSI Connections to a Live SCSI Bus:**

1. Could cause DM to negotiate from Wide to Narrow, which may not manifest itself until a server\_devconfig probe sends an electrical pulse down the ‘unplugged/replugged’ SCSI chain, creating errors on the Bus or Data Mover panic.
2. Breaking the SCSI Bus can leave the bus in an unstable electrical state, leading to commands failing and potential data corruption, and possibility for generating double bit errors.
2. Plugging Tape Devices generates a power surge from TERMPWR on the TLU/Tape device, which could affect DM SCSI circuitry.
3. Unplugging/plugging cables creates voltage surge as devices are ungrounded or grounded. This can lead to a short circuit on the SCSI cable, blowing the Pico Fuse [which is embedded on the SCSI cables so as to be the protective point for the SCSI bus]. A danger here is that further power surges, or short circuits, could damage both the HBA and the Tape Drive device. Connections are

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
designed so that the SCSI device makes a power ground and logic ground 1ms before device connector physically contacts the Bus.  
Hot swappable devices have specially extended ground pins to help protect the SCSI bus & electrical circuitry.

## **iSCSI PROTOCOL:**

- encapsulation protocol using TCP to transport scsi commands
- iSCSI links targets and initiators and presents SCSI LUNs to iSCSI initiators
- SCSI used as the presentation interface
- GigE required for throughput, and TOE's [TCP/IP Offload Engines in NIC Cards] are being used
- Remote disks appear as local drives with iSCSI, hanging off Controller, Target, Lun
- Some drawbacks to usage of iSCSI over WANs—latency, performance—don't use over public network
- iScsi uses basic transfer length of 0x800 or 2048 bytes (2k), equivalent to (4) 512 byte blocks as in SCSI
- iSCSI uses TCP port 3260

## **IMPORTANT CONCEPTS:**

iSCSI does not know about LUNs (uses CDB via SCSI), just Targets

## **COMMUNICATION LAYERS:**

Physical Layer

TCP Header

iSCSI BHS → Basic Header Segment--only layer different than using FC with SCSI. iSCSI device driver adds PDU data and passes packet to TCP stack

CDB

Data

## **iSCSI IDENTIFIERS:**

iSCSI Nodes have globally unique names and support 'eui' and 'iqn' names

iSCSI Sessions use Session IDs (ISIDs) using TCP between iSCSI Nodes

iSCSI Connections use Connection IDs to identify logical connections

iSCSI Portals manage IP addresses and TCP Port numbers

## **iSCSI INITIALIZATION:**

iSCSI Initiator sets up TCP connection to Target, then 'Discovers' SCSI Targets and Logs into Target via Session [Discovery or Normal]

Normal SCSI commands report on LUNs

## **TWO SESSIONS USED IN iSCSI:**

Discovery: Uses SendTargets and Logout commands

Normal: Accepts all commands

## **iSCSI SECURITY:**

IPSec and IKE encryption

Kerberos v5, SRP, SPKM1/2, and CHAP authentication

## **CELLERRA iSCSI:** Introduced NAS 5.3

### **NBS:**

Network Block Service (version 2.0) is the protocol that runs on top of the TCP, hence is connection-oriented, and can run over copper or optical media. Client side applications issue requests to Block Device Driver which are in turn passed to Block Server via TCP. Client device driver keeps an exported pseudo disk list that makes storage appear to be local, just as a local disk which can be partitioned, formatted, etc. Storage Objects on Celerra are file systems and files. Current implementation for authentication is CHAP 1-way from Client to Server.

## **TYPES OF iSCSI TOPOLOGIES:**

### **NATIVE iSCSI SAN**

→ Entire topology between iSCSI Client Initiator and iSCSI Target Storage array is connected from end-to-end over Ethernet, with an IP switch in the middle, and transmitting SCSI over IP

### **HETEROGENEOUS IP SAN**

→ A topology where some parts of the SAN are using IP to transmit SCSI, while others are using FC to transmit SCSI, with perhaps a Bridge to translate between iSCSI/Fibre Channel protocols

→ An iSCSI Host can use either a NIC or HBA device, while FC Target must use an HBA

## **MICROSOFT iSCSI:**

→ Microsoft iSCSI Initiator is a software driver installed on Windows 2000, 2003, & XP systems, but comes native in the OS for Windows 2008 & Vista (using Winsock Kernel mode interface—WSK)

**Note:** An example of a compatible version would be Initiator version 2.08

- An iSCSI Client is a Host or iSCSI Initiator which attaches to an IP network and sends requests and receives responses from iSCSI targets
- Microsoft recommends a minimum GbE Ethernet for iSCSI
- Typical iSCSI involves a Windows Host running the MS iSCSI software Initiator
- Microsoft recommends using Windows 2003 or 2008 Server for Microsoft Server Cluster (MSCS) support
- Use of dynamic disk volumes for iSCSI is only supported on Vista & Windows 2008
- Microsoft iSNS Server comes native with Windows 2008 (maintains iSNS registrations, de-registrations, and queries iSNS clients)
- MS Initiator supports CHAP (username and secret) and IPSEC (IKE protocol for authentication and packet encryption) on sw and hw initiators

### **iSCSI INITIATOR DISCOVERY MECHANISMS:**

SendTargets—SendTarget ports are statically configured to discover target portals at service startup, etc

iSNS—statically configure the iSNS address, Initiator can then obtain list of targets from iSNS server

HBA Discovery—MS Initiator obtains list of targets from HBA

Manually configured targets—manually configuring iSCSI targets

### **TARGETS vs. HOST INITIATORS:**

Celerra Servers are usually configured as “Targets”, hosts the iSCSI LUNs, and runs the iSCSI service, while the Host Windows platform act as the HBA “Initiator”, and once logged into the “target”, sees the iSCSI luns as if they were locally mounted partitions on the Windows system.

### **CELERRA MANAGER:**

Can configure & manage LUN, Target, iSNS management, and CHAP

### **iSCSI TOPOLOGIES:**

Bridging—use of iSCSI and Fibre Channel in mixed environment, not supported for Celerra

Native—use of only standard Ethernet switches, routers, and media for iSCSI→NO FIBRE CHANNEL (Default for Celerra)

### **iSCSI PROTOCOL:**

iSCSI is a open network protocol and network service running over TCP/IP. Network addressing is done via Portal Groups, which associate an iSCSI Target to an assigned network service port. Protocol allows for host initiators to communicate with iSCSI targets on storage devices using scsi commands between Host & storage. iSCSI Initiator takes SCSI CDB's and encapsulates them in PDUs, then wrapped in TCP & IP for transport to iSCSI Target.

### **iSCSI:**

--Celerra uses iSCSI in a solution in which there are Ethernet switches and routers. iSCSI is a Client/Server model. The iSCSI target and host initiators receive CDB's from SCSI Layer, encapsulated within iSCSI PDU (Protocol Data Units) that are then sent via IP to host

--iSCSI provides ‘block access’ vs. file access, an attribute which is helpful for databases (Oracle/Exchange)

--An iSCSI connection is a network portal connection that binds an Initiator with a Target (TCP port & IP address socket used by iSCSI node)

--Single iSCSI name per Host

--iSCSI can be managed from CLI or Celerra Manager

--Virtual LUN module in DART to support iSCSI LUNs

--DART runs iSCSI service to maintain updates for Internet Storage Name Service and Entity Status Inquiry [ESI] messages

--Build file systems >than double the capacity of iSCSI LUNs to support snaps

--Maximum of 256 LUNs allowed per Target

--iSCSI LUNs are based on File Storage Objects (SCSI-3 standard with direct-access disk emulation)

--Client access restricted using LUN masking for each LUN within a Target [auth. process makes LUNs avail. to hosts]

--iSCSI LUNs are created by specifying File System, Target, and LUN number [fs3\_T1\_LUN17]

--Default port is 3260

--Support for Digests, CHAP, iSNS, ImmediateData, FirstBurst, and Session-level 0 error recovery

--Celerra Host creates iSCSI Targets from within file systems as the access point (Initiators) to storage, using LUN Masking to restrict access from Initiators (Mask is assigned to each LUN within the Target, all Initiators denied access until mask is assigned)

--iSCSI memory requires contiguous physical memory

### **VMWARE ISSUES:**

#### **1. Memory Corruption Issues affecting iSCSI LUNS and VMWare machines**

**Note:** See emc196974 and AR115513 for 5.5 and 5.6 memory corruption issues on iSCSI LUNs, which especially affects VMWare Virtual Machines (VMs), as they will BlueScreen and run chkdsk if the lun is corrupt]. Fixed in full NAS versions 5.5.37 and 5.6.40.

#### **2. VMWare UUID changes after upgrading from NAS 5.5 to 5.6**

**Note:** See emc189395. Problem is that the Protocol Identifier Value (PIV) in the SCSI INQ response has changed format, which then triggers the VMWare UUID to change. The only way to recover the ESX Server access to iSCSI LUNs is to resignature the LUNs. In

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
extreme cases, the volume labels may need to be re-labeled and re-registered on all VMs. There is a special VMWare procedure to recover volumes: essentially, enable LVM Resignaturing on ESX Server & rescan, then disable LVM Resignaturing, re-label volumes in Datastores to their original names, rescan storage from all ESX Servers, re-register all VMs in VirtualCenter, etc.

### **iSCSI PDU (PROTOCOL DATA UNIT):**

The basic iSCSI packet is called the Protocol Data Unit, contains Header segment, additional header segment, Header digest, SCSI cmd & data, Header data digest

### **SCSI COMMAND SET USED WITH iSCSI:**

SPC-3→Command set used with printers, scanners, other SCSI devices, using Media Sense, Inquiry, Reset, etc.

SBC-2→Provides SCSI Block Commands for interacting with SCSI disk devices

**Note:** In combination, these command sets allows Hosts to manage LUNs as ‘local devices’

### **iSCSI STACK FROM HOST TO STORAGE:**

LAN→NIC Driver on Host→IP→TCP→Stream→Connection→Session→Target→Storage LUN→Virtual LUN→File Storage Object→File System→Logical Volume→CAM/SCSI/FC→SAN

### **iSCSI COMPONENTS:**

Logical Units; iSCSI Targets; LUN Masking; iSCSI Service, iSNS Name Service; CHAP Secret; Network Portals & Portal Groups

**iSCSI LUN MASKING NAMING FORMATS:** Each name is treated as an individual iSCSI initiator

iqn: iqn.1992-05.com.emc:mytarget-1

eui:eui.5006048DC7DFB1AF (16 digits hex)

### **TARGET ACCESS POINTS FOR INITIATORS:**

--Network Portal consists of IP Address & TCP Port number

-Portal Group is similar to a SCSI port and determines iSCSI sessions (identified by tags)

### **iSCSI TARGET FUNCTION:**

→Targets receive initiator requests via TCP

→Targets extract the SCSI commands from the iSCSI PDU's

→Targets route SCSI task to LUNs for execution

→Targets send status, data, and sense data back to Host Initiator

### **CELLERRA iSCSI:**

→As an iSCSI Target, Celerra supports Windows and Linux iSCSI initiators, receiving iSCSI commands

→Windows 2000 & 2003 can install a MS iSCSI SW initiator, while this feature is native in Windows 2008 & Vista

→Initiators encapsulate SCSI cmd's, data, status info into iSCSI PDUs

→Initiators log in and establish iSCSI sessions with iSCSI Targets

→Initiators discover iSCSI LUNs from Target

→Initiators send PDU's to Target on the Storage Device

→Once Initiator logs into a Target, the LUNs associated with the Target appears as a local SCSI disk device on the Host

→Initiators log into IQN's and export physical drives for LUNs that are masked to Celerra

→All endpoints are iqns

→Source iqn is the name of the iSCSI initiator

→Target iqn is name of ‘target’ on Celerra, which acts like a SCSI device and exports LUNs

**Note:** Targets are represented by software on the Celerra, with IQN's in EUI format

### **NETWORK PORTALS:**

→Initiators & Targets use network interfaces, called network portals, to communicate

→Initiator is identified by IP address while Target is identified by both IP address & TCP listening port (iSCSI port 3260)

→Rule, must have unique IP Address for each Network Portal bound to an Initiator

### **PORTAL GROUPS:**

→Collection of one or more network portals identified by a common tag (used for session control)

### **iSCSI NODE NAMES:**

→Each iSCSI Initiator and iSCSI Target node uses a globally unique permanent name

→Identifies Initiator & Target

→Independent of hardware location

→Allows for free movement of adapters and port addresses

→Uses iSCSI iQN or Qualified Naming Authority

→Use with DNS domain name and name becomes reserved

### **iSCSI QUALIFIED NAMES & ALIASES:**

#### **iSCSI Qualified Name (IQN):**

iqn.1987-05.com.cisco.1234abcdef987601267da232.betty

[Type; Date; Naming Authority; Subgroup Naming Authority]

#### **EUI NAME EXAMPLE:**

eui.02004567a425678d

[Type; EUI Identifier, ASCII Encoded Hex]

iSCSI ALIAS:

Unicode Alias for the long IQN or EUI names referenced above, i.e., a plain name used to identify the name

### **CELERRA iSCSI LUNS:**

→ Celerra iSCSI luns are dedicated file storage objects created from within a regular file system that emulate a SCSI disk device via software emulation. On the iSCSI Host initiator, the iSCSI LUN appears as a local disk with individual files, but cannot see any other iSCSI LUNs that might be built on the same file system

→ Celerra supports virtually provisioned LUNs or regular dense LUNs, but not both on the same file system

→ Automatic file system extension feature is supported only for virtually provisioned iSCSI luns

→ iSCSI LUNs can be extended on Celerra, but the iSCSI Hosts must run utilities to discover the expanded iSCSI lun

→ Luns interpret SCSI CDBs and executes SCSI commands—Celerra uses software for processing SCSI commands

→ Celerra uses regular dense Luns, or PBR (Persistent Block Reservation) to ensure that there is always space available for the Lun—server\_df may report that the Lun is full, eventhough it has not yet been written to. PBR reserves space in a Reservation Pool, but disks are not taken from the pool until needed

### **Types of iSCSI lun objects within the Celerra file system:**

PLU, SLU, TWS

→ PLU are Production Logical Units, or iSCSI luns, normally addressed from 0-127 (try not to use lun # 128-254), visible to Host Initiator, space is persistently reserved in the file system, reported as used space by Celerra

→ SLU are Snap Logical Unit Point-in-time copies of the PLU using pointers to blocks of data, normally addressed from 128-254 as temporary luns, promoted to LUN status for access. Size reported as % of the PLU size between 0-100%.

→ TWS are Temporary Writable Snaps, or a RW version of the SLU, used during Snap promote and restore operations, is deleted when demoted

→ LUN size is restricted to 2TB

### **LUN MASKS:**

Masks are used to control access to LUNs on iSCSI Targets

### **CELERRA iSCSI DEVICE DISCOVERY:**

**iSCSI (Internet Storage Name Service):** Discovery & naming protocol for discovery of iSCSI devices on TCP/IP network

#### **Two Types of iSCSI Discovery can be achieved:**

→ Celerra operates as an iSNS Client for updates from the iSNS Name Service, where auto discovery of initiators, targets, portals, & portal groups occurs [iSNS Discovery]

→ Celerra can be manually configured using SendTargets to configure Target's network portal (Target iSCSI name, TCP port & IP address) for the Initiator [SendTargetsDiscovery] to establish a discovery session with iSCSI service on the Target.

### **CONFIGURING iSCSI LUNS FOR CELERRA (5.3-5.5):**

1. Create and mount file system on Data Mover

**Note:** File System is called virtual\_iscsifs

2. Enable iSCSI License in Celerra Manager (not required for CLI)

3. Create iSCSI Target on DM, providing qualified name of DM (Port 3260)

# server\_iscsi server\_2 -target -alias target1 -Q iqn.1992-05.com.emc:test -create 1000:np=10.241.169.44 (1000 = portal group id, port 3260 used by default from CLI)

WebUI>iSCSI>Targets>New>Choose Data Mover: server-2 >Name: target1>Network Portals: 10.241.169.44:3260 (DM IP and Port 3260)

4. Create iSCSI LUN mask for the Target, specifying DM, target, Initiator, and LUN range

# server\_iscsi server\_2 -mask -set target1 -initiator iqn.1991-05.com.microsoft:nas81.celerra1.emc.com -grant 0-100

WebUI>iSCSI>Targets>rightclick target Properties>LUN Mask>New>Initiator: iqn.1991-

05.com.microsoft:NAS81.celerra1.emc.com (Open MS iSCSI Initiator application and copy the iqn to the WebUI box)>Grant LUNs: 0-100

5. Create iSCSI LUN specifying DM, target, LUN #, File System, and Size of LUN

# server\_iscsi server\_2 -lun -number 0 -create target1 -size 100M -fs virtual\_iscsifs -virtually\_provisioned yes

**Note:** Cannot create virtually provisioned iSCSI LUN from WebUI—all other iSCSI luns can be created from WebUI. Create multiple LUNs under same Target ID. Virtually Provisioned LUNs only supported in 5.5 via RPQ.

6. Enable iSCSI service on DM and point to iSNS Server [Windows 2000 Server with iSCSI Initiator installed]

# server\_iscsi server\_2 -ns isns -server 10.241.169.100

# server\_iscsi server\_2 -service -start

WebUI>Server\_2>iSCSI>Configuration>iSCSI Enabled

**Note:** iSCSI LUN will now be visible in WebUI

7. Add CHAP authentication:

# server\_security server\_2 -add -policy chap -name iqn.1991-05.com.microsoft:nas46.celerra2.emc.com

8. Install Microsoft iSCSI Initiator driver on Win2k/Win2k3 system [free download] and configure using iSCSI Initiator Properties

9. Add Target Portal IP addresses of Data Mover Target:

a.) iSCSI Initiator Properties>Discovery>Target Portals Add>IP Address or DNS name: 10.241.169.44:3260

**Note:** If the target is already listed, proceed to next step

**Target Portals:**

10.241.169.44 3260

10. Go to Targets tab>highlight iSCSI Target>select Log On to Target: check the box for “Automatically restore this connection when the system boots”

**Note:** If the target is already listed in the Status column as “Connected”, it has already been added to the “Persistent Targets” tab, and so you cannot select the “Automatically restore this connection...” box again.

11. Create iSCSI Partitions:

a) Start>Run>Diskmgmt.msc>If wizard is presented to convert to dynamic disk, Cancel>Rightclick the Unknown/Not Initialized Disk number>Initialize Disk>disk is now seen as Basic Online volume

b) Rightclick the Unallocated partition>New Partition>select Primary Partition>Confirm partition size>Assign a drive letter>Format this partition with the following settings: NTFS, FAT, FAT32, Provide a Volume Label designation, then select checkbox for Perform a quick format [Partition is formatted and marked Healthy]

## **DELETING iSCSI LUN:**

# server\_iscsi server\_2 -lun -list

server\_2 :

target: fs2\_lun  
lun size(MB) filesystem  
1 15360 fs2\_lun ( id=26 )

# server\_iscsi server\_2 -target -list

server\_2 :

alias\_name portal\_groups iscsi\_name  
fs2\_lun 1 iqn.1992-05.com.emc:sl7e10807001210000-1

# server\_iscsi server\_2 -lun -delete 1 -target fs2\_lun

server\_2 : done

## **THINGS TO KNOW ABOUT CONFIGURED SYSTEM:**

→Cannot change target's qualified name after creation

→Network portal modifications do not affect active sessions between Target & Initiator

→Deleting an iSCSI target deletes target and associated LUN masks, and deregisters from iSNS

→Cannot delete target, however, if iSCSI Initiator logged in or LUN is configured (delete all LUNs first)

**Async Param Problem:** 5.5.29.1

Default value for AsyncEvent is 1 to enable async event logging by the target.

When set to 0, the target sends async logout messages, and during addition of new iSCSI luns, a LUN Inventory is triggered which makes all the Celerra iSCSI luns unavailable.

Corrective action is to set param iscsi AsyncEvent=1

## **CELLERRA ISCSI FILES FOUND IN /etc:**

```
-rw-r--r-- 1 root bin      535 Jun  8 22:33 iscsi.conf (most importance config file, read by iSCSI svc on startup)
-rw-r--r-- 1 root bin      116 Jun  8 20:46 iscsi_initiator.conf
-rw-r--r-- 1 root bin      172 Jun  8 22:33 iscsi_lun.conf
-rw-r--r-- 1 root bin      270 Jun  8 20:46 iscsi_lunmask.conf
-rw-r--r-- 1 root bin       7 Jun  8 17:05 iscsi_pggroups.conf
-rw-r--r-- 1 root bin      21 Jun  8 17:05 iscsi_portals.conf
-rw-r--r-- 1 root bin      49 Jun  8 17:05 iscsi_targets.conf
-rw-r--r-- 1 root bin     210 Jun 23 21:24 chapdb
```

**drwxr-xr-x 2 root bin 1024 Jun 23 21:24 nbsdb →NBS directory**

**Note:** iSCSI LUNs are listed in this directory

```
lrwxr-xr-x 1 root bin      58 Jun  8 20:46 fs213_T1_LUN0_APM00050802675_0000 fsid=213 vlu=Disk
2675_0000 -> path=/fs213_T1_LUN0_APM00050802675_0000 fsid=213 vlu=Disk
lrwxr-xr-x 1 root bin      58 Jun  8 22:33 fs213_T1_LUN1_APM00050802675_0000 fsid=213 vlu=Disk
2675_0000 -> path=/fs213_T1_LUN1_APM00050802675_0000 fsid=213 vlu=Disk
lrwxr-xr-x 1 root bin     15 Apr 25 18:19 .nbsdbversion -> nbsdb version=5
```

**# cat isesi.conf**

```
# 5
iscsi target action=create name=vm iqn=iqn.1992-05.com.emc:apm000508026750000-1sn=1 1:np=137.148.5.28:3260
```

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
iscsi target action=mask name=vm initiator=iqn.1991-05.com.microsoft:auslander.csunet.csuohio.edu grant=0-1  
iscsi target action=mask name=vm initiator=iqn.1991-05.com.microsoft:outlander.csunet.csuohio.edu grant=0-1  
iscsi lun action=create number=0 target=vm fsid=213 size=75620 path=/fs213\_T1\_LUN0\_APM00050802675\_0000  
iscsi lun action=create number=1 target=vm fsid=213 size=75550 path=/fs213\_T1\_LUN1\_APM00050802675\_0000

### # cat iscsi\_initiator.conf

```
0    iqn.1991-05.com.microsoft:auslander.csunet.csuohio.edu
1    iqn.1991-05.com.microsoft:outlander.csunet.csuohio.edu
```

### # cat iscsi\_lun.conf

```
0    0    213    75620  0    /fs213_T1_LUN0_APM00050802675_0000fs213_T1_LUN0_APM00050802675_0000
0    1    213    75550  0    /fs213_T1_LUN1_APM00050802675_0000fs213_T1_LUN1_APM00050802675_0000
```

### # cat iscsi\_pggroups.conf

```
0    1    0
```

### # cat iscsi\_portals.conf

```
0    137.148.5.28  3260
```

### # cat iscsi\_targets.conf

```
0    vm    iqn.1992-05.com.emc:apm000508026750000-1      1
```

## CELLERRA iSCSI COMMANDS:

→ Celerra is the “target” and the Windows host is the “initiator”

### # server\_iscsi server\_2

server\_2 :

Error 2100: usage: server\_iscsi { <moveovername> | ALL }

```
-target <target_options>
| -lun <lun_options>
| -mask <lun_mask_options>
| -ns <name_service_options>
| -service { -start | -stop | -status }
| -snap <snap_options>
| -help
```

To get detailed options for target, lun, lun mask and name service, please type "-help" or only type "-target", "-lun", "-mask", "-ns" and "-snap" respectively after typing in the server name.

### # server\_iscsi server\_2 -target

```
-target {
-alias <alias_name> [-Q <iscsi_name>]
-create [<pg_tag>:np=<np_list> [<pg_tag>:np=<np_list> ...]]
| -delete <alias_name>
| -rename <old_alias_name> <new_alias_name>
| -bind <alias_name> <pg_tag>:np=<np_list>
| -unbind <alias_name> { <pg_tag> | np=<np_list> }
| -info { <alias_name> | -all }
| -stat { <alias_name> | -all }
| -list
```

### # server\_iscsi server\_2 -lun

```
| -lun {
-number <lun_number> -create <target_alias_name> -size <size>[M|G|T]
-fs fs_name [ -vp {yes|no} ] [ -readonly {yes|no} ]
| -extend <lun_number> -target <target_alias_name> -size <size>[M|G|T]
| -delete <lun_number> -target <target_alias_name>
| -info { <lun_number> | -all } [-target <target_alias_name>]
| -stat { <lun_number> | -all } [-target <target_alias_name>]
| -modify <lun_number> -target <target_alias_name> -readonly {yes [-Force] | no}
| -list [-target <target_alias_name>]
```

### # server\_iscsi server\_2 -mask

```
| -mask {
-list [<target_alias_name>]
| -info <target_alias_name> {-initiator <initiator_name> | -all }
| -set <target_alias_name> -initiator <initiator_name>
{ -grant <access_list>
| -deny <access_list> }
```

| -clear <target\_alias\_name> -initiator <initiator\_name>

## # server\_iscsi server\_2 -ns

| -ns isns

{ -info

| -set { -server <IP>[:<port>] | -esiport <port> }

| -clear { -server | -esiport | -all }

## # server\_iscsi server\_2 -service -status

server\_2 :

\*\*\*\* iSCSI Service \*\*\*\*

Status: Running

iSNS is configured to listen on ports:

TCP Port: 3260 -----output abridged-----

## **SETTING iSCSI LUN AS READ-ONLY FOR A TARGET:**

# server\_iscsi server\_2 -lun -modify 10 -target cel1 -readonly yes -Force

server\_2 : done

## # server\_iscsi server\_2 -lun -info 10

server\_2 :

Logical Unit 10 on target cel1:

(Production) fsid=31 size=5120MB alloc=134MB dense read-only

path=/cel1\_10/fs31\_T1\_LUN10\_APM00071600514\_0000/fs31\_T1\_LUN10\_APM00071600514\_0000

replication=available destination

max\_extension\_size=927MB

Healthy

**Note:** Not sure how effective the RO setting is. I did get a “write-protected” popup when trying to copy new files to the iSCSI lun, but in many cases, some folders and files were shown as copied before the popup error appeared. Bottomline, further testing showed that only rebooting the Data Mover would prevent files from being copied. Furthermore, after rebooting, the iSCSI lun was actually not accessible at all from the Windows Host, meaning that RO actually means no Host access. And, once the iSCSI lun was set back to RW, none of the files that were previously thought to have been copied over, were actually there. So, some quirks about this behavior.

## **CREATING iSCSI TARGETS WITHOUT & WITH QUALIFIED NAME:**

\$ server\_iscsi server\_2 -target -alias target1 -create 1000:np=192.168.25.122,192.168.25.123

\$ server\_iscsi server\_2 -target -alias target2 -Q iqn.1992-05.com.emc:test -create

## **CREATE LUNS on iSCSI TARGET:**

\$ server\_iscsi server\_2 -lun -number 2 -create target1 -size 1000 -fs iscsi02 (creates LUN 2)

## **SET LUN MASK ON TARGET:**

\$server\_iscsi server\_2 -mask -set target1 -initiator iqn.1991-05.com.microsoft:nas46.celerra6.emc.com -grant 0-100

## **SETTING iSNS SERVICE:**

\$ server\_iscsi server\_2 -ns isns -set -server 192.168.25.46

## **STARTING & VERIFYING iSCSI/iSNS SERVICE:**

\$ server\_iscsi server\_2 -service -start | -status

\$ server\_iscsi server\_2 -ns isns -info (iSCSI Client Service Status)

\$ server\_iscsi server\_3 -ns isns -info

server\_3 :

iSNS Client Service Status

-----

service : stopped

server : UNKNOWN →This is normal if customer is NOT running the Name Service for iSCSI

entity : UNASSIGNED

source :

ESI support : TRUE

ESI port : UNASSIGNED

timeout : 3000

trace : FALSE

PG support : TRUE

# server\_iscsi server\_2 -service -status (iSCSI Service Status)

server\_2 :

\*\*\*\* iSCSI Service \*\*\*\*

Status: Running

iSCSI is configured to listen on ports:

TCP Port: 3260

Header Digests are preferred

Data Digests are not preferred

### **iSCSI SERVICE LISTED IN LOCAL NETD FILE:**

\$ cat /nas/server/slot\_2/netd

**iscsi start**

**iscsi hiwat=0x60000 →TCP High Watermark reception tweak for clients using low MTU's with iSCSI—AR73389**

### **VERIFYING TARGET LIST & QUERYING SPECIFIC TARGET:**

**\$ server\_iscsi server\_2 -target -list**

server\_2 :

alias\_name portal\_groups iscsi\_name

SQL1\_Data 1 iqn.1992-05.com.emc:apm000520060760000-2

**\$ server\_iscsi server\_2 -target -info SQL1\_Data**

server\_2 :

Target Alias Name: SQL1\_Data

Target iSCSI Name: iqn.1992-05.com.emc:apm000520060760000-2

Serial Number: 2

Portal Group Tag: 1

Portal: 10.1.1.6:3260

Logical Units:

0 : (Production) fsid=24 size=102400MB alloc=3182MB dense path=/SQL1\_Data/fs

24\_T2\_LUN0\_APM00052006076\_0000

Connected Initiators:

iqn.2000-04.com.qlogic:qla4010.fs20520b01057

**\$ server\_iscsi server\_2 -t -s -a**

server\_2 :

Target Alias Name: 1850r8-iscsi

Target iSCSI Name: iqn.1992-05.com.emc:apm000424036380000-4

[Login/Logout Stats]

Login Accepted: 2

Login Failed: 0

Redirect: 0

Authorize: 0

Authenticate: 0

Negotiate: 0

Other: 0

Logout Normal: 0

Logout Other: 0

Last Login Failed: N/A

[Session Stats]

Initiator: iqn.1991-05.com.microsoft:epvt1850r-8.svtw2k.hop.emc.com TSIH: 1 ISID

: 400001370001

Command PDUs: 343

Read Command PDUs: 267

Write Command PDUs: 76

Response PDUs: 343

Data-In PDUs: 0

Data-Out PDUs: 0

R2Ts: 0

Tx Data Octets: 344806

Rx Data Octets: 317952

Digest Errors: 0

Connection Timeout Errors: 0

### **VERIFYING LUN LIST & QUERYING:**

**\$ server\_iscsi server\_2 -lun -list**

server\_2 :

target: SQL1\_Data

lun size(MB) filesystem

0 102400 SQL1\_Data ( id=24 )

**\$ server\_iscsi server\_2 -lun -info 0**

server\_2 :

Logical Unit 0 on target SQL1\_Data:

(Production) fsid=24 size=102400MB alloc=3182MB dense

path=/SQL1\_Data/fs24\_T2\_LUN0\_APM00052006076\_0000

**\$ server\_iscsi server\_2 -lun -info -all -target target1**

**\$ server\_iscsi server\_2 -lun -s -a**

server\_2 :

Logical Unit 0 on target 1850r8-iscsi:

(Production) fsid=25 size=1000MB alloc=0MB dense

path=/1850r8/fs25\_T4\_LUN0\_APM00042403638\_0000/fs25\_T4\_LUN0\_APM000

42403638\_0000 (snapped)

replication=source

max\_extension\_size=11787MB

Statistics:

Total SCSI CDBs: 136

Read: 63 Failed: 0 Blks read: 338

Write: 28 Failed: 0 Blks written: 248

## **VERIFYING LUN MASK LIST & QUERYING:**

**\$ server\_iscsi server\_2 -mask -list**

server\_2 :

target: SQL1\_Data

initiator\_name grant LUNs

iqn.2000-04.com.qlogic:qla4010.fs20520b01057 0

**\$ server\_iscsi server\_2 -mask -info SQL1\_Data -initiator iqn.2000-04.com.qlogic:qla4010.fs20520b01057**

server\_2 :

Initiator Lun Mask:

Initiator Name: iqn.2000-04.com.qlogic:qla4010.fs20520b01057

grant=0

**Note:** Where ‘SQL1\_Data’ is the alias name for the Target

**\$ server\_iscsi server\_2 -mask -list target1**

**\$ server\_iscsi server\_3 -mask -list**

server\_3 : no iscsi lun masking entry

## **ADDING CHAP SECURITY:**

**\$server\_security server\_2 -add -policy chap -name iqn.1991-**

**05.com.microsoft:auslander.csunet.csuohio.edu**

**# server\_security server\_2 -info -policy chap** (displays CHAP security info)

server\_2 :

chapdb name=iqn.1991-05.com.microsoft:auslander.csunet.csuohio.edu pass=\*\*\*\*\*

chapdb name=iqn.1991-05.com.microsoft:outlander.csunet.csuohio.edu pass=\*\*\*\*\*

**Note:** Use GUI to verify Portals, Available Targets, Active Sessions, Configure Partitions.

/slot\_2/etc

**# cat chapdb**

chapdb name=iqn.1991-05.com.microsoft:auslander.csunet.csuohio.edu pass=3539615135133561513D233719173D17

chapdb name=iqn.1991-05.com.microsoft:outlander.csunet.csuohio.edu pass=3539615135133561513D233719173D17

## **CELLERRA iSCSI HOST SOLUTIONS:**

→Exchange 2000 Integration with SnapSure module

→SnapSure Manager for Celerra iSCSI LUNs, using Client-side software to manage iSCSI LUN Snaps

→Eng. White Paper exists that describes possibility of using Celerra iSCSI luns as repository for boot images for Network diskless clients using PXE. Clients would boot using an iSCSI HBA [QLogic QLA4010—need to download specific bios separately] to boot

Windows 2003 Clients from the Celerra iSCSI LUN. Basically, the HBA bios establishes the target location of the boot image, establishes iSCSI session with Celerra target, sets the iSCSI LUN as the boot device, at which time the Windows system recognizes the boot device and goes through normal boot as if the LUN were local. Another boot method would be to use a separate PXE Server for the Client systems to download a winBoot/I bootstrap into RAM, at which point the program executes and connects to the iSCSI LUN. See “iSCSI Booting...” White Paper from Powerlink for more information.

### **EXCHANGE 2000 INTEGRATION:**

- Configure MS Exchange 2000 to use Celerra iSCSI LUNs for Exchange Logs and Databases
- Migrate Storage groups to Celerra iSCSI LUNs via wizard
- Use SnapSure to Restore & Delete storage group snaps, Delete Storage group Store, Promote Storage group snap by assigning iSCSI LUNs to make accessible to remote systems, Demote Storage group snap, Schedule a Storage group snap.
- Best Practices Guide suggests that Windows Server timeout for Exchange be set to 600 secs from 40 secs

### **EXCHANGE SNAPSURE LIMITATIONS:**

- Do not use MS Terminal Services to access SnapSure Manager
  - SnapSure temporarily mounts Exchange when creating iSCSI Snaps
  - Do not schedule Snaps during ‘At System Startup’ or ‘At Logon’
  - Each Information Store and Log Files for Exchange must be on separate iSCSI Disks
- LU 1 →Transaction Logs and System Files  
LU 2 →Priv Information Store  
LU 3 →Pub Information Store

### **QLOGIC 4010 MODEL iSCSI HBA:**

When using Qlogic 4010 iSCSI initiators, ensure min NAS 5.3.15.306. On Windows 2003 host, ensure the MS iSCSI Service is installed. Ensure initiator name is the same in both MS iSCSI GUI and in Qlogic Sansurfer. Run c:>mountvol /E to enable automount of promoted snaps on Windows in order to assign drive letters dynamically. LUN numbers cannot be >than 128 for this feature. Set the following param to 0 to disable ReportLunsOverflow:

**param iscsi ReportLunsOverflow=0**

### **INSTALLING CELERRA iSCSI APPLICATIONS:**

- SnapSure for Exchange 2000
  - SnapSure Manager for iSCSI
  - Celerra Logical Disk Service (CLD)
1. Ensure that Windows Host has Microsoft iSCSI Software Initiator installed first
  2. Install EMCCLDPack\_1\_4\_8.exe & select complete setup
  3. Configure DCOM Access permissions during InstallShieldWizard [add System and Administrator accounts to CLD service]
  4. Set Authentication Level to “Packet” in the General properties tab for CLD Service
  5. Set security access permissions in the Security properties tab
- \*Use custom access permissions and add SYSTEM & Administrator, and any other accounts required, to the service

### **USING SNAPSURE MANAGER for iSCSI:**

**Purpose:** Create Snap, Promote iSCSI Snap to SLU, on LocalHost or RemoteHost, Demote an iSCSI SLU, Restoring PLU from iSCSI Snap, Delete iSCSI Snap. Scheduling iSCSI Snaps

Programs>Celerra Tools>SnapSure Manager for iSCSI>should be able to see all Celerra PLUs (Production Logical Units)

1. Create Snap of Celerra iSCSI LU (Logical Unit) using SnapSure Manager for iSCSI
2. Rightclick iSCSI Logical Unit and choose new
3. Promote the iSCSI snap to an SLU on the local or a remote Host

**Note:** Celerra will automatically add the lun mask for the host. Use Explorer, Disk Manager, MS iSCSI Initiator programs to view SLUs

4. Verify promoted SLU from Celerra:

**# server\_iscsi server\_2 –target –info <target\_name>**

**Note:** SLU can be mounted and used as a Temporary Writeable Snap (TWS), but it's important to know that when the snap is demoted, the TWS information is lost!

5. Use EMC SnapSure Manager for iSCSI>Snap>Demote to demote the SLU

### **EXCHANGE MANAGEMENT iSCSI EXTENSIONS:**

- Migrate Exchange Stores and Logs to Celerra iSCSI Luns
- Create SnapSure Backups of Exchange Storage Groups [Create iSCSI Snap, Promote, Backup the SLU]
- Conduct Single Mailbox Restore using offline ‘Recovery Server’
- Conduct Point-in-Time or Up-to-the-Present Restores

## **GENERAL iSCSI SNAPSURE LIMITATIONS:**

- Snaps must not be in ‘Promoted’ Status
- PLU cannot be in “Restoring” status [Admins must quiesce PLU applications first]
- Snaps are created temporarily until changes are committed to PLU
- Snaps reserve space for allocated blocks on the source LUN and are called “sparse files” (Production LUNS are “dense files”)

## **iSCSI RESTRICTIONS/RULES/BEST PRACTICES:**

- Do not allow multiple initiator access to the same iSCSI Lun (but not a hard requirement)
- MS Initiators only support single connections per session
- Use IQN numbers in lowercase
- Max iSCSI LUN size NAS 5.3 is 1TB
- Max iSCSI LUN size NAS 5.4-5.6 is 2TB
- Linux iSCSI initiator’s can only work with 32 LUNS or less
- Do not use Lun #’s greater than 128-254 for PLU’s
- Do not run nas\_fsck on iSCSI luns without Eng. approval
- iSCSI Snapshots are only useful for DR purposes if saved off to tape
- Do not stop iSCSI Service if Initiators are logged in
- Multiple iSCSI luns can be configured on a single Celerra file system
- Initiators use only a single TCP connection to the Target
- iSCSI Luns are not supported for VDMs
- Checkpoints, IP Replications, Cava, MPFS Highroad, & DHSM are features not supported on Celerra iSCSI file systems
- IPsec & Dynamic Disks are not supported
- Native iSCSI supported for Celerra, not ‘bridging’ topologies
- MSCS (MS Cluster Service) is supported with limitations (no multiple connections per session)
- Both MS Exchange & SQL Server application timeout values on Host Servers should be set to 600 secs.

## **KNOWN iSCSI ISSUES:**

- NAS 5.5 Upgrade issue if iSCSI LUN doesn’t have 100MB free space, the lun will become totally corrupt and unavailable—needs enough space to write out new directory structure (see emc141515)

## **TROUBLESHOOTING iSCSI:**

→ For Windows systems, install DebugView Utility--download from [www.sysinternals.com/files/dbgvnt.zip](http://www.sysinternals.com/files/dbgvnt.zip)

**HKLM>Software>EMC>CLDSvc>Configuration**>Set Debug & Verbose values to 1 & reboot Windows System

→ Set following Debug Logging on Celerra [Log facilities are ISCSI, NBS, VLU, & IPSNS]

**# .server\_config server\_2 -v "logsys set severity ISCSI=LOG\_DEBUG" | NBS | VLU | IPSNS**

**Note:** Remember to turn off debug logging when done

**# .server\_config server\_2 -v "logsys set severity ISCSI=LOG\_PRINTF"**

→ Network Trace: Filter on Port 3260

## **USING HIDDEN iSCSI LOG:**

**# .server\_config server\_2 -v "iscsi log"**

Commands are:

pc → print current page  
pw → print page at write position  
pn → print next page  
pp → print previous page  
pw → print write position page  
pg=n → print page n  
nt → do not show time stamps  
ts → show time stamps  
cl → clear event log  
dis → disable logging  
en → enable logging

## **SET FOLLOWING PARAM TO GATHER ISCSI SESSION STATISTICS:**

**param iscsi.CollPerfStats=1**

**\$ .server\_config server\_x -v "iscsi eng showsessions"**

\*\*\* IscsiSessions \*\*\*

Sessions for Target: iqn.1992-05.com.emc:apm000433060400000-25

TPGT: 1 Initiator: 0 Name: iqn.1991-05.com.microsoft:eng16951 TSIH: 3 ISID: 40000

SessionType : Normal

MaxConnections : 4

MaxBurstLength : 262144

FirstBurstLength : 65536  
ImmediateData : Yes  
InitialR2T : No  
DefaultTime2Wait : 2  
DefaultTime2Retain : 60  
MaxOutstandingR2T : 16  
DataPDUInOrder : Yes  
DataSequenceInOrder : Yes  
ErrorRecoveryLevel : 0

Connections:

CID: 1 I: 172.24.169.51/1517 T: 10.169.0.15/3260  
HeaderDigest : CRC32C  
DataDigest : NONE  
MaxRecvDataSegmentLength(I) : 65536  
MaxRecvDataSegmentLength(T) : 65536

Total SCSI Command PDUs: 977592

Total SCSI Read Command PDUs: 637245

Total SCSI Write Command PDUs: 332111

Total R2Ts: 0

Total SCSI Response PDUs: 977591

Total Tx Data Octets: 1312311432

Total Rx Data Octets: 1730236928

Timing & Statistics:

Rolling Sampling Window: 53 sec (44% full)

Total SCSI Cmds: 59242

**Average iSCSI/SCSI Time: 182 usec (Time it takes Data Mover to complete an iSCSI command)**

Command Out of Order: 0%

Command Cluster:

Cluster Factor: 1.00

Max Cluster: 1

Command Execution Breakup (usec):

Conn Recv: 3, Conn Prep: 0

Sess Prep: 1

SCSI Prep: 1, SCSI Exec: 170

Sess Rply: 0, Conn Wait: 0

Conn Send: 3, Sess Post: 0

Outstanding Tasks:

Average number of tasks: 0.49

Normalized Std Dev (R^2): 0.01

## iSCSI ENG USAGE:

**# .server\_config server\_2 -v "iscsi eng"**

Usage: iscsi eng cmd

Where cmd is one of:

dumpdb [db=target|lun|portal|pgroup|lin|lunmask]

dumpviewfile - displays the view file

logout [name=aliasname] - logout target sessions

makeviewfile - generates a new version of the view file

setviewversion version=num - Sets the view file version

showluns [name=aliasname] - displays logical units

showmasks [name=aliasname] - displays target lun masks

showportals - displays portals

showsessions [name=aliasname] - displays target sessions

showtargets - displays a list of targets

targetinfo [aliasname] - displays target information

threads [add|kill=thrs] - add, kill, or display # of iSCSI threads

## BASIC TROUBLESHOOTING STEPS:

1. Verify IP connectivity
2. Check iSCSI service on DM and Initiator service on Host

3. Verify Initiator is logged into target
4. Verify Target & Initiator portal settings
5. Verify Target LUN mask, name, and access granted
6. Verify Initiator & Target iSCSI names for typos
7. Verify CHAP configuration
8. Verify status of Backend
9. Set debug logging on Celerra

## **REPAIRING/RECOVERING iSCSI CONFIGURATION FILES:**

1. The iSCSI configuration is located on the Control Station in /nas/server/slot\_x/iscsi file and on the Data Mover in the /etc/iscsi.conf file

**Note:** If for some reason, either of the locations is missing its complete file, then copy the file from the good location to restore, or find a valid copy in the NAS\_DB Backup

2. If the file is missing or corrupted in both locations, and the Data Mover has NOT been rebooted, then we can regenerate the iSCSI configuration by using the following command—simply copy the file to the proper location:

**# .server\_config server\_2 -v “iscsi eng makeviewfile”**

1126894703: UFS: 6: inc ino blk cache count: nInoAllocs 1: inoBlk a5076b04

1126894703: UFS: 6: inc ino blk cache count: nInoAllocs 2: inoBlk a5076a84

1126894703: ADMIN: 4: Command succeeded: iscsi eng makeviewfile

3. If various other iscsi\_\* files are corrupted, but the iscsi.conf or iscsi file is known to be good, do the following to rebuild the appropriate files:

a. Delete all iscsi\_\* files from .etc

b. Reboot data mover to clear cache

c. Replay the iscsi or iscsi.conf file to create the other files by doing the following, then associate all LUNs to Targets:

**#.server\_config server\_2 -v “iscsi target action=create name=itang iqn=iqn.1992-**

**05.com.emc:apm000517019120000-1 sn=1 1:np=10.6.3.203:3260”**

**#.server\_config server\_2 -v “iscsi target action=mask name=itang initiator=iqn.1991-**

**05.com,Microsoft:itangpc grant=1-2”**

Modify lun action=create line by removing fsid= & size= entries, and enter correct path=

Associate Lun to target itang, LUN id=2 using:

**# .server\_config server\_2 -v “iscsi lun action=create number=2 target=itang**

**path=/itangfs1/fs29\_T1\_LUN2\_APM00051701912\_0000”**

**\$ .server\_config id=1 -v “iscsi lun action=create number=0 target=dm2**

**path=/dm2\_iscsi0/fs44\_T1\_LUN0\_20051214\_165301”**

4. If more extensive recovery is required, engage EE for assistance

5. If iSCSI Luns can no longer be accessed via iSCSI protocol, but data must be obtained, use following Windows Utility to access:

a. Download Virtual Disk Driver from chitchat.at.infoseek.co.jp/vmware

b. c:\>vdk.exe install /auto

c. Start driver & verify: c:\>vdk.exe start c:\>vdk.exe driver

d. Access iSCSI file system on Server from CS0 and copy the LUN within the same file system, but as a different name

#cp fs29\_T1\_LUN2... fs\_recover

e. Map drive to Share containing the new LUN filename and read the file

c:\>vdk.exe view fs\_recover

f. Open file as a device and start copying data to safe location

c:\>vdk.exe open \* fs\_recover

## **PERSISTENT BLOCK RESERVATION:**

iSCSI LUNs use PBR to preallocate blocks for the LUNs. Use following command to verify PBR:

**\$ .server\_config server\_2 -v “file pbrcheck ufs checkonly <vol\_id\_of\_fs>”**

## **iSCSI PERFORMANCE BEST PRACTICES:**

→Use RAID 1 configurations for random I/O (SQL, Exchange)

→Use RAID 5 when access is via sequential I/O

→Configure DAE's as either all-Random I/O shelves or all-Sequential shelves

→Use AVM or MVM, depending on needs and complexity of disk & metavolume layout

→Use Dense LUNs for critical data

→Use sparse LUNs using file system auto extension and monitor

→Use snaps to restore data or to promote a Snap to PLU status

## **CELERRA DATABASE & VOLUMES CONFIGURATION:**

### **COMMAND OUTPUTS/TOOLS NAS 5.3 & HIGHER:**

**# nas\_fs -i fs03**

```
id      = 21
name    = fs03
acl     = 0
in_use  = True
type    = uxf
volume  = v111
pool    = symm_std
member_of = root_avm_fs_group_1
rw_servers= server_5
ro_servers=
rw_vdms =
ro_vdms =
stor_devs = 000187880333-0772,000187880333-0773,000187880333-0774,000187880333-0775,000187880333-0776,000187880333-0777,000187880333-0778,000187880333-0779
disks   = d3,d4,d5,d6,d7,d8,d9,d10
disk=d3 stor_dev=000187880333-0772 addr=c0t110-40-0 server=server_5
----output abridged-----
```

**# nas\_fs -E fs03 →Breaks down components used to build the File System**

```
0:befs:21
21:fs03:0:y:1:111:3:::0::19:          →File System fs03, fsid=21, client meta=v111
111:v111:0:y:3:2:21:110:             →Meta=v111, volumeID=111, sliceID=110 from volume list
110:s79:0:y:1:1:111:79:              →Slice=s79, volumeID=110
79:s79:0:y:100:110:10340:88156:180543488: →Slice s79 is a member of v100 stripe
100:v100:0:y:2:0:73,74,79:64:3,4,5,6,7,8,9,10: →Stripe VolumeID=100, contains d volumes 3-10
10:d10:0:y:4:1:100:10:           →dVolumeID=10 from nas_disk list, member of Stripe Volume 100
10:d10:0:y:12312:000187880333:1,2,3,4:10:0779:1: →dVolume is seen by Servers 2-5 [1,2,3,4]
```

**\$ nas\_disk -i d10 -opt all →Runs symcli commands to output all relevant information on the indicated volumes**

```
id      = 10
name    = d10
acl     = 0
in_use  = True
size (MB) = 54476
type    = CLSTD
stor_id = APM00023800386
stor_dev = 0013
volume_name = d10
servers = server_2,server_3
server = server_2      addr=c0t113
server = server_2      addr=c16t113
server = server_3      addr=c0t113
server = server_3      addr=c16t113
clarid      = APM00023800386
devname      = 0019
uid         = 60:6:1:60:79:6F:A:0:A5:42:8:43:2E:B:D9:11
RAID type   = RAID5
disk_group  = 0002
stripe_size = 128
offset      = 0
capacity    = 111568486
owner       = A
default_owner = A
auto_trespass = DISABLED
auto_assign   = ENABLED
write_cache   = ENABLED
```

```
read_cache      = ENABLED
variable_length_pf = ENABLED
idle_delay_time   = 20
idle_threshold    = 0
prefetch_idle_count = 40
max_prefetch     = 4096
prefetch_disable  = 4097
write_aside_size  = 1023
state            = BOUND
is_private       = DISABLED
failed_over      = False
spindle_type     = FC
bus              = 0
```

**\$ ./dbchk -p** →NAS\_DB consistency checker

nas version is 5.3

Checking probe scsi and database consistency...

**Note:** Earlier versions of 5.3 and 5.4 dbchk did not detect all problem conditions. There is a new dbchk and dbchk.jar file with NAS 5.4.21.4 that is far more robust and is available on DMS website

**\$ server\_devconfig server\_2 -l -s -fs fs1**

server\_2 :

Scsi Disk Table

|      | Director | Port |      |     |     |                |          |
|------|----------|------|------|-----|-----|----------------|----------|
| name | addr     | num  | type | num | sts | symm_id        | symm_dev |
| d16  | c16t118  |      |      |     |     | APM00023800386 | 0018     |
| d16  | c0t118   |      |      |     |     | APM00023800386 | 0018     |

**\$ server\_devconfig server\_2 -p -s -fs fs1**

server\_2 :

SCSI disk devices :

tid/lun= 1/8 type= disk sz= 54476 val= 16 info= DGC RAID 5 02071800180018NI

**\$ nas\_symm -l**

| id | acl | name         | serial_number |
|----|-----|--------------|---------------|
| 1  | 0   | 000187721055 | 000187721055  |

**# /nas/symcli/bin/symcfg verify -sid 55** [Use to verify Symm config to Celerra database]

Wed Sep 7 15:55:33 CDT 2005

The Symmetrix configuration and the database file are NOT in sync.

### **ADDITIONAL TOOLS ADDED 5.3.25.0/5.4.22.0/5.5:**

**# /nas/sbin/get\_clariion\_errors**

Collecting Clariion backend errors for the last 30 days

Errors for Clariion APM00023700172:

Service Processor A\_APM00023700172:

Found 1651 messages for SP A\_APM00023700172, written to the Event Monitoring log during the last 30 days

There were 0 unmatched messages in the Event Monitoring log of SP A\_APM00023700172

2,critical,A\_APM00023700172,904,1, Enclosure 0 Power B VSC Shutdown/Removed [0x04] 0 0

3,critical,A\_APM00023700172,906,1, Enclosure 1 Disk 5 Unit Shutdown [0x00] c80006 c80014

3,critical,A\_APM00023700172,908,1, SP A Fault - Cache Disabling [0x00] 0 0

1,critical,A\_APM00023700172,920,1, Enclosure 1 Disk 5 Hard Media Error [0x3f]0 0

2,critical,A\_APM00023700172,a07,1, Enclosure 1 Disk 5 CRU Powered Down [0x00]0 920c

1,warning,A\_APM00023700172,801,1, Enclosure 2 Disk 12 Soft SCSI Bus Error [0x07] 0 0

**Note:** Purpose of tool is to support monthly SYR reporting (called by log\_config), but can be used as standalone tool to collect last 30 days of SP errors--informational, warning, critical

**# /nas/tools/.whereisfs -all**

| RG    | FS's  |
|-------|-------|
| ----- | ----- |

APM00023801040-0008 [ 2] fs04 (d22) fs\_quota (d22)

FS Resources (RGs: [ total # of RG] {repeated for each RG} <RG\_ID> <LUNs used in the RG>)

-----

fs\_quota RGs: [ 4] APM00023801040-0016; LUNs:

RAID Groups in use:

RG LUN (dVols) FS list

-----  
APM00023801040-0008 0013 (d22 ) fs04 fs\_quota**Note:** Purpose of command to output file system info, Raid Groups, and Luns mapped to file systems, as well as fs, volumes, luns mapped to Raid Groups—Clarion backends only at this point**REPLACING SYMCLI LIBRARIES FILES ON CELERRA CONTROL STATION:****Caution:** Do not implement without TS2/Engineering approval. This procedure may be required to complete an SRDF Restore or a failing NAS Upgrade.

1. # cd /nas/opt/emc/SYMCLI/V6.0.0/shlib

2. Tar existing contents and copy to safe location:

# tar -cvf symcli600lib\_files.tar \*

3. Copy replacement symcli library files into /nas/opt/emc/SYMCLI/V6.0.0/shlib

# cp -ip symcli600lib\_files.tar /home/nasadmin

4. Untar the new files while in the /nas/opt/emc/SYMCLI/V6.0.0/shlib directory

# tar zxvf 60111.tar.gz

-r-xr-xr-x root/root 270475 2005-08-17 17:36:14 libstorpd.so

-r-xr-xr-x root/root 13264343 2005-08-17 17:36:13 libsymapi.so

-r-xr-xr-x root/root 805569 2005-08-17 17:36:13 libsymlvm.so

5. Verify Link on following file:

# cd /usr/lib

# ls -la

lrwxrwxrwx 1 root root 31 Sep 15 02:28 libstorpd.so -&gt; /usr/symcli/shlib/libstorpd.so

**Note:** If link does not exist, create link in /usr/lib directory:

#ln -s libstorpd.so /usr/symcli/shlib/libstorpd.so

6. Verify that symcli or symcfg commands work

7. Conduct # nas\_rdf –restore

**Note:** See Primus emc120887 & AR69029 for more details on Solutions Enabler Version 6.0.0 problem and Celerra SRDF operations**RECONCILING CELERRA AND SYMMETRIX DEVICE DATABASES:****VERIFYING WHAT DEVICES CELERRA SEES:****\$ nas\_diskmark -l -all** [Lists all SYMM devices, sizes, TID's, & GateKeepers for Celerra; same as devconfig]**\$ nas\_diskmark -m -all** [Use this to diskmark all disks for all servers]**\$ server\_devconfig server\_x -l -s -a** [Lists all Symm devices, Addresses, Director Number/Type, and GateKeeper]**\$ server\_devconfig server\_x -p -s -a** [Lists all Devices ‘discovered’ on each chain for a DataMover]**\$ server\_devconfig server\_x -c -s -a** [Adds new devices to Celerra Database; \$/nas/volume/volumes|disks]**# nas\_disk –delete d40 –perm** [Deletes volume and removes diskmark]**# nas\_diskmark -m -a -monitor y**

Discovering storage (may take several minutes)

server\_2:

chain 0 .....

chain 16 .

Restart scan due to reconfiguration

chain 0 .....

chain 16 .

Verifying disk reachability

Verifying file system reachability

Verifying local domain

Verifying disk health

Verifying gate keepers

Verifying device group

**USING SERVER DEVCONFIG:** Equivalent to the nas\_diskmark command**\$server\_devconfig server\_2 -c -s -a** [Running this command updates the following Celerra Database files]**DOES & DON'Ts BEFORE RUNNING DEVCONFIG –CREATE ON CELERRA:**

→don't run without first verifying paths &amp; bind tables via devconfig –probe and fcp bind show

→don't run if BCV devices are mirrored on, as these devices would not be seen by Celerra—BCV's are seen only when split, or mirrored off (aka, uxf state)—while devices are established, BCV diskmark is overwritten &amp; restored only when split back off

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
 →don't run for Celerra systems attached to Clariion backends without first verifying that there are no trespassed LUNs  
 →don't run –create without first making backup copies of camdisk files (-create will update the camdisk, disks, and scsidevs files)  
 All Volumes→ /nas/volume/volumes  
 All Disks→ /nas/volume/disks  
 Server GateKeepers→ /nas/server/slot\_x/scsidevs  
 Server Devices→ /nas/server/slot\_x/camdisk. Devconfig –create will update this file with new devices  
 All GateKeepers→ /nas/dev. Gatekeepers are used with Symm, not Clariion backends

### **BYPASSING NAS 5.4 HEALTHCHECK FAILURES:**

### **USING DEVCONFIG TO ADD NON-CELERRA FBA DEVICES AFTER UPGRADE TO NAS 5.4:**

**Note:** NAS 5.4 does a database healthcheck when running the devconfig –create command. It's possible to have some Celerra devices as FBA devices rather than Celerra\_FBA for older systems that are just upgraded. See Primus emc116641. Telltale messages are shown below. Run the server\_devconfig command to force the addition of "FBA" devices to Celerra database:  
 000185701868 018A d15 emulation is FBA, should be CELERRA\_FBA

```
# server_devconfig server_x -c -s -a -discovery n -Force y [Force option ignores healthcheck]
```

```
# nas_diskmark -m -a -d n -Force y
```

SL7E1081700022 storage API db out of sync, last updated 2009/03/29-02:34:19

Forcing discovery

Discovering storage (may take several minutes)

done

### **VERIFYING CAM TABLES BY SCSI CHANNEL:**

```
# .server_config server_5 -v "camshowconfig"
```

CAM Devices on scsi-0:

TID 00: 0:d0 1:d1 2:d2 3:d3 4:d4 5:d5 15:d6

TID 01: 0:d7 1:d8 2:d9 3:d10 4:d11 5:d12 6:d13 7:d14 8:d15 9:d16 10:d17 11:d18 12:d19 13:d20 14:d21 15:d22---abridged----

### **VERIFYING CELERRA VOLUMES DATABASE CONSISTENCY:**

1. Verifying Device mappings from Backend

```
/nas/tools/dbchk -p
```

**Note:** Script by default will verify database files located in /nas/volume directory and crosscheck with server\_devconfig probes  
 2. Run following commands as a second check—any mismatched entries will cause an eiffel dump

```
# nas_fs -i -all
```

```
# nas_volume -i -all
```

```
# nas_disk -i <d_name or id=>
```

### **KEY VOLUME DATABASE FILES:** /nas/volume

disks →List of diskmarked volumes that Celerra sees. Devconfig –create adds to this file, nas\_disk –d will remove entries from file

filesys →List of filesystems created on Celerra

slices →Slices used in creating metas

symms →Storage systems that Celerra is attached to

volumes →Volumes used in creating file systems

/nas/server/slot\_x/camdisk

/nas/server/slot\_x/ufs →Entries in this file removed when perm unmounting filesystem, likewise added when mounting

/nas/server/slot\_x/volume →Entries removed when perm unmounting fs, added when mounting

### **FILESYS FILE:**

22:IPrep\_src:0:y:1:113:2:::0:**26@1**:23: →FSID '22' is a backup of FSID '26' at Host ID 1, related to IP Replication or TimeFinder (the “@1” matches up with index number in the output of nas\_cel –list)

23:root\_avm\_fs\_group\_4:0:n:101:0::::0:::22,29,32,46,2,47,48,49,50:4:

29:pfs:0:y:1:**123**:2:::0:28@1,33:23: →Meta that file system is created on

32:vpfs32:0:n:11:134::::0:::23:33:

33:pfs\_rvfs:0:n:9:135::::**1098572929**:29:4:34:32: →Timestamp created whenever a backup filesystem is created

34:pfs\_rvfs\_ipfs1:0:y:5:0:::2:1098572938:33:9:**35**: →Vol pool that file system was created from

35:34\_APM000421031830000:0:n:100:0::::0:::34,31@1:

47:fb:0:y:**1**:158:2:::0:50:23: →Indicates type of filesystem: 1=uxfs 5=rawfs 7=ckpt

52:fb\_ckpt2:0:y:7:::**2**:1106676885::0::50: →Data Mover that filesystem is mounted on

**DESCRIPTION OF EACH COLUMN IN FILESYS FILE:****30:fs\_time1\_snap1:0:n:1:155:::1115308301:29:1:::**

1. 30 = filesystem id
2. fs\_time1\_snap1 = filesystem name
3. 0 = filesystem acl -- not used
4. n = in use, n=no, y=yes, f=frozen, c=corrupted
5. 1 = volume type, for timefinder 1=uxfs , 6=mirrored
6. 155 = meta volume id the filesystem is created on
7. blank = rw servers
8. blank = ro servers
9. 1115308301 = absolute backup time – time last backup was created or refreshed
10. = 29 = backup of , the volume id of the filesystem this is the backup of
- 11.1 = filesystem type-- 0=local, 1=bcv
- 12.blank = volume id of backup filesystems of this filesystem
- 13.blank = avm pool the filesystem is created from
- 14.blank = if this is an avm pool, the filesystems built from the pool
- 15.blank = avm pool type

**VALIDATING SERVER DEVICE DATABASE CONSISTENCY AND BINDING TABLES:****1. Use Devconfig Probe to List out Each SCSI Chain and the Devices associated with them:**

# server\_devconfig server\_5 -p -s -a

Chain 0 Devices 1-176

Chain 16 Devices 1-176

**2. Use Devconfig List to determine FA Associated with each Chain :**

Chain 0 Devices 1-176 →04B

Chain 16 Devices 1-176 →03B

**3. Use Devconfig List or Camdisk file to determine if there are extra devices listed in Camdisk that are not seen in Probe:****List:**

|     |   |      |                   |     |      |             |
|-----|---|------|-------------------|-----|------|-------------|
| 177 | n | 8631 | 000182502561-004B | STD | d177 | 5,3,1,2,4,6 |
| 178 | n | 8631 | 000182502561-004C | STD | d178 | 5,3,1,2,4,6 |
| 179 | n | 8631 | 000182502561-004D | STD | d179 | 5,3,1,2,4,6 |
| 180 | n | 8631 | 000182502561-004E | STD | d180 | 5,3,1,2,4,6 |

**Camdisk Contains Extra Devices:**

177:c32t5l10+52666104B320+,c48t5l10+52666104B010+:

178:c32t5l11+52666104C320+,c48t5l11+52666104C010+:

179:c32t5l12+52666104D320+,c48t5l12+52666104D010+:

180:c32t5l13+52666104E320+,c48t5l13+52666104E010+:

**4. Confirm whether extra devices not seen in probe are in use or not:**

#nas\_disk -l

|     |   |      |                   |     |      |             |
|-----|---|------|-------------------|-----|------|-------------|
| 177 | n | 8631 | 000182502561-004B | STD | d177 | 5,3,1,2,4,6 |
| 178 | n | 8631 | 000182502561-004C | STD | d178 | 5,3,1,2,4,6 |
| 179 | n | 8631 | 000182502561-004D | STD | d179 | 5,3,1,2,4,6 |
| 180 | n | 8631 | 000182502561-004E | STD | d180 | 5,3,1,2,4,6 |

**Note:** Determine from customer or field if devices are actually no longer assigned to Celerra and then remove using nas\_disk -delete**5. Compare persistent & dynamic binding tables to ensure that all Chains match up to correct FA port:**

# .server\_config server\_5 -v "fcp bind show"

\*\*\* Persistent Binding Table \*\*\*

Chain 0000: WWN 50060482bfd1fb13 HBA 0 **FA-04ba**Chain 0016: WWN 50060482bfd1fb12 HBA 0 **FA-03ba**

Chain 0032: WWN 0000000000000000 HBA 1 FA-04ab

Chain 0048: WWN 0000000000000000 HBA 1 FA-03aa

\*\*\* Dynamic Binding Table \*\*\*

Chain 0000: WWN 50060482bfd1fb13 HBA 0 ID 0 Inx 00:00 Pid 0000 S\_ID 021000 Sys

Chain 0016: WWN 50060482bfd1fb12 HBA 0 ID 0 Inx 01:01 Pid 0016 S\_ID 021f00 Sys

Chain 0032: WWN 0000000000000000 HBA 1 ID 1 Inx 02:81 Pid 0032 S\_ID 000000 Non

Chain 0048: WWN 0000000000000000 HBA 1 ID 1 Inx 03:81 Pid 0048 S\_ID 000000 Non

**Note:** Persistent & Dynamic binding tables match for those FA ports that have actual disk devices assigned [FA-4b and FA-3b]

## **DETERMINING IF CELERRA & SYMAPI DATABASES ARE IN SYNC, THEN FIXING DATABASE IF CELERRA VOLUMES HAVE BEEN REASSIGNED OR DELETED:**

**Note:** This is not the procedure for a 'straight' addition of 'new' volumes to Celerra--see other sections

Step 1. \$nas\_disk -l → Verify "d3-d101" list [Example Only]

Step 2. \$server\_devconfig server\_2 -p -s -a

Step 3. Compare nas\_disk output of "d" volumes against server\_devconfig -probe:

### **WHAT DOES OUTPUT OF SERVER DEVCONFIG PROBE SHOW?**

This column equates to the listing of "d" volumes: **val= 3** thru **val= 101**,etc

If there are "d" devices listed in nas\_disk that are not "seen" by the datamover's when conducting the "probe", then the extra nas\_disk "d" devices are no longer physically attached to the Celerra, either as a result of Volume Changes, Reassignments, etc. If this is determined to be the case, then the following would be required to cleanup the Celerra Databases and re-sync the SYMAPI db:

Step 4. Create list of "d" volumes that are candidates for deletion [Ensure that they are "not in use" = n in nas\_disk -l output]

Step 5. Make a copy of the /nas/volume/disks and /nas/volume/volumes file

Step 6. Edit each of the above files and delete the appropriate "d" volumes

Step 7. Re-run \$server\_devconfig server\_x -c -s -a on all datamovers [precaution to ensure that all devices are 'discovered']

Step 8. Rebuild the SYMAPI Database by first renaming it, then running the following command:

### **LOCATION OF SYMAPI DATABASE ON CELERRA:**

# mv /nas/symapi/db/symapi\_db.bin /nas/symapi/db/symapi\_db.old

Correct permissions for symapi\_db.bin file:

-rwxrwxr-x 1 nasadmin nasadmin 2715490 Jun 14 21:05 symapi\_db.bin

# /nas/sbin/nas\_rdf -localinit [This command should run clean with following output]

Discover local storage devices ...

This operation may take up to a few minutes. Please be patient...

### **Note:**

If having trouble creating FileSystems, check Device Table and Impl Bin Flag Settings. By default, Server\_2 is the server that actually creates File Systems, therefore you would want to verify what Devices it can see.

#server\_devconfig server\_2 -l -s -a [Shows what Server\_2 sees from the NAS Database tables]

#server\_devconfig server\_2 -p -s -a [Probes live SCSI chains to read what devices are out there]

So, the two listings should match. If not, may entail rebooting DataMover and then running #setup\_slot -i 2 to rebuild the database

## **TROUBLESHOOTING SYMAPI DB PROBLEMS:**

1. Collect files from /nas/symapi/\* directory & /nas/log/symapi.log

2. Obtain Full Debug Log if requested by EE:

a.) move symapi\_db.bin and rename

b.) # export SYMAPI\_DEBUG=-1

c.) # export SYMAPI\_DEBUG\_FILENAME=/home/nasadmin/symapi\_debug.log

d.) Run Symm Discover: #/usr/symcli/bin/symcfg discover >discover.out

e.) Exit the shell to revert back from debug logging

**Note:** Don't forget to export symapi\_debug to a file. In certain situations, with debug turned on, symcfg or symcli commands may fail and output debug info to terminal.

## **USING IPCS TO CHECK FOR SEMAPHORE LOCKS ON SYMAPI:**

# /usr/bin/ipcs -s

**Note:** Might use this command to identify semaphore locks if the nas\_rdf -localinit command fails, etc.

## **NAS DISK DISPLAY OF CONTROL LUNS ON VARIOUS CELERRA CABINETS:**

### **NS SYSTEMS--CONTROL LUNS DISKMARKED FOR SYSTEMS WITH NBSNAS:**

\$ nas\_disk -l

| id | inuse | sizeMB | storageID           | devID | type       | name | servers |
|----|-------|--------|---------------------|-------|------------|------|---------|
| 1  | y     | 11263  | APM00033500569-0000 | CLSTD | root_disk  | 1,2  |         |
| 2  | y     | 11263  | APM00033500569-0001 | CLSTD | root_ldisk | 1,2  |         |
| 3  | y     | 2047   | APM00033500569-0002 | CLSTD | d3         | 1,2  |         |
| 4  | y     | 2047   | APM00033500569-0003 | CLSTD | d4         | 1,2  |         |
| 5  | y     | 2047   | APM00033500569-0004 | CLSTD | d5         | 1,2  |         |
| 6  | y     | 2047   | APM00033500569-0005 | CLSTD | d6         | 1,2  |         |

**Note:** For NS systems with NBSNAS, Data Movers must diskmark Control Luns, as the Servers actually do the querying and writing to these volumes on behalf of the Control Station.

\$ server\_devconfig server\_3 -p -s -a

server\_3 :

SCSI devices :

chain= 0, scsi-0

```
symm_id= APM00033500569 celerra_id= APM000335005690000
tid/lun= 0/0 type= disk sz= 11263 val= 1 info= DGC RAID 5 009900000000000NI
tid/lun= 0/1 type= disk sz= 11263 val= 2 info= DGC RAID 5 00990100010001NI
tid/lun= 0/2 type= disk sz= 2047 val= 3 info= DGC RAID 5 00990200020002NI
tid/lun= 0/3 type= disk sz= 2047 val= 4 info= DGC RAID 5 00990300030003NI
tid/lun= 0/4 type= disk sz= 2047 val= 5 info= DGC RAID 5 00990400040004NI
tid/lun= 0/5 type= disk sz= 2047 val= 6 info= DGC RAID 5 00990500050005NI
chain= 16, scsi-16
```

```
symm_id= APM00033500569 celerra_id= APM000335005690000
tid/lun= 0/0 type= disk sz= 0 val= -5 info= DGC RAID 5 009900000000000NI
tid/lun= 0/1 type= disk sz= 0 val= -5 info= DGC RAID 5 00990100010001NI
tid/lun= 0/2 type= disk sz= 0 val= -5 info= DGC RAID 5 00990200020002NI
tid/lun= 0/3 type= disk sz= 0 val= -5 info= DGC RAID 5 00990300030003NI
tid/lun= 0/4 type= disk sz= 0 val= -5 info= DGC RAID 5 00990400040004NI
tid/lun= 0/5 type= disk sz= 0 val= -5 info= DGC RAID 5 00990500050005NI
```

**Note:** -5 means that device can be seen but is not owned by that SP

**\$ .server\_config server\_3 -v "fcp bind show" "fcp topology"**

\*\*\* Persistent Binding Table \*\*\*

Chain 0000: WWN 500601611060172f HBA 0 SP-a1 Bound

Chain 0016: WWN 500601691060172f HBA 1 SP-b1 Bound

**Note:** For systems running NBSNAS, Data Movers must see all Control Volumes, hence the diskmarks. Also note that Control Luns are supposed to be seen down SPA, hence when you probe, the values are seen on Chain 0 SPA, not SPB, as in above output. Captive systems do not use Name Servers, so bind table entries show up as ALPA connections.

### **GOLDEN EAGLE, EAGLE, ALL OTHER SYSTEMS:**

**Note:** Devices should not show up in nas\_disk output as diskmarked volumes

**\$ nas\_disk -l**

| id | inuse | sizeMB | storageID-devID     | type  | name       | servers |
|----|-------|--------|---------------------|-------|------------|---------|
| 1  | y     | 4095   | WRE00022000830-0000 | CLSTD | root_disk  | 1,2,3,4 |
| 2  | y     | 4095   | WRE00022000830-0001 | CLSTD | root_ldisk | 1,2,3,4 |

**# server\_devconfig server\_2 -p -s -a**

server\_2 :

SCSI devices :

chain= 0, scsi-0

```
symm_id= WRE00022000830 celerra_id= WRE0002200083000
tid/lun= 0/0 type= disk sz= 4095 val= 1 info= DGC RAID 5 00990000070B6CL
tid/lun= 0/1 type= disk sz= 4095 val= 2 info= DGC RAID 5 009901000070D7CL
tid/lun= 0/2 type= disk sz= 2047 val= -99 info= DGC RAID 5 009902000070F7CL diskerr= unmarked
tid/lun= 0/3 type= disk sz= 2047 val= -99 info= DGC RAID 5 00990300007117CL diskerr= unmarked
tid/lun= 0/4 type= disk sz= 2047 val= -99 info= DGC RAID 5 00990400007139CL diskerr= unmarked
tid/lun= 0/5 type= disk sz= 2047 val= -99 info= DGC RAID 5 00990500007166CL diskerr= unmarked
```

**Note:** The -99 should be seen via probe for Control Volumes for Golden Eagle, Eagle cabinets, & always indicates no diskmark.

**# tail -4 /nas/volume/disks**

```
-1000:rootd1000:0:yU:1:WRE00022000830::1000:0002:7:
-1001:rootd1001:0:yU:1:WRE00022000830::1001:0003:7:
-1002:rootd1002:0:yU:1:WRE00022000830::1002:0004:7:
-1003:rootd1003:0:yU:1:WRE00022000830::1003:0005:7:
```

**Note:** This information must be in the disks file to prevent the volumes from being diskmarked by the Servers. If these entries are missing, there is a configuration problem.

**# .server\_config server\_2 -v "fcp bind show"**

\*\*\* Persistent Binding Table \*\*\*

Chain 0000: WWN 50060160006002ab HBA 0 SP-a0 Bound

Chain 0016: WWN 50060168006002ab HBA 0 SP-b0 Bound

### **CAMDISK FILE:**

**#more /nas/server/slot\_2/camdisk** [File outlines Data Mover SCSI Channels]

Example: 1:c0t0l0 + 52656001A320+: [Verify Boot Port for DM as c0t0l0]

**Note:** Camdisk file is static file that is represented by the devconfig –list command. File is not read on reboots, but does play key role in data mover failover/failback.

## **SCSIDEVS FILE:**

**# more /nas/server/slot\_2/scsidevs**

Lists Gatekeeper & Tape Devices: 1:gk01:0:c0t15l15:526560030320:002804000060:

**Note:** This file is recreated when the devconfig –create command is run. For single Symm cabinets, only 1 gatekeeper is used, and should be lun 15. Clariion backends do not use the gatekeeper device.

## **REPAIRING CELERRA GATEKEEPER DATABASE: \$nas\_diskmark -mark -all**

Step 1. Make copies of following files & delete from directory: \$/nas/dev [scsi chains] /nas/volume/disks\volume\symms

**/nas/server/slot\_\*\scsidevs**

Step 2. **\$nas\_diskmark -m -a** [rewrites the above files with current information]

Step 3. Repair dynamic tables to match up with persistent tables:

**\$.server\_config server\_x -v “fcp bind show”**

## **DISKMARK OVERVIEW PROCESS:**

1. New disks are discovered via probe
2. Celerra diskmark is written
3. Checks performed to verify paths
4. NAS\_DB updated with new devices

## **PERMANENTLY DELETING DISKMARKS FROM CELERRA VOLUMES:**

**# nas\_disk -delete d11 -perm**

## **REPLAYING DISKMARKS TO RESTORE PROPER DISKMARKS:**

**Note:** Most important files are the camdisk and disks file, which need to matchup

1. Zap any incorrect diskmarks first

**\$.server\_config server\_x -v “zapdiskmark c0t1l2 128” | 129**

2. Recover correct camdisk file & disks file from SCCS

3. Replay or Restore the Diskmarks:

**# nas\_license -l**

| key      | status | value       |
|----------|--------|-------------|
| site_key | online | 43 fc b2 05 |

**# nas\_diskmark -R -a -m or # nas\_diskmark -m -a -R** (NAS 5.4.22, NAS 5.5 syntax change)

Enter password: 43fcb205

**Note:** Password is derived from site\_key in output of nas\_license -l

4. Run devconfig –probe to verify no mismatched TID's or conflicts

**Caution:** Never run command without verifying camdisk file is 100% correct. The –R command will probe each device to read its diskmark—if it cannot, it will then diskmark the volume with whatever is in the Camdisk file. So never run this command if you have not verified the camdisk file is what you wish to present or you could cause data loss. So, you would use this procedure to restore diskmarks to volumes based on what you had in your camdisk file.

## **DIFFERENCE BETWEEN “SYMCFG” DISCOVER AND –LOCALINIT COMMANDS:**

--Symcfg discover updates SYMCLI database

--nas\_rdf -localinit updates both SYMCLI and NASDB Files as shown below

### **/nas/server/slot\_2/:**

|            |   |      |      |                   |          |
|------------|---|------|------|-------------------|----------|
| -rw-rw-r-- | 1 | root | root | 923 Jun 21 15:18  | volume   |
| -rw-rw-r-- | 1 | root | root | 44 Jun 21 15:18   | scsidevs |
| -rw-rw-r-- | 1 | root | root | 127 Jun 21 15:18  | mount    |
| -rw-rw-r-- | 1 | root | root | 2451 Jun 21 15:18 | camdisk  |

### **/nas/volume/:**

|            |   |      |      |                   |         |
|------------|---|------|------|-------------------|---------|
| -rw-rw-r-- | 1 | root | root | 5419 Jun 21 15:18 | volumes |
| -rw-rw-r-- | 1 | root | root | 29 Jun 21 15:18   | symms   |
| -rw-rw-r-- | 1 | root | root | 4490 Jun 21 15:18 | disks   |

## **HOW TO DETERMINE SYMAPI VERSION ON CELERRA:**

**# /nas/symcli/bin/symcli SYMAPI Version 4.3.2.0**

**Note :** Symapi version 5.3.1 and lower only supported 128 hypers/drive. Symapi version 5.3.2 and higher supports 256 hypers, found in NAS 5.2.11-3 and higher.

**COMPARING DEVICE DATABASES BETWEEN SERVERS:**

Step 1. Output Results of Server Devconfig 'List' for both Servers to Files:

```
$server_devconfig server_4 -l -s -a > list4
$server_devconfig server_5 -l -s -a > list5
$diff list4 list5 [compare differences for possible database inconsistencies]
```

Step 2. Output Results of Devconfig 'Probe' for both Servers to Files:

```
$server_devconfig server_4 -p -s -a > probe4
$server_devconfig server_5 -p -s -a > probe5
$diff probe4 probe5 [compare differences]
```

Step 3. Conduct Devconfig 'Create' and Output to Files:

```
$server_devconfig server_4 -c -s -a > create4
$server_devconfig server_5 -c -s -a > create5
$diff create4 create5 [compare results]
```

**COMMENT:** Use this to compare datamovers that are connected to same SA's or FA's

**UPGRADE 5.1 TO 5.2 CREATES DISKMARKS ON CONTROL VOLUMES NS SERIES:**

**# nas\_disk -l** output shows dismarked volumes

|    |   |      |                     |           |         |
|----|---|------|---------------------|-----------|---------|
| 47 | n | 2047 | APM00024700388-0002 | CLSTD d47 | 1,2,3,4 |
| 48 | n | 2047 | APM00024700388-0003 | CLSTD d48 | 1,2,3,4 |
| 49 | n | 2047 | APM00024700388-0004 | CLSTD d49 | 1,2,3,4 |
| 50 | n | 2047 | APM00024700388-0005 | CLSTD d50 | 1,2,3,4 |

**# server\_devconfig server\_2 -p -s -a** [output shows diskmarked volumes]

```
tid/lun= 0/2 type= disk sz= 2047 val= 47 info= DGC RAID 5 02060200009190CL
tid/lun= 0/3 type= disk sz= 2047 val= 48 info= DGC RAID 5 020603000091BCCL
tid/lun= 0/4 type= disk sz= 2047 val= 49 info= DGC RAID 5 020604000091E7CL
tid/lun= 0/5 type= disk sz= 2047 val= 50 info= DGC RAID 5 02060500009213CL
```

**# cat /nas/volume/disks**

```
47:d47:0:n:2047:APM00024700388:1,2,3,4:233:0002:7:
48:d48:0:n:2047:APM00024700388:1,2,3,4:234:0003:7:
49:d49:0:n:2047:APM00024700388:1,2,3,4:235:0004:7:
50:d50:0:n:2047:APM00024700388:1,2,3,4:236:0005:7:
```

**Note:** Last two columns are incorrect and need to be changed as follows

→Add two zeroes in front of 02, 03, 04, 05 and change type from “1” to “7” in last column

**CORRECTIVE ACTIONS REQUIRED TO REMOVE DISKMARKS:**

1. Remove diskmarked “d” volumes from database: #nas\_disk -d d47 [thru d50]
2. Using output from devconfig -probe, zap diskmarks for each volume to blocks 128 & 129:  
\$.server\_config server\_2 -v “zapdiskmark c0t0l2 128” & for 129 [d47→repeat for d48-d50]
3. Edit /nas/volume/disks file to pad two zeroes in 2<sup>nd</sup> to last column and in last column change type from “1” to “7”
4. Run diskmark command and verify: #/nas/sbin/nas\_diskmark -m ALL

**CORRECT OUTPUT FROM DEVCFIG:**

```
tid/lun= 0/2 type= disk sz= 2047 val= -99 info= DGC RAID 5 02060200009190CL diskerr= unmarked
tid/lun= 0/3 type= disk sz= 2047 val= -99 info= DGC RAID 5 020603000091BCCL diskerr= unmarked
tid/lun= 0/4 type= disk sz= 2047 val= -99 info= DGC RAID 5 020604000091E7CL diskerr= unmarked
tid/lun= 0/5 type= disk sz= 2047 val= -99 info= DGC RAID 5 02060500009213CL diskerr= unmarked
```

**Note:** Optionally, use following command to remove disk volumes and dismarks: #nas\_disk -d dxx -perm

**SYMMETRIX DATABASE V. CELERRA DATABASE:**

Symmetrix database consists of the following databases: **/nas/symapi/db/symapi\_db.bin** & **/nas/dev**

Any changes to Celerra Volumes requires that the Symmetrix database be updated manually

**VERIFYING GATEKEEPER DEVICES ON CELERRA:**

1. Run \$server\_devconfig server\_x -p -s -a to see if the Gatekeepers are recognized correctly:  
tid/lun= 0/2 type= disk val= -99 info= 5266470AF050 diskerr= -1 (or unmarked) [Example of properly seen GateKeeper]
2. Verify that AS400 Bit is "on" for the Gatekeeper Volume [see BIN file>Edit Volumes]
3. To update Celerra DataBase, run following on every Server: \$server\_devconfig server\_x -c -s -a

4. To update Symmetrix DataBase, run following:

```
--#mv /nas/symapi/db/symapi_db.bin /nas/symapi/db/symapi_db.old  
--#/nas/dev [move all devices listed here to an archive directory]  
--Run #/nas/sbin/nas_rdf -localinit [this will update the Symm DB to match the Celerra DB]
```

## **DUPLICATE DEVICES--CASE STUDY:**

### **Symptoms:**

- Cannot Mount or Create a file system
- Duplicate symmID-devID's Listed When Running #nas\_disk -l

```
# nas_disk -l
```

| id | inuse | sizeMB | symmID           | devID | type      | name | servers |
|----|-------|--------|------------------|-------|-----------|------|---------|
| 1  | y     | 4153   | 000184500672-000 | STD   | root_disk |      | 1,2,3,4 |
| 3  | n     | 34526  | 000184500672-008 | STD   | d3        |      | 1,2,3,4 |
| 10 | y     | 34526  | 000184500672-008 | STD   | d10       |      | 2,1,4,3 |

**Error:** 2001-10-22 17:45:23 server\_4:0: server\_mount server\_4 ufs01/gs1: filesystem is unreachable

**Cause:** Volumes were removed from the Celerra without updating SYMM & CELERRA DATABASES correctly

### **Solution:**

- Verify the running BIN File to check Volume assignments
  - Ensure that the duplicate or ghost volumes are not in Use
- a.) vi /nas/volume/disks and /nas/volume/volumes and remove the offending entries
  - b.) Run #nas\_rdf -localinit

## **REMOVING DATAMOVER FROM CELERRA:**

1. If Standby Server: \$server\_standby server\_x -d mover
2. Ensure no Volumes Mounted or Exported:  

```
$server_export server_x -a -p -u          $server_umount server_x -p -a
```
3. Umount Root File System as well:  

```
$server_umount server_x -p root_fs_x
```
4. \$server\_cpu server\_x -halt now
5. \$/nas/sbin/t2reset pwroff -s 2 [Power down the datamover]
6. #/nas/sbin/setup\_slot -d 2 [Deletes the old slot number]
7. #/nas/symapi/db/symapi\_db.bin [rename to db.old]
8. Run Symapi Update: #/nas/sbin/nas\_rdf -localinit

## **VERIFYING DATAMOVER DEVICE DATABASE CONSISTENCY:**

1. \$server\_devconfig server\_x -p -s -a [Outputs what the DataMover can physically “see” down each SCSI Chain or Port]
2. \$/nas/server/slot\_x/camdisk  
Contents of Camdisk and Server\_devconfig Probe should be consistent.

### **Some Reasons for DataBase Inconsistencies:**

1. BIN File Change not properly applied to the Celerra
2. Use of VOLUME LOGIX to reassign SYMMETRIX Volumes that the Celerra sees
3. Use of SDDR on the SYMM side
4. Faulty Hardware Components: SCSI Cable on DataMover or to SYMMETRIX; Bad HBA; Bad Director
5. Replaced, Added, or Deleted SYMM disk drives that have not been properly Mapped to Celerra and/or Database updated
6. Rezoning HBA's without adding Devices to Database using \$server\_devconfig server\_x -c -p -s -a

## **HOW TO VERIFY VOLUME DATABASE AFTER NEW DRIVES ADDED:**

1. Add New Devices to Celerra Database: \$server\_devconfig server\_2 -c -s -a

2. Compare Databases for inconsistencies:

```
$server_devconfig server_2 -p -s -a  
$server_devconfig server_2 -l -s -a  
$nas_disk -l [ensure that there are no duplicate entries]
```

## **ADDING NEW VOLUMES TO EXISTING SINGLE CELERRA:**

1. After installation of new BIN file ensure new Volumes are 'write enabled'.
2. BackUp of the Celerra Database to a known dir: \$/nas/sbin/nasdb\_backup /nas /home/nasadmin/EMC1
3. Verify current disk list & note last “d” devices on list: #nas\_disk -l
4. Run \$server\_devconfig server\_x ALL -p -s -a and verify that all new volumes are being seen down the SCSI chains correctly! If not, refer to PRIMUS "emc7974" for guidance
5. Run \$server\_devconfig server\_x ALL -c -s -a [Verify Volumes have been added--\$more /nas/server/slot\_x/camdisk]
6. Rename following file: \$/nas/symapi/db/symapi\_db.bin to /nas/symapi/db/symapi\_db.old

**Note:** Conduct this step whenever adding, reassigning new volumes or adding new DMs because the SYMAPI database must first be "moved" before the command given in Step 7 can "update" the SYMAPI database and thus make it consistent with the Celerra's Camdisk Database.

**Caution:** *This step NOT required for NAS Codes > than 4.2.13.0 or if no TimeFinder or Checkpoints are used*

- Step 7. As root, run the following command: `#/nas/sbin/nas_rdf -localinit`

## **ADDING NEW VOLUMES TO CELERRA WHEN SRDF IS INVOLVED:**

1. Run a backup of NAS database: `#/nas/sbin/nasdb_backup /nas /home/nasadmin/nasdb.bak.date`
2. After new volumes having been added via BIN file, and synchronized, verify that all new devices are seen by all DM's on both R-1 and R-2 side and then add to both sides:  
`$server_devconfig -p -s -a`  
`$server_devconfig -c -s -a`
3. Rename old Symapi database on R-1 Celerra:  
`#mv /nas/symapi/db/symapi_db.bin /nas/symapi/db/symapi_db.date`
4. Run `nas_server -l` and note which DM's are local production [acl=1000] & which are RDF standby DMs [acl=2000]
5. Run `#/nas/sbin/nas_rdf -init` on R-1 Celerra to discover remote devices and create R1/R2 pairs in `symapi_db.bin`
6. Repeat steps 3 -6 on the Target R-2 Celerra
7. Conduct queries to ensure that RDF Groups are seen correctly on both R-1 and R-2 Control Stations:  
`# /nas/symcli/bin/symcfg list`  
`# /nas/symcli/bin/symdg list`  
`# /nas/symcli/bin/symrdf -g 1R1_1 query` [R-1 Side]  
`# /nas/symcli/bin/symrdf -g 1R2_500_1 query` [R-2 Side]

## **PRESENTING NEW STORAGE DEVICES/VOLUMES/DISKS TO CELERRA FROM CLARIION:**

1. EMC Navisphere Manager (version 6.19)

RAID Groups>RAID Group 2>Bind LUN>RAID Group for New LUN:

**Note:** Select existing RAID Group with correct RAID Type and freespace, or create new RAID Group at this point  
LUN ID: 30

Number of LUNS to Bind: 10

Default Owner \*Auto

LUN Size 100 \*GB >Apply

2. Assign LUNS to Celerra Storage Group and provide HLU's:

Storage Groups>Celerra\_ns2>rightclick Select LUNS:

Available LUNS → Selected LUNS (Move luns from left to right)

Assign HLU number for each LUN>Apply (Must be valid HLU numbers >than 15)

3. Perform rescan of storage system from WebUI, or run following command:

`#nas_diskmark -mark -all`

## **DELETING & REASSIGNING CELERRA VOLUMES SYMMETRIX BACKEND:**

- Step 1. Backup all affected file systems before making LUN change

- Step 2. Make backup copy of `/nas/server/slot_x/mount` file

- Step 3. Permanently unmount affected file systems

- Step 4. Make Bin file change on symmetrix side

- Step 5. Verify paths to backend on each data mover and inspect Bind Tables for path issues

- Step 6. Discover new disks on all Servers: `$server_devconfig server_x -c -s -a` [Updates Camdisk]

- Step 7. Verify that camdisk has been updated with the intended changes and matches `server_devconfig -probe`

- Step 8. Remount each affected file system

## **UPDATING DATAMOVER'S VOLUME DATABASE:**

**Solution:** Running the following command updates the "camdisk" file and does not require a Server reboot

**`$server_devconfig server_x -c -s -a`**

**Note:** This command dynamically updates the "camdisk" file and is recorded in the `/nas/log/cmd_log` file

## **REMOVING SYMM DEVICES FROM CELERRA DATABASE FILES:**

`$nas_volume -i strv34`

- Step 1. Ensure Volumes to be deleted are no longer in use [`$nas_disk -l` look for "n" and Symm Volume or Device #]

- Step 2. If devices are in use, must identify which Servers and/or File Systems are using the Devices:

`$nas_fs -i fs01 [fs02, etc.]`

Step 3. Conduct BackUp of NAS DataBase: \$/home/nasadmin/nasdb\_backup /nas /home/nasadmin

Step 4. Make BackUp copies of /nas/volume/disks and /nas/volume/volumes files

Step 5. Use \$/nas/sbin/rootnas\_disk -d roottd189 [command to remove BCV devices from Celerra database]

**Note:** Devices are dynamically removed from Disks and Volumes files

\$nas\_volume -l is /nas/volume/volumes

\$nas\_disk -l is /nas/volume/disks

### **Locations for Celerra Configuration & Volume Information:**

\$/nas/volume/volumes

\$/nas/volume/disks

\$/nas/server/slot\_x/camdisk

[scsi chains & devices listed]

\$/nas/server/slot\_x/volume

[hyper volumes and associated Disks/TID's]

\$/nas/server/slot\_x/ufs

[file systems for this slot]

\$/nas/server/slot\_x/scsidevs

[gatekeeper devices]

\$/nas/server/slot\_x/mounttab

[lists out mountpoints and volume ID numbers for file systems]

\$/nas/server/slot\_x/SCCS

[Backup databases of volume and disks files, etc.]

### **OBTAINING LISTING OF SCCS DIRECTORIES ON ALL SERVERS:**

**\$ ls -l \$NAS\_DB/server/slot\_\*/\***

### **EXTRACTING SCCS DB FILES FROM SERVER SLOT:**

**# tar cf -./SCCS/{\*cam\*,\*vol\*} | (cd /tmp; tar xpvf -)**

#cd /tmp;ls

p.camdisk.l p.volume. s.camdisk.l s.volume.l

#view s.volume.l

**Note:** Find appropriate date and version of desired file

#sccts get -r1.38 volume.l

#view volume.l

### **SCCS DATABASE RESTORE PROCEDURE: Celerra Database BackUp Files**

The SCCS program is invoked by the NAS Database Backup program and runs just after the top of each hour. This program will back up critical /nas/server, /nas/volume, and other configuration files that have been modified during the preceding hour and place them in their respective SCCS directory, usually with a name such as “s.netd.l” and “p.netd.l”. An important thing to remember about this limitation is that only the last change prior to the NAS\_DB backup, within the past hour, will be captured—so when troubleshooting database issues, keep this in mind as the complete record of all changes are not necessarily captured. The easiest way to verify the various changes made to any specific file is to simply go to the SCCS directory and cat the desired file:

**# cat s.netd.l**

h18535

s 00000/00000/00017

d **D 1.15** 05/02/07 14:04:11 nasadmin 15 14

c -rw-rw-r-- 1 nasadmin nasadmin 389 Feb 7 13:37 netd

e

s 00001/00001/00016

d **D 1.14** 04/12/12 20:03:06 nasadmin 14 13

c -rw-rw-r-- 1 nasadmin nasadmin 389 Dec 12 20:01 netd

e

s 00001/00001/00016

d **D 1.13** 04/11/14 20:02:32 nasadmin 13 12

c -rw-rw-r-- 1 nasadmin nasadmin 390 Nov 14 19:12 netd

**Note:** The above highlighted items show the various version levels for the netd file. Change directory back one level and run **#sccts get -r1.15 netd.l** to retrieve a copy of the contents found in Version 1.15 of this file—rename the file to a desired name and repeat the process to retrieve other versions of the file.

**# sccts info**

bmparams.l: being edited: 1.2 1.3 nasadmin 04/12/17 15:01:42

buff.l: being edited: 1.4 1.5 nasadmin 04/12/17 14:01:57

camdisk.l: being edited: 1.6 1.7 nasadmin 05/01/14 14:01:47

cluster.l: being edited: 1.1 1.2 nasadmin 04/09/20 17:01:31

**Note:** Use above command to list out database files available. Also note that when we retrieve a file, we call the “camdisk.l” file from the “s.netd.l” database by using the –r1.15 syntax. We do not recover files from the “p.netd.l” database.

## **RECOVERING FILES USING SCCS:**

### **1. Go into the SCCS directory for the slot in question—run List & Grep for appropriate file:**

```
# ls -la |grep file [/nas/server/slot_9/SCCS]
-r--r--r-- 1 root root 39 Apr 30 11:31 file.l
-rw-r--r-- 1 nasadmin nasadmin 35 Apr 30 11:01 p.file.bak.l
-rw-r--r-- 1 nasadmin nasadmin 35 Apr 11 23:01 p.file.l
-r--r--r-- 1 nasadmin nasadmin 195 Apr 30 11:01 s.file.bak.l
-r--r--r-- 1 nasadmin nasadmin 511 Apr 11 23:01 s.file.l
```

### **2. Select the file with the date that you are looking for & cat the contents to obtain version information:**

```
# cat s.file.l
```

```
h31713
s 00001/00001/00000
d D 1.3 03/04/11 23:01:43 nasadmin 3 2
c -rwxrwxr-x 1 nasadmin nasadmin 39 Apr 11 21:47 file
s 00000/00000/00001
d D 1.2 02/10/06 10:03:05 nasadmin 2 1
c -rwxrwxr-x 1 nasadmin nasadmin 40 Oct 6 07:44 file
s 00001/00000/00000
d D 1.1 02/09/17 11:01:18 nasadmin 1 0
c date and time created 02/09/17 11:01:18 by nasadmin
```

**Note:** Example here shows (3) versions

3. Retrieve a version of the file contents by version number while in /nas/server/slot\_9 directory:

```
# sccs get -r1.2 file.l
```

```
get: SCCS/s.file.l: Warning: No id keywords.
1.2
1 lines
```

**Note:** Even though you see version numbers listed under “s.file.l”, you retrieve different versions by calling “file.l”

```
# cat file.l
```

```
file initialize nodes=65536 dnlc=262144
```

**Note:** In above example was able to show that the nodes and dnlc settings were different in the past than now, probably due to a NAS upgrade or patch that reset the settings to the current, lower values seen on the system

```
# cat file
```

```
file initialize nodes=22528 dnlc=65536
```

4. If you find the correct version to restore, simply rename the “file.l” as file and reboot server

**Note:** Contact Support Center for information on trimming SCCS databases and correcting permissions problems that prevent SCCS backups from working on servers—emc167945.

## **Deleting a DataMover:**

```
# /nas/sbin/rootnas_server -delete server_5
```

## **CELERRA SCSI BOOT PORT ON DATAMOVER:**

Port A=Chain0=both Boot Port and Data and is the Lower Port on SCSI Card. So, if it fails, DM must failover to its Standby unit!!  
Port B=Upper Port, Data Only

**Server setup Command:** Used to setup type and protocol for a DataMover

```
-t {nas|nms|standby} -L specify the DART image to load at next boot -P {nfs|cifs -o start|stop|delete}
```

## **USING SETUP SLOT -INIT ON DATA MOVERS:**

```
#/nas/sbin/setup_slot -init -t nas 2
```

When to Run Setup\_Slot? Adding or replacing Hardware, rediscover disks and rebuild bind tables, or when boot.cfg file is damaged.

**Note:** NAS 5.5, # export TRACE\_LEVEL=9 to get debug trace output of all the functions that setup\_slot performs—logs to /var/tmp

```
-rw-r--r-- 1 root root 11776 Mar 26 12:08 setup_slot_trace.log
```

**Probing for Newly Added SCSI Disks:** \$server\_devconfig server\_4 -p -s -d [listing: -l -s -a ]

**Probing for New Non-Disk SCSI Devices:** \$server\_devconfig server\_4 -p -s -nondisks [-l -s -n to list]

**UPDATING MAC ADDRESSES ON IFCONFIG FILES FOR DM:****# /nas/sbin/setup\_slot -upgrade 2**

**Note:** This command is very useful in properly updating MAC addresses in the ifconfig files when the MAC addresses are either not present or have been deleted, or new devices have been added, etc.

**DISKMARKING NEW VOLUMES:**

# nas\_diskmark -a -m [updates camdisk, volumes files, &amp; diskmarks for all Servers]

# server\_devconfig server\_2 -c -s -a [Updates camdisk, volumes file, &amp; diskmarks for specific data mover]

**Adding & Verfying New SCSI Drives/Devices and Camdisk File:**

1. #server\_devconfig server\_x -probe -s -all [shows all the scsi devices on the scsi chain that Celerra can see]
2. #cat /nas/server/slot\_x/camdisk [shows Camdisk File configuration of a specific slot or DataMover]

3. If the Devconfig and Camdisk are different:

**Note:** Must be run on each Data Mover!!

#server\_devconfig server\_x -c -scsi -all [Makes Celerra re-probe SCSI chain and updates Camdisk config file]

4. Redo Steps (1) and (2) to verify

**CAMDISK ENTRY DEFINED:**

1:c0t010+526562001310+:

1=Device name; c0=DM scsi boot port A chain 0; c1=port B chain 1; t0=target; l0=Lun; +5265=SymmCode; 62=Last 2digitsSSN; 001=SymmDev.ID; 31=Symm SA#; 0=SA Port#

**CAMDISK EXAMPLE:**

130:c0t8115+566901206550

|      |                                                |
|------|------------------------------------------------|
| 130  | Celerra device ID                              |
| c0   | SCSI Chain                                     |
| t8   | Target number                                  |
| 115  | Lun number                                     |
| 5669 | microcode family                               |
| 01   | Last two digits of the Symmetrix serial number |
| 206  | Symmetrix device number                        |
| 55   | Symmetrix director number (breakdown below)    |
| 0    | Port number                                    |

**Note:** Determine Symm Director from chart below

| Director number | 1  | 2  | 3  | 4  | 5  | 6  | 7  | 8  | 9  | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
|-----------------|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| Processor A     | 01 | 02 | 03 | 04 | 05 | 06 | 07 | 08 | 09 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
| Processor B     | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 |
| Processor C     | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 40 | 41 | 42 | 43 | 44 | 45 | 46 | 47 | 48 |
| Processor D     | 49 | 50 | 51 | 52 | 53 | 54 | 55 | 56 | 57 | 58 | 59 | 60 | 61 | 62 | 63 | 64 |

**SERVER LOG MESSAGE: Failed to locate volume...**

1. Check Boot.cfg file to see if it contains Volume & Hyper information
2. Check /nas/server/slot\_x/volume file to see if Volume information is intact
3. Check /nas/server/slot\_x/disks file to see if Disk information is intact

**SERVER INODE STATS: \$ server\_df server\_x -i**

**Note:** This is a command that communicates directly with the DataMover, so is good for troubleshooting purposes! Server\_df calculates space of FS based on blocks in use—since we use 8k blocks, we may see more usage than a client system. Also, when doing async I/Os, Superblock is not flushed each time, meaning Free Block Count on-disk is incorrect. The On-Disk Superblock gets updated after reboots or after FSCK Phase 5, but In-Block Superblock values are accurate all the time.

**HOW TO VERIFY DISK USAGE/SIZE OF DIRECTORIES FROM CS0 ROOT:**

# cd /;du -sk \* |sort -g [Outputs Usage in kb under Root or in Current Path] 36 root

2232 boot

3628 nasmcd

5076 etc

5776 bin  
7976 sbin  
13308 home  
15680 var  
26764 lib  
286156 nas  
363684 usr

### **VERIFYING CURRENT PARENT DIRECTORY SIZE:**

# du -sk

### **VERIFYING DISK USAGE BY SPECIFIC FILE OR DIRECTORY:**

# du -sk /nas

286160 /nas

### **DISK USAGE BY PARENT DIRECTORY:**

# du -h --max-depth=1 -x

**VFSTAB File:** Virtual File System Table lists all file systems and swap devices:

# mount -p [Prints what's mounted in the /etc/vfstab directory]

### **DISK USAGE AND FILESYSTEM INFORMATION:**

#df ~ or df \$HOME [Lists out the number of free blocks in Home directory]

# df -h # df -k # du -h # du -sh /\* |grep M

# du -s # df system

#server\_df server\_2 [lists out the FileSystems and usage info]

#nas\_fs -size fs1 or nas\_fs -i fs1 [outputs info on a specific FileSystem]

#nas\_volume -i meta [outputs info on Meta Volumes]

#nas\_volume -i d3lstr1|mtv1 [details on a specific volume] \$nas\_slice -i slc1

# du -sh /\* →Best method for seeing actual sizes

### **EXAMPLE DIRECTORY USAGE:**

\$ df -h \$du -h

| Filesystem    | Size | Used | Avail | Use% | Mounted on |
|---------------|------|------|-------|------|------------|
| /dev/sdc3     | 1.5G | 1.4G | 84M   | 95%  | /          |
| /dev/sdc1     | 30M  | 2.8M | 26M   | 10%  | /boot      |
| none          | 251M | 0    | 250M  | 0%   | /dev/shm   |
| /dev/sde1     | 1.7G | 617M | 1.0G  | 37%  | /nas       |
| /dev/sda1     | 133M | 107M | 26M   | 80%  | /nas/dos   |
| /dev/sdf1     | 1.7G | 657M | 1.0G  | 39%  | /nas/var   |
| server_3:/mnt | 8.4G | 2.5G | 5.9G  | 30%  | /mnt       |
| server_3:/mnt | 8.4G | 2.5G | 5.9G  | 30%  | /mnt       |

### **Verifying Space Utilization on DataMover FileSystems:**

**Control Station:** \$df -k

**Servers:** \$server\_df ALL

**FileSystem:** #mount -F nfs server\_2:/mailstore /mnt \$cd /mnt;df -k

**Switches:** \$df . [space on current directory] \$df -a [space on all FileSystems] \$df -i [used and free inodes]

### **CELLERRA QUOTAS:**

**Intro:** File System quotas limit amount of data Users can store on a FileSystem. Tree Quotas limit amount of data Users can store in Subdirectories of a FileSystem. Quotas can be limited by Bytes of Data [Blocks], Number of Files [Inodes], or both. The limit on number of files is dependent on Inodes—inodes are used for Files, Directories, & Symbolic Links. Filesystem Quotas can be applied to individual Users or Groups of Users. Control Station manages via CLI by blocks, files, users, filesystems [NOT Groups for CIFS]. Windows 2000 Quotas are applied to the whole disk partition and cannot be applied to individual Folders. Quotas can be set using default values for all users, or set explicitly for Users/Groups, either via single fs basis or by DM basis.

### **DEFAULT FILES CREATED BY NFS QUOTAS IN SYSTEM FOLDER (Pre NAS 5.5):**

**Note:** Following files are created and stored in the “/fs20/.etc” directory when implementing Quotas. They are system files and should not be touched or edited ever. The .etc directory is created when Quotas is first configured from Unix side or after first mapping to CIFS FileSystem Share from Windows.

|                        |           |      |     |                                                 |
|------------------------|-----------|------|-----|-------------------------------------------------|
| config_file            | rw-r—r--  | root | bin |                                                 |
| config_file.InProgress | rw-r—r--  | root | bin |                                                 |
| quotas                 | rw-rw-rw- | root | bin | [User quotas file—created when first turned on] |
| quotas.config          | rw-rw-rw- | root | bin | [Other configuration information]               |
| quotas.group           | rw-rw-rw- | root | bin | [Group quotas file]                             |
| rpt_file               | rw-r—r--  | root | bin | [created after first report is run]             |
| rpt_file.InProgress    | rw-r—r--  | root | bin |                                                 |

**Note:** Verify that InProgress reports are completed or not by running octal dump against the files

```
# od -xc *InP*
```

## **CELERRA QUOTAS DATABASE:**

With NAS 5.5, the quotas, quotas.group, & quotas.config database files were moved from the hidden .etc directory of each production file system to the hidden /slashetc directory of the file system. A NAS Upgrade from 5.4 to 5.5 reflects the database change in the Server Log—Please note that the Quotas DB is now called Version 3 in NAS 5.5 because of this change:

**2007-03-27 10:38:37: UFS: 4: Quota Upgrade: Upgrading from quota version 2 to version 3 (fs /fs06)**

## **NAS 5.5 QUOTAS DB:**

**Note:** Following example shows the various files stored in the “.etc” and the “slashetc” directories on the file system

[root@nyip1 m3]# ls -la .etc

```
-rw-r--r-- 1 root bin 60 Jun 2 2006 config_file
-rw-r--r-- 1 root bin 9 Jun 2 2006 config_file.InProgress
-rw-r--r-- 1 root bin 0 Jun 2 2006 config_file.sids
-r--r--r-- 1 root bin 262144 May 10 10:28 gid_map
-rw-r--r-- 1 root bin 36 Jun 2 2006 rpt_file
-rw-r--r-- 1 root bin 9 Jun 2 2006 rpt_file.InProgress
-rw-r--r-- 1 root bin 6 Jun 2 2006 rpt_file.sids
```

/nasmcd/quota/slot\_2/m3/slashetc

[root@nyip1 slashetc]# ls -la

```
----- 2 root root 8192 May 10 15:34 ACLdata
----- 2 root root 8192 Jun 2 2006 ACLrecord
-rwxr-xr-x 2 root root 262144 May 10 10:28 gid_map
-rw-rw-r-- 2 root bin 1048832 Jun 2 2006 quotas →User Quotas database
-rw-rw-r-- 2 root bin 128 Jun 2 2006 quotas.config →Quotas configuration db
-rw-rw-r-- 2 root bin 64 Mar 28 2007 quotas.group →Group Quotas database
----- 2 root root 8192 May 10 10:28 ufsasyncDB
-rw-r--r-- 2 root bin 0 Jun 2 2006 viruschecker
```

## **PROCEDURE FOR REVEALING THE SLASHETC DIRECTORY NAS 5.5:**

**Note:** When exposing the slashetc directory to view, the link count is set to 2. Because of the link count behavior, we can then safely remove the slashetc directory, which actually only decrements the link count from 2 to 1, at which point the directory contents are invisible again.

1. Following example shows the method for revealing the hidden “slashetc” directory that contains the actual Quotas database files

**# .server\_config server\_2 -v "shadow showetc fs1"**

/nasmcd/quota/slot\_2/fs1

# ls -la

```
drwxrwxrwx 2 root bin 1024 Jan 5 2007 slashetc
```

2. Change to the slashetc directory:

# cd slashetc

# ls -la

```
----- 2 root root 8192 Nov 14 15:20 ACLdata
----- 2 root root 8192 Jan 5 14:08 ACLrecord
-rwxr-xr-x 2 root root 262144 Nov 3 12:55 gid_map
-rw-rw-r-- 2 root bin 1056000 Jan 5 13:35 quotas
-rw-rw-r-- 2 root bin 64 Jan 5 13:32 quotas.config
-rw-rw-r-- 2 root bin 64 Mar 28 2007 quotas.group
----- 2 root root 8192 Nov 3 12:55 ufsasyncDB
```

3. When done, remove the slashetc folder so as not to leave the QuotasDB exposed, as database files could be deleted:

**# rm -rf slashetc**

**Note:** Different syntax? **# .server\_config server\_x -v "shadow showetc \mnt1\danfstest"**

## **NAS QUOTAS COMMANDS:**

--Edit command is an Eiffel executable which creates (2) temp files called “mount-point-file” & “Uid-List-File” to store mountpoint information and UIDs/GIDs, respectively.

--A mac/config command is then sent to DM to create a report file (rpt\_file) [The DM removes the rpt\_file & rpt\_file.tmp in the .etc directory of the file system and creates a new “rpt\_file.tmp” in which is parses the UID list to report on all User with non-zero limits. Quota information is then compiled and the file renamed to “rpt\_file”].

--If quotas need to be edited, the edquota executable is called by Control Station, which reads the “rpt\_file” and opens vi editor—changes are then updated to DM by mac config commands to quotas for users and quotas.group for group quotas, and to in-memory copies of quota records.

## **QUOTA SUPPORT BY NAS CODE:**

1.2 Introduces User Quotas--also called Inodes Quotas

2.2 Introduces Group Quotas, with Quotas charged to a User’s primary group

4.0 Introduces Soft Quotas & Win2k Quotas, warning when threshold is reached, with grace period of 1 week by default

5.0 Introduces Tree Quotas, which can be used to limit Inodes or Blocks used, are set at the Directory level

**Note:** Using filesize quotas with DHSM migrated file systems—byte usage for tree or owner could be more or less than expected.

## **TYPES OF QUOTAS USED WITH CELERRA: USER/GROUP/DIRECTORY (Tree)**

→User quotas, applies to file system either explicitly, per User, or as default quota limits applied to all Users

→Group quotas, applies to file system explicitly, per Group, or as default limits

**Note:** Group quota applies to sum of storage for all users having the group as their default primary group

→Windows 2000 Quotas, applies to whole partition, and is based on User quotas

→Tree Quotas, applies to directory tree

→Quota db is stored in files named quotas & quotas.group in the .etc directory of each file system as binary structure dqblk form

→Quotas are managed from Control Station CLI using nas\_quotas commands or GUI

## **QUOTA FUNCTIONALITY/FACTS:**

--Quotas limit the bytes of storage used, the number of files used, or both

--Quota blocks are calculated in 8192 bytes, filesize by 1024 bytes

--Quotas are calculated for Blocks or Bytes Usages in (1k), meaning 1024 byte increments

--Default quota policy is still Block size, yet most CIFS applications should use Filesize

--User & Group quotas are based on UID/GIDs and are applied to the overall filesystem, whereas TreeQuotas apply only to a directory

--File limits count inodes for files, directories, or symbolic links

--Quota limits are tracked separately for User, Group, or TreeQuotas (aka QTrees)

--Nested tree quotas are not supported, i.e., cannot have multiple TreeQuotas within the same directory tree

--Default grace period (soft quota) is one week for both CIFS & NFS users, though only CIFS users will receive popup warning message (if configured). Once a grace period expires, the quota turns hard, and no further writes allowed.

--Quota records are indexed by UIDs

--MPD file systems allow tracking of root users for usage (countRootUsageInQuotaTree)

--Explicit quotas override global default quotas for users or groups

--Quota limits cannot be enforced for root user or root group

--Anonymous users UID=-2 can be set using “anon” with nas\_quotas

--Quotas database files are located in hidden .etc directory of the file system

--Can use both Group & User Quotas concurrently—whichever reaches limit first determines extent of write access

--Quotas support NFS & CIFS, though CIFS supports only Filesize quotas, while NFS supports both Block & File Size

--Win2k Quotas can work with either Passwd/Group file or Usrmapper access to map to Users/Groups

--Tree Quotas also have soft & hard limits, with filesize or blocksize policies

--Tree Quotas are limited to 2047 different Trees per file system

## **TWO TYPES OF QUOTA POLICIES:**

**Blocks**→default, calculates disk usage in terms of 8kb fs blocks and usage of all files (files, directories, symbolic links)

**Filesize**→disk usage calculated in terms of file sizes, as in CIFS environments, and does not include directories or symbolic links

**Note:** If a system started out with the default policy of blocks, then later set the param to filesize, the proper sequence of events should be to reboot server. User & Group quotas will recalculate properly to presently filesize. Tree Quotas, on the otherhand, would require manual intervention and use of the Quota Recalc Tool to fix 8k block to show 1k file. Please note that when looking at Quotas usage from reports and comparing to server\_df output, the two values will always be different

## **CIFS vs. NFS:**

NFS can use either block or filesize quotas

CIFS should use only filesize

## **USER QUOTAS RESTRICTS HOW LARGE FILE SYSTEM SIZE APPEARS TO CIFS USERS:**

If the file system is 1GB in size, but a User's quota is 100MB, the User will only see 100MB of total disk space from their Windows client. NFS Users, on the otherhand, will be able to see the total file system size of 1GB.

## **EFFECTS OF TURNING QUOTAS ON OR OFF:**

File system will be frozen and will affect CIFS & NFS users

Quota-checking is done when first turning quotas on (after being turned off), and after fsck when fs is mounted

## **SETTING DEFAULT USER/GROUP QUOTAS:**

#nas\_quotas -edit -config -fs fs1 (set the user and group quotas, and it will apply to the file system)

## **SETTING EXPLICIT ANONYMOUS USER QUOTAS:**

#nas\_quotas -edit -user -fs fs1 anon

## **TURNING ON/OFF BOTH USER & GROUP QUOTAS:**

#nas\_quotas -on | -off -both -fs fs1

## **REPORTING QUOTAS BY FILE SYSTEM:**

#nas\_quotas -report -user -fs fs1

## **VIEWING USER/GROUP QUOTA CONFIGURATION:**

#nas\_quotas -report -config -fs fs1

## **USER QUOTAS:**

1. Turn quotas on: \$nas\_quotas -on [-off] -fs hfs1
2. Edit User quota: \$nas\_quotas -e -u -fs hfs1 1556
3. Verify with Report: \$nas\_quotas -r -u -fs hfs1 1556

## **GROUP QUOTAS:**

Step 1. Set Param & Reboot: **param cifs useUnixGid=1**

Step 2. Setup Group quota: #nas\_quotas -g -e -fs ufs1 1001

**Note:** useUnixGid=0 (default) assigns GID to Windows User based on Primary Domain group; useUnixGid=1 assigns GID from GID field for the User in the /etc/passwd or NIS map. Quotas are charged to a User's Primary Group.

# server\_param server\_2 -facility cifs -i acl.useUnixGid -v

server\_2 :

```
name          = acl.useUnixGid
facility_name = cifs
default_value = 0
current_value = 0
configured_value =
user_action   = none
change_effective = immediate
range         = (0,1)
description   = Set the unix GID for CIFS created files (default:off)
detailed_description
```

Sets the GID mapping for files created on a Windows client. param cifs useUnixGid=0 Assigns the GID of the Primary Domain group to which the user belongs. param cifs useUnixGid=1 Assigns the Windows user's GID (as found in the GID field of the /etc/passwd file or NIS database entry).

## **QUOTAS SET FOR INDIVIDUAL USERS: Default="Blocks-Used"**

**Note:** To use Quotas on an NFS File System, you must be using (1) of the following (3) User-Based Authentication Methods:

a.) NIS Services for Unix    b.) Local Passwd & Group files on DataMover    c.) Usrmapper

1. If using local Passwd/Group files on Data Mover or Usrmapper, need to obtain UID's/GID's in order to set-up Quotas.
  - Lookup UID's, Edit Passwd file, and put file onto Datamover: \$server\_file server\_4 -p passwd passwd
  - Lookup UID's of users in /var/usrmapper/usrmap.db on Usrmapper Server
2. **Create or Edit a User 'Quota':** #nas\_quotas -e -u -fs ufs76 1005 [UID for user: todd]
  - This opens up the Quota file in the vi Editor: "Userid: 1005 FS ufs76 blocks (0) inodes (0)"
  - Enter your quota value in **blocks (0)** and leave inodes at (0), then save and exit the vi editor
3. **Turn On Quotas for the FileSystem:** #nas\_quotas -on {-off; -clear} -fs ufs76 {Using 'clear' erases all quotas & turns off}
4. **Verify Quotas by Report:** #nas\_quotas -r -u -fs ufs76 1005 [should show entries you made to the 'blocks' column]

**Note:** You can change Quotas at any time using the ‘nas\_quotas -e’ editor--changes are dynamic.

**QUOTAS SET FOR ANONYMOUS USERS:** \$nas\_quotas -e -u -fs ufs34 anon

**QUOTAS SET ON A DATAMOVER BASIS:** Quotas set for each User would apply to all FileSystems on that DataMover

1. #nas\_quotas -e -u -m server\_5 1005 [leave UID off to set for all users]
2. #nas\_quotas -on -m server\_5
3. #nas\_quotas -r -u -m server\_5 1005 [leave UID off for report on all users]

**FILESIZE QUOTAS:**

**param quota policy=filesize**

**Note:** Set quota policy and reboot data mover—quota tracking will change from “Blocks Used” to “Bytes Used”

4. Turn Quotas on using “\$nas\_quotas -on -m server\_4”

**Note:** You cannot have both *Blocks* and *FileSize* policy set for the same DataMover—only one Method or the other

**SETTING UP QUOTAS FOR PROTOTYPE USERS ON A USER BASIS:**

1. Create Quota for Prototype User: “\$nas\_quotas -e -u -fs ufs34 1000” -vi edit Blocks or Bytes (0)
2. Setup Quotas for other Users: #nas\_quotas -e -fs ufs34 -p 1000 1001 1002 1003 1004 1005 [creates quotas for others]
3. Turn quotas on: \$nas\_quotas -on -fs ufs34

**Note:** The prototype user is the first UID listed—all other UID’s inherit quota attributes of UID 1000. A benefit here is that you will not have to Edit the values separately for each of the other UID’s.

**Caution:** If you change Values for Proto User, then you must re-run Prototype -p command for other Users.

**SETTING UP QUOTAS FOR PROTOTYPE USERS ON A DATAMOVER BASIS:**

1. Create Quota for Prototype User: “\$nas\_quotas -e -u -m server\_4 1000” and edit Blocks or Bytes values.

Setup Other Users by UID: \$nas\_quotas -d -m server\_4 -p 1000 1002 1002 1003 1004 1005

Turn quotas on: \$nas\_quotas -on -m server\_4

Clearing Quotas on a Server: \$nas\_quotas -clear -m server\_4 [also turns Quotas off]

**TESTING QUOTAS FROM A UNIX CLIENT:**

1. Make sure your UID number is the same on both the UNIX client and the Celerra DataMover [passwd/group files or usrmapper]  
**Note:** Use Solaris “admintool” program to create Users—manually editing /etc/passwd file will not work on Solaris!
2. Mount Data Mover File System to Unix Client: #mount 193.1.21.183:/g3data1 /mnt
3. MakeFiles as ‘Root’ on Unix Client for Testing: #mkfile 2m one two three #chmod 777 one two three [allows everyone]
4. #cd /mnt;ls -la [to verify that you can view the mounted file system]
5. #cd / and login as the test user “todd” with UID 1005
6. #exec login login: todd password: nasadmin \$/usr/local/todd \$cd / \$whoami ‘todd’
7. Begin Copying Test to DM: \$cp one two three /mnt
8. Watch copy operation from NT Client mapped to same file system; Refresh shows status of copy operation

**Unix client:**

#quota -v tom [command to check quota information from NFS clientside]

**CLEARING QUOTAS:**

**\$nas\_quotas -clear -fs ufs1**

**Note:** Quotas cannot be cleared for single Users or Groups, only the entire fs!

**SETTING QUOTAS FOR A USER BY FILE SYSTEM:**

**\$nas\_quotas -u -e -fs ufs1 5005**

Edit either the Blocks or Files section and close vi editor session

**RUNNING QUOTAS REPORTS ON A FILE SYSTEM:**

**\$nas\_quotas -u -r -fs ufs1**

**TURNING ON/OFF QUOTAS:**

**\$nas\_quotas -u -on | -off -fs ufs1**

**TROUBLESHOOTING QUOTAS:**

**NETD File:** /nas/server/slot\_x/netd: rquotad action=start [Read into the Boot.cfg file upon shutdown]

Nested mountpoints can create problems using the "Editor" and "Report" commands [commands won't work]

**Server Log:**

2002-11-10 21:16:18: CFS: 3: write failed, startOffset = 0x0, status = 52

2002-11-10 21:16:18: CFS: 3: write failed, startOffset = 0x2000, status = 52

**Cause:** File\_QuotaExceeded, // 52 [From NAS 2.1.30.2]

**Disk Quota Exceeded Error:**

--Client receives when Hard Quota Limit reached

--Client receives when Soft Quota limit crossed and Grace Period expires

**Problem:** Quota Editing & Reporting does not work with nested mountpoints

**Solution:** Mount filesystem temporarily on 'non-nested' mountpoints

### **CANNOT APPLY QUOTA TO USER “ROOT”:**

**#nas\_quotas -u -e -fs fs20 0**

‘quotas not support for root (0)’

### **QUOTA PARAMETERS IN 5.4 CODE:**

**\$ .server\_config server\_2 -v "param quota"**

| Name | Location | Current | Default |
|------|----------|---------|---------|
|------|----------|---------|---------|

|                                 |            |            |            |
|---------------------------------|------------|------------|------------|
| quota.countRootUsageInQuotaTree | 0x0139efb8 | 0x00000001 | 0x00000001 |
|---------------------------------|------------|------------|------------|

**Note:** includes root users in tree quota usage

|              |            |            |            |
|--------------|------------|------------|------------|
| quota.maxuid | 0x013931d8 | 0x00000000 | 0x00000000 |
|--------------|------------|------------|------------|

|              |            |          |          |
|--------------|------------|----------|----------|
| quota.policy | 0x01393218 | ‘blocks’ | ‘blocks’ |
|--------------|------------|----------|----------|

**Note:** Default is ‘blocks’, based on number of 8kb fs blocks allocated. Policy using ‘filesize’ means quota usage by file usage in 1kb increments. Cannot implement this param without first having quotas turned off and rebooting data mover with new param value set

|                         |            |            |            |
|-------------------------|------------|------------|------------|
| quota.useQuotasInFsStat | 0x0139f008 | 0x00000000 | 0x00000000 |
|-------------------------|------------|------------|------------|

**Note:** When set to 1, displays free-space on a Quota basis to Unix users based on their User, Group, Tree quotas. Otherwise, returns total free-space in whole file system

**# server\_param server\_2 -f quota -l** (use –i –all for verbose output)

server\_2 :

| name | facility | default | current | configured |
|------|----------|---------|---------|------------|
|------|----------|---------|---------|------------|

|                           |       |   |   |  |
|---------------------------|-------|---|---|--|
| countRootUsageInQuotaTree | quota | 1 | 1 |  |
|---------------------------|-------|---|---|--|

|        |       |   |   |  |
|--------|-------|---|---|--|
| maxuid | quota | 0 | 0 |  |
|--------|-------|---|---|--|

|        |       |          |          |  |
|--------|-------|----------|----------|--|
| policy | quota | ‘blocks’ | ‘blocks’ |  |
|--------|-------|----------|----------|--|

|                   |       |   |   |  |
|-------------------|-------|---|---|--|
| useQuotasInFsStat | quota | 0 | 0 |  |
|-------------------|-------|---|---|--|

### **SOFT QUOTAS & “GRACE PERIOD”:**

--WIN2K Clients receive warning popup but can still write to File System and have an infinite warning state [grace period N/A]

--Unix Clients granted grace period when soft quotas reached but receive no Warning message and "grace period" expires after set amount of time—default is 1 week for Block/Bytes & Inode settings

--Win2k has infinite grace period setting whereby inode & block grace period = “-1”

--When Win2k has NoLimit set, values show up in nas\_quotas –edit as “-2”

--Warning/Event Logged whenever soft quota disk or file usage limit reached

--Use "Soft Quotas" as warning tool prior to reaching "Hard Quota" limits for Users or File Systems--default value is 100MB

### **CHANGING DEFAULT GRACE PERIOD FOR SOFT QUOTAS:**

**# nas\_quotas -c -e -fs fs16**

#### **File System Quota Parameters:**

fs fs16

**Block Grace: (-1.0 )** [By default, Grace Period for Soft Quotas is 7 days]

**Inode Grace: (-1.0 )** [By default, Grace Period for Soft Quotas is 7 days]

\* Default Quota Limits:

User: (soft = 8192, hard = 10240) inodes (soft = 0, hard= 0)

Group: (soft = 0, hard = 0) inodes (soft = 0, hard= 0)

Deny disk space to users exceeding quotas: (yes)

\* Generate Events when:

Quota check starts: (no)

Quota check ends: (no)

User's soft quota crossed: (yes)

User's block quota exceeded: (yes)

**Note:** Use the above command to Edit the Quotas Configuration file for Filesystem ‘FS16’ . Set the Block Grace: (-1.0) to change Grace period to infinity.

### **UNIX CLIENT BEHAVIOR WHEN SOFT & HARD QUOTA LIMITS REACHED:**

#### **SYSTEM CONSOLE MESSAGE:**

“/mnt/fs20/files/getreason”: Disk quota exceeded”

**Note:** Error seen on client when copying files to Celerra and Hard Quotas is reached—copy operation is aborted.

#### **QUOTA REPORT FOR FS20:**

|      | Bytes Used (1k) |      |          | Files |
|------|-----------------|------|----------|-------|
| Used | Soft            | Hard | Timeleft | Used  |

#201 6999 5000 7000 7.0days 300

### **SERVER LOG MESSAGES:**

UFS: 4: Block soft quota crossed (fs /fs20, uid 201)

UFS: 3: Block hard quota reached/exceeded (fs /fs20, uid 201)

CFS: 3: write failed, start Offset = 0x28000, status=52

**Note:** The “write failed” error is written to Server Log whenever a User that has exceeded their Hard Quota is trying to copy additional data to Celerra. This behavior is not duplicated from a Windows 2000 Client for Windows 2000 Quotas.

### **VARIOUS QUOTA ERROR MESSAGES:**

CFS: 3: write failed, startOffset = 0x2000, status = 52

**Note:** User Quota exceeded: File\_QuotaExceeded

CFS: 3: write failed, startOffset = 0x35, status = 63

**Note:** Group Quota Exceeded: File\_GroupQuotaExceeded

CFS: 3: write failed, startOffset = 0x33e99a000, status=67

**Note:** Tree Quota exceeded: File\_TreeQuotaExceeded

2006-03-15 20:26:04: CFS: 3: write failed, fsid = 224, startOffset = 0x3cb1000, status = 25

**Note:** Error when NFS client tries to write to full file system

2006-03-22 12:26:14: UFS: 3: Block hard quota reached/exceeded (fs /fs\_quota, uid 32774)

**Note:** Message showing hard limit reached for UID

## **WINDOWS REPORTS QUOTAS VALUES DIFFERENT THAN NAS QUOTAS REPORTS:**

Problem is that quota policy was not set to “filesize” [which needs to report file sizes in 1k blocks, not 8k blocks]

### **Resolution:**

1. Set param file: param quota policy=filesize
2. Turn off quotas: \$nas\_quotas -both -off -fs fs01
3. Reboot DM and turn quotas on: \$nas\_quotas -both -on -fs fs01

## **WINDOWS 2000 QUOTAS & CLIENT BEHAVIOR WHEN SOFT & HARD LIMITS REACHED:**

### **CLIENT ERROR—Soft Quota:**

No Client PopUp message seen if crossing just the “Soft Quota” threshold

### **SERVER LOG MESSAGE—Soft Quota:**

UFS: 4: 4: Block soft quota crossed (fs /fs21,uid 602)

### **WINDOWS QUOTAS TAB:**

fs21 on ‘fox’ (E): Rightclick mapped drive to Celerra>Properties>Quota>Quota Entries>

Warning [Yellow Warning Triangle appears beside User that has crossed Soft Quota limit]

### **CLIENT ERROR—Hard Quota:**

“Error Copying File or Folder

X Cannot copy celerra\_info: There is not enough free disk space. Delete one or more files to free disk space, and then try again.”

### **SERVER LOG MESSAGE—Hard Quota:**

UFS: 3: 5: Block hard quota reached/exceeded (fs /fs21, uid 602)

**Note:** Further attempts to write to disk by Client are denied but do not log any further Server Log messages

### **WINDOWS QUOTAS TAB:**

Warning Triangle message remains for User that has crossed either Soft or Hard Quota limits

### **TREE QUOTAS MESSAGES:**

UFS: 4: Block soft quota reached (fs /fs1 treeid 1)

UFS: 4: Inode hard quota exceeded (fs /fs1 treeid 1)

## **QUOTA MANAGEMENT:** Can use Unix, Win2k, or Control Station CLI

WIN2K Clients can only set quotas by "bytes" on Users using My Computer>Select Volume>Properties>Quota Tab [NOT blocks, NOT inodes, NOT Groups] and allows setting of soft quota warnings

UNIX Clients can only view Quota reports by use of "rquota" daemon that runs on datamover and UNIX client

### **COMMANDS:**

**#quota      #quota -v      #quota -v dave      #quota dave**

## **RQUOTA DAEMON ON DATA MOVER:**

**# /usr/sbin/rpcinfo -p 10.241.169.54 |grep quota**

100011 1 tcp 57687 rquotad

100011 1 udp 49166 rquotad

**\$quota -v user |<uid>**

## **USING QUOTA TO QUERY A UID FROM CONTROL STATION:**

**# quota 32769**

Disk quotas for user #32769 (uid 32769):

```
Filesystem blocks quota limit grace files quota limit grace
10.241.169.54:/fb
    41   5120  5120      16     0     0
```

**USING QUOTAS WITH WINDOWS 2000:****Caution:** To use quotas as Windows 2000 Quotas, Celerra Inode, Group, & Grace Periods for Soft Quotas should be disabled!

--Quota Policy should also be set to “filesize” for Windows 2000 Quotas

--Windows 2000 Quotas will work using Usrmapper V3 or Passwd/Group files on DMs, but not NIS!

--Unless Inodes limit Quotas, User will still be able to create files, but they will be ‘empty’ 0-byte files

--MPFS clients will not be able to mount a File System with Quotas enabled

--Windows 2000 Quotas uses “bytes” for quota calculation, while Celerra uses ‘blocks’ for amount used

--Minimum Block size used with Windows 2000 Quotas is 4k v. 8k Celerra Block size

Quotas are only applied to the upper level share or "partition", not individual folders! You can set Windows Quotas on any Share if you map a drive to it. Also note that if you set a hard quota limit for a particular user, the size that the user sees for the partition will match their 'hard limit' even though the partition itself may have far more disk space available.

**Note:** When setting up Quotas from Windows 2000, make sure DataMovers have following param set first!**param quota policy=filesize***Click on Mapped Share>Properties>General Security>Quotas Tab>Quota***Status:** Disk Quotas are disabled [default level] Enable quota management Deny diskspace to users exceeding quota limit**Important:** If set to ‘No’, will allow CIFS & NFS Users to use disk space beyond SOFT & HARD Quota limits**Select the default quota limit for new users on this volume:** \*Do not limit disk usage [default setting] Limit disk space to \_\_\_\_\_ Set warning level to \_\_\_\_\_**Select the Quota logging options for this volume:** Log event when a user exceeds their quota limit Log event when a user exceeds their warning level**Quota Entries>Quota>New Quota Entry>Add new User for Quotas:****Note:** Applying quotas from the Windows 2000 Server forces a Usrmapper call so as to map a new UID for the User to which the Quota has been established [if a new User]. The User then immediately appears in the Quotas Report on the File System.**QUOTA REPORTS: By File System, DataMover, or User****QUOTA REPORT BY UID:****# nas\_quotas -r -u -fs fs20 201****Note:** If using NIS, can run reports by Username or Groupname

Report for user quotas on filesystem fs20 mounted on /fs20

| User | Bytes Used (1K) | Files | Used  | Soft     | Hard | Timeleft | Used | Soft | Hard | Timeleft |
|------|-----------------|-------|-------|----------|------|----------|------|------|------|----------|
| #201 | 69991           | 50001 | 70001 | 6.4days1 | 3001 | 01       | 01   | 01   | 01   | 01       |

**# nas\_quotas -r -u -fs fs20**

Report for user quotas on filesystem fs20 mounted on /fs20

| User | Bytes Used (1K) | Files  | Used   | Soft     | Hard | Timeleft | Used | Soft | Hard | Timeleft |
|------|-----------------|--------|--------|----------|------|----------|------|------|------|----------|
| #201 | 69991           | 50001  | 70001  | 6.4days1 | 3001 | 01       | 01   | 01   | 01   | 01       |
| #602 | 01              | 100001 | 200001 | 1        | 01   | 01       | 01   | 01   | 01   | 01       |
| #603 | 01              | 110001 | 220001 | 1        | 01   | 01       | 01   | 01   | 01   | 01       |
| #604 | 01              | 120001 | 230001 | 1        | 01   | 01       | 01   | 01   | 01   | 01       |
| #803 | 01              | 01     | 01     | 01       | 11   | 01       | 01   | 01   | 01   | 01       |

**QUOTA REPORT BY GID:**

**\$nas\_quotas -report -g -fs fs01 801**

**#nas\_quotas -r -g -m server\_7**

Report for group quotas on filesystem fs20 mounted on /fs20

| Group | Bytes Used (1K) | Files |       |          |      |      |      |          |
|-------|-----------------|-------|-------|----------|------|------|------|----------|
|       | Used            | Soft  | Hard  | Timeleft | Used | Soft | Hard | Timeleft |
| #1    | 55              | 0     | 0     |          | 8    | 0    | 0    |          |
| #201  | 69991           | 20000 | 30000 |          | 297  | 0    | 0    |          |
| #803  | 0               | 20000 | 40000 |          | 0    | 0    | 0    |          |

### **QUOTA REPORT BY FILESYSTEM:**

**\$nas\_quotas -r -fs fs01**

### **QUOTA REPORT BY DATAMOVER:**

**\$nas\_quotas -r -m server\_2**

### **QUOTA CONFIGURATION BY FILESYSTEM:**

**\$nas\_quotas -config -report -fs fs01**

### **QUOTA CONFIGURATION BY SERVER:**

**#nas\_quotas -config -report -mover server\_7**

```
+-----+
| Quota parameters for filesystem fs16 mounted on /fs16:
+-----+
| Quota Policy: blocks
| User Quota: ON
| Group Quota: OFF
| Block grace period: (1.0 weeks)
| Inode grace period: (1.0 weeks)
| Default USER quota limits:
|   Block Soft: ( 8192), Block Hard: ( 10240)
|   Inode Soft: ( 0), Inode Hard: ( 0)
| Default GROUP quota limits:
|   Block Soft: ( 0), Block Hard: ( 0)
|   Inode Soft: ( 0), Inode Hard: ( 0)
| Deny Disk Space to users exceeding quotas: YES
| Log an event when ...
|   Block hard limit reached/exceeded: YES
|   Block soft limit (warning level) crossed: YES
|   Quota check starts: NO
|   Quota Check ends: NO
```

**Note:** Running a Quotas Configuration Report will also log an entry in the Server Log and update the respective /.etc/config\_file  
2004-11-29 17:00:26: ADMIN: 4: Command succeeded: file quota /data config report=/data/.etc/config\_file

### **ESTABLISHING OR EDITING QUOTAS:**

**\$nas\_quotas -u -e -fs fs01 602** [Brings up Vi session—Edit Soft & Hard values]

**\$nas\_quotas -u -e -fs fs01 -block 1000 -i 5000 602** [Sets up Quota directly from CLI]

**\$nas\_quotas -g -e -fs fs20 -block 10000 -i 40000 801**

**# nas\_quotas -e -g -fs m3 32780** (edit specific Group quota)

### **TURNING USER OR GROUP QUOTAS ON/OFF:**

**\$nas\_quotas -on -fs fs20** [Turns on User Quotas for FileSystem that has Users defined for Quotas]

**\$nas\_quotas -g -on -fs fs20** [Turns on Group Quotas for FileSystem that has Groups defined]

**-off** [Turns Quotas off]      **-clear** [Clears all Quota records]      **-both -on** [Turns On both User/Group Quotas]

**\$ .server\_config server\_3 -v "file quota \"/fs5\"type=user off"** [Be aware fs must be unmounted to turn on or off]

-----abridged-----

1111607313: UFS: 4: USER quota is now OFF.

**\$ .server\_config server\_3 -v "file quota \"/fs5\"type=user on"**

1111607064: VLU: 5: fs 0x100000024 is being unmounted

1111607064: VMCAST: 5: UnMountIPFSCopy():Begin

1111607064: VMCAST: 5: UnMountIPFSCopy():End

1111607064: CFS: 5: DNLC: erased 11 entries

1111607064: CFS: 6: executing VFS\_VnodeFreeList::syncFsVnodes

```
1111607064: CFS: 6: done VFS_VnodeFreeList::syncFsVnodes
1111607064: CFS: 4: fs 36 being unmounted. Waiting for quiesce ...
1111607064: CFS: 5: Calling waitForQuiesce, refCounter 37
1111607064: CFS: 6: executing File_RootDirNode::~File_RootDirNode
1111607064: DHSM: 6: unmount fs obj 0x7c30004, set optionsInitialized to FALSE
1111607064: MGFS: 6: unmount fs obj 0x7c30004
1111607064: CFS: 5: Calling waitForQuiesce, refCounter 35
1111607064: STORAGE: 5: closing volume 179, ref 2, ioObj 7be7c04
1111607064: CFS: 4: fs 36 unmounted
1111607064: STORAGE: 5: opening volume 179, ref 1, ioObj 0
1111607064: UFS: 5: Cleaners ino aa7d0b8, cg aa7d014, dirty aa7d15c
1111607064: UFS: 5: current revision is 2, upgrade not needed
1111607064: UFS: 4: Quota limits will be checked on block usage, with block size of 8192 bytes
1111607064: UFS: 4: Checking USER quotas ...
1111607064: UFS: 4: Quota Check started (fs /fs5)
1111607064: UFS: 4: Done.
1111607064: UFS: 4: Quota Check completed (fs /fs5)
1111607064: UFS: 4: USER quota is ON.
1111607064: UFS: 4: Tree quota for TreeId 1, is ON
1111607064: VLU: 5: fs 0x100000024 is being mounted
1111607064: USRMAP: 7: Unable to open file: ./etc/usrmapper/usrmap.settings
1111607064: VC: 4: Enable FileSystem 36 for Virus Checking
1111607064: ADMIN: 4: Command succeeded: file quota "/fs5"type=user on
```

## **CONFIGURING USER/GROUP QUOTAS:**

1. Set Block or Inode quotas on User, Group, FileSystem, or Data Mover basis
2. For Windows 2000 Quotas, or if Quota Policy = Files, set following parameters:

/nas/site/slot\_param [These params change Quotas from 8kb Block Size to 1kb File Size quotas]

**param quota policy=filesiz**

3. Turn Quotas on [by User, Group, or Both, or Datamover] | -off | -clear

**\$ nas\_quotas -u -on -fs fs01**

**\$ nas\_quotas -g -on -fs fs01**

**\$ nas\_quotas -both -on -fs fs01**

**\$ nas\_quotas -user -on -mover server\_2**

4. Edit Quotas by File System per DataMover:

**\$ nas\_quotas -config -edit -mover server\_2**

5. Turning Off and Deleting Quotas:

**\$ nas\_quotas -u -off -fs fs01 [Turns Quotas Off]**

**\$ nas\_quotas -u -clear -fs fs01 [Clearing Quotas; -m server\_2]**

**Note:** Turning quotas "off" retains existing Quotas configuration, but "clearing" Quotas resets all values to 0 and turns quotas 'off'

6. Creating a Quota for an Anonymous User:

**\$ nas\_quotas -edit -u -fs fs01 anon**

7. Creating Group Quotas:

**\$ nas\_quotas -group -edit -fs fs01 1101**

**Note:** Quotas are applied against a User's Primary Group

8. Using Proto User to Create Quotas:

**\$ nas\_quotas -edit -u -fs fs01 -p 1000 1001 1002** [Where 1000 represents the prototype User]

9. Using Proto User to Create Group Quotas:

**\$ nas\_quotas -edit -g -fs fs01 -p 1100 1101 1102** [Where 1100 represents the prototype Group]

## **SETTING UP WINDOWS 2000 QUOTAS (aka USER QUOTAS):**

1. Since the general recommendation is to use Quota Policy Filesize when using Windows, ensure the following param is set:

**# server\_param server\_2 -facility quota -modify policy -v filesize**

server\_2 : done

Warning 17716815750: server\_2 : You must reboot server\_2 for policy changes to take effect.

**# tail nas\_log.al**

2007-01-05 13:10:48.958 server\_2:0:9857:S: server\_param server\_2 -facility quota -modify policy -v filesize

# cat /nas/server/slot\_2/param

**param quota policy=filesize** →server\_param writes entry to param file

2. Reboot Data Mover

**Server Log messages:**

2007-01-05 13:32:11: UFS: 4: Quota limits will be checked on byte usage, in units of 1024 bytes

2007-01-05 13:32:11: UFS: 4: Checking USER quotas ...

2007-01-05 13:32:11: UFS: 4: Quota Check started (fs /mensha)

2007-01-05 13:32:11: UFS: 4: Quota Check completed (fs /mensha)

2007-01-05 13:32:11: UFS: 4: USER quota is ON.

3. Map to the toplevel Share of the Celerra file system Share&gt;Properties&gt;Quota tab

4. Check the ‘Enable Quota Management’ &amp; ‘Deny disk space to users exceeding quota limit’ boxes, then set values:

\*Limit disk space to 1 GB

\*Set warning level to 800 MB.

5. Select quota logging options if desired:

--Log event when a user exceeds their quota limit

--Log event when a user exceeds their warning level

6. Click on ‘Quota Entries’&gt;Quota&gt;New Quota Entry &amp; add Users to which to apply quotas to

**Note:** Only Users object types are allowed when configuring Windows quotas—Windows does not yet use the concept of Group quotas

7. Click o.k.—Status: Disk quotas are disabled should turn from Red light and message to Green light and following message:

Status: Disk quota system is active

8. Verify User Quotas from Control Station &amp; Data Mover perspective:

**# nas\_quotas -config -report -fs mensha**

```
+
+-----+
| Quota parameters for filesystem mensha mounted on /mensha:
+-----+
| Quota Policy: filesize
| User Quota: ON
| Group Quota: OFF
| Block grace period: (1.0 weeks)
| Inode grace period: (1.0 weeks)
| Default USER quota limits:
|   Block Soft: ( 819200), Block Hard: ( 1048576)
|   Inode Soft: (    0), Inode Hard: (    0)
| Default GROUP quota limits:
|   Block Soft: (    0), Block Hard: (    0)
|   Inode Soft: (    0), Inode Hard: (    0)
| Deny Disk Space to users exceeding quotas: YES
| Log an event when ...
|   Block hard limit reached/exceeded: YES
|   Block soft limit (warning level) crossed: YES
|   Quota check starts: NO
|   Quota Check ends: NO
```

**# nas\_quotas -report -fs mensha**

Report for user quotas on filesystem mensha mounted on /mensha

```
+-----+
|User | Bytes Used (1K) | Files |
+-----+-----+-----+-----+
|     | Used | Soft | Hard |Timeleft| Used | Soft | Hard |Timeleft|
+-----+-----+-----+-----+-----+-----+
#32770 | 01512000|1048576|  | 01 01 01 |  | 01 01 01 |  |
#32771 | 01512000|1048576|  | 01 01 01 |  | 01 01 01 |  |
#32999 | 44083|512000|1048576|  | 28| 01 01 |  |
```

**/nasmed/quota/slot\_2/etc/indications****# strings indications**

```
<Indication xmlns="http://www.emc.com/schemas/celerra/core_1.0"><Modify><TreeQuota fs="73" id="0" serverName="server_2"
treeQuotaIsOn="false" userQuotaIsOn="true" groupQuotaIsOn="false" spaceSoftLimit="0" spaceHardLimit="0" spacesUsed="0" s
paceTimeLeft="0" inodeSoftLimit="0" inodeHardLimit="0" inodesUsed="0" inodeTimeLeft="0" comment="" quotaPolicy="filesize"
hardLimitEnforced="true" spaceGracePeriod="604800" inodeGracePeriod="604800" checkStartEvent="false" checkEndEvent="fa
```

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
lse" crossedSoftEvent="false" exceededHardEvent="false" defaultUserSpaceSoftLimit="0" defaultUserSpaceHardLimit="0"  
defaultUserInodeSoftLimit="0" defaultUserInodeHardLimit="0" defaultGroupSpaceSoftLimit="0" defaultGroupSpaceHardLimit="0"  
defaultGroupInodeSoftLimit="0" defaultGroupInodeHardLimit="0" ></TreeQuota></Modify></Indication>

**Note:** Indications file on /.etc rootfs of Data Mover contains quota information in XML database

# pwd

/nasmcd/quota/slot\_2/mensha/etc

# ls -lrt

```
-r--r--r-- 1 root bin 262144 Nov 3 12:55 gid_map  
-rw-rw-r-- 1 root bin 0 Jan 5 13:05 quotas  
-rw-r--r-- 1 root bin 0 Jan 5 13:42 config_file.sids  
-rw-r--r-- 1 root bin 9 Jan 5 13:42 config_file.InProgress  
-rw-r--r-- 1 root bin 60 Jan 5 13:42 config_file  
-rw-r--r-- 1 root bin 135 Jan 5 13:44 rpt_file.sids  
-rw-r--r-- 1 root bin 9 Jan 5 13:44 rpt_file.InProgress  
-rw-r--r-- 1 root bin 108 Jan 5 13:44 rpt_file
```

## **WINDOWS 2000 QUOTA DEFAULTS:**

**Quota Policy:** filesize [turned on by default]

**User Quota:** ON [User quotas turned on by default—Windows 2000 does not use Group quotas]

**QUOTAS SUPPORT FOR MPFS:** Introduced in NAS 5.0

## **CELLERRA TREE QUOTAS:**

### **CELLERRA TREE QUOTAS (NAS 5.0+):**

Support for 'QUOTAS' based on Tree Directories [Quotas based on a directory tree structure by limiting number of blocks and/or inodes that can be created in a particular directory structure—treats every directory Tree as a separate "file system"]. Usages are computed on all files in the tree, irrespective of Owner. Besides allowing Soft & Hard Quotas, can also be used to "monitor" usage by setting limits to "0". File Handles identify the Tree Quotas using a unique identifier called 'TreId'. After turning Tree Quotas on, DART tracks usage and enforces limits.

→Max of 2047 Tree Quotas per file system, with max size of Tree Quotas capped at 4TB

**Note:** Number of Tree Quotas per fs raised to 8191 with NAS 5.6.45

### **TREE QUOTAS DEFINED:**

Tree Quotas are different and separate from User or Group Quotas or Windows 2000 Quotas in that quotas are set on specific directories or "trees". Tree quotas are tracked for all files and directories in the tree regardless of owner. Tree quotas are treated separately from user/group quotas in that files created in a Tree Quota will also count against user/group quotas. Use Tree quota function to create the directory and manage Tree Quotas through CLI. Tree Quotas can be moved within the File System using Unix move command.

**Correction:** Files owned by Users that have User Quotas, even within a Tree Quota structure, does apply to User quotas.

### **PURPOSE OF TREE QUOTAS:**

--Use to limit or monitor number of files/directories [inodes], or amount of diskspace [blocks] used in a directory tree.

--Tree Quotas are tracked only within the Tree for users/groups

--Tree Quotas can contain Hard or Soft limits, and be used with Filesize or Block Quota policies

### **COUNTING ROOT USAGE IN TREE QUOTAS:**

**param quota countRootUsageInQuotaTree=0 | 1**

**Note:** With NAS 5.2 and higher, if filesystem type is MPD, then Root User usage can be tracked against TreeQuota limits, however, root user will not be restricted from going over the limit—set to 1 to enable tracking.

### **REPORTING, LISTING, & VERIFYING TREE QUOTAS CONFIGURATION:**

#### **1. Run file system list to see tree quotas in effect:**

\$ nas\_quotas -tree -list -fs fs2

```
+-----+  
| Quota trees for filesystem fs2 mounted on /fs2:  
+-----+  
| TreId| Quota tree path (Comment) |  
+-----+
```

```
| 1 | /dir3 ()
| 2 | /dir4 ()
```

**Note:** Above output shows list of tree quotas contained on file system called “fs2”. The actual tree quotas paths are /fs2/dir3, /fs2/dir4

## **2. Run Tree Quotas Reports against whole file system or specific Trees:**

**\$ nas\_quotas -tree -report -fs fs2**

Report for tree quotas on filesystem fs2 mounted on /fs2

| Tree | Bytes Used (1K)  | Files      |
|------|------------------|------------|
| #1   | 936  1000  10504 | 31  50  90 |
| #2   | 0  0  16         | 1  0  0    |

done

**\$ nas\_quotas -tree -report -fs fs2 1**

Report for tree quotas on filesystem fs2 mounted on /fs2

| Tree | Bytes Used (1K)  | Files      |
|------|------------------|------------|
| #1   | 936  1000  10504 | 31  50  90 |

Done

## **3. Review Specific Quota Configuration:**

**\$ nas\_quotas -report -config -fs fs2 1**

```
+-----+
| Quota parameters for filesystem fs2 mounted on /fs2:
+-----+
```

### **| Quota Policy: blocks**

- | User Quota: OFF
- | Group Quota: OFF
- | Block grace period: (1.0 weeks)
- | Inode grace period: (1.0 weeks)
- | Default USER quota limits:

- | Block Soft: ( 0), Block Hard: ( 0)
- | Inode Soft: ( 0), Inode Hard: ( 0)

- | Default GROUP quota limits:

- | Block Soft: ( 0), Block Hard: ( 0)
- | Inode Soft: ( 0), Inode Hard: ( 0)

### **| Deny Disk Space to users exceeding quotas: YES**

- | Log an event when ...

- | Block hard limit reached/exceeded: NO

- | Block soft limit (warning level) crossed: NO

- | Quota check starts: NO

- | Quota Check ends: NO

**Note:** The ‘Deny Disk Space...’ value must be set to ‘YES’ in order for Tree Quotas to honor the hard limit. It does mention this in the 5.4 Quotas documentation, but it could be more clearly stated.

## **QUERYING TREE QUOTA INFO USING NAS\_FS:**

**# nas\_fs -query:"\*IsRoot==False:Type==uxfs" -fields:TreeQuotas -format:"%q" -query:"\*" -fields:ID,Filesystem,Path -format:"%s,%s, %s\n"**

```
2.coreprod,
1.coreprod, /options_new
2.linux, /lx7_sadmin
1.linux, /lx386_tp
```

## **DISPLAYING TREE QUOTAS DIRECTLY FROM BROWSER TO CONTROL STATION:**

<https://10.241.169.53/action/treeQuotaDisplay>

## **TREE QUOTAS RECALC TOOL (Quota Check Tool):**

**Problem:** Tree Quota usage may not match true fs usage--why?

-->The quota policy was changed after tree quotas created

-->The quotas.tree file is missing or corrupt

-->A software bug caused the calculation to be incorrect

**Previously, not an easy fix--Unlike user & group quotas, tree quotas can't be enabled or disabled when data exists—no easy way to recalculate usage--the only procedure was:**

-->Move all the data from the tree to a temporary location

-->Delete and recreate the tree

-->Repopulate the tree with the data

### **Normal Tree Quota differences that are not real problems:**

-->Quota policy when the tree was created was set to filesize. du reports block usage so it will likely be larger than what is returned by the quota report. If there are a large number of small files in the tree the du output and the quota report could differ greatly.

-->param quota countRootUsageInQuotaTree was set to 0 when the tree was created or the tree was created before 5.2.12 when root usage wasn't counted against the limit. Changing this param has no effect on the trees created earlier even if you run the recal tool.

## **USING THE TREE QUOTA RECALC/CHECK TOOL:**

**Note:** This tool only applies to issues related to space & inode usage within a Tree—compare the du output of a specific tree quota (/tree) to the Tree Quotas report to determine if a true size discrepancy exists. Functionally, the Tree Quota Check Tool traverses the quota tree & recomputes usage, as well as usages for User & Group quotas (if turned ON)—available in NAS 5.3.23.0 & 5.4.18.0.

### **1. Navigate to the file system through the rootfs:**

**# cd /nas/rootfs/slot\_2/fs01**

### **2. Locate appropriate Tree Quota in question:**

**# nas\_quotas -l -t -fs fs01**

| Quota trees for filesystem fs01 mounted on /fs01:

|                                   |         |         |
|-----------------------------------|---------|---------|
| +-----+                           | +-----+ | +-----+ |
| TreeId  Quota tree path (Comment) | +-----+ |         |
| +-----+                           | +-----+ |         |
| 1   /tree1 ()                     |         |         |
| 2   /tree2 ()                     |         |         |

### **3. Check the size of the Tree Quota using du command(Run during offpeak time):**

**# du -s tree1**

10712064 tree1

### **4. Run Tree Quota Report to determine if discrepancy exists:**

**# nas\_quotas -r -t -fs fs01 1**

Report for tree quotas on filesystem fs01 mounted on /fs01

|                                                              |                     |                     |
|--------------------------------------------------------------|---------------------|---------------------|
| +-----+-----+-----+                                          | +-----+-----+-----+ | +-----+-----+-----+ |
| Tree  Bytes Used (1K)  Files                                 | +-----+-----+-----+ |                     |
| +-----+-----+-----+                                          | +-----+-----+-----+ | +-----+-----+-----+ |
| Used   Soft   Hard   Timeleft  Used   Soft   Hard   Timeleft | +-----+-----+-----+ |                     |
| +-----+-----+-----+                                          | +-----+-----+-----+ | +-----+-----+-----+ |
| #1  10815144  01 01   01 01   01 01                          | +-----+-----+-----+ |                     |

### **5. If true discrepancy exists, it can be fixed by running the following recal command:**

**#.server\_config server\_2 -v "file quota /fs01 type=tree check path=/tree1"**

2005-09-27 22:16:32: UFS: 4: Performing treeQuotaCheck on directory tree

**Note:** Use following syntax for VDMs

**\$ .server\_config movername -v "file quota /vdm\_name/file\_system type=tree check path=/quota\_tree\_path"**

### **6. Verify effectiveness by rerunning Tree Quota report:**

**# nas\_quotas -report -tree -fs fs01 1**

Report for tree quotas on filesystem fs01 mounted on /fs01

|                                                              |                     |                     |
|--------------------------------------------------------------|---------------------|---------------------|
| +-----+-----+-----+                                          | +-----+-----+-----+ | +-----+-----+-----+ |
| Tree  Bytes Used (1K)  Files                                 | +-----+-----+-----+ |                     |
| +-----+-----+-----+                                          | +-----+-----+-----+ | +-----+-----+-----+ |
| Used   Soft   Hard   Timeleft  Used   Soft   Hard   Timeleft | +-----+-----+-----+ |                     |
| +-----+-----+-----+                                          | +-----+-----+-----+ | +-----+-----+-----+ |

**TREE QUOTA SIZE ISSUE (AR70615):**

At times, quotas.tree file can grow to be excessively large, though 2047 tree quotas limit per file system is supposed to prevent this from happening—use following procedure to truncate file:

1. Upload truncate executable to file system ./etc directory
2. Make backup copy of quotas.tree file if feasible
3. Run truncate command to truncate quotas.tree to 65536 bytes: #./truncate quotas.tree 65536
4. Verify that file was truncated to 64k
5. Unmount and then remount the file system
6. Run quota report against file system
7. Create test tree quota and view using tree report

**Hard Quota Events:** Server Log Error; QUOTA EXCEEDED (EDQUOT) Error returned

**Soft Quota Events:** Server Log Message; Grace Period Starts—once expired will act as a Hard Quota

Limiting Quota Event Notification: After a hard limit is hit, we enable the event generation only after the files are deleted below the soft limit. This is the behavior of bsd (UNIX) quotas and Celerra behaves the same.

**TABLE OF EVENT ACTIONS:**

| Action                                                         | Server_Log | Email |
|----------------------------------------------------------------|------------|-------|
| Soft quota hit                                                 | yes        | yes   |
| Soft quota hit after files deleted                             | yes        | no    |
| Hard Quota hit                                                 | yes        | yes   |
| Hard quota hit after files deleted                             | yes        | no    |
| Soft quota hit after files deleted<br>after hard quota reached | yes        | yes   |
| Hard quota hit                                                 | yes        | yes   |

**SETTING UP TREE QUOTAS:****1. Customer should decide on Policy to use: Blocks or FileSize**

Blocks-->Usage calculated in 8k blocks, the default

FileSize-->Bytes added or removed, recommended for cifs, limits & warnings based on 8k block

**param quota policy=filesize**

**2. Create directory path under mountpoint, to which TreeQuotas are going to be applied:**

**# nas\_quotas -on -tree -fs fs2 -path /proj1/user1**

done

**Note:** This creates TreeQuota called “user1” id=1 in the path /fs2/proj1/user1. The directory “user1” cannot exist prior to creating the TreeQuota and must be created by the TreeQuota command.

**# nas\_quotas -on -tree -fs fs2 -path /proj2/user2**

done

**3. Run List to verify new TreeQuotas created:**

**# nas\_quotas -list -tree -fs fs2**

```
+-----+
| Quota trees for filesystem fs2 mounted on /fs2:
+-----+
|TreeIdl Quota tree path (Comment) |
+-----+
| 1 |/proj1/user1 () |
| 2 |/proj2/user2 () |
```

**4. Edit the TreeQuotas and apply Hard & Soft Limits:**

**# nas\_quotas -edit -tree -fs fs2 1**

treeid : 1

fs fs2 blocks (soft = 1000, hard = 10000) inodes (soft = 0, hard = 0)

**Note:** Above sets soft quota at 1MB and hard quota at 10MB. Default values of 0 for inodes means that no values are set and are therefore unlimited.

**# nas\_quotas -tree -edit -block 20000:10000 -inode 20000:10000 -fs fs2 1**

done

**Note:** Above command edits Hard & soft quota, respectively, for both ‘Blocks/Bytes’ or ‘Inodes/Files’ without use of vi

**5. Export the TreeQuota directory for CIFS:**

# server\_export server\_2 -P cifs -n project1 /fs2/proj1/user1

server\_2 : done

# server\_export server\_2 -P cifs -n project2 /fs2/proj2/user2

server\_2 : done

# server\_export server\_2

share "project1" "/fs2/proj1/user1" maxusr=4294967295 umask=22

share "project2" "/fs2/proj2/user2" maxusr=4294967295 umask=22

## **6. Set NT permissions on the TreeQuota folders**

### **7. Make Quota configuration changes using the following:**

# nas\_quotas -config -edit -fs fs2

File System Quota Parameters:

fs fs2

Block Grace: (1.0 weeks)

Inode Grace: (1.0 weeks)

\* Default Quota Limits:

User: (soft = 0, hard = 0) inodes (soft = 0, hard= 0)

Group: (soft = 0, hard = 0) inodes (soft = 0, hard= 0)

Deny disk space to users exceeding quotas: (yes)

\* Generate Events when:

Quota check starts: (no)

Quota check ends: (no)

User's soft quota crossed: (no)

User's block quota exceeded: (no)

## **8. Run TreeQuota Reports or Obtain Config Information:**

# nas\_quotas -tree -list -fs fs2

# nas\_quotas -config -report -fs fs2

## **DELETING TREE QUOTAS:**

### **1. If NFS or CIFS exported, unexport first**

### **2. Obtain TreeQuotas list:**

\$ nas\_quotas -tree -list -fs fs2

```
+-----+
| Quota trees for filesystem fs2 mounted on /fs2:
+---+-----+
|TreeIdl Quota tree path (Comment)           |
+---+-----+
| 1 | /dir3 ()                           |
| 2 | /dir4 ()                           |
```

### **3. Delete data from the TreeQuotas & then delete the quota tree directory:**

\$ nas\_quotas -off -tree -fs fs2 -path /dir3

done

\$ nas\_quotas -off -tree -fs fs2 -path /dir4

done

**Note:** Deletes TreeQuotas directory called "dir3" and "dir4"--directory must be empty of all data before this command will succeed

### **4. Clear TreeQuotas for the file system:**

\$ nas\_quotas -clear -tree -fs fs2

done

## **TREE QUOTA COMMANDS:**

\$nas\_quotas -tree -on -fs fs06 -path /fs06/Tree [Turns Quotas On & creates dir. "Tree" in path /fs06/fs06/Tree]

\$nas\_quotas -tree -edit -fs fs06 -path /fs06/Tree [Edit Quotas]

\$nas\_quotas -tree -report -fs fs06 -path /fs06/Tree [Reporting]

\$nas\_quotas -tree -off -tree -fs fs06 -path /fs06/Tree [Turn quotas off first, then -clear to remove]

\$nas\_quotas -tree -clear -fs fs06 [Clears Tree Quotas from FS indicated]

**Note:** Command syntax changes slightly from NAS 5.0 to 5.1 in setting up Tree Quota path structure

\$ .server\_config server\_3 -v "file quota \"/fs5\"type=tree off"

\$ .server\_config server\_3 -v "file quota \"/fs5\"type=tree on"

### RUNNING TREE QUOTAS REPORTS:

\$nas\_quotas -tree -report -fs fs06 1 [Issues report for Tree Quotas for specified tree]

\$nas\_quotas -tree -report -fs fs06 [Generates Quota Report on all quota trees]

### LISTING TREE QUOTAS—VERIFYING:

# nas\_quotas -tree -list -fs treefs

```
+-----+
| Quota trees for filesystem treefs mounted on /treefs:
+-----+
|TreeId| Quota tree path
+-----+
| 1 | /tree1
| 2 | /tree2
| 4 | /tree4
| 5 | /tree5
```

### SETTING TREE QUOTAS LIMITS:

1. Set limits on Tree: \$nas\_quotas -tree -edit -fs fs06 1 [Opens up Tree Quota Editor for Hard & Soft limits]

  fs fs06 blocks (soft=2000, hard=5000) inodes (soft=2000, hard=5000)

2. # nas\_quotas -tree -edit -block 20000:10000 -inode 20000:10000 -fs fs2 1 [Example sets Hard & Soft Tree Quota Limits, resp.]

3. Set Grace Period: \$nas\_quotas -config -edit -fs fs06 [Opens editor to allow you to set Grace periods]

Note: Quotas limits are enabled with 0 limits by default. Use –edit command to setup block & inode hard & soft limits

# nas\_quotas -tree -edit -block 20000:10000 -inode 20000:10000 -fs fs2 1

done

# nas\_quotas -tree -report -fs fs2

Report for tree quotas on filesystem fs2 mounted on /fs2

```
+-----+
| Tree   | Bytes Used (1K) | Files
+-----+-----+-----+
|       | Used | Soft | Hard |Timeleft| Used | Soft | Hard |Timeleft|
+-----+-----+-----+
#1    | 0| 10000| 20000|      | 1| 10000| 20000|      |
#2    | 0| 0| 0|      | 1| 0| 0|      |
```

### SERVER LOG ENTRIES WHEN SOFT & HARD TREE QUOTAS ARE REACHED:

2005-07-13 16:13:22: UFS: 4: Block soft quota crossed (fs /fs2, treeid 1)

2005-07-13 16:23:11: UFS: 3: Block hard quota reached/exceeded (fs /fs2, treeid 1)

### NAS 5.3 FEATURES:

#### CREATING TREE QUOTA COMMENTS:

\$nas\_quotas -tree -on -fs fs06 -comment <comments> -path /fs06/Tree

Note: NAS 5.3 introduces Quota\_Descriptor feature, which are nothing more than comment fields

\$nas\_quotas -tree -on -fs fs06 -comment “ -path /fs06/Tree [removes comment]

#### EDITING TREE QUOTA LIMITS:

\$nas\_quotas -edit -tree -fs fs06 1

#### CHANGING OWNERSHIP AND PERMISSIONS:

#chown 6009 /quotaTrees/tree1

#chmod 755 /quotaTrees/tree1

Note: To set Quota Descriptor, must have QuotaTree itself created, though the last component of the tree (tree1) must not exist

### TYPICAL CONTENTS OF FS .ETC FOLDER FOR TREE QUOTAS:

```
# ls -l /nas/rootfs/slot_2/fs_A0/etc
-rw-r--r-- 1 root bin 60 Mar 12 09:08 config_file
```

```
-rw-r--r-- 1 root bin 9 Mar 12 09:08 config_file.InProgress  
-rw-rw-rw- 1 root bin 0 Mar 12 09:08 quotas  
-rw-rw-rw- 1 root bin 44 Mar 4 11:53 quotas.config  
-rw-rw-rw- 1 root bin 96 Mar 12 09:29 quotas.tree  
-rw-r--r-- 1 root bin 1028 Mar 12 09:32 rpt_file  
-rw-r--r-- 1 root bin 0 Mar 12 09:32 rpt_file.InProgress  
drwxr-xr-x 2 root bin 80 Mar 12 09:29 treeQuotaDB  
# ls -l /nas/rootfs/slot_2/fs_A0/.etc/treeQuotaDB  
lrwxr-xr-x 1 root bin 19 Mar 4 11:55 1 -> 10001.1a.b.40475f8a
```

## ARCHIVING, DELETING, CLEARING, RESTORING TREE QUOTAS:

**Note:** Might use the following method to delete a Tree Quota from a file system or as a way to correct a corrupted database

1. **#tar -tcvf tree1.tar.gz /fs1/tree1/\***
2. Remove Quota Tree Information: **#rm -rf /fs1/tree1/\***
3. Disable Quotas for FS1: **\$nas\_quotas -tree -off -fs fs1 -path /tree1**
4. Clearing Quotas for FS1: **\$nas\_quotas -tree -clear -fs fs1**
5. Recreate Quota Tree: **\$nas\_quotas -tree -on -fs fs1 -path /tree1** [May need to chown & chmod directory]
6. Restore Tree Database: **#tar -tzvf tree1.tar.gz /fs1/tree1/**

## HANDLING SPACES IN TREE QUOTAS PATH STATEMENTS:

**#nas\_quotas server\_2 -tree -on -fs filesystem1 -path "/department share"**  
or **#nas\_quotas server\_2 -tree -on -fs filesystem1 -path /department\share**

## TREE QUOTA LIMITATIONS:

- Max. of 2047 Quota Trees per file system
  - Tree Quotas cannot be nested
  - Tree Quotas can be moved within the same FS
  - Tree Quotas limited to 4TB due to 32bit O/S
  - Max. file size restricted to 4TB (vs. normal 16TB) when Quota policy set to ‘filesize’
  - Name of tree quota directory cannot exist on filesystem prior to setting up with Tree Quotas commands!
  - Cannot view a User’s Tree Quota Usage from Explorer Quotas Tab—Celerra returns total fs space information.
- Note:** NAS 5.1.16.0 introduces ability to return User, Group, or Tree Quotas information [to be changed with 5.1.18.2]

## ENABLING WRITES TO A 4TB TREE QUOTA AT 100%:

→Though the limitation of being able to report on a maximum of 4TB worth of Quota information still exists, the following parameter change allows customers to keep writing to a file system even when the 4TB limit has been reached. This change is effective with NAS 5.5.38 and 5.6.42:

```
# nas_quotas -tree -list -fs cifs_xp1  
+-----+  
| Quota trees for filesystem cifs_xp1 mounted on /cifs_xp1:  
+-----+  
| TreeId| Quota tree path (Comment) |  
+-----+-----+  
| 1 | /treed () |  
$ .server_config server_2 "file quota /cifs type=tree allowOverflow treeid=1"  
$ .server_config server_2 "file quota /cifs type=tree allowOverflow list"
```

## TREE QUOTAS LIMITATION FIXED FOR NFS AND CIFS CLIENTS:

Problem is that Users see only full file system quota and not the tree directory. AR29527 fixes for CIFS [5.1.16.0/5.2.1.0] and AR34993 fixes for NFS [5.1.20.0].

## CELLERRA FILE SYSTEM COMMANDS:

**FileSystem Commands:** #server\_df ALL #server\_export [NFS or CIFS] #server\_mount #server\_file [copies files from CS to DM and vice versa] #server\_archive [archiving] #nas\_fs -l #nas\_fs -n ufs1 -c vol1  
#nas\_volume -l #nas\_volume -n meta1 -c -M d4 #nas\_volume -s meta1 [verify volume size]

**# showmount -e 172.19.32.10** [shows the Exported File Systems on the indicated IP Address]

#server\_df server\_2 or #server\_df ALL #mount -F nfs server\_3:/mount1 /mnt [to mount remote filesystem]  
[NAS 4.0 introduces support for grouping of file systems for TimeFinder/FS operations]

**EXAMPLE:** \$fs\_group -n grp1 -c fs1 fs2 fs3

**Creates Group:** grp1 [Type 100 FileSystem]

## **CELERRA AUTOMATIC FILE SYSTEM USAGE NOTIFICATION HWM (File System Full):**

**Note:** By default, the Celerra will automatically log an event to /nas/log/sys\_log when any file system reaches 90% full

**Celerra Checks using dskMon and Cron Job every (3) Minutes:**

# cat /nas/site/cron.d/nas\_sys

3 \* \* \* \* root /nas/sbin/dskMon > /dev/null 2>&1

# .server\_config server\_2 -v "param file"

| Name                 | Location   | Current    | Default                                |
|----------------------|------------|------------|----------------------------------------|
| file.fsSizeThreshold | 0x01b890c0 | 0x0000005a | 0x0000005a → 5a hex = 90 decimal = 90% |

**Example /nas/log/sys log:**

Dec 11 08:40:33 2006 CFS:4:1 Slot 2: 1165855248: filesystem size threshold (90%) crossed (fs /sector)

**Note:** An event also shows up under Celerra Manager for CFS

## **CHANGING THE DEFAULT 90% FULL FS HWM USAGE NOTIFICATION:**

**Example of changing from 90% default to 80% full:**

**param file fsSizeThreshold=80**

**Using server param to modify value:**

\$ server\_param server\_2 -facility file -modify fsSizeThreshold -value 80

**Note:** Requires server reboot

**Using dskMon.cfg to set different file systems to different HWM thresholds:**

1. vi edit the /nbsnas/etc/dskMon.cfg file to specify file systems, change the default, disable the feature, etc.

# The default high-water mark for all FSeS is 90%. To change this

# default, uncomment the following line and change 90 to whatever

# you want it to be:

# DefaultHiWater 90 → Excerpt from the dskMon.cfg file

## **CREATING FILE SYSTEM HWM ALERTS FROM CELERRA MANAGER:**

Notifications>File System Usage>create new HWM notification and indicate email or SNMP trap

**Automatically creates a new JServer event file and loads on the Celerra:**

# nas\_event -L -i

3: /nas/jserver/event\_config/events.cfg

# pwd

**/nas/jserver/event\_config**

# ls -la

-rw-r--r-- 1 root root 109 Dec 11 10:41 events.cfg

# cat events.cfg

# This file has been automatically generated by JServer Facility

# JServer Facility

#

facilitypolicy 135, 4

**Events are logged /nas/log/sys log & appear on Celerra Manager status page:**

Mon Dec 11 09:50:21 PST 2006 Warning JServer Alert W0 (space); File System 'sector'; value=94.9% → Celerra Manager Status

Dec 11 09:50:21 2006 JServer:4:201 Alert W0 (space); File System 'sector'; value=94.5% → /nas/log/sys\_log

## **HOW TO CREATE FILE SYSTEMS USING A SPECIFIC DATAMOVER:**

**Note:** By default, all file systems will attempt to use "server\_2" to build the file systems

#nas\_fs -n mail5 -c mtv05 -o mover=server\_9

## **QUERYING FOR FILE SYSTEM USING XHMP:**

# nas\_fs -query:id==37 -fields:Name,type -format:"%s %s\n"

EPWxtra2\_DR uxfs

**Nas slice Command:** -l Slice Table -d delete -n name -c create -r rename -i details \$nas\_slice -n slv24 -d d24 4096

**\$nas\_slice -n slv25 -d d24 max** [to use the rest of the device for the slice volume]

**Nas volume Command:** -l Celerra Volume Table & names -d -n -r -c {-S stripe size -M meta[meta=default volume type\*]}

\***Note:** That is, if you do not specify a Volume Type, it will create a MetaVolume by default!!!!

-s total size -x extend a volume -i detailed info -acl -C to clone volume {volume\_name; disktype=; svol:dvol}

**\$nas\_volume -n str5 -c -S 8192 slv24,slv25** [Creating Stripe Volume]

**\$nas\_volume -n mtv1 -c -M str4,str5,str6** [Creating Meta Volume]

## **VERIFYING FILE SYSTEM SPACE ON DM:**

**\$server\_df server\_4 or \$server\_df server\_4 ufs201**

## **EXTENDING CELERRA FILESYSTEM:**

**CAUTION:** Recommend failing over to Standby DataMover to conduct File System Extend!

- Step 1. Locate Available Devices to Build New MetaVolume(s): \$nas\_disk -l \$nas\_volume -s d24 [specific info on volume]
- Step 2. Create new Slice or Stripe or Meta: In this case, need only a 2GB Slice from a 15GB d24 Volume:

**\$nas\_slice -n slv24 -c d24 2048**

- Step 3. Create New MetaVolume from Slice: \$nas\_volume -n mtv24 -c slv24
- Step 4. Reboot DataMover to flush RAM & thereby provide more resources for it to create “Extended” file system.
- Step 5. Have a Standby Server for the Primary system and test failover... [need to define why]
- Step 6. Extend File System: \$nas\_fs -x g3ufs7 mtv24 [yields g3mtv7]

**Note:** You can Extend File Systems using “MetaVolumes”; “Symm Device Volumes”; or “Slice Volumes”

**Example:** \$nas\_fs -xtend g1ufs13 d70                   \$nas\_fs -xtend g1ufs13 slc70

## **DATA MOVER ROOTFS EXTENSION PROCEDURE (extending rootfs):**

### **Introduction:**

With NAS 5.2, new design features were introduced [SecMap Caching, Internal Usermapper, VDMs] that did not fully take into account the space and inode requirements imposed on the rootfs of each Data Mover, in many cases leading to the disabling of SecMap Caching and inability to upgrade from External to Internal Usermapper. SecMap Caching requires (2) rootfs inodes for every UID or GID mapping made. Internal Usermapper requires (4) rootfs inodes for every UID or GID mapping made. The default size of the rootfs is 128MB, with a total of 16,000 inodes (1 inode/8k block) available for system use. With new installations of NAS 5.2.10.3 and higher, the inode density of the rootfs was increased from 16,000 to 130,000 inodes (1 inode/1k block), yet remains 128MB in overall size.

### **CAUTION!**

- If disks are not available, an emergency CCA should be submitted to add storage so that the rootfs extension can occur
- The extension of the rootfs should only be done by a qualified EMC Field Service person
- DO NOT extend the root\_fs with volumes built on ATA drives or any other drive technology than Fibre Channel RAID 5 4+1 disks (see enhancement AR145405)
- DO NOT extend the root\_fs with volumes built on anything other than the storage system where the Celerra boots from
- Use extreme care before issuing the rootfs extend command so as not to mistakenly overextend the rootfs
- With NAS versions below 5.2.19.0 and 5.3.13.0, extending the root file system to a size greater than 48G will prevent Usermapper from being able to calculate new uid and gids

### **ROOTFS FULL ISSUES:**

- Unable to create new Shares
- Unable to map new Users or Groups
- Unable to modify the localgroups database
- Unable to create new CIFS VDM container

**Note:** Keep in mind that Secmap and Internal Usermapper entries use symbolic links that are often larger than 83 bytes, resulting in allocation of extra 8kb blocks, and may result in more actual disk space being used than reported by the Linux Control Station du command. See AR81639.

## **EXTENDING DM ROOT FILE SYSTEM TO GAIN MORE SPACE & INODE CAPACITY:**

### **1) VERIFY CURRENT ROOTFS SPACE & INODE USAGE:**

**# server\_df server\_2**

```
server_2 :
Filesystem      kbytes   used   avail capacity Mounted on
root_fs_common  13624    2072   11552  15%   /.etc_common
root_fs_2       128592   119432  9160   92%   /
-->Ability to add to localgroups db may be disabled at 92% full mark
```

**# server\_df server\_2 -i**

```
server_2 :
Filesystem      kbytes   used   avail   capacity   Mounted on
shared          206515184 95339592 111175592 46%   /root_vdm_1/shared
                114592    1984    112608    2%   /root_vdm_1/.etc
root_fs_common  13624    5264    8360    39%   /.etc_common
root_fs_2       114592   108872   5720    95%   /
-->Inode threshold of 95% has been reached
```

### **2) VERIFY AVAILABILITY OF FREE STORAGE FOR EXTENSION:**

#### **a.) Find available d volume to use:**

**# nas\_disk -l**

```
id inuse sizeMB storageID-devID type name servers
1 y 11263 APM00023801040-0000 CLSTD root_disk 1,2
2 y 11263 APM00023801040-0001 CLSTD root_ldisk 1,2
7 n 108323 APM00023801040-0010 CLSTD d7 2,1
```

**# nas\_volume -s d7**

total = 108323 avail = 108323 used = 0 ( 0% ) (sizes in MB)

Note: d7 has 108GB of space available, so performing an extension is possible

**b.) Alternatively, identify a pool with space available for the extension:**

**# nas\_pool -list**

```
id inuse acl name
3 y 421 clar_r5_performance
```

**# nas\_pool -size clar\_r5\_performance**

```
id = 3
name = clar_r5_performance
used_mb = 900848
avail_mb = 2372920
total_mb = 3273768
potential_mb = 381137
```

**3) VERIFY CURRENT STATUS & NAME OF ROOT FILE SYSTEM:**

**# /nas/sbin/rootnas\_fs -s root\_fs\_2**

```
total = 111 avail = 110 used = 1 ( 0% ) (sizes in MB) ( blockcount = 262144 )
volume: total = 128 (sizes in MB) ( blockcount = 262144 )
```

**# /nas/sbin/rootnas\_fs -i root\_fs\_2**

```
id = 2
name = root_fs_2
acl = 0
in_use = True
type = udfs
worm = off
volume = root_volume_2
```

**# /nas/sbin/rootnas\_volume -i root\_volume\_2**

```
id = 12
name = root_volume_2
acl = 0
in_use = True
type = meta
volume_set = root_slice_2,root_s70_2
disks = root_disk,root_ldisk
clnt_filesys= root_fs_2
```

**4) USE OPTION I or II BELOW WHEN EXTENDING ROOTFS ON SYMMETRIX OR CLARIION**

**I. EXTENDING ROOTFS ON SYMMETRIX OR CLARIION STORAGE USING SLICE FROM CLI:**

**1.) Create Slice using root\_nas command for rootfs (Server 2 in this example):**

**# /nas/sbin/rootnas\_slice -n root\_slice\_2a -c d7 2048** (root\_slice\_2a is just the slice name to be used for the extension)

```
id = 1170284631
name = root_slice_2a
acl = 0
in_use = False
slice_of = d7
offset(MB)=0
size (MB)= 2048
volume_name = root_slice_2a
```

**2.) Extend Root File System from Slice created in previous Step:**

**# /nas/sbin/rootnas\_fs -x root\_fs\_2 root\_slice\_2a**

```
id = 2
name = root_fs_2
acl = 0
in_use = True
```

```
type      = uxf5
worm     = off
volume   = root_volume_2
pool      =
rw_servers= server_2
ro_servers=
rw_vdms  =
ro_vdms  =
auto_ext = no,virtual_provision=no
stor_devs = APM00023801040-0000,APM00023801040-0001,APM00023801040-0010
disks    = root_disk,root_ldisk,d7
```

**Note:** A 2GB extension takes a couple of minutes and can be done with the Server online. However, EMC recommends that this procedure be done during regular maintenance windows as CIFS access can be impacted if done 'online' on a busy server.

### **3.) Verify extension results:**

```
# /nas/sbin/rootnas_fs -s root_fs_2
```

```
total = 1935 avail = 1934 used = 1 ( 0% ) (sizes in MB) ( blockcount = 4456448 )
volume: total = 2176 (sizes in MB) ( blockcount = 4456448 )
```

## **II. EXTENDING ROOTFS ON SYMM OR CLARIION STORAGE USING STORAGE POOLS--AVM:**

### **1.) Use following command syntax to extend rootfs using CLARiiON or SYMMETRIX STORAGE POOLS:**

```
# /nas/sbin/rootnas_fs -xtend root_fs_2 size=2G pool=clar_r5_performance (pool=symm_std) -option slice=y
```

```
id      = 2
name   = root_fs_2
acl     = 0
in_use  = True
type    = uxf5
worm    = off
volume  = root_volume_2
pool    = clar_r5_performance
member_of = root_avm_fs_group_3
rw_servers= server_2
ro_servers=
rw_vdms  =
ro_vdms  =
auto_ext = no,virtual_provision=no
stor_devs = APM00023801040-0000,APM00023801040-0001,APM00023801040-0010,APM00023801040-002B,APM00023801040-0024,APM00023801040-0023,APM00023801040-0028
disks   = root_disk,root_ldisk,d34,d18,d30,d20
```

**CAUTION:** Use extreme care before executing command to ensure -option slice=y is correctly chosen

### **2.) Verify Extension:**

```
# server_df server_2
```

```
server_2 :
Filesystem      kbytes   used   avail capacity Mounted on
root_fs_2        1978504   1152   1977352  0%   /
# server_df server_2 -i
```

```
server_2 :
Filesystem      inodes   used   avail capacity Mounted on
root_fs_2        1985982   383   1985599  0%   /
```

## **USING ROOT LDISK RESERVE VOLUME OPTION\*:**

**root\_ldisk\_reserve (reserved volume built on Control LUN 1)**

**root\_disk\_reserve (reserved volume built on Control LUN 0)**

**Note:** root\_ldisk\_reserve begins on LUN 1 at the end of the (16) 64MB UFSLogs, offset 1024MB. The ldisk\_reserve volume is currently used during upgrades to extend rootfs file systems (to bring up to 128MB), and for increasing the size of the panic handlers

**\*CAUTION: This method is no longer sanctioned by Engineering & must not be used. If data volume cannot be used for an extension, contact TS2/EE—there are situations that may necessitate the use of the root\_ldisk\_reserve. Make sure that you are using the root\_ldisk\_reserve and NOT root\_disk\_reserve**

**1) Verify Root File System Space:****# server\_df server\_2**

```
server_2 :
Filesystem      kbytes   used   avail capacity Mounted on
root_fs_common  13624    2072   11552  15%   /.etc_common
root_fs_2       128592   119432  9160   92%   /

```

**2) Verify Root File System Inode Availability:****# server\_df server\_2 -i**

```
server_2 :
Filesystem      inodes   used   avail capacity Mounted on
root_fs_common  21822    18     21804  0%   /.etc_common
root_fs_2       15865   6875   8990   54%   /

```

**3) To use root\_ldisk\_reserve volume to extend the rootfs:****Note:** Determine if root\_ldisk\_reserve volume has capacity available to extend rootfs**# nas\_volume -l**

```
91      y 1 0  root_ldisk_reserve 0  69,70,71,72

```

**# /nas/sbin/rootnas\_volume -s id=91**

total = 10239 avail = 9999 used = 240 ( 2% ) (sizes in MB)

**Note:** Above output shows approximately 10GB of space available, which means that a 2GB extension can be done**4) Verify Naming Convention for Current Rootfs Volumes:****# /nas/sbin/rootnas\_fs -i root\_fs\_2**

```
id      = 2
name    = root_fs_2
acl     = 0
in_use  = True
type    = uxf
volume  = root_volume_2
#/nas/sbin/rootnas_volume -i root_volume_2
id      = 464
name    = root_volume_2
acl     = 0
in_use  = True
type    = meta
volume_set = root_slice_2,root_s70_2 -->Highlighted slice is the name of current slice used for original rootfs
disks   = root_disk,root_ldisk
Int_filesys= root_fs_2
```

**I. EXTENDING ROOTFS FOR SYMMETRIX STORAGE:****1) Create Slice using root\_s70 naming convention (SYMMETRIX STORAGE)--Make sure that you are using the root\_ldisk\_reserve and NOT root\_disk\_reserve as follows:****# /nas/sbin/rootnas\_slice -n root\_s70\_2a -c root\_ldisk\_reserve 2048** [2048 indicates 2GB slice]

```
id      = 84
name    = root_s70_2a
acl     = 0
in_use  = False
slice_of = root_ldisk_reserve
offset(MB)= 240
size (MB)= 2048
volume_name = root_s70_2a
```

**2) Extend Root File System from Slice (SYMMETRIX STORAGE):****# /nas/sbin/rootnas\_fs -x root\_fs\_2 root\_s70\_2a**

```
id      = 2
name    = root_fs_2
acl     = 0
in_use  = True
type    = uxf
volume  = root_volume_2
pool    = clar_r5_performance
member_of = root_avm_fs_group_3
rw_servers= server_2
```

```
ro_servers=
rw_vdms =
ro_vdms =
symm_devs = APM00040303779-0000,APM00040303779-0001,APM00040303779-0010
disks = root_disk,root_ldisk,d7
```

## **II. EXTENDING ROOTFS FOR CLARIION STORAGE:**

### **1.) Use following command syntax to extend CLARiiON-Based Celerra Systems:**

```
# /nas/sbin/rootnas_fs -xtend root_fs_2 size=2G pool=clar_r5_performance -option slice=y
```

### **2.) Verify Extension:**

```
# server_df server_2
```

```
server_2 :
Filesystem      kbytes   used   avail capacity Mounted on
root_fs_2       1047920  110368  937552  11%   /
# server_df server_2 -i
```

```
server_2 :
Filesystem      inodes   used   avail capacity Mounted on
root_fs_2       1047550  73673   973877  7%   /
# server_df server_2 -i
```

**Note:** Latest guidance is to extend a ROOTFS by at least (2) GB but may need to be larger, depending on default inode density and size of User/Group database. For Clariion Backends that do not have an 11GB root\_ldisk\_reserve volume available, there may be the need to create a new RAID 0 LUN from data volumes and add to a storagegroup before extending the rootfs.

**Caution:** Please be aware that in NAS versions before 5.2.19.0 or 5.3.13.0, an extension of the rootfs >than 48Gb will prevent Usermapper from being able to calculate new uid/gids.

## **NAS 5.3 ROOTFS INODE EVENTS:**

With NAS 5.3, the default threshold for logging CallHome events for inode usage on a rootfs is 90%. Parameter can be adjusted with following param:

```
param file fsInodeThreshold=
```

## **BLOCK SIZE HASH PARAMETER:**

Occasions where Celerra might have a problem traversing the indirect block hash list. The default indirect block hash size is 2, and if there are too many hits on the same hash list, the CPU could eat up all the CPU time trying to traverse the list. Increase default to 64:

```
param ufs indBlkHashSize=64
```

**Note:** Adjust to higher value when indirect block queue is very large. Problem is generally associated with very large files, as large files require many indirect blocks and the hash queues used to locate all of a file's indirect blocks can be consumed. NAS 5.3 changes the default hash size to 64.

## **FILE SYSTEM SYNC BEHAVIOR:**

Default sync thread behavior is to flush and cleanup file system modifications from memory every 30,000 msec. [30 secs]

With pre-510 data movers, could run out of memory for intensive fs operations, resulting in panic. Can increase frequency of file system sync behavior:

```
param ufs syncInterval=30000 [hex=7530]
```

**Note:** Default ufslog flush interval is that thread wakes up to flush every log changes every 30 seconds [30,000 msec]. In some cases, lowering the flush interval to 10,000 [10 secs] may help prevent panics on busy systems.

```
# .server_config server_2 -v "param ufs logHoldListSz" [use to reduce max pending transactions]
ufs.logHoldListSz INT 0x01b95558 1000 1000 (0,4294967295) FALSE REBOOT 'NA'
```

## **VALUES FOR UFSLOG & PERFORMANCE ISSUES—FEB 2007:**

**param ufs logHoldListSz=128 | 30 → 30** is the new default for NAS 5.6

```
param ufs syncInterval=15000
```

**Note:** See emc149588 for more information. With NAS 5.6, the new standard is to set the logHoldListSz to 30, thus allowing for more frequent ufslog flushes. Waiting for the queue to build to 128 has been found to cause too many forced flushes, which can cause performance issues.

## **WHY DO WE USE THE " T2RESET" COMMAND V. THE "SERVER CPU" COMMAND?**

1. Use the “t2reset” whenever you’ve edited the “boot.cfg” file, as in preparing to run an FSCK. The Server will reboot from the /nas/dos/slot\_x/boot.cfg and NOT rebuild the “boot.cfg” from scratch as it normally does on reboot.

2. Similarly, use the “server\_cpu” reboot command when you do want to rebuild the server’s configuration files [aka, “boot.cfg”]. The Server rebuilds the “boot.cfg” from the files located in the /nas/server/slot\_x.

## **CONDUCTING FSCK'S ON CELERRA SERVERS**

### **INTRODUCTION TO CELERRA FSCK:**

#### **Intro:**

FSCK's should be carefully planned, and in most cases, conducted using the Standby Data Mover--though depending on the situation, the Primary Server itself may be required. The basic methodology for using the Standby Server (or Production Server) is to temporarily or permanently unmount the production file system from the Server and then associate the Server with the file system to be FSCK'ed. This method is used when a single file system is involved and when the file system check is being conducted in a 'scheduled' manner.

**Note:** Other methods are available for recovering Servers from Rolling Panics and situations where the data mover will not boot, and where the file system(s) requiring FSCK are unknown—see other sections.

#### **EXAMPLE:**

```
# server_umount server_2 -t fs01 /fs01 →Production Server (temp unmount)
# server_mount server_3 -v fs01      →Standby Server (virtual mount puts metavolume info into Standby DART memory)
```

### **FSCK TRIVIA:**

#### **FAST FSCK & FIX INODE FUNCTION:**

With NAS 4.2.7.0 & higher, Celerra introduced multi-threaded 'fastfsck' using 32 Threads vs. single-thread. FSCK's now run extremely fast compared to previous codes. Beginning with 2.2.54.0, 4.0.18.0, 4.2.3.0, & 4.1.5.0, "fixinode" function was built into fsck. More recently, AR's 65330 & 57764 have been built into NAS 5.5.7.0+ to make ACLCHK multi-threaded instead of single-threaded, and also to improve Phase I of FSCK to redistribute workload among threads so that the process becomes faster. Also note that with NAS 5.4.18.0+, if DART detects a corrupted ACL database, the file system will be unmounted and marked for auto-fsck.

**Server Log Symptoms indicating need for FSCK:** Vnon lookup problem bad ino; Read Block failure referencing FS;

\$ls -laR [Running this list command against a FS may reveal problems]

**Locating Bad Inode:** \$find . -inum 1222877 -print

**Panic Asserts:** Asserts are Error Traps placed into the Code when certain thresholds or illegal conditions occur

### **MINIMUM NAS CODE VERSIONS TO USE FOR ALL FSCK's:** See AR45584, 56565, 72001, 75889

**5.1.26.0**  
**5.2.22.1**  
**5.3.25.2**  
**5.4.26.4**  
**5.5.24.2**

**Note:** The fsck information here can become dated--please make sure to use the latest approved nas.exe for the code family in which the FSCK is required.

### **FSCK REQUIREMENTS:**

1. Verify BackEnd & Fabric integrity before conducting FSCKs, especially after events which precipitate the need for an FSCK
2. For CLARiiON backends, make sure that there are no Uncorrectable Sector issues with LUNs prior to performing FSCK after IO Failure despite all retries panics
3. Verify that LUNs are not trespassed for Clariion Systems
4. Verify dual-path availability to BackEnd from the Server used for the FSCK
5. Verify that Write & Read cache are enabled on Clarion Systems
6. Verify most recent 5.1 or 5.2 nas.exe is used for FSCK due to recent issues with Alternate Data Streams/Named Streams corruption
7. Verify contents of lost+found directory prior to, and after each FSCK

**Note:** When appropriate, move the contents out of the lost+found directory to a salvage directory (lost+found limited to 18 blocks). Or, more easily, simply rename the lost+found directory to salvage1, repeat fsck, rename lost+found to salvage2, etc.—the fsck process will recreate the lost+found directory automatically

#### **EXAMPLE:**

```
# mkdir salvage →Create temporary directory at root of file system being FSCK'ed
# cd lost+found
# mv \#* ..\salvage &
```

8. Preferred FSCK method is to use the Standby Server to run FSCK against production file systems, while keeping Production Server up for all other undamaged file systems (there are new exceptions to this general practice, as seen below)
9. All file systems require ACLCHK procedure after the FSCK process is completed. ACLCHKS are mutually inclusive to FSCKs and are not run as a stand-alone operation.
10. All file systems except NFS-only systems, & MPD DIR3 file systems, require CIFS Update to rebuild Shadow files at each directory level of the file system. CIFS Updates are manually initiated.

11. Set FSTOOLS debug logging on the Server running the FSCK so as to observe and capture FSCK info from the Server Log

**Note:** Do NOT set FSTOOLS debug levels on NAS Versions 5.4 or 5.5 as it will pollute Server Log with excess information!

12. FSCK Rule of Thumb is to run FSCK subsequent times until the Server Log shows a complete and clean pass through all Phases

**Note:** (3) passes are not uncommon, sometimes more are required, sometimes less

12. Verify correct File System “volume” number before running .server\_config command—it is NOT the file system name or id number

#### **EXAMPLE:**

# nas\_fs -l

| id | inuse | type | acl | volume | name   | server |
|----|-------|------|-----|--------|--------|--------|
| 31 | y     | 5    | 0   | 171    | fstest | 2      |

**Note:** Look for “volume” column and use that number for the “file fsck uufs 171” command—remember to add +1 to Server Number seen to derive correct Server where file system is mounted—see comment below regarding changes here in NAS 5.4

### **SPECIAL FSCK RULES AND CONDITIONS REGARDING STANDBY SERVER:**

1. Do not use the Standby Server method for performing FSCK’s on file systems with IP Replication sessions established

#### **Performing FSCK on File System with Replication:**

a.) Suspend the replication session

**Note:** If using .server\_config to FSCK, unmount file system first, remount using \$ server\_mount -v to Production Server

b.) Use nas\_fsck from the Production Server, or .server\_config method

c.) Restart the replication session after fsck is completed

**Note:** Starting with NAS versions 5.4.27.x & 5.5.24.0, the server\_mount -v command will not allow a file system with replication running, to be mounted to any Server except the production Server—AR78568.

2. Starting with Napa 3 [NAS 5.5.23.2], if using the Rvector Tool to repair a file system that panics with an Uncorrectable Sector issue, the FSCK must be kicked off using the # nas\_fsck command on the Production Data Mover.

#### **Celerra FSCK checks.**

It is useful to know that an FSCK can only correct File & Directory references or links [unreferenced files; incorrect link count; bad directory entry; missing bitmap blocks; truncated inodes], NOT data. FSCK’s always update the superblock information and write a value to the block if FSCK is interrupted. Also, in rare situations, such as those involving the ADS issue, FSCK’s can actually cause file damage, especially to CAD/CAM, jpegs, and bitmap files.

#### **How long should an FSCK take?**

The time required to run an FSCK on newer multi-threaded NAS codes is generally much faster than in versions prior to 4.2.7.0.

Though FSCK times are always dependent on factors, such as, file system size, file system directory layout and depth, number of files in directories, and overall damage to the metadata of the file system itself, a good rule of thumb is that a 300-400GB file system should be able to complete the first FSCK in about an hour, give or take a half-hour. However, whenever considering the total time required to conduct file system maintenance, be aware that an FSCK may need to run up to (3) times in order to completely repair the metadata, and that an ACLCHK is also required, as well as CIFS Updates on pre-MPD file systems.

#### **What does an FSCK check for and do for repairs?**

--Checks integrity of Superblock and rebuilds if necessary from alternate superblock

--Checks and fixes link count for each inode in file system [Link count is number of references to inode, checked by counting the number of times the inode was referenced in the directory blocks. A reference count of zero will cause a valid inode to be moved to lost+found]

--Checks that block numbers in direct & indirect fields are unique to each inode [if two inodes share data block, both are removed]

--Verifies that all directories are connected to root

--Verifies that each directory entry has an inode [invalid inodes are cleared from directory entries]

--Verifies that “.” entry points to itself, and “..” entries point to parent of the directory

--Checks and fixes Cylinder Group summaries

### **FSCK PHASES:** FSCK contains three basic Phases: Validate Inodes, Validate Directories, & Summary Check

#### **FSCK PHASE 1: Validate Inodes**

Superblock checked. Inode & directory bitmaps are created, as are inode link count map [consisting of inodes with link count > than 1 and directories > than two]. For each unallocated inode, checks to ensure that data blocks have zero pointers. For each allocated inode, checks direct and indirect blocks and marks them as used in the used block bitmap. If multiple streams are found, a container bitmap is created, which is then link count & reference checked in a special phase 1M.

#### **SERVER LOG EXAMPLE:**

2004-08-15 16:23:20: FSTOOLS: 4: Phase 1: Validate Inodes

2004-08-15 16:23:21: FSTOOLS: 4: 0 % complete

2004-08-15 16:24:29: FSTOOLS: 3: INCORRECT BLOCK COUNT I=4205067 (304 should be 288)1092605069: FSTOOLS: 3: (CORRECTED)

2004-08-15 16:32:57: FSTOOLS: 3: INCORRECT BLOCK COUNT I=2097192 (304 should be 288)1092605577: FSTOOLS: 3:

(CORRECTED)

2004-08-15 16:33:13: FSTOOLS: 4: 50 % complete

2004-08-15 16:34:02: FSTOOLS: 4: 60 % complete

2004-08-15 16:34:04: FSTOOLS: 3: INCORRECT BLOCK COUNT I=6237193 (304 should be 288)1092605644: FSTOOLS: 3:  
(CORRECTED)

2004-08-15 16:35:11: FSTOOLS: 4: 70 % complete

2004-08-15 16:36:10: FSTOOLS: 4: 80 % complete

2004-08-15 16:37:07: FSTOOLS: 4: 90 % complete

2004-08-15 16:38:08: FSTOOLS: 4: 100 % complete

2004-08-15 16:38:08: FSTOOLS: 5: Inode Phase: took 888361490 Usec

### **FSCK PHASE 2: Validate Directories**

Checks all directories starting from root, for traversal, link count, and connectivity. Parent map can fix incorrect “..” entries.

Creates reference map for inodes in directory blocks, which should match with inode map created in Phase 1. If there are unreferenced inodes, they are moved to lost+found. Directory entries are added to traversal list and scanned. Connectivity of all directories checked from root-if disconnected are placed in lost+found.

### **SERVER LOG EXAMPLE:**

2004-08-15 16:38:08: FSTOOLS: 4: Phase 2: Validate Directories

2004-08-15 16:38:21: FSTOOLS: 4: 0 % complete

2004-08-15 16:38:28: FSTOOLS: 4: 10 % complete

2004-08-15 16:38:52: FSTOOLS: 4: 20 % complete

2004-08-15 16:39:23: FSTOOLS: 4: 30 % complete

2004-08-15 16:39:54: FSTOOLS: 4: 40 % complete

2004-08-15 16:40:32: FSTOOLS: 4: 50 % complete

2004-08-15 16:40:55: FSTOOLS: 4: 60 % complete

2004-08-15 16:41:39: FSTOOLS: 4: 70 % complete

2004-08-15 16:42:22: FSTOOLS: 4: 80 % complete

2004-08-15 16:42:50: FSTOOLS: 4: 90 % complete

2004-08-15 16:42:50: FSTOOLS: 4: 100 % complete

2004-08-15 16:42:50: FSTOOLS: 4: Invalid Link Count:I:495670L:5R:4

2004-08-15 16:42:50: FSTOOLS: 4: Adjust

2004-08-15 16:42:50: FSTOOLS: 6: Changing Reference Count of INODE 495670

2004-08-15 16:42:50: FSTOOLS: 4: Invalid Link Count:I:518591L:1R:0

2004-08-15 16:42:50: FSTOOLS: 4: Adjust

2004-08-15 16:42:51: FSTOOLS: 6: Changing Reference Count of INODE 518594

2004-08-15 16:42:52: FSTOOLS: 5: Directory Phase: took 284464733 Usec

### **FSCK PHASE 3: Summary Check**

Checks Cylinder Group summaries and rewrites if necessary.

### **SERVER LOG EXAMPLE:**

2004-08-15 16:42:52: FSTOOLS: 4: Phase 3: Summary Check

2004-08-15 16:42:54: FSTOOLS: 3: BLK(S) MISSING IN BIT MAPS CGNO:1787

2004-08-15 16:42:54: FSTOOLS: 3: BLK(S) MISSING IN BIT MAPS CGNO:3196

2004-08-15 16:43:02: FSTOOLS: 3: BLK(S) MISSING IN BIT MAPS CGNO:499

2004-08-15 16:43:02: FSTOOLS: 3: BLK(S) MISSING IN BIT MAPS CGNO:1684

2004-08-15 16:43:05: FSTOOLS: 3: BLK(S) MISSING IN BIT MAPS CGNO:2108

2004-08-15 16:43:06: FSTOOLS: 5: Summary Check: took 14232475 Usec

2004-08-15 16:43:06: FSTOOLS: 4: fs\_npreserved 0

2004-08-15 16:43:06: FSTOOLS: 3: FOLLOWING COUNT(S) WRONG IN SUPERBLK1092606186: FSTOOLS: 4: cs\_nbfree 8368928, fs->cs\_nbfree 8368917

2004-08-15 16:43:06: FSTOOLS: 4: cs\_nifree 39435667, fs->cs\_nifree 39435668

2004-08-15 16:43:06: FSTOOLS: 4: 4757740 files, Blocks: 36913064 in use, 8368928 free

2004-08-15 16:43:06: FSTOOLS: 6: fsck: 722088 reads, 354590 writes

2004-08-15 16:43:06: FSTOOLS: 4: fsck completed

## **USING PRODUCTION OR STANDBY SERVER TO FSCK FILE SYSTEMS:**

**Note:** Do not use Standby Server to run an FSCK on a production file system that has a replication session established, or a file system being repaired using the Rvector Uncorrectable Sector method—see comments outlined above.

### **I. RUNNING FSCK:**

#### **NEW WAY FOR DOING MOST FSCKs—NAS 5.4 & 5.5:**

Step 1. Kick off FSCK from Production Server:

```
# nas_fsck -start milla -mover server_2
id          = 75
name        = milla
volume      = v139
fsck_server = server_2
status      = In-progress
```

**Note:** FSCK automatically unmounts & remounts the file system. NAS 5.4 & 5.5 no longer requires setting FSTOOLS debug logging levels. ACLCHK is also run as part of the FSCK process.

Step 2. Monitor FSCK progress in Server Log:

```
2006-11-14 13:43:22: FSTOOLS: 4: 0: FsId: 75 Fsck Started (Manual).
2006-11-14 13:43:33: FSTOOLS: 4: 3: FsId: 75 Aclchk Started (Auto).
2006-11-14 13:43:34: FSTOOLS: 4: fsck completed.
2006-11-14 13:43:34: FSTOOLS: 4: 1: FsId: 75 Fsck Succeeded
```

Step 3. If the log shows FSCK corrections, run again until an FSCK pass runs cleanly

#### **TRADITIONAL SUPPORT METHOD AND FOR OLDER CODE VERSIONS:**

Step 1. Temporarily unmount Production FS from Server:

```
$ server_umount server_2 -temp m3 /m3
server_2 : done
```

Step 2. Virtually mount production file system to Standby Server:

```
$server_mount server_3 -v m3
```

**Caution:** Using this command does not actually 'mount' the filesystem, but does load the data structure into the memory of the Data Mover. If the site requires a large number of file system fsck's, an option would be to break the Standby relationship and convert the Standby to a Primary, and then do the FSCK as you would do normally. Since the data mover loads the file system structures into memory, these structures must be removed after performing an fsck by either rebooting the Server or running the umount command:

```
$server_umount server_3 -v m3
```

Step 3. Set FSTOOLS Logging on Standby Server:

```
$server_config server_3 "logsys set severity FSTOOLS=LOG_DBG3"
```

**Note:** This debug setting is no longer required with NAS Versions NAS 5.4 or 5.5 and may overrun the Server Log with superfluous messages.

Step 4. Identify correct File System and Volume ID:

```
# nas_fs -l
id  inuse type acl volume  name      server
30    y     1 0   133    m3           1 (fsid=30, volumeid=133)
```

Step 5. Start FSCK using Volume Number of File System—derived from #nas\_fs -l output:

```
$server_config server_3 "file fsck ufs 133=30"
```

*server\_3 : commands processed: 0*

**Error: RPC: Timed out**

**Note:** When running the FSCK command, you will see an "RPC Error" timeout message—the FSCK is running—do not issue command again or another FSCK will be queued. Please note that NAS versions 5.3.20, 5.4, and 5.5 use a newer syntax than previous versions of NAS.

#### **EXAMPLE FSCK SYNTAX FOR PRE-NAS 5.3.20 VERSIONS:**

```
$ .server_config server_3 "file fsck ufs 133"
```

Step 6. Monitor FSCK Progress:

```
$server_log server_3 -s -f |grep -i Phase [or Summary or FSTOOLS]
```

Step 7. When all FSCK passes are completed, remount File System to Production Server:

```
$server_mount server_2 m3 /m3
```

**Note:** Follow-through with ACLCHKs and CIFS Updates, as appropriate (no CIFS Update required on MPD file systems)

Step 8. Cleanup any leftover data structures on the data mover used to conduct the FSCK by running:

```
$server_umount server_3 -v m3
```

#### **NAS 5.4 ADDENDUM:**

##### **There are a number of significant changes with FSCK:**

→ You can run either the nas\_fsck or the .server\_config fsck commands when repairing file systems

→ Server may initiate an autofsck if corruption is detected—system will panic and run autofsck

→ Up to two concurrent nas\_fsck's can be run per DM

**param ufs skipFsck=1**

**Note:** Default NAS 5.4 & 5.5 value is 0, which is to perform autofsck on corrupted file system during reboots. Set this value to 1 to disable auto-fsck of corrupted file systems on reboot—instead, will leave file system unmounted to allow Data Mover to mount all other file systems during reboot, etc. Default value for NAS 5.6 was changed to 1, meaning that file systems will not be auto-fscked during bootup.

**# server\_param server\_2 -facility ufs -info skipFsck -v**

server\_2 :

```

name      = skipFsck
facility_name = ufs
default_value = 1 →NAS 5.6 default
current_value = 1
configured_value =
user_action = none
change_effective = immediate
range     = (0,1)
description = If enabled, fsck will not be run on a corrupted FS while booting up and FS will be left unmounted.

```

**Note:** User must manually run FSCK and then remount file system. Important to note that this is the default behavior if corrupted file system found on bootup.

**param ufs skipAclChk=1** [Use this option to run fsck without running the ACLCHK]

**Note:** Change can be loaded directly into memory prior to FSCK: **\$ .server\_config id=4 "param ufs skipAclChk=1"**

**SERVER\_CONFIG COMMAND LINE SYNTAX CHANGE:**

**\$ .server\_config server\_2 "file fsck ufs 203=22"** [Where 203=Vol ID and 22=FSID]

**Note:** NAS 5.4 or 5.5 will run an integrated ACLCHK during the FSCK process, whether using .server\_config or #nas\_fsck

**USING NAS FSCK TOOL:****# nas\_fsck -start id=93 -mover server\_2****# nas\_fsck -start fs30 -mover server\_5**

```

2005-11-28 15:05:24: FSTOOLS: 1: 0: FsId: 1282 Fsck Started (Manual).
2005-11-28 15:05:24: CFS: 4: fs 0x502 being unmounted. Waiting for quiesce ...
2005-11-28 15:05:24: CFS: 4: fs 0x502 unmounted
2005-11-28 15:05:24: FSTOOLS: 5: Starting page count is 0
2005-11-28 15:05:24: FSTOOLS: 7: Allocating page 0x6440c9 0x64c2c0 count is 0
2005-11-28 15:05:24: FSTOOLS: 7: Allocating page 0x6440db 0x64c2c0 count is 1
2005-11-28 15:05:24: FSTOOLS: 5: Doing consistency checks on superblock
2005-11-28 15:05:24: FSTOOLS: 5: PRIMARY SUPERBLOCK VALIDITY CHECK
2005-11-28 15:05:24: FSTOOLS: 5: SUCCESS
2005-11-28 15:05:24: FSTOOLS: 5: ALTERNATE SUPERBLOCK VALIDITY CHECK
2005-11-28 15:05:24: FSTOOLS: 5: SUCCESS
-----output abbreviated-----
2005-11-28 15:05:24: FSTOOLS: 4: ** Last Mounted on /fs30
2005-11-28 15:05:24: FSTOOLS: 4: Phase 1: Validate Inodes
2005-11-28 15:05:24: FSTOOLS: 4: Phase 2: Validate Directories
2005-11-28 15:05:24: FSTOOLS: 1: 3: FsId: 1282 Aclchk Started (Auto).
2005-11-28 15:05:24: FSTOOLS: 4: Aclchk fsid 1282: 100% complete
2005-11-28 15:05:24: FSTOOLS: 4: Aclchk fsid 1282 - phase2: Validate free acls
2005-11-28 15:05:24: FSTOOLS: 7: Allocating page 0x64e177 0x65047e count is 27
2005-11-28 15:05:24: FSTOOLS: 7: Allocating page 0x64e177 0x6504a6 count is 28
2005-11-28 15:05:24: FSTOOLS: 5: checkFreeAcls fsid 1282: Next free is 10
2005-11-28 15:05:24: FSTOOLS: 4: Aclchk fsid 1282 - phase2: 100% complete
2005-11-28 15:05:24: FSTOOLS: 1: 4: FsId: 1282 Aclchk Succeeded
2005-11-28 15:05:24: FSTOOLS: 4: Phase 3: Summary Check
2005-11-28 15:05:24: FSTOOLS: 4: fsck completed
2005-11-28 15:05:24: FSTOOLS: 1: 1: FsId: 1282 Fsck Succeeded

```

**# nas\_fsck -start fs8 -monitor**

```

id      = 38
name    = fs8
volume  = v119
fsck_server = server_2
status   = Done
..Done

```

**\$ nas\_fsck -list**

id type state volume name server

23 1 FSCK 134 ufs2 4

27 1 ACLCHK 144 ufs1 1

**\$ nas\_fsck -info ufs2**

name = ufs2

id = 23

volume = v134

fsck\_server = server\_5

inode\_check\_percent = 100

directory\_check\_percent = 100

used\_ACL\_check\_percent = 100

free\_ACL\_check\_status = Done

cylinder\_group\_check\_status = In Progress

**Note:** Use nas\_fsck -list and -info to keep track of fsck sessions and get specific info on fsck progress

**#nas\_fsck -aclchkonly** [Option for running aclcheck only—cannot be run with fs exported]

**#nas\_fsck -fsckonly** [Option for running fsck only]

## **II. RUN ACLCHK ON PRODUCTION SERVER—NOT STANDBY!:**

### **CRITICAL ACLCHK RULES:**

1. NAS 5.4/5.5 performs builtin aclchk and in most cases should not need to be run separately
2. Make a backup copy of the Server's mount & export/export.shares files
3. Always permanently unmount the Production file system and remount to a new, temporary mountpoint for the ACLCHK process to ensure that there is no access to the file system allowed during the ALCHK process
4. If allowed, stop CIFS as well to prevent any CIFS Admins from connecting to the C\$ share [If not an option, advise customer to not allow Admin users to connect to the Celerra Admin Shares].
5. NAS 5.1 and below will log errors during ACLCHK if users are accessing file system. NAS 5.2 and higher will lock the database during the aclchk and block user access—for any files in use, aclchk will wait for user access to be released.

### **BEGIN ACLCHK PROCESS:**

- Step 1. Permanently unmount and remount file system to temporary mountpoint:

**#server\_umount server\_2 -p fs1 /fs1**

**#server\_mount server\_2 fs1 /aclchk**

- Step 2. Start ACLCHK:

**#server\_config server\_2 "file aclchk /aclchk fix"**

**Note:** Similar to the FSCK command, it is not uncommon to see an RPC timeout error after issuing the command. Check Server Log to verify—depending on NAS version, should see an entry similar to the following:

**2003-07-16 21:16:40: ADMIN: 4: Command succeeded: file aclchk /aclchk fix**

**# server\_log server\_4 -a -s |grep CFS**

**2004-09-21 21:06:21: CFS: 4: aclchk: Done; status = 0**

**Note:** Grep for CFS when running NAS 5.1.19 and higher or ADMIN

**Caution: Do not use the -v flag when running aclchk—results will not print to the Server Log**

- Step 3. Open Server Log to Monitor Progress:

**\$server\_log server\_2 -s -f**

**Note:** The ACLCHK succeeded command and completion summary are written to the Server Log

### **EXAMPLE:**

2003-07-16 21:16:40: ADMIN: 4: Command succeeded: file aclchk /aclchk fix

2003-07-16 21:16:40: CFS: 4: aclchk: Done; status = 0

2003-07-16 21:16:40: ADMIN: 4: Command succeeded: file aclchk /aclchk fix

- Step 4. Permanently unmount File System and remount to original Mountpoint.

- Step 5. Re-export all exports:

**\$server\_export server\_2 -a**

## **III. RUNNING CIFS UPDATE:**

### **CIFS UPDATE RULES:**

1. Turn on SMB debug prior to running CIFS Update in order to see logging events: “**logsys set severity SMB=LOG\_DBG3**”
2. File System must be mounted, exported, and the CIFS service stopped in order to be effective when rebuilding the Shadow file.
- Note:** If CIFS cannot be stopped, the CIFS Update may not start, though in emergencies, this method can be tried without harm
3. Turn off group Quotas prior to running CIFS update

### **BEGIN CIFS UPDATE PROCESS:**

Step 1. Stop CIFS on Production Server:

**\$server\_setup server\_2 -P cifs -o stop**

Step 2. Run CIFS Update on mountpoint of File System:

**\$server\_config server\_2 "cifs update /export force level=0"**

**\$server\_config server\_2 "cifs update '/export/folder with spaces' force level=0"** [Syntax when folders have spaces]

**Note:** Verify progress in Server Log—CIFS updates are logged every 600 seconds until completed

**EXAMPLE:**

1003979370: SMB: 4: Update of : /mbs running since 600 s 340639 files, 42093 dirs have already been updated

1003979970: SMB: 4: Update of : /mbs running since 1200 s 780547 files, 94904 dirs have already been updated

1003980448: SMB: 4: Update of : /mbs completed after 1678 s

1003980448: SMB: 4: Update of : 1170072 files, 127689 dirs

Step 3. After Update is completed, Start CIFS:

**\$server\_setup server\_2 -P cifs -o start**

Step 4. Have customer verify CIFS access

## **STOPPING CIFS UPDATES:**

**Note:** Use the following command to stop any of the commands that use CIFS Update

**#server\_config server\_x "cifs update abort"**

## **CLEANING UP "LOST+FOUND" DIRECTORY ON FILESYSTEMS DURING FSCKs:**

**Symptom:** Server\_log shows error during FSCK: "2002-03-08 13:36:33: FSTOOLS: 3: SORRY. NO SPACE IN lost+found DIRECTORY1015612". Lost+found is a directory that is limited to 18 blocks of data.

**Cause:** During FSCK's, inodes that cannot be reconciled are removed and placed in the filesystem's "lost+found" folder. With extensive inconsistencies, this directory can rapidly be filled, preventing the FSCK from correcting any further problems.

**Fix:** Always check and cleanup the lost+found directory before and after each FSCK using following method

1. Create new directory underneath mountpoint of filesystem [#mkdir fsck]

2. Go into lost+found directory, output list of inodes to file, and move contents to temp directory:

**#cd lost+found**

**#du -h >/home/nasadmin/lostfound.fs02**

**#mv \#\* ..//fsck &**

## **LOST+FOUND:**

**# cd mt3/lost+found**

```
-rw-r--r-- 1 1000 203 1519 Mar 1 10:12 #00707548
-rw-r--r-- 1 1000 203 1110 Mar 1 10:12 #00707549
drw-r--r-- 1 1000 203 5342208 Nov 16 15:33 #01209752
```

**How to Access Lost+Found Entries:** #cd \#01209752 or #more \#00707548 to see file contents

## **OTHER FSCK METHODS:**

### **I. .Server Config FSCK Method Using either Production Server or Standby:**

**Note:** Used for Single FS FSCK while leaving DM up and running

Step 1. Verify File System Name, FSID, & Volid:

**# nas\_fs -I**

|    |       |      |     |        |      |        |
|----|-------|------|-----|--------|------|--------|
| id | inuse | type | acl | volume | name | server |
| 40 | y     | 1    | 0   | 150    | fs2  | 1      |

→Notice that FSID=40 and Volume ID=150

Step 2. Temp. unmount file system:

**# .server\_config server\_2 -v "file umount /fs2" or # server\_umount server\_2 -temp fs2 /fs2"**

Step 3. Virtually mount file system on either Production Server or Standby Server using -v:

**# server\_mount server\_3 -v fs2**

server\_3 : done

Step 4. Set debug logging and start FSCK:

**#server\_config server\_3 "logsys set severity FSTOOLS=LOG\_DBG3"**

**Note:** This debug setting is no longer required with current versions of NAS 5.4 or 5.5

**#.server\_config server\_3 "file fsck uxfs 150=40"**

**Note:** Syntax varies with NAS Version. Use above syntax for 5.3 and higher. Older NAS versions may use "file fsck uxfs 150".

Step 5. Verify completion of FSCK in Server Log and remount fs when completed:

**# server\_mount server\_2 fs2 /fs2**

Step 6. If virtually mounted to Server\_3, make sure to unmount it as well:

**# server\_umount server\_3 -v fs2**

## **II. BOOT.CFG FSCK METHOD FOR ROOTFS or PFS:**

**Note:** Used to FSCK Root or Production File Systems during t2reset. Example below shows FSCK of root and a production fs.

Step 1. Make backup copy of boot.cfg file:

**#cp /nas/dos/slot\_2/boot.cfg /nas/dos/slot\_2/boot\_ori**

Step 2. Identify FSID and Volume ID's of File Systems to be FSCK'ed:

**# nas\_fs -l**

| id | inuse | type | acl | volume | name      | server |
|----|-------|------|-----|--------|-----------|--------|
| 2  | y     | 1    | 0   | 12     | root_fs_2 | 1      |
| 38 | y     | 1    | 0   | 146    | fs1       | 1      |

Step 3. Edit boot.cfg and place following entries between the “mount begin” line and “file mount” lines:  
mount begin

**logsys set severity FSTOOLS=LOG\_DBG3**

**file fsck ufs 12=2**

**file fsck ufs 146=38**

file mount ufs rw / 12=2 rw

file mount ufs rw /fs1 146=38 rw

**Note:** Be aware of syntax change for performing fsck. Above entries are correct for NAS 5.3 and higher. NAS 5.2 or below requires use of “file fsck ufs 12” as the correct line syntax.

Step 4. Use t2reset to reboot DM without rebuilding the boot.cfg file from /nas/server/slot:

**#/nas/sbin/t2reset reboot -s 2**

**Note:** “server\_cpu” would rebuild the “boot.cfg” file from the /nas/server/slot files—we want to use the edited boot.cfg file

Step 5. Monitor FSCK progress in Server Log:

**#/nas/sbin/log\_slot -s -f 2**

2006-05-30 13:25:10: FSTOOLS: 4: 1: FsId: 38 Fsck Succeeded

2006-05-30 13:25:10: ADMIN: 4: Command succeeded: file fsck ufs 146=38

Step 6. Restore Boot File:

**#cp /nas/dos/slot\_2/boot\_ori /nas/dos/slot\_2/boot.cfg**

Step 7. Reboot DataMover:

**#/nas/bin/server\_cpu server\_2 -r -m now**

**Note:** Use this method to FSCK root filesystem of DM and to restore server to operation while selectively choosing which file systems to mount and which to keep unmounted. Use # to comment out any production file systems that you do not want mounted.

## **III. MOUNTING ALL FS EXCEPT FOR ONE BEING FSCK'ed:**

**Note:** Run FSCK on single fs while bringing all others back online—comment out mountpoint of file system in boot.cfg and run fsck just prior to eof line in boot.cfg file.

Step 1. Make backup copy of boot.cfg file:

**#cp /nas/dos/slot\_2/boot.cfg /nas/dos/slot\_2/boot\_ori**

Step 2. Identify FSID and Volume ID's of File Systems to be FSCK'ed:

**# nas\_fs -l**

| id | inuse | type | acl | volume | name      | server |
|----|-------|------|-----|--------|-----------|--------|
| 2  | y     | 1    | 0   | 12     | root_fs_2 | 1      |
| 38 | y     | 1    | 0   | 146    | fs1       | 1      |

Step 3. Comment out the file system to be fsck'ed & add fsck near end of file:

mount begin

file mount ufs rw / 12=2 rw

**#file mount ufs rw /fs1 146=38 rw**

---

logsys delete output console

**logsys set severity FSTOOLS=LOG\_DBG3**

**file fsck ufs 146=38**

**#verify\_eof**

Step 4. Use t2reset to reboot DM without rebuilding the boot.cfg file from /nas/server/slot:

**#/nas/sbin/t2reset reboot -s 2**

**# server\_mount server\_2**

server\_2 :

root\_fs\_2 on / ufs,perm,rw

fs1 on /fs1 ufs,perm,rw,<unmounted>

Step 5. Monitor FSCK progress in Server Log:

**#/nas/sbin/log\_slot -s -f 2**

2006-05-30 14:11:05: FSTOOLS: 4: fsck completed

2006-05-30 14:11:05: FSTOOLS: 4: 1: FsId: 38 Fsck Succeeded

2006-05-30 14:11:05: ADMIN: 4: Command succeeded: file fsck ufs 146=38

Step 6. Restore Boot File:

**#cp /nas/dos/slot\_2/boot\_ori /nas/dos/slot\_2/boot.cfg**

Step 7. Remount fs when ready, or reboot DataMover:

**# server\_mount server\_2 -a**

#### **IV. FSCK on Alternate Superblock:**

**#/nas/sbin/rootnas\_fsck fs2 -b 32**

**Note:** Command run recently at 5.3.21 site for following problem that was not being cleaned by regular fsck:

2005-11-02 23:51:47: FSTOOLS: 3: PERSISTENT RESERVATION BLK COUNT WRONG IN SUPERBLK

2005-11-02 23:51:47: FSTOOLS: 4: PBR count should be 14919, but it is 404 Correcting PBR count

### **CELERRA ACL SUPPORT:**

#### **Running ACLCHK Using “.server config” Command, followed by CIFS Update:**

Step 1. **#server\_umount server\_6 -p fs1 /export/htdocs** [Original production mountpoint]

**#server\_mount server\_6 fs1 /aclchk** [New mountpoint from which to run the ACL CHECK]

Permanently unmount PFS from original mountpoint. Create new mountpoint for aclchk purposes and permanently remount filesystem to this mountpoint. Conduct aclchk.

**Note:** Filesystem being aclchk'ed is NOT exported at all!

Step 2. Conduct ACL Check on Server\_6: **\$server\_config server\_6 "file aclchk /export/htdocs fix"**

Step 3. Open Server Log to Monitor: **\$server\_log server\_6 -s -f** [Will show ACL Check completion]

Step 4. Permanently unmount File System and remount to original Mountpoint.

Step 5. Re-export all exports: **\$server\_export server\_6 -a**

Step 6. Stop CIFS Service: **\$server\_setup server\_6 -P cifs -o stop**

Step 7. Run CIFS Update on FS: **\$server\_config server\_6 "cifs update /mntpoint force level=0"**

Step 8. Restart CIFS & Verify Customer Access

### **ACL DATABASE CORRUPTION AND OVERFLOW ISSUES (aclid overflow):**

#### **1. ACLEDUMP shows error:**

**\$ .server\_config server\_2 -v "acl dump=/ug\_fs\_001/sc52ld-BAK"**

1131472328: SMB: 3: Dump with no ThreadCtx

Dump of rights of /ug\_fs\_001/sc52ld-BAK

===== UNIX =====

USER 0x66c6 GROUP 0x67f7 mode=rwxrwxrwx

===== NT =====

aclId=606693

1131472328: UFS: 3: getAcl: Bad gen number 0x1, on disk gen is 0x3e for slot 0x941e5

Unable to load SD (28=InternalError)1131472328: ADMIN: 4: Command succeeded: acl dump=/ug\_fs\_001/sc52ld-BAK

#### **2. Server Log errors:**

# server\_log server\_2 -sltail

2005-11-08 11:58:06: UFS: 3: getAcl: Bad gen number 0x1, on disk gen is 0x44 for slot 0x8625

2005-11-08 11:58:06: UFS: 3: getAcl: Bad gen number 0x1, on disk gen is 0x3e for slot 0x8626

2005-10-17 10:57:42: CFS: 3: Setsd: setAcl failed 3

2005-10-17 10:57:42: CFS: 3: File\_NamingNode::synchronizeAcl Cannot store SD status : 3

2005-10-17 10:57:42: CFS: 3: File\_NamingNode::setAttr : cannot synchronize ACL status 3

2005-10-17 10:57:48: UFS: 3: Filesystem 35 ACL slotIndex goes beyond 1M

#### **3. Printstats ACLDB command shows error:**

**\$server\_config server\_2 -v "acl database=/fs23 printstats"**

"ACL Database of 131 fs locked in shared

1129660119 ACLUPD:3:parseAclDatabase: invalid database nextFree

AclDatabase of 131 fs unlocked

1129660119: ACLUPD:3: Run file aclchk <fs> to check and fix the acl database

End parse: RefCounter=20"

**Note:** For large databases, it may be better to run command without –v so as to output to Server Log.

## **EXAMPLE ACLDB CHECKANDCOMPACT COMMAND:**

**Note:** NAS 5.4 contains potential ACL database ID overflow problem when using the MIXED or MIXED\_COMPAT access policy. 5.4.17.504 & 5.4.18.300 have been released as a fix to prevent the acldb overflow issue from occurring on file systems, and also to compact the ACL database on existing file systems where it has not yet reached its maximum value of one million records. The acldb overflow issue occurs because MIXED or MIXED\_COMPAT policies do not make proper use of ACL tables, allowing them to quickly fill to their max. value of 1 million records. If a fs reaches 1 million records and becomes corrupted, then it's likely that all ACLs will be lost and a procedure to rebuild the ACLDB from scratch will need to be performed. See emc117671 & emc131087. The current "checkandcompact" tool has not been terribly successful and is being revamped. In addition to the previously mentioned MIXED & MIXED\_COMPAT issue, the ACLDB Overflow issue can also occur on file systems that are not using the aforementioned accesspolicies. Changes are in development to improve ACL Record availability by caching the on-disk ACL records into memory, as well as to retain the older ACL Cache mechanism. Tools are also being developed to assist in identifying and removing duplicate ACL Records from the ACL Database. See AR68014 & 76015.

## **ACLDDB CACHE MISS PROBLEM/ACLDDB GROWS:**

Celerra stores each file system's ACL database in the root of the file system (/slashetc/ACLdata | ACLrecord). The inode of each file or directory contains a link to an entry in the ACL database for its Windows ACLs. Ideally, each ACE, such as Everyone Full Control, is only referenced once in the ACLdb, and all files or directories that have the same ACL setting would reference the same ACLdb entry in its inode. However, due to required efficiencies, especially for new file writes, the Celerra keeps an LRU cache of the most recently accessed acls for files/directories in memory. When files are being written, the data mover searches cache for a match to the ACL—if a match is not made, then a new ACLdb entry is created. If a match is made, then the file is written with the inode link referencing the entry for the ACL in the ACLdb. For file reads, ACLs are read into cache so that if changes are made and a write is made, the ACL is properly linked to its original ACLdb entries. However, the cache is finite and if files are recycled very quickly, the ACL Cache will become 'flushed' and not contain the necessary record, thereby causing a new ACL record to be created. Over time this can cause the ACLdb to grow and eventually reach the 1 million record mark, at which point the file system ACLs are corrupted and need to be rebuilt, resulting in loss of all customer ACLs.

## **REASONS FOR ACL CACHE MISSES:**

--Heavy load on file systems where vnodes are quickly recycled, could lead to cache miss

--file system freeze and thaws can result in cache flush, then when write occurs (rely on the Read to put ACL record into cache), cache miss

--certain applications that Write without doing Reads first, can also cause cache miss

## **NAPA MAINTENANCE RELEASE STRATEGY FOR FIXING ACLDB ISSUES:**

The original ACL cache mechanism was built upon the VFS\_FileSystem object, tied to the Vnode/NamingNode cache in the VFS layer. A new ACL cache is being designed at the UFS layer that will be initialized upon file system mount and the ACL db cached. When a file write is being done, the setAcl() call will search the cache and only allocate a new aclid (ACL Record) if there is a cache miss. A built-in tool will compress the acl database to eliminate any existing duplicate entries. If the new ACL cache cannot be initialized (pre-510 Hardware, param disables the cache), then the VFS layer cache will be used.

**param ufs skipAclCache=0** →Default value 0 means that new UFS cache is enabled, and caches ACLs per maxAclCached  
**param ufs maxAclCached=** →Max cached entries per Blade (value between 2-8Million, with 4M the default)  
**param ufs defAclHashSz=256** →Cache table size for new aclCache—hidden param

## **NEW ACL CACHE, PARAMETERS, & CLEANUP TOOL WITH NAPA I (5.5.21.4):**

**Note:** See emc117671 for more details about ACL db overflow issues. It is possible in some cases to receive the following ACL db error, yet upgrade the code to the latest NAS 5.5 version, and run the nas\_fsck –aclchkonly option to clean up the ACL db without resorting to a complete rebuild of the ACL database.

**2007-03-16 13:33:01: UFS: 3: Filesystem 143 ACL slotIndex goes beyond 1M**

**# /nas/bin/nas\_fsck -start fs06 -aclchkonly**

ACLCHK: in progress for file system fs06

### **Server Log Entries:**

2007-03-27 10:39:27: FSTOOLS: 4: 3: FsId: 26 Aclchk Started (Manual).

2007-03-27 10:39:27: FSTOOLS: 4: Aclchk: Database of fsid 26 - volume id 100 is locked in exclusive

2007-03-27 10:39:27: CFS: 4: Resuming fs 26

2007-03-27 10:39:27: FSTOOLS: 4: Aclchk: fsid 26 - volume id 100 cleanup is done and aclchk is started

2007-03-27 10:39:27: UFS: 4: Aclchk fsid 26 - phase1: in-processing: 10% complete

2007-03-27 10:39:27: UFS: 4: Aclchk fsid 26 - phase1: in-processing: 20% complete

2007-03-27 10:39:27: UFS: 4: Aclchk fsid 26 - phase1: in-processing: 30% complete

2007-03-27 10:39:27: UFS: 4: Aclchk fsid 26 - phase1: in-processing: 40% complete

2007-03-27 10:39:27: UFS: 4: Aclchk fsid 26 - phase1: in-processing: 50% complete  
2007-03-27 10:39:27: UFS: 4: Aclchk fsid 26 - phase1: in-processing: 60% complete  
2007-03-27 10:39:27: UFS: 4: Aclchk fsid 26 - phase1: in-processing: 70% complete  
2007-03-27 10:39:27: UFS: 4: Aclchk fsid 26 - phase1: in-processing: 80% complete  
2007-03-27 10:39:27: UFS: 4: Aclchk fsid 26 - phase1: in-processing: 90% complete  
2007-03-27 10:39:27: UFS: 4: Aclchk fsid 26 - phase1: in-processing: 100% complete  
2007-03-27 10:39:27: UFS: 4: Aclchk fsid 26 - phase2: Validate free acls  
2007-03-27 10:39:27: UFS: 4: Aclchk fsid 26 - phase2: in-processing...  
2007-03-27 10:39:27: UFS: 4: Aclchk fsid 26 - phase2: 100% complete  
2007-03-27 10:39:27: UFS: 4: Aclchk fsid 26: uniqueIds 0, allocatedIds 0, duplication rate 0%  
2007-03-27 10:39:27: UFS: 4: Aclchk fsid 26: ACLrecord caculated sz 8192, current sz 8192  
2007-03-27 10:39:27: UFS: 4: Aclchk fsid 26: ACLdata caculated sz 8192, current sz 0  
2007-03-27 10:39:27: UFS: 4: Aclchk fsid 26 - phase3: Validate ACLrecord and ACLdata fileSize  
2007-03-27 10:39:27: UFS: 4: Filesystem 26 acl file size 8192  
2007-03-27 10:39:27: UFS: 4: Filesystem 26 acl firstfree 1  
2007-03-27 10:39:27: UFS: 4: end acl slot: 1  
2007-03-27 10:39:27: UFS: 4: file system 26 acl initialization done: 0  
2007-03-27 10:39:57: FSTOOLS: 4: 4: FsId: 26 Aclchk Succeeded.  
2007-03-27 10:39:57: CFS: 4: aclchk: Done; status = 0  
2007-03-27 10:39:57: ADMIN: 4: Command succeeded: file aclchk /fs06 fix

**# .server\_config server\_2 -v "aclInfoDump"**

1149013835: UFS: 4: Num of aclchk threads 32  
1149013835: UFS: 4: skipAclCache flag 0  
1149013835: UFS: 4: maxAclCached 4194304  
1149013835: UFS: 4: totalCachedEntries 13

**Note:** Default AclCache size is 4Million entries, or 4,194,304, but can be tuned between 2-8 million hash entries

**# .server\_config server\_2 -v "acl cache=fs2"**

Acl cache statistiques

in:7  
hits:362494  
miss:29  
stale:0

**MODIFYING ACL CACHE SIZE:**

**# .server\_config server\_2 -v "aclModCacheSz aclCacheSize=2000000"**

**NEW ACL CACHING PARAMETERS:** Home Depot Fix AR68014 5.3.23.3040

**\$server\_config server\_2 -v "param ufs skipAclCache=1"**

**Note:** Default value = 0, new cache is initialized. Set to 1 to disable the new cache, remount file system to put into effect.

**# .server\_config server\_2 -v "param ufs skipAclCache"**

ufs.skipAclCache INT 0x01067e94 0 0 (0,1) TRUE NONE 'NA'

**\$server\_config server\_2 -v "param ufs maxAclCached=2097152"**

**Note:** Where size of cache is a value from 2—8 hash records, with default being 4 million unique acl hashes [4194304]—AR76015.

**CLEANING UP ACL DUPLICATES USING -ACLCHKONLY:**

**# nas\_fsck -start fs1 -aclchkonly -mover server\_2**

Error 4423: fs fs1 is exported

**Note:** -aclchkonly can only be run against unexported file systems. Unexport file system or permanently unmount from original mountpoint and remount to temporary mountpoint in order to conduct -aclchkonly.

1. Permanently unmount fs:

**# server\_umount server\_2 -p fs1 /fs1**

server\_2 : done

2. Mount to temporary mountpoint:

**# server\_mount server\_2 fs1 /aclchkonly**

server\_2 : done

3. Run -aclchkonly to remove duplicates:

**# nas\_fsck -start fs1 -aclchkonly -mover server\_2**

ACLCHK: in progress for file system fs1

4. Verify progress in Server Log:

2006-05-30 15:23:32: FSTOOLS: 4: 3: FsId: 38 Aclchk Started (Manual).

2006-05-30 15:23:32: FSTOOLS: 4: Aclchk: Database of fsid 38 - volume id 146 is locked in exclusive

2006-05-30 15:23:32: CFS: 4: Resuming fs 38  
2006-05-30 15:23:32: FSTOOLS: 4: Aclchk: fsid 38 - volume id 146 cleanup is done and aclchk is started  
2006-05-30 15:23:32: UFS: 4: Aclchk fsid 38 - phase1: in-processing: 10% complete

-----output abridged-----

2006-05-30 15:23:32: UFS: 4: Aclchk fsid 38 - phase1: in-processing: 100% complete  
2006-05-30 15:23:32: UFS: 4: Aclchk fsid 38 - phase2: Validate free acls  
2006-05-30 15:23:32: UFS: 4: Aclchk fsid 38 - phase2: in-processing...  
2006-05-30 15:23:32: UFS: 3: Aclchk fsid 38: Invalid refcount for slot 2: inode refcount:18 / on disk acl refcount: 36  
2006-05-30 15:23:32: UFS: 4: Aclchk fsid 38 - phase2: 100% complete  
2006-05-30 15:23:32: UFS: 4: Aclchk fsid 38: uniqueIds 6, allocatedIds 6, duplication rate 0%  
2006-05-30 15:23:32: UFS: 4: Aclchk fsid 38: ACLrecord caculated sz 8192, current sz 8192  
2006-05-30 15:23:32: UFS: 4: Aclchk fsid 38: ACLdata caculated sz 16384, current sz 8192  
2006-05-30 15:23:32: UFS: 4: Aclchk fsid 38 - phase3: Validate ACLrecord and ACLdata fileSize  
2006-05-30 15:23:32: UFS: 4: Filesystem 38 acl file size 8192  
2006-05-30 15:23:32: UFS: 4: Filesystem 38 acl firstfree 5  
2006-05-30 15:23:32: UFS: 4: end acl slot: 8  
2006-05-30 15:23:32: UFS: 4: file system 38 acl initialization done: 6  
2006-05-30 15:24:02: FSTOOLS: 4: 4: FsId: 38 Aclchk Succeeded.  
2006-05-30 15:24:02: CFS: 4: aclchk: Done; status = 0  
2006-05-30 15:24:02: ADMIN: 4: Command succeeded: file aclchk /aclchkonly fix 010

**Note:** This tool is different than the built-in aclchk that is run in conjunction with an fsck and only removes acldb duplicates, which will be logged in Server Log. Running this tool with the latest NAS version can restore a file system to normal operation even if it has hit the 1M ACL index threshold.

#### **VERIFYING ACL HEALTH OF ACL DATABASE RECORDS:**

**\$server\_config server\_2 -v "acl database=/projects1 printstats"**

Acl database map/stats

-----  
Acl Database of 155 fs locked in shared

Acl Database of 155 fs unlocked

Statistiques of Acl database

-----  
Record file size: 0x138000 Bytes

Data file size: 0x4d9a000 Bytes

Total number of slots: 79460

Number of used slots: 79448 →Relates to total number of ACL records in use

Number of free slots: 12

Number of free chunks: 12

Average free chunk size:1 Slots

Number of SD: 79448

Total SD size: 14776968 Bytes

Bigest SD: 226 Bytes

Average SD size: 185 Bytes

Average SD size: 1 Slots

Total wasted space: 66577784 Bytes

Average wasted space: 838 Bytes

Updated SD: 0

Updated inode: 0

Completion time: 0 sec

**\$server\_config server\_2 -v "acl database=/projects1 checkAndCompact"**

**Note:** The checkAndCompact tool does not always work very well and is being replaced with an aclchk enhancement that will traverse file system and cleanup any duplicate ACL records

**\$server\_config server\_2 -v "acl database=/projects1 printstats"** (after compacting acldb)

Acl database map/stats

-----  
Acl Database of 155 fs locked in shared

Acl Database of 155 fs unlocked

Statistiques of Acl database

-----  
Record file size: 0x138000 Bytes

Data file size: 0x4d9a000 Bytes

Total number of slots: 79460

Number of used slots: 67102

Number of free slots: 12358

Number of free chunks: 69

Average free chunk size:179 Slots

Number of SD: 67102

Total SD size: 12480484 Bytes

Bigest SD: 226 Bytes

Average SD size: 185 Bytes

Average SD size: 1 Slots

Total wasted space: 56231964 Bytes

Average wasted space: 838 Bytes

Updated SD: 0

Updated inode: 0

Completion time: 0 sec

## **ACL DATABASE DELETE AND RECREATE PROCEDURE:**

**Note:** In certain cases, the ACL database may become corrupt and require deletion and re-initialization. Do not run this procedure without Eng. approval!

1. Typical indicators of ACL damage:

2002-07-12 16:23:38: UFS: 4: Cleaners ino 62d2f80, cg 62d3050, dirty 62d36d0

2002-07-12 16:23:38: UFS: 3: incorrect ACL magic number : 1275068417, possibly corrupted

2002-07-12 16:23:38: UFS: 3: error reading the ACL header

2005-08-22 21:59:13: UFS: 3: getAcl: Bad gen number 0x1, on disk gen is 0x37 for slot 0x1d

2005-08-22 21:59:57: UFS: 3: Invalid free acl: real gen: 55, aclid: 1048605, rent: 424

2005-08-22 21:59:57: UFS: 3: inative:releaseAcl failed with status 28

2. Create hardlink to hidden etc directory on affected file system:

**\$ .server\_config server\_4 "shadow showetc \fs1"**

1088518964: CFS: 3: Made link for name 'ACLrecord', fileId 8

1088518964: CFS: 3: Made link for name 'aCLdata', fileId 9

**Note:** Command creates a hardlink for above two files to a directory called "slashetc" at the root of the file system in question.

3. Mount DM root from CS and cd to newly created directory:

**#cd /mnt/fs1/slashetc**

4. Save copy of ACL database records before zeroing out:

**#server\_file server\_2 -get /fs1/\etc/ACLrecord ACLrecord | ...ACLdata ACLdata**

5. Zero out ACL database by touching new file called "foo", then copying that null file over the following files:

ACLdata | ACLrecord

# touch foo

# cp foo ACLrecord

# cp foo ACLdata

# ls -l

----- 2 root root 0 Jul 12 17:32 ACLdata

----- 2 root root 0 Jul 12 17:32 ACLrecord

----- 2 root bin 1071931392 Jul 12 10:48 NameInfo0

-rw-rw-r-- 1 root root 0 Jul 12 17:32 foo

6. Remove all files from slashetc, the slashetc directory, umount from CS0, reboot DM:

# rm \*

rm: remove `ACLdata'? y

rm: remove `ACLrecord'? y

rm: remove `NameInfo0'? y

rm: remove `foo'? y

# cd ..

# rmdir slashetc

# umount /mnt

# server\_cpu server\_2 -r -m now

7. Run .server\_config server\_2 "file aclchk /fs1 fix" to rebuild ACL database

1026510218: UFS: 4: Bad Acl for inode 66034, status 7, id 4996 →These entries are normal when the ACL db is being rebuilt

1026510218: UFS: 3: getAcl: error from readACLBLOCK: 0x7

8. Perform CIFS Update of affected file system:

**\$ .server\_config server\_2 "cifs update /fs1 resetacl"**

**Note:** This step is not an absolute. Some situations may not require the use of resetacl—if customer is using Roaming profiles, citrix, Terminal Services, etc.

9. Verify access to file system

### **Additional Tool to Consider:**

**# .server\_config server\_2 -v "inodeScan"**

1235617833: UFS: 3: inodeScan action={report|reset|truncate} fsid=xxx inum=xxx

### **VIEWING ACL DATABASE INFORMATION:**

**\$ .server\_config server\_5 -v "acl database=/fs23 printstats"**

Acl database map/stats

-----  
Acl Database of 343 fs locked in shared

Acl Database of 343 fs unlocked

Statistiques of Acl database

-----  
Record file size: 0x2000 Bytes

Data file size: 0x2000 Bytes

Total number of slots: 5

Number of used slots: 4

Number of free slots: 1

Number of free chunks: 1

Average free chunk size:1 Slots

Number of SD: 4

Total SD size: 1000 Bytes

Bigest SD: 362 Bytes

Average SD size: 250 Bytes

Average SD size: 1 Slots

Total wasted space: 3096 Bytes

Average wasted space: 774 Bytes

Updated SD: 0

Updated inode: 0

Completion time: 0 sec

1126560266: ADMIN: 4: Command succeeded: acl database=/fs23 printstats

**\$ .server\_config server\_5 -v "acl database=/fs23 printmap"**

Acl database map/stats

-----  
Acl Database of 343 fs locked in shared

Slots 0x1-0x1 size=0x6a Bytes refCount=1 gen=0x0

Slots 0x2-0x2 size=0xaa Bytes refCount=1 gen=0x1

Slots 0x3-0x3 free

Slots 0x4-0x4 size=0x16a Bytes refCount=1 gen=0x1

Slots 0x5-0x5 size=0x16a Bytes refCount=70 gen=0x1

Slots 0x6-0xffffffff free

Acl Database of 343 fs unlocked

Statistiques of Acl database

-----  
Record file size: 0x2000 Bytes

Data file size: 0x2000 Bytes

Total number of slots: 5

Number of used slots: 4

Number of free slots: 1

Number of free chunks: 1

Average free chunk size:1 Slots

Number of SD: 4

Total SD size: 1000 Bytes

```

Bigest SD:      362 Bytes
Average SD size:   250 Bytes
Average SD size:    1 Slots
Total wasted space: 3096 Bytes
Average wasted space: 774 Bytes
Updated SD:       0
Updated inode:    0
Completion time: 0 sec
1126560281: ADMIN: 4: Command succeeded: acl database=/fs23 printmap

```

## **NAS ACCESS CONTROL LEVELS(ACLs)**

Used to define 'Access Control Levels' for DataMovers, Users, FileSystems, or Volumes.

### **Creating a new User or Group for ACL Tables on Celerra:**

**# nas\_acl -name acme -create -user 202 level=2**

done

**# nas\_acl -l**

| index | type | level | num_id | name     |
|-------|------|-------|--------|----------|
| 1     | user | admin | 201    | nasadmin |
| 2     | user | admin | 202    | acme     |

**Note:** Example created a new user called “Acme” with UID of 202 and “admin” level rights

### **NAS ACL LEVELS AVAILABLE: 2-9 [Default levels are 2, 3, 4]**

2=Admin      Read—Write—Delete [Highest privilege level, has access to all ‘objects’]

3=Operator    Read--Write

4=Observer    Read Only [lowest access rights]

**Note:** Setting an ACL of 432 to an ‘object’ grants Observer=RO; Operator=RW; Admin=RWD, but no Ownership since the leftmost 4<sup>th</sup> digit has been omitted. ACL 444=All users RWD access; ACL 333=Level 2&3 RWD but no access for Level 4; ACL 222=Level 2 RWD only, no access. For Levels 3 & 4; ACL 5111=Ownership given but no others have access; ACL 7432

### **NAS ACL CHART:** Examples: Admin RWD=222

| Owner | Read | Write | Delete                        |
|-------|------|-------|-------------------------------|
| 0     | 0    | 0     | [Universal access, RWD]       |
| 2     | 1    | 1     | [Admin R, no access everyone] |
| 2     | 2    | 2     | [Admin RWD, no others]        |
| 3     | 3    | 3     | [Admin/Oper RWD]              |
| 4     | 3    | 2     | [Adm RWD, Oper RW, Obs Read]  |
| 4     | 4    | 4     | [Adm/Oper/Obs RWD]            |
| #     | 0    | 0     | [Only owner access]           |

All Users RWD=444

| Owner | Read | Write | Delete                   |
|-------|------|-------|--------------------------|
| 1     | 1    | 1     | [No access except owner] |
| 2     | 2    | 1     | [Admin RW, no access]    |
| 3     | 1    | 1     | [Admin/Oper Read]        |
| 4     | 1    | 1     | [Adm/Oper/Obs Read]      |
| 4     | 4    | 1     | [Adm/Oper/Obs RW]        |
| #     | 1    | 1     | [Only Owner access]      |
| 7     | 4    | 3     | O/A RWD, Op.RW, Obs Read |

**Note:** Leftmost number would indicate “Owner”, and is particularly important when setting ACLs for Data Movers. Owner value 1 indicates nasadmin. So, all of the above examples show 3-digit ACLs. Use 4-digit values when assigning ACLs to Servers.

### **NAS ACL RULES:**

→Root User has universal access

→When no 'Owner' is specified, then a 0 in R-W-D column gives universal access to all users

→If 'Owner' defined, then a 0 or 1 in R-W-D column means 'no-access' for anyone except 'Owner'

## **CELLERRA ACL's (Not Windows ACLs):**

### **CREATING ACL's ON FILE SYSTEM:**

**# nas\_fs -acl 444 fs01**

**# nas\_fs -i fs01**

|      |        |
|------|--------|
| id   | = 23   |
| name | = fs01 |
| acl  | = 444  |

### **CREATING ACL WITH OWNER, DEFAULT OWNER NASADMIN:**

**# nas\_fs -acl 1432 fs01**

|      |        |
|------|--------|
| id   | = 23   |
| name | = fs01 |

acl = 1432, owner=nasadmin, ID=201

### **REMOVING ACL ON FILE SYSTEM:**

**# nas\_fs -acl 000 FS1**

### **CREATING ACL's FOR DATAMOVERS:**

**# nas\_server -acl 1432 server\_4**

**# nas\_server -l** [Lists ACL's for Servers]

**Note:** Notice that ACLs for Data Movers require a 4<sup>th</sup> leftmost digit that signifies Server Ownership, in this case “1” = nasadmin

**DEFAULT ACL FOR DATAMOVERS:** Set at 1000 for User “nasadmin” at installation [nasadmin=owner (1)]

### **Allowing Other Users to Run Server Commands on DataMover:**

Step 1. Create New ACL User: \$nas\_acl -name tony -create -user 500 level=2

Step 2. Change ACL on DM: \$nas\_server -acl 432 server\_2

### **DELETING ACL's FOR DATAMOVERS:**

**# nas\_server -acl 000 server\_4** [Changes ACL back to 000, the default, which = Universal Access for Everyone]

### **CREATING ACL's FOR VOLUMES:**

**# nas\_volume -acl 432 vol65**

### **CREATING ACL's FOR USERS:**

**# nas\_acl -name test1 -create -user 1001 level=3**

**# nas\_acl -n <user> -c -user 1005 level=2**

### **VERIFYING ACL's FOR USERS:**

**# nas\_acl -l** [Lists ACL's for Users & Groups only]

**# nas\_acl -i -user 1005**

### **DISPLAYING ACLs FOR USERS:**

**# nas\_acl -info -user 201**

id = 1

name = nasadmin

level = admin

user\_id = 201

**# nas\_acl -info -user 501**

id = 2

name = mvadmin

level = 9

user\_id = 501

**Note:** If nas\_server or nas\_acl produces the following error (as logged to screen & in /nas/log/nas\_log.al.err), then check the /nas/site/acl\_param file—most likely it is corrupted—repair file from SCCS backup.

2006-09-26 15:26:54.982 db:0:16555:E: nas\_acl -info -user 201: Execution failed: not\_off: Precondition violated.

[EXTRACTOR.next]

**# cat acl\_param**

1:nasadmin:1:201:2:

2:mvadmin:1:501:9:

## ***CELLERRA MULTIPROTOCOL***

### **ACCESS & AUTHENTICATION IN MULTI-PROTOCOL ENVIRONMENT: CIFS v. NFS**

**USER AUTH:** \$server\_cifs server\_x -add security=NT | UNIX | SHARE [NFS & CIFS AUTH]

**NETWORK SECURITY:** \$server\_export server\_x -o access=, root=, anon=, rw=, ro=, /mntpoint [NFS Export & SHARE access]

**FILE/DIRECTORY:** \$server\_mount server\_x -o accesspolicy=NATIVE | NT | UNIX | SECURE fs01 /fs01 [NFS UGO & CIFS ACLs]

**LOCKING:** \$server\_mount server\_x -o noblock | rlock | rwlock fs01 /fs01 [NFS Locks & CIFS Deny Modes]

**I/O:** \$server\_mount server\_x -o rw | ro fs01 /fs01 [File System Mount]

### **NFS USER AUTHENTICATION:**

--Users authenticated as part of logon process to local system, but only as a lookup to a passwd file or NIS database

--NFSv2 & 3 Clients provide UID & GID credentials when accessing remote mounts

### **CIFS USER AUTHENTICATION:**

--Users logon to workstation with Username/Password/Domain & workstation passes to DC using RPC for the 'netlogon' process. DC passes to SAM (Security Account Manager), which compares Username & Hash against database. Access Token generated that includes SID of User & SIDs of Groups that are a member of. Netlogon process then hands back SAT to workstation's LSA (Local Security Authentication), a User Profile is selected or created and Explorer Interface is built for the User.

### **CIFS FILE OWNERSHIP/PERMISSIONS:**

--All Users assigned UID & default GID

--All files also have UID & GID assigned

**Note:** Default behavior is to assign the User with the GID of the Primary Windows Group (usually Domain Users). When setting useUnixGid=1, DART will assign the User's GID (Primary Group) based on the GID field in the local ./etc/passwd file or the NIS map for the User:

**param cifs useUnixGid=1**

## **MULTIPROTOCOL MODEL:**

### **Layer                      Security Entity**

|                |                                |                 |
|----------------|--------------------------------|-----------------|
| User Auth      | NFS Auth                       | CIFS Auth       |
| Network        | NFS Export                     | Share ACL       |
| File/Directory | NFS ugo-rwx                    | CIFS ACL        |
| Locking        | NFS Locks                      | CIFS Deny Modes |
| Structure      | File System Namespace for both |                 |

**Note:** NFS users present UID/GIDs to Server when authenticating, and are trusted by the server. CIFS users are authenticated by Domain Controllers or localgroups database for Local Users Support, SIDs mapped to UID/GIDs, and access-checking based on SIDs. In general, Celerra mappings between both protocols must use same User name identity, and SIDs are translated to UID/GIDs on the file system. Celerra uses default umask values for CIFS shares so that files and directories will have a default UNIX permission set. NFS Exports can provide some security levels. Translating ACLs to UNIX mode bits and back are not 100% reliable. Current recommended solution is to use MIXED accesspolicy, which will always use the ACL for access-checking, and allows the best translation of ACLs to UNIX, and NT Credentials for Unix if you need to be able to translate more than 16 UNIX groups into the ACL. Files will receive permissions based on directory permissions or inheritance. Celerra uses a single Lock Manager for both CIFS and NFS.

## **MAKING CIFS PERMISSIONS OBEY UNIX PERMISSIONS IN MIXED ENVIRONMENTS:** NAS

4.1.7.2 +

### **RULE ONE:**

***Do not use Usrmapper in mixed environments where the same User, whether from Unix or NT, needs to access the same files!***

**Comment:** Unix perms can be used to manage file systems when Unix-based administration & usage is more common than CIFS

### **General Eng. Guidance:**

- Ensure that all UNIX groups map cleanly into corresponding NT Groups
- Use accesspolicy=UNIX for mounting file systems
- Ensure that umask on all NT CIFS Shares matches those used by NFS Clients

### **Set the following parameters on Datamovers:**

/nas/site/slot\_param

**param cifs useUnixGid=1**

**Purpose:** Default DART behavior is to apply primary group membership as ‘Domain Users’. When changed to a value of 1, files created by a Windows user will apply a GID matched from either a local passwd file or the NIS database.

**param cifs useUnixGid=0**

**Note:** Windows User default GID of Domain Users is assigned to newly created files, directories—only applies to CIFS Users.

**param cifs acl.extendExtraGid=1**

**Note:** When set to 1, this parameter will map up to 15 additional Unix Secondary groups to NT Users, as found in the NIS or local Group files for that User. Data Mover will parse the NIS or ./etc/passwd file to match Username, then locate any Secondary Groups that a User is a member of from the NIS map or /etc/group file.

## **VIEWING UNIX PERMISSIONS FROM WINDOWS:**

**param cifs acl.extacl=2**

[Unix access rights are visible on Windows ACLs for files & directories—Windows Users can view Unix perms and change]

**param cifs acl.takegroupship=0**

**Note:** =1 Sets the new Primary group on files changed by NT user when useUnixGid=1 is in effect (0=default)

**param cifs acl.restrictedTakeOwnership=1**

**Note:** Enables assignment of GID based on source specified by acl.useUnixGid parameter. Purpose of param is to help determine source of GID for copied files from Windows User.

**param cifs acl.restrictedTakeOwnership=1**

**Note:** When set to 1 for UNIX or SECURE accesspolicy, only Root User (uid=0x0) can take Ownership of unix-permissioned files (except for Backup/Restore privilege)

**param cifs resolver=1**

**Note:** Using the “cifs resolver=1” enables DART to matchup a User’s Name from local passwd file or a NIS database. The first

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
preference will be to matchup the username without the domain extension applied. However, if a mapping cannot be made, a 2nd attempt to resolve UIDs is made by applying the ‘.domain’ extension to a User’s name.

### **param cifs resolver=0**

**Note:** Username.domain extension is always used when mapping UID/GIDs from passwd or NIS rather than just ‘username’.

#### **Example:**

murphys:x:35012:1000:Law, Murphys=murphys:/home/murphys:/bin/ksh [ordinary Unix-style & Celerra's new method with parameter "cifs resolver=1" enabled]

murphys.compucom:x:35012:1000:Law, Murphys=murphys:: [NT Domain style--old method]

**Comment:** User's name would be 'parsed' from Password file or NIS Database from the first line of the example if the "cifs resolver=1" option is selected. Otherwise, a 2nd line item would be required to define the "user.domain".

### **param cifs acl.checkacl=0**

**Note:** In this mode, only Unix permissions are checked—NT ACL’s or SIDs are not checked. This bit is set to 0 by default whenever CIFS is stopped.

### **param cifs acl.checkacl=1**

**Note:** Both Unix and Windows ACLs are checked. Default value for Celerra. Bit is set to 1 whenever CIFS is started.

## **SETTING ACL BEHAVIOR:**

**Note:** extacl parameter is used to manage special capabilities built around ACL management for two basic functions:

- 1.) Backup/Restore of UNIX attributes, symbolic links, access rights, UNIX modes using CIFS-based backup tools--NTBackup
- 2.) To view/change UNIX access rights from ACL management tools, like Windows Explorer

**param cifs acl.extacl=0** →(Default)Bit 0 param=0, DM presents UNIX meta-data for files/dirs for CIFS backups--EMC ACE Type →Bit 0 param=1, Celerra uses Standard ACE Type and encodes info in SID for the ACE

**param cifs acl.extacl=2** →Bit 1 set, param=2, CIFS users can view and change Unix perms; Unix perms show as (3) ACEs in ACL

**param cifs acl.extacl=4** →Bit 2 set, param=4, Unix perms presented in ACL so that CIFS backup applications can backup/restore

**param cifs acl.extacl=8** →Bit 3 set, param=8, CIFS apps. backup Unix symb. links with special ACL as 0-byte files

**Note:** When param is not set, CIFS Backups would follow the symbolic link to backup actual files

**param cifs acl.extacl=16** →Bit 4 set, param=16, CIFS apps. Backup UNIX names (all (3) names, M256, DOS 8.3, UNIX)

**Note:** UNIX names are encoded in special ACE in the file so that CIFS applications can backup/restore

**param cifs acl.extacl=32** →Bit 5 set, param=32, allows UNIX clients to View and Modify ACLs on Celerra using emcsetsd

**param cifs acl.extacl=64** →Bit6 set, param=64, UNIX rights are applied by granted rights and denied rights by DACL.

Normal behavior is for UNIX rights to files to be granted rights + rights not denied in DACL. Additionally, when this bit is set, requests to change files or directories are denied if they involve any of the (3) specil inheritance ACEs. In practice, this means that ACLs must be set using the Advanced options in Security>Properties from Explorer (i.e., uncheck Inheritance)

## **BINARY BIT SETTINGS TABLE:**

**2^6=64    2^5=32    2^4=16    2^3=8 (Bit3, etc)    2^2=4 (Bit2)    2^1=2 (Bit1)    2^0=1 (Bit0)**

Each of the above represents Bit0 through Bit6, as read from right-to-left. For example, if you wanted to set the Bit5 value for supporting the emcsetsd tool, the decimal value for the param would be 32 (or a range from 32-63). If you wanted to set the Backup/Restore symbolic link value, Bit3, the decimal value would be 15 (Base2 binary values for Bits 0, 1, 2, & 3 are combined). By having a decimal param value of 0 set means that none of the above Bits 0-6 are set, but a param value of 1 means that Bit0 is set.

#### **Shorthand Table:**

Bit0 = mask value 1 is set

Bit1 = mask value 2 is set

Bit2 = mask value 4 is set

Bit3 = mask value 8 is set

Bit4 = mask value 16 is set

Bit5 = mask value 32 is set

## **MOCKUP OF BIT TABLES AND BINARY VALUES:**

| Bit 6<br>Binary<br>Values | Dec.<br>Valu<br>es | Bit 5<br>Binary<br>Values | Dec.<br>Valu<br>es | Bit 4<br>Binary<br>Values | Dec.<br>Valu<br>es | Bit 3<br>Binary<br>Values | Dec.<br>Valu<br>es | Bit 2<br>Binary<br>Values | Dec.<br>Valu<br>es | Bit 1<br>Binary<br>Values | Dec.<br>Valu<br>es | Bit 0<br>Binary<br>Values | Dec.<br>Valu<br>es |
|---------------------------|--------------------|---------------------------|--------------------|---------------------------|--------------------|---------------------------|--------------------|---------------------------|--------------------|---------------------------|--------------------|---------------------------|--------------------|
| 1000000                   | 64                 | 0100000                   | 32                 | 0010000                   | 16                 | 0001000                   | 8                  | 0000100                   | 4                  | 0000010                   | 2                  | 0000000                   | 0                  |
|                           |                    | 0100001                   | 33                 | 0010001                   | 17                 | 0001001                   | 9                  | 0000101                   | 5                  | 0000011                   | 3                  | 0000001                   | 1                  |
|                           |                    | 0100010                   | 34                 | 0010010                   | 18                 | 0001010                   | 10                 | 0000110                   | 6                  |                           |                    |                           |                    |
|                           |                    | 0100011                   | 35                 | 0010011                   | 19                 | 0001011                   | 11                 | 0000111                   | 7                  |                           |                    |                           |                    |
|                           |                    | 0100100                   | 36                 | 0010100                   | 20                 | 0001100                   | 12                 |                           |                    |                           |                    |                           |                    |
|                           |                    | 0100101                   | 37                 | 0010101                   | 21                 | 0001101                   | 13                 |                           |                    |                           |                    |                           |                    |
|                           |                    | 0100110                   | 38                 | 0010110                   | 22                 | 0001110                   | 14                 |                           |                    |                           |                    |                           |                    |

|         |    |         |    |         |    |
|---------|----|---------|----|---------|----|
| 0100111 | 39 | 0010111 | 23 | 0001111 | 15 |
| 0101000 | 40 | 0011000 | 24 |         |    |
| 0101001 | 41 | 0011001 | 25 |         |    |
| 0101010 | 42 | 0011010 | 26 |         |    |
| 0101011 | 43 | 0011011 | 27 |         |    |
| 0101100 | 44 | 0011100 | 28 |         |    |
| 0101101 | 45 | 0011101 | 29 |         |    |
| 0101110 | 46 | 0011110 | 30 |         |    |
| 0101111 | 47 | 0011111 | 31 |         |    |
| 0110000 | 48 |         |    |         |    |
| 0110001 | 49 |         |    |         |    |
| 0110010 | 50 |         |    |         |    |
| 0110011 | 51 |         |    |         |    |
| 0110100 | 52 |         |    |         |    |
| 0110101 | 53 |         |    |         |    |
| 0110110 | 54 |         |    |         |    |
| 0110111 | 55 |         |    |         |    |
| 0111000 | 56 |         |    |         |    |
| 0111001 | 57 |         |    |         |    |
| 0111010 | 58 |         |    |         |    |
| 0111011 | 59 |         |    |         |    |
| 0111100 | 60 |         |    |         |    |
| 0111101 | 61 |         |    |         |    |
| 0111110 | 62 |         |    |         |    |
| 0111111 | 63 |         |    |         |    |

## **CONFIGURING DART TO APPLY PERMS TO USERS WITH SECONDARY UNIX GROUPS:**

Introduced with 4.2.13.100:

**param cifs acl.extendExtraGid=1**

**param cifs resolver=1**

**Note:** When using the new “extendExtraGid=1” parameter, must also use “cifs resolver=1” param.

Explanation:

Param allows Windows User to be able to add up to (15) Secondary Unix Groups to their NT Credential when logging in from CIFS, using either NIS or local /etc/passwd and group files. Customer was using NIS Service to map Unix Users and Groups, “accesspolicy=SECURE”, a umask of “027” on all Shares, & “param cifs resolver=1”. Whole point is that customer did not wish to manually create Windows 2000 Groups for every Unix Secondary Groups.

**Note:** There is a general NFS limitation, when using AUTH\_SYS RPC Security Calls, of (16) Groups that a Unix User can be a member of. Solaris 8 can be tuned to allow for a total of 32 Groups per User (Future NAS may allow 32 Groups).

## **USING EMCSETSD & EMCGETSD UTILITIES FOR VIEWING/SETTING ACLS FROM UNIX HOST (Introduced NAS 5.2):**

**Note:** These tools are unix-based and are designed to View and/or Modify NT ACLs when remote mounting a Celerra CIFS File System from either a Solaris, HPUX, or Linux Host. The emcgetsd and emcsetsd tools are found on the Celerra Applications & Tools CD>CifsTools>unixactools>hpxu | linux | solaris directories. You can also run the tool from the Control Station by simply copying the Linux binary files to /home/nasadmin.

1. Add the following param to the Data Mover and reboot (only required for emcsetsd tool, not emcgetsd):

**# server\_param server\_2 -facility cifs -modify acl.extacl -value 32**

**Note:** Using server\_param to change the acl.extacl value does not require reboot

**param cifs acl.extacl=32**

2. Create CIFS Share for file system to be setup/tested

share "emc" "/fs1" maxusr=4294967295 umask=22

3. Copy the "emcsetsd" and "emcgetsd" utilities from the Applications & Tools CD to a directory on the Control Station

4. Make the binary utilities executable:

**# chmod 755 emesetsd emcgetsd**

5. Mount the Data Mover from the CS (may need to export file system to CS IP address for RW access):

## # mount server\_2://fs1 /dm

6. Add a User to a Folder Path with permissions and flags desired:

# ./emcsetsd -D win2kemc -g tommie,RWXPDO,0x3 /dm/fs1/SD

Server=server\_2, Path in the server=//fs1/SD

1 ACE added

7. Verify by using "emcgetsd":

# ./emcgetsd /dm/fs1/SD

Server=server\_2, Path in the server=//fs1/SD

Dump of //fs1/SD Security Descriptor

Owner uid=32768 NT='WIN2KEMC\Administrator'

Group gid=32770 NT='WIN2KEMC\Domain Admins'

----- DACL -----

GRANT uid=32772 NT='WIN2KEMC\tommie'

Access RWXPDO 0xd013f

Flags 0x3

OBJECT\_INHERIT

CONTAINER\_INHERIT

**Note:** Path may be 'case sensitive'

# /home/nasadmin/emcgetsd

/nasmcd/quota/slot\_5//fs03/vol331/shots/ose/rad.b05/maya/projects/fastSkin/are\_all/temp/bake

Server=server\_5, Path in the server=//fs03/vol331/shots/ose/rad.b05/maya/projects/fastSkin/are\_all/temp/bake

Dump of //fs03/vol331/shots/ose/rad.b05/maya/projects/fastSkin/are\_all/temp/bake

Security Descriptor

Owner uid=9690 Unix='jvacanti' NT='UNIX UID=0x25da jvacanti"

Group gid=20 Unix='user' NT='UNIX GID=0x14 'user"

----- DACL -----

GRANT ??? NT='UNIX UID=0x25da jvacanti" -->getsd dump does NOT show that Owner can delete

Access RWXP-O 0x1e01ff

ReadExecute

Read

Write

ListFolderContents

Flags 0x0

GRANT ALL NT='Everyone'

Access RWXP-- 0x1601ff

ReadExecute

Read

Write

ListFolderContents

Flags 0x0

----- SACL -----

None

## USING EMCSETSD/EMCGETSD & MANAGING UNIX PERMISSIONS FROM WINDOWS HOST:

\$ server\_param server\_2 -facility cifs -modify acl.extacl -value 34

**Note:** Example of the decimal value setting required to use the Unix tools and to manage Unix permissions from a Windows host

## VIEWING WINDOWS ACLs FROM UNIX CLIENT:

# ./emcgetsd -v /dm/fs\_quota

Server=server\_2, Path in the server=//fs\_quota

Dump of //fs\_quota Security Descriptor

Owner uid=0 Unix='root' NT='UNIX UID=0x0 'root"

Group gid=0 Unix='root' NT='UNIX GID=0x0 'root"

----- DACL -----

GRANT ALL NT='Everyone'

Access RWXPDO 0x1f01ff

**Note:** Displays full ACL & Security Descriptor information—Ownership, Access Rights, ACL, Auditing information

# ./emcgetsd -a /dm/fs1 [Displays Access Rights of User logged in from Unix client]

Server=192.1.10.72, Path in the server=/fs1/fs1

Access of user uid=0, gid=0 groups=1,2,3,4,6,10 on /fs1/fs1

- FullControl
- Modify
- ReadExecute
- Read
- Write
- ListFolderContents

#### **EXAMPLE OF SETTING ACL WITHOUT acl.extacl SET TO 0x20 (decimal 32):**

```
# ./emcsetsd -D w2k -g nasadmin,fullcontrol /dm/fs1/lun_trespass.txt
```

Server=server\_2, Path in the server=/fs1/lun\_trespass.txt

Unknown Error

Set SD not allowed (param cifs.acl.extacl not set to 0x20) → 0x20 in decimal is 32

```
# server_param server_2 -facility cifs -modify acl.extacl -value 32
```

server\_2 : done

#### **ADDING ACE ENTRIES USING EMCSETSD:**

```
# ./emcsetsd -D w2k -g nasadmin,fullcontrol /dm/fs1/lun_trespass.txt
```

Server=server\_2, Path in the server=/fs1/lun\_trespass.txt

1 ACE added

```
# ./emcgetsd /dm/fs1/lun_trespass.txt
```

Server=192.1.10.72, Path in the server=/fs1/lun\_trespass.txt

Dump of /fs1/lun\_trespass.txt Security Descriptor

Owner uid=32768 NT='BUILTIN\Administrators'

Group gid=32770 NT='W2K\Domain Users'

----- DACL -----

GRANT uid=32769 NT='W2K\nasadmin'

Access RWXPDO 0x1f01ff

- FullControl

- Modify

- ReadExecute

- Read

- Write

- ListFolderContents

Flags 0x0

```
# ./emcsetsd -D w2k -g celadmin,rwx,0x3 /dm/fs1/lun_trespass.txt
```

Server=192.1.10.72, Path in the server=/fs1/lun\_trespass.txt

1 ACE added

```
# ./emcgetsd /dm/fs1/lun_trespass.txt
```

Server=192.1.10.72, Path in the server=/fs1/lun\_trespass.txt

Dump of /fs1/lun\_trespass.txt Security Descriptor

Owner uid=32768 NT='BUILTIN\Administrators'

Group gid=32770 NT='W2K\Domain Users'

----- DACL -----

GRANT uid=32770 NT='W2K\celadmin'

Access RWX--- 0x13f

Flags 0x3

- OBJECT\_INHERIT

- CONTAINER\_INHERIT

#### **PROBLEM OF UNRESOLVED SIDs: MIGRATIONS OR TAPE RESTORES**

Scenario is that in complex multi-domain environments, there is often the possibility that data will be copied to the Celerra containing SIDs that cannot be resolved. Default behavior is to either fail when copying data or to assign Everyone Full Control. Because of this, Engineering has developed the concept of the ‘delayed SID’—the concept of setting a SID to an ACL without completing authentication. This condition will occur if the SID is unknown on the DC or if there is no Unix mapping in the passwd file, NIS database, or Usrmapper server.

#### **BASIC PARAMETERS TO TRY WHEN MIGRATING UNRESOLVABLE/BROKEN SIDS:**

**Note:** First used for Emcopy issue

```
param cifs acl.mappingErrorAction=3
```

```
param cifs acl.retryAuthSid=600
```

**param cifs acl.FailOnSDRestoreError=0** [Consider setting to 1 if msg desired in Server Log]

acl.mappingErrorAction=3 → Store ACE & SID, but don't try again if resolution to DC fails

acl.mappingErrorAction=1 → Save ACE & SID, keep SID in pending state if resolution to DC fails

acl.mappingErrorAction=0 → Do not save SID & ACE if resolution to DC fails

#### **DEFINITIONS:**

**param cifs acl.mappingErrorAction=3** (means that Bits 0 & 1 are set)

**Note:** This parameter defines rules for unknown mappings between SID, UID, GID on ACL setting, either because DC does not know SID or User Name is not yet mapped to UID/GID. Max value is decimal 15 for this param in NAS 5.4

Bit 0 = stores unknown SIDs (decimal value=1)

Bit 1 = stores SIDs without Unix GID/UID mappings (decimal value=3)

Bit 2 = enable debug traces (decimal value=7, the max. value that can be set for this parameter)

Bit 3 = Only do lookup in cache (secmap or globalSidcache or per connection SID cache—decimal value=15)

**param cifs acl.retryAuthSid=600**

**Note:** This parameter sets the timeout value of 600sec. between authentication retries for unknown SIDs

**param cifs acl.FailOnSDRestoreError=0**

**Note:** This parameter suppresses any errors resulting from unresolved ACLs during migration of files to Celerra, migrating all resolvable SIDs properly and placing unresolved SIDs in the delayed SID state. Without this setting, errors will be returned when copying files with unresolvable SIDs, resulting in copying of files with default Everyone Full Control applied. Use these parameters when conducting an EMCopy/CDMS migration or any type of migration involving unresolvable SIDs.

#### **CHANGING DEFAULT CIFS TIMEOUT MAY HELP IN CASES OF UNRESOLVABLE SIDS:**

In certain situations, such as for orphaned SIDs, or SIDs that cannot be resolved in other domains because of Trusts or firewall issues, Clients may experience long “pauses” as directories are traversed—this is because of the 20sec CIFS timeout value. In extreme cases, setting this value to 1388=5000=5 secs may help speed things up.

cifs.NTsec.DCTimeout 0x015d6b54 0x00004e20 0x00004e20 [default value 4e20=20,000=20secs

#### **CELERRA AND NEGATIVE CACHING OF UNKNOWN SIDS:**

Prior to NAS 5.3.2, 5.2.11, and 5.1.23, Celerra did not cache unknown Sids. What would happen is that each time the unknown Sid was presented, we went off to DC's trying to lookup SID, eventually causing performance problems with large networks. Fix would be to ensure that Global Sid Caching is enabled on the DM. We should now cache negative lookups so that we don't continually try to resolve the same unknown SID. Besides Server Log messages, a key indicator of the health of authentication sessions are SessionSetups, as seen from server\_cifsstat. See AR41213.

**Server Log Example of Successful Negative Caching:** 2004-03-26 09:32:26: SMB: 3:

Usr='AUdoggetsm':Grp27=:SidTypeUnknown error=SUCCESS SID=S-1-5-15-8b33b8-7481765b-250e54d6-24b3

**Note:** This feature applies only to Group SIDs

#### **USING SETACL TO CLEANUP ORPHANED SIDS FROM CELERRA FILE SYSTEMS:**

Set the following data mover params, reboot, and then run Setacl command to flush unresolvable or unknown or orphaned SIDs from the Celerra File Systems—reboot after Setacl is completed. Always verify that UID/GID mapping database is correct, that correct Interface is specified for the respective Comname, and that CIFS is up and running during the Setacl process:

1. Set following params on Data Mover:

**param cifs acl.FailOnSDRestoreError=0**

**param cifs acl.mappingErrorAction=3**

2. Reboot Data Mover

3. Set SMB Debug Logging:

**\$server\_config server\_x -v "logsys set severity SMB=LOG\_DBG3"**

4. Run Setacl Against File System using Correct Interface Name to flush unresolvable, unknown, or orphaned SIDs:

**\$server\_config server\_x "cifs update /mntpoint force setacl if=fsn0"**

**# .server\_config pal1 "cifs update /pal\_fs2 force setacl if=pal2"** [Setacl using VDM Container name ‘pal1’]

**Caution:** When running setacl, do so from top-level share of each Comname, making sure to specify the correct interface name for the comname. Failure to do this could replace built-in localgroup information with the wrong comname. Also, you may need to set **SMB=LOG\_DBG3** in order to see cifs update entries in Server Log.

2005-07-01 08:47:17: ADMIN: 4:[pal1] Command succeeded: :1 cifs update /pal\_fs1 force setacl if=pal1

2005-07-01 08:53:00: SMB: 4:[pal1] Update of : /root\_vdm\_1/pal\_fs1 completed after 343 s

2005-07-01 08:53:00: SMB: 4:[pal1] Update of : 1452338 files, 119766 dirs, 1572093 Acl

5. Reboot Server after setacl is completed on all file systems

#### **ACL RECORDS:**

ACL records and database are stored in /etc directory of each file system, recording checksum, reference count, size of data.

ACL record has header with ACLID magic number, revision, and next free acl record slot.

ACL records and acl data are compared—acl record entries are deleted if there is a mismatch

## **RESETTING ACL's: \$server\_config server\_2 "cifs update /mountpoint/path resetacl"**

An important tool for "access denied" issues on specific CIFS files/folders is the "resetacl" command, which restores "Everyone Full Control" to the folder that it is run against. Use this as a quick way to restore administrative access to a subfolder or CIFS Share.

**Caution:** Administrator will have to reset NT Perms on the affected folder for the Users and Groups that were previously assigned.

**\$server\_config server\_2 "acl reset=/etc/PIPE/netlogon"**

**\$server\_config server\_2 "acl reset=/etc/PIPE/srvsvc" [or \$PIPE?]**

**Note:** Resetacl CIFS Update rebuilds the Shadow file and resets all CIFS attributes back to their default [Archive On, Read-Only, Hidden, and System Off], meaning that tape backups would run the equivalent of a full backup. Use resetacl only as a last resort and only if the customer agrees, as this command will open up all directories and files to Everyone Full Control access and would require extensive customer work in reapplying NTFS permissions. If required, the resetacl can be run on a Server that is not running a CIFS configuration as there is no need to consult with the Domain Controllers while running this command.

## **RUN “CIFS UPDATE” TO RE-LINK NFS FILE NAMES & CIFS 8.3 FILE NAMES:**

Step 1. Check CPU of Server prior to kicking off cifs update

Step 2. If there are Users connect to the Share that you are updating, stop and restart CIFS

**Note:** The cifs update will not start if Users are connected to the Share (SMB: 3: Error: Update cannot be done on share in use)

Step 3. Set SMB debug logging in order to see entries in Server Log:

**# .server\_config server\_2 "logsys set severity SMB=LOG\_DBG3"**

Step 4. Run CIFS Update Command:

**# .server\_config server\_2 "cifs update /sf02\_fs1 force level=0"**

command(s) succeeded

**Note:** Must see the “command(s) succeeded” message or the process is not running

Step 5. Verify progress of CIFS Update by checking command log—should see an entry for command succeeding, then an update every 600 seconds, and finally, an update completion message. Use server\_sysstat to check CPU usage during the update—a cifs update should use 5-15% CPU load.

2005-07-01 10:48:06: ADMIN: 4:[sf02] Command succeeded: :1 cifs update /sf02\_fs1 force setacl if=sf02

2005-07-01 10:58:06 :SMB: 4: Update of :/sf02\_fs1 running since 600 s 340639 files, 42093 dirs have already been updated

2005-07-01 10:60:50: SMB: 4:[sf02] Update of :/sf02\_fs1 completed after 164 s

2005-07-01 10:60:50: SMB: 4:[sf02] Update of : 418206 files, 55090 dirs, 473292 Acl

Step 6. Make sure to turn off SMB debug logging when completed:

**# .server\_config server\_2 "logsys set severity SMB=LOG\_PRINTF"**

## **USING SETACL TO UPDATE UID/GID VALUES ON-DISK:**

**Intro:** The “setacl” command forces the Datamover to read the ACL List from each Folder/File on the directory tree to obtain SID information. DM requests DC validation of each User & Group SID. Upon authentication by DC, DM searches queries Usrmapper for SID match, and reapplies UID/GID information associated with that SID to each User or Group’s ACE. If Usrmapper is not used, then the mappings change from SID match to Username/Groupname match, using 1.) Local DM “passwd” or “group” file 2.) NIS Database files 3.) Usrmapper Database

### **IMPORTANT CAVEATS PRIOR TO RUNNING SETACL:**

1. Please understand what the correct “source” of User and Group Mappings are before running “Setacl”. For instance, if local passwd and group files are populated on the data mover, but Usrmapper and SID History are in use, then you would want to rename the local passwd/group files prior to running setacl so as to get the correct UID-to-SID mapping from Usrmapper--this becomes an issue mainly when multiple SIDs are mapped for a User that may have the same name.

2. Please advise customer that a reboot may be necessary prior to starting setacl (and is actually preferred), as well as stopping & restarting CIFS, and that a reboot is always required after completing setacl on a given Server.

3. Please note that the CIFS process must be running during the setacl process, but that if many Users are actively connected, you may need to stop and restart CIFS in order to make setacl run.

4. Always verify that the correct Interface is matched up with the correct Compname before running setacl—failure to do this could adversely affect permissions on the entire file system. Use server\_export, server\_cifs, etc., to match-up compnames to interfaces to file systems.

5. Be aware that if local group permissioning is in use and the Domain Users or Domain Admins GID has changed, that you may need to run an “lg update force” before running the setacl.

6. Set SMB Debug logging prior to running setacl:

**# .server\_config sf02 "logsys set severity SMB=LOG\_DBG3"**

7. Start setacl and monitor Server Log. If the process seems to run slow, the DC that we are connected to may not be the best DC to use for the setacl process. Reboot the Data Mover to stop setacl & prepare for invalidate/validate. Alternatively, have customer reboot Domain Controller. Use sysstat to monitor cpu levels--normal CPU usage for a setacl should be between 10-20%:

```
$ .server_config server_2 -v "pdc invalidate=168.247.33.200"  
$ .server_config server_2 -v "pdc validate=168.247.34.250"  
$ .server_config server_2 "cifs update /mntpoint force setacl if=fsn0"
```

**Note:** Make sure “Command(s) succeeded” prompt returns. If prompt returns and complains about “share is in use”, then you will need to stop and restart CIFS and try the command again. Monitor server log for status of the update.

8. Make sure to turn off debug logging when completed:

2005-07-01 10:48:06: ADMIN: 4:[sf02] Command succeeded: :1 cifs update /sf02\_fs1 force setacl if=sf02

2005-07-01 10:50:50: SMB: 4:[sf02] Update of : /root\_vdm\_1/sf02\_fs1 completed after 164 s

2005-07-01 10:50:50: SMB: 4:[sf02] Update of : 418206 files, 55090 dirs, 473292 Acl

```
# .server_config server_2 "logsys set severity SMB=LOG_PRINTF"
```

## **HOW USERS/GROUPS AUTHENTICATE TO DATA MOVER FOR ACCESS:**

1. User conducts SessionSetup connection to Data Mover, presenting NT/Win2k Access Token [Access Tokens are built by Domain Controller every time a User successfully logs into the Domain]
2. If SecMap Cache is not enabled (NAS 5.2+) and SIDs are not stored locally, DM must validate SIDs of User and Groups
3. DM receives confirmation of User/Group SIDs, combines with Security Access Token & localgroups db information, & caches User credentials for that logon session
4. Dart checks User Credential against file system objects before allowing access to folders and file
5. Objects created by User combines UID/GID information with SID in the ACL, which gets written to the file or directory object as an ACE

## **EXAMPLE OF SETACL PROCEDURE AFTER REBUILDING USRMAPPER DATABASE:**

### **RUNNING SETACL COMMAND TO UPDATE FILE SYSTEM ACLs WITH NEW UID/GIDs:**

**EXAMPLE:** Duplicate UID/GIDs have been assigned due to Usrmapper issue. New Usrmapper database constructed and converted into Usrmapper Version 3.1.2 or higher

1. Remove old passwd/group files from all data movers
2. Replace Usrmapper database with newly converted or repaired database
3. Start Usrmapper and CIFS
4. Always run ACL DUMP against known directory to validate BEFORE SETACL & AFTER SETACL changes to UID/GIDs:

```
$ .server_config server_6 -v "acl if=fsn0 dump=/mnt4/its"
```

5. Update localgroups database on each Data Mover by running:

```
$ .server_config server_5 -v "lg update force"
```

**Note:** This will read each localgroups entry and update GIDs based on Usrmapper, or other mapping source

6. Reboot Server to refresh memory and disconnect all CIFS users

7. Set SMB Debug logging:

```
# .server_config sf02 "logsys set severity SMB=LOG_DBG3"
```

8. Leave CIFS running and immediately issue setacl command

```
$ .server_config server_5 "cifs update /fs01 force setacl if=interface0"
```

**Note:** If setacl fails because ‘shares are in use’, stop & restart cifs and issue command again

9. Track progress of setacl in the Server Log:

**2003-08-23 17:11:36: SMB: 4: Update of : /fs01 running since 9000 s 39854 files, 2554 dirs have already been updated**

2005-07-01 10:50:50: SMB: 4:[sf02] Update of : /root\_vdm\_1/sf02\_fs1 completed after 164 s

2005-07-01 10:50:50: SMB: 4:[sf02] Update of : 418206 files, 55090 dirs, 473292 Acl

10. After Setacls have completed, reboot data mover (will also turn off debug logging)

## **MONITORING SETACL PROGRESS:**

Setacl will update the Server Log every 600 secs. If you encounter the following errors in the log, it doesn’t mean that setacl has completely failed, but it could be indicating a problem with a Domain Controller or unresolvable SIDs. Setacl should rip through a 300GB file system in a half hour or less. To monitor progress, run server\_df and record the number of inodes used in the file system. Also, run server\_sysstat to see if CPU is being utilized—DM should use 10-20% CPU for a setacl. Then, at any of the update points, see how fast the setacl check is progressing. In the above example, a 69GB file system was taking an extremely long time to run and indicated a problem with a DC and Invalid SIDs that could not be resolved. The following corrective steps were taken:

## **SETACL FAILURE--DOES NOT COMPLETE--SERVER LOGS ERRORS:**

2003-08-23 14:41:40: SMB: 4: Update /fs01 cannot fix ACL for  
/fs01/fs01/usmt/nttoxp/userdata/SPlatts/USMT2I.UNC/C/DATA/NOTES/Lotusme1.dic c0000078

2003-08-23 14:41:41: SMB: 4: Update /fs01 cannot fix ACL for  
/fs01/fs01/usmt/nttoxp/userdata/SPlatts/USMT2I.UNC/C/DATA/NOTES/mail.box c0000078

**ERROR TRANSLATION:**

C:>err c0000078  
# for hex 0xc0000078 / decimal -1073741704 :

**STATUS\_INVALID\_SID ntstatus.h**

**CORRECTIVE STEPS FOR RUNNING SUCCESSFUL SETACL:**

1. Set following params to ignore Unknown SID mapping issues:

```
param cifs acl.mappingErrorAction=3
param cifs acl.retryAuthSid=600
param cifs acl.FailOnSDRestoreError=0
```

2. Reboot Data Mover to stop setacl and to set params

3. Force connection to another Domain Controller by issuing:

**\$ .server\_config server\_5 -v "pdc invalidate=168.247.33.200"**

**\$ .server\_config server\_5 -v "pdc validate=168.247.34.250"**

**RESULTS:**

**DC=PLANTDFS01(168.247.33.200) ref=2 time=3 ms INVALID [Invalidate this DC by using “pdc invalidate=” command]**

DC=\*SMBSERVER(168.247.33.73) ref=2 INVALID

DC=\*SMBSERVER(168.247.33.74) ref=2 INVALID

**>DC=PLANTDFS06(168.247.34.250) ref=13 time=0 ms [Connected to this DC using “pdc validate” command]**

**Note:** Stop and restart CIFS if needed to force DC switch

4. Alternatively, have customer reboot the Domain Controller

**Note:** Above steps were used at Kemper. After switching to new DC, setacls ran quickly and successfully:

**2<sup>nd</sup> SETACL ATTEMPT ON FS01:** Completed in just under 6 minutes

2003-08-23 18:42:46: SMB: 4: Update of : /fs01 completed after 359 s

2003-08-23 18:42:46: SMB: 4: Update of : 178322 files, 15214 dirs, 193533 Acl

**SETACL LIMITATIONS & PRECAUTIONS:**

--Always run ACL DUMP on a known directory where you expect GIDs or UIDs to be updated BEFORE & then AFTER “setacl”

**\$ .server\_config server\_6 -v “acl if=fsn0 dump=/mnt4/its”**

--Run setacl with Force flag:

**\$ .server\_config server\_6 “cifs update /mntpoint force setacl if=fsn0”**

--“Setacl” will not run against ATM interfaces and will panic the data mover—only known workaround is to have a fast Ethernet interface and add to cifs configuration using –Enable command and use this for the setacl

--If using “SIDHistory”, whereby Users may have multiple SIDs, this command may not function as expected

--Always run with CIFS started

**EXAMPLE OF ACLDUMP OUTPUT:** [with SIDHistory SID involved]

**# .server\_config server\_2 -v "acl if=fsn0 dump=/personal/its/dbarker"**

===== UNIX =====

1033850863: SECURITY: 4: Group file /.etc/passwd modified: T=3d9e2483/3d9e2462 S=58374/610413

USER osgadminje.suncor:2004 GROUP 301 mode=rwxr-xr-x

===== NT =====

Owner=USER osgadminje.suncor:2004 OSGADMINJE.SUNCOR:S-1-5-15-7417855-188b389a-30f211cf-528

1033850863: SECURITY: 4: Group file /.etc/group modified: T=3d9e2462/3d9e2483 S=610413/58374

Group=GROUP domain=20users.suncor:301 Domain Users.SUNCOR:S-1-5-15-7417855-188b389a-30f211cf-201

1033850863: SMB: 5: History SID

S-1-5-15-7417855-188b389a-30f211cf-697

**1033850863: SMB: 5: History SID**

**S-1-5-15-7417855-188b389a-30f211cf-697**

**Owner=USER 4607 dbarker.NETWORK:S-1-5-15-7417855-188b389a-30f211cf-697**

ALLOWED Flags=3 Mask=1301bf Rights=RWX-D-

1033850863: SMB: 5: ExtractNames:Usr='NETWORK\000017F6' RID=17f6 U=8 UID=0 T=3 (3)

1033850863: SMB: 5: lookupSIDs:bad reply=c0000073

1033850863: SMB: 5: MsError sendLookupSIDs=15 NTStatus=c0000073

Owner=USER 12012 S-1-5-15-f407b588-8201be4d-bda18164-17f6

ALLOWED Flags=3 Mask=1301bf Rights=RWX-D-

Owner=GROUP domain=20admins.suncor:302 Domain Admins.SUNCOR:S-1-5-15-7417855-188b389a-30f211cf-200

ALLOWED Flags=19 Mask=1f01ff Rights=RWXPDO

Owner=GROUP domain=20admins.suncor:302 Domain Admins.SUNCOR:S-1-5-7417855-188b389a-30f211cf-200

ALLOWED Flags=12 Mask=1f01ff Rights=RWXPDO

No SACL

1033850863: ADMIN: 4: Command succeeded: acl if=fsn0 dump=/personal/its/DParker

## **CELLERRA ACL FLAGS & ACCESS RIGHTS:**

**Note:** Use ACLDump to output Unix, NT, and ACL information

-->UNIX section describes user and group content (name:number), plus Unix access right

-->NT section describes Owner and Primary group of file

    Owner of file (Unix name: uid Nt name:Sid)

    Primary group (Unix name: gid Nt group name:Sid)

-->ACL section contain the following values for each ACE:

    Owner type (USER/GROUP) Unix name:uid/gid NT name:Sid DENIED/GRANT Flags Mask and Rights

### **Flag Values:**

|                                  |                                                                    |
|----------------------------------|--------------------------------------------------------------------|
| #define EFFECTIVE_ACE            | (0x0) // effective only                                            |
| #define OBJECT_INHERIT_ACE       | (0x1) // ACE used on File creation                                 |
| #define CONTAINER_INHERIT_ACE    | (0x2) // ACE used on Dir creation                                  |
| #define NO_PROPAGATE_INHERIT_ACE | (0x4) // ACE not propagated                                        |
| #define INHERIT_ONLY_ACE         | (0x8) // ACE used only for inheritance                             |
| #define INHERITED_ACE            | (0x10) // ACE inherited cannot be overwritten without spec. action |

### **Mask Values:**

|                                   |                                  |
|-----------------------------------|----------------------------------|
| #define FILE_READ_DATA            | 0x00000001                       |
| #define FILE_LIST_DIRECTORY       | 0x00000001                       |
| #define FILE_WRITE_DATA           | 0x00000002                       |
| #define FILE_ADD_FILE             | 0x00000002                       |
| #define FILE_APPEND_DATA          | 0x00000004                       |
| #define FILE_ADD_SUBDIRECTORY     | 0x00000004                       |
| #define FILE_CREATE_PIPE_INSTANCE | 0x00000004                       |
| #define FILE_READ_EA              | 0x00000008                       |
| #define FILE_WRITE_EA             | 0x00000010                       |
| #define FILE_EXECUTE              | 0x00000020                       |
| #define FILE_TRAVERSE             | 0x00000020                       |
| #define FILE_DELETE_CHILD         | 0x00000040                       |
| #define FILE_READ_ATTRIBUTES      | 0x00000080                       |
| #define FILE_WRITE_ATTRIBUTES     | 0x00000100                       |
| #define DELETE                    | 0x00010000L                      |
| #define READ_CONTROL              | 0x00020000L                      |
| #define WRITE_DAC                 | 0x00040000L // write/modify ACEs |
| #define WRITE_OWNER               | 0x00080000L                      |
| #define SYNCHRONIZE               | 0x00100000L // not used          |
| #define SECURITY_SYSTEM           | 0x01000000L // Only in Audit ACE |
| #define GENERIC_READ              | 0x80000000L                      |
| #define GENERIC_WRITE             | 0x40000000L                      |
| #define GENERIC_EXECUTE           | 0x20000000L                      |
| #define GENERIC_ALL               | 0x10000000L                      |

### **Generic Masks:**

Generic masks are stored in an ACE only if they are in an INHERIT\_ONLY ACE

### **Generic Masks Mapped On:**

```
#define GENERIC_ALL_MAP (FILE_READ_DATA | FILE_WRITE_DATA | FILE_APPEND_DATA | FILE_READ_EA |
FILE_WRITE_EA | FILE_EXECUTE | FILE_DELETE_CHILD | FILE_READ_ATTRIBUTES | FILE_WRITE_ATTRIBUTES
| DELETE | READ_CONTROL | WRITE_DAC | WRITE_OWNER | SYNCHRONIZE)
#define GENERIC_READ_MAP (FILE_READ_DATA | FILE_READ_EA | FILE_READ_ATTRIBUTES | READ_CONTROL |
SYNCHRONIZE)
#define GENERIC_WRITE_MAP (FILE_WRITE_DATA | FILE_APPEND_DATA | FILE_WRITE_EA |
FILE_WRITE_ATTRIBUTES | READ_CONTROL | SYNCHRONIZE)
#define GENERIC_EXECUTE_MAP (FILE_EXECUTE | FILE_READ_ATTRIBUTES | READ_CONTROL | SYNCHRONIZE)
```

### **Associated rights:**

R -> read access

W -> write access

X -> execute or traverse dir

P -> change permission

D -> delete

O -> take ownership

## **RESETSACL COMMAND:**

**\$server\_config server\_2 "cifs update /fs1/homedir resetsacl"**

**Note:** ACLs have two fundamental components: DACL & SACL [Directory & System Access Control Lists]

Certain situations where SACL may become corrupt if filesystems are being shared across multiple netbios names and ownership is taken by a User from a different netbios name. Well-known SIDs are replaced with larger User SID and ACL offsets can become distorted, leading to SACL corruption. Also, due to large offsets, datamover may run out of memory and panic. This condition is fixed in NAS 4.2.14.402/403 and 4.2.16.0, 5.0.16.0, 5.1.13.0 and higher. Resetsacl strips out auditing information, if set, and is harmless to run.

## **HOW TO VERIFY UNIX FILE NAMES & WINDOWS 8.3 NAMES:**

**Note:** The following command will allow you to inspect UNIX names v. Windows 8.3 names—may be useful when determining if a 0 byte file should actually be a folder—confirms whether there has been shadow file corruption

**\$server\_config server\_9 -v "shadow readdir \volta\f3u2\Victor"**

```
1041202451: CFS: 4: name: ADMIN ADMIN <DIR>
1041202451: CFS: 4: name: backup_error_ojv.txt BACKUP~1.TXT
1041202451: CFS: 4: name: tseprofile_ojv.txt TSEPRO~1.TXT
1041202451: CFS: 4: name: USERS USERS <DIR>
1041202451: CFS: 4: name: OJVTSEPROFILES OJVTSE~1 <DIR>
1041202451: CFS: 4: name: filesystem_ojv.txt FILESY~1.TXT
1041202451: CFS: 4: name: profile_ojv.txt PROFIL~1.TXT
```

**Note:** In some cases, the command will not run because of spaces or other peculiarities in the path—when all else fails, try to determine DOS 8.3 name for the next directory in the path and substitute that name:

**# .server\_config server\_7 -v "shadow readdir \data\userdata\SFISHE3\Sandy's files\Action Plans"**

```
1131570657: CFS: 3: shadow fix: getAlternateName failed: NotFound
```

**# .server\_config server\_7 -v "shadow readdir \data\userdata\SFISHE3\SANDY'~1\ACTION~1"**

## **HOW TO FIX SHADOW FILE FOR SPECIFIC FOLDERS:** [Use instead of CIFS UPDATE]

**\$server\_config server\_9 -v "shadow fix \volta\f3u2\Victor"**

**\$server\_config server\_9 -v "shadow update \ora\ImageArchives"** [Yet another version of update]

**Note:** NAS 5.1.20.401 no longer seems to log ordinary SHADOW messages. Turn on Shadow debug prior to running Cifs Update in order to see logging events: “**logsys set severity SHADOW=LOG\_DBG3**”

## **SHADOW FILE PROBLEM ON ROOTFS OF CIFS SERVER:**

### **RUNNING READDIR ON ROOTFS REQUIRES \\:**

**# .server\_config server\_3 -v "shadow readdir \\"** [Note the double slash required to run command]

```
1063307557: CFS: 4: modifyCount: 0x25
1063307557: CFS: 4: name: application.evt APPLIC~1.EVT
1063307557: CFS: 4: Found 1 entries
1063307557: ADMIN: 4: Command succeeded: shadow readdir \
```

### **RUNNING SHADOW FIX ON ROOTFS:**

**# .server\_config server\_3 -v "shadow fix \\"**

```
1063307726: ADMIN: 4: Command succeeded: shadow fix \ [zeroes out Shadow file & is rebuilt on next CIFS client access]
```

### **RESULTS:**

**# .server\_config server\_3 -v "shadow readdir \\"**

```
1063307750: CFS: 4: modifyCount: 0x27
1063307750: CFS: 4: name: . . <DIR>
1063307750: CFS: 4: name: .. .. <DIR>
1063307750: CFS: 4: name: Data DATA <DIR>
1063307750: CFS: 4: name: .etc .ETC <DIR>
1063307750: CFS: 4: name: BusDev BUSDEV <DIR>
1063307750: CFS: 4: name: HR HR <DIR>
1063307750: CFS: 4: name: temp3mc TEMP3MC
1063307750: CFS: 4: name: Scratch SCRATCH <DIR>
1063307750: CFS: 4: name: application.evt APPLIC~1.EVT
```

1063307750: CFS: 4: name: lost+found LOST\_F~1 <DIR>  
1063307750: CFS: 4: name: system.evt SYSTEM.EVT  
1063307750: CFS: 4: name: tempemc TEMPEMC <DIR>  
1063307750: CFS: 4: name: SpaceProg SPACEP~1 <DIR>  
1063307750: CFS: 4: name: security.evt SECURITY.EVT  
1063307750: CFS: 4: name: .etc\_common ~1.ETC <DIR>  
1063307750: CFS: 4: name: Operations OPERAT~1 <DIR>  
1063307750: ADMIN: 4: Command succeeded: shadow readdir \  
**CAUSE:** Rootfs was previously filled up, damaging the Shadow file!

### **SHADOW SHOW COMMAND:**

**# .server\_config server\_2 -v "shadow show \acl"**

**Note:** Creates a data file called ‘shadow’ in the path specified—only used by developers

### **GETTING CIFS FILE ATTRIBUTES:**

**\$ .server\_config server\_2 -v "shadow getattr \test\john.xls"**

1076189006: CFS: 4: j: archive

1076189006: ADMIN: 4: Command succeeded: shadow getattr \test\john.xls

**Note:** Command can show if files are RO, hidden, or archive files.

### **MISCELLANEOUS ACL RESET OF DM ROOT DIRECTORIES, FILES, PROCESSES:**

.server\_config server\_x "acl reset=/.etc/PIPE"  
.server\_config server\_x "acl reset=/.etc/PIPE/samr"  
.server\_config server\_x "acl reset=/.etc/PIPE/lsarpc"  
.server\_config server\_x "acl reset=/.etc/PIPE/srvsvc"  
.server\_config server\_x "acl reset=/.etc/PIPE/svcctl"  
.server\_config server\_x "acl reset=/.etc/PIPE/eventlog"  
.server\_config server\_x "acl reset=/.etc/PIPE/winreg"  
.server\_config server\_x "acl reset=/.etc/PIPE/wkssvc"  
.server\_config server\_x "acl reset=/.etc/PIPE/netlogon"  
.server\_config server\_x "acl reset=/.etc/PIPE/celerra"

### **ACL DUMP OF DM HIDDEN C\$ SHARE:**

**# .server\_config server\_2 -v "acl if=fsn0 dump=/.etc/\$C"**

===== UNIX =====  
USER 0 GROUP 1 mode=rwrxr-xr-x  
===== NT =====  
Owner=USER 0 UNIX UID=0x0 ".:S-1-5-12-1-0  
Group=GROUP 1 UNIX GID=0x1 ".:S-1-5-12-2-1  
Owner=ALL Everyone.:S-1-1-0  
ALLOWED Flags=0 Mask=1200a9 Rights=R-X---  
Owner=GROUP 1 UNIX GID=0x1 ".:S-1-5-12-2-1  
ALLOWED Flags=0 Mask=1200a9 Rights=R-X---  
Owner=USER 0 UNIX UID=0x0 ".:S-1-5-12-1-0  
ALLOWED Flags=0 Mask=1301bf Rights=RWX-D-  
No SACL

### **ACL DUMP OF DM HIDDEN PIPE SHARE:**

**\$ .server\_config server\_2 -v "acl if=fsn2 dump=/.etc/\$PIPE"**

===== UNIX =====  
USER 0 GROUP 1 mode=rwrxr-xr-x  
===== NT =====  
Owner=USER 0 UNIX UID=0x0 ".:S-1-5-12-1-0  
Group=GROUP 1 UNIX GID=0x1 ".:S-1-5-12-2-1  
Owner=ALL Everyone.:S-1-1-0  
ALLOWED Flags=0 Mask=1200a9 Rights=R-X---  
Owner=GROUP 1 UNIX GID=0x1 ".:S-1-5-12-2-1  
ALLOWED Flags=0 Mask=1200a9 Rights=R-X---  
Owner=USER 0 UNIX UID=0x0 ".:S-1-5-12-1-0  
ALLOWED Flags=0 Mask=1301bf Rights=RWX-D-  
No SACL

### **ACL DUMP OF DM ROOT FILESYSTEM:**

**\$ .server\_config server\_5 -v "shadow readdir \\\"**

```
1066684472: CFS: 4: modifyCount: 0xe
1066684472: CFS: 4: name: .etc .ETC <DIR>
1066684472: CFS: 4: name: lost+found LOST_F~1 <DIR>
1066684472: CFS: 4: name: fs28 FS28 <DIR>
1066684472: CFS: 4: name: system.evt SYSTEM.EVT
1066684472: CFS: 4: name: security.evt SECURITY.EVT
1066684472: CFS: 4: name: disc-mnt DISC-MNT <DIR>
1066684472: CFS: 4: name: fs21 FS21 <DIR>
1066684472: CFS: 4: name: fs40 FS40 <DIR>
1066684472: CFS: 4: name: .etc_common ~1.ETC <DIR>
```

## **MOUNTING/UNMOUNTING FILE SYSTEMS--CD-ROM's, Floppies, etc:**

**Server\_mount Command:** Filesystems, Mountpoints, type, Status, access rights, multiprotocols, accesspolicy

-perm for Permanent: remounts at bootup [Listed in DataMover's /etc/dfs/dfstab or /etc/dfs/sharetab files]

-a -o {rolrw; primary=; nonotify} disable CIFS notify option, which is on by default; nooplock—disable cifs opportunistic locks, which are on by default; accesspolicy=nolocklwlocklrwlock } -F forces RO to be RW for Timefinder/FS BCV's

**Note:** NFS & CIFS File Systems will unmount “temporarily” by default unless you use the –p option

## **CELLERRA FILE SYSTEM MOUNT/UMOUNT RULES:**

1. You can “Unmount” a File System by specifying either the ‘File System Name’ or its ‘Mountpoint.’
2. But, when you “Mount” a File System, must specify ***both*** ‘FS name’ and ‘Mountpoint.’
3. By default, File Systems are *soft* ‘unmounted’ <unmounted> if you do not specify –p for permanent!!
4. By default, File Systems are *hard* or permanently ‘mounted’, even if you do not specify the –p switch!
5. Most known mount options, such as nolock, accesspolicies, etc., are effective immediately upon issuing the command and do not require umount and remount of file system.
6. You cannot mount a file system to another data mover if it’s already mounted “RW”. You can only mount a file system to multiple data movers if the file system is mounted “RO” for all file systems.

**HARD MOUNTS DEFINED:** Permanent mounts that look like a ‘local’ file system or directory from User’s perspective

Hard mount becomes unavailable, Client will retry, eventually locking up Client machine processes

Hard mounting is the default method

**SOFT MOUNTS DEFINED:** Non-permanent mounts

Client will retry then fail, preventing local machine from locking up

Since this type of mount has “data loss” potential, recommended usage is only for READ-ONLY applications

### **Mounting All File Systems in Table:**

**\$server\_mount server\_4 -all**

### **Mounting Individual File Systems:**

**\$server\_mount server\_4 fs1 /fs1** [Default=permanent, RW]

### **Mounting FS with AccessPolicies:**

**\$server\_mount server\_4 -p -o accesspolicy=secure,rwlock ufs1 /ufs1**

**\$server\_mount server\_4 -p -o accesspolicy=NT,wlock fs\_01 /fs\_01**

**Note:** Accesspolicies are now Uppercase sensitive with NAS 5.5

### **Mounting FS with Virus Checking Scanning Disabled:**

**\$server\_mount server\_4 -o noscan fs34 /fs34**

### **Mounting SNAPSURE File System:**

**\$server\_mount server\_4 -p -o ro snaps /mntpoint**

### **Unmounting All File Systems:**

**\$server\_umount server\_4 -a -perm** [without –perm results in temporary unmount]

**\$server\_umount server\_3 -p /fs01** [Mountpoint for fs01]

**\$server\_umount server\_3 fs01** [Soft or Temp unmount; \$server\_mount server\_3 show this FS as <unmounted> but original Accesspolicies and other attributes are retained, such as Locking, etc.]

## **VERIFYING MOUNTED FILE SYSTEMS:**

**# .server\_config server\_2 -v "file mountdisplay"**

server\_2 : commands processed: 1

Current Mounted File Systems are:

ckpt ro /mnt10\_ckpt10 151=45 ro

ckpt ro /mnt10\_ckpt9 151=44 ro

ckpt ro /mnt10\_ckpt8 151=43 ro

```
ckpt ro /mnt10_ckpt7 151=42 ro
ckpt ro /mnt10_ckpt6 151=41 ro
ckpt ro /mnt10_ckpt5 151=34 ro
ckpt ro /mnt10_ckpt4 151=33 ro
ckpt ro /mnt10_ckpt3 151=32 ro
ckpt ro /mnt10_ckpt2 151=31 ro
(NULL)
ckpt ro /mnt10_ckpt1 151=29 ro
uxfs rw /mnt10 143=25 rw,accesspolicy=NATIVE
uxfs rw /gary 127=18 rw
uxfs ro /.etc_common 67=16 ro
uxfs rw / 39=2 rw
Total number of file systems is: 15
```

### **VOLUME INFO CACHING:**

NAS 5.3 caches volume structure to speed mount operations. An issue can occur when deleting file systems, however, in that the cache is not deleted, and can create problems when creating new file systems that reuse a volume id number previously ‘deleted’. NAS 5.4 deletes cached information for deleted file systems.

### **MOUNTING BCV “RO” FILESYSTEM “RW”:**

```
$server_mountpoint server_5 -c /fs2_snap1
$server_mount server_5 -F -p -o rw fs2_snap1 /fs2_snap1
```

### **LOCATING USERS/PROCESSES LINUX RED HAT 7.2:**

#### **/usr/sbin/safe\_finger**

| Login    | Name | Tty   | Idle | Login Time                  | Office | Office Phone |
|----------|------|-------|------|-----------------------------|--------|--------------|
| nasadmin |      | pts/1 |      | Mar 3 10:57 (10.240.16.84)  |        |              |
| nasadmin |      | pts/2 |      | Mar 3 11:19 (10.240.16.128) |        |              |
| nasadmin |      | pts/0 | 3d   | Feb 27 15:29 (172.24.80.35) |        |              |

**\$ /sbin/fuser -m /mnt/fs50**

/mnt/fs50: 6899c

**\$ /sbin/fuser -u** //:5975r(nasadmin) 6899r(nasadmin) 9521r(nasadmin) 16336r(nasadmin)

-v verbose -u uids -m mounted fs -k kill processes

accessing file -a display unused files

**#/sbin/fuser -k** [Kills processes from Users on a filesystem]

#### **Using the Fuser Command:**

**# fuser -u /tmp** [Lists out User processes by Username and Process ID for the Directory or File System specified]

**# fuser -f filename** [Lists out users of a particular file]

### **UNMOUNTING /NAS FILE SYSTEM IF IN USE:** ‘Resource Busy’ error when trying to unmount fs

1. Identify processes that are locking file system in use: **#/sbin/fuser -m /nas**

/nas: 1568 1575 1576 1577 1578 1579 1580 1581 1582 2220

c 2299c 2300c 2362 2362c 7719 7719c 7719e 7916c

2. Kill processes: #kill -9 1568 (.etc)

3. Unmount /nas: #umount /nas

### **MOUNTING CS1 FROM CS0 FOR COPYING FILES BACK & FORTH:**

**Note:** Only prerequisite is that you know what the path is for the root “/” partition of the control station that you want to mount. sdd3 & sdc3 are normal partition names for CS1 & CS0, respectively for Red Hat 7.2

#mount /dev/sdd3 /mnt [Mounts root partition of CS1]

#mount /dev/sdc3 /mnt [Mounts root partition of CS0]

### **The Unix Automounter Service:**

Linux uses “autofs”; Solaris uses “automounter” [/etc/auto\_master]; others use “amd”

#### **Purpose of Automounter?**

--Makes exported FileSystems available to Remote Hosts

- Use Wild-cards to mount remote directories to like-named mountpoints on local hosts
- Increases network reliability through use of multiple servers exporting large numbers of FS and Directories to Users
- Example: Common Home Directory server for Users to Login
- Defined by use of “map” files [Convert Sun automounter ‘map’ to “amd” map file using perlscript; automount2amd.pl script]
- Must have ‘portmapper’ service running

**Linux Automounter:** “autofs” is recommended and uses the “commandline” to configure

**\$ ps -ef |grep auto**

```
root 955 1 0 Nov17 ? 00:00:00 /usr/sbin/automount --timeout 1  
root 973 1 0 Nov17 ? 00:00:00 /usr/sbin/automount --timeout 1
```

**#/etc/rc.d/init.d/autofs start | stop**

**#/sbin/service autofs stop | start**

### **AUTOMOUNTER ISSUE NAS 5.3:**

**# ls -al /etc/auto.\***

```
-rwxrwxr-x 1 root root 117 Oct 14 17:45 /etc/auto.master → Problem is that access rights are set incorrectly  
-rwxrwxr-x 1 root root 575 Jul 14 14:17 /etc/auto.misc  
-rwxrwxr-x 1 root root 438 Oct 14 17:45 /etc/auto.nas  
# ps axlww | egrep auto  
040 0 31804 1 9 0 1500 644 pipe_w S pts/2 0:00 /usr/sbin/automount --timeout 1 /nasmcd/rootfs program /etc/auto.nas  
ro,,soft,intr,nosuid,noac  
040 0 31822 1 8 0 1496 640 pipe_w S pts/2 0:00 /usr/sbin/automount --timeout 1 /nasmcd/quota program /etc/auto.nas  
rw,,soft,intr,nosuid,noac
```

#### **SOLUTION:**

1. Stop automounter

**# /sbin/service autofs stop**

2. Change access rights to 644

**# chmod 644 /etc/auto.\***

**# ls -al /etc/auto.\***

```
-rw-r--r-- 1 root root 117 Oct 14 17:45 /etc/auto.master  
-rw-r--r-- 1 root root 575 Jul 14 14:17 /etc/auto.misc  
-rw-r--r-- 1 root root 438 Oct 14 17:45 /etc/auto.nas
```

3. Restart the automounter—process should be have “file” mode set, not “program” mode

**# /sbin/service autofs start**

**# ps axlww | egrep auto**

```
040 0 3466 1 8 0 1496 636 pipe_w S pts/2 0:00 /usr/sbin/automount --timeout 1 /nasmcd/rootfs file /etc/auto.nas  
ro,,soft,intr,nosuid,noac  
040 0 3495 1 9 0 1500 640 pipe_w S pts/2 0:00 /usr/sbin/automount --timeout 1 /nasmcd/quota file /etc/auto.nas  
rw,,soft,intr,nosuid,noac
```

4. Test automounter

**Note:** One symptom of this condition is that server\_export file becomes 0 bytes because it needs automounter to pull the exportdb from the Data Movers. Run #server\_export to verify fix.

### **CREATING AN AUTOMOUNT MAP FILE ON CELERRA:**

- Step 1. Create File: \$nas\_automountmap -create -out newmap

**Note:** This creates a map file of all exports on the Celerra & outputs lines similar to following:

**NFS8 -rw,intr,nosuid 192.10.2.25,192.10.2.24:/NFS8**

**legato -rw,intr,suid 192.10.2.23,192.10.2.22:/legato**

- Step 2. Use Text Editor to modify file

- Step 3. Merging an old mount and new mount map file: \$nas\_automountmap -create -in oldmap -out -newmap

- Step 4. Verifying automountmap list: \$nas\_automountmap -list\_conflict -in oldmap -out newmap

### **CELERRA AUTOMOUNTER:**

**AutoMounter Files:** /etc/auto.misc | auto.master & /nas/sys directories

**\$/etc/ls auto\*,file auto\***

auto.master auto.misc auto.nas

auto.master: ASCII text

auto.misc: ASCII English text

auto.nas: ASCII text

**Configuring Celerra With Auto-Mounter Entries:**

```
$ vi auto.misc
# $Id: auto.misc,v 1.2 1997/10/06 21:52:04 hpa Exp $
# This is an automounter map and it has the following format
# key [ -mount-options-separated-by-comma ] location
# Details may be found in the autofs(5) manpage
cd -fstype=iso9660,ro,nosuid,nodev :/dev/cdrom
# the following entries are samples to pique your imagination
#linux -ro,soft,intr ftp.example.org:/pub/linux
#boot -fstype=ext2 :/dev/hda1
#floppy -fstype=auto :/dev/fd0
#floppy -fstype=ext2 :/dev/fd0
#e2floppy -fstype=ext2 :/dev/fd0
#jaz -fstype=ext2 :/dev/sdc1
#removable -fstype=ext2 :/dev/hdd
```

**CELERRA SOFT MOUNTS:**

```
$ more auto.master
# timeout set to 2/10 of a second.
/nasmd /rootfs /etc/auto.nas ro,-t 1,soft,intr,nosuid,timeo=2,retrans=2,actimeo=1,nfsvers=3,noac,sync
/nasmd /quota /etc/auto.nas rw,soft,intr,nosuid,timeo=2,retrans=2
```

**Mount Options:**

-ro [Read Only] -blocksize=1024 -wsize=4096 [write buffer size] -intr [allow keybd interrupts to kill process]  
 -soft [Server fails to respond, return error after timeout period **-timeo=value** expires & do not retry]  
 -hard [Default Option: Client retries until Server responds—This can hang your network if not careful!!!]  
 -nouser [only root can mount FS] -noexec [do not execute files from this FS]  
 -sync [Use synchronous FS I/O] -timeout=60 [timeout after 60 seconds, for “softmounts”]

**NFS Mount Options:**

nosuid → Use this option for NFS file systems to prevent a user from running programs based on credentials of owner of file  
 noexec → forbids execution of any files from remote users

**Note:** #mount -timeo is used to change default RPC timeout values--may need to increase for RPC timeouts

**Sample Mount Problem:** Sometimes, df, sync, ls commands hang

**Cause:** Server using Hard Mounts and is down

**Solution:** Use soft mounts, timeout values, or restart NFS and Automount daemon

**Mounting All FS in Linux:** #mount -a

**Viewing All FS:** #showmount -e or #df

**Map File Built Using Syntax of “fstab” file:** /etc/fstab /etc/auto.master

| File System | Mountpoint | Options             |
|-------------|------------|---------------------|
| /dev/sdc1   | /cdrom     | iso9660,defaults,ro |

**MOUNTING DM FROM CONTROL STATION WITH PERSISTENT MOUNT:**

#mount -r server\_5:/ /dm

**Following Entry is Created in Dynamic Mount Table--/etc/mtab:**

server\_5:/dm nfs ro,addr=192.168.1.5 0 0

**Copy Above Entry into /etc/fstab:**

server\_5:/dm nfs ro,addr=192.168.1.5 0 0

**SYMPTOMS THAT AUTOMOUNTER IS NOT WORKING CORRECTLY:**

1. cannot cd to rootfs of Server
2. running server\_export lists only NFS exports, not shares
3. # nas\_quotas -r -fs stores report file not ready yet, retrying...done

4. Automounter errors seen in /var/log/messages or osmlog

Jun 11 07:48:37 go001cfs000 automount[880]: attempting to mount entry /nasmd/rootfs/slot\_4

Jun 11 07:48:37 go001cfs000 automount[6849]: >> nfs bindresvport: Address already in use

Jun 11 07:48:37 go001cfs000 automount[6849]: mount(nfs): nfs: mount failure on/nasmd/rootfs/slot\_4]

5. Output of netstat shows large number of ports in use by usrmount\_svc:

# **netstat -alp |grep -i usrmapper\_svc lwc -l** → Usrmapper Service was consuming 1010 ports

6. Automounter service still running:

```
$ ps -ef |grep auto
root    923    1 0 Jun06 ?    00:00:00 /usr/sbin/automount --timeout 1
root    941    1 0 Jun06 ?    00:00:00 /usr/sbin/automount /nasmcd/quot
```

### **SHOWING NFS EXPORT LIST ON DM:**

# **/usr/sbin/showmount -e server\_3 | 192.10.0.103**

Export list for server\_3:

```
/ 192.1.2.101,192.1.1.101,192.1.2.100,192.1.1.100
/root_ojv 192.1.1.100,192.1.1.100,192.1.1.100,192.1.1.100,192.1.1.100,192.1.1.100
/root_emc 192.1.1.100,192.1.1.100,192.1.1.100,192.1.1.100
/ojv_sa 192.1.1.100,192.1.1.100
```

**EXAMPLE:**

# **/usr/sbin/showmount -e 138.1.164.129**

Export list for 138.1.164.129:

```
/ 192.1.2.101,192.1.1.101,192.1.2.100,192.1.1.100
/nas019 rgmdebkp3-b,rgmdbs1-b
/nas022 rgmdebkp3-b,rgmdbs1-b
/nas021 rgmdebkp3-b,rgmdbs3-b
/nas056 rgmlnx03-b,rgmlnx02-b,rgmlnx01-b
```

### **SHOWING NFS HOSTS MOUNTED ON DM:**

\$ **/usr/sbin/showmount -a server\_3 | 192.10.0.103**

All mount points on server\_3:

```
192.1.1.100:/ojv_sa
192.1.1.100:/root_emc
192.1.1.100:/root_ojv
```

**EXAMPLE:**

# **/usr/sbin/showmount -a 138.1.164.129**

All mount points on 138.1.164.129:

```
rgmdbs1-b:/nas019
rgmdbs1-b:/nas022
rgmdbs3-b:/nas021
rgmlnx01-b:/nas056
rgmlnx02-b:/nas056
rgmlnx03-b:/nas056
```

**Other Files:** /etc/mtab [Shows active partitions mounted]

/etc/exports [Exported partitions]

#### **Indirect v. Direct Map Files:**

Indirect Map used for multiple Users with Relative Mountpoints, as in /home directories: user\_1 mach\_1:/export/home/user\_1

Direct Map used for direct mount statement: /home/research -rw filbert:/home/research

**Note:** Linux “autofs” does not support Direct Maps!

### **USE OF HARD vs. SOFT MOUNTS FOR NFS:**

Standard practice dictates that NFS filesystems be hard-mounted whenever conducting writes—otherwise, there is great potential for lost data if the Server crashes.

**HARD MOUNTS:** “NFS Server Not Responding” errors on screen or /var/messages file are indicators that Hard-mounts are used

**SOFT MOUNTS:** Soft-mounts are typically used for Read-Only mounts—a key indicator of a soft-mounted filesystem would be an NFS “RPC Timeout”—client retries a number of times before giving up.

**TroubleShooting Automounter:** /var/log/messages #df [will hang on client that is no longer active]

### **Mounting File Systems:**

**Quick Way to Remount or Temporarily Unmount all File Systems on a Server** [Provided that FS are in Mount Table]:

```
$server_mount server_3 -a          $server_umount server_3 -a
```

**Mounting the DataMover Root FileSystem to the Control Station:** #mount -F nfs server\_3:/ /mntptn

Or #mount -F nfs 172.16.21.183:/root\_fs\_3 /mnt

#cd /nas/rootfs/slot\_2 [another way to access root filesystem]

**Accessing DM Root From Windows:** \dm43\_ana0c\$

**MOUNTING DATAMOVER “RW” FROM CONTROL STATION:** “*Shortcut Method*”

#cd /nasmcd/quota/slot\_2/etc

**Permissions Problem When Trying to Mount to DataMover:**

If FileSystem is exported as “anon=0”, means that you are granted “user” Root access, but only “group” Other.

So, if the “mountpoint” was created by “root”, it has “user & group” as root and the “group” Other will be denied access.

Therefore, may need to “CHMOD” the mountpoint to which you want to remotely connect to:

#chmod 777 /mnt1

#mount -F nfs 172.19.32.233:/ /mnt1

**MOUNTING LINUX CONTROL STATION TO CD-ROM:**

#mount -ro /dev/cdrom /mnt/cdrom

**MOUNTING FLOPPY DRIVE WITH LINUX:**

#mount /dev/fd0 /mnt/floppy

**Mounting A Floppy Device to Sun:**

**Note:** If Sun has Volume Management installed, insert diskette and type \$volcheck at command line

Step 1. Insert floppy and from root, create mountpoint: #mkdir /floppy

Step 2. #mount -F pcfs /dev/diskette /pcfs [pcfs=PC DOS filesystem type]

**Note:** Purpose of “-F” switch is to signify a different File System format

**Mounting All FS for Linux:** #mount -a

**Mounting All FS for Solaris:** #mountall

**ARGUMENTS FOR MOUNTING NFS FILE SYSTEMS USING TCP vs. UDP:**

--Use NFS v3 over TCP, not UDP: If network is susceptible to dropped packets, UDP requires retransmission of all fragments in a datagram, whereas TCP would require retransmission of only a single packet. NFS v2 can only transfer in 8k datagrams.

Asynchronous writes using unstable flag and commit calls, and GETATTR calls are improvements over NFS v2.

--NFS/RPC timeouts for retransmissions take time, then requires complete retransmission

--NFS Streaming supported by V3 not V2

--TCP protocol allows for Congestion Window flow control, Receiver flow control based on Receive Window Size, and Congestion Avoidance and Slow Start Algorithms for recovery after packet loss.

--NFS requires retransmission at layer 5 RPC via timeout mechanism whereas layer 4 TCP uses Fast Retransmit and ACK

--In addition, NFS v3 over TCP allows for better handling of large data transfer sizes, say 32k, where UDP may overflow switches with large packet transfers

--TCP helps prevent buffer overflows in switches and pause frame flow control can be implemented

**EXPORTING/UNEXPORTING FILE SYSTEMS:**

**CELLERRA EXPORT COMMAND:**

\$ server\_export ALL [Displays exported and shared File Systems]

**Note:** Celerra shares can be created from the CLI or from the MMC console from a Windows system

→Computer Management>rightclick name>Connect to another computer>Celerra\_computername>System Tools>Shared

Folders>Shares>rightclick>New Share>Next>Folder Path: c:\m1 (enter upper mountpoint name or path for the share)>Share name: enter a name for the new Share>Select permission setting>Finish

**SCRIPT TO CREATE CIFS SHARES:**

#!/bin/bash

# script to create exports

#

NAS\_DB=/nas

for i in \$(seq 1 100); do server\_export server\_2 -P cifs -name fs3\_\$i -o netbios=dbms /fs3; sleep 1; done

for i in \$(seq 1 100); do server\_export server\_2 -P cifs -name fs4\_\$i -o netbios=dbms /fs4; sleep 1; done

for i in \$(seq 1 100); do server\_export server\_2 -P cifs -name fs5\_\$i -o netbios=dbms /fs5; sleep 1; done

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
for i in \$(seq 1 100); do server\_export server\_2 -P cifs -name fs6\_\$i -o netbios=dbms /fs6; sleep 1; done

### **SCRIPT TO CREATE DOMAIN USER ENTRIES ON CELERRA:**

```
#!/bin/bash
# script to create Internal Usermapper and Secmap Domain User entries
#
NAS_DB=/nas
for i in $(seq 30008 30999); do server_cifssupport server_2 -secmap -create -name w2ku$i -domain w2k; sleep 1; done
```

### **NFS EXPORT OPTIONS/RULES:**

--Root Users (UID=0) are considered unknown by NFS Servers unless specifically listed in export options  
--anon=uid--means that requests from unknown Users should receive ‘effective’ uid of ‘nobody’ or 65534 by default  
--User uids are checked first on data mover’s local ./etc/hosts file, then ./etc/netgroup, then NIS, and finally DNS  
--access=client--provides mount access to specified clients. Clients can be Hostname, Netgroup, Subnet, IP Address  
--root=client--provides root privilege to client in form of hostname, netgroup, subnet, IP Address. Default NFS behavior is that no Hosts have Root access  
--rw=client--exports pathname as read-mostly for specified clients, and read-only to everyone else [hostname, netgroup, subnet, IP]  
--ro option--exports pathname for NFS clients as Read-Only  
--sec=sys—SecureNFS sys security mode ro, rw=, ro=, root=, access=, wlan=, anon=, webroot, public  
--sec=krb5—SecureNFS security mode ro, rw=, ro=, root=, access=

### **EXAMPLES:**

```
export “/fs1” access=192.168.10.2 rw=192.168.10.1 root=192.168.10.1
ro rw=hostname; anon=uid root=hostname access=client {./etc/hosts or ./etc/netgroup—NIS}
access=hostname /mntpoint access=IP address /mntpoint access=Netgroup /mntpoint
access by Subnet→$server_export server_2 -o access=193.1.21.0/255.255.255.0 /mntpoint
```

**Note:** All hosts that need to connect as RW should be in “rw=” list; all hosts that should be RO should be in “access=” list;  
Using the “ro=” option forces all hosts to be Read Only, despite use of “root=”, “rw=”, for other hosts in export line. RO exports the whole export path as RO. Pathnames must be enclosed with quotations if extra spaces are in the path.

**NFS Export Note: Options [-o]:** ro,rw [default=rw]; anon=uid; root=client; access=client; rw=client ['client' means hostname', 'netgroup', 'subnetIP/mask', 'IPaddress/mask'; separate listings with colon :] Default Export for NFS if no switches used is “No Hosts Root Access”. Also, if using any of the methods listed, should ensure that there are entries in one of the following locations, in order of resolution:

1. Data Mover ./etc/hosts database
2. Data Mover ./etc/netgroup db
3. NIS Server
4. DNS Server

### **LINUX CLIENT ISSUE WITH CELERRA:**

Celerra responds to a Client NFS Write request with ERR\_STALE [Stale Filehandle] but inserts extra data into the error message, causing Linux client to hang. See AR49124,49188, 47569 DNLC entries.

### **VIEWING NFS EXPORTS FROM GIVEN IP ADDRESS:**

**# showmount -e 172.19.32.9**

### **COMMON NFS EXPORT ERRORS:**

--2002-06-12 21:23:59: NFS: 4: mount\_lookup: No match in export list, fs not exported: host 5785040a, ruid 320519936, handle 12.1017776002  
--/mnt Stale NFS file handle

**Comment:** Errors are harmless, indicating that an NFS FileSystem has been unexported but clients still have mounted

**Exporting NFS File Systems when having Client-side Write Issues:** There are certain rare occurrences whereby a network latency issue, or dropped packets [as in a failing switch] that can change the ability of a File System to be written to by remote hosts, say in the instance that the remote hosts were defined for RW access and permissions.

**Exporting the File System by Hostname and Giving Root Access also:** [As opposed to anonymous nobody access, which is RO]

**\$ server\_export server\_2 -i -o rw=acomp7:acomp8:acomp9,root=acomp7:acomp8:acomp9 /bvfs2**

### **Exporting File System by IP Address:**

**\$ server\_export server\_2 -i -o rw=192.25.88.20:192.25.88.21 /bvfs2**

Important point is that the user 'nobody' [anonymous] is given RO access to File Systems, so that if not properly defined, a user may come across thinking he has RW access, when in reality, he is assigned the “nobody” UID for RO access.

### **Common Error:**

Export error: Neither parent nor self conflict

**Note:** Error occurs when path still exists but is not currently exported.

## **EXPORTING ROOT “/” READ-ONLY ON LINUX CONTROL STATION FOR SPECIFIC HOST:**

**Caution:** Use this procedure with extreme caution—usually done for backing up Control Station partitions.

Step 1. Verify that “nfs” services are running on Control Station:

```
#/etc/rc.d/init.d/nfs status
rpc.mountd (pid 7328) is running...
nfsd (pid 7340 7339 7338 7337 7336 7335 7334 7333) is running...
rpc.rquotad (pid 7323) is running...
```

Step 2. If NFS services are not running, use following command:

```
#/etc/rc.d/init.d/nfs start | stop
```

Step 3. Verify current exports on Control Station:

```
#/usr/sbin/exportfs
```

Step 4. Add export entry to export Control Station “root” / :

```
#vi /etc(exports
/ 192.10.2.3 (IP Address of Remote Client to be allowed access to CS0 root)
```

Step 5. Refresh Exports on Control Station:

```
#/usr/sbin/exportfs -av
```

Step 6. Create mountpoint on Remote Unix Host:

```
#mkdir /cs0_romount
```

Step 7. From Remote Host, mount Control Station “/” root export:

```
#mount 192.24.80.11:/ /cs0_romount [CS0's IP Address, “/” root filesystem, and local mountpoint]
```

## **EXPORTING NFS UDP/TCP:**

Servers generally export both protocols and clients choose a default or indicate ‘proto=tcp’ from remote mounting

**NFS EXPORT TABLE LIMITATIONS:** Only 256 export entries can be displayed, though more actual exports can be applied

## **NFS EXPORT BY IP ADDRESS FOR ROOT ACCESS:**

```
$server_export server_3 -p -o root=193.1.21.210/255.255.255.0 /g1ufs1
```

## **EXPORTING ROOT FILE SYSTEM OF DATAMOVER FOR EXTERNAL CS ACCESS:**

**Note:** Procedure is listed for reference purposes only and is not condoned or supported for customer environments! Be aware that exporting the data mover over the internal interfaces from the Control Station may not show all directories. Exporting over external IP addresses and then mounting from CS will be a true NFS mount.

**Scenario:** Exporting Root File System of Server\_2 to external IP Address of CS0

1. #vi /nas/server/slot\_5/export [Add following line to DM root file system to external IP address of CS0]

```
export "/" anon=0 access=192.10.0.17:192.168.1.100:192.168.2.100:192.168.1.101:192.168.2.101
```

**Note:** “/” represents root mountpoint exported from DM. Access=IP is External IP address of CS0.

2. Reboot Server to rebuild exports

3. Mount root filesystem Read-Only from CS0:

```
# mount -r server_5:/ /dm
```

4. Following entry is made to /etc/mtab file:

```
server_5:/ /dm nfs ro,addr=192.168.1.5 0 0
```

5. Copy entry from /etc/mtab to /etc/fstab file

```
server_5:/ /dm nfs ro,addr=192.168.1.5 0 0
```

**Note:** This ensures that CS remounts the file system after a reboot, whereas just having the entry in mtab will not. Be advised that Tech Support does not condone or support exporting the DM root filesystem to any external IP address. A full package NAS Upgrade will also replace any ‘modified’ or added IP addresses in the Export file of all Servers for the “/” file system export with the default Internal IP addresses of the Control Station only.

## **MOUNTING FILE SYSTEMS OF DATA MOVERS FROM CS0:**

1. Create mountpoints for file systems to mount on CS0 [e.g., #mkdir /s2/fs1, #mkdir /s3/fs3, etc.]

2. Mount over external network so as to keep NFS functionality intact: #mount 10.64.25.14:/fs1 /s2/fs1

**Comment:** Mounts “fs1” on Server\_2 over external IP Address from CS0

## **MOUNTING DM OVER INTERNALS USING NFS V2:**

```
# mount -o vers=2 server_4:/ /mnt
```

**Note:** Use this to see all directories on datamover whereas NFS v3 may not properly display all directories.

## **DUMPING EXPORTS ON DATAMOVER:**

**\$ .server\_config server\_2 -v "export"**

1058207628: NFS: 4: /pol anon=0

1058207628: NFS: 4: / anon=0 access=10.64.25.10:192.168.1.100:192.168.2.100:192.168.1.101:192.168.2.101

## **CIFS SHARE OPTIONS:**

**Note:** Default is to export CIFS Shares as Read-Write to all Clients

ro—exports Share as Read-Only for CIFS clients

rw=client—exported read-mostly for those clients specified, but Read-Only for all others

umask=mask—allows setting Unix permissions to Share; CIFS default umask=022 or 755

**Note:** Directory permissions in octal expressed from 777; File permissions in octal expressed from 666

002=775, which is complete access to Owner & Group, and Read access to Other [includes directory search]

022=755, which is complete access to Owner & Read/Execute access to Group & Other [includes directory search]

### **Example File or Directory set to 755, or umask=022:**

-rwxr-xr-x 1 root root 0 Mar 21 10:57 file3

drwxr-xr-x 2 root root 4096 Mar 21 10:58 subfolder

### **Example File or Directory set to 775, or umask=002:**

-rwxrwxr-x 1 root root 0 Mar 21 10:52 file2

drwxrwxr-x 2 root root 4096 Mar 21 10:58 subfolder

**# server\_export server\_2 -P cifs -n home4 -o umask=002 /fs1**

Octal 777 gives rwx to User, Group, Other

Octal 666 gives rw to User, Group, Other

user==default\_user group=default\_group [Used to define User & Group access for Share-level Security]

ropasswd=share\_passwd rwpasswd=share\_passwd [Used to define RO & RW passwords for Share-level Security]

maxusr=maxusr—Sets number of simultaneous user connections to a Share [default is 4 billion]

netbios=compname—if a specific netbios name is not specified, Share is visible to all NetBIOS names on Server by default

**Other Options:** -l -P {cifs/nfs} -n {cifs sharename} -a -a-u -i -p [exports permanent for reboots] -t -u -o

## **SPECIAL CHARACTERS ALLOWED IN SHARENAMES:**

The following may be used in share names \*^%\$#@!()\_-+=?.,[]{}|

The following characters should not be used: \ and /

## **CIFS SHARES AND ACL CREATION BEHAVIOR:**

→new filesystem contains no ACLs until first exported for CIFS, then gets Everyone FC by default

→If not exported, the filesystem can only be accessed from CIFS by an Admin mapped to UID of 0

## **COMPNAMES AND NETBIOS NAMES:**

Default NetBIOS name of computer is assigned from first 15 characters of comp\_name. Use the netbios= option when creating a compname to specify a netbios name that needs to be different from the comp\_name.

## **CIFS SHARENAMES:**

Limited to 256 characters with Unicode UTF-8, or 12 if using ASCII mode

Sharenames are case-sensitive and may contain spaces

netbios=compname—CIFS share created specifically for a single netbios name

**Note:** Windows 9x Clients enumerate Shares using MSRAP Share Enumerate Calls to the Server. NT/Windows 2000 Clients use SRVSVC Share Enumerate Calls to obtain Shares listing.

### **Common Win9x Client Error in Server Log:**

“Max buffer size of the client to small (2850) 135 shares of 169 returned to the client”

## **LISTING DATA MOVER SHARES:**

**\$ .server\_config server\_2 -v “share” “sharedb info” “sharedb asc”**

1047060338: SMB: 4: HOME /

1047060338: SMB: 4: comment="Home Service"

1047060338: SMB: 4: umask=022

1047060338: SMB: 4: maxusr=4294967295 inUse=2

1047060338: SMB: 4: C\$ /

1047060338: SMB: 4: comment="Root Service"

1047060338: SMB: 4: umask=022

1047060338: SMB: 4: maxusr=4294967295 inUse=2

1047060338: SMB: 4: CHECK\$ /

```
1047060338: SMB: 4: comment="Virus Checking Service"  
1047060338: SMB: 4: umask=022  
1047060338: SMB: 4: maxusr=4294967295 inUse=2  
-----abridged-----
```

## **DUMPING ACLS ON CIFS SHARES:**

```
# .server_config server_2 -v "acl if=eand-fsn0 share=shared1"
```

Share Shared1 (EAND-FS001)

===== UNIX =====

USER ftpuser-emc:0 GROUP 1 mode=rw-----

===== NT =====

Owner=USER ftpuser-emc:0 UNIX UID=0x0 'ftpuser-emc':S-1-5-12-1-0

Group=GROUP 1 UNIX GID=0x1 ".:S-1-5-12-2-1

Owner=ALL Everyone.:S-1-1-0

ALLOWED Flags=0 Mask=1f01ff Rights=RWXPD0

No SACL

```
$ .server_config server_2 -v "acl share=home" (alternative syntax without interface)
```

## **CHANGING UMASK VALUES ON CIFS SHARES WITH ‘SERVER\_EXPORT’:**

```
$server_export server_6 -P cifs -n umask27 -o umask=027 /dawn
```

\$server\_export server\_6

```
share "umask27" "/dawn" maxusr=4294967295 umask=27
```

Comment: Our default umask value for NT Shares is “umask=022”

## **UMASK CALCULATION INTO PERMISSIONS:**

|                             |     |
|-----------------------------|-----|
| 777 UGO value               | 777 |
| 022 = umask                 | 002 |
| 755 = resultant permissions | 775 |

## **CHANGING UMASK VALUES USING .SERVER COMMANDS:**

DEFAULT UMASK: cifs.share.default.umask 0x0134eb30 0x00000012 0x00000012

Note: 0x12 is Octal value 022, which is the default value used for shares when no other specific Umask is defined. A Umask of 022 means rw-r—r—for file creation and rwx-r-xr-x for directories.

Change Umask: \$ .server\_config server\_2 -v "param cifs share.default.umask=0x2"

Note: This does not change the ‘umask=22’ listed in the output of server\_export, but changes the effective umask to the 0x2 value.

## **DAMAGED EXPORT TABLE:** [/nas/server/slot\_x/export] can have the following effects:

- Inability to mount the DataMover's rootfs either Permanently or Temporarily
- Inability of the DataMover to Failover
- Can produce error message when running Server\_Export command

## **DUMPING DATA MOVER EXPORTS FROM MEMORY:**

```
# .server_config server_6 -v "dumpexportdb"
```

## **EXPORTING/UNEXPORTING CIFS & HIDDEN SHARES:**

```
$server_export server_2 -P cifs -o user=todd,group=administrators,rwpasswd=write,ropasswd=read -n share6 /mnt1
```

```
$server_export server_2 -P cifs -u -n content$ -p [unexporting—do not need to specify mountpoint path!!]
```

```
$server_export server_2 -P cifs -n content$ -p /content/content [exporting to subdirectory level]
```

## **UNEXPORTING FILESYSTEMS:**

NAS codes >than 4.2 no longer support temporary unexporting of either NFS or CIFS—despite switch, exports will be permanent

### **Permanently Unexporting ALL Exports:**

```
$server_export server_4 -u -a
```

Un-exporting NFS File Systems: \$server\_export server\_4 -p nfs -u -all [or /path for a single FS]

Un-exporting CIFS File Systems Permanently: \$server\_export server\_4 -P cifs -a -p -u

Un-exporting Single CIFS File System Permanently: \$server\_export server\_4 -P cifs -n cifsfs -p -u

Ignore Export Table: \$server\_export server\_4 -p -ignore

Note: The –ignore option will re-export all exports or the specific export and replace all previous export options used

## **VERIFYING PATH OR SHARE EXPORTS ON DATAMOVER:**

**\$server\_export server\_2 -l -P cifs -n share**

### **RECOVERING CIFS SHARES:**

**Note:** Starting with NAS 4.x and higher, the export.shares file was created on the Control Station. Under ordinary circumstances, this file is kept up-to-date with the shares database, which is now hosted on the data mover. In some situations, there may be a need to resync the CS database to the Server db, located in [/etc/shares/@import](#) | @Global | @system

→ Make backup copy of export.shares file and run the recovery command:

**\$server\_export server\_2 -r**

### **DNS AND SHARES ISSUE:**

**Situation:** Under certain circumstances, if deleting a sharename for a specific share on a DM, then re-exporting the same sharename to a different mountpoint and DM with a different IP address, Clients will connect and see wrong information underneath the “Share” that they are accustomed to working on.

#### **Resolution:**

→ Verify problem using \$.server\_config server\_2 -v “dns dump” and compare to \$server\_cifs and \$server\_export output

→ Unexport and re-export the “Share” to the correct mountpoint and NetBIOS Server

→ Flush Client System dns cache

→ Flush DM cache using \$server\_dns server\_2 flush

### **CASE EXAMPLE:**

\$ server\_export server\_2

server\_2 :

share "shares" "/shares" netbios=DIANAS01 maxusr=4294967295 umask=22

share "edocs" "/gissshare" netbios=DIAGIS03 maxusr=4294967295 umask=22

**Note:** “shares” was mistakenly re-exported for /gissshare –o netbios=DIAGIS03 before correcting, but left following problem and client access problems to ‘shares’

**\$ server\_cifs server\_2**

CIFS Server (Default) DIANAS01[DIA]

Full computer name=dianas01.dia.dnvr realm=DIA.DNVR

Comment='EMC-SNAS:T4.2.20.100'

if=fsn0 l=169.133.30.25 b=169.133.31.253 mac=0:6:2b:4:20:15

FQDN=dianas01.dia.dnvr (Updated to DNS)

CIFS Server DIANAS02[DIA]

Full computer name=dianas02.dia.dnvr realm=DIA.DNVR

Comment='EMC-SNAS:T4.2.20.100'

if=fsn1 l=169.133.30.26 b=169.133.31.253 mac=0:6:2b:4:20:16

FQDN=dianas02.dia.dnvr (Updated to DNS)

CIFS Server DIANAS03[DIA]

Full computer name=dianas03.dia.dnvr realm=DIA.DNVR

Comment='EMC-SNAS:T4.2.20.100'

if=fsn2 l=169.133.30.27 b=169.133.31.253 mac=0:6:2b:4:20:17

FQDN=dianas03.dia.dnvr (Updated to DNS)

CIFS Server DIAGIS03[DIA]

Full computer name=diagis03.dia.dnvr realm=DIA.DNVR

Comment='EMC-SNAS:T4.2.20.100'

if=fsn3 l=169.133.30.28 b=169.133.31.253 mac=0:6:2b:4:20:18

FQDN=diagis03.dia.dnvr (Updated to DNS)

**Note:** Single IP address for each NetBIOS Server

**\$ .server\_config server\_2 -v "dns dump"**

1061496646: LIB: 4: DNS cache size for one record type: 64

1061496646: LIB: 4: DNS cache includes 16 item(s):

1061496646: LIB: 4: dianas02.dia.dnvr

1061496646: LIB: 4: Type:A TTL=536 s dataCount:1

1061496646: LIB: 4: 169.133.30.26 (local subnet)

1061496646: LIB: 4: ---

1061496646: LIB: 4: dianas01.dia.dnvr

1061496646: LIB: 4: Type:A TTL=536 s dataCount:2

1061496646: LIB: 4: 169.133.30.28 (local subnet)

1061496646: LIB: 4: 169.133.30.25 (local subnet)

1061496646: LIB: 4: ---

1061496646: LIB: 4: dianas03.dia.dnvr

1061496646: LIB: 4: Type:A TTL=547 s dataCount:1

1061496646: LIB: 4: 169.133.30.27 (local subnet)

1061496646: LIB: 4: ---

1061496646: LIB: 4: diagis03.dia.dnvr

**[Note:** Two IP Addresses cached for dianas01 when should be only .25]

1061496646: LIB: 4: Type:A TTL=552 s dataCount:1  
1061496646: LIB: 4: 169.133.30.28 (local subnet)

## **\$ server dns server 2 -o flush**

server\_2 : done

## **\$ .server config server 2 -v "dns dump"**

```
1061497044: LIB: 4: DNS cache size for one record type: 64
1061497044: LIB: 4: DNS cache includes 4 item(s):
1061497044: LIB: 4: dianas02.dia.dnvr
1061497044: LIB: 4: Type:A TTL=975 s dataCount:1
1061497044: LIB: 4: 169.133.30.26 (local subnet)
1061497044: LIB: 4: ---
1061497044: LIB: 4: dianas03.dia.dnvr
1061497044: LIB: 4: Type:A TTL=990 s dataCount:1
1061497044: LIB: 4: 169.133.30.27 (local subnet)
1061497044: LIB: 4: ---
1061497044: LIB: 4: dianas01.dia.dnvr
1061497044: LIB: 4: Type:A TTL=991 s dataCount:1
1061497044: LIB: 4: 169.133.30.25 (local subnet)
1061497044: LIB: 4: ---
1061497044: LIB: 4: diagis03.dia.dnvr
1061497044: LIB: 4: Type:A TTL=1186 s dataCount:1
1061497044: LIB: 4: 169.133.30.28 (local subnet)
```

**Note:** DM DNS Cache is now normal, along with flush of Client DNS, normal share access restored

## **ROAMING PROFILES & CLIENT/SERVER BEHAVIOR--TIMESTAMP ISSUE:**

The Celerra does not yet calculate time in nanoseconds, as do MS Clients. Therefore, when Windows 2000 & XP Clients use Roaming Profiles that are stored on the Celerra, they compare timestamps of the locally stored profile with those on the Data Mover, and because the Data Mover rounds off all timestamps to the next lowest second by default behavior, the timestamps are always different. This does not usually present a problem during Client logon because timestamps are usually newer on the client (because of the default behavior of the Celerra just mentioned). However, on Client logoff there is a substantial performance exposure because the Client then compares the timestamps of all files and considers its local copy as newer than those on the Server and so initiates a write of all files back to the Celerra. Under ordinary circumstances this does not cause a problem--however, when User profiles grow to hundreds of megabytes in size, or there are many User profiles stored on a single data mover, the potential is for a perceptible performance hit on the Data Mover during logoff periods, and Clients may see that files take several minutes or longer to open. Network traces can usually confirm this by showing the names of the files being processed during the affected timeperiod--they may show up predominantly as User Profile files.

AR39503 (primus emc77448) documents this issue, and with NAS codes 5.3.1.0, 5.2.10.0, and 5.1.21.0 and higher, a parameter and mount option was introduced to force the data mover to round up files to the next whole second. Once the param or mount option is implemented, the Client behavior requires two logon and logoff sessions to get to a point where the Client begins to recognize that the files stored on the Server are newer than the Client, that is, except for files that are genuinely changed and updated by the Client on any given day.

Specifically, after the change is implemented (param & reboot), when a Client logs in, all files are downloaded, just as always. Then, when Client logs off, Client will force rewrite of all files to the Server because the two timestamps will still be different (Client's time will be newer), but this time the Celerra will round up all files to next nearest whole second. Client will log in for the 2nd time, this time downloading all files containing the rounded up Celerra timestamps, because the times have all changed. During the 2nd logoff, things will be different--the Client will now compare timestamps and see that for the most part, the timestamps are the same as those on the datamover, and will only rewrite those files that have been worked on during the day to the Celerra, thereby decreasing the size and length of the write operations.

## **OPTIONS WHEN IMPLEMENTING NANOSECOND ROUNDUP:**

**param cifs nanoroundoff=1** [Rounds-up date of files to next second on Data Mover or globally via slot-param file]

**# server\_mount server\_2 -o cifsnanoroundup -n davefs /davefs** [use option to change timestamp on a single fs]  
davefs on /davefs uxf5,perm,rw,cifsnanoroundup

**Note:** One final comment, if a User's profile is very large, and the data is always being changed, then changes will still be rewritten constantly to the Celerra--if this is the case, then the size and validity of the data in the User's profile should be verified. In one example, a User's profile was close to 1GB in size!

**# du -sh /\* |grep M**

4.6M ./Personal

**766M ./Settings**

2.9M ./Templates

## **CREATING FILE SYSTEMS--Configuring, Mounting, Exporting:**

**Note:** 2048=2GB; 4096=4GB

- I. **Create (4) 2GB Slice Volumes:** #nas\_slice -n slv01 -c d34 2048 #nas\_slice -n slv02 -c d34 2048 2048  
#nas\_slice -n slv03 -c d34 2048 4096 #nas\_slice -n slv04 -c d34 2048 6144
- II. **Create (4) Stripe Volumes:** #nas\_volume -n stv01 -c -S 8192 d3, d4, d5 #nas\_volume -n stv02 -c -S 8192 d6, d7, d8  
#nas\_volume -n stv03 -c -S 8192 d9, d10, d11 #nas\_volume -n stv04 -c -S 8192 d12, d13, d14
- III. **Create (7) MetaVolumes (comprised of Slice, Stripe, and Disk Volumes):**  
#nas\_volume -n mtv01 -c -M slv01 #nas\_volume -n mtv02 -c -M stv01  
#nas\_volume -n mtv03 -c -M slv02, stv02 #nas\_volume -n mtv04 -c -M d15  
#nas\_volume -n mtv05 -c -M d16, slv03 #nas\_volume -n mtv06 -c -M d17, stv03  
#nas\_volume -n mtv07 -c -M d18, stv04, slv04
- IV. **Create Unix File Systems:** \$nas\_fs -i -all  
#nas\_fs -n ufs01 -c mtv01 #nas\_fs -n ufs02 -c mtv02 #nas\_fs -n ufs03 -c mtv03  
#nas\_fs -n ufs04 -c mtv04 #nas\_fs -n ufs05 -c mtv05 #nas\_fs -n ufs06 -c mtv06  
#nas\_fs -n ufs07 -c mtv07
- V. **Create Files System Mount Points for DataMover:**  
#server\_mountpoint server\_2 -c /s2-mp01 #server\_mountpoint server\_2 -c /s2-mp02  
#server\_mountpoint server\_3 -c /s3-mp01 #server\_mountpoint server\_3 -c /s3-mp02
- VI. **Mount the File Systems:**  
#server\_mount server\_2 -p ufs01 /s2-mp01 #server\_mount server\_2 -p ufs02 /s2-mp02  
#server\_mount server\_3 -p ufs03 /s3-mp01 #server\_mount server\_3 -p ufs04 /s3-mp02
- VII. **Exporting File Systems:**  
#server\_export server\_2 -p -o anon=0 /s2-mp01 #server\_export server\_2 -p -o anon=0 /s2-mp02  
#server\_export server\_3 -p -o anon=0 /s3-mp01 #server\_export server\_3 -p -o anon=0 /s3-mp02
- VIII. **Mounting a FileSystem from a Remote Client System:**  
1. Create local mount\_point first or use /mnt for Solaris  
2. Mount remote FileSystem by IP: #mount -F nfs 193.1.21.173:/remotemntpt /mnt  
3. Mount remote FileSystem by Hostname: #mount -F nfs server\_2:/mountnew /mnt [requires /etc/hosts file on Server]

## **CREATING FILE SYSTEM FROM WEBUI USING AVM PROFILE:**

**nas\_fs -name cdms -type ufs -create size=50000M pool=clar\_r5\_performance storage=SINGLE worm=off -option slice=y,mover=server\_2** (NAS 5.4)

**Note:** Above output is from cmd\_log and shows how NAS 5.4 AVM creates file system using Storage Pools

## **CREATING FILE SYSTEM FROM CLI USING AVM PROFILE:**

**# nas\_fs -name migrate -create size=1000M pool=clar\_r5\_performance -o slice=y**

## **CREATING MULTIPLE FILE SYSTEMS VIA CLI:**

**# for i in \$(seq 1 20); do nas\_fs -name fs\$i -create size=10G pool=clar\_r5\_performance; done**

## **INTRODUCTION TO CIFS [Common Internet File System]:**

CIFS is a file sharing and network access protocol based on the Server Message Block (SMB) protocol, adopted by Microsoft to run on top of NetBIOS (Network Basic Input/Output System)--the building blocks upon which traditional Windows-based Microsoft Networks are built. NetBIOS served as a 'Session Layer API' that provided (3) basic functions for a Microsoft Network: 1) Name Service host resolution 2) Datagram Delivery Service using UDP 3) Session Service that established & maintained point-to-point connection-oriented NetBIOS sessions over TCP/IP between 'networked' computer systems.

Most recently, Microsoft has redesigned CIFS to be the native file sharing protocol for Windows 2000 and implements SMB directly over TCP/IP using Port 445 (traditional networks used SMB over NBT over TCP/IP using Port 139) and makes CIFS 'transport independent.'

## **CIFS/SMB SUPPORT:**

CIFS is implemented as a Client/Server request & response protocol that supports File Sharing Services, Messaging, File Locking, Security, Authentication, & Short/Long File Naming conventions using communication mechanisms related to File Sharing, API's, Named Pipes, & Mail Slot communications. Client systems issue SMB commands that are handled by the local 'redirector' to connect to 'remote systems' or resources using connection establishment messages. After authentication & security considerations, clients receive access to Shares for opening, reading, writing, and executing files. Other forms of messages between Client & Server are Mail Slots and Named Pipes [Mail Slots is a connectionless broadcast delivery service often used in Browsing services, while Named Pipes is a form of connection-oriented messaging via a virtual circuit or 'pipe', established between client & server, so that two applications or processes can communicate with each other. Data is passed as output from (1) host process to input on the other host].

Remote Procedure Calls (RPC) are the basic mechanism by which a Client computer can make a network request to use the processing capabilities of a Server--both Client/Servers use an End Point Mapper service to listen on Port 135 for TCP/IP connection

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
requests. Clients must then conduct an RPC-bind to an interface before it can begin making SMB or other types of procedure calls.

Celerra implements the CIFS file sharing protocol to allow network access for Windows-based Users--CIFS is implemented inside the Operating System kernel for better performance and is layered on top of DART (Data Access in Real Time).

**Comment:** Celerra supports “SMB Signing”, which is a form of host-based authentication [See Q161372], with NAS 5.2 +.

## **CELERRA CIFS SERVICE:**

The Celerra CIFS Service runs as an agent on DART. The ‘service’ uses the CIFS Protocol to provide for access to ‘network resources’ or ‘application processing’ in a Windows environment [other Protocols are NFS, FTP].

## **DOMAIN CONTROLLER DISCOVERY & SELECTION:**

1. Every 15 minutes, DM broadcasts on each enabled interface requesting a list of DCs for each Domain configured
2. Every 15 minutes, DM also sends request for list of DCs from each interface to WINS server, if configured.
3. Answer to above two actions is a list of IP Addresses
4. DM builds this list as \*SMBSERVER for each IP
5. DM then queries for name of each DC on list using NBTSTAT -A 192.10.3.2 or LDAP ping for Windows 2000 systems
6. For each answer, \*SMBSERVER is replaced by name and response time

### **Notes:**

--If all DC's on list become marked as unavailable, then Steps 1 & 2 are initiated

### **What Marks a DC as Unavailable?**

1. TCP connection cannot be established, or transmission fails
2. NETLOGON request cannot be established [IPC\$ pipe broken]

**Note:** DCs marked unavailable are not used until list is refreshed

## **WINDOWS 2000 DC DISCOVERY & SELECTION:**

1. DART queries DNS for list of all DC's in the domain using \_ldap.\_tcp.dc.\_msdcs.DOMAINNAME as the DNS Service Locator
2. DART then conducts CLDAP netlogon RPC Ping call to each DC, reviewing bit 0x80 in ADS\_CLOSEST field, choosing the DC with the best response time.
3. MS clients differ at this point and checks both the response time and the site name, then queries for \_ldap.\_tcp.SITENAME.\_sites.DOMAINNAME to select DC's, choosing best priority/weight.
4. DNS is queried again for list of Kerberos Servers to use, \_kerberos.\_tcp.SITENAME.\_sites.DOMAINNAME

**Note:** CLDAP is only used in AD to discover which Site a client belongs to, then using DNS to find the right DC to use. CLDAP uses LDAP over udp on Port 389. Used to query AD domain controllers. C:>nlttest /dsgetdc:domainname

## **CLIENT USER CONNECTS TO DATA MOVER:**

1. Upon client connection, DM uses first available DC in list to authenticate
2. DM uses quick UDP NETBIOS request for name of DC
3. If successful, then establishes IPC\$ named pipes connection to DC using NETLOGON Secure Channel service
4. Client is authenticated through this Secure Channel pipe [Failure to establish IPC\$ results in DC unavailable status]

## **NFS V. CIFS ENVIRONMENTS:**

NFS is to Unix what CIFS is to Windows. Celerra bridges the two worlds by providing for a multi-protocol interface to access a filesystem via either Network OS protocol.

## **PRIMARY DIFFERENCES BETWEEN NFS & CIFS:**

### **STATEFUL V. STATELESS:**

NFS is stateless-Error recovery is the responsibility of the Server--clients continue to query until a Server

CIFS is stateful-Error recovery is responsibility of applications and generally require restarting from Client

### **HOST V. USER-BASED AUTHENTICATION:**

NFS uses Host-based authentication-machine Hostname + UID/GID of User + NIS Netgroup or Password files

CIFS uses User-based authentication-User logs on providing valid Username/Password to PDC [Challenge & Response Authent.]

### **NFS V. CIFS PERMS FOR AUTHORIZATION TO NETWORK RESOURCES:**

NFS uses Unix file permissions [RWX-UGO] to determine file access or authorization

CIFS uses Extended file permissions [RWXPDO] & ACL's to determine level of file access or authorization

### **ADVISORY V. OPPORTUNISTIC LOCKS:**

NFS uses advisory locking mechanisms to handle file integrity

CIFS locks are mandatory & employ Opportunistic & Client caching mechanisms to support enhanced file access & performance

### **UNIX FILE NAMES V. CIFS:**

NFS supports use of Unix long names only

## **CIFS FILESYSTEM FEATURES:**

File access; File & Record locking; Safe caching read ahead & write behind; file change notification; Extended attributes; Server name resolution independence; Batched requests; Unicode file names

**File Access:** Open, close, read, write, seek

**File & Record Locking:** File & record locking, as well as unlocked access to files. Applications that lock cannot interfere with applications that do not lock.

**Safe caching, Read-Ahead/Write-Behind:** Caching & Read Ahead can be done by multiple clients to files while Write Operations have to be handled by the Server.

**File Change Notification:** Applications register with Server to be notified when File or Directory contents changed

**Extended Attributes:** Such as author's name, content description, etc

**Server Name Resolution:** Use of DNS, WINS, etc allowed

**Batch Operations:** Allows for multiple requests to be served in one message, improves efficiencies

## **CIFS/NFS AUTHORIZATION OR ACCESS:**

### **UNIX ACCESS PERMISSIONS:**

Access Classes [User, Group, Other] + Access Modes [read, write, execute] = Unix Authorization

Unix uses (3) bits each to define USER, GROUP, and OTHER “rwx” permissions [RWX→UGO]

Unix file permission sets contains Unix UIDs and GIDs

**User=Owner**

**Group=Groups**

**Other=all other users except root**

Read Permissions--list files only, in directories

Execute Permissions--utilize files that you can name

Read/Execute Permissions--list and utilize files

Read/Write/Execute Permissions--list files in directories, utilize files, create/delete files

### **CIFS ACCESS PERMISSIONS:**

CIFS access is based on ACLs

NT Access mask is more verbose than Unix [contains 32 bits v. 3 bits for Unix]--many bits are used to describe Directory SDs, etc  
User-Level ACL's [RWXPDO] + Share-Level ACL's = Authorization

SD's [Security Descriptors] describe Access Rights for Users on Files, Folders, or other Objects

SD's contain Access Control Lists [ACL's] which are constructed from Access Control Entries [ACE's] and typically contain SID's [Security Identifier Descriptors] for Users and Groups & their associated rights.

**Authorization** is granted or denied after comparing the Security Access Token [SAT] to an Object's SD [security descriptor] to determine permissions overall permissions to the object. A SD is made up of an ACL containing User and Group ACE entries.

**Authentication or User Access** is determined by a Security Access Token constructed for each User based on the User SID, Group SID, and Username/Password.

### **THREE TYPES OF PERMISSION SETS USED BY NT:**

**Share-level:** Permissions set on network shares → No Access; Read; Change; Full Control

**Directory-level:** NTFS permissions set on directories → No Access; List; Read; Add & Read; Change; Full Control; Special Directory and Special File Access [RWXDPO]

**File-level:** NTFS permissions set on files → No Access; Read; Change; Full Control; Special Access [RWXDPO]

## **CIFS/NFS AUTHENTICATION:**

### **UNIX AUTHENTICATION:**

**Host-based:** Trusted Hosts or machines authenticate to each other over the network

UID/GID is the basis for User authentication for NIS Netgroup and Password/Group files

### **NFS SECURITY:**

→For every NFS request by default (including mount requests), Unix operating systems traditionally deployed the RPC security mechanism called AUTH\_SYS, which contains a set of User credentials containing UID & a list of relevant GIDs. The Server in turn uses the presented credentials for permission/access-checking for Unix file access.

→Special use cases are Root User; Anonymous User; and Users with too many Groups

### **AUTH\_SYS sends 3 important things:**

→A 32 bit numeric user identifier (/etc/passwd file)

→A 32 bit primary numeric group identifier (/etc/passwd file associated with the UID)

→A variable length list of up to (16) 32-bit numeric supplemental group identifiers (/etc/group file)

## **CIFS AUTHENTICATION:**

### **User-based:**

CIFS Server → NETLOGON process to log into NT Domain as a member server using Secure Channel IPC\$ connection.

CIFS User → Attempts access to CIFS Share

CIFS Server → Obtains UID/GID Mapping of User from Password File or Usrmapper Service

CIFS Server → Interrogates DC for security permissions [SAT] of User

CIFS User → Granted access to filesystem

**Authentication** occurs when a User connects to a DataMover and the Datamover obtains the User's NT credentials from the Domain Controller—this consists of the Security Access Token—User's SID, Group ID & SIDs, and Access Rights.

**→ Authentication is the process of establishing one's identity whereas Authorization is the process of limiting actions taken to a managed object through use of permissions and Access Tokens.**

## **CIFS AUTHENTICATION V. CIFS ACCESS:**

**Authentication** is the process of successfully logging into the datamover and obtaining a Security Access Token, which has the User's SID, Group ID & SIDs, and Access Rights.

**Access** is when the Security Access Token [SAT] is compared to an Object's SD [security descriptor] to determine a User's access.

## **MICROSOFT AUTHENTICATION METHODS:**

**LM→LAN Manager** Challenge/Response—uses weak security and sends hash of user password for WinLogon (Win95 era).

Passwords are case insensitive, restricting number of characters for use, with long passwords divided into 7-char. chunks

**NTLM→NT LAN Manager or NTLMv1** Introduced with NT, uses LanMan hash of password—adds case sensitivity and removes password division problem that LM had, no signing or encryption available

**NTLMv2→NT LAN Manager version 2** Introduced with NT SP4, native to Win2k, added via DS Client to Win9x.

Increases password space key to 128bits; Establishes secure channel Signing and/or Encryption between Client & Server to conduct Challenge/Response; Uses encryption of messages. NTLMv2 is used for Session security and also for Authentication.

**Note:** Celerra only supports the password length used with NTLMv2 in that if a client uses such a password, Server can forward to DC for authentication. See AR44353 & 41591. Native NTLMv2 support provided in 5.3.18.0/5.4.16.0—see AR56266 and primus emc108209.

**MS-KRB→Microsoft Kerberos 5** Windows 2000/XP/2003

## **SMB SIGNING:**

NAS 5.1.18.8 can co-exist in Win2k3 environments (that use SMB Signing) with modifications outlined in emc77391—versions prior to this cannot. NAS 5.2+ provides native SMB Signing support. Win2k3 uses SMB Signing by default. See AR75020 to enable SMB signing in registry by default for when SMB Signing is not governed by GPO policy. See AR85147 for SMB Signing issue related to TCP reset on connections where Client was using SMB Signing, then failed to use, at which point Celerra conducted reset to TCP connection. AR85147 will disable SMB Signing as the default in the registry of the DM.

### **SMB SIGNING ISSUE NAS 5.5:**

AR75020 made Data Mover capable of using SMB Signing with 5.5.23.0 & 5.4.25.0 when GPO Policies are left undefined—this capability was extended in the DM registry. This AR introduced a regression whereby the Data Mover will disconnect a TCP session with the Client, if the client had been using SMB Signing, and then inexplicably began not using it. AR85147 disables the registry setting on the Data Mover so that in situations where GPO Policies for SMB Signing are not defined, clients will not be forced to use SMB Signing.

## **CHANGING SPECIFIC CIFS SERVER SYSTEM REGISTRY:**

**# .server\_config server\_2 -v "ntreg**

**find=system\currentcontrolset\services\lanmanserver\parameters\enablesecuritysignature set=0"**

**Note:** SMB Signing for Data Movers that rely on registry settings (for example where GPO Settings for SMB Signing are undefined) will now be disabled by default in the registry with NAS Versions 5.4.28.0 & 5.5.25.2. The Data Mover will still support and enforce SMB Signing whenever GPO Policies are defined.

## **VERIFYING CURRENT SMB SIGNING SETTING ON DM REGISTRY:**

**# .server\_config server\_2 -v "ntreg**

**find=SYSTEM\CurrentControlSet\Services\LanmanServer\Parameters\enableSecuritySignature"**

1173795874: SMB: 3: FOUND:\System\CurrentControlSet\Services\LanmanServer\Parameters\EnableSecuritySignature=1 (**0x1**)

## **VIEWING GLOBAL DATA MOVER SMB SIGNING SETTINGS:**

**# server\_param server\_2 -facility cifs -info smbsigning**

server\_2 :

name = smbsigning

```

facility_name      = cifs
default_value     = 1
current_value     = 1
configured_value  =
user_action       = restart Service
change_effective  = restart Service
range             = (0,1)
description       = Controls SMB signing on the data mover

```

Used to enable or disable SMB signing "globally" on the Data Mover. param cifs smbsigning=1 enables SMB signing, and is the default setting param cifs smbsigning=0 disables SMB signing. Disabling this parameter will not work unless it is also disabled on the Windows 2003 clients.

#### **Symptoms of SMB Signing When Not Supported:**

Problems could be seen with Join issues that require SMB-signing of Kerberos tickets, Client connection failures, timeouts, hangs, or Password Challenge box.

#### **SMB SIGNING GPO:**

**To configure SMB message signing Windows Server 2003, Windows XP, and Windows 2000, use the following GPO options:**

#### **DEFAULT POLICIES WIN2K:**

|                                                     | <b>Local</b> | <b>Effective</b> |
|-----------------------------------------------------|--------------|------------------|
| Digitally sign server communication (when possible) | Disabled     | Enabled          |
| Digitally sign server communication (always)        | Disabled     | Disabled         |
| Digitally sign client communication (when possible) | Enabled      | Enabled          |
| Digitally sign client communication (always)        | Disabled     | Disabled         |

#### **Domain Security Policy**

|                                                     |             |
|-----------------------------------------------------|-------------|
| Digitally sign server communication (when possible) | Not defined |
| Digitally sign server communication (always)        | Not defined |
| Digitally sign client communication (when possible) | Not defined |
| Digitally sign client communication (always)        | Not defined |

#### **Domain Controller Security Policy**

|                                                     |             |
|-----------------------------------------------------|-------------|
| Digitally sign server communication (when possible) | Enabled     |
| Digitally sign server communication (always)        | Not defined |
| Digitally sign client communication (when possible) | Not defined |
| Digitally sign client communication (always)        | Not defined |

#### **DEFAULT POLICIES NATIVE WIN2K3:**

#### **Local Computer Policy**

|                                                                           |          |
|---------------------------------------------------------------------------|----------|
| Microsoft Network server: Digitally sign communication (if client agrees) | Enabled  |
| Microsoft Network server: Digitally sign communication (always)           | Enabled  |
| Microsoft Network client: Digitally sign communication (if server agrees) | Enabled  |
| Microsoft Network client: Digitally sign communication (always)           | Disabled |

#### **Domain Security Policy**

|                                                                           |             |
|---------------------------------------------------------------------------|-------------|
| Microsoft Network server: Digitally sign communication (if client agrees) | Not defined |
| Microsoft Network server: Digitally sign communication (always)           | Not defined |
| Microsoft Network client: Digitally sign communication (if server agrees) | Not defined |
| Microsoft Network client: Digitally sign communication (always)           | Not defined |

#### **Domain Controller Security Policy**

|                                                                           |                                     |
|---------------------------------------------------------------------------|-------------------------------------|
| Microsoft Network server: Digitally sign communication (if client agrees) | Enabled                             |
| Microsoft Network server: Digitally sign communication (always)           | Enabled [Not defined in mixed mode] |
| Microsoft Network client: Digitally sign communication (if server agrees) | Not defined                         |
| Microsoft Network client: Digitally sign communication (always)           | Not defined                         |

#### **SMB SIGNING TABLE FOR WINDOWS & DART:**

#### **WINDOWS O/S**

#### **W2kPro SP4**

```

LanmanWorkstation enableSecuritySignature=1
LanmanWorkstation requireSecuritySignature=0
LanmanServer enableSecuritySignature=0
LanmanServer requireSecuritySignature=0

```

#### **W2kServer SP4**

```

LanmanWorkstation enableSecuritySignature=1
LanmanWorkstation requireSecuritySignature=0

```

LanmanServer enableSecuritySignature=1

LanmanServer requireSecuritySignature=0

### **XPPro**

LanmanWorkstation enableSecuritySignature=1

LanmanWorkstation requireSecuritySignature=0

LanmanServer enableSecuritySignature=0

LanmanServer requireSecuritySignature=0

### **W2k3 Server**

LanmanWorkstation enableSecuritySignature=1

LanmanWorkstation requireSecuritySignature=0

LanmanServer enableSecuritySignature=1

LanmanServer requireSecuritySignature=1

### **Vista**

LanmanWorkstation enableSecuritySignature=1

LanmanWorkstation requireSecuritySignature=0

LanmanServer enableSecuritySignature=0

LanmanServer requireSecuritySignature=0

### **DART O/S**

#### **DART 5.5.23/5.4.25.1 Prior to AR85147**

LanmanWorkstation enableSecuritySignature=1

LanmanWorkstation requireSecuritySignature=0

LanmanServer enableSecuritySignature=1

LanmanServer requireSecuritySignature=0

#### **DART 5.5.27.5 after AR85147 Fix**

LanmanWorkstation enableSecuritySignature=1

LanmanWorkstation requireSecuritySignature=0

LanmanServer enableSecuritySignature=0

LanmanServer requireSecuritySignature=0

### **NTLMv2 AUTHENTICATION:**

MS04-011 Hot Fix changes default to use NTLMv2. Celerra Versions 5.1.24.0, 5.2.12.0, & 5.3.4.0 can provide NTLMv2 password response to DC's. Native NTLMv2 support provided in NAS 5.3.18.0 & NAS 5.4.16.0

#### **Symptoms of Enforced NTLMv2 (settings 3, 4, or 5):**

Session Setups to Data Mover will fail

#### **NTLMV2 REGISTRY SETTINGS:**

**HKLM>System>CurrentControlSet>Control>Lsa>LMCompatibilityLevel:** REG\_DWORD

0 – Default→Send LM & NTLM responses, do not use NTLMv2 Session Security

1 – Send LM & NTLM responses, use NTLMv2 if negotiated

2 – Send NTLM responses only

3 – Send NTLMv2 responses only

4 – Refuse LM responses, Send NTLMv2 response only

5 – Refuse LM & NTLM responses, Send NTLMv2 response only

### **STRONG LDAP SIGNING:**

Native support tentatively scheduled for NAS 5.5. In an effort to further bolster network security for Windows domains, Microsoft has introduced the ability to require LDAP Signing when performing LDAP Bind queries to AD Servers.

#### **Symptoms of Strong LDAP Signing:**

With LdapServerSecurity=2 set in the registry of each Domain Controller, LDAP Bind requests to AD will fail for any authentication mode that does not use SASL or TLS/SSL, which is what occurs when trying to conduct a Join.

#### **LDAPSERVERINTEGRITY REGISTRY SETTINGS:**

**HKLM>System>CurrentControlSet>Services>NTDS>Parameters** LDAPServerIntegrity=2 [Strong LDAP Signing or SASL required]

**LDAP BIND SETTINGS:** Values are as follows for LDAP server (ldapagnt.lib) handling of LDAP bind command requests

**LdapServerIntegrity=1** Default or not defined—LDAP Server supports client request for LDAP traffic signing when handling a LDAP bind command request which specifies a SASL authentication mechanism.

**LdapServerIntegrity=2** LDAP Server requires LDAP traffic signing, unless LDAP bind request is already protected with TLS/SSL. It rejects the LDAP bind command request if other types of authentication are used.

### **DATAMOVER NETLOGON PROCESS—NT 4.0:**

## **NT AUTHENTICATION & CELERRA:**

Through the NETLOGON SERVICE, whenever the Celerra CIFS Server is started, it logs into the NT Domain as a member server and communicates to the Domain Controller using an encrypted Secure Channel for account password authentication upon startup, and then thereafter for User Authentication. During the initial session setup by the Client User, as a request for resources located on the CIFS server, the CIFS Server processes the Challenge & Reply request from the Domain Controller to the User using NTLM 0.12 Challenge & Response. If successfully authenticated by Username & Password, the User obtains a Security Access Token [SAT] which contains their SID, UID/GID, and Access Rights. This SAT is then compared to each CIFS object's Security Descriptor [SD] to determine level of access.

**Note:** See emc108209 for Celerra NTLMv2 Session Security (Signing and Sealing) or Authentication, introduced in NT 4.0 SP4  
**Two-Part Process:**

Step 1. User's username & password are authenticated by Celerra against a Domain Controller via the Netlogon Service.

Step 2. User is mapped to a valid UID/GID from a Password/Group file, NIS, or Usermapper database.

## **USER AUTHENTICATION PROCESS & THE DATA MOVER:**

--User authentication is mostly a single threaded operation on the DM, meaning that if there is a queue of 100 requests, each one is serviced in a linear fashion

--Invalid Users should normally be rejected immediately so as to not delay other requests

--However, in the case of Trusted Domains, if a User is trying to authenticate but the trusted DC is not available, originating DC will continue to probe until the 15 second Windows timeout value.

--Data Mover's timeout value is 20 seconds

## **PARAMETER TO SWITCH DC WITH ‘NO LOGON SERVER’ ERROR:**

**param cifs Ntsec.SwitchDC\_OnNoLogonServer=1**

**Note:** 5.1.18.5, 4.2.21.0 disables DC switching with default value of 1. Set 0 to enable switching of DC's with NoLogon Error. This param is no longer in 5.5 code.

## **WINDOWS NT TERMINOLOGY:**

### **SIDs**

SIDs are unique values that are used to identify a “trustee”. A ‘trustee’ is a User, Group, or Logon Session to which an ACE applies. Each ACE in an ACL has one SID associated with a ‘trustee’. Well-known SIDs identify generic Groups and Users.

### **Well-Known Sid Types:**

SID: S-1-1-0

Name: Everyone

Description: A group that includes all users, even anonymous users and guests. Membership is controlled by the operating system.

SID: S-1-5-domain-512

Name: Domain Admins

Description: A global group whose members are authorized to administer the domain. By default, the Domain Admins group is a member of the Administrators group on all computers that have joined a domain, including the domain controllers. Domain Admins is the default owner of any object that is created by any member of the group.

SID: S-1-5-domain-513

Name: Domain Users

Description: A global group that, by default, includes all user accounts in a domain. When you create a user account in a domain, it is added to this group by default.

SID: S-1-5-domain-514

Name: Domain Guests

Description: A global group that, by default, has only one member, the domain's built-in Guest account.

SID: S-1-5-11

Name: Authenticated Users

Description: A group that includes all users whose identities were authenticated when they logged on. Membership is controlled by the operating system.

**Note:** Above listing are examples only, list is very long. Current Celerra code [5.2.11.301 and lower] does not support the ability to add Well-Known Sid types to the LocalGroups database.

### **SUMMARY OF ISSUE FROM CELERRA PERSPECTIVE:**

In the process of obtaining SID information from the DC when a User connects to the Celerra, we also parse the localgroups file to see if the user belongs to any localgroups. Unfortunately, the current mechanism for retrieving SID information from the DC does not return the Everyone and other 'Well-known Sid' types such as 'Authenticated User'.

### **Access Control List (ACLs)**

An Access Control List is a table of entries (ACEs) for Objects that contain a Security Descriptor, such as files or directories, that further defines access rights for Users or Groups. Collectively, each User or Groups' Access Control Entry creates the ACL Table.

### Access Control Entries (ACEs)

Access Control Entries contain a set of access rights & SID (Security Identifier) for a User or Group to whom the rights are allowed, denied, or audited. An ACE is an entry in the ACL Table for each “trustee” (User, Group, SID, domain\trustee). Windows 2000 adds object-specified ACE’s and Automatic Inheritance. ACEs contain trustees, 32-bit access masks, ACE type (deny, audit), inheritance flag. Three types of ACEs are: Deny, Grant, and Audit

#### ACE Directory Inherit Flag (0x2)

Directory and all inherited sub-directories

#### ACE Files Inherit Flag (0x9)

Only inherited files

#### ACE Taking Ownership of Directory (0x3)

Inherit on files and directories

#### ACE FIELDS WITHIN SD:

ace\_type;ace\_flags;rights;object\_guid;inherit\_object\_guid;account\_sid

#### Security Descriptor (SDs)

Security Descriptors are the basic security information and structure for a “securable object”. The SD further identifies an Object’s Owner and Primary Group. An Object can also have a DACL (Discretionary Access Control List) and SACL (System Access Control List) associated with it to control access to the Object and control security auditing, respectively. The structure of an SD includes: SIDs of Owner & Primary Group; DACL that specifies Access Rights allowed or denied to Users or Groups; SACL that specifies access events that generate audit records; set of control bits that qualify the meaning of SD. The SDDL (Security Descriptor Description Language) uses ACE strings in the DACL or SACL components of a Security Descriptor string.

#### FOUR MAIN COMPONENTS OF THE SECURITY DESCRIPTOR:

O:owner\_sid → SID string identifying Object owner

G:group\_sid → SID string identifying Object’s Primary Group

D:dACL\_flags(string\_ace1)(string\_ace2)... (string\_acen) → Control flags for DACL and ACE strings

S:SACL\_flags(string\_ace1)(string\_ace2)... (string\_acen) → Control flags for SACL and ACE strings

#### SID STRINGS USED IN THE SD:

Owner

Primary group

The trustee in an ACE

#### SD BUFFER SIZE ISSUE: fixed NAS 4.2.13.3

**getMaxSmbSize:** Param to increase buffer size when reading SD’s from files created on other NetBIOS servers

#### DACL

Discretionary Access Control List controls access to an Object and is determined by the Object’s Owner

#### SACL

System Access Control List controls auditing for security of an Object and is determined by the Administrator

#### SECURITY ACCESS TOKENS (SATs)

Security Access Tokens are created & updated for a User each time they log onto an NT Domain. This Token identifies the User, their Groups, and privileges to resources and processes on the domain. The Operating System uses this Token to allow or deny access to resources or processes.

#### Access Rights

Access Rights are bit flags for operations that a thread can perform on a securable object (i.e., an object with an SD)

#### Access Masks Encoding

ACCESS\_MASKs are 32-bit values containing 'standard', 'specific', & 'generic' rights used to build ACE’s [Access Control Entries] on objects. This is the primary means for granting access to objects in Windows NT. Access Masks specify allowed or denied Access Rights controlled by ACEs. In addition, access masks are grouped into a number of categories:

#### Generic Access Rights

GENERIC\_ALL is used to indicate Read, Write, & Execute access

GENERIC\_EXECUTE is used to indicate Execute access

GENERIC\_READ indicates read access

GENERIC\_WRITE indicates Write Access

#### Standard Access Rights

DELETE is used to delete an object

READ\_CONTROL is the right to read an object’s SD information

SYNCHRONIZE allows the object to use synchronization and allows threads to wait for the correct state

WRITE\_DAC is the right to modify DACL in the SD of the object

WRITE\_OWNER is the right to change the owner in the SD of the object

#### SACL Access Rights

Used to get or set the SACL in an object’s SD

#### Directory Services Access Rights

Various AD object SD’s, as follows: ACTRL\_DS\_OPEN, ACTRL\_DS\_CREATE\_CHILD, ACTRL\_DS\_DELETE\_CHILD,

#### Access Mask Table:

Access Masks represent Windows permissions built with one or more bits as 32-bit values. Processes that request access to objects also present an Access Mask for the access that is requested. The O/S compares the access mask to the masks contained in the ACE entries, looking for bits that match.

31 30 29 28 27 26 25 24 23 22 21 20 19 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1 0

#### Notes:

Bits 0-15 are considered low-order bits used for object specific access rights

Bits 16-22 are the bits used to define Standard Access Rights

Bit 23 is for the right to access an objects' SACL—Audit Access privilege

Bits 24-27 are reserved

Bits 28-31 are used to define Generic\_All, Generic\_Execute, Generic\_Write,& Generic\_Read, respectively, and are considered the high-order bits.

## **CIFS FILE NAME TYPES:**

**M83 → DOS 8.3 format consists of two parts:** Basename of up to (8) characters; Extension of up to (3) characters

Basename & Extension are separated by a 'period'.

All characters are legal in 8.3 format except for the following: space character "0x20" " . / \ [ ]: + | < > = ; , \* ?

**M256 → Long File Names:** Supports up to 255 Characters; indicated by setting bit2 in Flags2 field of SMB Header

Wildcards: SMB allows use of wildcards for both 8.3 Naming and Long File Naming

#### DOS 8.3 Wildcards: \* ? .

\* matches 0 or more characters until reaching the . in the name

? matches any single character, or until encounters a . in the name string. Can use multiple ??z for searches.

. matches a . or an empty extension string

#### Long File Name Wildcards:

Files may now have none, one, or more than one . within the name. 'Spaces' within file names are also allowed.

? matches single characters; can use multiple ??

\* matches entire name

## **ENFORCING STRICT/LOOSE WILDCARD MATCHING:**

**param cifs ntWildcardMode=0** e:\celshare>dir ?????? [Strict wildcard matching, only filenames with six characters are returned; Win9x style]

**param cifs ntWildcardMode=1** [Loose Wildcard matching; all filenames with six characters or less would be returned; NT]

## **CELLERRA FILE LOCKING POLICIES:**

Most Operating Systems prevent "concurrent" writes to a single file from more than one process at a time. This is called the DOS & NT "deny-mode".

**CIFS Clients:** CIFS Locks--'nolock' is the default; 'wlock', 'rwlock' [multi-protocol modes]

Lock requests are made, locks respected. I/O request made, ONLY CIFS locks respected.

CIFS Mandatory Locks--byte-

**UNIX Clients:** NFS Locks--Read, Exclusive Write Locks, & Advisory Locks

Lock requests are made, locks respected. I/O request made, (3) possible options:

Server allows client lock until another lock request is served, then updates changes before allowing other client lock.

NFS Locks considered 'cooperative', as they allow other clients to access locked files, as long as 2<sup>nd</sup> client does not try to lock file

Advisory Locks inform other Clients that a file is in use, and does not affect Read or Write access

Exclusive Locks can still be Read, but no other client can lock the file again

Read Locks can be taken out by multiple clients or processes

## **HOW CELERRA USES CIFS LOCKING MECHANISMS:**

**Note:** Byte Range and Sharing Locks are usually handled by "applications", whereas "Oplocks" in general are handled by the Client's Operating System, though they are still generally initiated as a result of Client Applications. For interoperability between NFS & CIFS File Locking, the basic rule is that: NFS Locks are Advisory & CIFS locks are mandatory.

Locking policies are Implemented during the "Mounting" of a CIFS File System using "server\_mount".

**File Locking Defaults:** "nolock", "oplocks"on, "notify" on.

**File Locking Options:** "wlock", cifs "rwlock", "nooplocks", "nonotify"

#### NFS Locks:

CIFS clients respect only CIFS locks—ignores NFS locks

Nolock→default where NFS locks are advisory & NFS client can Read & Write to a file

Wlock→NFS reads allowed, but no writes

Rwlock→NFS writes denied, read denied for read locks

“Deny Read-Write” mode request is translated into an NFS exclusive RW lock

### **Celerra Opportunistic Locks (Level I):**

Clients request oplocks in the NT\_CREATE\_ANDX or NT\_TRANSACT\_CREATE when opening a file. If the Server grants as OplockLevel or 1, Client holds a Level II Oplock and caches portion of file locally, with changes flushed to Server only after work completed or if Oplock is revoked. "Oplocks" allows Server to revoke a lock at any time, while allowing Client the opportunity to cache file data safely, and to acknowledge the oplock break request.

**Note:** Celerra does not support Level II Oplocks in NAS 5.x or below. Level I Oplocks supported NAS 4.2.14 +

### **CIFS OPPORTUNISTIC FILE LOCKING:** Client always requests, but Server always controls & grants or doesn't **EXCLUSIVE OPLOCKS:**

Client that requests an 'oplock' will receive either an Exclusive, Level II, or No Oplock grant from Server. The Client process that holds an 'exclusive oplock' is only process that has file open--this allows for safe data read & write caching, metadata caching, and record lock caching. Clients specify a lock by Length and Offset values within a file. Server will revoke this type of lock if other clients request access to same file. Client writes cached data back to Server prior to other Clients accessing file.

→Server informs client of exclusive oplock until client file is closed, at which time server is updated

### **BATCH OPLOCKS:**

Client requesting a 'batch' oplock will receive a Batch, Level II, or No Oplock from Server. Purpose of this type of oplock is to allow Clients to keep a file open that otherwise is opened and closed repeatedly by an Application. Revoked by server as above.

→Same function as Exclusive Oplock except that it will allow the Client to repeatedly Open & Close the file

### **LEVEL II OPLOCKS:**

Clients never request this, but Server may grant in response to an 'Exclusive' or 'Batch' Oplock request. This type of Oplock allows Client to cache 'read' data while files are opened on other Clients. Server revokes without waiting for response.

→Server informs client that multiple clients are accessing a file, but no modifications have been made. Client can perform Read-ahead or cached read or file attribute operations without server updates.

**Note:** Celerra supports Level II Oplocks at NAS 5.1.15.3 and higher. Default Oplocks timeout value is 10 seconds. CIFS Oplocks are always on by default, and configured on a per-filesystem basis. Oplock file changes remain local to Client system cache.

**No Oplock:** Client will always receive a "nooplock" if requested. “No Oplock” turns off Oplocks. For example, you might turn off OPLOCKS for CIFS File Systems using IIS or Database Applications such as SQL, Oracle, ClearCase, Access & MPFS file systems

**Note:** Oplocks only support "files", not "directories" or "named pipes"

### **ENSURING SYNCHRONOUS WRITES FOR DATABASES:**

**Note:** Set mountpoint option per following example—option is actually relevant for both CIFS and NFS file systems. Option was introduced to guarantee synchronous writes to critical databases.

**\$server\_mount server\_5 -o cifssyncwrite fs1 /fs1**

### **CHANGING MAX TIMEOUT VALUE FOR CIFS LOCKX REQUESTS:**

Default time that CIFS will process an incoming lockX request—idea here is to avoid locking all CIFS threads in infinite lockX requests: **param cifs maxLockXTimeout=30000** (30 seconds default)

### **HOW OPLOCKS ARE HANDLED BY AN NT SERVER:**

--Client (1) holds oplock on file

--Client (2) requests 'open' to same file

--Server revokes oplock of Client (1)

--Client (1) flushes data to Server & acknowledges revocation of oplock

--Server then responds to the open request for the file for Client (2)

**Note:** **Oplocks are a function of Client's O/S & clients always initiate request**

**OplockLevel 0:** Client has no oplocks on file

**Oplock Level 1:** Client has Level II oplock [Not supported by Celerra 5.x or below]

### **NFS LOCKING--NLM MANAGER:**

→NFS uses Network Lock Manager [NLM] for file locking by running a Lock Daemon [lockd] on each DataMover

→CIFS also uses a Lock Manager daemon

→File Locks are configured on a FileSystem basis

NLMv1: Unix style file locking

NLMv3: MS-DOS style file locking for NFS v2

NLMv4: Support for large files greater than 4GB [NFS v3]

### **NFS LOCKING\RECOVERY:**

Server reboots, clients reclaim locks via statd/NLM after Server checks against its /var/statmon/sm directory

Server keeps track of client locks using statd daemon

Celerra DataMover writes locking info its root FS; ./etc/registry/Dmhostname

If DataMover has multiple host names, add following to **/nas/server/slot\_x/param** file; “param statd” “hostname=”name1, name2”

Locking is done from Server using Datamover Host name, NOT IP address!!

### **NLM ADVISORY & COOPERATIVE NFS LOCKING:**

→ NFSv2 and NFSv3 use cooperative locking rules (aka, advisory locks), which means that multiple clients can access the same files for Reads and Writes, with clients being notified that file is in use

→ NFSv4 incorporates functionality of NLM within the protocol, and provides mandatory locking, which means client applications can be blocked if another client has the file locked

### **NLM SHARE RESERVATIONS WITH NFS:**

→ Share reservations grant a process Read and/or Write access to a file and ability to deny other processes access to the same file, hence ‘reserving’ the use of the file

→ NLM provides ‘Advisory Share Reservations’ for NFSv2 & v3 and does not prevent access

→ NFSv4 provides ‘Mandatory Share Reservations’ with OPEN operation, and reservations are enforced for file access

### **NLM RANGE LOCKING:**

→ A range lock is a range of bytes in a file, or entire file, that can have “Shared” read locks or “Exclusive” write locks—if an exclusive range lock is taken, no other process can access any range in the file until the exclusive lock is removed. Read and Exclusive locks can be changed back and forth.

### **NFSv4 LEASE LOCKING:**

→ NFSv2 & v3 grants locks by the Server until the Client releases, even if the Client crashes or goes away

→ NFSv4 implements locks for a duration period that need to be renewed by the client. If the Client fails, after the lease expires and a grace period expires, then the file is available for other clients.

### **NFSv4 DELEGATION:**

→ NFSv4 Servers can delegate actions on the client to buffer file data and metadata locally to do work before sending updates to the Server, which improves network performance

→ Delegation is configured on a file system basis with RW delegation as the default

**Note:** Turn off delegation if data is frequently shared between clients, or in the cases of databases or mission-critical data where a client failure could impact data integrity

→ Celerra supports No Delegation; Read Delegation only; Read & Write Delegation

## **SETTING UP FILE LOCKING ON CELERRA FILE SYSTEMS:**

**# server\_mount server\_2 -o noblock [nolock|wlock|rwlock]**

**Note:** Disabling noblock disables CIFS locks

### **I. NoLock-Default Mode**

NFS clients ignore CIFS locks-i.e., NFS Reads-Writes can occur unhindered

- a. When accessing via NFS or FTP, clients can Read and Write to files ‘locked’ by CIFS!
- b. When accessing via CIFS, only CIFS locks are honored.
- c. NFS or CIFS client can request to “lock” a file.
- d. Neither NFS or CIFS client can “lock” a file already locked...

### **II. Wlock-NFS Reads allowed, Writes Denied**

- a. When accessing via NFS or FTP, Client can READ but not WRITE to a file locked by CIFS.
- b. When accessing via CIFS, only CIFS locks are honored.
- c. NFS or CIFS client can request to “lock” a file.
- d. Neither NFS or CIFS client can “lock” a file already locked...

### **III. RWLock-NFS Reads & Writes denied if file locked by CIFS [safest method]**

- a. NFS or CIFS client can request to “lock” a file.
- b. Neither NFS or CIFS client can “lock” a file already locked...
- c. When accessing via CIFS, only CIFS locks are honored.
- d. When accessing via NFS or FTP, if locked, will deny all access to the file.

## **INVESTIGATING LOSS OF NFS ACCESS:**

**Symptom:** Server Log error ‘LOCK:3: lockd\_Monitor: unable to contact local Statd daemon...’

Checking Lockd Stats on Advisory, Mandatory Range, Oblocks, Break Oblocks:

**# .server\_config server\_2 -v "lockd"**

lockd stat [reset]

lockd list

```
lockd info fldp=<address>
lockd info delegp={<address>} 10
lockd remove fldp=<address>
  {type=NFS ip=<addr> pid=<pid>}
  | {type=PCNFS ip=<addr>}
  | {type=CIFS cnxp=<ptr> fid=<fid>}
```

### \$server\_config server\_2 -v 16384 "lockd stat"

```
Lockd statistics since 00:45:08 12/12/2005:
Advisory range locks Granted Blocked Denied
Shared:      22929    0    0
Exclusive:   127737614  97190   517
-----output abridged-----
```

### \$server\_config server\_2 -v 16384 "lockd list"

```
FileSystem: ufs rw /cel9dm2fs1 142=33 rw
  inum=2 NtName(utf8)=cel9dm2fs1
  fldp=0x13de80dc Opens_cnt=1 RangeLocks: Mandatory_cnt=0 Advisory_cnt=0
  inum=9428 NtName(utf8)=COLMAN
```

```
$ .server_config server_2 -v "logsys set severity LOCK=LOG_DEBUG"
```

### DISABLING OPLOCKS AND FILE CACHING ON WINDOWS CLIENTS:

HKLM>System>CurrentControlSet>Services>MRXSmb>Parameters

OplocksDisabled=1 (REG\_DWORD)

NoDeleteOnClose=1 (REG\_DWORD)

### TRADITIONAL NFS LOCKING MECHANISMS:

```
Client 1 → NLM_LOCK_MSG request → NFS Server
Client 1 ← NLM_LOCK_RES granted ← NFS Server
Client 2 → NLM_LOCK_MSG request → NFS Server
Client 2 ← NLM_LOCK_RES blocked ← NFS Server [because Client 1 already has file locked]
Client 1 → NLM_UNLOCK_MSG → NFS Server
Client 1 ← NLM_UNLOCK_RES granted ← NFS Server
Client 2 ← NLM_GRANTED_MSG ← NFS Server
Client 2 → NLM_GRANT_RES accepted → NFS Server
```

### TRANSLATING CIFS TO NFS AND NFS TO CIFS LOCKS:

--CIFS deny RW lock translates into an NFS exclusive RW lock

--NFS shared Read lock translates into a CIFS deny Write lock

### CIFS FILE SYSTEM ACCESSPOLICIES: [Policies apply only if NT Security is set]

Purpose: Accesspolicies are used in ‘multi-protocol’ environments to control Unix access permission bits and Microsoft ACL rights access for files and folders. When referring to accesspolicies, we are referring to “access-checking” based on UNIX or CIFS credentials, or both. Actual access is controlled by “user authorization” on the Server.

UGO—UNIX mode bits, rwxrwxrwx

ACL—Access Control List, contains list of ACEs for file system object

ACE—Access Control Entry

### ACCESSPOLICY OPTIONS:

→**Accesspolicy=NT|UNIX|SECURE|NATIVE|MIXED|MIXED\_COMPAT**

→Implemented on File System basis using server\_mount command

Note: Accesspolicy syntax is case-sensitive with NAS 5.2 and below—use all caps for these earlier NAS versions and also for all versions after 5.5. Accesspolicies can be changed on-the-fly by running the server\_mount command with the revised accesspolicy.

### Server Log:

```
2009-07-08 15:35:22: ADMIN: 6: Command succeeded: serverMount ufs rw /cifs1 122=32 rw accesspolicy=MIXED_COMPAT hwm=90 maxsize=0 (modifyMount)
```

```
2009-07-08 15:35:47: ADMIN: 6: Command succeeded: serverMount ufs rw /cifs1 122=32 rw accesspolicy=NATIVE hwm=90 maxsize=0 (modifyMount)
```

### EXAMPLE:

```
# server_mount server_2 -p -o accesspolicy=MIXED_COMPAT cifs1 /cifs1
```

### **ACCESSPOLICY=NATIVE (Default Policy for Celerra):**

- Access to files or directories via NFS/FTP is granted if UNIX permissions allow access
- Access to files or directories via CIFS is granted if Windows permissions on the object allows access
- ACLs and UNIX permissions are maintained on all files & folders
- File object permission changes from NFS do not affect CIFS permissions or access, and permission changes from CIFS does not affect UNIX permissions or access (each protocol is completely independent of the other)
- Users are only checked for access against the protocol being used to access the Celerra

### **ACCESSPOLICY=NT:**

**Note:** Windows access rights are controlled for each network object by Security Descriptors, which contain information about the object's owner & primary group, and also the Access Control List DACL (Discretionary ACL) & SACL. Each Access Control Entry (ACE) defines an access right for a User or Group specified in the ACL list.

- Access to files or directories via NFS/FTP is granted if both UNIX & CIFS permissions allow access
- Access to files or directories via CIFS is granted if Windows permissions on the object allows access (Unix perms are not checked)
- ACLs and UNIX permissions are maintained on all files & folders
- File object permission changes from NFS do not affect CIFS permissions or access, and permission changes from CIFS does not affect UNIX permissions or access (each protocol is completely independent of the other)
- Only UNIX users are checked against both NFS & CIFS protocols when accessing file system objects on the Celerra

### **ACCESSPOLICY=UNIX:**

**Note:** UNIX access rights are referred to as mode bits that define access for every file system object. These 'mode bits' are represented by bit strings that represent access mode/privilege granted to each User-owner, Group, Other users (rwx triplet)

- Access to files or directories via NFS/FTP is granted if UNIX permissions allow access (CIFS perms are not checked)
- Access to files or directories via CIFS is granted only if both NFS & Windows permissions on the object allows access
- ACLs and UNIX permissions are maintained on all files & folders
- File object permission changes from NFS do not affect CIFS permissions or access, and permission changes from CIFS does not affect UNIX permissions or access (each protocol is completely independent of the other)
- Only CIFS users are checked against both NFS & CIFS protocols when accessing file system objects on the Celerra

### **ACCESSPOLICY=SECURE:**

- Access to files or directories via NFS/FTP or CIFS is granted only if UNIX & CIFS permissions allow access
- ACLs and UNIX permissions are maintained on all files & folders
- File object permission changes from NFS do not affect CIFS permissions or access, and permission changes from CIFS does not affect UNIX permissions or access (each protocol is completely independent of the other)
- Both CIFS & NFS users are checked against CIFS & NFS protocols before access to files is granted
- Highest level of access security

### **CIFS & UNIX INHERITANCE RULES FOR NT, UNIX, NATIVE, SECURE:**

- When CIFS client creates file, the ACL is inherited from the directory (if there is one)
- When CIFS client creates file, Unix permissions are determined by umask of Share from server\_export
- When NFS client creates file, ACL is inherited from the directory (if there is one)
- When NFS client creates file, UNIX permissions are determined by User's umask

**Note:** Umask octal value 644 for files, 755 for directories. Celerra default umask 022 (full access to User/Owner, Read & directory search to Group & Others)

### **ACCESSPOLICY =MIXED (NAS 5.4):**

- \* Access to a file or directory through either NFS/FTP or CIFS is always determined by the ACL.
- \* Both Windows ACLs and UNIX mode bit permissions are maintained for every file and directory.
- \* ACLs for files or directories are created based on the NFS or CIFS protocol that last changed or set the permissions (i.e., if NFS client changes or sets perms on file/folder, the ACL is rebuilt based on UNIX mode bits. If CIFS client changes or sets perms on file/folder, the ACL is built based on normal Windows perms).
- \* In all cases, the ACL is used to control file or directory access regardless of whether the Client is using the NFS or CIFS protocols.
- \* ACL permissions are more granular than UNIX mode bits, consequently not all permissions set in an ACL can be translated to UNIX mode bits. In some cases the mode bits may show more permissions than a user actually has.
- \* Celerra uses default CIFS User's Windows Group as the Unix primary group when assigning group attributes to created objects

### **MIXED ACCESSPOLICY TRANSLATION OF UNIX MODE BITS INTO ACLS:**

- (3) ACL entries are created based on the UNIX mode bits of Owner, Group, Other: File Owner, Group, Everyone
- Delete/Change permissions & Take Ownership are set for the Owner but not for Everyone or other Groups

### **MIXED ACCESSPOLICY TRANSLATION OF ACLS INTO UNIX MODE BITS:**

- Windows ACL Deny option is not translated, but the Allow ACL option is (same for both MIXED & MIXED\_COMPAT)
- UNIX Owner mode bits are built from Owner entry, the ACE of the file/directory Owner, and the Everyone Group ACE
- UNIX Group mode bits are built from Group entry, the ACE of the primary group Owner, and the Everyone Group ACE
- UNIX Other mode bits are built from Other ACE, all ACE's different from User/Group, and the Owner/Group ACE

## **CIFS & UNIX INHERITANCE RULES FOR MIXED & MIXED COMPAT ACCESSPOLICY:**

- When CIFS client creates file, if inheritance flag is set, and object's parent has an ACL, the file object will inherit the ACL, and UNIX Mode bit permissions will be created based on the ACL translation.
- When CIFS client creates file, and the parent directory does not have an ACL, then Unix permissions are set according to umask values—644 Octal for Files and 755 Octal for directories
- When NFS client creates file, Unix mode bits are based on umask value and ACLs are created based on the Unix mode bit translation

## **ACCESSPOLICY =MIXED COMPAT (NAS 5.4):**

- \* Access to a file or directory through NFS/FTP/CIFS is determined by which protocol (NFS or CIFS) last set the permissions
- \* Both Windows ACLs and UNIX mode bit permissions are maintained for every file and directory.
- \* If a file or directory object had its permissions set from a CIFS client then access is determined by the ACL (aka, EXPLICIT ACL).
- \* If a file or directory object had its permissions set from a Unix client then UNIX mode bits dictate access (An ACL is still created but is not used for access checking).

**Note:** In otherwords, the last protocol that has changed or set permissions on a file or folder is the protocol that is used to govern access-checking of the object

\* ACL permissions are more granular than UNIX mode bits, consequently not all permissions set in an ACL can be translated to UNIX mode bits. In some cases the mode bits may show more permissions than a user actually has.

## **NFS ACCESS RIGHTS FOR MIXED COMPAT DEPENDS ON WHETHER DIRECTORY PERMS WERE SET BY UNIX OR CIFS:**

If Directory Set with UNIX Mode bits

If File Set with EXPLICIT Windows ACL

Then NFS User will ignore file ACL and exercise access rights based on UNIX directory Mode bits

---

If Directory Set with EXPLICIT Windows ACLs

If File Set with UNIX Mode bits

Then NFS User would be checked for access rights based on Windows directory ACL

## **MIXED COMPAT ACCESSPOLICY TRANSLATION OF ACLS INTO UNIX MODE BITS:**

- Windows ACL Deny option is not translated, but the Allow ACL option is (same for both MIXED & MIXED\_COMPAT)
- Builds None, Owner, and granted ACEs into Group and Other UNIX mode bits (different than MIXED policy)

## **MIXED COMPAT ACCESSPOLICY TRANSLATION OF UNIX MODE BITS INTO ACLS:**

→ Only two entries are created in the ACL: Owner & Everyone (MIXED policy builds (3) entries in ACL)

→ An Everyone Group ACE is created from Group mode bits, since other groups are not translated with this policy

→ The Other mode bits are ignored and not used when building the ACL

→ Delete/Change permissions and Take Ownership are set for the File Owner but not for the Everyone group

**Note:** Accesspolicies can be changed on-the-fly by issuing the server\_mount command. For example, if accesspolicy was unix, issue command to change to native and the effects are immediate—no further action required.

## **CELLERRA ACL/PERMISSION CHECKING BEHAVIOR:**

### **USING CHECKACL PARAM TO ENFORCE UNIX PERMISSIONS:**

#### **param cifs acl.checkacl=0**

**Note:** In this mode, only Unix permissions are checked—NT ACL's or SIDs are not checked. This bit is set to 0 by default whenever CIFS is stopped.

#### **param cifs acl.checkacl=1**

**Note:** Both Unix and Windows ACLs are checked. Default value for Celerra. Bit is set to 1 whenever CIFS is started.

### **FILE CREATION:**

→ When NFS client creates file or folder, ACLs are inherited from directory (if there are ACLs), UNIX permissions are determined by User's UMASK

→ When CIFS client creates file or folder, ACLs are inherited from directory and UNI permissions are determined by UMASK value on CIFS share

## **CELLERRA SECURITY MODES:**

Change Security mode on DM using “server\_cifs -a security=”. Default is “**security=nt**” when the CIFS service is started.

### **(3) CELERRA SECURITY MODES:**

**security=nt:** Default mode supports ACL's & Domain Controller Authentication. Requires /etc/passwd, NIS service, or Usrmapper/NTMigrate service to map Unix UID's & GID's to DataMover CIFS “Shares”

**security=unix:** Requires clear-text passwords. CIFS usernames checked against Datamover's “/etc/passwd” file or by NIS service AccessPolicies and NT Domain Controller Authentication do not apply to this mode!!

**security=share:** Read-Write or Read-Only access passwords on shares

**Changing CIFS Security Mode:** Changing from NT Security to Unix Security

1. Stop CIFS Service: \$server\_setup server\_3 -P cifs -o stop
2. Change Security Mode: \$server\_cifs server\_3 -a security=unix
3. Start CIFS Service: \$server\_setup server\_3 -P cifs -o start
4. Verify CIFS: \$server\_cifs server\_3 [Should see “Security Mode = UNIX” listed in output]

**Note:** If CIFS is stopped when you reboot a DataMover, it does not restart automatically. Must restart manually!

## **TYPICAL CIFS CLIENT--SERVER SMB TRAFFIC:**

### **SMB COM NEGOTIATE:**

Client sends message to Server with Dialects it supports

### **SMB COM SESSION SETUP ANDX:**

Client sends Username & credentials to server for verification

### **SMB COM TREE CONNECT ANDX:**

Sends name of disk export share to access.

### **SMB COM OPEN ANDX:**

Name of file to open

### **SMB COM READ:**

Client supplies TID, FID, file offset, and number of bytes to read to the Server—Server responds with the data

### **SMB COM CLOSE:**

Client closes file

### **SMB COM TREE DISCONNECT:** Client disconnects from network resource TID.

**TRANS2 DFS GET REFERRAL:** For DFS Shares, Server will send a referral list of Server & Shares to try for access

**Note:** ANDX client commands can chain multiple I/O commands together in a single request. TRANSACTION & TRANSACTION2 commands support large data transfers (>than 64kb).

### **TRANSACTION/TRANSACTION2:**

**Transaction:** Used as transport for older protocols such as Netlogon, Browsers, etc.

**Transaction2:** Used as transport for DCERPC interfaces such as LSA, SAMR, etc.

**NT Transaction:** Handles Audit info, Quotas, etc. [Not used for large calls]

These functions are used for large CIFS requests, capable of supplying more than 64kb data. Sub-functions of these commands are: Create directory; Find first2; Find next2; Get DFS referral; Query FS infor; Query Path info; Report DFS inconsistency; Set file info; Set path info

**Note:** Find subfunctions do folder searches; Set file info sets ACLs. Broken or orphaned ACLs make these functions take longer

**SMB TRAFFIC:** Clients exchange messages with servers in CIFS environments using Server Message Blocks (SMB)

Tid Field: Represents an authenticated connection to a server for a client 'instance'

Pid Field: Unique identifier for the Client's process id with the Server for the current session

Uid Field: Server assigned number for the Client during the session

**Note:** After authenticating a User, Server provides a Uid identifier that is all for all subsequent requests during the session.

### **INCREASING ENTRIES FOR Fid/Tid/Uid Blocks:**

#### **param cifs.listBlocks=128**

**Note:** Default=64. On certain occasions, clients may generate and not release ID's quick enough so that Celerra runs out of the 64k memory bucket caching for these Identifier Handles. Param allows for 16k FID/TID/UIDs at same time. Typical Server Log entries:  
SMB: 4: No more search handle available  
SMB: 4:increase param cifs.listBlocks (64 now)

**SMB DATA:** Typically contains data to be read or written, directory paths, etc.

**Note:** Prior to NTLM 0.12, all strings encoded using ASCII; With NTLM 0.12 & later, can include UNICODE strings that include file, resource, and user names.

## **CIFS COMMUNICATIONS:**

Traditional legacy Clients used NetBIOS Session Requests to Port 139 [SMB over NetBIOS over TCP]

Windows 2000 Clients use ‘direct hosted’ SMB over TCP

### **SMB & ACCESS MASK ENCODING:**

ACCESS\_MASK is a 32-bit value that contains 'standard', 'specific', & 'generic' rights used to build ACE's [Access Control Entries] on objects. **This is the primary means for granting access to objects in Windows NT.**

Bits 0-15 contain access mask for a specific object type.

Bits 16-23 contain the object's 'standard' access rights.

Bits 24-31 contain an object's generic access rights.

### **STANDARD SMB FILE ATTRIBUTE ENCODING:** Encoded as 16-bit value for message exchanges regarding files.

Read Only File: 0x01  
Hidden File: 0x02  
System File: 0x04  
Volume: 0x08  
Directory: 0x10  
Archive File: 0x20

### **EXTENDED SMB FILE ATTRIBUTE ENCODING:**

ATTR\_ARCHIVE: 0x020 {used to mark files for backup or removal}  
ATTR\_COMPRESSED: 0x800 {file or directory compressed. Files inherit from directory}  
ATTR\_NORMAL: 0x080 {File has no other attributes set}  
ATTR\_HIDDEN: 0x002 {File is hidden}  
ATTR\_READONLY: 0x001 {File is Read Only}  
ATTR\_TEMPORARY: 0x100 {File is temporary}  
ATTR\_DIRECTORY: 0x010 {File is a directory}  
ATTR\_SYSTEM: 0x004 {File part of or used by Operating System}  
WRITE\_THROUGH: 0x80000000 [O/S should write through intermediate cache to access file]  
NO\_BUFFERING: 0x20000000 {Server requested to open file with no intermediate buffering or caching}  
RANDOM\_ACCESS: 0x10000000 {application intends to access the file randomly}  
SEQUENTIAL\_SCAN: 0x08000000 {file to be accessed sequentially--used to optimize caching}  
DELETE\_ON\_CLOSE: 0x04000000 {server should delete file upon close}  
BACKUP\_SEMANTICS: 0x02000000 {File being opened or created for a Backup or Restore operation & should be allowed}

### **I. SMB NEGOTIATION REQUESTS:** Client -- Server interaction

SMB Negotiate dialect request is sent by Client to Server. Server responds with authentication 'dialect' to use.

**USHORT SecurityMode:** This field indicates User or Share security (1 or 0 respectively)

**USHORT MaxBufferSize:** Size of largest message which Client can send to Server

**Challenge/Response Protocol:** Challenge sent to client, response is encrypted with 168-bit session key computer from User's password--Server validates response by performing same computation and matches to cached User's Password.

**MAC (Message Authentication Code):** Messages can be authenticated using MD5 encryption similar to IPSec.

**DES** is a block encryption mode using 7-byte Key and 8-byte Data block [E (K, D) or Extended DES Ex(K,D)]

### **ASCII VALUES:**

First 128 characters are 0x00 -- 0x7F

Remaining 128 characters (0x80 -- 0xFF) map to OEM Character sets, or Code Pages as used in I18N UNICODE

### **II. SMB SESSION SETUP:** Between Client & Server

Primary purpose here is to "logon" the Client, which presents a ***Userid*** for AccountName & ***Password*** for AccountPassword

#### **Typical Client Information Sent:**

USHORT MaxBufferSize (to let Server know what its maximum receive buffer size is)

USHORT PasswordLength (account password size)

STRING AccountName (Account Name used)

STRING PrimaryDomain (Domain)

STRING NativeOS (O/S of Client)

STRING NativeLanMan (Native LAN Manager type)

#### **Typical Server Responses:**

STRING NativeOS (Server's O/S)

STRING NativeLanMan (Server's LAN Manager type)

STRING PrimaryDomain (Server's primary Domain)

If Challenge/Response Authentication used, Client supplies AccountName and AccountPassword

#### **Typical Errors:**

ERRSRV/ERRerror (no NEG\_PROT issued)

ERRSRV/ERRbadpw (password incorrect)

ERRSRV/ERRtoomanyuids (max. number of users per session exceeded)

#### **Typical Services the Client Seeks to Use:**

A: (Disk shares)  
LPT1: (Printer)  
IPC (Named pipe)  
COMM (communications device)  
????? (any type of device)

**User LogOff:**

Uid in SMB header is logged off. Server closes all files opened by this User & invalidates any further requests from that "uid"

**Tree Connects:**

Typically involve the Client's "*Tid*" and syntax to a resource such as `\server\share`

**Tree Disconnects:**

Client's resource connection, Tid, is disconnected from Server & then invalidated for further use. Open files, locks, etc., are released.

File System Information Requests:

TRANS2\_QUERY\_FS\_INFORMATION (FileSystem identified by Tid in SMB header)

SMB\_INFO\_ALLOCATION

SMB\_INFO\_VOLUME

SMB\_QUERY\_FS\_VOLUME\_INFO

SMB\_QUERY\_FS\_SIZE\_INFO

SMB\_QUERY\_FS\_DEVICE\_INFO

SMB\_QUERY\_FS\_ATTRIBUTE\_INFO

**File or Directory Creation/Open Request:**

SMB\_COM\_NT\_CREATE\_ANDX

FILE\_OPEN (open an existing file)

FILE\_CREATE (create file)

FILE\_OPEN\_IF (open if exists, or create if not)

**When Using Security Descriptors (SD's):**

NT\_TRANSACT\_CREATE

ACCESS\_MASK (access desired)

ULONG SecurityDescriptorLength (Length of SD in bytes)

ULONG NameLength (length of file in characters)

STRING Name (Name of the file)

**Client Read Requests:**

USHORT Fid (File Handle)

ULONG Offset (Offset in file to begin the Read operation)

USHORT MaxCount (Max number of bytes to return)

**Client Write Requests:**

USHORT Fid (File Handle)

USHORT WriteMode (Write mode bits)

UCHAR Data (data to write)

**Other Client Operations:**

SMB\_COM\_LOCKING\_ANDX (client file locking requests)

SMB\_CLOSE & TREE\_DISCONNECT (Closing a File & Tree Disconnect)

SMB\_COM\_DELETE (Delete a File)

USHORT Flags [Flags for Copy Operations: bit 0=file; bit1=directory; bit2=copy target as binary (0) or ASCII (1); bit3=copy source; bit4=verify writes; bit5=tree copy operation]

**Checking Directory:**

SMB is used to verify that path exists and is a directory--if not, following error returned to Client

STATUS\_NOT\_A\_DIRECTORY or SMB\_ERR\_BAD\_PATH

**File Attributes expressed by AccessFlags:**

FILE\_READ\_DATA 0x00000001 (data can be read from file)

FILE\_WRITE\_DATA 0x00000002 (data can be written to file)

FILE\_APPEND\_DATA 0x00000004 (data can be appended to file)

DELETE 0x00010000 (file can be deleted)

READ\_CONTROL 0x00020000 (Access Control List & Ownership of file can be read)

WRITE\_DAC 0x00040000 (ACL's & Ownership can be written to file)

**Checking or Setting Security on Files:**

NT\_TRANSACT\_QUERY\_SECURITY\_DESC

NT\_TRANSACT\_SET\_SECURITY\_DESC

**COMMON SMB ERRORS:**

SUCCESS 0 (request was successful)

|        |      |                                                                                             |
|--------|------|---------------------------------------------------------------------------------------------|
| ERRDOS | 0x01 | (Error is from core DOS O/S)                                                                |
| ERRSRV | 0x02 | (Error is from Server Network Manager) {ERRbadpw; ERRinvnetname; ERRbadClient;ERRnosupport} |
| ERRHRD | 0x03 | (Error is from hardware) {ERRdata; ERRbadcmd; ERRnotready--drive not ready; ERRread, etc}   |
| ERRCMD | 0xFF | (Command not in 'SMB' format)                                                               |

## **SPECIFIC SMB ERRORS:**

ERRseek (seek error); ERRbadmedia (unknown media type); ERRread (read fault); ERRgeneral (general failure); ERRsharebufexc (sharing buffer exceeded); ERRbadcmd (unknown command); ERRbadreq (Bad request structure length); ERRbadshare (An Open conflicts with an existing file); ERRlock (lock request conflict); ERRtimeout (operation timed out); ERRbadLogonTime (cannot access server at this time); ERRaccess (Client does not have necessary access rights to perform function)

## **NT 4.0 DIRECTORY LISTING BEHAVIOR:**

NT “DIR” command retrieves 16,384 bytes or 170 files for each SMB “transact2 findnextfile” transaction. Known issue where ‘DIR’ requests can cause 200 ms delays for SMB responses. ACKs are normally sent for every other TCP segment received, or unless the delayed ack time of 200ms is reached.

## **SUPERUSER ACCOUNT [aka root]:**

When accessing remote file systems over NFS, the "root" user's UID is normally mapped to the anonymous user account called "nobody", which has a UID of -2 and a "nogroup" GID of -2 [For non-negative unix systems, these values translate into a 16-bit unsigned representation of 65534].

**Note:** NFS handles requests for Users without valid credentials by mapping them also to the "anonymous" user account

## **USING ANON EXPORT FOR NFS: /etc/exports**

Set NFS access mappings to specific UID's using the "anon=100" syntax when "exporting" file systems for NFS

Set NFS access mappings to specific HOSTS using "access=taco,root=taco"

Further, if you export NFS to a specific HOST as ReadWrite [rw=taco:mountpoint,access=100] then all others will mount RO

## **CIFS FILE SYSTEM SECURITY:**

SD's describe Access Rights on NT Objects—files, directories, etc. SD's contain ACL's. ACL's are constructed from ACE's, which also describe umask permissions. Celerra uses the Domain Controller to compare a User's SAT against local SD information for CIFS objects, such as Files and Directories. Users obtain access based on their SAT v. SD of an object. All Directories and Files have Security Descriptors, which define ownership and Discretionary Access Control List information.

**EXAMPLE:** NT Folder “Shared” contains two ACE's. (1) for Everyone [SID S-1-1] READ-ONLY and (1) for Administrators [SID S-1-5-32-544] FULL CONTROL. These ACE's makeup the ACL list and resultant Security Descriptor Information [SD]

## **SEQUENCE OF EVENTS WHEN STARTING CIFS SERVICE & READING CIFS OUTPUT:**

1. When starting CIFS, server Queries WINS using combination of Broadcasts and directed IP communication
2. Receives Listing of IP's for Domain Controllers from WINS Server or Broadcasts & writes them initially as \*SMBSERVER [DMs query WINS & broadcasts every 15 minutes thereafter to produce a 'refreshed' DC list]
3. Server tries to connect to each IP address using “NBTSTAT -A 192.10.2.2”
4. Receives DC reply, and \*SMBSERVER becomes ‘DC=BIWMST-DC3 (167.150.37.3) R=1 T=437382ms S=-1,c0000001/-1’
5. When clients attempt to connect to a “SHARED” CIFS File System, a client negotiate message is sent to DC and an 8-byte Challenge key is returned. Client then uses Session Setup to authenticate to DC using Netlogon IPC\$

Usermapper[2]=[139.21.204.245] --Brackets indicate the 'Master Usermapper'

>DC indicates an active TCP connection and in new code most likely indicates the DC to which DM is logged onto

ref=163 indicates the number of internal references to an object, with ref=2 being a minimum for a valid DC

RC=3 is also an internal reference count & is fairly meaningless at a practical level

## **NETBIOS ALIASES—NT & WIN2K STYLES:**

**Purpose:** Allow the Data Mover to have more than one name [besides the Default NetBIOS name] for the same IP Address, similar to Windows O/S. Clients connect to Aliases via SMB by name. Aliases are registered with Browse Service and WINS but not AD or DNS.

--Max. (10) Netbios names or Aliases per CIFS Service Interface, with max. (29) CIFS Services per DataMover, 15 characters length  
--alias name cannot begin with @ or – characters or include white space or tab characters, or any of the following characters:

/ \ : ; , = \* + | [ ] ? < > “

--Aliases share Localgroups & Shares of Netbios Primary, but do not need NT Machine accounts

--Primary 'default' NetBIOS name still used for Domain Controller registration

--Aliases are registered in WINS and BROWSE Services, therefore must be unique on the Network [Not registered with DNS/AD]

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
--Default Netbios name is the primary machinename used for DC Registration, and maintains ownership of “Shares”, “Localgroups”, and “Configuration Settings” [i.e., Domain Controllers communicate only with primary NetBIOS names]

### **COMPNAME ALIAS EXAMPLE:**

```
# server_cifs server_2 -a compname=mary, domain=mouse.com, interface=ana0, alias=twodoors
server_2 : done
# server_cifs server_2 -Join compname=mary, domain=mouse.com, admin=backdoor
server_2 : Enter Password:*****
# server_cifs server_2
CIFS Server MARY[MOUSE] RC=2
Alias(es): TWODOORS
Full computer name=mary.mouse.com realm=MOUSE.COM
# cat netd
cifs add compname=MARY domain=MOUSE.COM interface=ana0,alias=twodoors
```

## **CIFS Configuration & Troubleshooting:**

### **Fundamental Concepts for Mixed Unix and NT Environments:**

**I. Mounting File Systems:** Mounting determines File Locking [Nolock/Wlock/RWlock] and File System Access Policies [accesspolicy=NATIVE;=UNIX; =NT; =SECURE] for CIFS. Unix locking is ‘advisory’ and transparent for NFS.

#### **Mounting File Systems for NFS or CIFS:**

a.) **Mount Command for NFS File System:** \$server\_mount server\_5 -perm -o rw fs1 /fs1 [Default=perm, RW]

b.) **Mount Command for CIFS File System:** \$server\_mount server\_5 -p -o accesspolicy=nt,rwlock fs1 /mount1  
[Default: accesspolicy=NATIVE]

**II. Exporting File Systems:** Exporting determines accesspolicies for NFS--with options like “anon=0; anon=uid; access=; rw=; root=” specified for “hosts” or “netgroups” during the \$server\_export command.

**Note:** NFS accesspolicies are based on “hosts”

#### **Exporting File Systems for NFS or CIFS:**

a.) **Export Command for NFS:** \$server\_export server\_5 -p -o anon=0 /fs1

b.) **Export Command for CIFS:** \$server\_export server\_5 -P cifs -n sharename /mount1

c.) **Unexporting:** \$server\_export server\_5 -u -p /fs1 [path]

**Note:** NFS Exporting defines Unix access policies, such as “anon=0”, “anon=uid”; “root=”; “access=”; “rw=”{=hostname, netgroup, subnetIP/mask, or IPAddr/mask}. If options are not specified, then the Default is “**No Hosts Root Access**”. CIFS Exporting is the Windows equivalent of creating “Shares”--done after setting up the CIFS service on the DM.

**III. CIFS-Server Security Modes:** When setting up CIFS, unless otherwise specified, the default security mode is “security=nt,dialect=nt1”. To change modes, use the \$server\_cifs -a security=unix [share] command.

**Note:** “security=nt” is the only mode that supports CIFS “accesspolicies” that were defined during the server\_mount command for a File System! Therefore, setting either ‘**Unix**’ or ‘**Share Security**’ renders “accesspolicies” useless.

**IV. NIS Server:** The data mover can be a NIS client, and can retrieve user and group name mappings using ‘yp match’ calls. Be aware of “cifs resolver” param and when you would use default value of 0 and value of 1. Also note that Celerra requires that all NIS entries be in lowercase.

**V. Usrmapper Service:** The “Usermapper Service” runs on Solaris, SCO, or Linux. Primary purpose is to provide mappings for NT Users/Groups with Unix UIDs/GIDs in order to access Celerra CIFS Shares over the network. Data Movers ‘point’ to the Usrmapper Server so that requests for access can be dynamically mapped to the Usrmapper Server’s usrmapper.db and usrmapper.grp.db databases. Usrmapper does not automatically synchronize either NT/PDC or Unix/NIS user-accounts databases!

#### **Usrmapper Provides NT-to-Unix User Mapping [access] to resources, and “authentication” control from an NT PDC!**

**VI. NT Users:** For Windows-based access to Celerra, the CIFS Service must be running, and (1) of the following NT Usernames to UID Mapping Services must be implemented: Password/Group files, NIS Service, Usrmapper Service

## **CIFS SERVICE CONFIGURATION RULES:**

### **CIFS SERVICE RULES:** NetBIOS Name, NT Domain Name, NIC Interface Name, and WINS Service [optional]

#### **\$server\_cifs server\_6 -a netbios=celerradm1, domain=t2dom1, interface=ana0, wins=192.10.2.2**

--CIFS Service consists of a NetBIOS and Domain Name assigned to an Interface Port

--Can Configure Multiple “CIFS Services” on a single DataMover--however, first “Cifs Service Netbios” name assigned becomes the “Default” name, which Windows Networking lists in Server Manager, User Manager, Network Neighborhood, or Explorer [i.e., when connecting to ‘Shares’ using UNC naming or Explorer: <\\celerradm1\ntshare1>]. It also becomes the “Default” security mode.

## **CIFS NETBIOS NAMING RULES:**

--You can have a very large number of IP Configurations and Interface Names given to a DataMover for both Single or Multiple NIC Ports. You assign Netbios Names to physical “ports” [ana0,ana3,etc], not to Interface names [dmzana1,dmzana2,etc]

--The same NETBIOS name can be used on ‘multiple’ NIC ports using different IP configurations

--Only one NETBIOS name can be the “default” name for a given NIC Interface [i.e., \$server\_cifs server\_6 output]

**Note:** Additional IP Configurations and different Interface Names can be created using the \$server\_ifconfig server\_6 -c -D.

Additional CIFS Services [i.e., netbios names] can then be created for each different Interface name. Therefore, a DataMover can be configured to belong to multiple NT Domains via the use and assignment of the ‘CIFS Service’ to different Interface Ports.

--Multiple netbios names can also be assigned to a Single Interface port, but only if additional IP configurations have been created using the \$server\_ifconfig server\_6 -c -D command

## **CELERRA WINS SERVICE:**

**Intro:** Required only for Data Movers that are member servers of NT 4.0 Domains. As long as the DM Wins service is defined, a CIFS Server can operate on a network that is different from Domain Controllers. The DM can also function without a WINS Server if the Domain Controller is on the same local network as the DataMover—it does this through the use of “Broadcasts” to locate Wins and also to request a list of 1C records from the Wins Server. Otherwise, if Wins servers are located on separate networks, an entry for Wins should be added to the DM so that direct IP registration and queries can be made to the Wins Server. Disable unused interfaces to prevent registration with WINS servers.

## **PORTS USED BY WINDOWS FOR WINS SERVICES:**

|                                       | UDP | TCP |
|---------------------------------------|-----|-----|
| WINS Manager                          | --- | 135 |
| WINS NetBios over TCP/IP name service | 137 | --- |
| WINS Proxy                            | 137 | --- |
| WINS Registration                     | --- | 137 |
| WINS Replication                      | --- | 42  |

### **Example One:**

#### **CIFS LOGON USING WINS/BROADCAST:**

##### **DOMAIN SMSTEST**

**DC=PLNT025(173.55.21.10) ref=1 time=96933 ms**

CIFS Server (Default) NAS2[SMSTEST]

Comment='EMC-SNAS:T2.2.35.4'

if=ana2 l=173.55.25.200 b=173.55.25.255 mac=0:0:d1:20:1:e3

**Note:** The pointer “>” & SID won’t show up until ‘touched’ by a User account, such as Administrator, from Server Manager, Explorer, or Network Neighborhood if this is the first time that CIFS has been setup on this Server

### **Example Two:**

This Output Also Shows that CIFS has Logged Into the Domain [via WINS or Broadcasts]—Only change is that if this is the first CIFS Service set up on the Celerra, you may see a slightly different output:

##### **DOMAIN SMSTEST**

**DC=\*SMBSERVER(173.55.21.20)ref=1 time=0 ms**

### **Example Three:**

This Output Shows CIFS Logged Into the Domain & also “accessed” by a User Account:

##### **DOMAIN SMSTEST**

**SID=S-1-5-6f2a2ed4-7dd53e1-24fe7268-ffffffffff** [SID of DOMAIN “SMSTEST”]

**>DC=PLNT025(173.55.21.10) ref=2 time=0 ms**

CIFS Server (Default) NAS2[SMSTEST]

Comment='EMC-SNAS:T2.2.35.4'

if=ana3 l=173.55.21.200 b=173.55.21.255 mac=0:0:d1:20:1:e4

**Note:** The SID is obtained once a valid User on the NT Domain is successfully authenticated to the Celerra—as indicated above, this could be when accessed from either Network Neighborhood, Explorer, or Server Manager. The authentication mechanism is usually provided automatically by the Usrmapper Service or from some form of Unix Passwd/Group file checking.

## **TROUBLESHOOTING CELERRA WINS:**

**Intro:** If Celerra CIFS servers are located on different logical networks than the Domain Controllers, then manual configuration of WINS addresses for CIFS is required. By design, Celerra does not work the same as a native Windows Server in the registration and renewal of its WINS netbios name. Celerra Servers register their netbios name with WINS upon each startup of CIFS, and then every 15 minutes thereafter, using broadcast and direct IP [if configured with an IP address of the WINS server]. Also at startup and every 15 minutes, Celerra queries and updates its DC list. Stopping the CIFS service will de-register the Netbios names, and starting the CIFS service will also initiate a Wins Registration and Query for 1C list.

**param cifs wins.Refresh=384**

**Note:** Default behavior is to register our Workstation and Server services with the WINS server every 15 minutes via direct IP [when configured] and broadcasts. Celerra also sends query to WINS servers by direct IP and broadcast to obtain new list of <1c> records

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
[Domain Controllers]. 384 is hex for 900 seconds or 15 minutes. Also note that for multiple Wins Server entries [in CIFS], the Data Mover registers and queries only the first WINS Server on its list [it will failover to the next Server if it does not receive a reply]

**Example from Server Log with ‘Traces=2’ enabled:**

```
2003-11-13 17:15:48: SMB: 4: Id:0x298 Send C-Registration for WINS<00> to 192.10.20.10 [direct IP for workstation service]
2003-11-13 17:15:48: SMB: 4: Id:0x299 Send C-Registration for WINS<20> to 192.10.20.10 [direct IP for server service]
2003-11-13 17:15:48: SMB: 4: Id:0x29a Send C-Registration for WINS<00> to 192.10.255.254 [broadcast....]
2003-11-13 17:15:48: SMB: 4: Id:0x29b Send C-Registration for WINS<20> to 192.10.255.254 [broadcast....]
2003-11-13 17:15:48: SMB: 4: Id:0x29c Send C-Query for SEINFELD<1c> to 192.10.20.10
2003-11-13 17:15:48: SMB: 4: Id:0x29d Send C-Query for SEINFELD<1c> to 192.10.255.254
2003-11-13 17:15:48: SMB: 4: Id:0x298 WINS<00> Registration by 192.10.20.10
2003-11-13 17:15:48: SMB: 4: Id:0x299 WINS<20> Registration by 192.10.20.10
2003-11-13 17:15:48: SMB: 4: Domain=SEINFELD<1c>, RR Flags=0x8000, Host=192.10.20.10
```

**DISABLING DNS & USING WINS WHEN BUILDING DC LIST FOR 2K3 SERVERS:**

# server\_param server\_2 -f cifs -info NTsec.getDCfromADServices -v

```
server_2 :
name      = NTsec.getDCfromADServices
facility_name = cifs
default_value = 1
current_value = 1
configured_value =
user_action = restart Service
change_effective = restart Service
range      = (0,1)
description = Use DNS to get list of Domain Controllers
detailed_description
```

When the param is set to False (0), WINS are used to build the list which is sorted using response time (SamLogon) and subnet info.

**TROUBLE GETTING RID OF INVALID DC’s FROM CIFS LIST:**

→AR87585, emc218088, NAS 5.6 and above

**Typical Message:**

Discovered from: WINS

DCd660a80c 10.110.255.228[WRO] INVALID DC=0xd660a804 shd=0x00000000 name=\*SMBSERVER

**Workaround→Set following parameter to 60 seconds to remove non-advertised DCs from Server DC Listing:**

# server\_param server\_2 -facility cifs -info DC.autoRemovalSeconds -verbose

```
server_2 :
name      = DC.autoRemovalSeconds
facility_name = cifs
default_value = 604800
current_value = 604800
configured_value =
user_action = none
change_effective = immediate
range      = (0,4294967295)
description = Time to remove obsolete Domain Controllers (no longer announced) from domain's DC list
detailed_description
```

All DCs no longer announced by WINS/DNS are automatically removed from domain's DC list after this time specified in seconds. 0 value disable to removal process.

**WINS INFORMATION & DEFINITIONS:**

Windows Internet Naming Service (WINS) is used to resolve a Netbios name on a network to its respective IP Address, and also to Register and release Netbios names with the WINS Servers. Windows Explorer, or DOS prompt "net" commands will invoke the NBT interface, using NetBIOS over TCP using Port 137. The idea behind WINS is that each Client will be configured with the WINS Server IP Addresses so that direct communication can occur, as opposed to network broadcasts.

**WINS NAME RESOLUTION MECHANISMS:**

**B-node:** Uses UDP broadcast datagrams to Register and perform WINS resolution activities

**P-node:** Called Peer-to-Peer because it uses direct IP to a defined WINS Server IP address

**M-node:** Mixed mode—Broadcasts used first, then P-node directed IP

**Note:** This is the default mode of registration by Data Mover for its Workstation and Server Services

**H-node:** Hybrid mode—P-node used first, then Broadcasts

**NETBIOS SERVICE NAMES:**

00 Redirector or Workstation Service for a netbios name

20 Server Service for a netbios name

1B PDC of a Domain (Primary Domain Controller)

1C Domain Controller records

1D Domain Master Browser service

1E Browser service on other systems than Domain Master

**WINS DATABASE PROPERTIES FOR REGISTERED NETBIOS RECORDS:**

**Renewal Interval:** Timeframe in which a client must re-register its name before being "released" by WINS server

**Extinction Interval:** A period of time after the Renewal Interval has expired in which WINS server will wait for a "released" record before it tombstones that record

**Extinction Timeout:** A period of time after Extinction Interval in which WINS will remove the tombstoned entry from its database

**Verification Interval:** Time period that an active record exists before being verified

**TROUBLESHOOTING WINS:**

**1.) Increase Server Logging for WINS**

**`$server_config server_5 "param cifs wins.Traces=2"`** → Starts debug logging

**Note:** With 5.6, you may need to use `$server_log server_x -i` option to see trace output in the server log

**`$server_config server_5 "param cifs wins.Traces=0"`** → Stops debug logging

**Note:** Sets debug logging for Wins, check Server Log for logging information. Set 'Traces=0' to stop WINS debug logging.

**2.) Query each defined WINS Server to see if records are registered for CIFS Servers:**

**`#server_config server_5 -v "wins addr=192.10.20.10 name=server"`**

1068755669: SMB: 4: SERVER <00> 4000 (Unique NetBios Name, M Node) Addr=192.10.0.109

**`# .server_config server_5 -v "wins addr=192.10.20.10 name=**"`** → Dumps records from WINS server

**Note:** Command can cause panics prior to 5.2.16.x—AR44550

**3.) Examine Server Log for Registrations, Queries, and Replies from Wins:**

**DEBUG LOG MESSAGES:**

2003-11-20 23:24:03: SMB: 4: Id:0x5b17 Send C-Registration for AP3CDM3<00> to 3.130.163.255  
2003-11-20 23:24:03: SMB: 4: Id:0x5b18 Send C-Registration for AP3CDM3<20> to 3.130.163.255  
2005-01-07 23:41:05: SMB: 4: Id:0xc03 Send C-Registration for CELPFS52<00> to 160.5.56.41  
2005-01-07 23:41:05: SMB: 4: Id:0xc04 Send C-Registration for CELPFS52<20> to 160.5.56.41  
2005-01-07 23:41:05: SMB: 4: Id:0xc05 Send C-Registration for MMBINTD1<00> to 160.5.56.41  
2005-01-07 23:41:05: SMB: 4: Id:0xc06 Send C-Query for MMBINTD1<1c> to 160.5.56.41  
2005-01-07 23:41:05: SMB: 4: Id:0xc03 CELPFS52<00> Registration by 160.5.56.41  
2005-01-07 23:41:05: SMB: 4: Domain=MMBINTD1<1c>, RR Flags=0x8000, Host=160.5.56.18  
2005-01-07 23:41:05: SMB: 4: Domain=MMBINTD1<1c>, RR Flags=0x8000, Host=160.3.4.16  
2005-01-07 23:41:05: SMB: 4: Domain=MMBINTD1<1c>, RR Flags=0x8000, Host=160.21.32.21  
2005-01-07 23:41:05: SMB: 4: Domain=MMBINTD1<1c>, RR Flags=0x8000, Host=160.5.56.18  
2005-01-07 23:41:05: SMB: 4: Domain=MMBINTD1<1c>, RR Flags=0x8000, Host=160.128.67.246  
2005-01-07 23:41:05: SMB: 4: Domain=MMBINTD1<1c>, RR Flags=0x8000, Host=161.2.56.237  
2005-01-07 23:41:05: SMB: 4: Domain=MMBINTD1<1c>, RR Flags=0x8000, Host=160.3.4.142  
2005-01-07 23:41:05: SMB: 4: Domain=MMBINTD1<1c>, RR Flags=0x8000, Host=160.129.14.56  
2005-01-07 23:41:05: SMB: 4: Domain=MMBINTD1<1c>, RR Flags=0x8000, Host=160.129.14.33  
2005-01-07 23:41:05: SMB: 4: Domain=MMBINTD1<1c>, RR Flags=0x8000, Host=160.5.56.17  
2005-01-07 23:41:05: SMB: 4: Id:0xc04 CELPFS52<20> Registration by 160.5.56.41

**WINS BUG—AR55681:**

If multiple WINS servers are assigned & written to netd file as a separate line, instead of to each individual CIFS Server, when communications are lost to the first WINS server on the list, DART will not cycle to the next Wins Server. Workaround would be to add WINS to each separate Netbios Server: [server\_cifs server\_2 -add

netbios=XXX, domain=YYY, interface=ZZZ, wins=x.x.x.y.y.y]—DART will then cycle to the next server on the list. Issue resolved with NAS 5.5.1.0, 5.4.10.4, & 5.3.16.0.

**4.) WINS configuration for VDM CIFS Servers:**

Our documentation does not clearly discuss how Wins should be configured with VDM's. The recommended approach would be to add WINS to the VDM and verify in the vdm.cfg file:

**`$ server_cifs vdm1 -add wins=10.241.168.92`**

**Common WINS Suffixes:**

1Ch = Group Entry for list of Domain Controllers

1B = Address of PDC or in Windows 2000, address of PDC Emulator

## **HOW TO IDENTIFY & RESOLVE WINS ISSUES WITH CELERRA CIFS SERVER:**

|                    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
|--------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Situation:</b>  | WINS & CIFS Servers on different Networks, with Celerra using direct IP communication with WINS                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| <b>Problem(s):</b> | CIFS Server cannot authenticate to NT Domain when using WINS on different Network than CIFS Server<br>Missing Entries on WINS Server for NT PDC/BDC leads to loss of connectivity with CIFS Server service<br>Incorrect IP configuration on WINS Server could lead to loss of connectivity with CIFS server service                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| <b>Causes:</b>     | PDC/WINS Server missing WINS IP address in the IP Configuration Tab--CIFS server will not be able to communicate, or authenticate CIFS Users, with the NT Domain Controller.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| <b>Actions:</b>    | Missing entries in the WINS database service that defines the PDC/BDC & NT DOMAIN<br>Ensure the WINS Server has its own IP information in the following location: Rightclick "Network Neighborhood">>PROPERTIES>PROTOCOLS>TCP/IP>PROPERTIES>WINS ADDRESS<br>[verify IP address entries are present and reboot WINS server if entries are added or changed]                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| <b>Actions:</b>    | Verify WINS Database Entries by opening WINS MANAGER>"Mappings">Show Database<br>You should see entries "similar" to the following:<br><b>NT DOMAIN NAME REGISTRATIONS:</b> NT Domain Name in this Example is "SMSTEST"<br>SMSTEST 1Ch [DC list for Domain with IP of PDC]<br>SMSTEST 1Eh [Browser Service registration with IP of PDC]<br>SMSTEST 1Bh [Domain Master Browser with IP of PDC]<br>SMSTEST 00h [Workstation Service with IP of PDC]<br><b>NT COMPUTER NAME REGISTRATIONS</b> for DC: NT DC in this Example is "PLNT025"<br>PLNT025 00h [Workstation Service]<br>PLNT025 03h [Messenger Service]<br>PLNT025 20h [Critical entry--defines the Server Service of the DC, to which the Celerra communicates]<br><b>Actions:</b> CORRECTING WINS DATABASE ENTRIES: You may need to add "STATIC" MAPPINGS for the NT DOMAIN CONTROLLER and NT DOMAIN to the WINS Server Database by opening WINS MANAGER>MAPPINGS>STATIC MAPPINGS>Add DOMAIN "SMSTEST" & IP Address of PDC<br>WINS MANAGER>MAPPINGS>STATIC MAPPINGS>Add NT PDC "PLNT025" & IP Address |

**Cautionary:** These fixes are not all-inclusive--there are Network Switch/Router configurations that may also prevent the CIFS Service from communicating properly with the Domain Controller across Subnets.

## **TROUBLESHOOTING WINS ENTRIES FROM DM:**

**#.server\_config server\_3 -v "wins addr=131.99.75.60 name=compucom type=0x1c"** [Wins Records for DCs]

**Note:** Plug in different WINS records for more info; 0x1b; 0x1d, etc.

## **EXAMPLE OF WINS RECORDS FOR DC's:**

**# .server\_config server\_2 -v "wins addr=172.19.3.254 name=network type=0x1c"**

1050334973: SMB: 4: NETWORK <1c> 8000 (Group NetBios Name, B Node) Addr=172.19.3.252

1050334973: SMB: 4: NETWORK <1c> 8000 (Group NetBios Name, B Node) Addr=172.19.3.253

## **TOMBSTONED ENTRIES FOR WINS:**

Server Log error SMB: 3: WINS ERROR \*<00>: No answer from 10.8.1.137

**Note:** Server recognizes that a Netbios name is tombstoned by the specified Wins server IP address. This message may also occur if a DC is not running the messenger service.

## **TROUBLESHOOTING TOMBSTONED ENTRIES:**

--Verify that Wins Servers are replicating their databases correctly

**Note:** In the case of multiple Wins entries for CIFS, the Server only registers and queries the Wins server at the first IP address on its list and relies on Wins to replicate its entries to other Wins Servers.

--Use Static Wins entries, if necessary, in order to keep Data Mover names registered in Wins [Adding the IP Address and name of the DM will create a Workstation <00>and Server Service <20>record

--Use nbtstat –a command from DC to register DM netbios name on a temporary basis—valid only for the current logon session

**c:>nbtstat -a 192.168.0.103** [Data Mover IP]

## **SET FOLLOWING ENTRY IN NETD FILE AND REBOOT SERVER:**

**1. #vi /nas/server/slot\_2/netd**

pdc trace=1

cifs start

## **2. Reboot Data Mover and Observe Server Log:**

2005-01-04 19:46:34: SMB: 4: >DC=MICKEY(10.241.169.16) R=5 T=1 ms S=0,1/-1

2005-01-04 19:46:34: SMB: 4: last message repeated 2 times

41daf29a DCd52ae0c MICKEY[mouse.com] 4 setCurrentDC Ctx=daee204 Old=d52ae04 New=0

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
2005-01-04 19:46:34: SMB: 4: DCT=1104867994546 dom=0xd37ff84, BrowseBroadcastsDCcount=1  
2005-01-04 19:46:34: SMB: 4: DCT=1104867994546 localMaster setDC=MICKEY respTime=0 mode=0 BROADCAST  
2005-01-04 19:46:34: SMB: 4: DCT=1104867994546 preconnectDC addDC  
2005-01-04 19:46:34: SMB: 4: DCT=1104867994546 getdc184 setDC=MICKEY respTime=0 mode=3 UNKNOWN ??  
2005-01-04 19:46:34: SMB: 4: DCT=1104867994546 preconnectDC addDC  
41daf29a DCd52ae0c MICKEY[MOUSE] 5 W2KgetInitialDC: DC=MICKEY found after refresh domain=mouse.com (0,0)

### **CREATING ‘NON-GLOBAL’ SHARES—i.e., for specific Netbios or Compcodes only:**

**\$ server\_export server\_3 -P cifs -o netbios=homely -n HomeComingQueen /mnt13**

### **DELETING SPECIFIC NETBIOS SERVER SHARE:**

**\$ server\_export server\_3 -P cifs -u -p -o netbios=biwusrmap -n Projects /fs13**

**Note:** This immediately removes entry from export.shares file & /.etc/shares/BIWUSRMAP folder

### **LOSS OF ACCESS WHEN USERS HAVE RESTRICTED MACHINE LOGONS:**

**Note:** In situations where DM is not Joined to same Domain as Users, and where User accounts have Restricted Logon rights, users may not be able to map to the Share

“Mapped drive could not be created...An extended error has occurred”

SMB: 4: authenticate NA\useaorap S=19 SamLogonInvalidReply

Fixed in NAS 4.2.12.0 & 5.0.12.0 by querying Client name from SMB thread

### **SHAREDB:**

Shares are managed via the ShareDB, a relatively new part of code since 4.0 [in support of Windows 2000]

ShareDB relative to Share permissions are stored in the hidden .etc folder in NAS 4.2 and has been moved to the roots in NAS 5.1

With Nas 5.x, ShareDB directly manages and handles share Security descriptors

### **HOW IT IS POSSIBLE TO LOSE/CHANGE SHARE PERMS ON CELERRA:**

#### **Background:**

By default, if Shares are created on a file system, the hidden .etc directory in NAS 4.2 will contain the name of the share. If there are no share names in 4.2, then it's likely that the default Everyone FC perm has been applied, or it could mean that inheritance for the .etc directory has been set.

#### **When Share Perms Can be Lost or Changed:**

1. If the hidden .etc directory is exposed, that is, if the upper mountpoint of a file system is ‘Shared’, it is possible for someone to inadvertently deleted this folder. If anything other than the default Share perm of ‘Everyone FC’ was applied, the perms would be lost and after a data mover reboot, the .etc folder would be recreated with Everyone FC for all Shares on that file system.

2. If the inheritance bit is set on the .etc folder, it is possible that an Administrator could change or apply Share/NTFS permissions on the Upper Mountpoint, which could also then change the Share permissions to those of the underlying NTFS permissions of the upper mountpoint. In the latter case, the application of a new ACL may actually remove the existing ACL—acts as a replace vs. an edit.

#### **Resolution:**

There is no engineering solution for this problem as the system is operating as designed.

--Always create nested Shares after creating the initial mountpoint Share so as not to expose the .etc folder

--Never apply permissions on the Upper Mountpoint of the file system

### **LOCATION OF CIFS SHARES ON CELERRA—NAS 4.2:**

#### **CIFS Shares are now stored in the following locations:**

1. /nas/server/slot\_x/export.shares
2. /nas/rootfs/slot\_3/.etc/shares/@global [If Shares are not created specifically for one netbios name, all show up here]
3. /nas/rootfs/slot\_3/.etc/shares/netbios\_name [Each compname will have a default folder created in the “shares” directory, which in turn will contain “Shares” if exported to the specific compname only]

#### **Folders Stored in /.etc/shares:**

/nas/rootfs/slot\_3/.etc/shares:

@global: Folder contains HOMEDIR & all global Shares for the DataMover

@import: No entries unless \$.server\_config server\_x -v “sharedb backup” has been run, then stores shares.bak file in this folder

@system: Contains default System Shares for each Server→C\$, CHECK\$, HOME, IPC\$

/HOMELY: Each “compname” will have its own folder created & populated with Shares if specific to the Compname

### **HOW CIFS SHARES BEHAVE WITH NAS 4.x/5.0:**

#### **CIFS SHARE CREATION WITH GLOBALSHARES=0 SET:**

**param cifs srvmgr.globalShares=0**

## I. Exporting new “share” from CommandLine:

**\$server\_export server\_2 -P cifs -n CLI /mnt**

### Results:

- Since not specified in the export command, Share will be created as a “Global” Share and is written to /.etc/shares/@GLOBAL
- Share is immediately written to the “/nas/server/slot\_2/export.shares” file

## II. Exporting new “share” from Windows 2000 “Computer Management” Console:

### Results:

- Becomes a new “Share” for only the compname connected to using ‘Computer Management’ [i.e., will not become a “Global” share for all compnames & is written to [/.etc/shares/BRUCE](#)
- Share is not written to “/nas/server/slot\_2/export.shares” file

**Comment:** Neither Stopping, Starting CIFS, nor Rebooting DataMover will write new “sharename” to “export.shares” file. Though the ‘Share’ is seen in the “server\_export” output, the only way to get this new share into the “export.shares” file is to run the following command:

**\$server\_export server\_2 -r**

**Explanation:** AR28311 states that this behavior is by design in NAS 4.x and above as the CS no longer runs a “server\_mgr” daemon. In the pre-NAS 4.x model, the MMC Console created shares through CS daemon by request to Data Mover, therefore the “shares” were always kept in sync. With NAS 4.x and above, CS only maintains cache information and DART saves new Shares created from MMC Console in its Share Database in the /.etc/shares directory. Also, Shares were removed from Boot.cfg file at the 4.x and above families. Only present workaround is to use \$server\_export server\_x -r to retrieve ShareDB information from Data Mover and update the CS Cache file, “export.shares”.

## III. Exporting new “share” from CommandLine and Specifying Netbios Name:

**\$ server\_export server\_2 -P cifs -o netbios=bruce -n newcli /mnt**

### Results:

- Becomes a new “Share” for only the compname called “Bruce”
- New Share is written immediately to “export.shares” file
- New Share is also written to /.etc/shares/BRUCE [Does not populate the @GLOBAL directory]

## IV. Deleting Share from Commandline and Specific Netbios Server Name:

**\$ server\_export server\_3 -P cifs -u -p -o netbios=biwusrmap -n Projects /fs13**

## V. Exporting new “share” from Celerra WebUI Web Manager Interface:

- Becomes new share for compname to which it was created
- New Share is written immediately to the “export.shares” file
- New Share is written to /.etc/shares/REX [Does not populate the @GLOBAL directory]

## **RECOVERING SHARedb ON SERVERS RUNNING NAS 4.2 OR HIGHER:**

**Problem:** Database corruption occurred, resulting in complete loss of all shares in the /.etc/shares directory, including @system hidden administrative shares.

**Solution:** As long as there is a recoverable SCCS export.shares file, or if the export.shares file in the /nas/server/slot\_x directory is intact, recovery is fairly straightforward:

- 1.) Reboot datamover [This will rebuild the Hidden Administrative shares “C\$, CHECK\$, & IPC\$”]
- 2.) Do #server\_export server\_5 -a [This repopulates the “sharedb” into the proper folders in the /.etc/shares directory]

## **BACKING UP AND RESTORING SHARES FROM ONE DATA MOVER TO ANOTHER:**

### 1. Backup Sharedb on Server:

**# .server\_config server\_5 -v "sharedb backup"**

**Note:** Creates “shares.bak” file in [/.etc/shares/@import](#)

### 2. Copy Sharedb from Slot 5 to Destination Slot:

#cp /nasmcd/quota/slot\_5/.etc/shares/@import/shares.bak /nasmcd/quota/slot\_4/.etc/shares/@import/shares.bak

### 3. Import Sharedb into Slot 4:

**# .server\_config server\_4 -v "sharedb restore=shares.bak"**

### 4. Verify:

**# .server\_config server\_4 -v "share"**

**# .server\_config server\_4 -v "sharedb info"**

## **CIFS SHARE CREATION WITH GLOBALSHARES=1 SET:**

- I. By default, when exporting from Command Line, the CIFS Share will be created as a “Global” Share unless you specify the “Netbios” name in the export options [/.etc/shares/@GLOBAL]

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
II. However, when using the Windows 2000 Computer Management to ‘manage’ the Celerra, all Shares created will become “Global” shares because the param=1.  
III. Finally, if specified in the export command, will create the Share as a specific Share for the Netbios name [/.etc/shares/BRUCE]

## **COMMON CIFS ERROR/INFORMATION MESSAGES IN SERVER LOG:**

**Server Log:** 2002-04-25 14:59:44: SECURITY:3: /etc/group does not exists and NIS not started

**Reason:** Actual message is harmless & means that there is no "group" file located on the /etc directory of the datamover. An entry is made each time a User touches a CIFS Share.

**Resolution:** #touch group #server\_file server\_x -put group group [Error message will no longer be logged]

**Server Log:** NT\_Access\_Credential::RequestFromSID:primary nt group not mapped, use unix primary

**Note:** User is accessing DM but Windows Primary group cannot be mapped.

### **HIDDEN SHARE NOTE:**

If exporting CIFS Shares at the root FS level, the system admin shares are exposed to view [/.etc, c\$, Lost + Found, etc]. There are situations where this may not be desirable. As an alternative, export the File System with a subdirectory, such as “fs04\bruno”—then, when accessing the File System from NT, the “admin” shares will remain “hidden”.

## **DELETING CIFS SERVICE FROM DATAMOVER:**

- Step 1. Stop the CIFS service
- Step 2. Delete CIFS Service: **\$server\_cifs server\_x -d netbios=dm2** [Deletes all three: Netbios, Domain, & Interface Name]
- Step 3. Verify that the netd file has been updated correctly: /nas/server/slot\_x/netd
- Step 4. If permitted, reboot DataMover

## **DELETING CIFS SERVICE USING SINGLE COMMAND:**

**\$server\_setup server\_x -P cifs -o delete** [Removes all entries from "netd" file--much easier way to remove CIFS Service]

## **MIGRATING FROM ONE CELERRA TO ANOTHER--RETAINING ORIGINAL**

**NETBIOS NAMES:** [Within the same NT Domain]

### **PART I. MIGRATION SETUP:**

**ALPHA CELERRA CS0:** 172.24.80.10

**SERVER 4:**

**NETBIOS NAME:** SOURCE [old]

ana0 192.10.2.24

**FILESYSTEM:** origserv /origserv

**CIFS SHARE:** "sourceshare" /origserv [SOURCE]

**USRMAPPER:** Linux CS0 172.24.80.10

**WINS SERVICE:** 192.10.2.2

**NT DOMAIN:** T2DOM1

**MIGRATION NT ACCOUNT:** "T2DOM1\newby"

**Migration Servers:** SOURCE [old] → TARGET using NT 4.0 WORKSTATION [NTWK1DOM1]

**Important Note:** NT User "newby" was added to the LOCAL GROUP called "Administrators" on both Celerra Servers and the NT Workstation using User Manager interface. User "newby" was also given the following (4) specific USER RIGHTS on both Celerra Servers and the NT Workstation: **Backup Files & Directories; Generate Security Audits; Manage Auditing & Security Logs;**

**Restore Files & Directories**

**SOURCE PATH:** \SOURCE\sourceshare\migration

**TARGET PATH:** \TARGET\targetshare\destination

**METHODOLOGY:**

**\SOURCE\sourceshare\Migration\Folder1, 2, 3, 4:**

--Permissions on folder "Migration" set for Administrators & Domain Admins Full Control, and Everyone Group as Read Only.

--Created (4) LOCALGROUPS on CIFS Server "SOURCE", each with (2) Users from NT DOMAIN "T2DOM1".

--Created (4) subfolders on CIFS Server "SOURCE" underneath the Directory called "Migration".

--Assigned localgroup "LOCAL1" [T2DOM1\amato | betty] to "\FOLDER1" giving Ownership to "SOURCE\Administrators" and Full Control Permissions to Administrators | LOCAL1 [localgroup,etc.]

--Assigned localgroup "LOCAL2" [T2DOM1\nasadmin | newby] to "\FOLDER2" giving Ownership to "SOURCE\Administrators" and Full Control Permissions to Administrators | LOCAL2

--Assigned localgroup "LOCAL3" [T2DOM1\bing | bob] to "\FOLDER3" giving Ownership to "SOURCE\Administrators" and Full Control Permissions to Administrators | LOCAL3 & Read Permissions to T2DOM1\Todd1

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
--Assigned localgroup "LOCAL4" [T2DOM1\david\elf] to "\FOLDER4" giving Ownership to "T2DOM1\Administrators" and Full Control Permissions to Administrators | LOCAL4 & Read Permissions to T2DOM1\thm

#### **TARGET\targetshare\Destination:**

--Granted Ownership to Administrators and Full Control Permissions to Administrators & Everyone

#### **MIGRATION PLAN:**

Plan is to "migrate" the "localgroups" information, then all FileSystem data, from the CIFS Server called "SOURCE" on Original Celerra, to the destination Celerra CIFS Server called "TARGET" using the "external IP Address" of the Usrmapper Service on original Celerra. After the "migration", the original Usrmapper Service needs to be stopped, the database files copied over to the new Celerra Control Station, destination CIFS servers pointed to the New Usrmapper IP Address, and the new Usrmapper Primary Service started. The Plan is then to remove the original "SOURCE" server share, then its CIFS Service, and its entry in the WINS Servers database, while leaving the Netbios name "SOURCE" in the NT Domain, and finally to create the CIFS Service for "SOURCE" on the new Celerra Server\_4 on a different IP Address.

#### **PART II. CONDUCTING MIGRATION USING EMCOPY TOOLS:**

a.) Run LGDUP.EXE from NT 4.0 Workstation called "NTWK1DOM1" while logged in as the Migration User Account called "Newby" who has DOMAIN ADMINS membership in addition to the Steps outlined in Part I:

**C:\emcopy>lgdup.exe -v \source \target > lgdumpmigration.txt**

**Note:** Examine output file called "lgdumigration.txt" to see if migration of LOCALGROUPS information was successful. If not, re-verify the setup of "newby's" User Rights on all (3) Servers involved in the migration, as well as its membership as Administrator.

b.) Run EMCOPY.EXE from "NTWK1DOM1":

**C:\emcopy>emcopy.exe \source\sourceshare\migration \target\targetshare\destination /o /a /lg /s /secfix /d > emcopymigrat.txt**

**Note: Use EMCOPY Version 2.04b to conduct the EMCOPY Migration of Data, and User Ownership/Permissions!**

#### **PART III. POST MIGRATION TESTING:**

- a.) Tested access to the "destination" folder on Server "TARGET" as a Domain Administrator User. Verified that all Ownership, Permissions, and LOCALGROUP entries were migrated over identically and successfully from the "SOURCE" Server.
- b.) After migration was verified as successful and complete, saved a copy of the original Localgroups.db file from each Celerra Server involved in the migration [#/home/nasadmin/server\_file server\_4 -get localgroups.db localsource.db, etc.]
- c.) Unmapped all drive shares to "SOURCE" CIFS Server
- d.) Unexported and unmounted Shares/FileSystem from "SOURCE" [\$server\_export server\_4 -P cifs -n sourceshare -p -u]
- e.) Stopped CIFS Service on "SOURCE"
- f.) Stopped Linux Usrmapper Service on 172.24.80.10 and FTP'ed "usrmap.db"; "usrmapgrp.db", and "usrmap.cfg" database files to FTP Server and then pulled down to Linux Usrmapper Server 172.24.80.11 on the new Celerra [Tarred the whole original /var/usrmapper directory, FTP'ed, then untarred on destination Celerra Usrmapper]
- g.) Deleted CIFS Service for "SOURCE" on original Celerra called "ALPHA" [\$server\_cifs server\_4 -d netbios=source]

#### **PART IV. COMPLETING SETUP & IDENTITY CHANGE FOR NEW "SOURCE" SERVER ON CELERRA "BETA":**

- a.) Stopped CIFS Service on Server\_4 on new Celerra called BETA
- b.) Deleted entry for old Usrmapper Server: \$server\_cifs server\_4 -d usermapper=172.24.80.10
- c.) Added entry for new Usrmapper Server: \$server\_cifs server\_4 -a usermapper=172.24.80.11
- d.) Created new IP on ana1, 192.10.3.25, for new "SOURCE" server on Beta Server\_4
- e.) Created new CIFS Service for "SOURCE": \$server\_cifs server\_4 -a netbios=source, domain=t2dom1, interface=ana1
- f.) Started Usrmapper Service on BETA Control Station [Usrmapper on ALPHA must not be running!!]
- g.) Started CIFS Service on Server\_4 and verified CIFS access
- h.) Unshared the destination folder from Server\_4 "TARGET" called "targetshare" on "/temperv"
- i.) Created a new share on "SOURCE" called "sourceshare" on "/temperv"
- j.) From NT Client, mapped drive to \source\sourceshare, verified access to folders, and permissions once again
- k.) Retrieved copy of Localgroups.db database from Server\_4: \$server\_file server\_4 -get localgroups.db newlocalsource.db
- l.) Stopped CIFS Service on Server\_4 and the Usrmapper Service on BETA Control Station
- m.) Made BackUp copy of "newlocalsource.db" as "newlocalsource.bak"
- n.) Using Vi Editor: #vi newlocalsource.db

**Note:** During vi session, change all entries that have word "TARGET" to word "SOURCE" in the first part of the LocalGroups File. Then, in the bottom half of the file, change all entries that were "SOURCE" to "TARGET". Once vi session is completed, place the revised file back onto Server\_4: \$server\_file server\_4 -put newlocalsource.db localgroups.db

o.) Restart Usrmapper Service, then the CIFS Service, and verify access.

#### **PART V. OPTIONAL: MAKING "SOURCE" SERVER THE DEFAULT CIFS SERVICE ON SERVER 4:**

**Intro:** Because the "Target" Server existed first, on the new Celerra, it became the default CIFS Service for Server\_4. After the Migration Steps I-IV. are completed, do the following to make "Source" the default CIFS Service configuration on Server\_4. Delete the CIFS Services on Server\_4 and readd back the "Source" CIFS Service:

1. Stop CIFS on Server\_4

2. Stop Usrmapper Service on BETA CS0 172.24.80.11
3. Delete WINS entries from WINS Database for both "TARGET" & "SOURCE" CIFS Servers
4. Stop & Restart the WINS Service on NT
5. Delete the CIFS Configuration for SOURCE & TARGET: \$server\_cifs server\_4 -d netbios=source | target
6. Readd the CIFS Configuration for SOURCE: \$server\_cifs server\_4 -a netbios=source,t2dom1,interface=ana0 [or keep ana1]

### **RENAMING CELERRA COMPNAMES [Compname rename Netbios Rename, etc] (-rename):**

**Note:** There is no direct way to simply "rename" a Celerra compname from one name to another, but must use the following procedure, which allows a Compname to be removed from the Domain, creates a new CIFS Netbios name, undergoes a Netbios rename operation that re-indexes the Server database so as to preserve the Data Mover's Shares, LocalGroups security context, & Folder/File permissions on file systems, deletes the Netbios name, then reads and Joins the new Compname back to Active Directory. This procedure can be safely applied to regular CIFS Servers and CIFS Servers that are members of VDM containers. When using the following procedure, please note that Netbios names cannot be longer than 15 characters.

#### **1. Unjoin the original Compname from the Domain:**

→Localgroups file not touched

# server\_cifs server\_2 -Unjoin compname=W2k3, domain=abc.com, admin=Administrator

#### **2. Delete the Compname from the Data Mover's CIFS configuration:** →Localgroups file not touched

# server\_cifs server\_2 -delete compname=W2k3

#### **3. Add the name back to the Data Mover's CIFS configuration as a Netbios name:** →Localgroups file not touched

# server\_cifs server\_2 -add netbios=W2k3, domain=abc, interface=fsn01

#### **4. Rename the Netbios Server to the new name:** →Localgroups file updated with new name

# server\_cifs server\_2 -rename -netbios W2k3 W2k3prod

#### **5. Delete the old Netbios name from the Data Mover's CIFS configuration:**

# server\_cifs server\_2 -delete netbios=W2k3prod

#### **6. Add & Join the new Compname to the CIFS Configuration & AD Domain:**

# server\_cifs server\_2 -add compname=W2k3prod, domain=abc.com, interface=fsn01

# server\_cifs server\_2 -Join compname=W2k3prod, domain=abc.com, admin=Administrator

**Note:** CIFS Shares that are shared to a specific "netbios" name are updated with the new name and all Windows permissions remain intact. The Localgroups database retains its internal SID and ordinal position in the file, with only the name changed and updated. Please note that all persistent Client drive mappings will need to be deleted and readded to handle the new Compname.

### **CELERRA LOCALGROUPS.DB FILE:** NAS 2.2.x

When creating the "CIFS Service" on a DataMover, the "localgroups.db" file is automatically produced. There are several different ways to access this database file directly, or indirectly:

#### **I. NT Server Administrator:** Map directly to the root hidden C\$ share of the CIFS Server:

Start>Run box: \\polk\c\$ [cd to /etc directory and open the localgroups.db with wordpad]

#### **II. Control Station:** #mount -F nfs server\_2:/ /mnt #cd /mnt #cd .etc #ls -la #more localgroups.db

#### **III. Control Station:** \$cd /nas/rootfs/slot\_2/.etc \$ls [localgroups.db; .db.bak]

#### **IV. User Manager>User>Select Domain>\polk** [Indirectly accesses the Server's "localgroups.db" file]

Changes made while using this Interface will change and update the Server's localgroups.db

**Note:** Since all changes to the localgroups database are first stored in "/.etc/.db.cache", you may need to stop and restart the CIFS service in order to initially 'populate' or to update an existing localgroups database file. Additionally, the "/.etc/.db.bak" is an automatic backup of the "localgroups.db" file, first created whenever any changes are made and the CIFS service is shutdown.

#### **V. Control Station Screen Output :**

\$server\_config server\_x -v 16384 "lg list"

### **EXPLANATION OF LOCALGROUPS FILE SIDS:**

**Intro:** The SIDs listed at the beginning of the "localgroups.db" file refer to the various "netbios" or "comnames" [i.e., CIFS Services] created on the DM and are local machine SIDs only. These do not represent Domain SIDs. Each "machine account" contains its own SAM database [Security Accounts Manager] just as any native NT 4.0 or Windows 2000 System. Celerra is unique, however, in that each "Server" can contain multiple SAM databases, one for each "CIFS Service" created on the DM.

### **EXAMPLE OF BEGINNING OF LOCALGROUPS FILE FOR SERVER 4:**

# Localgroups database

\$RELEASE:02\$ →Indicates ASCII mode

@FILE008:1:1000:S-1-5-15-32434d45-2385-9c64f114-ffffffffff

@FSDATA:2:1560:S-1-5-15-32434d45-4103427e-479af894-ffffffffff

@FSR5T:3:1000:S-1-5-15-32434d45-310332f2-a20d8e2a-ffffffffff

@OSG96ST:4:6181:S-1-5-15-32434d45-2a033da3-abcd4245-ffffffffff

@OSG96S:6:1000:S-1-5-15-32434d45-2a033662-d0f26db9-ffffffffff

# Localgroups of server FILE008

### **I18N MODE LOCALGROUPS ENTRY:**

```
# head .db.localgroups
#
# Localgroups database
#
$RELEASE:4$ →4$ indicates Unicode mode localgroups database
```

### **MULTIPLE LOCALGROUPS DATABASE FILES:**

localgroups.db →Normal localgroups file contained on DataMover

.db.localgroups →New localgroups file created after turning on I18N and subsequently in use by DataMover

.db.cache → Some versions of NAS used this file to cache “changes” to the Localgroups file until they are flushed to file

.db.5.localgroups →Indicates that Local Users Support feature has been enabled

### **HOW TO OUTPUT SPECIFIC DM LOCALGROUPS TO SCREEN:**

**# .server\_config server\_6 -v 16384 "lg list vs=osgdata"**

### **HOW TO OUTPUT COMPLETE SERVER LOCALGROUPS TO SCREEN:**

**# .server\_config server\_6 -v 16384 "lg list"**

**Note:** Use “server\_file server\_6 –get localgroups.db localgroups.db.s6” to obtain working copy of db file

### **REMOVING LOCAL GROUPS ENTRY FROM VDM CIFS CONTAINER:**

**\$server\_config server\_2 "lg remove vs=TEST3 forever"**

### **FORCING AN UPDATE OF THE LOCALGROUPS FILE ON DM:**

**\$server\_config server\_6 -v "lg update force"**

**Caution:** Do not run this command without Engineering or TS2 approval

**Purpose:** Forces Datamover to query Domain Controllers and Usrmapper for updates to all Groups that match the “localgroups.db” file. An example of when you would use this command would be if you have rebuilt Usrmapper and now have different GIDs assigned to all Groups. This command will take the new GID from Usrmapper and update the entry in the Local Groups file. Be aware that in the case of multiple Compnames, the update does not always work and some manual editing may be required to update GIDs, but this is a rare occurrence.

### **LOCALGROUPS FILE CREATED WHEN USING LOCAL USERS SUPPORT:**

-rw-r--r-- 1 root bin 4156 Oct 18 12:24 **.db.5.localgroups**

**Note:** Introduced with NAS 5.4

### **WELL-KNOWN SID'S:**

SID: S-1-0 Null Authority S-1-5-domain-512 Domain Admins

SID: S-1-0-0 Nobody S-1-5-domain-513 Domain Users

SID: S-1-1-0 Everyone S-1-5-domain-519 Enterprise Admins

SID: S-1-3-0 Creator/Owner SID: S-1-5-domain-500 Administrator

### **NESTED GROUPS & USRMAPPER MAPPINGS:**

**Note:** You must be using Native Mode Active Directory to get this to work:

1. Create user called "nestee" -->By default, a member of Domain Users group
2. Create a Global Group called "nest1" and a Global Group called "nest2"
3. Add the user "nestee" to Group "nest2"
4. Add Global Group "nest2" to "nest1"
5. Log off computer and back on to obtain new Security Token and map new drive letter to Celerra.

**Results:** Both Global Groups "Nest1" and "Nest2" were mapped as groups by Usrmapper because the user "nestee" is a member of "nest2", while "nest2" is a member of "nest1". The user "nestee" was not a direct member of "nest1".

### **WINDOWS MIXED MODE vs. NATIVE MODE DOMAINS:**

Mixed Mode is an interim mode when preparing to move from NT 4.0 Domains to Windows 2000 domains. Win2k domains default to mixed mode when created. The basic premise is that at least one NT 4.0 BDC (Backup Domain Controller) still exists, requiring that NTLM authentication be allowed. W2k uses the RID Operations Master & PDC Emulator service to provide security descriptor services for NT, as well as a master browser. Once you switch to Native Mode operation, there is no reversing the step.

Some of the major advantages to running W2k (or higher) in Native Mode is the ability to use Universal Groups, have Nested Groups, and to provide automatic transitive 2-way trusts between domains within the forest.

#### **Windows 2003 supports (4) different domain levels:**

Windows 2000 Mixed Mode—default, supports DC's running W2k3, W2k, or NT 4.0

Windows 2000 Native—supports DC's running W2k3 & W2k

Windows 2003 Interim—supports DC's running W2k3 and NT 4.0—can upgrade directly from 4.0 to 2k3

Windows 2003—Supports only W2k3 DC's (aka, Native Mode)

**Note:** MS has replaced the terminology for W2kK3 domains from “mixed” mode to “Raise Functional Level”. If you change the functional level of a W2K3 domain in “Active Directory Domains and Trusts>Raise Domain Functional Level”, you would be setting the domain to native W2K3 mode and no DC's other than W2k3 will be able to function.

#### **QuarantinedDomains: SID Filtering**

This is a Microsoft concept to allow the ability to isolate certain trusted domains so as to avoid potential exploits. If SID filtering is enabled in the registry, or via the use of the netdom tool, all SIDs will be discarded from the named Trusted domain except for those related to the Trusted Domain SID itself.

HKLM>System>CurrentControlSet>Services>Netlogon>Parameters

1. Create new Key “QuarantinedDomains” with value of 0 or use c:\netdom /filtersids no <Trusted\_domain\_example>
2. Add Netbios name of the Domain to filter or ‘quarantine’
3. Stop and start the Netlogon service

**Note:** Side effects will affect migrations that use SIDHistory, and Universal Group SIDs will be discarded.

#### **CELERRA HOMEDIRS—Celerra Home Directory Service:**

**Introduction:** Creates Home directories on a Celerra file system, preferably underneath a nested mountpoint so as to hide folders from other Users. Homedirs requires the use of a ‘homedir’ map file, a valid network share to the file system, and the Homedir service running on the DM. User folders must be created & permissioned manually with older NAS versions, and can be set to AutoCreate if using MMC Homedir snapin. Home Directories functions the same for both NT 4.0 and Win2k Domains. Map file is parsed from the top down.

**Note:** NAS 5.2 now supports the use of multiple wildcard entries per domain. Dart will parse each line from the top down until it finds a Home Directory mapping for a User. You can have multiple trusted domains listed in the Homedir file. Older code had restrictions in that multiple wildcard entries could not be used for the same domain.

#### **HOMEDIR FILE PARSING:**

Home Directories are read and evaluated from top to bottom in the config file, with the last applicable entry applied. Because of this, use Wildcard matching at the top of the file and more specific matching at the bottom. The more specific matches would take precedence over the wildcard matches.

```
*.*:/fs1/guest  
galaxy:user1:/fs1/users  
galaxy:user2:/fs1/users
```

#### **STOPPING & STARTING HOME DIRECTORY SERVICE:**

**\$server\_cifs server\_x -o homedir            \$server\_cifs server\_x -o homedir=NO**

#### **CONFIGURING HOMEDIRS USING CELERRA MANAGEMENT SNAPIN:**

→NAS 5.5.31.6, CelerraCIFSManagement v4.2.00

1. Configure and start CIFS
2. Export path to the Home Directory Share
3. Set permissions on Home Directory path
4. Install CelerraCIFSManagement snapin on Windows Server
5. Open Celerra Management>Homedir (disabled)>Rightclick to Enable: “You are attempting to start the home directory service with an empty database. Before the service can be started the database must have at least one entry. Do you want to create new entry and now start the service?” yes

#### **Home Directory properties**

New home directory entry on da30

#### **Domain**

Dwarfs

User

\*

Path

/fs30/da30

Options

--Autocreate Directory [Only creates directory in the Homedir file, not on the share?]

Umask>modify (use default umask or set umask override for User Group Other

Resulting Homedir File is Automatically pushed to ./etc directory:

**dwarfs:\*:/fs30/da30:create**

## **NAS 5.4 HOME DIRECTORY IMPROVEMENTS:**

DART does dynamic mapping of User and Domain to homedir map db [consisting of domain, username, path]

DART uses last entry in its database to map user—therefore, last entry in the file would be used rather than first entry

→Automatic directory creation can be configured from MMC snapin for Home Directories, but still requires use of Profile path in ADUC for \cifs\home to be able to create the User directory

**dwarfs:\*:/fs30/da30:create** (the ‘create’ entry on the line indicates that automatic directory creation was enabled from the MMC)

→Regular expression parsing now used, based on Extended Regular Expression Characters (ERE), tools such as grep, awk, sed

Example:

**stunet\*:\*/homedir\_stu/users/<u>:regex:create**

**Note:** For any domain name beginning with string "stunet", map all Users to the path indicated and then automatically create a directory for the user by name, using regular expression <u> for the user. Note that in testing 5.5.31, the actual directory on the file system is not automatically created.

### **VALID EXTENDED REGULAR EXPRESSIONS (ERE):**

**^ . [ \$ () | \* + ? { \**

**Note:** Valid if found outside of bracket expressions

### **SUPPORT FOR REGULAR EXPRESSIONS IN HOMEDIR CONFIG FILE:**

**DENVERNT:^[a-k]:/Home-root1/Home1:regex**

**Note:** must add the word ‘regex’ at the end of each line in the Homedir file to recognize use of Regular Expressions

Single Character ERE Mapping:

. Matches any one character “user.”

... Matches any one of characters listed between brackets “[2ez]”

\* Matches preceding element zero or more times “\*” matches any string

+ Matches preceding element one or more times “(00)+” or “a+”

{num} Matches preceding element by the num times indicated between the brackets

^ Matches beginning of the line

\$ Matches end of the line

“^high\$” Matches single line containing nothing but text “high” in this example

| Separates two EREs and accepts a match to either the left ERE or the right ERE, or both

<d> <u> When used in path, represents Domain and User, respectively

**Highest to Lowest Precedence:** Regular expressions are evaluated with the following precedence

**[ ] Bracket Expression () Group \* + {m} Single-char. Quantifier ^ \$ Anchor | Alternator** (either or match)

→Management done via MMC Snapin on 5.4 Applications & Tools CD—Does NOT extend AD Schema

→Requires use of IE 6.0 SP1 or later

→RO option for Homedir access

## **EXAMPLE OF CREATING HOMEDIR ENTRY USING MMC SNAPIN:**

### **I. Programs>Administrative Tools>Celerra Management>Homedir (enabled)>rightclick, New>Home directory entry:**

Domain

w2k\* [Enter regular expression--asterisk matches any domain beginning with w2k]

User

\* [Asterisk matches any user name]

Path

\fs3\students

Options

Enable \_\_Auto Create Directory and \_\_Regular Expression

### **II. Open Active Directory Users and Computers><select a user>Properties>Profile**

Home folder

Select \* Connect Z: and enter UNC path to the built-in “home” directory share: \\dbms\home

### **III. When user logs into the Celerra, a directory is automatically created for the User**

## **I. CONFIGURING CELERRA HOME DIRECTORIES (HOMEDIRS):**

### **Step 1. Create Homedir Map File and push to Root File System of DM (/etc):**

#vi homedir

```
mouse:*:/webui/Users  
mouse:*:/webui/Users/2004  
mouse:*:/webui/Users/2005
```

**#server\_file server\_2 -p homedir homedir**

**Note:** Where ‘MOUSE’ = Domain Name; \* = any User—standard wildcard; /webui = upper level share or “mountpoint” of filesystem that hosts the Home Directories; /users = subfolder that contains a collection of user Folders by name

### **Step 2. Create a file system share via CLI or Computer Management/Server Manager for “/webui”:**

#### **Step 3. Create Home Directory Folders and apply permissions to each Users’ Folder**

#### **Step 4. Create User Logon Profiles via Win2k ADUC or NT 4.0 User Manager Interface:**

**NT 4.0 DOMAINS:** → USER MANAGER FOR DOMAINS

**User>nas1>profile>”Connect” Z: \\mouse\home**

**Note:** You may see an error, ‘...Could not create the HOME directory...for this user’—Ignore this message!

**WINDOWS 2000 DOMAINS:** → Active Directory Users and Computers (ADUC)

**User>nas1>profile>”Connect” Z: \\mouse\home**

**Note:** Ignore following message: “The \\mouse\home home directory was not created because it already exists...”

### **Step 5. Start Home Directory Service on DM & Verify:**

**\$server\_cifs server\_2 -o homedir | homedir=NO** [to stop service]

**\$server\_cifs server\_2**

**Home Directory Shares ENABLED, map=/.etc/homedir**

**Note:** The default CIFS share called “Home” is created automatically when you start the HomeDir service:

/etc/shares/@system/HOME --> HOME. Please ensure that the Customer does not create any other Celerra Shares with the name “Home”—reserved only for Homedirs!

### **Step 6. Log on as a User—the User’s Home Directory should be mapped in Explorer automatically**

**Note:** With Windows 2000 environments, replication may need to occur before a User’s Home Directory is automatically mapped

### **Step 7. Alternatively, log on as a User and manually map to Home Folder: START>Run: \\mouse\home**

**Note:** This method does not require that the User’s Logon Profile be enabled in User Mgr or ADUC

## **Wildcard \* Limitations for Home Directory Configurations [HOMEDIRS]:**

**Note:** Home Directories will not work when using multiple Wildcard Lines in the HOMEDIR Map File

**HOMEDIR Map Files Using Wildcard \*:** [\* symbol means any valid NT user from that Domain]

**VALID HOMEDIR MAP FILE EXAMPLE (1):** Use of Single Wildcard Line to Map all Users to a Common Home Directory

sycamore:\*/fs2/Students

**Comment:** Where "sycamore" = NT Domain; where \* = Wildcard Mapping for any valid User in the "sycamore" domain; where /fs2 = Celerra Mountpoint for Home Directory File System; where /fs2/Students = Unix Path for Home Directories

**Note:** In this example, Users would be mapped automatically to their Individual Home Directories if using the \\sycamore\home UNC path at the START>Run prompt as follows: \\sycamore\home [where "sycamore" in this case is the Netbios name of the CIFS Server] Individual User Folders must be physically created, and permissions assigned, under the "Students" Directory, as follows:

/fs2/Students/2001usr1

/fs2/Students/2002usr1

**VALID HOMEDIR MAP FILE EXAMPLE (2):** [Placing Wildcard \* Entry on Top Line only!]

sycamore:\*/fs2/Students/2001

sycamore:2002usr1:/fs2/Students/2002/2002usr1

sycamore:2002usr2:/fs2/Students/2002/2002usr2

**Note:** This configuration will only work for the Wildcard Line with the asterisk \* --all other lines above will NOT be "parsed" Wildcard entries must be on the top line of the HomeDir Map File! As per the example, you can use a wildcard entry once on the top line, but then must map all other paths explicitly by user name if the path is different than the wildcard line. This applies only to NAS codes 2.2 and lower.

## **MORE HOMEDIR SETUP EXAMPLES:**

Homedit File: t2dom1:\*/mnt30/cifs1  
Server Mgr Share: c:\mnt30 [File System ufs30 is “shared” for NT as “CIFS”]  
Server\_export: share “cifs” “/mnt30”  
**Note:** Mountpoint ‘/mnt30’ is ‘shared’ for NT as ‘cifs’  
User Mgr Profile: Z \* <\\monitor\\home>  
Explorer Folders: Subfolder “cifs1” is created under the ‘share’ called “cifs”  
User folders are created under “cifs1” [e.g., user1; user2; user3]

## **TROUBLESHOOTING HOME DIRECTORIES:**

1. Verify Service is started using server\_cifs
2. Verify that “homedit” map file contains correct mapping

## **EXAMPLE OF INCOMPLETE OR INCORRECT HOME DIR MAP FILE:**

<\\brighton\\home>

The network name cannot be found.” [Incorrect path or other configuration issue with Home Dir config file]

3. Try manually mapping to the Home Directory from START>Run: \\brighton\home

**Note:** A failure to map manually indicates that the Home Directory Service or Path are not setup correctly!

4. If automatic drive mapping letter does not appear when a User logs on, verify that NT or Windows 2000 replication has occurred
5. Set Debug Logging on Server:

**param cifs homedirTraces=1**

**Note:** With 5.6, you may need to use \$ server\_log server\_x -i option to see trace output in the server log

## **CELLERRA SECURITY MODES: NT, UNIX, SHARE**

### **UNIX SECURITY MODE:**

#### **Key Concepts:**

--NT Domain Controller not used

--UID/GID MAPPING FOR UNIX USERS PERFORMED BY NIS, OR BY LOCAL DATAMOVER PASSWD FILE

--Cannot implement UNIX Security if I18N has already been enabled

**Note 1:** Password File should include valid encrypted “password” but not the NT Domain name as part of the username

**Note 2:** Requires enabling of PlainText Passwords on the NT or WIN2K Clients that are accessing the Celerra because users are *NOT* authenticated to an NT Domain Controller [i.e., NT ACL’s are *NOT* used]

--Accesspolicies” do not apply and are not effective while using Unix Security Mode

--Consequently, must use either a local DataMover “encrypted” Passwd File or a NIS Server DataBase for File Access/Authentication

--You still need to run the CIFS service

--You export NT “shares” via the server\_export command

--Unix Users would use the NIS or local PASSWD/GROUP files for access [Client sends username and plain-text password to CIFS server and is authenticated against the local ./etc/passwd file or NIS Server Database]

--Still considered a method of “user-level authentication”, similar to NT Security [as opposed to Share-Level Security], yet no File or Directory-level ACL checking occurs.

**Implementing Unix Security in an NT Workgroup environment:** NT Registry entry to allow clear-text passwords

**Note:** Normally, you would employ Unix Security in an NFS environment if you had an NT workgroup that needed to access Unix FileSystems but did not have a Domain Controller. Typically, you would have a NIS server that would authenticate the NT User against the NIS database using Username & clear-text Passwords.

#### **Setting Up Unix Security:**

Configure DataMover Interface>Create MetaVolumes>Create FileSystems>Create mountpoints>Mount file System without accesspolicies>Export File Systems for NFS>Setup CIFS Service>Create passwd/group file & place locally on DM or use NIS Database>Set Unix Security: \$server\_cifs server\_2 -a security=unix>Start CIFS service>Export FileSystems as “shares” for CIFS

**Note:** Only need the Server’s “Netbios” name if you intend to use NT Browsing to list Celerra shares: \$server\_cifs server\_2 -a netbios=server\_2, domain=comics, interface=ana0

### **NT SECURITY MODE:**

Unix UID/GID Conversion for NT Users mandatory!

**Note:** UID/GID MAPPING FOR NT USERS DONE WITH PASSWD & GROUP FILES

[Local Passwd/Group files on DM; Using NIS Service to map to NIS db; Using Usermapper Service; or NTMigrate Utility]

--NT Security is the Default Mode. If the CIFS Service is configured and started, NT Security is the default mode.

--Default Accesspolicy is NATIVE, and is configured using the Mount command

--Eventhough you can assign more than one “NT Netbios” name for a Server for multiple NIC Interfaces, the first name used will become the default Netbios Name used for NT Security on the DataMover.

--NT Security Supports Three Types of ACL's [Access Control Lists]: NTFS File, Directory, and Share Level ACL's

**NTFS ACL Permissions**—set at the File or Directory level

**NT Share ACL Permissions**—set at the NT “share” when exported as a share

**Note:** NT will employ the most restrictive of the above permissions collectively when checking access!

--NT Clients will use NT Domain Controller for authentication

--Additionally, use of a WINS Server is mandatory when the Domain Controller is on a separate network from the DataMover

**Note:** WINS registration cannot occur across a router using Netbios Broadcasts alone—must use directed IP address of WINS Server to route across a Router

--Must use a Passwd/Group file locally on the DM or a NIS Server passwd database service

## **SHARE SECURITY MODE:**

**Note:** When exporting Share, assign password of DEFAULT\_USER

→Requires enabling of Clear-Text Passwords on the NT or WIN2K Clients that are accessing the Celerra because users are *NOT* authenticated from Domain Controllers

→Accesspolicies do not apply and are not in effect when using Share Security Mode

→Use a local Passwd & Group file on the DataMover Access by using a “default” user—INDIVIDUAL UID/GID accounts are not used. Only a single “default” user account can be assigned and must be in the local DM password file. Usernames must be limited to 20 characters and Groupnames limited to 256 characters.

**Note:** You can Export the NT “Shares” so that a User and Password are required, or you can Export “Shares” so that no User and Password box appears [i.e., would allow direct UNC connection: <\\server\share>].

→CIFS service is required

→Setup Netbios name for NT Browser service if required

### **Create mountpoints/directories on DataMover from Control Station:**

**Example:** #mkdir newshare #chmod 777 newshare

**PASSWD FILE:** admin:\*:100:300::: [where admin=user; \*=passwd field, not checked in this case; 100=UID; 300=GID]

**GROUP FILE:** share:\*:300: [where share=Group Name; \*=any user; 300=GID]

**Note:** CHMOD the file system mountpoint to allow “share” users access to the Sharepoint [**#chmod 777 newshare**]. Ordinarily, should not need to run “chown” on Sharepoints “unless” you are having trouble writing or deleting the Share.

If so, use this syntax to change ownership of the Sharepoint to the UID of the “Default User” and the GID of the Group:

**#chown 100:300 newshare** [where 100=UID of the Default User “admin” and 300=GID of Group “share”]

### **User-Level v. Share-Level Authentication: chmod 755 the ‘shared directory’**

NFS can use either User-level [server\_export; server\_cifs UNIX] or Share-level Authentication [server\_export command]

CIFS uses either User-level [server\_export; server\_cifs NT] or Share-level authentication [server\_cifs SHARE command]

Unix authenticates using Passwd or Group file or NIS Server—No ACL checking!!

DataMover [for Unix] can implement Share-level security via the “server\_export” command, defining the default UID/GID for the share. File locking or Access Checking policies do not need to be specified.

NT authenticates using NT Domain Controller supporting User, Group, Rights, File, Directory, and Share ACL's

Share Authentication: A default user is assigned to a share and can be assigned a Password or No Password.

## **NULL SESSION SUPPORT**

### **USING NULL LOGON SESSIONS (Anonymous Logons) FOR CELERRA SHARES:**

A Null Session is a NetBIOS connection made to a Server by an anonymous Client without the use of authentication—client's id field is null. Windows 2000 allows null sessions to enumerate local SAM accounts and Shares on Servers by Clients, while XP and Win2k3 allow only enumeration of Shares via null session.

#### **WINDOWS 2000:**

HKLM>System>CurrentControlSet>Control>Lsa>RestrictAnonymous: DWORD

RestrictAnonymous=0 →Default Win2k setting, no restrictions—enumerate Shares and local SAM accounts

RestrictAnonymous=1 →No enumeration of Shares or SAM accounts allowed

RestrictAnonymous=2 →Enumeration allowed only for explicit anonymous permissions

**Comment:** An example of Null Sessions support would be using the AT Scheduler to open a command prompt session and connect to remote network share using “net use”, etc. Commands run via the AT Scheduler typically use the local System Account context, not the locally logged in User account context.

**Symptom:** Logon Failure: the user has not been granted the requested logon type at this computer

#### **CONFIGURING NULL SESSION SUPPORT:**

1. Enable Null Session Support by setting the following params:

**# server\_param server\_2 -facility cifs -modify nullSession -value 1**

**# server\_param server\_2 -facility cifs -modify nullSessionNotOnFS -value 1**

2. Use Celerra Management snapin GUI from Windows Server to set following privilege on the Data Mover:

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
Programs>Administrative Tools>Celerra Management>Action>Connect to Compname>Data Mover Security Settings>User Rights Assignment>double-click ‘Access this computer from the network’>Add ‘Everyone’ group

## **PARAMETERS INVOLVED:**

**param cifs nullSession=1**

**param cifs nullSessionNotOnFS=0**

**Note:** nullSession=1 allows clients to connect to IPC\$ on the DM without credentials. nullSessionNotOnFS=0 opens up access to Shares for all unauthenticated Users. Control actual file & directory access using NTFS permissions, which are still effective.

## **SETTING GUEST ACCESS TO SHARES WITHOUT AUTHENTICATION:**

1. Enable Guest account in AD
2. Set following param to allow write access to file system

**# server\_param server\_2 -facility cifs -modify nullSessionNotOnFS -value 0**

3. Use ntrights.exe utility to enable login access for Everyone group on Compname

**C:\>ntrights.exe +r SeNetworkLogonRight -u Everyone -m \\datamover\_netbios\_name**

## **CELERRA UID/GID MAPPING METHODS FOR CIFS USERS/GROUPS:**

In pre-Windows 2000 support CIFS configurations, Celerra supported UID/GID mappings by names—the search order for mappings included the local DM passwd/group file, then NIS if the NIS client was configured, and Usermapper last. In more recent NAS code, say approximately from 4.2 and higher, we use additional mapping mechanisms and with Usermapper, map by SIDs as opposed to names, though if NIS and local passwd/group files are used, we revert to name lookups.

## **DART MAPPING RESOLUTION ORDER FOR USER/GROUP AUTHENTICATION:**

### **MAPPING RESOLUTION ORDER NAS 5.1 & BELOW:**

1. Global Data Mover Sid Cache (DART resolves Group SIDs and maps to GIDs)
2. Local Password & Group files (DART resolves by User/Group names and maps to UIDs/GIDs, respectively)
3. NIS Client Service (DART resolves by User/Group names and maps to UIDs/GIDs, respectively)
4. Active Directory Mapping Utility (DART resolves by User/Group SIDs and maps to UIDs/GIDs, respectively)
5. UserMapper Service (External-only) (DART resolves by User/Group names to UIDs/GIDs, respectively)

### **MAPPING RESOLUTION ORDER NAS 5.2 & ABOVE:**

1. SecMap Persistent Cache [/etc/secmap] (DART resolves by Group/User SIDs and maps to GIDs/UIDs, respectively)
2. Global Data Mover Sid Cache (DART resolves Group SIDs and maps to GIDs)
3. Local Password & Group files (DART resolves by User/Group names and maps to UIDs/GIDs, respectively)
4. NIS Client Service (DART resolves by User/Group names and maps to UIDs/GIDs, respectively)
5. LDAP iPLANET\*—optional configuration using NSSWITCH.CONF file on DM, ordering UID/GID mapping searches between LDAP Service, Files (local passwd/group), or NIS
6. Active Directory Mapping Utility (DART resolves by User/Group SIDs and maps to UIDs/GIDs, respectively)
7. UserMapper Service (External or Internal) (DART resolves by User/Group SIDs and maps to UIDs/GIDs, respectively for Internal Usermapper, and by Name-to-UID/GID prior to NAS 5.2)

**Note:** Do not use uppercase characters in passwd files or NIS database—use lowercase ASCII for Celerra. Also note that param cifs resolver=1 should be set when using NIS or local passwd/group files. With param set to 0, DART will request a mapping based on name.domain, and NIS will not return a mapping unless its an exact match. Keep in mind that the Global Group SID cache is kept in cache permanently, only being removed by mappings that are new and/or re-used at some frequency (LRU mechanism). When configured & populated, SecMap Cache will be the very first cache interrogated by DART for User & Group mappings.

### **\*MAPPING RESOLUTION ORDER NAS 5.4 & ABOVE:**

**Note:** LDAP iPlanet Directory Services support was introduced with NAS 5.4, via configuration of nsswitch.conf file, order of mapping Users/Groups can be specified as any combination of LDAP, Local Files, or NIS lookups

**\$ cat /nasmcd/quota/slot\_2/etc/nsswitch.conf**

passwd: files ldap nis (can be in any order that customer wishes)

group: files ldap nis

## **DEFAULT HOSTNAME SEARCH ORDER FOR DART:**

“files nis dns”

Use /etc/nsswitch.conf file to modify the default search order behavior for DART

## **I. LOCAL SID CACHE ON DM:**

From 5.2 versions and higher, DART now caches group SIDs from sources such as local group files or from other mapping mechanisms. This cache is kept in memory for as long as the Data Mover is up and running. However the number of cache entries are finite and can be increased with a parameter setting. Once the cache is filled, the next new entry is added on an LRU basis, that is, the oldest unused entry will be replaced by the newer entry. This cache is used as an efficiency mechanism to help prevent

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
authentication logjams by speeding up the mapping validation—in otherwords, if an entry is already in cache, DART does not have to go back out to the Domain Controller to conduct lookup SIDs for each Group in the domain.

**param cifs sidcache.globalSidCacheSize=1901** [Should be a prime number]

## **II. SECMAP CACHING:**

DART 5.2 introduced a persistent UID/GID-to-SID cache that is stored in the ./etc/secmap folder of the data mover. SecMap will cache the first mapping for a particular SID and note where the mapping originated: i.e., local passwd file, NIS, Usermapper, etc. Subsequent data mover access by an owner of the SID will be immediately mapped & authenticated from this cache, meaning that DART does not have to query the DC's to validate the SID again. This mechanism was introduced as an efficiency mechanism to cut down on the single-threaded authentication calls made to the DC that often translated into authentication backlogs and slow DM performance during heavy logon periods, etc.

## **III. LOCAL DATA MOVER PASSWD/GROUP FILES:**

The use of local passwd & group files on the data mover has long been the most recommended way to cut down on data mover traffic to the DC's and provide a redundant mapping database in the event that another mapping method, such as Usermapper, were ever unavailable to serve mapping requests. With the introduction of Internal Usermapper and SecMap caching, local passwd/group files are not now normally used and should only be done so as a last resort and with knowledge of the various mapping components and interactions. In addition, Version 3 style (SID in first field) passwd/group files should never be used on the data mover.

### **TRADITIONAL UNIX PASSWD FILE:**

**user.emc : \* : 501 : 100 : User Bob : :**

**Note:** User & Domain Name appended in first column; passwd field; UID; Domain GID; Comment Field]. Use of the \* asterisk means that NT Security is in use and Unix passwd field is ignored. If using UNIX Security, you would use a valid password and NOT append domain name to user entry.

### **TRADITIONAL UNIX GROUP FILE:**

**emc: \* :100**

**domain=20admins.emc: \* : 101**

**domain=20users.emc: \* : 102**

**Note:** First entry is for the NT Domain—default GID for domain is 100. Additional entries are for Groups. \*Asterisk indicates to map and domain User or Group. Please note that UTF8 characters are not supported in either the passwd or group files, only ASCII values. However, you can use and =20 to represent a space between words in a name, or ==hexahexa characters to translate UTF8 characters into ASCII readable characters. Use a tool like uni2ascii.

**# uni2ascii -Z ==%04x < group**

clearcase:x:554:

==65e5==672c==8a9e:x:555:

==5e83==544a==63b2==8f09:x:556:

## **CUSTOMER SAMPLE GROUP/PASSWD FILES:**

**\$ head -5 group**

```
001a=20dept=20accounts.corp:*:93066:S-1-5-15-8f5db4-14de7199-375b3a1a-9bc6:  
001m=20dept=20accounts.corp:*:91986:S-1-5-15-8f5db4-14de7199-375b3a1a-b56a:  
003l=20dept=20accounts.corp:*:92991:S-1-5-15-8f5db4-14de7199-375b3a1a-9bc8:  
003m=20dept=20accounts.corp:*:94453:S-1-5-15-8f5db4-14de7199-375b3a1a-21c9f:  
004l=20dept=20accounts.corp:*:92720:S-1-5-15-8f5db4-14de7199-375b3a1a-9bc9:
```

**\$ head -5 passwd**

```
1808dcc.corp:/*:95225:90000:S-1-5-15-8f5db4-14de7199-375b3a1a-2015a:/usr/S-1-5-15-8f5db4-14de7199-375b3a1a-2015a:/bin/sh  
a27lab.corp:/*:96102:90000:S-1-5-15-8f5db4-14de7199-375b3a1a-26c5b:/usr/S-1-5-15-8f5db4-14de7199-375b3a1a-26c5b:/bin/sh  
aa02.corp:/*:96515:90000:S-1-5-15-8f5db4-14de7199-375b3a1a-27061:/usr/S-1-5-15-8f5db4-14de7199-375b3a1a-27061:/bin/sh  
aa03.corp:/*:95912:90000:S-1-5-15-8f5db4-14de7199-375b3a1a-27063:/usr/S-1-5-15-8f5db4-14de7199-375b3a1a-27063:/bin/sh  
aa04.corp:/*:94463:90000:S-1-5-15-8f5db4-14de7199-375b3a1a-27065:/usr/S-1-5-15-8f5db4-14de7199-375b3a1a-27065:/bin/sh
```

## **IV. DATA MOVER AS NIS CLIENT:**

See NIS section for more details. Essentially, after configuring proper data mover params, DART can map Users/Groups from NIS maps if the proper entries have been made to the databases. This method is specifically recommended for ‘mixed’ Unix/CIFS environments.

## **V. DATA MOVER AS LDAP CLIENT:**

With NAS 5.4 and above, can use LDAP client to query an iPlanet directory server to resolve users from a passwd database, or groups from the group database. Need to setup the ldap connection to the LDAP Directory Server, and have Users and Groups mapped in the LDAP database.

**# cat nsswitch.conf**

**#**

**# /etc/nsswitch.conf**

**passwd: ldap files nis**

group: ldap files nis  
hosts: dns files nis  
netgroup: ldap files nis  
**# server\_ldap server\_2 -info**  
server\_2 :  
LDAP domain: hosts.pvt.dns  
State: Configured - Connected  
NIS domain: hosts.pvt.dns  
Proxy (Bind) DN: uid=admin,ou=Administrators,ou=TopologyManagement,o=NetscapeRoot  
Profile Name: celerra1  
Profile TTL: 43200 seconds  
Next Profile update in 8258 seconds  
Connected to LDAP server address: 192.1.4.209 - port 389

**# server\_ldap server\_2 -lookup -user w2ku45685**

server\_2 :  
user: W2kU45685, uid: 45685, gid: 65000  
**# .server\_config server\_2 -v "lsarpc if=cge user=W2kU45685"**  
1147785620: LDAP: 7: LdapClient::searchOn: client @ = 0xd87cb4c4, DN = , scope = 0  
1147785620: LDAP: 7: LdapUdpXportImpl::read: select timeout = 200  
1147785620: LDAP: 7: LdapService::shutdown: @ = 0xd8351a44, connection = 0x25b04  
1147785620: UFS: 6: inc ino blk cache count: nInoAllocs 1: inoBlk dd633184  
Finding User SID from Name=SUCCESS  
Interface 'cge' Address=192.1.6.203  
User0='2K3\W2kU45685' (0) use=1 nameType=0  
S-1-5-15-bfc56af5-d6cf8701-5f67b1a3-4c0 W2kU45685  
UNIX ID=45685 Type=0

**# view seemap\_ldap**

sid S-1-5-15-bfc56af5-d6cf8701-5f67b1a3-4c0 mapping  
    **Mapped from ldap** on Tue May 16 09:16:37 2006  
    user=b275(fde8)  
    Name=2K3\W2kU45685

## **VI. CELERRA USRMAPPER SERVICE:**

Celerra introduced an automatic User and Group UID/GID mapping mechanism that was traditionally a service run on a Unix-based system. From NAS 5.2 and higher, the Usermapper Service became integrated into the DART code and its functionality is augmented with a SecMap Caching mechanism—be aware that SecMap caching and Usermapper are not mutually inclusive. SecMap caching will add User and Group mappings based on the first mapping mechanism used on the Celerra. Usermapper is designed to be used only for pure Windows environments and should NOT be used in mixed Unix/CIFS situations.

## **CELERRA USRMAPPER 3.x:**

→Prior to 3.x, Usermapper mappings were made based on NT Usernames/Groupsnames when mapping to UID/GIDs  
→With 3.x, Usermapper mappings are based on NT SIDs for Users and Groups  
→3.x was designed to support Windows 2000, last version of Usermapper was called 3.1 for External & Internal Usermapper  
Note: Internal Usermapper is no longer ‘versioned’ per se and is an integral part of DART

## **SAMPLE “USRMAP.CFG” FILE WITH SID HISTORY LINE:**

suncor:12345:2001:9999:301:9999 [NT 4.0 Domain]  
sgho:23456:10000:10999:10000:10999  
sun\_ref\_net:34567:11000:11999:11000:11999  
network,network.lan:45678:12000:19999:12000:14999 [Windows 2000 Domain, using shortened Netbios name of domain]  
**\_history\_sid\_range\_:600:20000:24999:15000:21000**

Note: With all External Usermapper implementations, the use of the \_history\_sid\_range is mandatory, but often left out.

## **USING SORT,etc, TO VERIFY USRMAPPER DATABASE INTEGRITY:**

### **Dumping Usrmapper Database & Sorting:**

**#!/usrmap\_control dump |cut -d : -f 3,4,5 |sort**

**Sorting the Usrmap.passwd or Usrmap.group Files:** [from ‘usrmap\_control dumpfiles 1’ command]

#sort -t: -k 3 usrmap.passwd > usrmap.db

### SORTING VERSION 2.0 USRMAPPER DATABASES PRIOR TO 3.0 CONVERSION:

#sort -t: -n -k 3 usrmap.db >usrmap.passwd.sorted

#sort -t: -n -k 3 usrmapper.grp.db >usrmap.group.sorted

Comment: Use this if having problem with “preconvert.ksh” script

#sort -t: -k 3,3n usrmap.db >passwd.sorted

#sort -t: -k 3,3n usrmapper.grp.db >group.sorted

### SORTING UNIX-STYLE USRMAPPER DATABASES TO REMOVE MULTIPLE UIDS/GIDS:

#sort -k1,1 -t: -u -o check\_pass usrmap.passwd

#sort -k1,1 -t: -u -o check\_group usrmap.group

### PUTTING SPACES INTO COMMENT FIELD:

Note: Adding characters to fill spaces in comment field of “usrmap.db” file so that “correct\_dups.ksh” does not remove comments

#vi usrmap.db esc :., \$/ /xxx/g

Removing the characters from comment field:

#vi usrmap.db.new esc :., \$s/xxx/ /g

### VERIFYING USRMAPPER DATABASE RECORDS:

1. Dump Users and Groups from Usrmapper database:

#./usrmap\_control dump |grep “:user “>users [Outputs complete User database records to file called ‘users’]

#./usrmap\_control dump |grep -v “:user “>groups

[Outputs all records except ‘users’—note that two additional lines are also in the output—“UID” and “GID”, so take that into consideration when doing a word count list of the files].

2. Verify that UID and GIDs are in consecutive and numerical order:

#cat users lawk -F: '{print \$3;}' >users.sorted

Note: This should output only the column of UID numbers for review of numerical order,etc. Not sure this will work for multiple domains and ranges.

3. #view users.sorted esc : set list :set nu [Observe beginning line & ID number, then go to end of file ‘shift’ + G & compare]

4. #cat users | wc -l [compare with any previous known dumps of the ‘users’ or passwd files]

5. #cat users lawk -F: '{print \$3;}' >userlist

6. #cat groups lawk -F: -v '{print \$3;}' >grouplist

### REMOVING DUPLICATE ENTRIES FROM USER OR GROUP FILE:

# cat groups\_final | awk -F: '{print \$1":"\$2":"\$3":"\$4":";}' >groups\_sorted

Note: When running this for the passwd file, the number of \$columns will be different

# cat users\_oct lawk -F: '{print \$1":"\$2":"\$3":"\$4":"\$5":"\$6":"\$7;}' >users\_awk

### SORTING OUT DUPLICATE NAMES FROM USRMAPPER OUTPUT FILES:

Purpose: If the customer has deleted and readded user or group accounts from NT, we end up with multiple records in the usrmapper database for the same name, eventhough the SIDs are unique—if placed onto the DM, the DM will only parse down to the first name that it is looking for—and based on the way that records are added, the first of two duplicate “names” will not be the correct one.

**Step 1.** #cat usrmap.passwd lsrt -k 3,3 -g -t: -r -o passwd.rsorted

**Step 2.** #cat passwd.rsorted lsrt -k 1,1 -t: -u -o passwd.unique

**Step 3.** #cat passwd.unique lsrt -k 3,3 -g -t: -o passwd.final

Comment: Repeat steps with Group file. Purpose of these steps are to take a passwd database and reverse sort by UID from Highest to Lowest Value. Reason for doing this is that if you have multiple entries that share the same name [different SID, i.e., the User or Group was deleted and then readded on the NT Side with the same name], the duplicate name that we want to keep will be parsed first in the file. Then, the second sort goes through the file and preserves the first record in column one as the “unique” item and deletes any other ‘duplicate’ names. Finally, the last sort simply restores the file to its original sorted state, in ascending UID/GID value.

### SORTING OUT COLUMN (3) FROM UNIX PASSWD OR GROUP FILE IN ASCENDING ORDER:

#cat usrmap.db |cut -d: -f3 | sort > /tmp/aaa

## **TROUBLESHOOTING EXTERNAL USERMAPPER:**

1. Check Server Logs for entries related to Usermapper Primary or Secondaries
2. Check “usrmapper.log” for errors
3. On Windows Systems, check Event Viewer Application Logs
4. Test functionality of Usrmapper by outputting database files
5. Delete temporary \_\_db\* files if having trouble restarting Usrmapper on Linux
6. Debug Usrmapper entries using “usrmapper” commands:

## **EXTENDING UID/GID RANGES:**

**Note:** This procedure is required for changing the usrmap.cfg file for any External Usermapper configuration. It is also required for an Internal Usermapper running in CONFIG mode. Using the comma, you should be able to add multiple separate UID or GID ranges, as in the following example:

### **ORIGINAL RANGE :**

t2dom3,t2dom3.local:100:200:300:400:500

**EDITED RANGE :** Creates two additional UID ranges and one additional GID range

t2dom3,t2dom3.local:100:200:300,600:615,75000:76999:400:500, 80000 :81000

**Note :** The proper way to implement a range change would be to stop Usrmapper, edit the ranges, start Usrmapper, then verify that the ranges have been “read” into the Usrmapper database properly by running the following command:

# strings usrmappc.db

100:200:300,600:615,75000:76999:400:500,80000:81000:202:401

t2dom3

**Note:** The above entry shows that the new ranges have been recognized properly, and the last two columns show the next available UID & GID that will be used by Usrmapper, respectively.

**Warning!:** The UID/GID Ranges of a Usrmapper 3.0.6 – 3.0.10 can only be extended once from that of the original range. A third extension attempt will fail.

**RANGE EXTENSION issue is resolved by Usrmapper 3.1.1**

## **SERVER CIFS OUTPUT--LAST ACCESS=0/1 FOR USERMAPPER:**

# server\_cifs ALL |grep -i usrmapper

**Usermapper[0] = [10.10.28.46] last access 1**

**Usermapper[0] = [10.10.28.46] last access 0**

**Note:** Last access 1 means that last mapping request to the DM was a failure. A new connection request should change this status. Last access 0 means that last mapping request to DM was a success.

## **CONTROL STATION ALIASING FOR USERMAPPER REDUNDANCY:**

### **ASSIGNING AN “IP ALIAS” TO CS0:**

**Purpose:** When using dual control stations in a High Availability CS0-to-CS1 failover environment, you would assign a separate IP Address to CS0’s “eth2” interface as an “alias”, and then point all CIFS Servers to this “alias” IP Address. Once an IP Alias is in place, upon Control Station failover, the IP Alias will failover to CS1. As long as all Data Movers point to the alias IP, then CIFS communication with the Usrmapper Service will not be interrupted.

**Note:** The IP Alias is only assigned to the external interface “eth2” on CS0.

## **CONFIGURING CS0 WITH AN ‘IP ALIAS’ FOR USRMAPPER:**

### **1. Create the Alias on CS0 for External Interface ‘eth2’:**

a.) #/nas/sbin/nas\_config -IPalias -c -n eth2 0

**Note:** IPalias should be a ‘user defined number’ from 0 through 255—can have up to 255 aliases

b.) Do you want slot\_0 IP Address <172.24.80.11> as your alias [Yes or No]: No

c.) Please enter an IP address to use as an alias: 172.24.80.9

done

**Note:** Creates the following file called /nas/site/cs\_aliases/

# cat /nas/site/cs\_aliases/ifcfg-eth2:alias

DEVICE=eth2:alias

IPADDR=10.19.20.167

NETMASK=255.255.252.0

BROADCAST=10.19.23.255

### **2. Verify Status of IP Alias:**

a.) # /nas/sbin/nas\_config -IPalias -list

```
alias           IPAddr      state
eth2:0         172.24.80.9  UP
```

**b.) #server\_ping server\_x 172.24.80.9**

```
PING 172.24.80.9 (172.24.80.9) from 172.24.80.9 : 56(84) bytes of data.
64 bytes from 172.24.80.9: icmp_seq=0 ttl=255 time=52 usec
```

**c.) # /sbin/ifconfig -a**

```
eth2:0 Link encap:Ethernet HWaddr 00:00:D1:20:56:C9
      inet addr:172.24.80.9 Bcast:172.24.80.255 Mask:255.255.255.0
          UP BROADCAST RUNNING MULTICAST MTU:1500 Metric:1
          Interrupt:9 Base address:0x1c00
```

**3. Point CIFS Servers to New IP Alias Address:**

- a.) #server\_cifs server\_x -d usrmapper=172.24.80.11
- b.) #server\_cifs server\_x -a usrmapper=172.24.80.9

**4. Test Failover of CS0 to CS1 to verify IP Alias Configuration & CIFS Access:**

**a.) #/nasmcd/sbin/cs\_standby -failover**

**Note:** After issuing command, verify CIFS access

**(3) KEY DATABASE FILES IN EXTERNAL USRMAPPER:**

**usrmapc.db** → Key database that mirrors usrmap.cfg file and shows current increment number for UID/GIDs in last two columns  
# strings usrmapc.db [Use the strings command to examine this file]

10:30:40:50:60:34:58

t2dom3

**VIEWING DOMAIN RANGES USING STRINGS:**

```
# strings --bytes=1 usrmapc.db
*
80000:80001:89999:80001:89999:80001:80001
na
*
70000:70001:79999:70001:79999:70121:70044
dsigus
```

**Note:** Traditional strings of usrmapc.db file does not output Netbios Domain name ranges that are (3) characters or less in length without using above switch

**usrmapusrc.db** → Key database file that reflects UIDs assigned

**#strings usrmapusrc.db**

```
S-1-5-15-74b49ff8-2a1f6232-7ff7ed83-4b3:*:16:10:user boyo2 from domain t2dom3:/usr/S-1-5-15-74b49ff8-2a1f6232-7ff7ed83-4b3:/bin/sh S-1-5-15-74b49ff8-2a1f6232-7ff7ed83-4b3
S-1-5-15-74b49ff8-2a1f6232-7ff7ed83-4b2:*:15:10:user boyo1 from domain t2dom3:/usr/S-1-5-15-74b49ff8-2a1f6232-7ff7ed83-4b2:/bin/sh S-1-5-15-74b49ff8-2a1f6232-7ff7ed83-4b2
```

**usrmapgrpc.db** → Key database file that reflects GIDs assigned

**#strings usrmapgrpc.db**

```
S-1-5-15-74b49ff8-2a1f6232-7ff7ed83-200.*:52:domain=20admins.t2dom3:S-1-5-15-74b49ff8-2a1f6232-7ff7ed83-200
S-1-5-15-74b49ff8-2a1f6232-7ff7ed83-65d.*:51:ct4.t2dom3:S-1-5-15-74b49ff8-2a1f6232-7ff7ed83-65d
```

**IMPORTANT USRMAPPER DATABASE FILES:**

Usrmap.cfg----→ Represents domain configuration file with Domains, Domain GIDs, and User and Group ID ranges defined

Usrmapc.db---→ Represents database conversion of the usrmap.cfg file and also tracks next UID/GID to be assigned

Usrmapusrc.db→ Represents User database records

Usrmapgrpc.db→ Represents Group database records

Sidname.db----→ Represents SID-to-Name mappings, used in outputting passwd and group files with correct information

**CELLERRA INTERNAL USRMAPPER:** Introduced NAS 5.2 with Usermapper version 3.1.4

**Note:** Usermapper 3.1.4 was released in NAS 5.1.20.4 (External only) and NAS 5.2.8.x (External or Internal)

Usermapper automatically starts and runs in Universal Mapping mode on Server\_2 for new installations. No manual configuration is required and other data movers will discover the Usrmapper service via broadcast over the internal IP network. No usrmap.cfg file is required, as all mappings come from a universal mapping space, regardless of the domain name. External Usrmapper configurations can also be migrated, or imported, to an Internal Usrmapper configuration [but not prior to NAS 5.2.14-0], in which case it would be

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
called Config Mode—this mode is completely managed manually similar to the rules applied to External Usermapper, hence is less desirable. You can also migrate Usermapper databases to Internal Usermapper without using a usrmapper.cfg file, and thereby maintain the ease of use provided by Universal Mode. Please note that Internal Usermapper is now integrated into DART and is no longer ‘versioned’.

## **VARIOUS INTERNAL USERMAPPER STATES:**

1.) Service Initialized but not currently running:

**# server\_usrmapper server\_2**

server\_2 : Usrmapper service: Initialized

Service Class: Primary

**# cat /nasmcd/quota/slot\_2/etc/usrmapper/usrmap.settings**

Identity=192.168.1.2,192.168.2.2

**Started=off**

2.) Service Running and Enabled:

**# server\_usrmapper server\_2**

server\_2 : Usrmapper service: Enabled

Service Class: Primary

**# cat usrmapper.settings**

Identity=192.168.1.2,192.168.2.2

**Started=on**

3.) Service Not Initialized or Setup on Server\_2:

**# server\_usrmapper server\_2**

server\_2 : Usrmapper service: Uninitialized

4.) Secondary Usermapper Service:

**# server\_usrmapper server\_2**

server\_2 : Usrmapper service: Enabled

Service Class: Secondary

Primary = 10.241.169.26 (c)

**# cat usrmapper.settings**

Identity=192.168.1.2,192.168.2.2

**Master=10.241.169.26**

Started=on

## **INTERNAL USRMAPPER MODES:**

### **Universal Mode & Config Mode**

→Universal Mode is the default mode for all new installs, and requires no manual setup or intervention—user & group assignments are pulled from a common pool of numbers starting at 32768 for Users and 32770 for Groups

→Config Mode is the mode used when a usrmapper.cfg file is used, as in an External-to-Internal Usermapper conversion—as such, the domains & ranges are pre-defined by the config file, and the user and group databases are imported from the External Usermapper system. Config Mode requires manual maintenance of Domains and UID/GID ranges similar to External Usermapper. Usermapper databases can also be imported to Internal Usermapper without the use of a usrmapper.cfg file, and therefore allows the service to run automatically in Universal Mode.

## **INTERNAL USRMAPPER CONFIGURATIONS/RESTRICTIONS:**

→Unlike external Usermapper, Internal Usermapper is completely integrated in DART code

→Only a single Primary Usermapper service is allowed in any namespace environment

→Do not run both Internal & External Usermapper services in the same environment

→Configure Primary Usermapper service first, then Secondary Usermapper services

→Configure only a single IUsrmapper instance on any given Celerra cabinet, with a single overall Primary—one Secondary can be configured for each additional Celerra cabinet in the same namespace environment

→Other Data Movers in a cabinet locate the Primary service via the Internal network using the autobroadcast feature

→For most environments, a single Usermapper service is adequate

→All new installations start IUsrmapper on Server\_2, but a different physical server could be configured manually if desired

→Usrmapper Service cannot be configured on VDM's [configured on a physical server basis only]

→Importing usrmapper.passwd & usrmapper.group files from external files can be done on both Primary & Secondary Usrmapper Servers

→Converting from External to Internal Usermapper is allowed

→Usrmapper will stop mapping UID/GIDs if rootfs reaches 95% capacity

→NASDB backs up the User and Group databases

→Use of IUsrmapper and other mapping methods, such as NIS, are NOT recommended, though can be done

**LOCATION OF INTERNAL USRMAPPER FILES:**

/.etc/usrmapper

**NAS 5.6 CONVERTS USERMAPPER FROM NAMEDB TO BERKELEY DATABASE FORMAT:**# **server\_dbms server\_2 -db -list Usermapper**

server\_2 : done

BASE NAME : Usermapper

Version : 1

State : Opened

Backup enabled : YES

Comment : Usermapper database. It allows to assign a new uid or gid to a given SID.

Size : 57344

Modification time : Wed May 20 14:02:15 2009

Creation time : Wed May 20 14:02:15 2009

TABLE NAME : aliases

Version : 1

State : Opened

Comment : This table allows to retrieve a domain name from one of his aliases

Size : 8192

Modification time : Wed May 20 14:02:15 2009

Creation time : Wed May 20 14:02:15 2009

-----output abridged, other Tables not shown-----

**INTERNAL USRMAPPER BACKUPS:**

**Note:** While the complete directory structure found in /.etc/usrmapper is not backed up, the NAS\_DB backup script does include the following commands [run hourly] to export the User and Group entries into respective files, from which a recovery could be made if required using the server\_usrmapper –Import option—the only problem that might be seen are for sites that have unique ranges defined outside of the “automatic domain mapping” space where a usrmap.cfg file would also need to be imported:

**/nas/server/slot\_2**

-rwxrwxr-x 1 nasadmin nasadmin 9157 Jan 29 14:37 um\_group

-rwxrwxr-x 1 nasadmin nasadmin 35689 Jan 29 14:37 um\_user

\$ **NAS\_DB/bin/server\_usrmapper \$server\_id -Export -user \$NAS\_DB/server/slot\_\$slot/um\_user**\$ **NAS\_DB/bin/server\_usrmapper \$server\_id -Export -group \$NAS\_DB/server/slot\_\$slot/um\_group****DUMPING INTERNAL USERMAPPER DATABASE TO SCREEN:**# **.server\_config server\_2 -v "usrmap dump"**

1101999939: USRMAP: 4: Usrmapper Database

1101999939: USRMAP: 4:

1101999939: USRMAP: 5: Usrmapper dump started.

1101999939: USRMAP: 4: UID DUMP:

1101999939: USRMAP: 4: S-1-5-15-42f831d9-66417ccd-28a68b82-45a:/:32768:32768:user nasadmin from domain mouse:/usr/S-1-5-15-42f831d9-66417ccd-28a68b82-45a:/bin/sh

1101999939: USRMAP: 4: GID DUMP:

1101999939: USRMAP: 4: S-1-5-15-42f831d9-66417ccd-28a68b82-203:/:32770:domain=20computers.mouse:

**VERIFYING USERMAPPER CONFIGURATION:**# **.server\_config server\_2 -v "usrmapserv info"**# **.server\_config server\_2 -v "usrmapper display"**

1166040224: SMB: 4: Usrmapper[0] = [127.0.0.1] state:active (auto discovered)

# **.server\_config server\_2 -v "usrmapper uid=32768"**

1173888525: SMB: 4: usrmapper uid/gid=0x8000 name=backdoor

S-1-5-15-7a50bbc6-60415a8a-6b635f23-46bc

**EXPORTING DB FILES TO DATA MOVER:**# **.server\_config server\_2 -v "usrmapserv export"**

1165608647: USRMAP: 5: Usrmapper dump started

1165608647: UFS: 6: inc ino blk cache count: nInoAllocs 1: inoBlk d7161304

1165608647: UFS: 6: inc ino blk cache count: nInoAllocs 2: inoBlk de1ba284

1165608647: USRMAP: 5: Dumping passwd in /.etc/usrmapper/usrmap.passwd.out

## **INTERNAL USRMAPPER EVENTS AND ERROR MESSAGES:**

USRMAP\_CreateEvent = 1 [DM sends to Control Station when database is created]  
USRMAP\_EnableEvent = 2 [Usermapper Service enabled]  
USRMAP\_DisableEvent = 3 [Usermapper Service disabled]  
USRMAP\_DestroyEvent = 4 [Database is destroyed]  
USRMAP\_FSQuotaExceedEvent = 7 [Usrmapper exceeds quota on rootfs: ‘running out of free inodes’ or ‘running out of free space’]

## **SERVER LOG ERRORS:**

2,000,000,001 [No more UID mappings available—only in usrmap.cfg file mode, not universal mode]  
2,000,000,002 [No more GID mappings available—same as above]  
2,000,000,006 [Primary Usermapper service is unavailable]  
2,000,000,010 [Mapping for GID or UID cannot be found]  
2,000,000,011 [Unknown request received, unsupported]  
2,000,000,013 [Invalid input error, V3 request is malformed]

## **INTERNAL USRMAPPER SERVER LOG ERROR LIST:**

|                           |                                                        |
|---------------------------|--------------------------------------------------------|
| INVALID_UID_REQUEST       | -2                                                     |
| INVALID_GID_REQUEST       | -3                                                     |
| REQUESTED_UID_IN_RANGE    | -4                                                     |
| REQUESTED_GID_IN_RANGE    | -5                                                     |
| INVALID_DOMAIN_RECORD     | -6                                                     |
| INVALID_USER_RECORD       | -7                                                     |
| INVALID_GROUP_RECORD      | -8                                                     |
| INVALID_RANGE             | -9                                                     |
| CONFIG_FILE_NOT_FOUND     | -10                                                    |
| MULTIPLE_CONFIG_LINES     | -11                                                    |
| INVALID_CONFIG_RECORD     | -12                                                    |
| OBSOLETE_OPTION           | -13                                                    |
| DEFAULT_GID_CLASH         | -14                                                    |
| UID_CLASH                 | -15                                                    |
| GID_CLASH                 | -16                                                    |
| DOMAIN_RANGE_INCONSISTENT | -17                                                    |
| INVALID_INTERVAL          | -18                                                    |
| INTERNAL_ERROR            | -19                                                    |
| DB_FULL                   | -20 [Example of this problem would be rootfs 95% full] |

## **EXAMPLE OF SERVER LOG ERRORS:**

SMB: 3: Usermapper[127.0.0.1](map2unix): error: -17 for group gid request  
2004-09-08 08:04:13: USRMAP: 3: gID: Cannot put mapping for S-1-5-15-2f61266b-649d4088-496b15a5-3fb (err=-17)

## **SERVER LOG ERROR FOR ROOT INODE 95% FULL:**

**SMB: 3: Usermapper[127.0.0.1](map2unix): error: -20 for user uid request**

**Note:** New Users cannot be mapped to the Usermapper database because the rootfs is 95% full

## **CLIENT ERROR WHEN TRYING TO APPLY PERMISSIONS FOR ROOT INODE 95% FULL:**

Security ID structure invalid error when attempting to assign security for the new user

## **SERVER LOG ERROR FOR MAPPING FAILURES:**

2005-10-10 09:01:19: USRMAP: 3: sid::name: Cannot found sid->name mapping (err=1)

**Note:** Error is logged when DART cannot determine the user or group name related to a SID

## **MAPPINGS CAN FAIL FOR USERS/GROUPS WITH MULTIPLE “.” WITHIN NAMES:**

User or Group Names with leading or multiple dots within the name can fail to map via Usermapper to the Celerra:

S-1-5-15-4862e393-28a68b82-6b635f23-a7170::\*:32780:.amsd.rfs002.imsource\_c.noe:

## **Server Log Errors:**

2005-06-20 16:57:01: USRMAP: 3: Cannot insert gid mapping for group94.bidon-5.bidon-5 (err=28)

1132589685: SMB: 3: Usermapper[192.168.1.2](map2unix): error: -8 for group gid request

**Note:** See primus emc122618

## **INTERNAL USRMAPPER AND NAS 5.2 NEW INSTALLS:**

New installs automatically configure server\_2 as the primary internal usrmapper service for the entire Celerra & enables the service—if data movers have not been configured with ‘-add usermapper=192.10.0.17’, Servers will broadcast over internal network to discover the Usrmapper service that should be used—use Server\_cifs command to verify. Internal Usrmapper will automatically failover properly to a Standby Server in the event of failover.

## **SETTING UP A SECONDARY INTERNAL USERMAPPER SERVICE:**

**Note:** A reason for setting up a Secondary Usermapper Service would be if there are two or more Celerra cabinets sharing the same database space. In this case, leave Server\_2 as the Primary on one of the sites and make Server\_2 of any additional sites as the Secondary.

1. Secondary Site: **\$server\_usermapper server\_2 -enable primary=172.169.2.5** [Point Secondary to IP address of Primary]
2. Primary Site: **\$server\_usermapper server\_2 -disable**
3. Primary Site: **\$server\_usermapper server\_2 -enable secondary=172.143.3.10** [Point Primary to Secondary]

**Note:** Pointing the Secondary Usermapper Service to the primary is all that is required. The Secondary database will update properly as it receives mappings requested from the Primary. In later codes, it no longer appears to be a requirement to specify the Secondary server to the Primary.

## **CONVERTING SECONDARY USERMAP TO PRIMARY ROLE:**

### **BEFORE CHANGE:**

**# .server\_config server\_2 -v "usrmapsvc info"**

```
1146154656: USRMAP: 4: Usrmapper service: Enabled
1146154656: USRMAP: 4: Service Class: Secondary
1146154656: USRMAP: 4: Primary = 192.1.6.202
1146154656: USRMAP: 4: Mapping mode: universal
1146154656: USRMAP: 4:
1146154656: USRMAP: 4: Usermapper addresses
1146154656: SMB: 4:
1146154656: SMB: 4: Usermapper[0] = [127.0.0.1] state:active (auto discovered)
1146154656: USRMAP: 4:
1146154656: USRMAP: 4: Cached enviroments
1146154656: USRMAP: 4:
1146154656: USRMAP: 4: 1 0x205c900 Env ???: Enabled
1146154656: USRMAP: 4:
1146154656: USRMAP: 4: Cached domains
1146154656: USRMAP: 4:
1146154656: USRMAP: 4: \(32768): 2
1146154656: USRMAP: 4: UID ffffffff :[0 8010-7fffffff] [1 10-7fff]
1146154656: USRMAP: 4: GID ffffffff :[0 8010-7fffffff] [1 10-7fff]
1146154656: USRMAP: 4: @(32768): 2
1146154656: USRMAP: 4: UID ffffffff :[0 8000-800f]
1146154656: USRMAP: 4: GID 8004 :[0 8001-800f]
1146154656: USRMAP: 4:
1146154656: USRMAP: 4: w2k(32768): 1
1146154656: USRMAP: 4: 2k3(32768): 1
1146154656: USRMAP: 4: lab(32768): 1
```

**# .server\_config server\_2 -v "usrmapsvc registry"**

```
1146154681: USRMAP: 4: Registry:
1146154681: USRMAP: 4:           Master = 192.1.6.202
1146154681: USRMAP: 4:           Identity = 192.168.2.2,192.168.1.2
1146154681: USRMAP: 4:           Started = on
1146154681: ADMIN: 4: Command succeeded: usrmapsvc registry
```

**# cat usrmap.settings**

```
Master=192.1.6.202
Identity=192.168.2.2,192.168.1.2
Started=off
```

**# ls -la usrmapc.db**

|            |   |      |     |                                                                              |
|------------|---|------|-----|------------------------------------------------------------------------------|
| lrwxr-xr-x | 1 | root | bin | 39 Feb 6 16:24 @ -> 32768{-1,32772}:32768:32783:32769:32783                  |
| lrwxr-xr-x | 1 | root | bin | 57 Feb 6 16:23 \-> 32768:32784:2147483647,16:32767:32784:2147483647,16:32767 |
| lrwxr-xr-x | 1 | root | bin | 13 Feb 6 16:23 2k3 -> 32768:0:0:0                                            |
| lrwxr-xr-x | 1 | root | bin | 13 Feb 6 16:23 lab -> 32768:0:0:0                                            |
| lrwxr-xr-x | 1 | root | bin | 13 Feb 6 16:24 w2k -> 32768:0:0:0                                            |

### **AFTER CHANGE:**

**# .server\_config server\_2 -v "usrmapsvc info"**

1146154883: USRMAP: 4: Usrmapper service: Enabled  
1146154883: USRMAP: 4: Service Class: Primary  
1146154883: USRMAP: 4: Database File system: Root  
1146154883: USRMAP: 4: Mapping mode: universal  
1146154883: USRMAP: 4:  
1146154883: USRMAP: 4: Usermapper addresses  
1146154883: SMB: 4:  
1146154883: SMB: 4: Usermapper[0] = [127.0.0.1] state:active (auto discovered)  
1146154883: USRMAP: 4:  
1146154883: USRMAP: 4: Cached enviroments  
1146154883: USRMAP: 4:  
1146154883: USRMAP: 4: 1 0x205c900 Env ??: Enabled  
1146154883: USRMAP: 4:  
1146154883: USRMAP: 4: Cached domains  
1146154883: USRMAP: 4:  
1146154883: USRMAP: 4: \(32768): 2  
1146154883: USRMAP: 4: UID ffffffff :[0 8010-7fffffff] [1 10-7fff]  
1146154883: USRMAP: 4: GID ffffffff :[0 8010-7fffffff] [1 10-7fff]  
1146154883: USRMAP: 4: @(32768): 2  
1146154883: USRMAP: 4: UID ffffffff :[0 8000-800f]  
1146154883: USRMAP: 4: GID 8004 :[0 8001-800f]  
1146154883: USRMAP: 4:  
1146154883: USRMAP: 4: w2k(32768): 1  
1146154883: USRMAP: 4: 2k3(32768): 1  
1146154883: USRMAP: 4: lab(32768): 1

### # .server\_config server\_2 -v "usrmapsvc registry"

1146154953: USRMAP: 4: Registry:  
1146154953: UFS: 6: inc ino blk cache count: nInoAllocs 1: inoBlk de57d984  
1146154953: USRMAP: 4: Identity = 192.168.2.2,192.168.1.2  
1146154953: USRMAP: 4: Started = on  
1146154953: ADMIN: 4: Command succeeded: usrmapsvc registry

### # cat usrmap.settings

Identity=192.168.2.2,192.168.1.2

Started=on

### # ls -la usrmappc.db -->After accessing with new Users

|            |   |      |     |                                                                              |
|------------|---|------|-----|------------------------------------------------------------------------------|
| lrwxr-xr-x | 1 | root | bin | 42 Apr 27 2006 @ -> 32768{32772,32774}:32768:32783:32769:32783               |
| lrwxr-xr-x | 1 | root | bin | 57 Feb 6 16:23 \-> 32768:32784:2147483647,16:32767:32784:2147483647,16:32767 |
| lrwxr-xr-x | 1 | root | bin | 13 Feb 6 16:23 2k3 -> 32768:0:0:0                                            |
| lrwxr-xr-x | 1 | root | bin | 13 Feb 6 16:23 lab -> 32768:0:0:0                                            |
| lrwxr-xr-x | 1 | root | bin | 13 Feb 6 16:24 w2k -> 32768:0:0:0                                            |

## CHANGING PRIMARY USERMAPPER SERVICE FROM SERVER 2 to SERVER 3

1. Verify state of current Usermapper Service:

\$ server\_usermapper server\_2

\$ .server\_config server\_2 -v "usrmapsvc info"

2. Disable Usermapper service on server\_2:

\$ server\_usermapper server\_2 -d

3. Set Debug Logging on Server\_3:

\$ .server\_config server\_2 -v "logsys set severity USRMAP=LOG\_DBG3"

4. Export the User and Group databases from Usermapper & do wordcount for number of lines:

\$ server\_usermapper server\_2 -Export -user users\_nov23

\$ server\_usermapper server\_2 -Export -group groups\_nov23

\$ wc -l users\_nov23 groups\_nov23

2415 4405

5. Create Tar Backup of Present Usermapper Directory:

# tar -zcpf /home/nasadmin/bak/usrmap.bak.tar.gz usrmapper

**Note:** Run command from /etc directory--tars & zips usrmapper directory to the location & filename indicated

6. Remove the Usermapper service and associated directory from server\_2:

**\$server\_usrmapper server\_2 -remove -all**

server\_2 : Warning: This operation will erase all user/group mappings.

7. Setup Primary Usrmapper service on Server\_3:

**\$server\_usrmapper server\_3 -enable**

**\$server\_usrmapper server\_3 -Import -user users\_nov23**

**\$server\_usrmapper server\_3 -Import -group groups\_nov23**

**\$server\_usrmapper server\_3**

server\_3 : Usrmapper service: Enabled

Service Class: Primary

8. Export Database files on Server\_3 and validate that line count is at least as great as the imported files:

**\$ server\_usrmapper server\_3 -Export -user users\_exp**

**\$ server\_usrmapper server\_3 -Export -group groups\_exp**

**\$wc -l users\_exp groups\_exp users\_nov23 groups\_nov23**

2415 4405 2415 4405

9. Delete all Internal Usrmapper Settings from netd files, including 127.0.0.1 from Server\_2:

**\$ cat /nas/server/slot\_\*/netd |grep -i usrmapper**

**\$ server\_cifs ALL -d usrmapper=192.168.1.2,192.168.2.2**

**\$ server\_cifs server\_2 -d usrmapper=127.0.0.1**

**Note:** Newer code versions may not require steps 9 or 10. The new primary Usrmapper service will be discovered as mapping requests are issued from each CIFS Server.

10. Add back all Internal Usrmapper settings for Server\_3 to Netd files:

**\$ server\_cifs ALL -add usrmapper=192.168.1.3,192.168.2.3**

**Note:** Delete server\_3's entry and readd as 127.0.0.1 (\$server\_cifs server\_3 -add usrmapper=127.0.0.1)

11. Verify Usrmapper service on Server\_3 by mapping new User/Group,export database if required to verify

12. Turn off debug logging USRMAP on Server\_3

## **INTERNAL USRMAPPER AND UPGRADES—IMPORTING FROM EXTERNAL USRMAPPER:**

**Caution:** External to Internal Usrmapper Upgrades are allowed for an existing “usrmap.cfg” file—see other section for more details. UIDs 1-32768 are reserved—first assignment would be 32768, GIDs start at 32770.

→Users will have a choice to continue using External Usrmapper Service [Defined as preferably a Primary-only service running on the Linux Control Station] or to Upgrade to the new Internal Usrmapper configuration via a procedure outlined in NAS 5.2 Tech Module, “Upgrading From External to Internal Usrmapper” May 2004.

**Migrations both from or to an External Usrmapper configuration are allowed using the following commands:**

**#server\_usrmapper server\_2 -export -user 1 -group <path>** [Import only to a primary Usrmapper server]

**#server\_usrmapper server\_2 -import -user 1 -group <path>** [Import database entries to the primary Usrmapper server]

**Caution:** Use –force only if importing and erasing existing entries—does not append!

**#/usrmap\_control dumpfilesall 3** [Used to output passwd/group files from External Usrmapper database]

## **EXPORTING INTERNAL USRMAPPER DATABASE:**

**# server\_usrmapper server\_2 -Export -user users**

server\_2 : done

**# server\_usrmapper server\_2 -Export -group groups**

server\_2 : done

→Used when migrating from the primary internal service to another source

→Used as a means to backup the database

→Used to collect information for troubleshooting

**Note:** Exports only from Primary Internal Usrmapper into version 3 [SIDs] “group” or “passwd” files.

## **PROCEDURE TO DISABLE SECMAP CACHING AND INTERNAL USRMAPPER:**

1. Disable Secmap Caching and Internal Usrmapper Autobroadcast by setting following params:

**#vi /nas/site/slot\_param**

**param usrmap autobroadcast=0**

**param cifs secmap.enable=0**

2. Create Backup Copy of User and Group Database from Internal Usrmapper:

**#server\_usrmapper server\_2 -Export -user /home/nasadmin/usrmap.passwd** [Export database to file as backup]

**#server\_usrmapper server\_2 -Export -user /home/nasadmin/usrmap.passwd** [Export database to file as backup]

**Note:** Must specify filename as usrmap.passwd and usrmap.group for export to succeed

3. Disable and remove the Internal Usrmapper Configuration:

**#server\_usrmapper server\_2 -disable**

**Note:** Command may not indicate that Internal Usermapper is disabled--ignore this and execute next command

**#server\_usrmapper server\_2 -remove -all**

**Caution:** This command deletes the Internal Usermapper directory from ./etc/usrmapper rootfs of DM. Please do not do this unless you have an alternate form of UID/GID Mapping in place in order to access Celerra CIFS Servers. In this example, customer reverted back to an already populated External Usermapper database. Trying to convert the Internal Usermapper passwd and group files into an acceptable format to accomodate importing into a new External Usermapper configuration is untested and unsupported at this time.

4. Stop CIFS Service

**#server\_setup server\_2 -P cifs -o stop**

5. Reboot Data Mover to put new params into effect and to disable SecMap caching

6. After Server is back up, manually delete the ./etc/secmap directory and all contents [#rm -Rf secmap]

7. Restart CIFS Service on Data Mover

8. Implement alternate means for mapping Users and Groups by UID/GID to the Celerra CIFS Server

**Note:** This procedure should no longer be the norm for NAS 5.2 or 5.3 installations or upgrades.

## **UNIVERSAL MODE OUTPUT:**

**\$ .server\_config server\_4 -v "usrmapserv info"**

1121362991: USRMAP: 4: Usrmapper service: Enabled

1121362991: USRMAP: 4: Service Class: Primary

1121362991: USRMAP: 4: Database File system: Root

1121362991: USRMAP: 4: **Mapping mode: universal**

1121362991: USRMAP: 4:

1121362991: USRMAP: 4: Usrmapper addresses

1121362991: SMB: 4:

1121362991: SMB: 4: Usermapper[0] = [127.0.0.1] state:active (auto discovered)

1121362991: USRMAP: 4:

1121362991: USRMAP: 4: Cached environments

1121362991: USRMAP: 4:

1121362991: USRMAP: 4: 1 0x9296df04 462 ??: Enabled

1121362991: USRMAP: 4: 1 0x1726d40 Env ??: Enabled

1121362991: USRMAP: 4: 1 0x74bc4704 353 ??: Enabled

1121362991: USRMAP: 4: 1 0x914e5104 461 ??: Enabled

1121362991: USRMAP: 4:

1121362991: USRMAP: 4: Cached domains

1121362991: USRMAP: 4:

**1121362991: USRMAP: 4: \(32768): 2**

**1121362991: USRMAP: 4: UID ffffffff :[0 834e-7fffffff] [1 10-7fff]**

**1121362991: USRMAP: 4: GID ffffffff :[0 a90d-7fffffff] [1 10-7fff]**

**1121362991: USRMAP: 4: @(32768): \*\* 31**

**1121362991: USRMAP: 4: UID 8344 :(0 8000-833d) [1 833e-834d]**

**1121362991: USRMAP: 4: GID a8ff :(0 8001-a8fc) [1 a8fd-a90c]**

1121362991: USRMAP: 4:

1121362991: USRMAP: 4: cs-group(32768): 1

1121362991: USRMAP: 4: csfb(32768): 1

1121362991: USRMAP: 4: globalusers(32768): 1

1121362991: USRMAP: 4: \_history\_sid\_range\_(32768): 1

1121362991: USRMAP: 4:

1121362991: ADMIN: 4: Command succeeded: usrmapserv info

## **WHAT DO THE \(32768) and @(32768) FIELDS REPRESENT FOR INTERNAL USERMAPPER:**

### **"usrmapserv info" output:**

**1096316892: USRMAP: 4: \(32768): 2** →This section represents unallocated or unused UID/GID ranges—i.e., ranges avail.

1096316892: USRMAP: 4: UID ffffffff :[0 8010-7ffffef] [1 10-7fff] →Section is where next UID range to be assigned will begin

1096316892: USRMAP: 4: GID ffffffff :[0 8030-7ffffef] [1 10-7fff] →Section is where next GID range to be assigned will begin

**1096316892: USRMAP: 4: @(32768): \* 2** →This section represents any Filled ranges and current UID/GID ranges in use

1096316892: USRMAP: 4: UID 8006 :(0 8000-800f) →Next UID to be assigned; currently allocated UID/GID ranges in use

1096316892: USRMAP: 4: GID 802e :(0 8001-801f) [1 8020-802f] →Next GID to be assigned; GID range filled; GID range used

**HOW DOES USRMAPD.DB FILE ENTRIES CORRESPOND TO USRMAPSVC INFO OUTPUT:****/nasmcd/quota/slot\_2/.etc/usrmapper****# ls -la usrmappc.db**

```
lrwxr-xr-x 1 root bin      54 Feb  7 15:30 @ -> 32768{32774,32814}:32768:32783!32769:32799,32800:32815
lrwxr-xr-x 1 root bin      57 Dec 12 09:01 \-> 32768:32784:2147483647,16:32767:32816:2147483647,16:32767
lrwxr-xr-x 1 root bin      13 Dec  6 13:44 2k3 -> 32768:0:0:0
lrwxr-xr-x 1 root bin      13 Dec 12 09:14 lab -> 32768:0:0:0
lrwxr-xr-x 1 root bin      13 Nov  9 20:06 ntdomain -> 32768:0:0:0
lrwxr-xr-x 1 root bin      13 Dec 12 09:01 w2k -> 32768:0:0:0
```

**Comments:** The @ line represents the default domain @ -> 32768, followed by next UID/GIDs to be mapped, respectively {32774,32814}, followed by currently allocated in-use range for UIDs (:32768:32783), then a separator !, followed by a filled range for GIDs (32769:32799), and the currently allocated in-use range for GIDs (32800:32815). The \-> 32768 line represents the unallocated or unassigned ranges still available for Internal Usermapper's default domain 32768, with :32784:2147483647 and 32816:2147483647 representing the available UID and GID ranges, respectively. The !, :, and ; all represent delimiters in the range lines.

**REPAIRING BROKEN RANGES IN USRMAPC.DB DIRECTORY:****GOOD RANGE ENTRIES:**

```
# ls -alrt /nasmcd/quota/slot_2/.etc/usrmapper/usrmapc.db
lrwxr-xr-x 1 root bin      13 Mar 13 21:07 lab -> 32768:0:0:0
lrwxr-xr-x 1 root bin      57 Mar 13 21:07 \-> 32768:32784:2147483647,16:32767:32784:2147483647,16:32767
lrwxr-xr-x 1 root bin      39 Mar 13 21:07 @ -> 32768{-1,32770}:32768:32783:32769:32783
```

**BAD RANGE ENTRIES:**

```
# ls -alrt
lrwxr-xr-x 1 root bin      21 Feb 28 12:49 \-> 32768:23232:26999:0:0 →This line is incomplete, no value for fields
lrwxr-xr-x 1 root bin      51 Mar  5 15:26 @ -> 27000{23224,-1}!22000:23215,23216:23231!27001:29999
```

**FIXING BAD RANGE LINE:****# mv \ bad-range** →This moves the bad range line called “\” and renames it to “bad-range”**# rm -f bad-range****# ln -s 32768:23232:26999:30000:32999 \** →Adds range back with complete ranges

```
# ls -alrt
lrwxr-xr-x 1 root bin      51 Mar  5 15:26 @ -> 27000{23224,-1}!22000:23215,23216:23231!27001:29999
lrwxrwxrwx 1 root  root    29 Mar  6 00:17 \-> 32768:23232:26999:30000:32999
```

**SETTING UP INTERNAL USERMAPPER IN CONFIG MODE (vs Universal Mode):**

1. Verify that current Internal Usermapper service has not been configured:

**# server\_usermapper server\_2**

server\_2 : Usrmapper service: Uninitialized

2. Import the usrmap.cfg file that defines Domains and Ranges in use for UIDs/GIDs (if needed, define path to the file):

**# server\_usermapper server\_2 -enable config=usrmap.cfg**

server\_2 : done

3. Import the User and Group database files, as originally exported from an External Usermapper configuration &amp; verify:

**# server\_usermapper server\_2 -Import -user usrmap.passwd**

server\_2 : done

**# server\_usermapper server\_2 -Import -group usrmap.group**

server\_2 : done

**# server\_usermapper server\_2**

server\_2 : Usrmapper service: Enabled

Service Class: Primary

4. Review Usermapper Configuration by using the following commands:

**# .server\_config server\_2 -v "usrmapserv info"**

1109605023: USRMAP: 4: Usrmapper service: Enabled

1109605023: USRMAP: 4: Service Class: Primary

1109605023: USRMAP: 4: Database File system: Root

**1109605023: USRMAP: 4: Configuration = usrmap.cfg****1109605023: USRMAP: 4: Mapping mode: config**

1109605023: USRMAP: 4:

1109605023: USRMAP: 4: Usermapper addresses

```
1109605023: SMB: 4:  
1109605023: SMB: 4: Usermapper[0] = [127.0.0.1] state:active (auto discovered)  
1109605023: USRMAP: 4:  
1109605023: USRMAP: 4: Cached enviroments  
1109605023: USRMAP: 4:  
1109605023: USRMAP: 4: 1 0x11f9c484 498 ??: Enabled  
1109605023: USRMAP: 4: 1 0x16e9738 Env ??: Enabled  
1109605023: USRMAP: 4: 1 0x11f9c284 506 ??: Enabled  
1109605023: USRMAP: 4:  
1109605023: USRMAP: 4: Cached domains  
1109605023: USRMAP: 4:  
1109605023: USRMAP: 4: \(32768): 2  
1109605023: USRMAP: 4: UID ffffffff :[0 8000-7fffffff] [1 10-7fff]  
1109605023: USRMAP: 4: GID ffffffff :[0 8000-7fffffff] [1 10-7fff]  
1109605023: USRMAP: 4:  
1109605023: USRMAP: 4: disney(5000): 1  
1109605023: USRMAP: 4: UID ffffffff :[0 1770-17d4]  
1109605023: USRMAP: 4: GID ffffffff :[0 1b58-1bbc]  
1109605023: USRMAP: 4: mouse(1000): 21  
1109605023: USRMAP: 4: UID ffffffff :[0 3e9-7cf]  
1109605023: USRMAP: 4: GID ffffffff :[0 7d0-c1c]  
1109605023: USRMAP: 4: _history_sid_range_(40000): 1  
1109605023: USRMAP: 4: UID ffffffff :[0 9c40-124f7]  
1109605023: USRMAP: 4: GID ffffffff :[0 9c41-afc7]  
1109605023: USRMAP: 4:  
1109605023: ADMIN: 4: Command succeeded: usrmapsvc info  
$ .server_config server_2 -v "usrmapserv registry"  
1109618346: USRMAP: 4: Registry:  
1109618346: USRMAP: 4: Config = usrmap.cfg  
1109618346: USRMAP: 4: Identity = 192.168.1.2,192.168.2.2  
1109618346: USRMAP: 4: Started = on  
1109618346: ADMIN: 4: Command succeeded: usrmapsvc registry
```

### **INTERNAL USERMAPPER: ADDING OR EXTENDING RANGES FOR CONFIG MODE:**

**Note:** Purpose is to edit the usrmap.cfg to add a new domain range or edit an existing range. The same rules apply to editing the usrmap.cfg as existed for External Usermapper. User or Group ranges can be “extended” using a single comma “,” between the original range and the new range being added.

1. Verify Usermapper Service is running prior to maintenance activity:

**\$ server\_usermapper server\_2**

```
server_2 : Usrmapper service: Enabled →Enabled means Usermapper Service is running normally  
Service Class: Primary
```

2. Export Users and Groups from database prior to making edits of usrmap.cfg file:

**# server\_usermapper server\_2 -Export -user users**

**# server\_usermapper server\_2 -Export -group groups**

3. Disable Usermapper Service prior to any editing of usrmap.cfg file:

**# server\_usermapper server\_2 -disable**

**# server\_usermapper server\_2**

```
server_2 : Usrmapper service: Initialized →Initialized means Usermapper Service has been disabled  
Service Class: Primary
```

**# server\_usermapper server\_2**

```
server_2 : Usrmapper service: Enabled
```

Service Class: Secondary

Primary = 192.1.4.250

**# .server\_config server\_2 -v "usrmapserv info"**

```
1146154500: USRMAP: 4: Usrmapper service: Initialized  
1146154500: USRMAP: 4: Service Class: Secondary  
1146154500: USRMAP: 4: Primary = 192.1.6.202  
1146154500: USRMAP: 4: Mapping mode: undefined  
1146154500: USRMAP: 4:
```

1146154500: USRMAP: 4: Usermapper addresses  
1146154500: USRMAP: 7: Broadcast internal usermapper addresses  
1146154500: USRMAP: 7: Interface broadcast addr #0: 192.168.2.255  
1146154500: USRMAP: 7: Interface broadcast addr #1: 192.168.1.255  
1146154500: USRMAP: 7: Send the usrmapper broadcast request  
1146154502: USRMAP: 4: Broadcast timeout, No answer received  
1146154502: USRMAP: 7: Usrmapper broadcast completed: 0/2

4. Navigate to the Data Mover's `/etc/usrmap` directory, then run the following commands, and compare—though the format is different, the `usrmap.cfg` file should match the `usrmapc.db` output:

**\$ cat usrmap.cfg**

RJ02:130000:130001:139000:139001:139999

**\$ ls -la usrmapc.db**

lrwxr-xr-x 1 root bin 34 Mar 17 11:39 rj02 -> 130000:**130001:139000:139001:139999**

**Note:** Above example shows the UID/GID ranges for the domain called “rj02”. 130001:139000 represents the UID range, while 139001:13999 represents the GID range.

5. Make a backup copy of the `usrmap.cfg` file before editing, then add the new domain range or extend an existing UID/GID range. In the following example we are going to extend the UID range for domain “rj02” using the “,” to separate the original UID range from the new UID range being added. Keep in mind that Usermapper rules do not allow you to use a UID range that overlaps any other UID ranges that may already be defined—same rule applies to GIDs:

**# cp usrmap.cfg usrmap.cfg.bak**

**# vi usrmap.cfg**

**Note:** Add new UID range, 600000-650000, to the domain “rj02”

6. Cat the `usrmap.cfg` file to verify that the edits look correct:

**# cat usrmap.cfg**

RJ02:130000:130001:139000,600000:650000:139001:139999

7. Set Debug Logging for Usrmapper Service:

**\$ .server\_config server\_2 “logsys set severity USRMAP=LOG\_DBG3”**

8. Start the Internal Usermapper Service and verify that it is ‘enabled’:

**# server\_usermapper server\_2 -enable**

server\_2 : done

**Note:** Do not use the ‘config=’ syntax when restarting Usermapper. Usermapper will automatically use the `usrmap.cfg` file in the `/etc/usrmap` directory.

**\$ server\_usermapper server\_2**

server\_2 : Usrmapper service: Enabled

Service Class: Primary

**Caution:** If Service does not restart, verify that the edits were correct and that none of the ranges added overlap. Turn on Debug Logging and run command again. If a range conflict occurs, the Server Log will show it with the following message:

**# server\_log server\_2 -s |grep -i usrmapping |tail**

2005-03-17 12:07:13: USRMAP: 3: Domain: uid, \_history\_sid\_range\_ clashes with the that of domain rj02

2005-03-17 12:07:13: USRMAP: 3: Cannot initialize database context

2005-03-17 12:07:13: USRMAP: 3: Usermapper service not started

9. With Usermapper Enabled, verify output of `usrmapc.db` file:

**# ls -la usrmapc.db**

lrwxr-xr-x 1 root bin 48 Mar 17 12:08 rj02 -> 130000:130001:139000,600000:650000:139001:139999

10. Verify Server Log for any errors or messages related to USRMAP, and have customer test by mapping a new test Windows User to the modified range.

11. Use additional commands to validate newly edited changes:

**# .server\_config server\_2 -v "usrmapping info"**

1111079909: USRMAP: 4: Usrmapper service: Enabled

1111079909: USRMAP: 4: Service Class: Primary

1111079909: USRMAP: 4: Database File system: Root

1111079909: USRMAP: 4: Configuration = `usrmap.cfg`

1111079909: USRMAP: 4: Mapping mode: config

1111079909: USRMAP: 4:

1111079909: USRMAP: 4: Usermapper addresses

1111079909: SMB: 4:

1111079909: SMB: 4: Usermapper[0] = [127.0.0.1] state:active (auto discovered)

1111079909: USRMAP: 4:

1111079909: USRMAP: 4: Cached enviroments  
1111079909: USRMAP: 4:  
1111079909: USRMAP: 4: 1 0x16e9738 Env ???: Enabled  
1111079909: USRMAP: 4:  
1111079909: USRMAP: 4: Cached domains  
1111079909: USRMAP: 4:  
1111079909: USRMAP: 4: \(32768): 2  
1111079909: USRMAP: 4: UID ffffffff :[0 8000-7fffffff] [1 10-7fff]  
1111079909: USRMAP: 4: GID ffffffff :[0 8000-7fffffff] [1 10-7fff]  
1111079909: USRMAP: 4:  
1111079909: USRMAP: 4: pegasus(100000): 47  
1111079909: USRMAP: 4: UID ffffffff :[0 186a1-1a9c8]  
1111079909: USRMAP: 4: GID ffffffff :[0 1a9c9-1adaf]  
1111079909: USRMAP: 4: am-sd-nt(50000): 25  
1111079909: USRMAP: 4: UID ffffffff :[0 c351-e678]  
1111079909: USRMAP: 4: GID ffffffff :[0 e679-ea5f]  
1111079909: USRMAP: 4: telemar(20000): 10396  
1111079909: USRMAP: 4: UID ffffffff :[0 4e21-7148] [1 8823-892f] [2 249f0-26d18]  
1111079909: USRMAP: 4: GID ffffffff :[0 7149-752f] [1 8000-8085] [2 26d19-270ff]  
1111079909: USRMAP: 4: rj06(40000): 160  
1111079909: USRMAP: 4: UID ffffffff :[0 9c41-bf68]  
1111079909: USRMAP: 4: GID ffffffff :[0 bf69-c34f]  
1111079909: USRMAP: 4: \_history\_sid\_range\_(500): 45  
1111079909: USRMAP: 4: UID ffffffff :[0 61a80-6ddd0]  
1111079909: USRMAP: 4: GID ffffffff :[0 7a120-7dfa0]  
-----abridged-----  
1111079909: USRMAP: 4: rj01(30000): 6761  
1111079909: USRMAP: 4: UID ffffffff :[0 7531-8822] [1 493e0-4baf0]  
1111079909: USRMAP: 4: GID ffffffff :[0 9859-9c3f] [1 4baf1-4c2c0]  
1111079909: USRMAP: 4: rj02(130000): 9  
1111079909: USRMAP: 4: UID ffffffff :[0 1fb01-21ef8] [1 927c0-9eb10]  
1111079909: USRMAP: 4: GID ffffffff :[0 21ef9-222df]  
1111079909: USRMAP: 4:  
1111079909: ADMIN: 4: Command succeeded: usrmapsvc info  
**# .server\_config server\_2 -v "usrmapserv registry"**  
1111079838: USRMAP: 4: Registry:  
1111079838: USRMAP: 4: Config = usrmap.cfg  
1111079838: USRMAP: 4: Identity = 192.168.1.2,192.168.2.2  
1111079838: USRMAP: 4: Started = on  
1111079838: ADMIN: 4: Command succeeded: usrmapsvc registry

## **USERMAPPER DB PROBLEM OUTLINED IN PRIMUS EMC104375:**

**# .server\_config server\_2 -v "usrmapserv info"**  
1130447354: USRMAP: 4: Usrmapper service: Initialized  
1130447354: USRMAP: 4: Service Class: Primary  
1130447354: USRMAP: 4: Database File system: Root  
1130447354: USRMAP: 4: **Configuration = usrmap.cfg**  
1130447354: USRMAP: 4: **Mapping mode: undefined**  
1130447354: USRMAP: 4:  
1130447354: USRMAP: 4: Usermapper addresses  
1130447354: USRMAP: 7: Broadcast internal usermapper addresses  
1130447354: USRMAP: 7: Interface broadcast addr #0: 192.168.1.255  
1130447354: USRMAP: 7: Interface broadcast addr #1: 192.168.2.255  
1130447354: USRMAP: 7: Send the usrmapper broadcast request  
1130447356: USRMAP: 4: Broadcast timeout, No answer received  
1130447356: USRMAP: 7: Usrmapper broadcast completed: 0/2  
1130447356: USRMAP: 4:  
1130447356: USRMAP: 4: Cached enviroments  
1130447356: USRMAP: 4:  
1130447356: USRMAP: 4: 1 0x1725708 Env ???: Initialized

1130447356: USRMAP: 4:  
1130447356: USRMAP: 4: Cached domains  
1130447356: USRMAP: 4:  
1130447356: USRMAP: 4: \(32768): 2  
1130447356: USRMAP: 4: UID ffffffff :[0 8000-7fffffff] [1 10-7fff]  
1130447356: USRMAP: 4: GID ffffffff :[0 8760-7fffffff] [1 10-7fff]  
1130447356: USRMAP: 4:  
1130447356: USRMAP: 4: pegasus(100000): 1  
1130447356: USRMAP: 4: UID ffffffff :[0 186a1-1a9c8] →Brackets would normally indicate Config Mode range  
1130447356: USRMAP: 4: GID ffffffff :[0 1a9c9-1adaf] →But, UID/GID ffffffff means that no mappings have yet been made  
1130447356: USRMAP: 4: am-sd-nt(50000): 1  
1130447356: USRMAP: 4: UID c36a :[0 c351-e678] →c36a = 50,026 & is the next UID to be assigned from [ ] range  
1130447356: USRMAP: 4: GID e694 :[0 e679-ea5f]  
1130447356: USRMAP: 4: telemar(20000): 1  
1130447356: USRMAP: 4: UID 26169 :[0 4e21-7148] [1 8823-892f] [2 249f0-26d18]  
1130447356: USRMAP: 4: GID ffffffff :(0 7149-752f) (0 8000-8085) (0 8000-875f) (0 26d19-270ff) →This particular line shows the bug. When bug is present, the GID range has become automatically extended. From previous usrmap.cfg file, we know that the other ( ) ranges are correct—the fact that they now have a ( ) around the ranges vs. a [ ] is another indicator of the bug and means that Usrmapper has been stopped and now cannot be restarted, for the GID range goes to ffffffff as it no longer knows what the next GID should be. Meanwhile, the damage here is that 8000-875f has been dynamically assigned by Usrmapper, and conflicts with previous range definitions in the usrmap.cfg file, meaning that range rules have been violated and usrmapper cannot restart:

#### **Server Log Error:**

2005-10-27 14:43:43: USRMAP: 3: Decreasing the max uid or gid for domain telemar below already allocated ranges is not allowed

### **REPAIRING INTERNAL USERMAPPER EMC104375:**

1. Verify current Internal Usermapper state:

**# server\_usermapper server\_2**

server\_2 : Usrmapper service: Initialized -->Indicates Service is not running  
Service Class: Primary

2. Attempt to restart Service:

**# server\_usermapper server\_2 -enable**

3. Review Server Log if service fails to go to "Enabled" state

2005-04-01 15:38:57: USRMAP: 4: Starting usermapper service

2005-04-01 15:38:57: USRMAP: 3: Decreasing the maximum uid or gid for domain karma below already allocated range is not allowed: 200{1122,33128}:1000:10000!1000:2000;32768:33119,33120:33135

2005-04-01 15:38:57: USRMAP: 3: Cannot initialize database context

2005-04-01 15:38:57: USRMAP: 3: Usermapper service not started

2005-04-01 15:38:59: USRMAP: 4: Broadcast timeout, No answer received

**Note:** Output indicates that Internal Usermapper cannot restart because of range violation. See emc104375 for a more complete description—problem is triggered only after Primary Usermapper Server has been rebooted. Once a User or Group range has been filled, Usermapper will revert to Universal Mapping mode and begins assigning UIDs or GIDs from the 32768 range, regardless of whether other domain ranges have been defined to use this range, therefore leading to the problem described herein. Usermapper will continue to run in this condition until the service is stopped. Once stopped, it will not restart.

4. Review output of "usrmapsvc info" to confirm problem:

1112393920: USRMAP: 4: Usrmapper service: Initialized

1112393920: USRMAP: 4: Service Class: Primary

1112393920: USRMAP: 4: Database File system: Root

1112393920: USRMAP: 4: Configuration = usrmap.cfg -->CONFIG Mode

1112393920: USRMAP: 4: Mapping mode: undefined -->Mapping mode undefined indicates problem

5. Export Users and Groups databases from Internal Usermapper:

**\$server\_usermapper server\_2 -Export -user users**

**\$server\_usermapper server\_2 -Export -group groups**

6. Download Users, Groups, & usrmap.cfg files. Take output from "usrmapsvc info" and compare to contents on usrmap.cfg and determine which ranges have been affected by the current problem. Determine whether database can be repaired:

1112393922: USRMAP: 4: karma(200): 1

1112393922: USRMAP: 4: UID 462 :[0 3e8-2710]

1112393922: USRMAP: 4: GID 8168 :(0 3e8-7d0) (0 8000-815f) [1 8160-816f]

**# cat usrmap.cfg**

karma,karma.com:200:1000:10000:1000:2000

**Note:** The range that has reverted to Universal Mapping mode is indicated by the ( ) parentheses, indicating that ranges have been filled, followed by [ ] brackets, indicating that a new range has been dynamically assigned. Checking the hex value for "8168" shows that the last GID assigned was 33,128. In this example, GIDs 32,768 - 33,128 have been assigned to the 'Karma' domain--review the Group file to validate. Also in this example, the fixed domain range "lesskarma.com" has been violated by Universal Mapping mode and is the reason that Internal Usermapper cannot restart. As a matter of record here, the "lesskarma.com" range had not previously been used, meaning that a database repair was feasible without deleting any Group mappings.

7. Modify the usrmap.cfg file to reflect the actual mappings that occurred prior to Usermapper failing and to separate the ranges so that no overlapping occurs in the config file--Usermapper will not start if ranges overlap:

### # vi usrmap.cfg

```
karma,karma.com:200:1000:10000:1000:2000,32768:33127,54801:56000
```

```
lesskarma.com:980:206000:207999:32700:32767,33128:33200
```

**Note:** This example shows a corrected config file and remaps the ranges to reflect what 'Karma' was assigned while removing those assignments from the 'Lesskarma' range.

8. For any User or Group ranges that have been automatically extended, review the appropriate database file for duplicate entries. In this casestudy, all mappings that occurred when Usermapper went into Universal Mapping mode, were duplicated into two separate sections in the groups file. Validate and then remove the duplicated entries.

9. Once the Users, Groups, and usrmap.cfg files have been validated and made ready for a fresh import, test on a lab box to validate the Import process, the usrmap.cfg file, and that Usermapper will 'Enable' properly. Set debug logging for USRMAP and conduct the following steps:

a. # .server\_config server\_2 -v "logsys set severity USRMAP=LOG\_DBG3"

b. # server\_usermapper server\_2 -enable config=usrmap.cfg

server\_2 : done

c. Check Server Log for Success:

```
2005-04-01 19:21:19: ADMIN: 4: Command succeeded: usrmapsvc enable config=usrmap.cfg
```

d. Dump Usermapper Configuration, review output to verify Config Mode, and examine all domain ranges:

# .server\_config server\_2 -v 32768 "usrmapserv info"

```
1112401312: USRMAP: 4: Usrmapper service: Enabled
```

```
1112401312: USRMAP: 4: Service Class: Primary
```

```
1112401312: USRMAP: 4: Database File system: Root
```

```
1112401312: USRMAP: 4: Configuration = usrmap.cfg
```

```
1112401312: USRMAP: 4: Mapping mode: config
```

# .server\_config server\_2 -v 32768 "usrmapserv registry"

```
1112401322: USRMAP: 4: Registry:
```

```
1112401322: USRMAP: 4:           Config = usrmap.cfg
```

```
1112401322: USRMAP: 4:           Identity = 192.168.1.2,192.168.2.2
```

```
1112401322: USRMAP: 4:           Started = on
```

```
1112401322: ADMIN: 4: Command succeeded: usrmapsvc registry
```

e. Import Users & Groups databases & Verify Service:

# server\_usermapper server\_2 -Import -user users

server\_2 : done

# server\_usermapper server\_2 -Import -group groups

server\_2 : done

# server\_usermapper server\_2

server\_2 : Usrmapper service: Enabled

Service Class: Primary

f. Verify Imports in Server Log:

```
2005-04-01 18:49:29: USRMAP: 4: 2884 new users imported
```

```
2005-04-01 18:49:43: USRMAP: 4: 2544 new groups imported
```

```
2005-04-01 18:49:43: ADMIN: 4: Command succeeded: usrmapsvc import
```

g. Re-export databases & verify against original Files:

# server\_usermapper server\_2 -Export -user user\_post

server\_2 : done

# server\_usermapper server\_2 -Export -group group\_post

server\_2 : done

h. Verify Usermapper Service, Check Server Log, ouput Usrmapsvc Info & Server\_Cifs:

\$ server\_usermapper server\_2

server\_2 : Usrmapper service: Enabled

Service Class: Primary

**Server Log:**

2005-04-04 11:04:25: USRMAP: 4: Usermapper[127.0.0.1] now available

**\$ .server\_config server\_2 -v "usrmapserv info" |head -20**

1112703824: USRMAP: 4: Usrmapper service: Enabled

1112703824: USRMAP: 4: Service Class: Primary

1112703824: USRMAP: 4: Database File system: Root

1112703824: USRMAP: 4: Configuration = usrmap.cfg

1112703824: USRMAP: 4: Mapping mode: config

**\$ server\_cifs server\_2**

Usermapper[0] = [127.0.0.1] state:active (auto discovered)

10. Upload modified database & usrmap.cfg files & Rebuild Internal Usermapper on Production System:

a. **# cp -Rf usrmapper /home/nasadmin** (save copy of original database)

b. **# server\_usermapper server\_2 -disable**

c. **# server\_usermapper server\_2 -remove -all**

Yes

**Note:** This will delete the usrmapper directory and completely uninstall Usermapper

d. Follow re-import & verification steps outlined in Step 9 above

11. Verify Internal Usermapper operation with customer

## **NAS 5.2.14 EXTERNAL TO INTERNAL USRMAPPER UPGRADE PROCEDURE**

**Note:** Minimum NAS 5.2.14.0 and External Usrmapper 3.1.4. Upgrade to NAS 5.2.14.0 or higher in order to use External-to-Internal Usrmapper conversion procedure. For NAS Upgrades from 5.1 to 5.2, External Usrmapper will continue to run normally under the control of the nas\_mcd daemon and, by default, will NOT initialize the Internal Usrmapper service on Server\_2.

### **EXTERNAL-TO-INTERNAL USERMAPPER CONVERSION CAVEATS:**

→ Ensure that Data Mover rootfs contains sufficient space and Inodes to support both SecMap Caching and Internal Usrmapper. Each User and Group mapped to the Celerra by Internal Usrmapper requires (4) rootfs inodes to support Internal Usrmapper mappings & (2) rootfs inodes to support SecMap Caching.

**Notes:** Do not convert to Internal Usrmapper without calculating size of current Usrmapper database and available inodes on rootfs of data mover! Be aware that you can disable secmap caching and still use Internal Usrmapper. Be advised that NAS 5.6 eliminates the need for Inodes for Secmap or Usermapper entries, and therefore, the space and Inode limitations are no longer a factor. 5.6 will convert the flat inode-based “namedb” databases to a Berkeley database.

→ Do not convert from External to Internal Usrmapper with NAS Versions prior to 5.2.14-0 [See example of error message seen below] or Usermapper Versions prior to 3.1.4 [With 3.1.3 the -disable command does not work properly]

**#server\_usermapper server\_2 -enable config=/nas/cifs/usrmapperV3/linux/usrmap.cfg**

server-2 :

Error 2225 : server\_2 -enable config=/nas/cifs/usrmapperV3/linux/usrmap.cfg: invalid operation

→ Beginning with NAS 5.2, for installations that are not yet running Internal Usermapper configurations, customers will have the option to continue to use the existing External Usermapper configuration or to upgrade to an Internal Usermapper configuration. However, please note that there are rules, procedures, and limitations to be considered when performing an "External to Internal" Usermapper migration that must be reviewed and understood. It should be understood that the primary documentation for the External-to-Internal Usermapper upgrade is the published Engineering Technical Module--"Upgrading from External Usermapper 3.1 to Internal Usermapper" Version 5.3 August 2004 Rev A01--use this doc as reference material only--use the Engineering Tech Module steps to perform the Upgrade. The Tech Module information should be referenced & available via the CCA [Change Control Activity] Review Process.

## **UPGRADING (CONVERTING) FROM EXTERNAL TO INTERNAL USERMAPPER SERVICE:**

### **1. Verify Status of Current Configuration Prior to Converting:**

**#ps -ef |grep usrm** [verifies that External Usrmapper Service is running]

root 7816 7336 0 May13 ? 00:00:00 /nas/cifs/usrmapperV3/linux/usrm

**#server\_usermapper server\_2**

server\_2: Usermapper Service: Uninitialized

**Note:** Internal Usermapper Service should still be “uninitialized” at this point. If not, run #server\_usermapper server\_2 -disable

**#server\_cifs server\_2**

Usermapper[0] = [10.241.169.43] state:available

**Note:** Verify that DM's point to External Usermapper Service, in this case '10.241.169.43' represents CS0's External IP address

### **2. Enable Server\_2 as a Secondary Usermapper and point to IP Address of External Usermapper Service:**

**#server\_usermapper server\_2 -enable primary=10.241.169.43** [syntax to use when not specifying an existing usrmap.cfg file]

#server\_usrmapper server\_2 -enable primary=10.241.169.43 config=/nas/cifs/usrmapperV3/linux/usrmap.cfg [syntax when using specific usrmap.cfg file settings]

server\_2 : done

**Comment:** You can import usrmap.passwd and usrmap.group databases from an External Usermapper source and allow Internal Usermapper to automatically assign new UID/GIDs from the default 'Universal Mapping' ranges, or, you can reference the path to the external usrmap.cfg file to retain the original ranges and databases defined in the External Usermapper configuration, and then import the usrmap.group and usrmap.passwd files from the External Usermapper database as shown in Step 6 below. Best practices dictates that you DO NOT use the usrmap.cfg file unless absolutely required by the Customer. Usermapper maintenance is virtually eliminated if not using the usrmap.cfg file (CONFIG Mode) and instead running in UNIVERSAL Mode.

### 3. Verify Functionality of Secondary Usermapper Service:

#server\_usrmapper server\_2

server\_2 : Usrmapper Service : Enabled

Service Class: Secondary

Primary = 10.241.169.43 <c> → "c" means that Secondary has contacted the Primary service

### 4. Delete the External Usrmapper Reference from ALL CIFS Data Movers:

#server\_cifs server\_2 -delete usrmapper=10.241.169.43

### 5. Dump the entire External Usrmapper database in Version 3 format (SIDs listed first):

#/nas/cifs/usrmapperV3/linux/usrmap\_control dumpfilesall 3

### 6. Take the Version 3 dumpfiles [usrmap.passwd & usrmap.group] and Import the files one at a time into Internal Usrmapper's database:

#server\_usrmapper server\_2 -Import -user /nas/cifs/usrmapperV3/linux/usrmap.passwd

#server\_usrmapper server\_2 -Import -group /nas/cifs/usrmapperV3/linux/usrmap.group

server\_2 : done

**Note:** If importing to an already existing database, you must use the "-force" switch if you wish to overwrite--otherwise, entries are appended to the existing database. Check Server Log after importing is complete to check for import errors.

### 7. Change Usrmapper Service on Server\_2 from Secondary to Primary to complete conversion to Internal Usrmapper:

#server\_usrmapper server\_2 -disable

#server\_usrmapper server\_2 -enable primary=127.0.0.1

server\_2 : done

**Note:** Above command enables Server\_2 as the Primary Internal Usermapper service and uses its loopback address to reference itself as the Primary. All other Data Movers would 'discover' the Primary Internal Usermapper service via broadcast across the internal IP network.

### 8. Verify Primary Service on Server\_2:

#server\_usrmapper server\_2

server\_2 : Usrmapper service: Enabled

Service Class: Primary

#server\_cifs server\_2

Usermapper auto broadcast enabled

Usermapper[0] = [127.0.0.1] state:active (auto discovered)

### 9. Disable External Usrmapper Service on CS0:

- comment out "Usermapper Service" from /nas/sys/nas\_mcd.cfg file
- Stop & Restart NAS Services and verify that External Usrmapper does not restart
- Rename External Usrmapper executables so that the process does not inadvertently restart
- Conduct tar backup of external usrmapper directory and archive to safe location
- Ensure local passwd and group files are removed from data movers--stop & restart CIFS to flush cache.

### 10. Check SecMap Caching and Internal Usermapper Broadcast Parameters:

**Note:** Run the appropriate commands or check the param files to verify the current settings for the SecMap caching mechanism and the Internal Usermapper Broadcast mechanism. Due to certain limitations of earlier NAS 5.2 versions, either one or both of these parameters may have been disabled--review the 'Upgrade Caveats' outlined at the beginning of this solution prior to changing any params. If the following params are set in the param files, remove them and reboot the Data Mover(s).

param usrmap autobroadcast=0 [Indicates that Internal Usermapper Autobroadcast feature is disabled; =1 to enable broadcast]

param cifs secmap.enable=0 [Indicates that SecMap caching mechanism is disabled; =1 to enable SecMap caching]

### 11. Verify Internal Usrmapper Configuration Files:

- Go to root of datamover and verify contents of "usrmapper" directory [/nasmd quota/slot\_2/etc/ usr mapper]:

#### Typical Internal Usrmapper Directory Contents after Converting from External Usrmapper:

/nasmd quota/slot\_2/etc/ usr mapper

```
drwxr-xr-x 2 root bin 1024 May 14 13:42 aliases.db
drwxr-xr-x 2 root bin 80 May 14 12:52 extidxname.db
drwxr-xr-x 2 root bin 80 May 14 12:52 extuidgidxsid.db
drwxr-xr-x 2 root bin 1024 May 14 13:52 groupmapnamesid.db
```

```
drwxr-xr-x 2 root bin 1024 May 14 13:53 idxname.db
drwxr-xr-x 2 root bin 2048 May 14 13:53 sidname.db
drwxr-xr-x 2 root bin 80 May 14 12:52 uidgidsid.db
drwxr-xr-x 2 root bin 1024 May 14 13:53 usrrpmapnamesid.db
drwxr-xr-x 2 root bin 1024 May 14 13:53 usrmapc.db
-rw-r--r-- 1 root bin 117 May 14 12:52 usrmap.cfg
drwxr-xr-x 2 root bin 1024 May 14 13:52 usrmappc.db
-rw-r--r-- 1 root bin 62 May 14 13:05 usrmap.settings
drwxr-xr-x 2 root bin 1024 May 14 13:53 usrmapusrc.db
```

**Note:** All User and Group mappings that were imported can be found in “usrmapusrc.db” and “usrmapgrpc.db”, respectively.

b.) Verify entries in Usrmap.Settings file:

```
# cat usrmap.settings
```

Config=usrmap.cfg

Identity=192.168.1.2,192.168.2.2

Started=on

c.) Export Internal Usermapper Database and compare beginning and ending entries between External and Internal Usermapper:

```
#server_usrmapper server_2 -Export -user /home/nasadmin/usrmap.passwd
```

```
#server_usrmapper server_2 -Export -group /home/nasadmin/usrmap.group
```

## 12. Test Internal Usermapper with New User Account & Verify UID/GID Counters are incrementing correctly:

```
# ls -la usrmapc.db
```

```
lrwxr-xr-x 1 root bin 57 May 14 12:52 \-> 32768:32768:2147483631,16:32767:32768:2147483631,16:32767 [default ranges]
```

```
lrwxr-xr-x 1 root bin 27 May 14 13:42 _history_sid_range_-> 600:60000:65000:51001:52999 [imported History Sid range]
```

**lrwxr-xr-x 1 root bin 22 May 14 13:53 ts2 -> 10{50,97}:40:80:90:120 ->Imported domain range-next available UID/GID counters are in red**

**Note:** Testing access with New User--UID/GID counters should increment properly, starting at UID/GID number in sequence as found in the imported usrmap.passwd or usrmap.group files. This step is not necessary when using the 'Universal Mapping Mode'. In otherwords, this step would only be prudent if importing an existing usrmap.cfg file when upgrading from External to Internal.

### EXAMPLE OF LAST ENTRY IN EXPORTED PASSWD FILE (From step 11c):

```
S-1-5-15-1f6d0078-63f6f323-65d637a8-488:*:48:10:user mapper8 from domain ts2:/usr/S-1-5-15-1f6d0078-63f6f323-65d637a8-488:/bin/sh
```

**AFTER ACCESS WITH NEW USER:** # ls -la usrmapusrc.db |tail -4

```
23-65d637a8-489 -> S-1-5-15-1f6d0078-63f6f323-65d637a8-489:*:49:10:user mapper9 from domain ts2:/usr/S-1-5-15-1f6d0078-63f6f323-65d637a8-489:/bin/sh
```

```
lrwxr-xr-x 1 root bin 130 May 14 13:53 S-1-5-15-1f6d0078-63f6f3
```

**Note:** Usrmapc.db directory shows that next UID to be assigned for “ts2” domain is {50,97} and next GID is 97

## **DISABLING & REMOVING AN EXISTING INTERNAL USRMAPPER CONFIGURATION:**

1. Verify Existing Usermapper operation:

```
# server_usrmapper server_2
```

server\_2 : Usrmapper service: Enabled

Service Class: Primary

2. Disable Internal Usermapper & Verify:

```
# server_usrmapper server_2 -disable
```

server\_2 : done

```
# server_usrmapper server_2
```

server\_2 : Usrmapper service: Initialized →Note that service goes from “Enabled” to “Initialized”

Service Class: Primary

3. Remove the Usrmapper Database from the Server & Verify:

```
# server_usrmapper server_2 -remove -all
```

server\_2 : Warning: This operation will erase all user/group mappings.

CIFS users may lose access.

Continue(Y/N):

done

**Note:** Prior to deleting, export the Users and Groups from Usermapper, and copy usrmap.cfg, since the –remove –all command will delete the entire usrmapper directory.

```
# .server_config server_2 -v "ls -l .etc" |grep -i usrmapper
```

<directory is not found>

4. Since server\_cifs output still shows Autobroadcast feature enabled, disable by entering the following param and rebooting the Server:

## Usermapper auto broadcast enabled

#vi /has/site/slot\_param

### param usrmap autobroadcast=0

5. Delete the Internal Usermapper Reference from CIFS Configuration on all Data Movers:

**Usermapper[0] = [127.0.0.1] state:available (internal)**

**# server\_cifs server\_2 -delete usermapper=127.0.0.1**

server\_2 : done

6. Reboot Servers as allowed to disable the autobroadcast feature

7. As a final note, disable and delete the SecMap Cache if required by disabling the SecMap feature in the param file, rebooting the Server, then manually removing the ./etc/secmap directory (keep an archive copy if possible)

## USRMAPPER 3.1.4: NAS 5.2.8-2006

**./usrmap\_control -disable** [Disables and stops Usermapper service, temporarily]

The Usermapper server mapping functions are now DISABLED!

**Note:** Usrmapper mappings are disabled and Users cannot map to Celerra. However, the process still shows as running:

#ps -ef |grep usrm

### Usrmapper Log Entry:

“User has DISABLED mapping functions of the Usermapper server.”

**./usrmap\_control status**

Status of local Usermapper Server:

IP Address \_\_\_\_\_ 143.242.14.72

Primary/Secondary \_\_\_\_\_ Primary

Secondary Usermapper IPs\_\_\_\_\_

State \_\_\_\_\_ Enabled

**./usrmap\_control -enable**

The Usermapper Server mapping functions are now ENABLED!

Usermapper Service will now restart...

**Note:** Recognizes that usrmapper is still running under the control of the NAS MCD. Once ‘enabled’, actually stops the Service and lets NAS restart it. If External Usrmapper has been disabled, and Internal Usrmapper enabled for first time, creates following:

**/etc/usrmapper**

(11) default db directories & creates usrmap.settings file

**/etc/secmap**

@gid [Contains GIDs that have been mapped into Secure Map]

S-1-5-15....[Contains SIDs of Users and Groups that have been mapped]

@uid [Contains UIDs that have been mapped]

**Note:** Once Users or Groups have accessed the Celerra and obtained a “secmap” mapping, the Usrmapper Service can be disabled, Users and Groups will still have access to CIFS. The “secmap” database permanently maps a User or Group SID to the data mover and is persistent on Server reboots.

## INITIALIZING INTERNAL USRMAPPER DB WITHOUT “ENABLING” ROLE (i.e., Primary or Secondary):

**Note:** Use following procedure to create Secondary Usermapper server with fully populated database from Primary side

1. Stop CIFS on Secondary Usermapper server

2. Disable Secondary Usrmapper service

**#server\_usrmapper server\_2 -disable**

3. Remove Secondary Usermapper database configuration:

**#server\_usrmapper server\_2 -remove -all**

4. Remove, copy, or tar Secmap directory from ./etc/secmap to /home/nasadmin as a backup

**#rm -Rf secmap**

**# cp -Rip secmap /home/nasadmin/tm/**

**#tar -zcpf /home/nasadmin/bak/usrmap.bak.tar.gz usrmapper** [Run from ./etc directory]

5. Reboot Server\_2 to clear cache

6. Rebuild Secondary Usermapper using “usrmapsrv init” command to create the database structure without starting the Usermapper Service, which in this case is crucial so that the -Export functions will work later on

**# .server\_config server\_2 -v "usrmapsrv init"**

1100566384: USRMAP: 6: Opening the database...

1100566384: USRMAP: 6: Domain cache unloaded

```
1100566384: USRMAP: 7: \domain loaded
1100566384: USRMAP: 6: Usermapper domain manager started
1100566384: USRMAP: 3: Usermapper service not started
1100566384: ADMIN: 4: Command succeeded: usrmapsvc init
# server_usermapper server_2 -Import -user pass_pri [files were uploaded from Primary Usermapper side]
server_2 : done
# server_usermapper server_2 -Import -group gru_pri
server_2 : done
# server_usermapper server_2 -enable primary=161.156.232.22 [Point to IP Address of Primary Usermapper Service]
server_2 : done
# server_usermapper server_2
server_2 : Usrmapper service: Enabled
Service Class: Secondary
Primary = 161.156.232.22 (c) → "c" means that Secondary has contacted the Primary service
7. Start CIFS and verify access using server_cifsstat
8. Export Secondary Usermapper database and verify format of passwd and group files
```

## **REBUILDING INTERNAL USERMAPPER & IMPORTING USERS/GROUPS:**

1. Backup information as needed, then remove any existing traces of Internal Usermapper from your system

**#server\_usermapper server\_2 -disable**

**#server\_usermapper server\_2 -remove -all**

2. Set debug logging for Usrmapper:

**#.server\_config server\_2 -v "logsys set severity USRMAP=LOG\_DBG3"**

3. Re-initialize Internal Usermapper using following command:

**# .server\_config server\_2 -v "usrmapserv init"**

4. First do wordcount and then import the users and groups into the existing internal usermapper db:

**#cat users lwc -l #cat groups lwc -l**

**#server\_usermapper server\_2 -Import -user users**

**#server\_usermapper server\_2 -Import -group groups**

5. Enable the usermapper service:

**#server\_usermapper server\_2 -enable**

6. Check Server Log to see if Imports succeeded and run \$server\_usermapper to see if 'Enabled'

7. Output Usermapper database to see if output files wordcount matches what was imported:

**#server\_usermapper server\_2 -Export -user users\_after**

**#server\_usermapper server\_2 -Export -group groups\_after**

**#wc -l users\_after groups\_after**

8. Run following command to check health of Internal Usermapper, and verify ranges in hex output.

**\$ .server\_config server\_4 -v "usrmapserv info"**

1121362991: USRMAP: 4: Usrmapper service: Enabled

1121362991: USRMAP: 4: Service Class: Primary

1121362991: USRMAP: 4: Database File system: Root

1121362991: USRMAP: 4: Mapping mode: universal

9. Turn off debug USRMAP=LOG\_ERR

10. Map new Users as test & re-export database to verify new mappings

## **SERVERT USERMAPPER COMMAND OPTIONS:**

\$server\_usermapper ALL |-enable primary= | -disable | -remove -all | -import | -export -user -group -force <pathname>

**-enable | -disable:** Starts or stops Usermapper service to allow for configuration changes, maintenance, removal, etc.

**primary=:** Used to create a Secondary Usermapper Service and specifies the Primary Server

**secondary=:** Used by Primary to learn about the newly created Secondary Usermapper Server

**config=:** Used only by Primary Usermapper when migrating from an External to Internal Usermapper configuration

**-remove -all:** Permanently removes Usrmapper directory from Primary or Secondary Usermapper server

**-Import -user | -group | -force:** Imports either Users/Groups to the Data Mover

**Note:** Use the -force option to overwrite the existing /etc/usrmapper database--By default, imported Users or Groups are appended to the existing database.

**-Export -user | -group:** Exports users & groups file from Primary or Secondary Usermapper databases

## \$server\_usrmapper server\_2

server\_2: Usrmapper service: Enabled

Service Class: Primary

**Note:** Used to verify the usrmapper service on a data mover

**\$server\_usrmapper server\_2 -disable** [Used to temporarily disable the Usermapper service]

## INTERNAL USRMAPPER PARAMS:

### NAS 5.6 USRMAPPER PARAMS:

**# server\_param server\_2 -facility usrmap -list**

server\_2 :

|               |          |            |            |            |
|---------------|----------|------------|------------|------------|
| param_name    | facility | default    | current    | configured |
| maxuid        | usrmap   | 2147483647 | 2147483647 |            |
| mingid        | usrmap   | 16         | 16         |            |
| autobroadcast | usrmap   | 1          | 1          |            |
| maxgid        | usrmap   | 2147483647 | 2147483647 |            |
| minuid        | usrmap   | 16         | 16         |            |

**Note1:** Please note that when changing min or max params with NAS 5.5 or 5.6, the Server must be rebooted in order for the change to go into effect. This behavior is different from pre-5.6 NAS versions.

**# server\_param server\_2 -facility usrmap -modify maxgid -value 65535**

server\_2 : done

Warning 17716815750: server\_2 : You must reboot server\_2 for maxgid changes to take effect.

**Note2:** Ranges are only effective when initially creating the database—they have no affect when modifying after a database has already been created.

## CONFIGURING INTERNAL USRMAPPER PORT:

**Note:** Internal Usermapper uses port 12345 by default, hex 3039. Use following procedure to change port.

1. Verify existing port:

**# .server\_config server\_2 -v "param usrmap" |grep -i port**

usrmap.port 0x01721988 0x00003039 0x00003039

**Note:** 3039 is hex value for port 12345

2. Disable Usermapper Service:

**# server\_usrmapper server\_2 -disable**

3. Change usermapper service port:

**# .server\_config server\_2 -v "param usrmap port=1012"**

4. Re-enable Usermapper:

**# server\_usrmapper server\_2 -enable**

5. Verify port changed & service is listening:

**# .server\_config server\_2 -v "param usrmap" |grep -i port**

usrmap.port 0x01676b78 0x00003039 0x00003039

**Note:** Port 3039 is hex for 12345

**# server\_netstat server\_2 -a |grep 12345**

tcp \*.12345 \*.\* LISTEN

udp \*.12345

## INTERNAL USRMAPPER CONFIGURATION OUTPUT:

**# .server\_config server\_2 -v "usrmapserv info"**

1096316892: USRMAP: 4: Usrmapper service: Enabled

1096316892: USRMAP: 4: Service Class: Primary

1096316892: USRMAP: 4: Mapping mode: universal

1096316892: USRMAP: 4: Usermapper addresses

1096316892: SMB: 4: Usermapper[0] = [127.0.0.1] state:available (internal)

1096316892: USRMAP: 4: Cached environments

1096316892: USRMAP: 4: 1 0x218d2504 45 ?: Enabled

1096316892: USRMAP: 4: 1 0x14a8258 Env ?: Enabled

1096316892: USRMAP: 4: Cached domains

**1096316892: USRMAP: 4: \(32768): 2** →This section represents unallocated or unused UID/GID ranges—i.e., ranges avail.

1096316892: USRMAP: 4: UID ffffffff :[0 8010-7ffffef] [1 10-7fff] →Section is where next UID range to be assigned will begin

1096316892: USRMAP: 4: GID ffffffff :[0 8030-7ffffef] [1 10-7fff] →Section is where next GID range to be assigned will begin

**1096316892: USRMAP: 4: @ (32768): \* 2** →This section represents any Filled ranges and current UID/GID ranges in use

1096316892: USRMAP: 4: UID **8004** :[0 **8000-800f**] →Next UID to be assigned; currently allocated UID/GID ranges in use  
1096316892: USRMAP: 4: GID **8029** :(0 **8001-801f**) [1 **8020-802f**] →Next GID to be assigned;GID range filled; GID range used  
1096316892: USRMAP: 4: operations(32768): 1  
1096316892: USRMAP: 4: ops(32768): 1  
1096316892: USRMAP: 4: ohiodom004(32768): 1  
1096316892: ADMIN: 4: Command succeeded: usrmapsvc info

**Note:** Above command represents internal information stored for usrmappc.db file. When in Universal Mode [i.e., not using an imported usrmap.cfg file from a previously used external Usermapper configuration], the domain names do not have a range defined. Usermapper takes the “\” available range to allocate ranges for UID/GIDs, and assigns values to the “@” universal domain as assigned ranges—both \ and @ ranges are complementary and when added together represent the Global mapping range in use.

## **INTERNAL USERMAPPER DEBUG LOGGING:**

**\$ .server\_config server\_2 -v "logsys set severity USRMAP=LOG\_DBG3"**

**\$ server\_usermapper server\_2 -disable**

server\_2 : done

**\$ server\_usermapper server\_2**

server\_2 : Usrmapper service: Initialized

Service Class: Primary

**Note:** Notice that the –disable command changes the service from “Enabled” to “Initialized”. Check Server Log for more details.

**\$ server\_log server\_2 -a |grep -i USRMAP**

1105582533: USRMAP: 5: Usrmapper threads stopped

1105582533: USRMAP: 6: Environment cache unloaded

1105582533: USRMAP: 4: 3: Usermapper service disabled

1105582533: ADMIN: 4: Command succeeded: usrmapsvc disable

## **VARIOUS COMMANDS TO DEBUG USERMAPPER MAPPINGS OF USERS, GROUPS, SIDS:**

**\$ .server\_config server\_2 -v "usermapper if=cge0 user=nas10"**

1105585082: SMB: 4: UserName0='MOUSE\nas10' (0) use=1

S-1-5-15-42f831d9-66417ccd-28a68b82-498 nas10

1105585082: VC: 5: abortCheckWait(smb\_share=0x0)

1105585082: SMB: 4: User mouse\nas10 **UID=32779** →Lookup by User Name returns UID value

**\$ .server\_config server\_2 -v "usermapper if=cge0 group='domain admins'"**

1109767775: SMB: 4: UserName0='MOUSE\domain admins' (0) use=2

S-1-5-15-42f831d9-66417ccd-28a68b82-200 domain admins

1109767775: VC: 5: abortCheckWait(smb\_share=0x0)

1109767775: SMB: 4: Group mouse\domain admins **GID=2000** →Lookup by Group Name returns GID value

**Note:** Use above command to validate that Usermapper can map a User or Group.

**\$ .server\_config server\_2 -v "usermapper if=cge0 group='domain admins'"**

1109767775: SMB: 4: UserName0='MOUSE\domain admins' (0) use=2

S-1-5-15-42f831d9-66417ccd-28a68b82-200 domain admins

1109767775: VC: 5: abortCheckWait(smb\_share=0x0)

1109767775: SMB: 4: Group mouse\domain admins **GID=2000** →Lookup by Group Name returns GID value

**\$ .server\_config server\_2 -v "usermapper if=cge0 gid=2002"**

1109768035: SMB: 4: usermapper uid/gid=0x7d2 name=**domain=20users** →Lookup by GID returns name associated with GID

S-1-5-15-42f831d9-66417ccd-28a68b82-201

1109768035: ADMIN: 4: Command succeeded: usermapper if=cge0 gid=2002

**\$ .server\_config server\_2 -v "usermapper if=cge0 uid=1010"**

1109768104: SMB: 4: usermapper uid/gid=0x3f2 **name=nas1** →Lookup by UID returns name associated with UID

S-1-5-15-42f831d9-66417ccd-28a68b82-45d

1109768104: ADMIN: 4: Command succeeded: usermapper if=cge0 uid=1010

**\$ .server\_config server\_2 -v "usermapper if=cge0 gsid=S-1-5-15-42f831d9-66417ccd-28a68b82-45a"**

1109768384: SMB: 4: S-1-5-15-42f831d9-66417ccd-28a68b82-45a=**2010** →Lookup using SID results in GID associated with SID

## **NAS 5.2 AND ROOTFS INODE LIMITATIONS:**

I. NAS Codes prior to 5.2.10.3 support inode densities of 1 inode per 8k block for the rootfs of the Data Mover

II. NAS Codes and new installs of 5.2.10.3 and higher support inode densities of 1 per 1k block

## **INODE USAGE REQUIREMENTS BY USRMAPPER AND SECMAP DATABASES:**

--Each User and Group mapped to the Celerra by Internal Usrmapper will require (4) rootfs inodes

--Each User and Group mapped to the Celerra by the Secmap Cache will require (2) rootfs inodes

### **INODE LIMITATION PROBLEM:**

Default size of the data mover's root file system is 128MB, yielding a total of about 15,464 inodes with a density of 1/8k block

Default size of new file system is still 128MB, yielding 130,871 inodes with a density of 1/1k block.

Systems that are upgraded from any level lower than 5.2.10.3 will not be able to support inode densities of 1/1k block

### **INODE LIMITATION SOLUTION:**

Short and Long-term solutions for sites running out of inodes for the rootfs are as follows:

- I. Cases where customers use very large numbers of shares [8000+], try moving customer to Homedir service to free rootfs space
- II. Extend root file system by 2GB or higher if required—see emc87826

### **CELERRA SECMAP CACHING:**

→DART 5.2 introduced a persistent UID/GID-to-SID cache that is stored in the /.etc/secmap folder of the data mover or /root\_vdm\_x/.etc/secmap for VDM containers

→Secmap Cache was introduced as an efficiency mechanism to cut down on the single-threaded authentication calls made to the DC that often translated into authentication backlogs and slow DM performance during heavy logon periods, etc.

→SecMap Cache consumes (2) Inodes on the Data Mover rootfs for each UID or GID entry

→SecMap is used with Active Directory, NT Credentials for Unix, and SecureNFS features, but not with NFSv4

→Starting with NAS 5.6, Secmap is converted from a nameDB inode reference structure to an internal Berkeley database format

### **NAS 5.6 Secmap Upgrade Issue:**

Upgrade may fail on Task 68 if the secmap service is disabled, and is now checked by the PreUpgrade script:

E000123: Secmap is disabled on Server (server\_2).

**Fix:** Enable secmap on the Data Mover and run the upgrade script again

### **SERVER CIFSSUPPORT SECMAP TOOL:** Introduced with Napa 4 release

**# server\_cifssupport server\_2 -secmap -list | -create | -verify | -update | -delete | -export | -import | -report**

**Note:** Beginning with NAS 5.5.24.2, the server\_cifssupport command allows for reporting and editing of secmap database entries.

### **SECMAP RULES/REQUIREMENTS:**

→Each SID mapping will contain only one UID or GID value

→Each entry is based on the NameDB database and represents a symbolic link file, where the target of the link is the content (applies to all NAS versions prior to NAS 5.6)

→Unix UID/GID Reverse Mappings could have multiple SIDs mapped for each UID or GID

→When Secmap is enabled, the Global Sid Cache is not used

→Secmap uses (3) tables: Domain SID Table, UID Table for Reverse Mappings, GID Table for Reverse Mappings

→Secmap Cache entries are the first source that DART consults when Users/Groups connect to the Celerra

→Data Mover rootfs and all VDM CIFS containers will contain a Secmap cache as CIFS Servers are defined

→SecMap caches each mapping for SIDs upon the first access by that SID to the Celerra, and notes where the mapping originated: i.e., local passwd/group file, NIS, Usermapper, LDAP, etc.

→Subsequent access by the same SID will be mapped & authenticated from the cache, meaning that DART does not have to query the DC's to validate the SID again

→Reverse Mappings are created when using local passwd/group files, NIS, or older versions of Usermapper that use names (DART must take name and query DC's for match for the name to obtain SID), though once cached, subsequent reconnections will map User/Group from SecMap cache based on SIDs-to-UID/GID [See CIFS param quotas.queryNames]

→Reverse Mappings are used by NTCredential for Unix User feature, and for W2K Quotas name lookups

### **CIFS PARAMETER quotas.queryNames:**

**# server\_param server\_2 -facility cifs -info quotas.queryNames -v**

By default, the cifs quotas.queryNames param is set to 0, meaning that User/Group name queries to obtain valid SIDs for W2K Quota enumeration, will use only Usermapper, and once cached in Secmap, then only Secmap to perform the lookup. By setting the param value to 1, DART would use NIS or Local password/group files to perform User/Group name lookups for W2K Quotas

### **TWO TYPES OF MAPPINGS USED IN SECMAP:**

→SID to UID or GID Mapping (most commonly used mapping whenever CIFS users connect to Celerra)

→Reverse UID/GID to SID Mapping (NTCREDENTIAL for UNIX & W2K Quotas)

**Note:** SecMap Caching requires (2) inodes on the DM rootfs for every UID or GID mapping made. One Secmap entry is for the main SID-to-UID/GID mapping and the other entry is for the Reverse Mapping UID/GID-to-SID, which is used when building an NT Credential for UNIX users or when retrieving Users and Groups to get UIDs/GIDs for W2K Quotas.

### **/.etc/secmap**

drwxr-xr-x 2 root bin 1024 May 14 13:52 @gid →GID-to-SID mappings, with GIDs in Hex—no names

drwxr-xr-x 2 root bin 1024 May 14 13:53 S-1-5-15-1f6d0078-63f6f323-65d637a8-ffffffffff →SIDs

drwxr-xr-x 2 root bin 80 May 14 13:53 @uid →UID-to-SID mappings, with UIDs in Hex—no names

**Note:** With secmap caching enabled, default 5.2 installs, directories are populated with hex UID/GID mappings for User/Group SIDs as Users access the system for the first time

# ls -la \*

@uid:

lrwxr-xr-x 1 root bin 39 May 14 13:52 **31** -> S-1-5-15-1f6d0078-63f6f323-65d637a8-489

**Note:** Secmap lists the UID/GID's in HEX [**31 hex = UID 49**], along with the respective Windows SID, but NOT the name.

#cd S-1-5-15-1f6d0078-63f6f323-65d637a8-ffffffffff

# ls -la S\*

lrwxr-xr-x 1 root bin 36 May 14 13:42 **200** -> 162?5a?5a?40a5050b?TS2\Domain Admins →RID & Group Names

lrwxr-xr-x 1 root bin 35 May 14 13:52 **201** -> 162?5b?5b?40a50759?TS2\Domain Users “ “

lrwxr-xr-x 1 root bin 29 May 14 13:52 **489** -> 161?31?a?40a50759?TS2\mapper9 →RID & User Names

**Note:** RID [Relative Identifier—last part of SID 32 bits] of User or Group listed in red

## DUMPING SECMAP CONTENTS TO FILE, LISTING SECMAP DATABASES, INODE USAGE:

# .server\_config server\_2 -v "secmap if=foo list=secmap\_dmp"

# .server\_config server\_2 -v "secmap info"

1153419719: SMB: 4: SecMap: Secure Mapping Database server\_2

1153419719: SMB: 4: SecMap:

1153419719: SMB: 4: SecMap: S-1-5-15-7a50bbc6-60415a8a-6b635f23-ffffffff

1153419719: SMB: 4: SecMap: S-1-5-15-bfc56af5-d6cf8701-5f67b1a3-ffffffff

1153419719: SMB: 4: SecMap: S-1-5-15-242a3a09-6bc5c62-3f32a78a-ffffffff

1153419719: SMB: 4: SecMap:

1153419719: SMB: 4: SecMap: 3 tables defined

# .server\_config server\_2 -v "secmap usage"

1153420014: SMB: 4: SecMap: Database usage

1153420014: SMB: 4: SecMap: Nodes: 101

1153420014: SMB: 4: SecMap: Size: 0

## SECMAP PRIMARY LOOKUP: SID-to-UID/GID:

# .server\_config server\_2 -v "secmap get sid=S-1-5-15-bfc56af5-d6cf8701-5f67b1a3-4d6"

1153420211: SMB: 4: SecMap: sid S-1-5-15-bfc56af5-d6cf8701-5f67b1a3-4d6 mapping

1153420211: SMB: 4: SecMap: Mapped from usermapper on Wed Jun 21 12:14:13 2006

1153420211: SMB: 4: SecMap: user=800b(8000)

1153420211: SMB: 4: SecMap: Name=2K3\backup

**Note:** Please be aware that this command only searches cache for a SID match and retrieves the info as is

## SECMAP REVERSE LOOKUP: UID/GID-to-SID:

# .server\_config server\_2 -v "secmap get uid=0x800b"

1153420363: SMB: 4: SecMap: S-1-5-15-bfc56af5-d6cf8701-5f67b1a3-4d6

## WINDOWS ACCOUNT RENAMED--ACCOUNT MAPPED IN USERMAPPER & SECMAP—

### FORCING SECMAP CACHE UPDATE:

**Note:** Names are no longer considered as important for Usermapper or Secmap as the SID. Therefore, if a Windows account was renamed, the SID renames the same, but the account name is NOT automatically updated in SecMap Cache or Usermapper db.

# .server\_config server\_2 -v "secmap get sid=S-1-5-15-bfc56af5-d6cf8701-5f67b1a3-4d6"

1153420211: SMB: 4: SecMap: sid S-1-5-15-bfc56af5-d6cf8701-5f67b1a3-4d6 mapping

1153420211: SMB: 4: SecMap: Mapped from usermapper on Wed Jun 21 12:14:13 2006

1153420211: SMB: 4: SecMap: user=800b(8000)

1153420211: SMB: 4: SecMap: Name=2K3\backup

**Note:** This command only searches the secmap cache database for matching SID

# .server\_config server\_2 -v "secmap check sid=S-1-5-15-bfc56af5-d6cf8701-5f67b1a3-4d6"

1153421283: SMB: 7: ProcessBindReply:cMsRPC\_Bind Pipe,lsarpc to fid=8002

1153421283: SMB: 7: ProcessBindReply:maxFragSizes: Cl=10b8/10b8 Set=10b8/10b8

1153421283: SMB: 7: Request\_enumSIDs:enum 44 4036 1009,1=1

1153421283: SMB: 7: Request\_enumSIDs:done 8988160,895d190 1,1 1,1

1153421283: SMB: 6: ExtractDomain: Domain(0/1)=2K3

1153421283: SMB: 7: ExtractNames:Usr='2K3\backup\_renamed' RID=4d6 U=1 UID=0 T=-1 ((NULL))

1153421283: SMB: 7: Unix user 'backup\_renamed' found with 'SECMAP:backup\_renamed.2k3' UID=32779

1153421283: SMB: 6: FindUserId:Access\_Password 'backup\_renamed',1=800b T=0 (OK)

**Note:** The “check sid” command does a SID-to-Name lookup to the DC and returns the renamed account with the original UID for 2k3\backup, but DOES NOT update the SecMap Cache

**# .server\_config server\_2 -v "secmap fix sid=S-1-5-bfc56af5-d6cf8701-5f67b1a3-4d6"**

1153422090: SMB: 7: ProcessBindReply:cMsRPC\_Bind Pipe,lsarpc to fid=8003

1153422090: SMB: 7: ProcessBindReply:maxFragSizes: Cl=10b8/10b8 Set=10b8/10b8

1153422090: SMB: 7: Request\_enumSIDs:enum 44 4036 1009,1=1

1153422090: SMB: 7: Request\_enumSIDs:done 895b548,8964d28 1,1 1,1

1153422090: SMB: 6: ExtractDomain: Domain(0/1)=2K3

1153422090: SMB: 7: ExtractNames:Usr='2K3\backup\_renamed' RID=4d6 U=1 UID=0 T=-1 ((NULL))

1153422090: SMB: 7: Unix user 'backup\_renamed' found with 'SECMAP:backup\_renamed.2k3' UID=32779

1153422090: SMB: 6: FindUserId:Access\_Password 'backup\_renamed',1=800b T=0 (OK)

**Note:** This command is supposed to validate SID-to-Name lookups with DC and then update SecMap cache with any changed info—did not work with 5.5.22.0.

## **DART MAPPING RESOLUTION ORDER FOR USER/GROUP AUTHENTICATION: MAPPING RESOLUTION ORDER NAS 5.2 & ABOVE:**

1. SecMap Persistent Cache [/etc/secmap] (DART resolves by Group/User SIDs and maps to GIDs/UIDs, respectively)
2. Global Data Mover Sid Cache (DART resolves Group SIDs and maps to GIDs)
3. Local Password & Group files (DART resolves by User/Group names and maps to UIDs/GIDs, respectively)
4. NIS Client Service (DART resolves by User/Group names and maps to UIDs/GIDs, respectively)
5. Active Directory Mapping Utility (DART resolves by User/Group SIDs and maps to UIDs/GIDs, respectively)
6. UserMapper Service (External or Internal) (DART resolves by User/Group SIDs and maps to UIDs/GIDs, respectively for Internal Usermapper, and by Name-to-UID/GID prior to NAS 5.2)

**Note:** Do not use uppercase characters in passwd files or NIS database—use lowercase ASCII for Celerra

## **DISABLING & DELETING SECMAP CACHING (Pre-NAS 5.6):**

1. Set param to disable SecMap caching:

**param cifs secmap.enable=0**

**Note:** Secmap is set to ‘enable=1’ by default.

2. Stop CIFS Service
3. Reboot Data Mover to removed cache from memory
4. Delete SecMap directory, or move to Control Station for archiving:

**# rm -Rf secmap**

5. Start CIFS Service

## **REMOVING ENTIRE SECMAP CACHE WITHOUT STOPPING CIFS:**

**\$ server\_cifssupport server\_x –secmap –list |grep “S-1” | sed ‘s/S-1@/ /’ | awk –F’@’{printf “S-1%\$1\n”, \$2}’ |xargs –i server\_cifssupport server\_x –secmap –delete –sid {}**

**Note:** New method available with 5.5.24.2

## **DELETING SECMAP DB WITH NAS 5.6:**

**Note:** You can edit and delete individual entries using server\_cifssupport, but in order to delete the entire secmap database, use the following steps

1. Stop the CIFS service on the Data Mover in order to “close” the dbms secmap database
2. Delete the secmap dbms database

**# server\_dbms server\_2 –db –delete Secmap**

3. Restart the CIFS service

## **ENABLING SECMAP (Pre NAS 5.6):**

1. Set the param manually in the appropriate /nas/server/slot\_x/param file, or use the server\_param facility:

param cifs secmap.enable=1

**# .server\_config server\_2 -v "param cifs secmap.enable=1"**

**# server\_param server\_2 –facility cifs –modify secmap.enable –value 1**

2. Stop & restart CIFS service to enable the Secmap module

### **Server Log:**

2005-10-27 15:01:42: SMB: 4: SecMap: server\_5 secure mapping DB started

3. After User access, check /.etc directory--Secmap folder should now exist

```
# ls
secmap
```

**DELETING SPECIFIC MAPPINGS FROM SECMAP CACHE (Pre server\_cifssupport 5.5.24):**

```
# .server_config server_2 -v "secmap if=laip1-2a del sid=S-1-5-15-bfc56af5-d6cf8701-5f67b1a3-200"
```

1140705748: SMB: 4: SecMap: Mapping removed

1140705748: ADMIN: 4: Command succeeded: secmap if=laip1-2a del sid=S-1-5-15-bfc56af5-d6cf8701-5f67b1a3-200

**Note:** Need SecMap enabled and CIFS running in order to run this command. Engineering states that SecMap Cache is checked for GID SID mappings prior to consulting Global Sid Cache. The latter cache does not “expire”, but fills up to capacity, and then uses an LRU algorithm to add new entries while deleting “old” entries. May need to run the following procedure to clean up everything.

**REMOVING SINGLE SECMAP ENTRIES FROM CACHE (Pre server\_cifssupport 5.5.24):**

- Dump secmap cache & check for GID mapping for the Group SID in question

```
# .server_config server_3 -v "secmap if=cge0 list=sec_dump"
#server_file server_3 -get ./etc/secmap/sec_dump sec_dump
```

- Delete SID for GID in Question: →[Domain Guests example]

```
# .server_config server_3 -v "secmap if=cge0 del sid=S-1-5-15-42f831d9-66417ccd-28a68b82-202"
```

1113946407: SMB: 4: SecMap: Mapping removed

- Dump sec\_map again

```
# .server_config server_3 -v "secmap if=cge0 list=mod_sec"
```

- Check to see if mapping has been removed from SecMap cache

- If it has, then stop CIFS to flush Global Sid Cache

- Restart CIFS

**DUMPING & READING SECMAP DATABASE:**

- # .server\_config server\_2 -v "secmap if=172\_25\_9\_200 list=dump.1st" →dumps secmap db to .etc directory

- # server\_file server\_2 -get ./etc/secmap/dump.1st dump.1st

```
# more dump.1st
```

# Secure Mapping Database for server\_2

#

# S-1-5-15-5155eb8a-c0ded414-a9b42a3c-ffffffffff

#

sid S-1-5-15-5155eb8a-c0ded414-a9b42a3c-200 mapping

**Mapped from usermapper** on Wed Jun 30 00:59:19 2004

group=8002 →**Group mapped from Usermapper, hex 8002 = 32770 decimal**

Name=SC\Domain Admins

sid S-1-5-15-2f1e7626-3115e3fc-28a68b82-252d mapping

**Mapped from usermapper** on Mon Aug 23 12:08:23 2004

user=b870(1f40) →**User mapped from Usermapper, hex b870 = 47,216 decimal, group = 1f40 (8000 dec.)**

Name=ALLISON\rd\_tzn2y3

sid S-1-5-15-3e525c04-3373819-101d7ffe-5b3 mapping

**Mapped from etc** on Sun Aug 22 05:42:42 2004

user=13ac(1392) →**User mapped from local Passwd file (uid=13ac; gid=1392 hex)**

Name=GMNMDOMATD02\xz33xp

sid S-1-5-15-8bb001b-20e65643-252c19b2-7bc mapping

**Mapped from nis** on Tue Sep 14 21:21:22 2004

user=8968(1388) →**User mapped from NIS database**

Name=EDSRDATD01\tnz2y3

sid S-1-5-15-bfc56af5-d6cf8701-5f67b1a3-4c0 mapping

**Mapped from ldap** on Tue May 16 09:16:37 2006 →**User mapped from iPlanet Directory Server database**

user=b275(fde8)

Name=2K3\W2kU45685

**lrwxr-xr-x 1 root bin 36 May 16 09:16 4c0 -> 171?b275?fde8?4469d0b5?2K3\W2kU45685**

**Note 1:** The “171?b275?fde8?4469d0b5?” shows that the mapping was derived from Ldap iPlanet.

**Note 2:** Reviewing file shows origins of each SecMap mapping (i.e., NIS, Passwd/Group file, Ldap, Usermapper, ADMapping, etc.)

**DUMPING AND RETRIEVING SECMAP DATABASE FROM CIFS VDM:**

```
# .server_config CLKPTDM04 -v "secmap list=sec_dmp"
```

```
# server_file CLKPTDM04 -get ./etc/secmap/sec_dmp sec_dmp
```

CLKPTDM04 : done

**Note:** In this example, NAS 5.4.17.5, the VDM names are all UpperCase, perhaps a result of creation by WebUI

## **DETERMINING ORIGIN OF MAPPING METHOD TO CREATE A SECMAP ENTRY:**

1. Go to Secmap and SID directory for a given Domain SID:

```
#cd /nasmcd/quota/slot_2/.etc/secmap/S-1-5-15-5155eb8a-c0ded414-a9b42a3c-ffffffffff
```

2. Run ls -la:

```
[root@wdseccmcs01 S-1-5-15-5155eb8a-c0ded414-a9b42a3c-ffffffffff]# ls -la lhead
```

```
lrwxr-xr-x 1 root bin 34 Oct 25 08:34 1000 -> 161868d8000417d1cfSC\lozano_v
```

```
lrwxr-xr-x 1 root bin 36 May 16 09:16 4c0 -> 171?b275?fde8?4469d0b5?2K3\W2kU45685
```

**Note:** The “161868d8000417d1” output tells us that this was a mapping derived from a Usermapper call. The “171?b275?fde8?4469d0b5?” shows that the mapping was derived from Ldap iPlanet.

**0 - unknown**

**1 - secmap**

**2 - localgroup**

**3 - etc local files**

**4 - nis**

**5 - active directory**

**6 - usrmapper**

**7 - ldap**

## **USING SECMAP DATABASE TO RECREATE THE INTERNAL USERMAPPER DATABASE:**

**Note:** Use the following procedure to export the User and Group mappings from Secmap, create User and Group import files, then import to Internal Usermapper to repopulate a lost db.

1. Export the Secmap Mappings

```
# server_cifssupport server_x -secmap -export -file secmap_now
```

**Note:** The second colon-delimited field identifies Groups and Users

```
S-1-5-15-32ea1e67-607d8090-79df51a3-a36:2:96:8007:8007:BRCSLAB\LG Teste >> This is a domain group
```

```
S-1-5-15-32ea1e67-607d8090-79df51a3-458:1:96:8000:8000:BRCSLAB\gustavo >> This is a domain user
```

2. Use the following script to parse the output of the exported secmap database to the format the internal usermapper understands. Using vi, create a script file named as usrmap\_imp.sh with the following contents, and use chmod 755 to make the script executable:

```
#!/bin/sh
awk -F: '{
split($6,fld,"\\\"");
if($2!=2){
printf("printf \"%s:*:%%d:%%d:user %s from domain %s:/usr/%s:/bin/sh\\n\" 0x%s 0
x%s\\n",$1,fld[2],fld[1],$4,$5);
} else{
printf("printf \"%s:*:%%d:%s.%s:\\n\" 0x%s \\n",$1,fld[2],fld[1],$4);
}
}'|bin/sh
```

3. Next, use grep to parse the secmap\_now output to create a User file, then a Group file:

```
cat secmap_now |grep :1:./usrmap_imp.sh > usrmap_user.tmp ( This generates the user file )
```

```
cat secmap_now |grep :2:./usrmap_imp.sh > usrmap_group.tmp ( This generates the group file )
```

**Note:** The format of the internal usermapper database files that will be imported should be like this :

```
$ cat user_now
```

```
S-1-5-15-32ea1e67-607d8090-79df51a3-1f4:32769:32768:user Administrator from domain BRCSLAB:/usr/S-1-5-15-32ea1e67-607d8090-79df51a3-1f4:/bin/sh
```

```
$ cat group_now
```

```
S-1-5-15-32ea1e67-607d8090-79df51a3-200:32770:Domain Admins.BRCSLAB:
```

```
S-1-5-15-32ea1e67-607d8090-79df51a3-201:32771:Domain Users.BRCSLAB:
```

```
S-1-5-15-32ea1e67-607d8090-79df51a3-202:32776:Domain Guests.BRCSLAB:
```

4. Use the following steps to import the user and group files into usermapper :

a) Disable internal usermapper :

```
# server_usrmapper server_x -d
```

b) Clean the database :

```
# server_usrmapper server_x -remove -all
```

c) Enable internal usermapper :

```
# server_usrmapper server_x -e
```

d) Import the user and group files :

```
# server_usermapper server_2 -I-user usrmap_user.tmp  
# server_usermapper server_2 -I-group usrmap_group.tmp
```

e) Check the server\_log for any errors. A clean import should look like this :

```
# server_log server_x -a -s |grep USRMAP
```

```
2008-07-31 20:04:10: USRMAP: 4: Starting usermapper service  
2008-07-31 20:04:10: USRMAP: 4: 2: Usermapper service enabled  
2008-07-31 21:58:15: USRMAP: 4: Starting usermapper service  
2008-07-31 21:58:15: USRMAP: 4: 2: Usermapper service enabled  
2008-09-23 11:17:58: USRMAP: 4: 3: Usermapper service disabled  
2008-09-23 11:18:08: USRMAP: 4: 4: Usrmapper database destroyed  
2008-09-23 11:18:18: USRMAP: 4: 1: Usermapper database created  
2008-09-23 11:18:18: USRMAP: 4: 2: Usermapper service enabled  
2008-09-23 11:18:27: USRMAP: 4: Import completed, 60 new users imported  
2008-09-23 11:18:47: USRMAP: 4: Import completed, 883 new groups imported  
2008-09-23 11:50:37: USRMAP: 4: Export last 361 ms
```

## **COMPARING NFS & CIFS PROTOCOLS:**

### **NFS:**

Stateless  
Single filename space  
Host-based “Authentication”; UID’s/GID’s; /etc/Hosts  
Advisory Locking scheme  
DNS & NIS nameservices and User database  
NIS Authentication DB  
RWX file attributes and umask values—no ACL’s

### **CIFS:**

Stateful (i.e., client must recover after server crash)  
Multiple namespace; 8.3/256 [max.path length]/posix\*  
User-Based “Authentication”  
Mandatory locking scheme  
DNS & Wins nameservices  
NT Authentication DB  
Extended file attributes {rwxpdo, file creation date & ACL’s}  
Browsing; Hidden shares; encrypted passwords; HomeDirs

\*Posix is a Unix-based standard called Portable Operating System Interface for computer systems.

**Note:** TCP can be enabled for NFS in /nas/server/slot\_x \*/netd by appending “tcp=1” to “nfs start” line

## **CIFS INTEROPERABILITY ISSUES:**

**Symbolic Links**—an alias, or shortcut for NT; a copy or pointer for Unix

**File Naming**—NT is case preserving, NOT case sensitive

**DOS Attributes**—DOS has file creation date, NOT unix

**Access Checking Policies**—NT or Native affords any NT client access; Secure=Highest security

**File Locking Mechanisms**—Three Options: File Locking occurs when Multiple File System protocols are involved and for CIFS is determined by Applications. Notepad, Wordpad, and Vi Editor do not practice locking, while Word & Excel do.

- 1.) rwlock: A CIFS file lock denies any further access. An NFS file lock will allow access, no 2<sup>nd</sup> lock.
- 2.) wlock: NFS can open an already opened file in Read Only mode.
- 3.) nolock: Either NFS or CIFS can open and write to a ‘locked’ file

## **CIFS ACL's: TROUBLESHOOTING & REPAIR OPTIONS**

**Symptoms:** Loss of access to NT Shares. Specific “Access Denied” message from NT Client when trying to access Celerra Shares through Network Neighborhood or Server Manager, either as Administrator or Ordinary User.

### **Tools to Use:**

View Server\_log--Look for SHADOW file rebuild or other irregularities

Use ACL Dump tool

### **I. Run ACL Check Using .server config Command:**

**Note:** Do not use "aclchk" utility on nas versions under 2.1.32.2

**Revised Procedure:** There is a known issue with running ACLCHK on filesystems when “Exports” are temporarily unexported. Rather than dealing with Exports at all, permanently Unmount the PFS [after noting mount options], create a new mountpoint, and remount the PFS on the new mountpoint to conduct ACLCHKs without interference from Exports.

Step 1. Permanently unmount PFS from original mountpoint. Create new mountpoint for aclchk purposes and permanently remount filesystem to this mountpoint. Conduct aclchk.

**Note:** Filesystem being aclchk’ed is NOT exported

Step 2. Conduct ACL Check on Server\_6:

```
$server_config server_6 "file aclchk /export/htdocs fix"
```

Step 3. Open Server Log to Monitor:

```
$server_log server_6 -s -f [Will show ACL Check completion]
```

Step 4. Permanently unmount File System and remount to original Mountpoint.

Step 5. Re-export all exports:

**\$server\_export server\_6 -a**

## **II. Running the “CIFS UPDATE” on Mountpoint path of affected Share or File System:**

Step 1. Leave the Share Exported for CIFS

Step 2. \$server\_setup server\_2 -P cifs -o stop [Recommended method for Update]

Step 3. Verify CPU Usage: \$server\_sysstat server\_2 [i.e., if heavily utilized, will take more time & possibly panic DM]

Step 4. **\$server\_config server\_2 “cifs update /emcfset12 force level=0”**

Step 5. Monitor CIFS Update in Server Log: \$server\_log server\_2 | tail

**Note:** 'Command Succeeded' will be echoed at commandline--results of Update will be logged:

2001-11-06 23:22:24: SMB: 4: Update of : /emcfset12 completed after 437 s

2001-11-06 23:22:24: SMB: 4: Update of : 219475 files, 21229 dirs

**Note 2:** If Update fails with following error, Stop & Restart CIFS to knock Users off file system:

1005105482: SMB: 3: Error: Update cannot be done on share in use

## **RUNNING CIFS UPDATE ON A FILE SYSTEM:**

**\$server\_cifs server\_6 -update /emcfset12 mindirsize=0 force**

**Note:** Need to run this update if seeing errors in Server Log regarding issues when name info file returns “not found” messages. This command completely rebuilds the Shadow file! A potential drawback of this method is that short filenames may be renamed, depending on the order of the file in the directory as opposed to when they were created:

### **EXAMPLE OF SHORTNAMES:**

**# .server\_config server\_5 -v "shadow readdir \fs28\shortnames"**

1068044303: CFS: 4: name: shortname7.txt 69331000.TXT

1068044303: CFS: 4: name: shortname1.txt SHORTN~1.TXT

1068044303: CFS: 4: name: shortname3.txt SHORTN~3.TXT

1068044303: CFS: 4: name: shortname9.txt 89331000.TXT

### **ALTERNATE SYNTAX:**

**# server\_cifs server\_5 -update /fs40 mindirsize=0 force verbose**

2003-11-05 09:23:34: ADMIN: 4: Command succeeded: cifs update /fs40 mindirsize=0

2003-11-05 09:23:34: SMB: 4: Update of : /fs40 completed after 0 s

2003-11-05 09:23:34: SMB: 4: Update of : 2 files, 3 dirs

### **OTHER SHADOW FIX/CIFS UPDATE METHODS:**

**\$server\_config server\_9 -v “shadow update \ora\ImageArchives”** [Yet another version of update]

**\$server\_config server\_9 -v “shadow fix \volta\f3u2\victor”**

### **The following commands can be useful in troubleshooting and resolving Shadow File problems:**

Step 1. Run ‘Shadow Readdir’ against suspect folder in the file system path—notice the use of backward slashes to define the path:

**\$server\_config server\_x -v “shadow readdir \fs01\users\martin”**

**Note:** If you do not see a listing of all files that you would expect to see, it’s possible that the Shadow File is damaged and needs to be manually repaired. Or, if known subdirectories are seen as 0 Byte files, this also could be an indicator of a damaged Shadow File.

Under certain conditions, such as a full filesystem, the algorithms that dictate how many times the hashing index can be changed before requiring a rebuild operation, can be broken, requiring manual intervention to repair.

Step 2. Use ‘Shadow Fix’ to repair suspect folder, then run the ‘Shadow Readdir’ again to see if the directory contents are ‘visible’.

**\$server\_config server\_x -v “shadow fix \fs01\users\martin”**

**Note:** In some cases, the command will not run because of spaces or other peculiarities in the path—when all else fails, try to determine DOS 8.3 name for the next directory in the path and substitute that name:

**# .server\_config server\_7 -v “shadow readdir \data\userdata\SFISHE3\Sandy's files\Action Plans”**

1131570657: CFS: 3: shadow fix: getAlternateName failed: NotFound

**# .server\_config server\_7 -v “shadow readdir \data\userdata\SFISHE3\SANDY'~1\ACTION~1”**

### **Shadow File Symptoms:**

Loss of Access to Shares by Users

Server\_Log Indicator: SHADOW; forceRebuild; shadow assert failed on line 1740 file ..\shadow\_internal.cxx: 7

### **Server Log Example:**

2001-10-31 10:21:11: SHADOW: 3: ESHADOW: 3: Error from write block: 25 [Error 25 often means File System full]

2001-10-31 10:21:11: CFS: 3: @@@ commit failed: status = 25 [File\_NoSpace]

2001-10-31 10:21:11: SHADOW: 3: Error 25 from writeBlock

Intermittent access to CIFS Shares from different NT Users

CIFS Users see Directories as Files

**Causes:** NT users may lose the ability to access or see files or subdirectories within a file system, yet NFS can still see those files.

**Reasons:** Shadow File has become corrupted in FS and requires 'relinking' the Unix & Cifs filenames. File Systems are at 100% and SHADOW File cannot rebuild itself--File System should be reduced in size, then run the CIFS Update--ENGINEERING recommends that File Systems be kept at 90% or lower. Whenever these conditions occur, the "Shadow File" for a specific NT Share or Upper Level Mountpoint may need to be updated using the following commands:

### **Traditional CIFS Update Fix to Shadow File:**

**\$server\_config server\_2 "cifs update /emc force level=0"**

**\$server\_config server\_2 "cifs update /emc"**

**Purpose of Server CIFS Update Command:** Rebuilds the “shadow” file by relinking all Files and Directories to both their DOS 8.3 and NT Long Names to their Unix counterpart

### **WINDOWS CLIENT FILE SYSTEM FULL “POP-UP” MESSAGE:** 2.2.49.2 +

/nas/server/slot\_x/param

**param cifs sendMessage=0** [Turns off PopUp messages]

**param cifs sendMessage=1** [Default value. Send message on Error only]

**param cifs sendMessage=2** [Send message on Warning only]

**param cifs sendMessage=3** [Send message for both Warnings & Errors]

**param cifs minFreeFS=0** [Default Value is 10, which means that when FS is 90% Full, Pop-Up message is generated. By setting this value to 0, Clients will not be sent PopUp message based on file system capacity]

**Comment:** With both parameters set, a popup message will be sent to any Client connecting to the Share when the minimum Free Space setting is reached. MinFreeFS setting can be modified to any value. “Pop-Up” may not show on Win95 Users unless winpop.exe is turned on

### **CELLERRA CIFS SHADOW FILE:**

The Shadow File is an index containing Unix-to-NT name mappings, consisting of linkages between Unix names and Windows Long Names [M256] and DOS Short Names [8.3]. A separate Shadow File Index is located in every directory on a CIFS file system and is created when CIFS Users access a directory for the first time. Subsequently, the Shadow File is kept up-to-date dynamically as changes to files & subdirectories occur. The Shadow File itself is not visible from ordinary NT or Unix client access.

**Note:** Important to know that a CIFS Shadow File rebuild, whether done by the file system automatically or from a manually initiated CIFS Update, uses the UNIX file name to rebuild and then rename the Windows Long and Short names.

### **SHADOW FILE SYNCRONIZATION:**

One of the limitations of the current file system architecture is that the disk space required to accomodate the shadow file is dependent on the changing structure of the file system itself and cannot be pre-allocated through the use of reserved space or placed in some other location. The effective result of this design means that once a file system reaches the 100% capacity mark, the dynamic mechanism used to keep the shadow file up-to-date and intact can deteriorate. The index is considered out of sync when a file exists in a directory but the corresponding index entry is not yet updated to the Shadow File. This condition triggers a “Shadow Rebuild” operation and under normal circumstances, is quickly rebuilt.

### **THINGS THAT AFFECT SHADOW FILE HEALTH:**

Though events like an unexpected shutdown of a datamover can trigger an out of sync condition, the most common scenario is that of a file system reaching 100% capacity. When file systems reach 100%, it is possible for the shadow file update mechanism to begin to breakdown, especially within directories that support large numbers of files that are rapidly changing. Each time a CIFS host attempts to access files that are determined to be invalid because of changes to the files in the existing directory, a shadow file rebuild operation is initiated for the directory. Under ordinary circumstances, and with adequate free space available, shadow rebuild operations conclude quickly and without visible impact to CIFS clients. However, when file system capacities have been reached and there is limited free-space available, the Shadow rebuild process may force a directory lock while the rebuild is in progress. In extreme cases the rebuild operation can lock-out both NFS and CIFS clients. If the rebuild process is successful, the lock is released and normal access is returned to all Clients. However, if the rebuild is not successful, at some point the lock will be released, allowing NFS clients to access the directory, but denying CIFS users 'visibility' to the impacted directory. Further access attempts will repeat the lock and unlock cycles, and it may be possible that the shadow file cannot regenerate itself under these conditions. The net effect of a failure to rebuild the shadow file is the loss of cifs access from the perspective of Windows Clients. In fact, the overall directory or files within a directory may appear to be "missing" completely, when in reality the files still exist, but are masked from CIFS visibility due to the broken linkages between the underlying Unix file system names and the Windows names--without the linkages intact, files 'disappear' from CIFS view.

### **FUTURE DESIGN CHANGES ELIMINATE THE SHADOW FILE:**

Because of the limitations inherent with the Shadow File design, the Multiprotocol Directory [MPD] structure was been introduced with NAS 5.2—GA on February 9, 2004. The underlying file system improvement eliminates the Shadow File in its entirety, meaning that shadow file information is no longer stored in an auxiliary database, but in the file system inode structure itself, eliminating both the performance bottlenecks and file system full limitations that the previous architecture experienced.

**MAINTAINING A HEALTHY FILE SYSTEM & SHADOW FILE:**

The most effective method for minimizing shadow file issues are to maintain file systems with adequate space to consistently handle the ebb and flow of system and user operations. Ideally, file systems should be monitored so as to maintain 90% or less capacity. As a preventive monitoring measure, current NAS versions have the ability to configure SNMP Traps and Mail Notification alerts in the event that file systems reach a pre-determined plateau.

**KNOWN SHADOW FILE LIMITATIONS:**

--The Shadow File algorithm for calculating rebuilds may fail if the Shadow File tree/leaf structure is greater than 15 levels in depth, causing Shadow Asserts and incorrect computation of hash values. This condition can happen if there are hundreds of thousands of file entries within a file system directory.

--In NAS 4.2 and 5.1.9.4, a problem can occur in the Shadow File Index when files or directories are renamed with the same name [e.g., a Windows client renames a file from all lowercase to uppercase—this really isn't a name change]. Our Unix name mangling function does not check to make sure that the name did not really change and adds a ~ tilde character to the Unix name. Subsequently, a system Shadow rebuild or manual CIFS Update uses the Unix name to rebuild the Shadow file and will rename the NT Long and Short names with the ~ tilde character at the end of the file. See AR30194 and Primus EMC81938.

**SHADOW FILE STRUCTURE: 3 NAME LINKAGES FOR ALL FILES****Shadow file stores following information for each file in a directory:**

**1. DOS 8.3 short names & long M256 names**

**2. UNIX Names**

**3. CIFS Attributes: (5) windows attributes for a file (Archive, System, Hidden, Read-Only, Compression)**

**Symptoms:** Server\_Log: SHADOW; forceRebuild; shadow assert failed on line 1740 file ..//shadow\_internal.cxx: 7  
Loss of Access to Shares by Users [Error 7 means FileNotFound]

**Server Log Example:**

2001-10-31 10:21:11: SHADOW: 3: ESHADOW: 3: Error from write block: 25 [Error 25 means File System full]  
2001-10-31 10:21:11: CFS: 3: @@@@ commit failed: status = 25  
2001-10-31 10:21:11: SHADOW: 3: Error 25 from writeBlock  
2001-11-01 18:08:38: SMB: 4: Update of : /cifs2/share completed after 1940 s  
2001-11-01 18:08:38: SMB: 4: Update of : 1083559 files, 90786 dirs

**Shadow Rebuild Limitation:**

Prior to NAS 2.2.58.1, a file in the Shadow file could only be changed 256 times. Once this limit is reached, file becomes invisible to Client and CIFS Update must be run. Following Log message is indicative of this:

SHADOW: 4: Collision 0xa5,invhash;0x95f1

**HOW TO VERIFY UNIX FILE NAMES & WINDOWS 8.3 NAMES:**

**Note:** The following command will allow you to inspect NT Long Names and DOS 8.3 Short Names—may be useful when determining if a 0 byte file should actually be a folder—confirms whether there has been shadow file corruption

**\$server\_config server\_9 -v "shadow readdir \volta\f3u2\Victor"**

| <b>LONG 256 NAME</b>                           | <b>SHORT 8.3 NAME</b> |             |
|------------------------------------------------|-----------------------|-------------|
| 1041202451: CFS: 4: name: ADMIN                | ADMIN <DIR>           | [Directory] |
| 1041202451: CFS: 4: name: backup_error_ojv.txt | BACKUP~1.TXT          | [File]      |
| 1041202451: CFS: 4: name: USERS                | USERS <DIR>           |             |
| 1041202451: CFS: 4: name: filesystem_ojv.txt   | FILESY~1.TXT          |             |
| 1041202451: CFS: 4: name: profile_ojv.txt      | PROFIL~1.TXT          |             |

**HOW TO FIX SHADOW FILE FOR SPECIFIC FOLDERS:** [Use instead of CIFS UPDATE]

**\$server\_config server\_9 -v "shadow fix \volta\f3u2\Victor"**

**Note:** Shadow collision messages in the Server Log are a good indicator of a hash table problem occurring in the shadow file. If the hashing index reached 0x100 (decimal 256), the algorithm would stop and could no longer rebuild itself, therefore requiring the manual “shadow fix”. You can use this command with CIFS running & shares exported. This command zeroes out the Shadow directory and is rebuilt during next CIFS Client access.

**III. OBTAINING ACL DUMP ON PROBLEM SHARE:**

Step 1. **\$server\_config server\_x -v "acl if=ana0 dump=/export/hs1\_it"** [Full path to Share and Interface Used]

**Note:** If command does not execute, try only the 'sharename' itself in command; "acl if=ana0 dump=/export"

**ACL's are stored on the root of every CIFS file system**

**IV. FIXING ACL's USING SETACL:**

**Purpose:** One of the primary reasons for running the “setacl” command would be if users were experiencing “access denied” problems to CIFS Shares that they previously had access to. Probably the most common example of this problem would be if the Usrmapper databases were changed/deleted, and Users and Groups received totally new UID/GID mappings. Typically, this would be seen by a User as a loss of access to their “Home Directory”, or to a group of users, as a loss of access to a “Share” or “Directory” who’s permissions were based on a Global NT Group for access.

**Note:** When running “setacl”, ensure that CIFS is running—but have an administrator disconnect users from Server Mgr or Computer Management.

### **How does the “SETACL” command work?**

When the “setacl” command is issued against a filesystem mountpoint or CIFS directory, the ACLs are read by the DART to obtain SID information for all Users and Groups included in the ACL List. The DART then queries the DC or AD Server for each User and Group to verify SID information. Once this information is obtained, the DART obtains the latest mapping information from Usrmapper for each User and Group, and then updates the ACL list for that directory, matching User/Groupname/SID to UID/GID.

## **METHODS FOR UPDATING ACL LIST ON CELERRA FILE SYSTEMS:**

### **1<sup>st</sup> METHOD--COMMANDLINE SETACL COMMAND FROM DATAMOVER:**

**\$server\_config server\_2 “cifs update /mnt4 force setacl if=ana0”**

**Note:** This command realigns the UID/GID of a User or Group with the appropriate SID & updates the ACL records on the FS

**Caution:** *Do not run the “setacl” based on this section—review the full SETACL procedure in previous sections.*

### **2<sup>nd</sup> METHOD--WINDOWS NT/2000 GUI:**

As Administrator, view the properties of a particular folder and then “re-apply” the “permissions”—this also does the same thing as issuing the setacl command, though on a less global basis.

**Comment:** Use the ACL Dump command to verify ACLs on a folder BEFORE & AFTER running this command!!!

**\$server\_config server\_2 “acl if=ana0 dump=/mnt4”**

**Caution:** From experience, you may need to reboot the DataMover after each FileSystem after running the “setacl” command. If the command fails to execute [Error 4020], reboot the Server and try again.

**PREFERRED SYNTAX:** **\$server\_config server\_x “cifs update /mnt4 force setacl if=fsn0”**

**ALTERNATE SYNTAX:** **\$server\_config server\_x “cifs update /mnt4 setacl if=fsn0”**

**2002-10-01 18:20:53: SMB: 4: Update of : /shared completed after 630 s**

**2002-10-01 18:20:53: SMB: 4: Update of : 463771 files, 30447 dirs, 494218 Acl**

### **STOPPING CIFS UPDATES:**

**Note:** Use the following command to stop any of the commands that use CIFS Update

**#server\_config server\_x “cifs update abort”**

### **OUTPUTTING ACL DUMP INFORMATION:** Directory or File Permissions on Celerra File Systems

**\$server\_config server\_6 -v “acl if=fsn0 /personal/its/DParker”**

### **USING INLINE SCRIPTS TO RECURSIVELY DUMP ACLS:**

**\$for i in `find ftest /testdata/RMS`;do .server\_config server\_4 -v “acl if=fsn0 dump=\$i” >> /tmp/fptest.acldump; done**

## **V. RESETTING ACL’s ON CIFS SHARE OR SUBFOLDER TO ‘EVERYONE’ FULL CONTROL:**

**Note:** Apply to the “toplevel” CIFS Share only as it will reset the acl’s to “Everyone FULL CONTROL”. Otherwise, you can use this command discretely by running against a specific subfolder. The permissions would then have to be manually reassigned and built by the system admin. If you have the opportunity, conduct an “acl dump” prior to resetting ACL’s so that an analysis can be done on the share in question.

--If Administrator has Mask=1f01ff & Rights=RWXPDO, they have Full Control, no problem

--If Everyone has Mask=0x100000 & Rights=RWX, they have “No Access”

**\$server\_config server\_x “cifs update /mntpoint resetacl”**

**Note:** Can also run against subfolder

**\$server\_config server\_x “cifs update /share/subfolder resetacl”**

### **“SECURITY ID STRUCTURE INVALID”:**

Windows Client Pop-Up Error Message

**Cause:** ACL’s are corrupt, missing, or not populated with User/Group permissions

--Use ACL DUMP to examine ACLs on Folder or File

--Set Server Logon Traces: “param NTsec logonTraces=6” & have User retry

--Examine Server Log for SMB Error Code message for User & debug using “ntstatus 0xc0000007”

--Look at Celerra Security Event Log for “failure” messages in Event Viewer

## **CIFS ACL CHANGES IN NAS CODE:**

**Effective Rights:** RWXPDO    NAS Code 2.1.5 thru 2.2.46 uses algorithms to convert Effective Rights to Extended

**Extended Rights:** Bit-masked 4-byte value: NAS Code 2.2.46+ uses only Extended Rights, as does Win2k

|                                  |                                   |                                        |
|----------------------------------|-----------------------------------|----------------------------------------|
| Bit 1 = List/Read                | Bit 2 = Create files/Write Data   | Bit 3 = Create folders/Append data     |
| Bit 4 = Read Extended attributes | Bit 5 = Write Extended Attributes | Bit 6 = Traverse Folders/Execute Files |
| Bit 7 = Delete Subfolders/files  | Bit 8 = Read Attributes           | Bit 9 = Write Attributes               |
| Bit 10-16 Unused                 | Bit 17 = Delete                   | Bit 18 = Read permissions              |
| Bit 19 = Change permissions      | Bit 20 = Take Ownership           | Bit 21-24 Unused                       |

## **CELERRA AND ACCESS CONTROL LISTS:**

For Celerra, each file or directory object or Share contains an ACL [Access Control List], which consists of ACE's for Users and Groups. The ACE contains the permission set for the User or Group, the SID, Effective, and Extended NT Rights, as depicted in the values above.

**--ACL's are stored on the root of every CIFS file system!!**

## **VIEWING CIFS PARAMETERS:**

**\$server\_config server\_2 -v 32768 "param cifs"** [from 2.2.46.6 code]

Note: With NAS 5.3, the .server\_config commands are wrapped in XML so as to be able to output the full contents of commands to screen, as the older RPC had 16k limitation. Beginning with NAS 5.4, the 32k output will be the default.

**TO DUMP COMPLETE PARAM LIST TO SERVER LOG: \$server\_config server\_2 "param cifs"**

## **CIFS PROTOCOL[Common Internet File System]:**

CIFS is a file sharing and network access protocol based on the Server Message Block (SMB) protocol, adopted by Microsoft to run on top of NetBIOS (Network Basic Input/Output System)--the building blocks upon which traditional Windows-based Microsoft Networks are built. NetBIOS served as a 'Session Layer API' that provided (3) basic functions for a Microsoft Network: 1) Name Service host resolution 2) Datagram Delivery Service using UDP 3) Session Service that established & maintained point-to-point connection-oriented NetBIOS sessions over TCP/IP between 'networked' computer systems.

More recently Microsoft redesigned CIFS to be the native file sharing protocol for Windows 2000 and implemented SMB directly over TCP/IP using Port 445 (traditional networks used SMB over NBT over TCP/IP using Port 139), making CIFS 'transport independent.'

CIFS is implemented as a Client/Server request & response protocol that supports File Sharing Services, Messaging, File Locking, Security, Authentication, & Short/Long File Naming conventions using communication mechanisms related to File Sharing, API's, Named Pipes, & Mail Slot communications. Client systems issue SMB commands that are handled by the local 'redirector' to connect to 'remote systems' or resources using connection establishment messages. After authentication & security considerations, clients receive access to Shares for opening, reading, writing, and executing files. Other forms of messages between Client & Server are Mail Slots and Named Pipes [Mail Slots is a connectionless broadcast delivery service often used in Browsing services, while Named Pipes is a form of connection-oriented messaging via a virtual circuit or 'pipe', established between client & server, so that two applications or processes can communicate with each other. Data is passed as output from (1) host process to input on the other host].

Remote Procedure Calls (RPC) are the basic mechanism by which a Client computer can make a network request to use the processing capabilities of a Server--both Client/Servers use an End Point Mapper service to listen on Port 135 for TCP/IP connection requests. Clients must then conduct an RPC-bind to an interface before it can begin making SMB or other types of procedure calls.

Celerra implements the CIFS file sharing protocol to allow network access for Windows-based Users--CIFS is implemented inside the Operating System kernel for better performance and is layered on top of DART (Data Access in Real Time).

**Note:** With NAS 5.2, Celerra now supports "SMB Signing", a form of host-based authentication [See MS--Q161372].

## **SMB2:**

Celerra began supporting Microsoft's new and improved SMB2 CIFS protocol with NAS 5.6.42

## **TROUBLESHOOTING CIFS:**

**# server\_cifssupport server\_2** [Excellent CIFS troubleshooting utility added to 5.5.24.2—Napa 4]

**# server\_checkup server\_2** [New CIFS configuration troubleshooting utility to be released with GrandNapa 5.5.27.x]

**\$server\_config server\_3 "cifs stop"** [Use command to stop cifs when other command fails]

**\$server\_cifs server\_3 -o audit** [Returns open connections to data mover]

**\$server\_cifsstat server\_3 -full** [-summary; -z to zero] {SMB Statistics--Open Connections to DM, SMB calls, ops per second; total operations}

**\$server\_cifsstat server\_x -s** [Provides shortened version of stats]

**\$ server\_cifsstat ALL -s |grep -A 1 -i open**

Open connection Open files

130 340

### \$server\_config server\_3 -v "cifs audit full" or "cifs audit full debug"

```
1147962378: SMB: 4: ||| AUDIT Ctx=0xd827c004, ref=1, XP Client(XP1) Port=1811/445
1147962378: SMB: 4: ||| IP[2K3] on if=cge
1147962378: SMB: 4: ||| CurrentDC 0xd8950004=GEORGE
1147962378: SMB: 4: ||| Proto=NT1, Arch=Win2K, RemBufsz=0xffff, LocBufsz=0xffff
1147962378: SMB: 4: ||| 0 FNN in FNNlist NbUsr=1 NbCnx=2
1147962378: SMB: 4: ||| Uid=0x3f NTcred(0xd8280004 RC=3 NTLMSSP Capa=0x10801) '2K3\tmatta'
1147962378: SMB: 4: ||| Cnxp(0xd8951e04), Name=IPC$, cUid=0x3f Tid=0x3f, Ref=1, Aborted=0
1147962378: SMB: 4: | readOnly=0, umask=22, opened files/dirs=0| Absolute path of the share=\.etc
1147962378: SMB: 4: ||| Cnxp(0xd7979c04), Name=ip_share, cUid=0x3f Tid=0x41, Ref=1, Aborted=0
1147962378: SMB: 4: | readOnly=0, umask=22, opened files/dirs=1
```

Note: Dumps information on User connections and open files, but only if Users are connected to Server

### \$server\_config server\_3 -v "param NTsec logonTraces=6" [Increases logon levels; set 'logonTraces=3' when done]

Note: With 5.6, you may need to use \$ server\_log server\_x -i option to see trace output in the server log

\$server\_sysstat server\_3 # of threads in Use--should be a low number

\$server\_config server\_3 -v "printstats cifs"

\$server\_config server\_3 -v "param cifs"

\$ .server\_config server\_2 -v "cifsThrd" [new facility with NAS 5.4]

1110213490: SMB: 4: cifsThrd: checking blocked threads

minimum: 0/256 threads, 5 seconds

\$server\_config server\_3 -v "printstats meminfo" or "printstats mem"

\$server\_config server\_3 -v "acl if=ana0 dump=/fsdata/prod"

\$server\_config server\_3 -v "ntcred client=0x384dom user=robin" [Dumps perms for user 'robin' to server\_log]

\$server\_config server\_3 -v "ntcred client=168.244.127.17" [or machinename]

\$server\_config server\_3 -v "ntcred user=watson"

\$server\_config server\_3 -v "ntcred credp=0x6b87c04" [Dumps perms for user 'robin' to server\_log]

\$ .server\_config server\_2 -v "ntcred all"

# .server\_config server\_2 -v "samr if=10\_241\_169\_49 user=nas1"

1099322602: SMB: 4: RID=0201 GID=2002 A:7 U:2 ='Domain Users'

1099322602: SMB: 4: RID=0462 GID=2009 A:7 U:2 ='NAS Users'

\$server\_config server\_3 -v "sidcache global status" [Look for High Cache Miss rate]

Sid cache status

Enable:1

DefaultSize:53

Hits:0

Miss:2

Count:0 Collisions:0

Total of collisions:0 Total of hash OK:0 Compare:0

\$ .server\_config server\_5 -v "sidcache if=ana0 global dump"

\$server\_config server\_5 -v "srwpwd dump" [Checks machine account trust relationship between Server and Domain by verifying machine passwd against DC—used for IPC\$ netlogon and membership in domain]

\$server\_config server\_2 -v "tcp audit" [Useful to see what Ports datamover listening to & IP Addresses]

Caution: Use this command in conjunction with "cifs audit" output to identify the hex value for the User account in question. An incorrect hex value entered here could panic the datamover!

\$server\_config server\_3 "logsys set severity SMB=LOG\_DEBUG" [Debugging SMB problems]

\$server\_config server\_3 "logsys set severity SMB=LOG\_PRINTF" [Turn off SMB debug mode]

### VERIFYING CURRENT LOGGING LEVEL FOR CATEGORIES:

# .server\_config server\_2 -v "logsys get severity SMB"

1158591814: LIB: 4: Server log severity for facility SMB is 4

### IMPORTANT CIFS TROUBLESHOOTING TOOLS:

1. \$ server\_cifssupport utility added with NAS 5.5.24.2

2. \$ server\_checkup utility to be added with NAS 5.5.27.x (Spring 2007)

3. Set Server Log Debugging: "logsys set severity SMB=LOG\_DBG3" [\_ERR to turn off]

4. Increased Logon Traces: "param NTsec logonTraces=6" [=3 to turn off]
5. Audit User Accounts: 16384 "cifs audit full"
6. Debug Server Log SMB Errors: "ntstatus 0xc0000073" →NONE\_MAPPED
7. Dump Folder/File Permissions: "acl if=ana0 dump=/fsprod/prod2"
8. Dump User Credentials: "ntcred client=MACHINENAME or IP user=thomas" | "ntcred all"
9. View Event Viewer "Security Event Logs" on DataMover and increase Log Size [Access via Event Viewer NT Program]
10. Verify Usrmapper Process, Database entries, check Usrmapper.log, Server Log for errors
11. Verify PDC Connections: -v "pdc dump"
12. Look at Server CIFS output to verify CIFS configuration [crosscheck with /nas/server/slot\_x/netd file]
13. Observe User Connections/Activities: "server\_cifsstat ALL -s" [SMB Load, Open Files & Connections]
14. Observer CPU/Memory Activity: "server\_sysstat ALL"
15. Run setacl, resetacl, cifs update of Shadow File if required
16. Verify Shadow File and/or Fix using: "shadow readdir \personal\its" & "shadow fix \personal\its"
17. Windows commandline Utility to debug NT Errors: err.exe
18. Windows DOS prompt help utility: C:>net helpmsg 1208  
An extended error has occurred.
19. MS Windows Errors—link to website:  
<http://msdn2.microsoft.com/en-us/library/ms681381.aspx>
20. Checking for blocked CIFS Threads: "cifsThrd"

### **DEBUGGING SMB SERVER LOG ERRORS:** SMB: 3: NetLogon::buildSecureChannel=7 E=0xc0000022

**Run Following Command:** **\$server\_config server\_2 -v "ntstatus 0xc0000022"**

1023977569: SMB: 4: NT status 0xc0000022: **ACCESS\_DENIED**  
1023977569: ADMIN: 4: Command succeeded: ntstatus 0xc0000022

**\$server\_config server\_2 -v "ntstatus 0xc000023"**

1108730617: SMB: 4: NT status 0xc0000233: DOMAIN\_CONTROLLER\_NOT\_FOUND

**\$server\_config server\_2 -v "ntstatus 0xc0000bb"**

1108730807: SMB: 4: NT status 0xc00000bb: NOT\_SUPPORTED

### **COMMON SMB ERROR MESSAGES:**

"ntstatus 0xc0000073" → NONE\_MAPPED  
"ntstatus 0xc0000016" → MORE\_PROCESSING\_REQUIRED  
"ntstatus 0xc000015b" → LOGON\_TYPE\_NOT\_GRANTED  
"ntstatus 0xc000006d" → LOGON\_FAILURE  
"ntstatus 0xc0000022" → ACCESS\_DENIED  
"ntstatus 0xc000018c" → TRUSTED\_DOMAIN\_FAILURE  
"ntstatus 0xc0000078" → INVALID\_SID  
"ntstatus 0xc00000e5" → STATUS\_INTERNAL\_ERROR

### **CIFS STATISTICS:**

**\$server\_cifsstat ALL -s \$server\_cifsstat ALL -z**

Compare "cifsstat" 'State info:' & 'Open connection Open files' to the cifs audit output [server\_cifs server\_x -o audit]  
→server\_cifsstat reports Open Connections to Data Mover, not total number of "SMB Sessions"—each Session can involve use of multiple connections, such as a Client connecting to several different Shares on the same Server. SessUsers param defines the number of SMB Sessions that can be established from a specific client system—observe Sessions from Computer Manager.

#### **Note "SessSetupX" Times:**

**SessSetupX 167997 0.04 13869.96 4.13** [Indicates healthy and fast Session Setups]

"Opened files/dirs=0" Indicates an "Open connection" but not an 'Open files'

"Opened files/dirs=1" Indicates an "Open connection" and an "Open files"

"Total Trans2Smb/NTSmb"

--Trans2 and TransNT are special commands--Trans2 are used mainly to get/set information on filesystem and files. TransNT are used to Create/open files with NT clients and send special request like IOCTL and Notification requests.

--TransNTCreate requests are not specifically outlined in Cifsstat Output, but are used in backup/restore applications

### **SERVER CIFSSTAT OUTPUT:** # server\_cifsstat server\_6

#### **SMB statistics:**

| proc    | ncalls   | %totcalls | maxTime | ms/call |
|---------|----------|-----------|---------|---------|
| Mkdir   | 103      | 0.00      | 47.91   | 4.88    |
| Rmdir   | 18       | 0.00      | 84.56   | 25.11   |
| Close   | 33936449 | 12.82     | 8738.58 | 0.05    |
| Flush   | 81789    | 0.03      | 1263.04 | 2.62    |
| Unlink  | 213000   | 0.08      | 9965.19 | 2.05    |
| Rename  | 30384    | 0.01      | 756.07  | 1.34    |
| GetAttr | 51363    | 0.02      | 164.85  | 0.34    |

|               |           |       |          |       |
|---------------|-----------|-------|----------|-------|
| SetAttr       | 122       | 0.00  | 93.64    | 1.97  |
| Read          | 22477     | 0.01  | 33.93    | 0.11  |
| Write         | 426095    | 0.16  | 5731.77  | 0.51  |
| Unlock        | 3037      | 0.00  | 17.57    | 0.38  |
| ChkPath       | 123269    | 0.05  | 921.37   | 3.34  |
| Lseek         | 12062     | 0.00  | 0.02     | 0.01  |
| ReadBlockRaw  | 485114    | 0.18  | 1001.34  | 0.65  |
| WriteBlockRaw | 161       | 0.00  | 54.09    | 13.28 |
| LockingX      | 2090094   | 0.79  | 30026.35 | 2.55  |
| Trans         | 270749    | 0.10  | 12361.55 | 0.51  |
| Echo          | 216070    | 0.08  | 0.03     | 0.01  |
| OpenX         | 18604     | 0.01  | 10007.94 | 5.64  |
| ReadX         | 36395296  | 13.74 | 8999.47  | 0.39  |
| WriteX        | 2257952   | 0.85  | 6695.34  | 0.59  |
| Trans2Prim    | 115904021 | 43.77 | 12897.62 | 0.25  |
| FindClose2    | 121979    | 0.05  | 1.06     | 0.00  |
| TreeDisko     | 71725     | 0.03  | 9.31     | 0.04  |
| NegProt       | 17883     | 0.01  | 0.21     | 0.12  |

**SessSetupX 93729 0.04 284718.61 4764.62 [Indicates long session setups]**

|              |          |       |          |        |
|--------------|----------|-------|----------|--------|
| UserLogoffX  | 43230    | 0.02  | 27.27    | 0.40   |
| TreeConnectX | 73576    | 0.03  | 601.78   | 0.53   |
| DiskAttr     | 837      | 0.00  | 29.10    | 0.08   |
| Search       | 27       | 0.00  | 1959.60  | 576.04 |
| TransNT      | 32210613 | 12.16 | 4969.74  | 0.04   |
| CreateNTX    | 39619456 | 14.96 | 13026.27 | 0.33   |
| CancelINT    | 26321    | 0.01  | 8.96     | 0.02   |

**Trans2 SMBs:**

| proc        | ncalls   | %totcalls | maxTime  | ms/call |
|-------------|----------|-----------|----------|---------|
| FindFirst   | 73548722 | 63.46     | 12897.60 | 0.28    |
| FindNext    | 546687   | 0.47      | 3948.37  | 4.29    |
| QFsInfo     | 403226   | 0.35      | 53.62    | 0.05    |
| QPathInfo   | 20199195 | 17.43     | 2865.52  | 0.15    |
| QFileInfo   | 20835106 | 17.98     | 1427.06  | 0.05    |
| SetFileInfo | 370689   | 0.32      | 2530.81  | 1.32    |

**NT SMBs:**

| proc         | ncalls   | %totcalls | maxTime | ms/call |
|--------------|----------|-----------|---------|---------|
| Create       | 1        | 0.00      | 14.19   | 14.00   |
| SetSD        | 1002891  | 3.45      | 4969.73 | 0.80    |
| NotifyChange | 154493   | 0.53      | 9.21    | 0.01    |
| QuerySD      | 27884178 | 96.01     | 17.05   | 0.01    |

**Performance info:**

| Read     | Re/s    | Write   | Wr/s    | All       |
|----------|---------|---------|---------|-----------|
| Ops/sec  |         |         |         |           |
| 36902887 | 2554.03 | 2684208 | 1730.17 | 412931843 |
| 761.20   |         |         |         |           |

**State info: [SMB Connections and Open Files on DataMover]****Open connection      Open files****199      9****Shadow info:**

| Reads      | Writes  | Splits | Extinsert | Truncates |
|------------|---------|--------|-----------|-----------|
| 1676555672 | 1059217 | 38     | 0         | 0         |

**SMB total requests:**

| totalAllSmb | totalSmb  | totalTrans2Smb | totalTransNTSmb (unsupported) |
|-------------|-----------|----------------|-------------------------------|
| 412931843   | 264817605 | 115903625      | 32210613    3169050           |

**Max number of CIFS open file handles on a DataMover: 10,000****Max UID's [User ID's] on a DataMover: 4,000,000,000****Max GID's [Group ID's] on a DataMover: 65,534****AUDITING CIFS USERS: \$server\_cifs server\_4 -o audit [Shows active clients connected to Celerra]****\$server\_config server\_4 16384 "cifs audit full"** [verbose version]**\$server\_config server\_x -v "cifs audit"** [abbreviated version]**\$server\_cifs server\_x -o audit** [Generates list of Users connected, Files opened, and dumps into Server\_4's Log]**\$server\_config server\_x -v "ntcred user=xxx" | "ntcred all"****\$server\_cifssupport server\_x**

## **SETTING UP CELERRA EVENT LOG AUDITING:**

1. From the Application and Tools CD, install the Celerra Management MMC snapins on the Windows management station
2. From Windows, go to Programs>Administrative Tools>Celerra Management, to open the snapin
3. Rightclick 'Data Mover Management', select 'Connect to Data Mover...', and enter compname, "Name: Harry"
  - a.) Under Data Mover Security Settings, select Audit Policy, rightclick and "Enable auditing"
  - b.) Doubleclick the Events to Audit, such as "Audit Object Access", and set policy for 'Success and/or Failure' of the event
  - c.) Close interface
4. Go to a drive mapped to the Celerra Share where file or directory objects are to be audited, rightclick, select Properties>Security tab>Advanced>Auditing>Add [Add the Group or User to be audited, such as the group Everyone, if auditing access on objects for all Users]
5. Select the type of attributes to audit [such as Read, Write, Delete, Change Ownership, etc] and apply
6. Auditing is now in place.
7. To verify & view auditing events for Celerra "objects", open Event Viewer on Windows: Start>Run: eventvwr>highlight Eventviewer>Rightclick and "Connect to Another computer">Type in name of CIFS Server>Click on "Security" and observe Events that are recorded in the righthand window pane

**Note:** In Event Viewer, look at the "Properties" of the "Security" item to observe the default settings and location of the security.evt file [Default is Overwrite Events older than 10 days, 512kb is Security Log size, and default location of "security.evt" log for Celerra Server is c:\security.evt]. The size of the security event log can be increased, but only when moving the log off the roots of the Data Mover to a dedicated file system.

## **CHANGING THE LOCATION AND SIZE OF THE CELERRA SECURITY EVENT LOG:**

**Note:** This procedure illustrates how to create a dedicated file system for use with Celerra Event Logging, how to change the current location of the security.evt log file from the roots of the Data Mover to a dedicated file system, and finally, how to change the default Log size from 512KB to some larger size.

1. Using Celerra Manager, create a file system called "event\_log" and a custom mountpoint called "/winnt"

Custom

Pathname: **/winnt**

2. As Administrator, create a hidden share to the /winnt mountpoint called "winnt\$" on CIFS server 'Harry'

3. Map to the hidden share "winnt\$" from a Windows system, and create the following directory structure underneath the \winnt mountpoint:

\harry\winnt\$

\system32\config

4. Open the Registry Editor from a Windows system (Start>Run>Regedit), go to Registry>Connect Network Registry>type in the name of the Celerra CIFS server name (Harry)

**Note:** Ignore the popup that says "Unable to connect to all the roots of the computer's registry"

5. Navigate from the CIFS server name registry hive to the Security key:

HKEY\_LOCAL\_MACHINE>System\CurrentControlSet\Services\Eventlog\Security

6. Doubleclick and edit the "File" data location, changing the location from the default "c:\security.evt" to the new file system location and path:

**Default location:**

**c:\security.evt**

**New location:**

**c:\winnt\system32\config\security.evt**

7. Close the Registry Editor and open Event Viewer (Start>Run>eventvwr)

8. Highlight Event Viewer, rightclick and "Connect to another computer", type in CIFS Server name>Highlight the "Security" item, rightclick and "Clear all events" (Click o.k. to close the dialogue without saving any events)

**Note:** The new log location will not go into effect without first clearing all events from the old log

9. Rightclick "Security" >Properties, and change the 'Maximum log size' from the default 512kb to the desired new value (e.g., 512000KB = 512MB)

10. Finally, for the section "When maximum log size is reached:", select the appropriate setting to overwrite or not to overwrite events, etc.

## **VIEWING CELERRA SECURITY EVENT LOG:**

1. Copy msaudite.dll from Windows system to Control Station and push file to .etc of Data Mover

2. Connect to Data Mover registry using regedit & and modify the following keys:

HKLM>System>CurrentControlSet>Services>Eventlog>Security>Security>CategoryMessageFile

HKLM>System>CurrentControlSet>Services>Eventlog>Security>Security>EventMessageFile

Add c:\etc\msaudite.dll to each key

3. Close registry editor

## **GPO POLICY ALLOWS DEFAULT EVENT LOG SIZE TO BE > THAN 512kb:**

The historical size restriction for all Event Logs (Security, Application, System) for DART has been defined as 512kb when the logs are located on the roots of the Data Mover. The reason for this restriction is because the rootfs is traditionally very small in size (typically between 16 and 128MB). Even with rootfs extensions, this limitation has not been lifted. Under normal circumstances, changes made to increase the size of the log from 512kb to a higher value (when the logs reside on the rootfs) are restricted by DART code for Event Log or Registry changes. However, for GPO policies, DART fails to enforce the size restriction and it becomes possible to create a log size larger than the 512kb default on the rootfs of the Data Mover.

**Note:** See emc123515 and AR71688—issued resolved with 5.4.21.0

## **\$server\_security server\_3 -info -policy gpo**

server\_3 :

Server compname: tyos00510002

Maximum security log size (Kilobytes): 256000

## **ACCESSING WINDOWS 2003 GPO EDITOR:**

Start>Run>gpedit.msc>Group Policy Object Editor>Computer Configuration>Administrative Templates>DNS Client

## **CHECKING SECURITY LOG STATE:**

### **\$ .server\_config server\_5 -v "ntlog security state"**

Security log file status:

PATH:/winnt/jt1/security.evt

MAX SIZE:0x7d00000 (Next=0x6b70000 reg=0x6b70000 gpo=0x7d00000)

RETENTION:0xffffffff (reg=0xffffffff gpo=0xffffffff(Not Defined))

GUEST ACCESS RESTRICTED: No

LOG:opened

SIZE:0x1c18

FLAGS:0x1

OLDEST EVENT:1

NEXT EVENT:1d

START OFFSET:30

END OFFSET:1bf0

AUDIT POLICY (0xffffffff: not defined 4:disable 1:audit success, 2:audit failure)

Current Settings

-----

Mode=1

system=3 logon=3 objectAccess=3 privilegeUse=3

policyChange=3 accountManagement=3 DirectoryService=3 accountLogon=3

Registry Settings

## **SETTING USER RIGHTS ON CELERRA IN WINDOWS 2000 ENVIRONMENTS:**

Programs>Administrative Tools>Celerra Management>EMC Celerra Management>Data Mover Management [Rightclick to connect to compname]>Data Mover Security Settings>Select ‘User Rights Assignment’ and set the appropriate right:

Take Ownership of files and other objects

Backup files and directories

Restore files and directories

Bypass traverse checking

Generate Security audits

Adjust memory quotas for a process

Manage auditing and security log

EMC Virus Checking

Access this computer from the network

**Note:** User Rights available NAS 5.3.14

## **NTRIGHTS TOOL:**

c:>ntrights -u administrators -m <\\dm2-cge0-1> -e add +r seAuditPrivilege [Syntax to set Generate Security Audit privilege for Administrators group]

## **SETTING VIRUSCHECKING PRIVILEGE FOR AV USER USING NTRIGHTS UTILITY:**

<C:\Documents and Settings\Administrator>ntrights -u europe\cavauser -m 10.241.183.85 +r SeVirusChecking>

Granting SeVirusChecking to europe\cavauser on 10.241.183.85... successful

## **SETTING CIFS “SMB” USER LOGON TRACES ON DM:**

**\$server\_config server\_2 -v "param NTsec logonTraces=6"** [Enables SMB Logging]

**\$server\_config server\_2 -v "param NTsec logonTraces=3"** [Resets to normal]

**Example of Log Entries with ‘Traces=6’ Enabled:**

2002-10-05 14:28:18: SMB: 4: replyNT1 capa=8801/8000e3fd

2002-10-05 14:28:18: SMB: 4: SSXAKsc host=host@fsdata.network.lan

2002-10-05 14:28:18: SMB: 4: SSXAK: accept\_sec\_context status:NT=0 0,0

**2002-10-05 14:28:18: SMB: 4: SSXAK map UID=12152 GID=45678 for U=rthomas D=NETWORK**

2002-10-05 14:28:18: SMB: 4: SSXAuth\_KERBEROS=0

2002-10-05 14:28:18: SMB: 4: SSXAuth\_SERVER\_EXT15 aT=5 mT=0 0

**2002-10-05 14:28:29: SMB: 4: User NETWORK\rthomas primary GID set to 2ef3 (Unix=b26e)**

2002-10-05 14:28:29: SMB: 4: addNewNTGids 15/84=-16

**DEFAULT TIMEOUT VALUES FOR DC CONNECTIONS (CIFS TIMEOUTS):**

**NTsec.DCConnectTimeout=1388 = 5000 decimal = 5 secs.** [length of time waiting to connect to DC]

**NTsec.DCTimeout=4e20 = 20000 = 20 seconds** [length of time waiting for reply from DC before trying another]

cifs.NTsec.DCTimeout 0x015d6b54 0x00004e20 0x00004e20

**Note:** DataMover will reselect next DC on its list after the 20-second timeout value if no further communication is occurring. In certain situations, such as for orphaned SIDs, or SIDs that cannot be resolved in other domains because of Trusts or firewall issues, Clients may experience long “pauses” as directories are traversed—this is because of the 20sec CIFS timeout value. In extreme cases, setting this value to 1388=5000=5 secs may help speed things up, but could have adverse impact & cause DC cycling.

**SECURE CHANNEL COMMUNICATIONS AND NTLMv2 SUPPORT:** NAS 5.4.16.0/5.3.18.0 +

With the implementation of native NTLMv2 support and subsequent fixes provided in AR59648, the first CIFS server of the Data Mover will open a secure channel connection to a Domain Controller. Other CIFS servers will open a separate secure channel whenever the first client connects only if that client is using NTLMv2--this behavior is different from the default behavior in the past, when the Celerra only opened a single secure channel to the Domain Controller that was used as the security context for all CIFS servers on the Data Mover. If NTLMv2 is not being used by the Client, then client connections to other CIFS servers on the Data Mover will piggyback over the first secure channel session that was opened without opening a separate secure channel. Clients that connect to the Data Mover using WINS name resolution, or direct IP address mappings, will default to NTLM, or NTLMv2 if set.

**EXAMPLE:**

**# .server\_config server\_2 -v "pdc dump"**

DC\*DC9a6c600c USADC50[USA](157.184.147.198) ref=42 time(getdc184)=0 ms

Pid=100c Tid=0800 Uid=0800

Cnx=NO\_TRUST\_SAM\_ACCOUNT,17 logon=Authenticate2InvalidReply 3 SecureChannel(s):

[LEXFILE16] Fid=0x801 CallID=0xa Status=SUCCESS/SecureChannelOK

[LEXFILE1] Fid=0x802 CallID=0xb Status=SUCCESS/SecureChannelOK

[LEXFILE13] Fid=0x803 CallID=0xa Status=SUCCESS/SecureChannelOK

**USING PDC DUMP TO VERIFY DM CONNECTIONS TO DC's:**

**\$server\_config server\_x -v 16384 "pdc dump"**

**Note:** Windows 2000 datamovers can have multiple “>DC” connections established. (1) for Kerberos AD Services, (1) for LDAP Directory Services, etc. PDC Dump can also tell you how the DataMover discovered the DC--BROADCAST, WINS, DNS

**EXAMPLE:** [Server Cifs Output and PDC Dump Output]

DOMAIN NETWORK FQDN=network.lan

SID=S-1-5-15-f407b588-8201be4d-bda18164-ffffffffff

>DC=UTIL002(172.19.3.254) ref=136 time=3 ms

DC=UTIL003(172.19.3.253) ref=240 time=0 ms

>DC=UTIL004(172.19.3.252) ref=33 time=0 ms

**# .server\_config server\_2 -v 16384 "pdc dump"**

1033865308: SMB: 4: Dump DC for dom='NETWORK' OrdNum=0

>DC=DC12c9a104 UTIL002[NETWORK](172.19.3.254) ref=137 time=3 ms

Pid=1003 Tid=5007 Uid=3004 Cnx=2,0 logon=1 [TID & UID > than 0 means connections established to DC]

refCount=137 newElectedDC=12c9a104 dis=0 inv=0 [dis=0 means DC isn't disable; inv=0 means DC isn't invalid]

Discovered from: WINS DNS

>DC=DC13269f04 UTIL004[NETWORK](172.19.3.252) ref=33 time=0 ms

Pid=1002 Tid=1003 Uid=1803 Cnx=2,0 logon=1

refCount=33 newElectedDC=13269f04 dis=0 inv=0

## Discovered from: BROADCAST WINS DNS

### PDC INVALIDATE/VALIDATE COMMAND:

**\$ .server\_config server\_5 -v "pdc invalidate=10.64.25.101"**

1058881433: SMB: 4: TSWORKSTATION DC invalidate  
1058881433: SMB: 4: DCInvalidate TSWORKSTATION from cmdLine  
1058881433: ADMIN: 4: Command succeeded: pdc invalidate=10.64.25.101

**Note:** Use command to invalidate or validate a specific DC

### SERVER CIFS OUTPUT AFTER RUNNING INVALIDATE COMMAND:

DOMAIN TS\_NAS

SID=S-1-5-15-209b5669-55a118ba-7e4b2f8f-ffffffffff

>DC=SNEEZY(10.64.25.106) ref=4 time=0 ms

**DC=TSWORKSTATION(10.64.25.101) ref=2 time=136 ms INVALID**

**\$ .server\_config server\_5 -v "pdc validate=10.64.25.101"**

**Note:** Removes the “INVALID” designation for DC if successful

### VALIDATING MACHINE ACCOUNT PASSWD IN THE WINDOWS DOMAIN:

**\$ .server\_config server\_5 -v "srwpwd dump"**

1068222796: SMB: 4: srwpwd V=1 fileV=1 1 records  
1068222796: SMB: 4: Entry:4[tech] @ 12[SEINFELD.COM]  
1068222796: SMB: 4: Time: 0x3f79a7f7 Now=0x3fabc94c D=0x322155  
1068222796: SMB: 4: Account: 5[TECH\$]  
1068222796: SMB: 4: Password:15[LWbegMgTBxzB48r]  
1068222796: SMB: 4: GUID: 5de1e21f-9b6a-426b-b971-0ff657eba7d3  
1068222796: ADMIN: 4: Command succeeded: srwpwd dump

**Note:** Verifies machine account trust relationship between Server and Domain by verifying machine passwd against DC—used for IPC\$ netlogon and establishes membership in domain

### SERVER PASSWORD COMMANDS (Pre-NAS 5.5):

**\$ .server\_config server\_5 -v "srwpwd dump"**

**\$ .server\_config server\_5 -v "srwpwd display"**

**\$ .server\_config server\_4 -v "srwpwd compname=DM4 change"**

1095873609: SMB: 7: DomainJoin::modifyServerPassword: Reset server password dm4@MOUSE.COM using 'DM4\$' -  
machineAccount: DM4\$

**\$ .server\_config server\_4 -v "srwpwd compname=DM4 minutes=8640"**

**Note:** Changes frequency of password change from default 10020 minutes to 8640 minutes, or every 6 days.

### DATA MOVER MACHINE ACCOUNT PASSWORD (NAS 5.2):

→Windows NT/2000/2003 computers use a machine-account password to establish & maintain a ‘trust’ relationship with the Domain  
→The computer account password is used to authenticate the server during communication with the DC  
→Password is also the hash base for generating the Server’s Kerberos service keys  
→AD stores the password for each computer and generates Kerberos Service tickets for clients that wish to access the Server  
→NT 4.0 systems automatically change their passwords every 7 days while Windows 2000 changes every 30 days  
→Password is generated during Join process using either Kerberos set password protocol, or MSRPC SAMR set password call  
With the introduction of NAS 5.2, Celerra CIFS Servers automatically change their machine-account password every 7 days by authenticating to a Domain Controller, changing the password, then recalculating their Kerberos Encryption/Decryption keys based on the new password. Typically, Clients cache their Kerberos Service Tickets and may need to request new tickets after the CIFS Server has changed its password. A problem can occur if the Clients connect to a DC that does not yet have the Data Mover’s latest password change (i.e., replication not completed yet), meaning that the Client’s Service Ticket will still be invalid and DART will reply with MORE\_PROCESSING\_REQUIRED.

### NAS 5.5:

Please be aware that NAS 5.5 has completely revamped the machine account password system on DART so that a configurable number of passwords can be retained on the Data Mover in a history file, and that by default, the actual password change mechanism itself is not enabled. To put machine account password change in effect, provide a value in minutes to the param updMinutes.

### WORKAROUND FOR CIFS SERVER PASSWORD CHANGE ISSUES:

**param cifs srwpwd.updtMinutes=0**

**Note:** Set the above parameter to disable the weekly CIFS password change, then reboot the Data Mover to put into effect—this is the only way to make this param effective. See Primus emc92464 for more details. The automatic password change can also induce a race condition panic with DART, and more recently has been directly tied to a loss of domain trust as a result of the passwd change attempt failing, requiring Rejoining the compname to the domain via Unjoin/Rejoin or resetserverpasswd command.

### # server\_param server\_2 -facility cifs -i srvpwd.updtMinutes -v (NAS 5.5)

server\_2 :

```

name          = srvpwd.updtMinutes
facility_name = cifs
default_value = 0
current_value = 0
configured_value =
user_action    = reboot DataMover
change_effective = reboot DataMover
range         = (0,4294967295)
description    = Time interval between password changes
detailed_description

```

Defines the time interval between two server password changes in minutes, by default, it is 0 (not enabled). The Windows default value is 7 days - 1 hour (0x2724).

## TROUBLESHOOTING CIFS OUTAGE PROBLEM WITH CFS 2000 DM: NAS 4.2.9.0

**Intro:** Case 8715047, AR28374

**Scenario:** Customer experiences periodic CIFS Outages because DC is running out of resources

**Problem:** Data Mover uses its netlogon connection to IPC\$ share on the DC using NULL SMB session, then opens communication channel via RPC ‘named pipes’ in order to bind to DC’s LSA Interface [Local Security Authority for authentication calls]. However, rather than conducting multiple “\lsarpc” calls [OpenPolicy2/LookupSIDs2] down a single pipe, we are opening pipes for each command, which eventually consume the DC’s memory resources, and a CIFS outage occurs. Problem could be exacerbated by DM calls to DC to try and resolve local machine account SIDs. DC’s, however, do not know local machine account SIDs for individual Windows systems and cannot answer DM correctly. DM “switching” to alternate DC could then occur, as DM seeks to find a DC that can resolve the unknown SID [STATUS\_NO SUCH\_USER].

### How DM Queries DC When Client Makes Network Request for Resources:

As with any member Server, DM must query DC for username associated with SID of Client making Network Resource request. Data Mover uses its connection to default IPC\$ share on DC to communicate & issue commands via named pipes. DM sends SMB NTCCreateAndX calls to open the \lsarpc “magic” file via names pipes, which is used to talk to the DCERPC/LSA authority on the DC. DM queries DC for username of SID by opening “magic” \lsarpc file via LSA service and issues “OpenPolicy2” command to DC, and “OpenPolicy2/LookupSIDs2()” command via SMB-Transaction/Named Pipes to obtain username from SID. DM opens a new pipe for each call if required.

### DM-to-DC AUTHENTICATION CALLS:

DM→connected to DC via IPC\$ share

DM→binds to LSA interface on DC

DM→sends \lsarpc call OpenPolicy2 to DC \\AUS

DM→sends \lsarpc call LookupSIDs2 for SIDs of users/groups

### Symptoms:

1. Most obvious is a CIFS outage on DM, as well as port 445 “Too many connections open in listen queue”
2. Windows 2000 ADUC Interface sees hundreds of Pipe\LSARPC connections to DC
3. Netmon Traces show “STATUS\_PIPE\_NOT\_AVAILABLE” & “STATUS\_IO\_TIMEOUT” errors, which are indicators that there is a resource issue with the affected DC.

### Troubleshooting Commands:

1. \$server\_config server\_x -v “pdc dump”
2. \$server\_cifs server\_x
3. \$server\_config server\_x “param NTsec logonTraces=6” | Traces=3 [to turn off trace after done collection]
4. **\$server\_config server\_3 -v “pdc trace=1”** [Trace of DC actions] | trace=0 [to turn off PDC tracing after collection]
5. \$ .server\_config server\_3 -v “sidcache if=ana0 global status” [Shows health of SID Cache for User/Group lookups]

## HOW DM CHOOSES FASTEST-RESPONDING DC or LOCAL SUBNET DC:

Normally, DataMover will select the fastest-responding Domain Controller, but in 4.1, this may not always work. Set following parameter to make DataMover select DC on local Subnet, if available:

### /nas/server/slot\_x/param

**param cifs DC.useFastest=0** [Default Value is 1 for selecting “fastest-responding” DC]

**Note:** NAS 4.1 does not recompute DC List every 15 minutes as previous CIFS code did. NAS 4.2 does recompute DC list every 15 minutes and uses new mechanism for selecting DC on local Subnet, and then only (1) DC for all AD Services required

### **DUMPING ACLS [permissions] ON FOLDERS/FILES:**

**.server\_config server\_2 -v "acl if=ana0 dump=/mnt30/andy"**

### **DUMPING LOCALGROUPS FILE TO SCREEN:**

**.server\_config server\_2 -v 16384 "lg list vs=nas46"** [where nas46 = netbios name]

**.server\_config server\_2 -v 16384 "lg list"** [Dumps all Localgroups for DM]

### **RUNNING NTCRED TO VERIFY USER CREDENTIALS:**

**\$server\_config server\_4 -v "ntcred client=machinename user=rparsons" | "ntcred all"**

**\$server\_config server\_4 -v "ntcred client=172.19.8.10"**

### **COMMAND TO FIND CLIENT IP ADDRESS:**

**\$ server\_cifs cldm2 -o audit**

|||| AUDIT Ctx=0x8beae04, ref=0, Client (168.244.127.17) Port=1340

||| MTESTNET[LOWES] on if=int1

||| CurrentDC 0x8b0d904=DCCSC

||| Proto=NT1, Arch=Win2K, RemBufsz=0xffff, LocBufsz=0xffff

||| Uid=63 NTcred(0x8bc7904 RC=1) 'LOWES\mstestblephew\$' [User logged into Celerra]

||| Uid=64 NTcred(0x8bc6a04 RC=1) 'LOWES\celerraav' CHECKER [User logged into Celerra]

|| Cnxp(0x8c2f404), Name=IPC\$, Tid=63, Ref=1

|| readOnly=0, umask=22, opened files/dirs=0

|| Cnxp(0x8b90004), Name=CHECK\$, Tid=64, Ref=1

|| readOnly=0, umask=22, opened files/dirs=0

**Note:** Meaning of some values seen with cifs audit feature

Ctx = address in memory of the Stream Context

Arch = type of client O/S

RemBufsz = max buffer size client

LocBufsz = local buffer size negotiated

FNN/FNNlist = number blocked files not yet checked by VC

NbCnx = number connections to shares for this TCP connection

cUid = User ID who has opened the connection first

Tid = Tree ID representing share connection

### **CIFS RIGHTS & PRIVILEGES:**

Logon Rights: Controls who is authorized to log onto a computer and how [allow logon and deny logon]

Privileges: Controls access to system resources and override permissions on the local computer—manage via group membership

Permissions: Shares, Directories, Files

**\$ .server\_config server\_2 -v "ntcred user=administrator"**

1190081463: SMB: 4:

VDM=server\_2 CIFS SERVER=FESTUS[LAB] dump cred at 0xe96d8404

1190081463: SMB: 4: USER 'LAB\administrator'

Auth=NTLMSSP CredCapa=0x10801

SID = S-1-5-15-7a50bbc6-60415a8a-6b635f23-1f4

1190081463: SMB: 4: PRIMARY = S-1-5-15-7a50bbc6-60415a8a-6b635f23-201

1190081463: SMB: 4: **Priv=0x7f,0x3** DefOpt=0xe Adm=1, Backup/Adm=1 (Bkp=0) NTCapa=0x10801 →0x7 represents

Privileges, and 0x3 Logon Rights

**The full listing of privileges is as follows:**

|                   |   |      |                                                     |
|-------------------|---|------|-----------------------------------------------------|
| - SeTakeOwnership | = | 0x1  | →0x7 is a combination of the first three privileges |
| - SeBackup        | = | 0x2  |                                                     |
| - SeRestore       | = | 0x4  |                                                     |
| - SeChangeNotify  | = | 0x8  |                                                     |
| - SeAudit         | = | 0x10 |                                                     |
| - SeIncreaseQuota | = | 0x20 |                                                     |
| - SeSecurity      | = | 0x40 |                                                     |
| - SeVirusChecking | = | 0x80 |                                                     |

**The full listing of logon rights is as follows:**

- SeOpenLocally = 0x1 → 0x3 is a combination of the two logon rights
- SeNetworkLogonRight = 0x2

## **COMMENT ON OWNERSHIP & MEMBERSHIP IN LOCAL ADMIN GROUP:**

If a User is a member of the local Administrators group on a Server, the SID S-1-5-20-220 is used in the Owner field, and not the UID of the User. This mimics Windows behavior. So, file ownership is held by the Administrator group and not the User.

### **NTCRED OUTPUT:**

→Shows the cumulative credentials for a User, based on domain and local group memberships

**\$ .server\_config c1dm2 -v "ntcred client=168.244.127.17 user=lowes"**

CIFS SERVER=MTESTNET[lowes.com] dump cred at 0x8bc7904

**1039819890: SMB: 4: USER 'LOWES\celerraav'** [User connected to Celerra]

Auth=KERBEROS

SID = S-1-5-15-3814076-70c33b37-3b2108ea-14daf

1039819890: SMB: 4: PRIMARY = S-1-5-15-3814076-70c33b37-3b2108ea-201

1039819890: SMB: 4: Priv=0xff,0x3 DefOpt=0xe Adm=1, Backup/Adm=1 (Bkp=0) NTCapa=0x2

1039819890: SMB: 4: NT2UNIX: 5 groups

1039819890: SMB: 4: gid=0x186a8 S-1-5-15-3814076-70c33b37-3b2108ea-201

1039819890: SMB: 4: gid=0x80400220 S-1-5-20-220

1039819890: SMB: 4: gid=0x804003e9 S-1-5-15-32434d45-3e71-656745e8-3e9

1039819890: SMB: 4: gid=0x80400221 S-1-5-20-221

1039819890: SMB: 4: gid=0x1 S-1-5-12-2-1

**1039819890: SECURITY: 4: cred at 0x8bc7904: uid=0x3ed, gid=0x64, inuid=1005 [celerraav user's UID]**

1039819890: SECURITY: 4: 5 other gids:

0x186a8 0x80400220 0x804003e9 0x80400221 0x1

### **VIEWING SID CACHING STATISTICS ON DM:**

**\$ .server\_config server\_4 -v "sidcache global status"**

Sid cache status

Enable:1

DefaultSize:53

Hits:723851

Miss:31759

Count:1897 Collisions:-9644

Total of collisions:20213 Total of hash OK:11541 Compare:1238587

**Comment:** Look for high miss rate to hit rate ratio

### **Data Mover Kerberos Services:**

**\$ server\_kerberos server\_2 -l**

Kerberos realms configuration:

realm name: NETWORK.LAN

KDC: util002.network.lan

admin server: util002.network.lan

default domain: network.lan

kpasswd server: util002.network.lan

End of Kerberos realms configuration.

**\$ server\_kerberos server\_2 -keytab**

Dumping keytab file

keytab file major version = 5, minor version 2

realm: NETWORK.LAN, principal: host, host: FILE014,

principal type 1, key version: 1, encryption type 3

realm: NETWORK.LAN, principal: host, host: DM2.NETWORK.LAN,

principal type 3, key version: 1, encryption type 3

realm: NETWORK.LAN, principal: cifs, host: FILE014,

principal type 1, key version: 1, encryption type 3

realm: NETWORK.LAN, principal: cifs, host: DM2.NETWORK.LAN,

principal type 3, key version: 1, encryption type 3

realm: NETWORK.LAN, principal: user, host: dm2,

principal type 1, key version: 1, encryption type 3

End of keytab entries.

**Note:** A Keytab is a Key Table file that contains “keys” that can be used by a Host or Service

**\$ server\_kerberos server\_2 -ccache**

**Note:** Displays credential cache on Data Mover

### # .server\_config server\_2 -v "kerberos listrealms"

1158590830: KERBEROS: 4: Kerberos realm configuration:

1158590830: KERBEROS: 4:

1158590830: KERBEROS: 4: realm name: 2K3.PVT.DNS

1158590830: KERBEROS: 4: kdc: george.2k3.pvt.dns

1158590830: KERBEROS: 4: admin server: george.2k3.pvt.dns

1158590830: KERBEROS: 4: kpasswd server: george.2k3.pvt.dns

1158590830: KERBEROS: 4: default domain: 2k3.pvt.dns

### # .server\_config server\_2 -v "kerberos"

1158591016: KERBEROS: 4: usage:

kerberos <listrealms | addrealm | delrealm | keytab> listrealms

addrealm realm=<realm name> kdc=<fq kdc name>[:port][kdc=...] [kadmin=<kadmin fqdn>] [kpasswd=<kpasswd fqdn>]  
[domain=<dns domain>] defaultrealm delrealm realm=<realm name> addrel realm=<realm name> domain=<dns domain>  
delrel domain=<dns domain> [realm=<realm name>] keytab

## **VIEWING USER INFO & GROUP MEMBERSHIP ON NT 4.0/WIN2K DOMAIN:**

**C:>net user dparker /DOMAIN**

## **HOW TO VIEW GROUP MEMBERSHIP NT 4.0/WIN2K:**

**C:>net group Everyone /DOMAIN**

## **OTHER CIFS TOPICS:**

### **Creating Trust Relationship between NT 4.0 Domains (1-way trust):**

Example: Netinfo (“trusting”) {aka, Resource...}domain -----→trusts Netcentric (“trusted”)

Step (1): Netcentric adds Netinfo in UserMgrDomains to the “Trusting Domain” box and assigns Password

Step (2): Netinfo then adds Netcentric to its “Trusted Domain” box using the provided password to establish the trust.

**Accessing PDC from NT Machine:** run: \\superman\c\$\winnnt\system32\srvmgr {usrmgr,winsadm,cmd,etc}

## **USING USER MANAGER INTERFACE ON WIN2K STYLE CIFS SERVER:**

### **# server\_param server\_3 -facility cifs -info majorVersion**

server\_3 :

name = majorVersion

facility\_name = cifs

default\_value = 0

current\_value = 0

configured\_value =

user\_action = restart Service

change\_effective = restart Service

range = (0,4294967295)

**Note:** Set value to 4 on Celerra to allow usrmgr interface to manage CIFS server

## **MULTIPLE NT DOMAINS:**

Each DataMover Interface can be a member of different NT Domains [Create interfaces & CIFS Services]

Use VLAN IDs to provide security & separation between multiple DOMAINS defined on a single Data Mover [VLAN Tagging]

**NT 4.0 & Celerra:** Celerra uses NTLM 0.12, or NT1, a dialect of CIFS using SMB’s [MD4 encryption]. Celerra supports NTLMv2 with NAS 5.4.18/5.3.16 and higher

## **LOTUS DOMINO/NOTES R5.0.3:**

**Intro:** Domino & Notes applications are sensitive to network disruptions, with CIFS outages of 10 seconds and NFS outages of 30 seconds causing potential database corruption. However, automatic recovery can be done with transaction logging and restarting of the Lotus database.

### **Lotus R5.0 Support in NAS Environments:**

1. Enable Transaction Logging for databases
2. Turn off CIFS Oplocks [Use nooplocks for mounts]
3. Use Lotus Administrative Client to remotely administer Celerra resource, as we do not support ‘null sessions’ local db account
4. Stop Domino Server before making snapshot or checkpoint

## **SQL SERVER SUPPORT:**

Microsoft generally recommends only running SQL databases over SAN or direct-attached storage. If you have to use NAS solution, however, Microsoft requires specific “write ordering” and “write-through” guarantees. By default, you cannot create SQL database link over the network—a special setup is required using the ‘Trace flag 1807’ option. Key I/O considerations for SQL are I/O Aggregated Bandwidth (measured in MBps); I/O Latency (measured in ms); and CPU Cost (Host cpu in microseconds/operation).  
--CIFS oplocks should be disabled—that is, use “nooplocks” to help prevent corruption if using SQL databases  
--SQL 7.0 is intolerant to network interruptions and a file system unavailability > than 10 secs. will cause database corruption  
--SQL 2000 is more tolerant to network disruptions—up to 45 secs, but “nooplocks” rule still applies

## **RATIONAL CLEARCASE SUPPORT: 4.1**

Supported only via links from ClearCase Server’s storage for Unix/NFS & NT/CIFS—Celerra hosts the VOB (Versioned Object Database) pools [Special setup & Clearcase User account required]

## **IIS 4.0 (NT 4.0, IIS 5.0 (Windows 2000) IIS 6.0 (Windows 2003) SUPPORT:**

**Intro:** There are unique challenges presented when running an IIS Server with Celerra. Because Internet Information Servers rely on “content caching” of web pages to work properly, it requires that “NAS file servers” [Celerra, NetApps] notify IIS whenever Files or Directory content changes—hence the terms “File Change Notification”, “Directory Watch”, “Notify”, “WatchTree Bit”, “WatchSubTree”. IIS 4.0 can only cache ASP files—IIS 5.0 caches both ASP & HTML files. In older versions of NAS that did not provide for a “notify” function and “WatchSubTree” bit, the only way that IIS could update its cache would be to stop & start the IIS Service. IIS 6.0 introduces file caching from remote NAS servers using last-modified times (mtime) for files, as well as the traditional file state change notification feature.

### **IIS 6.0 & Windows 2003 Constrained Delegation:**

IIS 6.0 also supports a concept for 2003 native domains called “Constrained Delegation” which allows a Domain Administrator to delegate specific services on a target NAS Server, for access by individual Users connecting through an IIS 6.0 Server. Essentially, Constrained Delegation (aka, Service4User2Proxy) allows an IIS server to pass User credentials (e.g., Kerberos) to the remote NAS Server (pass-through authentication) so that NTFS permissions can be used to control file access, as opposed to the traditional method of allowing a “designated” user access via IIS (can configure this for NTLM, Basic Digest, Kerberos, etc.). The IIS Server is also responsible for controlling access to a specific list of Services on the Remote NAS Servers. Remote Servers authorize access based on the delegated Service Tickets sent by IIS and grants access based on ACLs. Kerberos Integrated Windows Authentication, is an example of a protocol that supports delegation, allowing for application of NTFS permissions through the use of Domain Users or Domain Groups.

Constrained Delegation is a Windows 2003 capability that allows Active Directory to delegate, via Administrative configuration, certain Services for access from an end User via an intermediary Server, to a remote Server (i.e., NAS) using Kerberos Pass-through Authentication. In the case of NAS and IIS 6.0, there is an MS Kerberos Extension called "S4U2Proxy" that allows a Web Server to present an end User's Service Ticket for the connection to the Web Server, to the KDC, in order to obtain access to a remote NAS Server. Essentially, the KDC returns the TicketGrantingService ticket to the Web Server, which contains the end User's authentication data. The Web Server then presents the Service Ticket to the remote NAS Server on behalf of the end User. One of the key reasons for using 'Constrained Delegation' is that this allows Customers to secure NTFS permissions on remote NAS devices using Domain-based Users and Groups, as opposed to traditional IIS designation of a special User account with unrestricted UNC access. See emc142136 & emc132863 for more information.

### **Constrained Delegation has been tested and approved, though with certain caveats:**

--Use only Compname and not IP Address when configuring Celerra Virtual Web Directory (Kerberos requires Compname)  
--Only the Shortname & FQDN for the CIFS & HOST Services should appear in the Delegation properties tab for the IIS Server—if there are extra instances present, delete them

## **CONFIGURING W2K3 DELEGATION:**

ADUC>rightclick Domain>Raise Domain Functional Level (only W2k3 and is irreversible)

ADUC>Computers>IIS Server properties>Delegation tab>select ‘Trust this computer for delegation to specified Services only’ & add other authentication schemes such as Kerberos.

Click Add>Add Services>Select Users or Computers>enter NAS Server

Select services for the NAS Server that are registered with SPN's, select Host & CIFS for Service Type list

## **CONSTRAINED DELEGATION CONFIGURATION PROCEDURE FOR CELERRA WITH IIS 6.0:**

1. Configured NAS 5.5\* CIFS Server running with exported Shares
2. Windows 2003 Domain must be running in native mode

3. Configure Virtual NAS Directory to Celerra using IIS Manager:

a.) Create new Virtual Directory or use existing Home Directory tab on Default Web Site Properties tab

Default Web Site>new>Virtual Directory

b.) Enter alias for the UNC location

c.) Enter UNC path for the remote NAS Share path

**Note:** Do not use IP address here--must specify the Compname of the Data Mover since Kerberos is required

d.) Keep the setting, 'Always use the authenticated user's credential when validating access to the network directory'

e.) Set the appropriate Virtual Directory access permissions to be allowed

4. Set Directory Security on Virtual NAS Directory using IIS Manager:

a.) Highlight Virtual NAS Directory>Directory Security>Authentication and access control>Edit>Uncheck 'Enable anonymous access' and ensure that 'Integrated Windows authentication' is checked

5. Configure Constrained Delegation using Active Directory for Computers (ADUC) MMC Interface:

a.) Open ADUC>highlight IIS Server>properties>Delegation tab>select 'Trust this computer for delegation to specified services only' and 'Use Kerberos only'>Click on 'Expanded' box and add CIFS & HOST services of the Data Mover using 'Add' button (if services not already listed)

**Note:** One known issue is that multiple entries may appear in the Delegated Services box--remove any duplicated entries.

**Examples of Correct Service Entries for NAS--Shortname & Fully Qualified Domain Names required:**

Service Type User or Computer

cifs dm2

cifs dm2.pvt.dns

host dm2

host dm2.pvt.dns

## **IIS 6.0 TUNING PARAMETERS:**

|                                                                             | Default Values | Max Values |
|-----------------------------------------------------------------------------|----------------|------------|
| HKLM>System>CurrentControlSet>Services>LanmanWorkstation>Parameters>MaxCmds | 50             | 65535      |
| HKLM>System>CurrentControlSet>Services>LanmanServer>Parameters>MaxMpxCt     | 50             | 65535      |
| HKLM>System>CurrentControlSet>Services>LanmanServer>Parameters>MaxWorkItems | 4096           | 65535      |

## **CELLERRA IIS SUPPORT:**

Only NT Security Mode is supported (security=NT)

**Note:** Not supported with Unix or Share Security modes

## **SETTING UP CELERRA WITH IIS:**

IIS User account must be defined for IIS and also the Passwd/Group file for the DataMover (or Usrmapper database)—not the IUSR Anonymous User account created by default for IIS Servers. Map Celerra Shares for IIS using 1.) Network Drive 2.) UNC Path

## **FILE CHANGE NOTIFICATION COMPONENT: FileChangeNotify**

File Change Notification is implemented by Windows Applications using SMB that allows for Applications to setup a “watch” on single or multiple directories [or whole subdirectories] so that any File or Directory changes that are made are then sent back to the IIS Server from the Celerra File Server as a ‘file change notification’. Main purpose here is to have Applications “register” their requests for change notification with the File Server so that constant ‘polling’ of the File Server is not required. Changes are usually based on creation or modification times.

### **NAS 4.2 FCN EXAMPLE:**

IIS Server sends file change notification request for each virtual directory to monitor, to the File Server using SMB sessions. File Server will reply only when a change occurs by using a watch bit subdirectory flag.

‘Changes’ to IIS Web Server content [hosted on Celerra Shares] are buffered so that multiple notifications can be satisfied with a single response. By default, the file system level for ‘notifications’ is 512 subdirectories down from the root. This can be changed by using the Server Mount command to set the “triggerlevel” option:

**\$server\_mount server\_2 -o “triggerlevel=0x000000f” fs01 /fs01**

**Note:** Example sets notification down 15 directories from root

### **DISABLING TRIGGER NOTIFICATION ON FILE SYSTEM:**

**\$server\_mount server\_2 -o “triggerlevel=-1” fs01 /fs01**

**Note:** Negative value disables file change notification feature

### **Directory Notifications OnAccess & OnWrite:**

CFS Windows Environment Guide [NAS 4.2 page 4-9,], suggests that the "notifyonaccess" and "notifyonwrite" values are disabled by default for file systems, meaning that they would have to be setup if you wanted to have notification by access time of a modification and notification of write access to a file system [same for NAS 5.2].

**\$server\_mount server\_x -o notifyonaccess,notifyonwrite fs1 /fs1**

**IIS Limitations:** When using "Share Security" with IIS, you might encounter an "HTTP Error 401" from the Clientside. This is due to 'unauthorized' access on ACL. To correct this condition, convert to "NT Security" on CIFS Server. Celerra also does not recognize the built-in IIS anonymous User account. The workaround here is to specify a regular NT User as the "anonymous user" account.

**WatchSubTree Functionality:** Setting this "bit" allows for applications to notify Celerra of directory changes, such as in IIS Only supported with NT Security, not 'Share-level'

**NAS 2.2.40 & Below:** NO SUPPORT FOR WatchSubTree function

**NAS 2.2.40-46.x:** Limited support for “Directory Change Notification” by using WIN32 API. Monitors (8) directories below root by default & can be specified on File Systems using the ‘mount’ command to monitor down (15) levels: \$server\_mount server\_2 -o "triggerlevel=0x0000000f" fs01 /fs01

**NAS 2.2.47 & Above:** Trigger level extended down to 512 directories [also the default setting]

### **"Notify" WatchSubTree 'bit' Restrictions:**

- NOT SUPPORTED on NFS, FTP, or MPFS [ONLY FOR CIFS]
- SHARE-LEVEL SECURITY MODE NOT SUPPORTED
- ONLY NT "User-level" SECURITY MODE SUPPORTED

### **SETTING TRIGGER LEVELS FOR FILE CHANGE NOTIFICATION:**

\$server\_mount server\_2 -o "triggerlevel=0x0000000f" fs01 /fs01 [Hex value specifying (15) directory levels]

### **DISABLING FILE CHANGE NOTIFICATION:**

\$server\_mount server\_2 -o nonotify fs01 /fs01

### **DISABLING IIS DYNAMIC AND STATIC CACHING:**

**Disabling Dynamic Cache:** Open IIS Manager, rightclick ‘Computername’>Properties>Edit WWW Service Master Properties>Go to ‘Home Directory’>Click on ‘Configuration’>Process Options’ tab and select “Do not cache ASP files”>Apply>o.k., Stop & Restart the IIS Service to put changes into effect.

**Disabling Static Cache:** run: regedt32>HKLM>System>CurrentControlSet>Services>Inetinfo>Parameters>Add following value to registry with value of 1: “DisableMemoryCache”: REG\_DWORD: 1

**Note:** Must reboot IIS Server to make changes effective

### **Disabling Caching on IIS 4.0 Web Servers:**

1. Open Internet Service Manager on IIS Web Server
2. Rightclick Server name>Properties>WWW Service Master properties>Edit>Home Directory>Configuration>Process Options: “Do no cache ASP files”
3. Apply change
4. Stop & Restart IIS Service
5. Make following Registry Change on IIS Server: HKLM>System>CurrentControlSet>Services>Inetinfo>Parameters>Add following value: “DisableMemoryCache” with value of REG\_DWORD:1
6. Reboot IIS Server

### **SOME REASONS FOR FILE CHANGE NOTIFICATION TO FAIL:**

- Loss of network connectivity
- lack of permissions
- invalid user accounts
- MaxCmds limit has been exceeded on Client Redirector
- MaxWorkItems context has been exceeded on the Server
- File Server not returning notifications completely
- File Server not returning any notifications at all

### **IMPLEMENTING NONOTIFY FOR ROOT FILESYSTEM ON DM:**

#### **root\_fs\_2 on / ufs,perm,rw,nonotify**

**Purpose:** Also called the NetShareEnum issue in that Windows 9x clients can only browse 64k in buffer size for Shares. Problem can occur when a Client browses the root filesystem of the data mover, placing a ‘notify’ request for any changes to the .etc directory. This in turn causes poor performance on the data mover and locks up browse and notify threads. Symptoms are poor browse performance.

**Fix:** Apply the “nonotify” option to the root filesystem of the DM

1. #vi /nas/server/slot\_x/mount
- ufs rw / 48=2 rw,nonotify
2. Reboot server

**Note:** This fix should not be applied with Engineering/GNS diagnosis and recommendation

### **COMMON CIFS PORTS & PROTOCOLS:**

#### **PORT USAGE**

| PORT | USE      | Transport               |
|------|----------|-------------------------|
| 53   | DNS      | TCP [DNS over TCP]      |
| 88   | Kerberos | UDP [Kerberos over UDP] |

|     |               |     |                                                                                                     |
|-----|---------------|-----|-----------------------------------------------------------------------------------------------------|
| 138 | Netbios-DGM   | UDP | [Netbios datagram service using SMB over Netbios over UDP-broadcasting [SMB Browser Svc]            |
| 137 | Netbios-NS    | UDP | [NetBIOS Name Service, broadcast or directed IP—NBT WINS over UDP]                                  |
| 139 | MicroSoft-SSN | TCP | [Netbios sessions using SMB over Netbios over TCP [typical Win9x, NT client authentication/traffic] |
| 389 | LDAP          | UDP | [Default LDAP Port over UDP or TCP]                                                                 |
| 389 | LDAP          | TCP | “ “                                                                                                 |
| 445 | Microsoft-DS  | TCP | [SMB encapsulated on top of TCP [Win2k client port and Win2k forwarding Directory Services]]        |
| 464 | DNS           | UDP | [DNS over UDP]                                                                                      |
| 135 | Endmapper Svc | RPC | Similar to Unix Portmapper service, maps services to ports                                          |

## **WINDOWS 2000 & CELERRA SUPPORT:**

### **MIXED MODE SUPPORT:**

**Intro:** By definition, "Mixed Mode" means that there is still a legacy NT 4.0 Domain Controller on the network that can replicate with the Active Directory Mixed Mode Server. However, all that is required for a CIFS NT 4.0 Server to be established on a Windows 2000 domain is the running of the PDC Emulator Service, which runs by default on the first Active Directory Server installed. There is one PDC Emulator for every Domain in the Forest. PDC Emulator supports NT Clients, Member Servers, Domain Controllers, Windows 9x Clients via the following:

- Windows 9x & NT 4.0 Clients are supported for Password Changes & Checking & AD Writes via the PDC Emulator
- NT 4.0 BDC's replicate from PDC Emulator
- PDC Emulator servers as NT 4.0 Domain Master Browser (0x1B)

**Note:** With the installation of the ADClient package, Windows NT & Windows 9x Clients can write to AD to any AD Server

## **CIFS & WINDOWS 2000 ACTIVE DIRECTORY SUPPORT:**

**Intro:** First of all, it is important to note that NAS 4.0 and above will support CIFS as either a legacy "NT 4.0-style" Server or as a native "Windows 2000-style" Server without regard to "Mixed v. Native Mode". The common misconception is to refer to this behavior as "Mixed-Mode" v. "Native-Mode". Strictly speaking, "Mixed-Mode" means only that there are still NT 4.0 Domain Controllers operating within the Windows 2000 Domain. "Native-Mode" means that there are only Windows 2000 Active Directory Servers in the Windows 2000 Domain space.

### **CIFS 2000 SERVER SUPPORT:**

- Supports Kerberos MD5 (128-bit hash output), NTLM, or NTLMSSP [NTLM Security Support Provider] Domain Authentication Modules
- CIFS Service created from commandline [\$server\_cifs server\_x –add compname=]
- Computer Name is "Joined" to Active Directory Domain from commandline with the use of Dynamic DNS Support as a 'native' Windows 2000 type server
- Uses Dynamic DNS for Name Resolution and Active Directory LDAP queries for locating network resources
- By definition operates only within the Windows 2000 Domain

### **CIFS 4.0 SERVER SUPPORT:**

- Supports only NTLM Version 1.2 Domain Authentication [Which AD Servers support by default for NT 4.0-style Clients]
- CIFS Service created from commandline [\$server\_cifs server\_x –add netbios=]
- Requires manual creation of CIFS Computer Account [i.e., Netbios name] using Server Manager or AD Users & Computers Snap-in & participates as an 'NT-4.0' type server
- Requires use of NetBIOS protocol for WINS Name Resolution and Network Browsing
- Participates in either an NT 4.0 or Windows 2000 Domain environment
- Data Mover broadcasts name and comment field every 15 minutes to Browse Service [initial startup broadcasts 5x in first minute]

## **CELERRA AUTHENTICATION SUPPORT FOR MICROSOFT WINDOWS NT, 2000, 2003:**

### **NTLM    NTLMSSP    NTLMV2**

Typically, NT 4.0 Domains/Clients use NTLM authentication on Port 139 → Celerra Supports NTLM

Typically, Windows 2000 Domains/Clients use NTLMSSP (SPNEGO) authentication on Port 445 → Celerra Supports NTLMSSP

**Note:** SSP is a Security Support Provider, and SSP Negotiate is a special authentication protocol based on the Simple and Protected GSS-API Negotiation Mechanism (SPNEGO). By default, Win2k SSP Negotiate will select Kerberos authentication, but if no NDS is available, or a service SPN name is not registered, will fall back on NTLMv2, NTLMv1, LAN Manager, etc.

If Windows 2003 Domain Security Policy is set for 'NTLMV2' only, Celerra does not support & cannot negotiate until 5.3.18+

**Note:** 5.1.24.1/5.2.12.0/5.3.4.0 allows NTLMSSP even if Port 139 is used by Clients.

## **MIGRATING CELERRA SERVERS TO WINDOWS 2000 ACTIVE DIRECTORY: CONVERTING NT 4.0 CFS DM's TO CIFS 200:**

**Background:** Customer running Usrmapper with “\_SID\_HISTORY” turned on. Both NT 4.0 Domain Groups and Windows 2000 Domain Groups are defined in Usrmapper & the DataMover’s localgroups file

Step 1. Disconnected all Users from CIFS Shares using NT 4.0 Server Manager interface

**Note:** For busy systems, if this first step was not done, CIFS could hang when trying to stop

Step 2. Stopped CIFS Service on DataMover to be converted

Step 3. Deleted WINS entries from CIFS

**\$server\_cifs server\_2 -d wins=172.19.3.254,172.19.3.253**

Step 4. Deleted NETBIOS Name from CIFS:

**\$server\_cifs server\_2 -d netbios=file014**

**Note:** Make sure that the Netbios name is also deleted from the NT 4.0 Domain using Server Manager!

Step 5. Verify netd file: /nas/server/slot\_2/netd [In some cases, discovered that WINS entries were not removed].

**#vi /nas/server/slot\_2/netd**

**cifs add wins=172.19.3.254 wins=172.19.3.253 [Deleted this line]**

Step 6. Point DataMover to NTP Server:

**\$server\_date server\_2 timesvc start ntp 172.24.80.11**

**\$server\_date server\_2 0209041130** [Run to correct Time if not in sync—Year, Mnth, Day, Time]

Step 7. Setup or Add Windows 2000 DNS Domain & Servers—leave entry for old DNS Domain:

**\$server\_dns server\_2 -p tcp network.lan 172.19.10.254,172.19.10.253**

**Note:** You can point to multiple DNS Domains per DM—limited to 3 different domains from GUI, & 3 servers per domain.

Step 8. Enable ASCII Filtering for Non-UTF-8 [Non-I18N] Implementations:

**#vi /nas/server/slot\_2/param**

**param shadow asciifilter=1**

Step 9. Reboot DataMover—always reboot DataMover after “editing” either the “netd” or “param” files!

**Note:** For those Customers that can, turn I18N On!

**\$/nas/sbin/uc\_config -setup**

**\$/nas/sbin/uc\_config -on -mover server\_2**

Step 10. Add Computer Account to Active Directory:

**\$server\_cifs server\_2 -add compname=dm2,domain=network.lan,netbios=file014,interface=fsn0**

Step 11. Join CIFS Server to Active Directory Domain:

**\$server\_cifs server\_2 -Join compname=dm2,domain=network.lan,admin=migratory**

**Password: \*\*\*\*\***

Step 12. Start CIFS Service

**\$server\_setup server\_2 -P cifs -o start**

Step 13. Reboot Datamover—verify CIFS Service & access to Shares

**Comments:** In this example, did not have to edit usrmap.cfg file because the “network,network.lan” entry was already added. Nor did the localgroups.db file require editing or updating because both the NT 4.0 & WIN2K Domain Groups had already been added.

## **MIGRATING NETBIOS SERVER FROM NT 4.0 TO WIN2K—“ALIAS METHOD”:**

**Prerequisites:** NAS 4.2 and higher. Migrating from old to new Server while retaining old Server’s name.

Trust relationship between NT 4.0 & Windows 2000 domains. Sid History configuration on Usrmapper.

1. Setup Windows 2000 CIFS Server with a different netbios name than the NT 4.0 Server to be migrated

2. Migrate localgroups database from NT 4.0 to Windows 2000 CIFS Server using LGDUP

3. Migrate data using EMCopy

4. Migrate Shares using SHAREDUP

5. Delete original Netbios name from NT 4.0 domain

6. Create alias on Win2k server with old netbios name from NT 4.0 domain

## **MIGRATING NETBIOS SERVER FROM NT TO WIN2K WITHOUT ALIAS:**

**Prerequisites:** Migrating from old to new Server, but retaining old server’s name. Ensure that Usrmapper is configured.

1. Setup Windows 2000 CIFS temp Server with a different netbios name than the NT 4.0 Server to be migrated [not as the default]

2. Migrate localgroups database from NT 4.0 to Windows 2000 CIFS Server using LGDUP

3. Migrate data using EMCopy

4. Migrate Shares using SHAREDUP

5. Delete original Netbios name from NT 4.0 domain

6. Create CFS 2000 Service using original netbios name on same Windows 2000 Data Mover

7. Migrate localgroups, data permissions, and Shares from temp Win2k to new service with original netbios name

8. Delete temporary name [i.e., if not the default CIFS service]

## **TROUBLESHOOTING CIFS NT 4.0 & WINDOWS 2000 'JOIN' ISSUES:**

**Note:** Computer Name is created using "-a compname=mickey, domain=t2dom3.local, interface=ana0" command

**AD Account & DNS Updates occur during "-J compname=mickey, domain=t2dom3.local, admin=tmatta" command**

--Are WINS entries properly registered on WINS Server for CIFS netbios name?\* [Workstation & Server Service: 00h/20h]

--Are WINS entries properly registering the NT Domain & Domain Controller that the DataMover uses for authentication?\*

[NT DOMAIN: 00h [Workgroup service]/1Bh [Domain Master Browser]/1Ch [Domain Controller]/1Eh [Normal Group Name]  
[PDC/AD Server: 00h/20h/03h]

--Are Broadcast & Subnet Masks correct for Celerra? [Incorrect entries will not allow for proper network communications]

--Are switches configured to allow for Port 137 [Netbios Name Service], 138 Netbios UDP Service], 139 Netbios TCP Service], 445 [Secure SMB over TCP Service], and 135 [TCP Port Service Locator Service] for Netbios-related services?

--Is Windows 2000 Active Directory Server Security Template set to 'Highly Secure'?\*

[Will prevent DM from communicating because of lack of support for digital encryption communications]

--Does AD Server have '\*Enable netbios over TCP/IP' set?\* [Netbios will not communicate with AD Server]

--Verify PDC Emulator Service is configured by checking 'Operations Master' properties on Active Directory Users & Computers\*

--Does AD Server have multi-homed Network Interface Card?\* [Can cause netbios name resolution to fail]

--Run NBTSTAT -r on AD Servers

--Ensure Reverse & Forward LookUp Zones for DNS are setup properly and using 'Allow Dynamic Updates--Yes'\*\*  
[Cannot properly Join domain without]

--Ensure Reverse Lookup Zone is created for DataMover

--Ensure Time Service setup on DataMover to allow for Kerberos participation\*\* [Without will not properly Join Domain]

--Ensure valid Domain Admin account is used for 'Join' command, as well as a valid password\*\*

\*CIFS NT 4.0 Legacy Server only

\*\*CIFS WIN2K Native-mode Server only

### **KERBEROS ENCRYPTION ERROR:**

...KDC has no support for encryption type

LDAP: 3: LDAP bind: Security negotiation failed

**Note:** KDC fails to issue TGS ticket with DES-cBC-MD5 encryption & DM cannot Join Domain. Run following command to reset:  
c:>nlttest.exe /server:kdc\_computer\_name /SC\_CHANGE\_PWD:DC\_domain\_name

### **JOIN PROBLEM IF DC USES UPNs (Universal Principal Names):**

Join fails with error: 'Message Stream modified' [LDAP fails to bind during Join command]

**Note:** Use LDP to verify UPN name; Use adsedit utility to delete UPN name; reset computer account using nlttest

### **JOIN ISSUE REGARDING KERBEROS UDP PACKET SIZE LIMITATION OF 2K:**

Microsoft limitation for Kerberos over udp is 2k. Often, a Server's Kerberos packet information [User's group information, compnames, IP Addresses] may exceed the 2k size limit and prevent a Join from working for a new compname. NAS 5.2 supports Kerberos over TCP, which will resolve this issue.

### **KERBEROS OVER UDP 2K SIZE LIMITATION FOR JOIN COMMAND:**

Prior to NAS 5.2, we used UDP protocol for transmitting Kerberos data to DC's. If a User is a member of many groups, or has a large number of interfaces already defined, a "Join" for a new compname might fail with following Server Log message due to Kerberos Ticket Size:

**2004-04-02 09:40:50: KERBEROS: 3: DomainJoin::getAdminCreds: gss\_acquire\_cred\_ext failed: Miscellaneous failure  
KRB5 error code 52**

### **POSSIBLE WORKAROUNDS:**

→Create new Admin User and make member of only Domain Admins so as to have fewest possible Group memberships

→If Server contains large number of defined interfaces, 'down' all other interfaces except for Compname being Joined

→Upgrade to NAS 5.2 so as to take advantage of Kerberos Over TCP, which removes Kerberos ticket size limitation

### **JOIN FAILS WHEN STRONG LDAP AUTHENTICATION IS SET ON AD SERVERS:**

**Note:** NAS 5.5 fully supports LDAP Signing and LDAP Privacy

### **SERVER LOG LDAP ERRORS IN DEBUG MODE:**

2004-05-26 16:14:57: ADMIN: 3: Command failed: domjoin compname=ph11s domain=my.domain.com admin=paasol password=7\_init

2004-05-26 16:15:03: SMB: 7: DomainJoin::execute: domjoin invoked with ' compname=ph11s domain=my.domain.com admin=paasol password=7\_init '

2004-05-26 16:15:03: SMB: 7: DomainJoin::findServer: try to connect ph11s101.my.domain.com@131.82.241.1 port 389.

2004-05-26 16:15:03: SMB: 7: DomainJoin::execute: Server NB name PHL1S

2004-05-26 16:15:03: SMB: 7: DomainJoin::initPrincs: admin princ=paasol@MY.DOMAIN.COM

2004-05-26 16:15:03: SMB: 7: DomainJoin::initCifsSrvPrincs.

2004-05-26 16:15:03: SMB: 7: DomainJoin::verifyDomain: Connected to LDAP server'ph11s101.my.domain.com' (@131.82.241.1), port 389

2004-05-26 16:15:03: SMB: 7: DomainJoin::verifyDomain: List of entries and attributes:

2004-05-26 16:15:03: SMB: 7: Attributes (and values):

2004-05-26 16:15:03: SMB: 7: defaultNamingContext

2004-05-26 16:15:03: SMB: 7: ('DC=my,DC=DOMAIN,DC=COM')  
2004-05-26 16:15:03: SMB: 7: supportedSASLMechanisms  
2004-05-26 16:15:03: SMB: 7: ('GSSAPI')  
2004-05-26 16:15:03: SMB: 7: ('GSS-SPNEGO')  
2004-05-26 16:15:03: SMB: 7: ('EXTERNAL')  
2004-05-26 16:15:03: SMB: 7: ('DIGEST-MD5')  
2004-05-26 16:15:03: SMB: 7: DomainJoin::connect: attempting to connect to server 'phl1s101.my.domain.com' (@131.82.241.1) via SASL;port=389, admin DN='cn=paasol,cn=Users,dc=my,dc=domain,dc=com'  
2004-05-26 16:15:03: SMB: 3: DomainJoin::connect: Unable to connect to the LDAP service on phl1s101.my.domain.com (@131.82.241.1) - result code Strong auth required, error message 00002028: LdapErr: DSID-0C090169, comment: The server requires binds to turn on integrity checking if SSL\TLS are not already active on the connection, data 0, vece.  
2004-05-26 16:15:03: ADMIN: 3: Command failed: domjoin compname=phl1s domain=my.domain.com admin=paasol password=7\_init

## **CAUSE:**

### **LDAP Bind ‘Strong Authentication’ is set on AD Servers:**

**HKLM>System>CurrentControlSet>Services>NTDS>Parameters>LdapServerIntegrity=2** [LDAP Enabled]

**Note:** Customer upgraded to Windows 2003 Domain Controllers and hardened LDAP Bind security settings for Active Directory Services. Celerra does not currently support hardened LDAP Bind Security for either NAS 5.1 or 5.2. With LdapServerSecurity=2 set in the registry of each Domain Controller, LDAP Bind requests to AD will fail for any authentication mode that does not use SASL or TLS/SSL, which is what occurs when trying to conduct a Join. Celerra does not support the use of SASL (Simple Authentication and Security Layer Protocol) or TSL/SSL (Transport Layer Security). The only current solution for this problem is to disable 'Strong LDAP Authentication' on all Domain Controllers in which the Celerra CIFS Servers are a member of--emc87712.

**Note:** TLS is SSL version 3.1.

### **DISABLING STRONG LDAP AUTHENTICATION ON AD SERVERS:**

**HKLM>System>CurrentControlSet>Services>NTDS>Parameters>LdapServerIntegrity=1** [LDAP disabled]

**LDAP BIND SETTINGS:** Values are as follows for LDAP server (ldapagnt.lib) handling of LDAP bind command requests

**LdapServerIntegrity=1** Default or not defined—LDAP Server supports client request for LDAP traffic signing when handling a LDAP bind command request which specifies a SASL authentication mechanism.

**LdapServerIntegrity=2** LDAP Server requires LDAP traffic signing, unless LDAP bind request is already protected with TLS/SSL. It rejects the LDAP bind command request if other types of authentication are used.

## **HOW TO CREATE CORRECT COMPUTER ACCOUNT AND JOIN ACTIVE DIRECTORY IF DNS NAME AND WINDOWS 2000 DOMAIN NAMES ARE DIFFERENT:**

**Intro:** By default, when creating Windows 2000 Domain, the DNS Domain is the same. However, this does not have to be the case and there are legitimate times when a Customer will have a different DNS FQDN than WIN2K FDQN. If this is the case, in order to properly Join AD, use following (2) Steps and Syntax:

### **Step 1.**

**\$server\_cifs server\_2 -add compname=fei0018s009, domain=ds.ferg.com, interface=fsn2, dns=hq.ferg.com -comment "dm2"**

### **Step 2.**

**\$server\_cifs server\_2 -Join compname=fei0018s009, domain=ds.ferg.com, admin=backup**

### **CIFS OUTPUT EXAMPLE OF DIFFERENCES:**

CIFS Server (Default) FEI0018S009[FEI-NT]

Full computer name=**fei0018s009.ds.ferg.com** realm=DS.FERG.COM [FQDN with AD]

Comment='dm2'

if=fsn2 l=172.16.104.41 b=172.16.104.255 mac=0:60:cf:20:a6:9e

FQDN=**fei0018s009.hq.ferg.com** (Updated to DNS) [FQDN with DNS]

### **Error Message for Unsuccessful Join:**

If above syntax is not used, Server CIFS output will show a “DNS Opcode Failure/Error” message

## **CIFS AND WINDOWS 2000’S REVERSE LOOKUP ZONE FOR DNS:**

NAS 4.0 & 4.1 Releases depend on an MIT implementation of Kerberos, which relies on determining the real FQDN of hosts for Kerberos services based on using the IP Address of the Host to perform a Reverse Zone Lookup [request for PTR record]. In actuality, the Join process may work without this Zone being configured, as the DNS query will generally still provide the correct FQDN—In short, best practices require that this Zone be setup.

**Change Update: NAS 4.2.6.x + no longer requires that Celerra use the Reverse Lookup Zone “PTR” Record. Only the “A” Record for Hostname & IP in the Forward Lookup Zone is required.**

**Caution:** If both Zones are configured, however, the Join process will not work if the two Zones contain conflicting records for the Host as a Kerberos Ticket will not be issued to the DataMover.

### **Disabling Dynamic Reverse Lookups for Celerra:**

Set the following parameter to disable Reverse Zone Lookups if using “Static DNS” records:

/nas/server/slot\_x/param or /nas/site/slot\_param

**param dns updatePTRrecord=0 | 1=enable** [enables reverse zone lookups]

**Note:** NAS 4.2 & higher are no longer dependent on Reverse Lookups as DART implements DNS the same as native Win2k Client.

### **REVERSE LOOKUP ZONE PTR RECORD RULE:**

--Reverse Lookup FQDN & IP Address must be the same as in Forward Lookup Zone

--Ensure that multiple/old entries are removed if using Static DNS entries!

--Remove DNS ‘suffix’ option from AD Server—allows interface to register real DNS name

--Use alternate DNS suffix for DM Interfaces by using “dns=” option in server\_cifs [Use this if WIN2K name is different than DNS!]

**Change Update: NAS 4.2.6.x no longer requires that Celerra use the Reverse Lookup Zone “PTR” Record. Only the “A” Record for Hostname & IP in the Forward Lookup Zone is required.**

### **Symptoms:**

Users cannot access Shares & DC List changes to “\*SMBSERVER”

**Server cifs Output:** “Update of “PTR” record failed during update”

**Server Log:** DNS update ERROR: Opcode not implemented in name server while registering..; Primary name server not found...”

## **NORMAL SERVER LOG/SCREEN MESSAGES FOR SUCCESSFUL MACHINE**

### **ACCOUNT CREATION AND JOIN COMMAND:**

#### **SERVER LOG RESULTS WHEN ADDING NEW COMPUTER NAME FROM COMMANDLINE:**

\$server\_cifs server\_5 -add compname=dopey, domain=t2dom2.local, interface=ana0

2002-03-14 14:34:34: SMB: 4: CIFS Server DOPEY[T2DOM2] created (0)

2002-03-14 14:34:34: SMB: 4: Full computer name todd.t2dom2.local, Realm T2DOM2.LOCAL

2002-03-14 14:34:34: SMB: 4: Interface ana0 added into CIFS Server DOPEY[T2DOM2]

2002-03-14 14:34:34: SMB: 4: Interface ana0 added into CIFS Server DOPEY[T2DOM2]

2002-03-14 14:34:34: ADMIN: 4: Command succeeded: cifs add compname=DOPEY domain=T2DOM2.LOCAL interface=ana0

### **JOINING WINDOWS 2000 AD/DNS FROM COMMANDLINE:**

# server\_cifs server\_5 -Join compname=dopey, domain=t2dom2.local,

admin=tmatta

server\_5 : Enter Password:\*\*\*\*

done

### **Corresponding Log Entry:**

2002-03-14 14:41:25: SMB: 4: Dart account todd not registered with DNS.

2002-03-14 14:41:25: ADMIN: 4: Command succeeded: domjoin compname=dopey domain=t2dom2.local admin=tmatta password=7?!? init

2002-03-14 14:41:25: LIB: 3: dns\_updateMessage: unable to get hostname of DNS server : Domain not exists

2002-03-14 14:41:25: LIB: 3: DNS update ERROR: Domain not exists while registering dopey.t2dom2.local

**Comment:** DNS Update Error is normal until CIFS service is started!!

## **USING NON-ADMIN USER ACCOUNT TO JOIN COMPUTERS (OR DMs) TO DOMAIN:**

### **Step 1. Create User account to be used for this process:** [User account called “Join”]

a.) As Administrator, open Active Directory Users and Computers & highlight “EMC Celerra” container

**Optionally:** Use the default “Computers” container

b.) Rightclick “EMC Celerra” container and select “Delegate Control”

c.) Select “Next”, then choose “Add” under “Selected users and groups”

d.) Add User called “Join” (A User account used for this test) and click on “o.k.”

e.) Select “Next” and then: \*Create a custom task to delegate

f.) Keep radial button selection: “Delegate Control of: \*This folder, existing objects in this folder, and creation of new objects in this folder

g.) Select “Next” and then check all three of the following boxes:

\_x\_General; \_x\_Property Specific; \_x\_Creation/deletion of specific child objects

h.) On same window, select “\_x\_Full Control” under section for “Permissions:”

i.) Select “Next” and “Finish”

**Note:** The “Computers” container underneath “EMC Celerra” will inherit the rights delegated above to the user “Join”

### **Step 2. Force AD replication update—Test user account “Join” by creating Computer account in ADUC MMC:**

### **Step 3. Create CIFS service and add compname from CLI using “Join” user account:**

a.) #server\_cifs server\_3 -add compname=phoenix, domain=t2dom3.local, interface=ana1

b.) #server\_cifs server\_3 -Join compname=phoenix, domain=t2dom3.local, admin=join

**Step 4.) Verify Status of new Compname:**

- a.) Check Server Log
- b.) Check CIFS output
- c.) Map drive to new compname “Shares” and access data
- d.) Highlight compname in Active Directory Users and Computers interface and select "manage"

**Note:** By default, only Domain Admins can create new compnames or join computers to AD Domains

**SYSTEM/MACHINE-ACCOUNT PASSWORD:**

Native Win2k changes machine password every (7) days. Celerra passwd set to infinity by default with NAS versions prior to 5.2 (NAS 5.2 defaults to weekly password changes). NAS 5.5 has re-implemented passwd change feature and a passwd history cache so that the Data Mover can decrypt Kerberos tickets that may still be sent to it based on the previous passwd.

**param cifs srvpwd.updtMinutes=10020** [Sets param from default to 6 days, 23 hours]

**Note:** Use above param to set different password interval change than the default for DM's in Windows Domain

**PASSWORD CHANGES AND VIRUSCHECKER PROBLEM:**

Issue in NAS 5.2 where password change would cause VC to stop working and create a CIFS outage. See AR45264.

**param cifs srvpwd.updtMinutes=0** & stopping/restarting CIFS will disable Server Password change as workaround.

**CFS 2000 SERVER: COMMON ERRORS WHEN JOINING WINDOWS 2000 DOMAIN CLOCK TIME TOO GREAT BETWEEN AD SERVER & CFS SERVER:**

Kerberos Security Mechanisms will fail if clock skew is greater than 5 minutes

**COMPUTER NAME CREATION COMMAND FAILS:**

2002-04-08 10:50:48: ADMIN: 3: Command failed: domjoin compname=monday domain=t2dom2.local admin=tmatta password=##=?=7%- init

2002-04-08 10:50:55: SMB: 4: Dart account monday not registered with DNS.

2002-04-08 10:50:55: KERBEROS: 3: krb5\_sendto\_kdc: unable to send message to any KDC in realm T2DOM2.LOCAL.

2002-04-08 10:50:55: KERBEROS: 4: krb5\_gss\_release\_cred: kg\_delete\_cred\_id failed: major GSS\_S\_CALL\_BAD\_STRUCTURE or GSS\_S\_NO\_CRED, minor G\_VALIDATE\_FAILED

2002-04-08 10:50:55: KERBEROS: 3: DomainJoin::getAdminCreds: gss\_acquire\_cred\_ext failed: Miscellaneous failure

Preauthentication failed

**JOIN COMMAND FAILS DUE TO TIME SERVICE ISSUE:**

2002-04-08 10:50:55: ADMIN: 3: Command failed: domjoin compname=monday domain=t2dom2.local admin=tmattha password=Å1!3575 init

2002-04-08 10:51:05: SMB: 4: Dart account monday not registered with DNS.

1969-12-31 19:01:05: KERBEROS: 3: DomainJoin::getAdminCreds: gss\_acquire\_cred\_ext failed: Miscellaneous failure  
Clock skew too great

**RESOLUTION:** Sync datamover to NTP server [troubleshoot & check Time Service, etc]

**KERBEROS DECRYPT INTEGRITY CHECK FAILURE:**

--Following error can result from a TIME SYNCHRONIZATION problem between DM & AD Server, resulting in an expired KERBEROS Authentication Key, or it could indicate a COMPNAME PASSWORD DESYNCHRONIZATION ERROR—either situation would prevent communicating or integration into Active Directory & loss of CIFS Access or failure to “Join” Domain:

**SERVER LOG:**

2002-04-08 11:01:32: SMB: 3: SSXAK=c000006d origin=600 stat=d0000,-1765328351

2002-04-08 11:01:32: SMB: 3: SSXAK=c000006d origin=600 stat=d0000,-1765328351

2002-04-08 11:01:32: SMB: 3: SSXAK=c000006d origin=600 stat=d0000,-1765328351

**Note:** Entry is populated in Server Log everytime a User tries accessing via Explorer or AD Computer Management Interface, etc

**RESOLUTION:**

--Verify and correct Time Desynchronization problem [Verify NTP Service, DM Timesvc, AD Server Timezone to CS0 Timezone]

**REJOINING\RESETTING DATAMOVER SERVER PASSWORD ACCOUNT:**

**\$server\_cifs server\_4 -Join compname=suncor, domain=network.lan, admin=migratory -o  
resetserverpasswd**

**Note:** This option can be used when it becomes clear that the data mover can no longer connect to AD with valid Kerberos credentials. Resets the server's password and encryption keys with the principal. Please note that if using Celerra Manager to perform the Join, it will automatically use the reuse or resetserverpasswd option and complete the Join—NAS 5.4 & 5.5.

**SERVER LOG ENTRIES FOR SERVER THAT CAN LONGER AUTHENTICATE TO DOMAIN:**

2004-09-27 19:33:10: SMB: 3: Srv=OHIONFS3 buildSecureChannel=Authenticate2InvalidReply E=0xc0000022

2004-09-27 19:33:10: SMB: 3: NLogon\_SecureChannel not OK=Authenticate2InvalidReply  
2004-09-27 19:33:10: SMB: 4: DCT=1096313590103 browse unsetDC=198.151.186.53  
2004-09-27 19:33:10: SMB: 4: SessSetupX failed=c0000233  
2004-09-27 19:33:10: SMB: 3: KC\_GetCreds:gss\_acquire\_cred\_ext failed; majStatus=0xd0000, min=-1765328360  
2004-09-27 19:33:10: SMB: 3: DC\_GetBlob OPS.GXS.COM\ohionfs3 OHIONTMGR4HOST=d0000 -1765328360  
2004-09-27 19:33:14: SMB: 7: ActiveDirectoryServer::connectToDc: trying connection to ohiontmgr3.ops.gxs.com  
(198.151.186.52), port 389  
2004-09-27 19:33:14: LDAP: 3: LDAP authentication: Unable to acquire credentials for principal: OHIONFS3\$@OPS.GXS.COM. -  
GSS-API major error: Miscellaneous failure  
2004-09-27 19:33:14: LDAP: 3: LDAP authentication: Unable to acquire credentials for principal: OHIONFS3\$@OPS.GXS.COM. -  
GSS-API minor error: Key table entry not found  
2004-09-27 19:33:14: LDAP: 3: LDAP bind: Security negotiation failed.  
2004-09-27 19:33:14: SMB: 4: Unable to connect to Active Directory server ohiontmgr3.ops.gxs.com (198.151.186.52), port 389

### **PERFORMING REJOINRESETSERVERPASSWD:**

#### **cmd log:**

**2004-09-27 20:09:09.914 server\_2:201:27833:S: server\_cifs server\_2 -Join compname=ohionfs3  
domain=ops.gxs.com admin=root resetserverpasswd -o resetserverpasswd**

#### **SERVER LOG ENTRIES FOR SUCCESSFUL REJOIN(Abbreviated):**

2004-09-27 20:09:09: SMB: 7: DomainJoin::execute: domainjoin invoked with ' com  
pname=ohionfs3 domain=ops.gxs.com admin=root password=99212F2D21231725311B resetserverpasswd '  
2004-09-27 20:09:09: SMB: 7: DomainJoin::execute: Realm OPS.GXS.COM  
2004-09-27 20:09:09: SMB: 7: Attempting connection to Active Directory server ohiontmgr4.ops.gxs.com @ 198.133.251.4 port 389.  
2004-09-27 20:09:09: SMB: 7: DomainJoin::buildHostName: \_hostname = ohionfs3 - \_dnsname = ohionfs3.ops.gxs.com - domain  
DN - admin DN  
2004-09-27 20:09:09: SMB: 7: DomainJoin::execute: Server NetBIOS name OHIONFS3  
2004-09-27 20:09:09: SMB: 7: DomainJoin::initPrincs: \_uprinc = host/ohionfs3.ops.gxs.com@OPS.GXS.COM  
2004-09-27 20:09:09: SMB: 7: DomainJoin::initPrincs: admin princ=root@OPS.GXS.COM  
2004-09-27 20:09:09: SMB: 7: DomainJoin::initCifsSrvPrincs:  
2004-09-27 20:09:09: SMB: 7: DomainJoin::verifyDomain: Connected to LDAP server 'ohiontmgr4.ops.gxs.com'  
(@198.133.251.4), port 389  
2004-09-27 20:09:10: SMB: 7: DomainJoin::search: acctName = OHIONFS3  
2004-09-27 20:09:10: SMB: 7: DomainJoin::search: filterStr = sAMAccountName=OHIONFS3\$, workbuf = OHIONFS3  
2004-09-27 20:09:10: SMB: 7: DomainJoin::search: Performing ldap search for account; dn = 'DC=ops,DC=gxs,DC=com', filter  
= 'sAMAccountName=OHIONFS3'\$  
2004-09-27 20:09:10: SMB: 7: DomainJoin::search: Found account 'OHIONFS3'.  
2004-09-27 20:09:10: SMB: 7: updateAccountToRc4: 'userAccountControl' present value = 69632 (0x11000)  
2004-09-27 20:09:10: SMB: 7: DomainJoin::DJ\_setServerPassword: removing key from keytab for service principal  
cifs/ohionfs3.ops.gxs.com@OPS.GXS.COM  
2004-09-27 20:09:10: SMB: 7: DomainJoin::DJ\_setServerPassword: added key to keytab for service principal  
cifs/ohionfs3.ops.gxs.com@OPS.GXS.COM  
2004-09-27 20:09:10: SMB: 7: DomainJoin::getAccountGuid: sAMAccountName = OHIONFS3\$  
2004-09-27 20:09:10: SMB: 7: DomainJoin::getAccountGuid: Found account 'OHIONFS3\$'.  
2004-09-27 20:09:10: ADMIN: 4: Command succeeded: domjoin compname=ohionfs3 do  
main=ops.gxs.com admin=root password=99212F2D21231725311B resetserverpasswd \*  
4158736a DCba34500c OHIONTMGR4[OPS] 6 ConnectReset  
2004-09-27 20:09:15: SMB: 4: checkDCBlob:TicketFlags doesn't match 17/3  
2004-09-27 20:09:15: SMB: 4: checkDCBlob: accept\_sec\_context stat=0,0 NT=0  
2004-09-27 20:09:15: SMB: 4: >DC\*OHIONTMGR4(198.133.251.4) R=7 T=0 ms S=0,1/-1  
2004-09-27 20:09:15: SMB: 4: sendW2KEnumTrustedDomains for srv=OHIONFS3 on domain 'OPS'  
2004-09-27 20:09:15: SMB: 4: processLogonReply:SidCount=2  
2004-09-27 20:09:15: SMB: 4: authenticate OPS\strakar S=0 UserAuthenticated

### **COMMON JOIN ERRORS FROM SERVER LOG:**

**CommandLine Error:** Failed to complete command

**Client not found in Kerberos Database**→Computer account created in AD but not yet replicated to KDC; Or, could mean that Kerberos Server is different than AD Server that DataMover is communicating with [Use DC= command to JOIN]

**Server not found in Kerberos Database**→Common with 4.0.12.0; Need to force DM to use specific KDC; or, install DNS Service for DataMover

**DomainJoin::connect:** Unable to connect to the LDAP service on win2k.t2dom3.local - result code Sasl protocol violation. → Again, try using DC=FQDN command to Join Domain

**DomainJoin::getAdminCreds:** gss\_acquire\_cred\_ext failed: Miscellaneous failureKDC has no support for encryption type → Reset Administrator's password and try again

**Note:** When AD Server is first setup, it's Admin password is encrypted using DES encryption. The password needs to be changed after Installation so that Kerberos encryption is used.

## **PREVIOUSLY KNOWN JOIN ISSUES:**

- DNS Reverse Lookup Zones & Forward Lookup Zone entries for DataMover FQDN must match or else Join process will fail!
- Once DataMover is “Joined” to Domain, Users will be able to map to Celerra Shares, but Administrator may not be able to “manage” the DM when using the Active Directory Users & Computers MMC Interface
- Windows 2000 Groups may not be mapped in Usrmapper Database. DM will obtain User UID's without issue, but may not serve up Group GID's. Workaround at present is to use Localgroups.db file.
- DNS Domain is usually same as WIN2K domain but can be different! Use “dns=” option in server\_cifs to specify DNS suffix for datamover interfaces
- Known issue where Kerberos ticket is too large due to the number of Groups that the admin user is a member of, causing Join to fail
- Known issue where tpc stack mishandles Kerberos tickets and ignores certain bytes, rejecting the Kerberos ticket and causing Join to fail (turn param tcp doRFC1323=0 to zero to disable)

## **COMMON JOIN FAILURES & CORRECTIVE ACTION REQUIRED:**

### **JOIN FAILURE:**

Invalid User account used to conduct Join

#### **Server Log Error:**

2006-04-03 11:59:53: KERBEROS: 4: Warning: Client principal "schmuck" not found in Kerberos realm "2K3.PVT.DNS" database - KDC 0.0.0.0.

2006-04-03 11:59:53: SMB: 3: DomainJoin::getAdminCreds: gss\_acquire\_cred\_ext failed: Miscellaneous failure. Client not found in Kerberos database.

2006-04-03 11:59:53: ADMIN: 3: Command failed: domjoin compname=bestbuy domain

=2k3.pvt.dns admin=schmuck password=\*\*\*\*\* ou="ou=Computers,ou=EMC Celerra" init

#### **Corrective Action:**

Select a valid account and conduct Join.

### **JOIN FAILURE:**

Correct Admin User account used to conduct Join, but password wrong or entered wrong

#### **Server Log Error:**

2006-04-03 12:12:08: SMB: 3:[VDM\_UPPERCASE] DomainJoin::getAdminCreds: gss\_acquire\_cred\_ext failed: Miscellaneous failure. Preauthentication failed.

2006-04-03 12:12:08: ADMIN: 3:[VDM\_UPPERCASE] Command failed: :4 domjoin compname=bestbuy domain=2k3.pvt.dns admin=tmatta password=\*\*\*\*\* ou="ou=Computers,ou=EMC Celerra" init

#### **Corrective Action:**

Verify password or reset User's password in ADUC and retry Join.

### **JOIN FAILURE:**

Correct Admin User account and password used, but the wrong fully qualified domain name was entered

#### **Server Log Error:**

2006-04-03 12:28:48: SMB: 3: DomainJoin::findServer: Unable to locate a Domain Controller in domain 2k3.svt.dns

2006-04-03 12:28:48: SMB: 3: DomainJoin::findServer: Unable to contact any domain controller in domain 2k3.svt.dns

2006-04-03 12:28:48: ADMIN: 3: Command failed: domjoin compname=bestbuy domain

=2k3.svt.dns admin=tmatta password=\*\*\*\*\* ou="ou=Computers,ou=EMC Celerra" init

#### **Corrective Action:**

Delete CIFS compname for “2k3.svt.dns” using GUI, then attempt new Join using correct FQDN, “2k3.pvt.dns” for the Compname.

### **JOIN FAILURE:**

Compname already exists in the netd file on Data Mover

#### **Server Log Error:**

2006-04-03 12:34:04: SMB: 3: Cifs error: Bad domain name

2006-04-03 12:34:04: ADMIN: 3: Command failed: cifs add compname=BESTBUY domain=2K3.PVT.DNS netbios=bestbuy interface=vdm\_test

#### **Corrective Action:**

Delete compname from CIFS Service and re-attempt Join.

### **JOIN FAILURE:**

DNS Service not running, not configured, or misconfigured on the Data Mover

#### **Server Log Error:**

2006-04-03 12:41:09: SMB: 3: DNS is required for domain w2k.pvt.dns, it is actually not configured or it is not running, please configure DNS

2006-04-03 12:41:11: SMB: 3: DomainJoin::findServer: Unable to locate a Domain Controller in domain 2k3.pvt.dns

2006-04-03 12:41:33: SMB: 3: ActiveDirectoryServer::connectToDc: No Domain Controllers found for domain 2k3.pvt.dns.interface=laip1-2a

2006-04-03 12:41:33: SMB: 4: ActiveDirectoryServer::initiate: Couldn't open anonymous LDAP connection to domain 2k3.pvt.dns.interface=laip1-2a

**Corrective Action:**

Verify DNS Service configuration, correct configuration, start DNS client on Data Mover. Conduct Join.

**JOIN FAILURE:**

NTP Service may not be running Data Mover, or, Service is running but time is still out of sync, wrong NTP Service, etc. Clock skew too great

**Server Log Error:**

2006-04-02 12:49:31: SMB: 3: DomainJoin::getAdminCreds: gss\_acquire\_cred\_ext failed: Miscellaneous failure. Clock skew too great.

2006-04-02 12:49:31: ADMIN: 3: Command failed: domjoin compname=bestbuy domain

=2k3.pvt.dns admin=tmattha password=\*\*\*\*\* ou="ou=Computers,ou=EMC Celerra" init

**Corrective Action:**

Verify NTP service and make sure Data Mover time and AD Time are in sync. Conduct Join.

**JOIN FAILURE:**

A valid, but non-admin User tried to Join compname to domain

**Server Log Error:**

2006-04-03 13:18:21: SMB: 3: DomainJoin::addHostToDomain: AD Account creation for computer account 'bestbuy' failed. LDAP response code = Insufficient access rights, error message = '00000005: SecErr: DSID-03151E04, problem 4003 (INSUFF\_ACCESS\_RIGHTS), data 0'.

2006-04-03 13:18:21: ADMIN: 3: Command failed: domjoin compname=bestbuy domain=2k3.pvt.dns admin=dodo3 password=\*\*\*\*\* ou="ou=Computers,ou=EMC Celerra" init

**Corrective Action:**

Verify that user account is member of Domain Admins (or has correct delegated privileges), and attempt new Join.

**LEGACY NT 4.0 CIFS SERVER PREREQUISITES FOR WINDOWS 2000:**

**I. Requires at a minimum the PDC Emulator Service:**

[Also could include an NT 4.0 Domain Controller, but is not required]:

**Verifying the PDC Emulator Service:**

Programs>Administrative Tools>Active Directory Users & Computers>Rightclick AD Server>Select "Operations Masters">PDC Tab>Operations Master: win2ksrvdom2.t2dom2.local [PDC Emulator Server is listed in this box]

**II. Requires that NetBIOS Communications be Enabled for TCP/IP for Legacy Servers on AD Server:**

Network & Dial-up Connections>Local Area Connection>Properties>Select "Internet Protocol (TCP/IP)"

Properties>Advanced>WINS Tab>\*Enable NetBIOS over TCP/IP [Enabled by default]

**III. Windows 2000 Security Template Cannot Be Set to 'Highly Secure' for Machines Using NTLM:**

Using Windows 2000 "Highly Secure" templates are designed for pure Windows 2000 networks that use digitally signed and encrypted network communications and are not supported by legacy NT 4.0 Servers. The default 'Security Template' for a Windows 2000 Domain is "Basic", but there are also "Secure" and "Compatible" templates as well.

**CIFS & WINDOWS 2000 /WIN2K/ ACTIVE DIRECTORY SUPPORT:**

**NAS CODES SUPPORTING WIN2K:** NAS 2.2.35.4; NAS 2.2.49.0; NAS 4.0.12.1; NAS 5.0, 5.1

**NAS 4.0.12.1:**

When looking for DC, LDAP, or Kerberos Services, will use the first DC in the list as returned from DNS

For multi-Domain environments, requires Enterprise Admins rights—needs to have authority to obtain a User's Group credentials from other Domain Controllers in the Forest

This version often suffered problems in multi-domain environments, especially if there were synchronization issues

**NAS 4.0.16.100:**

Tries to use DC on same subnet first, then first one listed in DNS query

Requires use of Enterprise Admins rights to interoperate in multi-domain environments

**NAS 4.0.18.0/4.1.6.0/4.2.4.1:**

No longer requires Enterprise Admins rights in multi-domain environments

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
Reverse DNS Zone entries no longer required to validate canonical name of KDC upon lookup—Celerra obtains correct information for real host name from forward DNS queries.

Locates services in same manner as native Win2k systems—using “use\_ds” file for local KDC, LDAP services

### **NAS 4.2 & Above:**

Data Mover decodes the PAC [Privilege Attribute Certificate] for ‘joins’ as do Win2k systems

**Best Practices:** Try to have DC on same subnet as any single NetBIOS names on an interface  
Use the SNTP Service on the AD Server for Time Synchronization

## **HOW DATAMOVER DISCOVERS DC’s IN ACTIVE DIRECTORY ENVIRONMENT:**

--DNS Query for \_ldap.tcp.REALM SRV records

--NetBIOS Name Query Broadcast for 0x1C & 0x1B or directed IP to WINS Server

**Note:** The NetBIOS Name Query method is the default method for Windows NT CIFS Servers but is still used with the CFS 2000 implementation. The DC List is recomputed every 15 minutes by default. The “Join” process uses only DNS & will also use the closest DC determined from LDAP-UDP ping.

### **DATA MOVER-to-DC COMMUNICATION:**

--Celerra uses LDAP-UDP queries to all DC’s that are returned from DNS

--In the absence of Site information from AD, Celerra will use the closest DC as determined from the LDAP-UDP ping RTT

--DM will also use closest DC for KDC service, unless the AD Server does not know the “principal” that the DM is trying to service, then the DM will resort to querying other KDC’s in the realm.

**Rule:** Data Mover will always use the closest DC for all LDAP operations!

--Join requests use LDAP over TCP

## **HOW DART DISCOVERS AND SORTS LIST OF AD SERVERS:**

**param ds useDCLdapPing=1** [default DM setting]

**param ds useDCLdapPing=0** [Sorts by order that DNS returns AD List, or as set by DM in directoryservices file]

**Note:** Dart uses Ldap Ping to AD Servers returned from DNS in order to sort the DC’s. Default behavior is to sort based on site information from LDAP queries to AD and on response time to the ping. The LDAP ping is used every 15 minutes to help data movers maintain current DC preference list and whenever CIFS service is started. When CLOSEST\_SITE bit is set, 0x8000 is ADDED to the final preference calculation. The response time is SUBTRACTED from the final preference calculation (So if you have 2 DCs with the same preference number at this point, the response time will be the tie break). See DS Dump example:

1082056734: LIB: 4: IP:19.59.116.46 server:fresca.ford.com

1082056734: LIB: 4: **Preference:0x00057ffc**

**Note:** Data Mover will not be truly “Site” aware until NAS 5.4

## **WINDOWS 2000 ACTIVE DIRECTORY MIGRATION TOOL (ADMT.EXE):**

### **Intro:**

Tool used for migrating User, Group, Computer, Service Accounts, and other Network Resources from one Domain to another. The “admt.exe” executable is installed on the Target Domain’s PDC Emulator. The Target Server must be running Windows 2000 Active Directory in Native Mode. For our purposes, we are mainly concerned with migrating from an NT 4.0 to a Windows 2000 Domain. ADMT is used to migrate “security principals” in the following situations:

- a.) From one Forest to another Forest (Copying or Cloning accounts from Source to Target)
- b.) From one Domain to another within the same Forest (Moving accounts from Source to Target Domain)

### **SID History:**

The ADMT process migrates the User or Group SIDs from the Source Domain and assigns a new SID in the Windows 2000 Domain. The new WIN2K SID contains an attribute called “SIDHistory” that stores the User’s original SID and access rights [Security Principals] to the Source Domain’s resources as long as proper 2-way Trusts have been established. In otherwords, until the original Source Domain is deactivated, a User should be able to log in to either “Source” or “Target” domain after the migration and have equivalent access rights to the original Source Domain resources.

**Note:** NAS 4.2.13.0 and higher now integrates the SID History function correctly into the Usrmapper database

### **Utilities:**

To view or delete SIDHistory entries after a “migration”, use “ldp.exe” from Resource Kit. To view SID identities, use “getsid.exe”. Other useful Utilities include apimon.exe, dumpel.exe, xcacls.exe. The “netdom” utility is useful for managing computer accounts, such as Adding, Removing, or Querying, and also to establish and manage Domain Trusts. It also can be used to verify and reset a computer’s “secure channel” communication to the DC.

## **SECURITY ACCESS TOKENS (SAT):**

The ‘SAT’ serves as the basis for allowing a User ‘Network Access’ to resources. The Access Token itself is constructed each time a User logs into the Domain and consists of a User’s SID & all the Group SIDs to which the User belongs. The basis of “permissions”

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
is determined by the Operating System when comparing a User's SAT to the Network Resource's Security Descriptor, which consists of an Access Control List (ACL) and Access Control Entries.

## --Access Tokens equate to User Authentication to Network Resources based on User's SID, Group SID, and SIDHistory...aka 'access'.

## --Security Descriptors are attached to all Network Resources and equate to Authorization based on ACLs (and ACE entries) and SIDs and Access Rights. Authorization = verification of SAT SID to ACL on SD.

**Note:** Access Tokens are limited to 1023 SIDs. Users with membership in 500 or more Groups may experience log on or access problems under WIN2K. Corrective action may be to consolidate or reduce the number of Groups a User is a member of; Turn off the SID History function after the migration is completed; Remove the SIDHistory entries.

## **CELERRA SID HISTORY SUPPORT:**

Celerra Supports SIDHistory with the following "usrmap.cfg" entry, which must always be the last line of this file:

**\_history\_sid\_range\_:600:17500:18500:18600:19600**

**Comment:** SID History range optional in NAS prior to 4.2.13.0. 4.2.13.0 and higher requires SID History range configured in all "usrmap.cfg" files. Usrmap will no longer work properly without this entry. The SID History range will now map correctly, whereas in versions prior to 4.2.13.0, the SIDs were mapped from regular Domain ranges & not the SID History range. SID History was never implemented correctly in any NAS version prior to 4.2.13.0 due to a Netbios name limitation of 15 characters and a real "**\_history\_sid\_range\_**" of 19 characters.

## **RUNNING THE ACTIVE DIRECTORY MIGRATION TOOL:**

1. Download "ADMT.EXE" from Microsoft: 9 Mar 2000 Windows NT to Windows 2000 Active Directory Migration Tool (ADMT) Version 1.0 2543kb. Installed on PDC Emulator (Creates new MMC Snap-in)
  - a.) Verifying PDC Emulator: ADUC>Rightclick "T2DOM2.LOCAL">Operations Master>PDC:
- Operations Master: win2ksrvdom2.t2dom2.local**
2. Must have 2-Way Trusts established between Target & Source Domains
  - a.) User Manager for Domains>Policies>Trust Relationships: Add T2DOM2 to "Trusting Domains" on T2DOM1 NT 4.0
  - b.) Active Directory Domains and Trusts>Rightclick "T2DOM2.LOCAL">Properties>Trusts>Add T2DOM1 to "Domains trusted by this domain" [Listed as an External Trust]
  - c.) Add T2DOM1 to "Domains that trust this domain" on AD Server
  - d.) Add T2DOM2 to "Trusted Domains"
- Note:** If having difficulty establishing Trusts, make sure that WINS defines the Windows 2000 Domain Controller and the Domain Name for the NT 4.0 Domain.
3. Choose a User account that has Domain Admin privileges on Source Domain
  - a.) Add Target "Domain Admins" group to local group "Administrators" on NT 4.0 Domain
  - b.) Add Target "Domain Users" group to local group "Users" on Source Domain
4. Run the Active Directory Migration Wizard: Programs>Administrative Tools>Active Directory Migration Tool>Rightclick>User Migration Wizard; Group Migration Wizard; Computer Migration Wizard; Reporting Wizard; Service Account Migration Wizard; Trust Migration Wizard; Group Mapping & Merging Wizard
  - a.) User Migration Wizard>Welcome to the User Account Migration Wizard>Next
  - b.) Test or Make Changes → \*Test the migration settings and migrate later?  
→ o Migrate now?
  - c.) Domain Selection>Source Domain: T2DOM1 Target Domain: T2DOM2.LOCAL
  - d.) User Selection>Add>T2DOM1>select Users to migrate>Next
  - e.) OU Selection>Users>Next: LDAP://T2DOM2/CN=USERS,DC=T2DOM2,DC=LOCAL
  - f.) Password Options: \* Complex Passwords \*Same as User name
  - g.) Account Transition Options: \*Disable Source Accounts \*Disable Target Accounts \*Leave both accounts open  
\_\_x\_\_Migrate User SIDs to target domain [This is SID History]

## **PREREQUISITES HANDLED BY ADMT MIGRATION WIZARD:**

- 1.) --Auditing is not currently enabled on Target domain. Enable? Yes

**Note:** NT Auditing must be turned on for both Source & Target Domains

Source: User Manager>Policies>Audit>Select Audit These Events: \*Success \*Failure User and Group Management

Target: AD Users & Computers>DC Container>Group Policy tab>Default DC Policy>Edit>Computer Config>Group

Policy>Windows Settings>Local Policies>Audit Policy>Audit Account Management: Select "Define these Policy Settings" \*Success \*Failure [Use SECDEDIT to refresh policy: c:>secedit /refreshpolicy machine\_policy]

- 2.) --The Local Group 'T2DOM1\$\$\$' does not exist on T2DOM1. This group is required to migrate SIDs. Would you like to create it? Yes

- 3.) --TcpipClientSupport registry key is not set on the Source domain. Would you like to add this registry key? Yes

**Note:** This entry is required to allow the Lsa of the Source Domain to communicate with the Target Domain

## **HKLM>System>CurrentControlSet>Control>Lsa>TcpipClientSupport: REG\_DWORD: 0x1**

- 4.) --Would you like to reboot the PDC to insure changes will take effect? Yes.
  - h.) User Account To add SID history, you must supply a user account with the proper permissions:  
Username: nasadmin  
Password: \*\*\*\*\*  
Domain: T2DOM1
  - i.) User Options. You can migrate user groups, profiles, and security settings. Please select the following options to customize your migration task.  
 Translate roaming profiles [Leave blank]  
 Update User Rights  
 Migrate associated user groups (recommended)  
 Update previously migrated objects  
Select how all migrated accounts should be moved.  
 \*Do not rename accounts  
 \*Rename with prefix:  
 \*Rename with suffix:
  - j.) Naming conflicts. To select how you like to resolve group account name conflicts, click an option below.  
 o Ignore conflicting accounts and don't migrate [Change to select this]  
 \*Replace conflicting accounts [Default choice, unselect this]  
 \_x\_Remove existing User rights  
 \_x\_Remove existing members of groups being replaced  
 o Rename conflicting accounts by adding the following: o Prefix: o Suffix:
  - k.) You have successfully completed the User Account Migration Wizard. Finish. Migration Progress... Status: In progress. Completed. The "Viewlog" will document results of the Migration.

**Migrating Groups:** Rule of thumb—always migrate Users and Global Groups together. Groups—migrate Global Groups as “closed sets” when migrating Users    Ignore conflicting accounts and do not migrate    Replace conflicting accounts    Rename conflicting accounts

**Group Mapping & Merging Wizard:** Migrates Group membership from Source to Target; Merge multiple Groups into one

**Tools:** usrstat; showmbrs; showgrps

## **WIN2K ACTIVE DIRECTORY PREREQUISITES and CELERRA CIFS:**

Nas 2.2.35.4/2.2.49.0/4.0.12.1 [See notes below for details on specific support per code level]

Kerberos Authentication for Win2k network client/server authentication using secret key cryptography

Dynamic DNS--mechanism which DataMover uses to register on the WIN2K domain, as well as for Host Name resolution

**Important Point:** Celerra CIFS must use the "reverse zone lookup" feature of DNS, which is not setup by default for DNS--therefore, ensure that it is setup--this is the only way a CIFS DataMover can join the WIN2K Active Directory Domain as a "Windows 2000 CIFS Server". If running Windows 2000 native mode & you have multiple Domains, then DNS must be able to resolve the Global Catalog Service [port 3268] in order to authenticate users [the Gobal catalog authenticates all Users]. Each Domain runs its own Active Directory and has its own namespace in DNS. Trusts are created automatically if within same 'forest' or by specifying each other as the secondary DNS server for the other's zone DC's TCP/IP properties; configure to allow Zone Transfers.

Network Time Protocol [NTP]--Standard Internet Time Service required for implementing Kerebos Network Authentication

Active Directory--requires use of LDAP [Lightweight Directory Access Protocol] protocol between Clients and WIN2K Servers

**Note:** Active Directory is stored in NTDS.DIT and can hold up to 1.5 million objects

Windows NT 4.0 & Celerra CFS 2.2.49.2 +/4.0.12.1 cannot use IPSec for encrypted network communications

## **FORCING ACTIVE DIRECTORY REPLICATION:**

Programs>Administrative Tools>Active Directory Sites & Services>doubleclick the Domain Controller>Select NTDS Settings & rightclick in right window pane on <auto generated> Replicate Now

## **FORCING AD REPLICATION FROM RUN BOX:**

**c:> repadmin /syncall <dest dsa> [<naming context>] [<flags>] <dc name>**

**C:\2k3tools>repadmin /syncall george.2k3.pvt.dns**

CALLBACK MESSAGE: SyncAll Finished.

SyncAll terminated with no errors.

**c:> repadmin /showconn /showmeta <object dn> [DSA] [/nocache]**

## **Replication Monitor GUI:**

**c:>replmon** [Add AD Servers—use tool to verify Operations Master and Global Catalog Roles]

## **WINDOWS 2000 SUPPORT TOPICS:**

**Organizational Unit:** A container used to organize and store objects in the domain

### **Forest:**

AD Schema contains all forest objects & attributes, stored in Global Catalog of Domain Controllers. Multiple Trees can be created to form a “Forest”. Multiple forests can be trusted.

### **Tree:**

Multiple domains combined into hierarchical trees. All Domains within a Tree & Forest share a common Schema & Global Catalog. Domain Structure usually goes from a Parent to Child domains.

### **Sites:**

Multiple domains can be in a single site; Single Domains can be in multiple sites

Used to service client requests for network resources

Used for AD replication among Domain Controllers

Generally based on geographic locales & IP Subnets

Minimum of (1) DC running the Global Catalog Service in each site

### **Global Catalog:**

Separate db from AD and contains partial RO replica of AD objects in entire forest. Used to provide Universal group membership information during logon process and also when using UPN name to logon

### **Operations Master Roles:**

Schema Master (updates & mods to AD Schema) and Domain Naming Master (domain name space changes) in the forest  
PDC Emulator, RID Master, and Infrastructure Master in the domain

## **ACTIVE DIRECTORY:**

AD database uses ESE (Extensible Storage Engine) and is similar to Exchange transactional logging & online db maintenance. AD database contained in a file called Ntds.dit

Comprised of multiple sites

Schema contains list of object classes & attributes in the Forest--stored in Global Catalog; View Schema using MMC AD Schema snap-in after registering a dll: run: regsvr32 schmmgmt.dll. Also edit AD objects using ADSI Edit Snap-in; register the dll: run: regsvr32 adsedit.dll

Contains Domain, Schema, and Configuration data directory partitions

Information about Users, Groups, Machines in a logical and hierarchical tree structure

User Name Format Changes from NT 4.0 Style to Win2k Style: UPN (Universal Principal Name): [tmattha@t2dom2.local](mailto:tmattha@t2dom2.local)

AD Replication conducted by KCC (Knowledge Consistency Checker) Service in AD using IP or SMTP protocols

Global Catalog represents central repository of all AD Objects in Tree or Forest—required for network logon

REPADMIN tool used to monitor and diagnose replication of AD

### **AD Names:**

Logon Names: UPN=User Principal Name ([user1@t2dom2.com](mailto:user1@t2dom2.com)) SAM=legacy (user1)

GUID: Globally Unique Identifier number

LDAP Names: DN=Distinguished Name (/DC=com /DC=t2dom2 CN=users /CN=user1)

RDN=Relative Distinguished Name (CN=user1);

URL=Universal Resource Locator Name (LDAP: //t2dom2.com/cn=user1,cn=users)

Canonical Name=t2dom2.com\users\user1

SID=Security Identifier for Users, Groups, & Computers

Security Principal Name (user1)

LDAP name components commonly used in AD: DC (Domain Component), OU (Organizational Unit), CN (Common Name)

### **UPN (Universal Principal Name):**

→Domains have default current Domain and root Domain User Principle Name suffixes for User accounts. Companies can use alternative UPN suffixes to provide more logon security and to simplify the logon name for a User.

**Active Directory Domains and Trusts>Properties:** Alternative UPN suffixes (Add)

→The alternative UPN suffixes do not need to match the current or root domain naming structure

### **Example of UPN Use:**

Internal Windows Domain Name: iq.no

External Internet Domain Name: iqas.no

Pre-Win2k Name: iq\pin (uses NTLM authentication)

FQDN: [pal.innvik@iqas.no](mailto:pal.innvik@iqas.no) (uses Kerberos authentication)

→UPN's for Users in the base domain “iq.no” can be assigned any suffix name: [tm@network.com](mailto:tm@network.com), and still maps to the same underlying User SID in the “iq.no” domain

## **WIN2K Trusts:**

Within “forest” trusts are 2-way Transitive between “parent” and “child” domains

“Shortcut Trusts between child domains within forest are 1-way trusts, but can have 1-way trust in each direction

Between “forests” of other external domains, trusts are “external”, with 1-way trusts in each direction

## **WIN2K Groups:**

### **Local Groups:**

Contains Local, Global, Universal groups from parent domain; Global & Universal Groups from Trusted domains; Global Groups from NT 4.0 Domains

### **Domain Local Groups:** Cross forest oriented

Contains Users, Universal & Global Groups from any domain in Forest; other Domain Local Groups from same domain only

Nesting of Domain Local Groups within Local Groups now supported at 4.2.14 +

**NAS 5.1.9.4 Bug:** Adding “Domain Local Groups” to Administrators group on DM fails—‘...Member has wrong account type.’

### **Global Groups:** Domain oriented (Nest Global Groups into Universal Groups for replication reasons)

Contains Users or other Global Groups from same domain

### **Universal Groups:** Forest oriented

Contain Users, Universal and Global Groups from any Domain in Forest; other Universal Groups

## **GROUP POLICY OBJECT (GPO):**

Manages software configurations based on Group Policy Objects & applied during User Logon and Machine Startup

Stored on a Domain basis

GPO’s are applied to Sites, Domains, OUs, are cumulative, and are inherited from Top down

Domain Admins, Enterprise Admins, local system have Full Control permissions to modify GPO’s by default

Celerra supports machine account GPO’s in NAS 5.0 and higher—use \$server\_security server\_x -i -p gpo

## **TWO TYPES OF GPO’s:**

### **COMPUTER SETTINGS:**

Settings processed when domain member system starts up

### **USER SETTINGS:**

Settings processed when user logs onto domain member

### **GPOs ARE LINKED TO AD CONTAINERS:**

AD Site

AD OU

AD Domain

### **ORDER OF PRECEDENCE FOR GPOs:**

Local GPO, Site GPO, Domain GPO, OU GPO (gPLink & gPOptions attributes)

### **DFS PATH REFERRALS SYSVOL SMB SHARE:**

[\\domainname\sysvol\domainname\Policies\GPO\\_GUID & gpt.ini](\\domainname\sysvol\domainname\Policies\GPO_GUID & gpt.ini) file \Machine\registry.pol; \User\registry.pol

## **GPO LOOPBACK POLICY:**

Certain special-use computers, such as Terminal Servers, can have the Loopback Policy applied so that all Users logging into the computer will have GPO’s applied to their User accounts. GPO’s can be applied in a Merge Mode (User gpo’s are gathered using GetGPOList call, and computer’s GPO’s are called last and added to the User’s list) or Replace Mode (only computer GPO’s are used). For applying GPO’s to Terminal Servers, you might chose to put the computers into a separate OU or use the Loopback Policy settings mentioned here.

## **DATAMOVER & WINDOWS 2000 JOIN PROCESS:**

1. Celerra needs to discover which DC’s are closest in proximity—uses LDAP over UDP queries to DC’s to accomplish this
2. Will use ping if needed to help determine closest DC’s [shortest round-trip]
3. Actual Join to AD is done using closest ‘available’ DC using LDAP over TCP
4. Celerra will also attempt to contact closest KDC server first, but if the KDC does not know the principal “administrator” or ‘server’ name, will contact others in the KDC ‘realm.’ Kerberos & Kpasswd services are used to obtain a Kerberos Ticket used in Join Process & setting server password, and thereafter for renewals.
5. List of DC’s seen in “server\_cifs” output is regenerated every 15 minutes [Not in 4.1, but in 4.2 & 5.0]
6. > carat symbol in “server\_cifs” means only that DataMover has a current TCP connection with the DC

## **DATAMOVER JOIN PROCESS:**

1. DM queries DNS for list of DCs and returns fastest DC list
2. DM queries DNS for LDAP “SRV” records and replies with Service & listening Port number
3. DM queries DC’s using LDAP null query to make sure that LDAP service is up
4. DM queries DNS for Kerberos and Kpasswd services
5. DM binds and unbinds to LDAP again to verify that authentication works
6. DM binds to LDAP to find out supported security layer to use for authentication, then unbinds
7. DM logs in to AD with ‘Guest’ account to make ADSI calls [DM establishes defaultNamingContext & encryption type]

8. DM obtains TGT for specified User in Join (KRB\_AS\_REQ to KDC to obtain Kerberos ticket and receives KRB\_AS REP?)
  9. DM sends KRB\_TGS\_REQ for ‘Service Ticket’ for LDAP service to prepare to modify LDAP db
  10. DM binds using SASL and initiates direct calls to ADSI using LDAP bind [Checks for existence of compname and creates OU’s]
  11. DM checks for name and creates new one if does not exist.
  12. Service principals for compname are modified and AD is populated with DM compname and other information such as SAM account and ObjectGUID
  13. Account password is changed for compname using kpasswd protocol & unbinds from LDAP service
- Note:** /etc directory populated with krb5.account, krb5.conf, krb5.conf.old, & krb5.keytab
- krb5.account: database file holding Kerberos information
- krb5.conf: holds encryption support types or TGS
- krb5.conf.old: backup file
- krb5.keytab: file contains information from current Authentication session with KDC

## **HOW DATAMOVER DISCOVERS DC & SRV SERVICES IN AD:**

- DM uses a locator service, similar to Windows 2000, to query Windows 2000 DNS Domain/Realm & determine Site
- DM requests list of services (SRV records) for Domain from DNS
- DC’s are validated via LDAP\_Ping, then sorted based on SITE and LDAP ping response times [server\_cifs list]

## **WINDOWS NT/2000 LOGON AUTHENTICATION:**

NTLM/NTLM v2 for NT/Kerberos & NTLM [NTLM & NTLM V2] for Windows 2000

**Note:** Legacy NT 4.0 systems use traditional NTLM for User logon authentication, while Windows 2000 uses MIT’s version 5 (RFS1510) implementation of Kerberos. Authentication protocols allow computers to mutually identify each other on the network.

## **WINDOWS 2000 MACHINE LOGON PROCESS:**

- Client→Boots Up & Obtains IP configuration from DHCP Server [APIPA—Auto Private IP Addressing]
- Client→Locator “DsGetDcname” Service collects information needed & passes to Local NETLOGON Service
- Client→NETLOGON uses “DsGetDcName” service to “DNSQuery” to obtain A & **SRV records** for LDAP Server [ldap.tcp.dc.\_msdcs.win2kdomain.com]
- Client→NETLOGON Service retrieves list of Domain Controllers & caches list
- Client→Uses **locator service** to find Site and list of DC’s; ldap.tcp.Default-First-Site-Name.\_sites.dc.\_msdcs.win2kdomain.com]
- Client→Queries list of DC’s using LDAP UDP datagrams to determine which DC to log into
- Client→Uses **netlogon** process to create Secure Channel for ‘logon’ & ‘authentication’ to DC [SMB & RPC End Port Mapper svc]—LDAP protocol is used to establish logon via local SAM to Directory Service Agent
- Client→Kerberos Session Ticket from KDC to then create an “IPC\$” connection to DC [Queries KDC service and negotiates Ticket]
- Client→Queries **LDAP RootDSE** for standard directory information
- Client→Uses LDAP query to locate Group Policy Objects & then loads policies from DC’s “SYSVOL” share & ‘autoenrolls’ cert.
- Client→Time synchronization over Port 123 with NTP Server
- Client→Updates Dynamic DNS with its name
- Client→Breaks down connections with DC using SMB Tree disconnect

## **CLIENT LOGON ABBREVIATED:**

DHCP Networking→Site & DC Info→Secure Channel to DC→Kerberos AS/TGS Tickets→Load Group Policies→Enroll Client Certificates→Time Sync→Dynamic DNS Updates→Completion

## **WINDOWS 2000 USER LOGON PROCESS:**

User→initiates logon using SAM Account Name, Domain, & Password; User Principle Name (UPN); or FQName--user@win2kdom.com]

**Note:** UPN logon only supported in Native Mode and uses Global Catalog to lookup User Account

User→Logon process generates Kerberos Authentication request for Session Ticket from KDC via Ticket Granting Service

User→Group Policy Object information downloaded from “SYSVOL” share

User→Closes SMB Port 445 connection to DC after logon process completed

## **UPNs & SPNs (User Principal Names & Service Principal Names):**

UPNs are unique in a forest, and consists of a User’s Principal name, @ symbol, and DNS domain name. SPNs are also unique in the forest and represent a unique service for each machine account in the domain—consists of ServiceClass, Host, Port, Service name. Use the setspn –list Resource Kit tool to see registered SPN service names for Servers.

## **PROBLEM WITH SPNs, DNS ALIASES, and WINDOWS 2003:**

→Users unable to connect to Celerra shares after DCs upgraded from W2K to W2K3, when the client is trying to use a DNS alias and SPNs are in use in AD. Users can access if they use the real “compname”. Problem is that with W2K3, the KDC may not provide the canonical compname of the CIFS server to the client requesting the access ticket, and the connection to the DM will fail. See emc158629, AR92558 for more detail.

**Resolution:** Upgrade to NAS 5.5.29 or higher and set the following parameter to 1

## # server\_param server\_2 -facility cifs -info LanmanServer.disableNameChecking

server\_2 :

```
name      = LanmanServer.disableNameChecking
facility_name = cifs
default_value = 0
current_value = 0
configured_value =
user_action = reboot DataMover
change_effective = reboot DataMover
range      = (0,1)
description = Disables checking of the server's principal name of the client's kerberos ticket.
```

## **COMMON TCP/IP PORTS USED FOR WINDOWS 2000:**

|           |                                                                                                           |
|-----------|-----------------------------------------------------------------------------------------------------------|
| Port 42   | WINS Replication                                                                                          |
| Port 53   | DNS Service                                                                                               |
| Port 68   | DHCP (UDP only)                                                                                           |
| Port 88   | Kerberos                                                                                                  |
| Port 135  | RPC; RPC EndPoint Mapper; Wins Manager                                                                    |
| Port 137  | NetBIOS Name Service: Logons; Browsing; NT 4.0 Trusts, Secure Channel (UDP only); WINS Registration (TCP) |
| Port 138  | NetBIOS Datagram Service: Logons, NT 4.0 Trusts, NetLogon; Browsing (UDP only)                            |
| Port 139  | NetBIOS Session Service: NTB; SMB; File Sharing; NT 4.0 Secure Channel/Trusts; Logons (TCP)               |
| Port 389  | LDAP                                                                                                      |
| Port 445  | SMB or CIFS                                                                                               |
| Port 464  | Kerberos Kpasswd (UDP)                                                                                    |
| Port 636  | LDAP over SSL                                                                                             |
| Port 3268 | Global Catalog Search for LDAP                                                                            |

## **KERBEROS BASICS:**

- In networking parlance, Kerberos is a protocol designed for ‘distributed security’
- Kerberos is implemented as an SSP (Security Support Provider) using SSPI interface
- NTLM & SSL are also implemented as SSP authentication solutions on MS networks, using SSPI interface
- Basic Kerberos v5 protocol implements 56bit DES (Data Encryption Standard) encryption by default
- MS networks use 128bit RC4 encryption with their version of Kerberos
- Kerberos offers delegation and mutual authentication, which NTLM does not
- Kerberos assures security for authentication and network services using encryption based mostly on secret key encryption (aka private key, symmetric keys, or shared secrets)
- Password-based keys are only used during login process, and even then, the Principal’s (user) logon is a hash of their password, aka message digest or checksum. The bitstring cannot be used to recover the original input value, i.e., the password itself, and is called a one-way hash. On the KDC, the passwords of all users and computers are hash values and not the actual password itself!
- Users, servers, and other workstations all have passwords, known as security principal
- Kerberos tickets contain encrypted & unencrypted information

### **Unencrypted portion of ticket:**

Name of Windows 2000 domain (realm) and name of principal of the ticket (user or computer)

### **Encrypted portion of ticket:**

Encryption session key, used to encrypt data being exchanged; encrypted copy of principal name & domain name; start & end times for ticket validity; IP addresses identifying user system; User authorization data (SIDs, etc); other fields

### **BASIC KERBEROS SERVICES:**

- Authentication of Clients & Servers
- Data integrity and data privacy

**Note:** Checksums use HMAC with RC4 encryption, session keys between Client & Server are shared

## **CELERRA WIN2K AUTHENTICATION SUPPORT:**

- With NAS 5.5, Celerra uses Kerberos v5, NTLM, or NTLMv2 as valid authentication packages on MS domains
- In general, Windows 2000/XP/2003 clients should use Kerberos authentication

→ NT Clients will be authenticated using NTLM v. 1.2 if within the same or trusted Domain as the Celerra

**Note:** Celerra does not provide native NTLMv2 support until NAS 5.3.17.x & NAS 5.4

→ In addition, Celerra supports SMB & LDAP Signing

## **CELERRA DOES NOT SUPPORT SChannel (SSL/TLS):**

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
Celerra does not support the ability to use SChannel (SSL/TLS, Secure Socket Layer/Transport Layer Security) as an authentication SSP module (Security Support Provider). Celerra supports only the Kerberos and NTLM (LAN Mgr, NTLM v1, NTLM v2) SSP authentication mechanisms.

## **PURPOSES FOR DATA MOVER COMMUNICATES WITH DOMAIN CONTROLLERS:**

- Kerberos service ticket validation
- Pass-through authentication, as in case of NTLM clients
- Retrieving GPO information for the domain
- Listing out trusted domains to domain Joined to
- Resolving names to Windows SIDs, resolving SIDs to Windows names

## **KERBEROS VERSION 5:**

- Security principal name for Windows domain is called krbtgt, and is used by KDC
  - A network authentication and security protocol in use by Windows 200/2003 Domains [Kerberos V5—based on the MIT Version 5] which replaces the NT 4.0 NTLM Netlogon Services protocol---supported by 4.0.12.1
  - Kerberos allows for mutual authentication between computers on a network (NTLM does not)
  - Kerberos provides for faster network connections between C/S (NTLM requires communication to DC each time Client connects)
  - Uses UDP/TCP over IP Port 88 for legacy computers and W2k/W2k3 systems for both Client & Server
  - Kerberos V5 is default authentication protocol for Windows 2000—see RFC 1510
- Note:** Celerra supports only Kerberos over UDP until NAS 5.2
- Kerberos Tickets are built to include User name & User SID
- “Realm” serves a single Kerberos database and set of KDCs. Realm names are displayed in uppercase and usually correlate to the Windows 2000 Domain name. A Realm consists of the Kerberos db of all security principals in the domain, therefore the Realm is usually equivalent to the Windows domain.

## **TYPICAL CLIENT WORKSTATION KERBEROS LOGON PROCESS:**

- I. Client logs into AD with password and sends “AS” [Authentication Service] request to KDC AS Service, receiving Ticket Granting Ticket [TGT]
- II. Client requests access to Service or Resource on network by sending TGS [Ticket Granting Service] to TGS on KDC. Individual TGS is returned to the Client.
- III. Client submits TGS to Server using AP messages, with SIDs and PAC as part of the ticket sent to the Server. If Client requests mutual authentication for this exchange, Server will respond with Authenticator Timestamp.

## **KERBEROS AUTHENTICATION SERVICE:**

Only portion of Kerberos communications that requires a password, or pre-shared secret, to be verified between Server and Client, typically invoked during logon. A TGT (Ticket-Granting Ticket) is provided to the client logging into the domain, from the KDC, which can then be used to obtain additional service tickets. AS exchange uses KRB\_AS\_REQ from client to server. Preauthentication data is included in the Client’s secret key timestamp and KDC realm must match. Applications, or the O/S, can specify the type of SSP to use [i.e., NTLM or Kerberos], over Port 88. Clients secret keys are 64-bit hash of the client’s password and are replicated to all DC’s, whereas private keys are not shared with other DC’s. Server’s KRB\_AS REP consists of encrypted data that cannot be read by a Client except for server name and realm. Client presents server name and realm when presenting Authorization Ticket to Server. Data block encrypted to client’s secret key and holds session key for client.

## **WHAT IS MUTUAL AUTHENTICATION?:**

- System whereby communication partners share crypto keys, called ‘Shared Secrets’ in order to verify each other
- Shared Secret Keys are symmetric, meaning that a single key can be used to both encrypt and decrypt keys
- Use of Authenticator contains info based on name & timestamp and encrypted based on client’s master key, which receiving system can decrypt to prove identity, sending back an Authenticator with portion of key encrypted with its master key and original timestamp, proving its identity to Client.

## **(3) KERBEROS PROTOCOLS IN OPERATION FOR WINDOWS NETWORKS:**

### **I. Authentication Service Exchange (AS Service):**

Provides logon session key & TGT’s to Clients, using the TGS exchange (Ticket Granting Service)--KRB\_AS\_REQ & KRB\_AS REP

Used by the KDC to provide a Client a logon session key and a Ticket Granting Ticket (TGT) for further communication between Client & KDC. When a User logs on, their password is encrypted using DES-CBC-MD5 encryption, from which a Master Long-term Key is derived. Client sends KRB\_AS\_REQ message, containing User’s information, name, and service requested, as well as pre-authentication data (Preauth Data is usually the timestamp encrypted with Client’s longterm key). The KDC retrieves a cached copy of the User’s Master key from AD to decrypt pre-authentication data and review timestamp, and creates logon session key and

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
credentials for the user. KDC returns logon session key encrypted with User's longterm key, and another copy of the logon session key in the form of the TGT, which contains the user's Session Key encrypted with the User's Master Key, which is also encrypted by the KDC's long-term master key & returned to client via KRB\_AS REP. Client decrypts logon session key using User's master key, and stores in ticket cache. TGT is also extracted and stored in local ticket cache. For further communications to KDC for the current logon session, Client uses the Session Key. The TGT has a special authorization data field that contains the User's SID, SIDs from Domain Groups, and in multiple domains, SIDs from Universal Groups (queries the GC for this info).

## **II. Ticket Granting Service Exchange (TGS):**

### **Provides Service Tickets for Network Resources—KRB\_TGS\_REQ/\_REP**

Client sends Kerberos TGS (KRB\_TGS\_REQ) to KDC for access to a network service (supplying Session Key, Authenticator, TGT). Request include's User's name, Authenticator encrypted with logon Session Key, and TGT obtained during original AS Exchange, name of service requested, and Preauthentication data to prove User's ID, encrypted timestamp from Client's Master Key. KDC decrypts the TGT using master key, extracts User's logon session key, which is then used to decrypt the User's authenticator. KDC then creates new Session Key to be used for the TGS, encrypting one copy of the Session Key with Client's logon session key, and embeds another copy of Session Key in the TGS, that contains User's authorization data (SIDs, etc) encrypted with Server's longterm Session Key (KRB\_TGS REP). User decrypts target Server session key using logon session key, and stores session key in local cache, along with target Server's ticket.

### **Multiple Domain Communications:**

If a Client User exists in a Child domain, and needs to communicate to another child domain Server, user would request TGS from KDC, which would in turn refer User to KDC in the other domain, and process may need to be repeated until the KDC in the Server's domain in contacted, which can then provide the TGS to the Client.

## **III. Client Server Authentication Exchange (CS):**

### **Client & Server mutual authentication—KRB\_AP\_REQ/\_REP**

Client sends Application Request (KRB\_AP\_REQ) to Server, which contains an Authenticator encrypted with Session Key for the service and the TGS ticket from the KDC, encrypted with Server's master key. Server retrieves authorization data and decrypts session key with secret key, extracts the Session Key, and decrypts Authenticator, reviewing timestamp for validity. Server returns timestamp encrypted with Session Key in form of Application Reply (KRB\_AP\_REPLY) to authenticate itself. The Session Ticket contains Domain name, User name, and time of ticket validity.

## **KERBEROS PREAUTHENTICATION:**

KerberosV5 uses preauthentication using encrypted timestamps by clients to prove their password when going to the KDC for a TGT.

## **KERBEROS AND GPO POLICIES:**

GPO settings for Kerberos can only be defined once, at the AD Domain level—never at the OU or SITE level

## **KERBEROS TICKET DEFAULTS:**

Forwardable flag is set, used for delegation (Client's authentication key can be forwarded to multiple systems using same credentials)

Ticket renewal flag is set

Preauthenticaion flag is set

TGT has initial flag set

## **DEFAULT KERBEROS TICKET PARAMETERS FOR WINDOWS 2000:**

**Note:** Enforced by default Domain policy template

MaxServiceTicketAge: 10 hours (Service Ticket)

MaxTicketAge: 10 hours (User Ticket TGT)

MaxRenewAge: 7 days (Service or TGT Ticket)

MaxClockSkew: 5 minutes (time skew between Kerberos authenticator timestamp and current KDC or Resource Server time)

Enforce Logon Restrictions: On by default—KDC checks validity of User account every time a ticket request is submitted

## **DEFAULT CELERRA KERBEROS PARAMETERS:**

|                                  |                                                                             |
|----------------------------------|-----------------------------------------------------------------------------|
| Enforce User Logon Restrictions: | Default=enabled [Check validity of User each time ticket request submitted] |
| MaxServiceTicketAge:             | Default=600 minutes [Kerberos service ticket]                               |
| MaxTicketAge:                    | Default=10 hours [User Ticket lifetime for the TGT]                         |
| MaxRenewAge:                     | Default=7 days auto-renewal for User TGT per GPO if User remains connected  |
| MaxClockSkew:                    | Default=5 minutes clock synchronization, or per Domain GPO                  |

## **WINDOWS 2000 SUPPORTED KERBEROS ENCRYPTION TYPES:**

RC4-HMAC 128-bit [Default encryption type for Windows Clients]

DES-CBC-MD5 56-bit [When using DES encryption, Win2k Clients use this encryption type for session tickets]

DES-CBC-CRC 56-bit [Used only for MIT Kerberos, not Win2k Clients]

**Note:** DES stands for the Data Encryption Standard, a block cipher algorithm using 56-bit keys. Triple DES (TDES) is a more secure combination of three sets of 56-bit DES encryption keys [168-bits]. AES (Advanced Encryption Standard) is largely replacing DES and has much better performance characteristics than TDES. The block cipher is 128-bits, but has different key sizes of 128, 192, or 256 bits. Celerra does not use TDES or AES encryption at this time.

## **KERBEROS CACHE ON WINDOWS CLIENTS:**

Tickets are cached for User logon sessions by the LSA

Use ResKit Utilities klist.exe or kerbtray.exe to verify

**C:\>klist tgt | tickets**

Cached TGT:

ServiceName: krbtgt

TargetName: krbtgt

FullServiceName: mattat

DomainName: CORP.EMC.COM

TargetDomainName: CORP.EMC.COM

AltTargetDomainName: CORP.EMC.COM

TicketFlags: 0x40e00000

KeyExpirationTime: 256/0/29920 0:100:8048

StartTime: 7/22/2005 7:21:19

EndTime: 7/22/2005 17:21:19

RenewUntil: 7/29/2005 7:21:19

TimeSkew: 7/29/2005 7:21:19

## **KERBEROS TROUBLESHOOTING UTILITIES:**

mytoken.exe →Display contents of User's access token, including user rights & group memberships

klist.exe →Examines local ticket cache, purges cache

kerbtray →GUI version of ticket cache

netdiag →netdiag /test:Kerberos

netdom →manages domains and trusts; reset member accounts; verify, reset time, etc.

setspn →View current SPNs, reset, add or delete

Network Monitor: nltest →discover trusted domains, DC's, DC for a particular Trusted Domain User

## **SUPPORTED KERBEROS INTERFACES USED BY WINDOWS:**

SSPI (Security Support Provider Interface)—very similar to GSSAPI (General Security Service Application Prog. Interface)

LsaCallAuthenticationPackage—used to retrieve tickets from Kerberos cache. Windows 2000 uses the SSP as its Authentication Protocol for Kerberos, and also supports an SSP for NTLM authentication. Windows 2000 AD Servers stores Tickets and Keys in a credential cache managed by Kerberos SSP (LSA authority).

## **HOW TO RESTRICT AUTHENTICATION TO KERBEROS ONLY:**

**\$server\_cifs server\_x -add compname=swift, domain=mover.com, authentication=kerberos**

**Note:** Use this option to restrict authentication to Kerberos only—most secure mode for Celerra.

## **MICROSOFT SP3 ISSUES:**

**MICROSOFT KERBEROS MAXTOKENSIZE ISSUE & HOT FIX:** Primus emc62537, emc74207, & emc80211

Celerra code versions prior to 4.2.14.403; 4.2.16.0; 5.0.16.0; 5.1.12.0; 5.1.9.100/300; and 5.2.1.0 all have potential to cause Rolling Data Mover panics after application of SP3 or SP4 to Windows 2000 Domain Controllers, or any of the MS Hot Fixes for Kerberos “MaxTokenSize” that allow for greater numbers of Group SIDs in the Kerberos Tokens of Users—prior to the fix, a User could only be a member of about 70-80 groups before running out of buffer. The main impact of this limitation is related to GPO’s, and certain GPO’s could not be applied properly. From Celerra perspective, the original corrective action required removing Hot Fix, Unjoining & Rejoining DM to Domain. Latest guidance suggests that applying the latest 4.2.17.2 nas.exe is sufficient to stabilize environment without further action.

**Cause:** Q263693/Q280830 Microsoft HotFix addresses Kerberos Token size limitation and increases large group support.

**Microsoft Registry Fix to Increase MaxTokenSize on AD Servers:**

HKLM>System>CurrentControlSet>Control>Lsa>Kerberos>Parameters> Add new parameter called "MaxTokenSize" and set in DECIMAL to 100000

**Note:** Limitation is approx 70-80 groups but can vary considerably

**Symptoms:** Data Movers panic continuously; Server\_log errors→SSXAK=c000006d origin=600 stat=d0000; Mapping drives to CIFS Shares causes DM Panics

**Common Panic Headers:**

A.) Page Fault Interrupt..

0x95cf80: 0x290a3 \_in\_pcblklookup+0x37

0x95cfb8: 0x2e7d11 \_tcp\_pcblkLookup\_\_FP5inpcbUIUsUIUsi+0x9d

0x61cec: 0x2df7b4 \_tcp\_input\_\_FP13IPInterface\_tP4msgbi+0x330

0x61e80: 0x2c2fa4 \_ip\_input+0x3c4

0x61ea4: 0x2be36d \_ip\_mxrrput+0x1dd

0x61ec0: 0x2136a3 \_alread\_FPV\_P4msgbUsUsUs+0x333

0x61ee8: 0x21e59a \_FillJumboRecvRing\_\_FP7altinst+0x96a

**B.) Page Fault Interrupt...**

0x95d1f4: 0x2e9729 \_sbspace\_\_FP5tcpcb+0x11

0x95d218: 0x2e5946 \_tcp\_output\_\_FP5tcpcb+0x11a

0x61f88: 0x2ea160 \_tcp\_fastimo\_\_Fv+0xc4

0x61fa0: 0x13faa3 \_RTtimer\_sched\_\_FiPv+0x57

0x61fb4: 0x140a0d \_start\_\_15SchedRT\_Initial+0xe9

0x61fe4: 0x1434d1 \_Sthread\_startThread\_internal+0x11

**C.) PANIC in file: ./malloc.cxx at line: 312 : Memory corrupted**

0x14f41d28: 0x13fe57 \_PANIC+0x3b

0x14f41d48: 0x17aec9 \_free+0x55

0x14f41d70: 0x3923d4 \_krb5\_free\_ticket+0x28

0x14f41d84: 0x789306 \_SSXAuth\_KERBEROS\_\_13smb\_threadCtxR16cSessionS

0x14f41e20: 0x7897bd \_SSXAuth\_SERVER\_EXT\_\_13smb\_threadCtxR16cSessionS

0x14f41e68: 0x7842a2 \_replySessSetupX\_\_13smb\_threadCtxP4msgb+0x5d2

0x14f41f54: 0x778ba9 \_process\_\_13smb\_threadCtxP4msgb+0x611

0x14f41fc4: 0x771cda \_start\_\_13smb\_threadCtx+0x3f6

**Solution:** Apply NAS code to stabilize Servers. In some cases, may need to Unjoin and Join Data Movers back to Domain. In one case, fsn's needed to be un-joined and re-joined to AD Domain & the Hot Fix backed out of AD Servers. A workaround might entail disabling the default FSN Interface, then un-joining and re-joining the other interfaces, followed by re-enabling the default interface. In rare cases this problem has resulted in memory corruption and subsequent file system corruption, requiring fsck.

**MICROSOFT KERBEROS MAXPACKETSIZE LIMITATION FOR UDP:**

Microsoft has a well known limitation of 2000 bytes when transmitting Kerberos information over UDP, resulting in various authentication and GPO issues. Kerberos tickets should use TCP by default if the ticket size is larger than 2kb, which can easily occur because Kerberos tickets carry PAC data.

**DATA MOVER SYMPTOM:** Cannot Join Domain

2004-02-20 14:27:31: ADMIN: 3: Command failed: domjoin compname=mercury domain=aba.ad.abanet.org admin=administrator password=19asee init

2004-02-20 14:27:49: KERBEROS: 3: DomainJoin::getAdminCreds: gss\_acquire\_cred\_ext failed: Miscellaneous failure KRB5 error code 52

**TYPICAL CELERRA WORKAROUND:**

Create temp User account and add only to Domain Admins Group [i.e., delete from Domain Users], then do the Join.

**EMC & MS INFORMATION:**

Primus Articles that address this issue are: emc74207 and emc80211

**Microsoft Articles that address this issue:**

Q-244474 Max Packet Size Limitation when using Kerberos over UDP

Q-320903 addresses a subsequent issue with using Kerberos Over TCP that is fixed in Win2k SP4

**Registry Fix:**

HKLM>System>CurrentControlSet>Control>Lsa>Kerberos>Parameters> Add new value called "MaxPacketSize" and set to "1" to force Kerberos Over TCP

**SMS & WINDOWS 2000 SP4:**

Changes have been made to the way clients and DC's communicate. Clients that use local SMS accounts cause DM to try to authenticate them with AD Server, causing delayed logon performance, as these are not domain accounts. Disable ability of SMS Servers to find DM in MS registry.

**USERS CONNECT TO DM FROM LOCAL CLIENTS WHILE LOGGED IN AS A NON-**

**DOMAIN/LOCAL USER OR SMS ACCOUNT:**

AR30449 identified the problem where Users would log into a local machine domain with a local user account, then attempt to connect to the Celerra. The effects of this were that the DM would try to have the AD server authenticate the local User account & match the local machine's Domain with a known AD domain, resulting in a 20 second timeout, which, when combined with many User logon attempts, would slowdown logons to the Celerra.

**Following new parameters were added to code stream:**

**param NTsec.logonOptions=1** (default) →Check to avoid bad local domain and use DC timeout value

**param NTsec.logonOptions=0** [Prevent Celerra from treating local machine name as a domain but allow local machine User-accounts to log into the Celerra if the Guest Account param is enabled]

**param NTsec.guestLogon=1** (default) →Allow Guest account to logon local Users from local machine domains

**param NTsec.guestLogon=0** [Disable the guest logon feature to the Celerra, especially for Guest accounts]

Fixed in 4.2.19.0, 5.1.16.0, and 5.2.1.0

AR33286 identified a memory leak regression introduced by AR30449 that caused Data Mover panics when local User accounts such as SMS attempted to log onto the Celerra. Fixed in 5.1.18.8, 4.2.22.2, 5.2.1.0

**DEFAULT BUILTIN USER ACCOUNTS IN WINDOWS 2003:**

Administrator

Guest

Help Assistant (used for Remote Assistant sessions)

Support\_388945a0 (used in running signed scripts from Help and Support Services)

**Note:** The default Domain and builtin Guest user account is disabled in Windows 2000, 2003, & XP. Be advised that the account, if enabled, does not have any password assigned.

### **(3) TYPES OF WINDOWS USER ACCOUNTS:**

Local User accounts defined in SAM on local system only

Domain user accounts

Default builtin User accounts

### **DEBUGGING USER LOGON ISSUES:**

**\$server\_config server\_3 -v "param NTsec logonTraces=6"** [=3 to turn off]

**\$server\_config server\_3 -v "pdc trace=1"** [=0 to turn off]

### **PAC DECODING FOR CLIENT AUTHENTICATION:**

With NAS 4.x, Data Mover will use Kerberos authentication field (PAC—Privilege Attribute Certificate) to obtain information about client Users or Groups, as opposed to using LDAP queries to AD for the same information. This more closely emulates native Windows behavior.

Clients use the Kerberos PAC to obtain Kerberos tickets from KDC for the Server that they are connecting to.

PAC Certificates are the Kerberos Tickets used for system/user calls to access network resources. ‘Marshalling’ is a term applied to the manner in which variable data [such as User.name, User.password, User.SID, User.GroupID] is serialized into a data stream for transmission across the network. Unmarshalling is the decoding of this variable data stream on the receiving end and reconstructing the appropriate fields of data. Celerra has had issues with the ‘unmarshalling’ process because the format was not as the DM expected to see, resulting in a filling in of the variable data fields with incorrect data. This would later result in a system crash.

### **MS PRIVILEGE ATTRIBUTE CERTIFICATE (PAC):**

PAC credentials are created by the KDC during an AS request that has been validated by a pre-authentication or by a TGS request from a client that has no PAC and the target resource is a service in the domain or a ticket granting service (referral ticket). PAC information contains a digitally signed User’s logon ID and group membership list, and is unique for the User’s logon session—i.e., a new one would be created only if User logged off domain and logged back on.

PAC is part of the Kerberos V5 authorization data field that is used by the Windows O/S to provide User Logons and to create Access Tokens. The PAC is of C data type with integers encoded in little-endian order. PAC\_INFO\_BUFFERS are structures that hold PAC\_LOGON\_INFO information for the client’s credentials inside the Kerberos ticket. The PAC contains two digital signatures, one using master key of the Resource Server and one using the master key of the KDC.

### **RECENT CODE ISSUES WITH NEW PAC RECORDS:**

Recent PAC changes in Windows 2003 have caused a number of Celerra issues. See emc132863 (AR76179) & emc126516 (AR73457) for KRB and ASN errors seen in Server Logs and network traces, resulting in loss of CIFS access and panics.

#### **EXAMPLE:**

CIFS Server (Default) FILE026[NETWORK]

Full computer name=dm6.network.lan **realm=NETWORK.LAN**

**KDC:** Key Distribution Center—trusted agent in the protocol—holds account information for network users and resources  
KDC runs on every Domain Controller and is started by LSA—provides Authentication Service and Ticket-Granting Service  
Obtains account database from AD

Registered in DNS as as SRV Record (\_kerberos.\_udp.t2dom2.com)

LSA=Local Security Authority—loads the SSP (Security Support Provider) DLL’s which support authentication protocols (Kerberos/NT LAN Manager)

SSP—manages “credential cache” on AD server (stores tickets & keys from KDC)

LTK=Long-Term Key—used to establish security between KDC and “Security Principal” (User)—built during log-on process by passing password through 1-way hashing function

STSK=Short-Term Session Key—used for normal network communications between parties

TGT=Ticket Granting Ticket—a special STSK, usually granted for a 10-hour period

***Note: Can view Client Tickets using the ‘Kerbtray’ Resource Kit Utility!***

### **HOW CLIENTS & KDC SERVER COMMUNICATE DURING USER LOGON PROCESS:**

1. User on Client computer logs on with Username & Password using Kerberos client, which sends an Authentication Service (AS) request to KDC using pre-shared secret known only to Server and Client—AS Service validates the user and then returns a “Ticket Granting Ticket” (TGT) that validates the user’s credentials.
2. Client then requests session to a “service” or “resource” using TGS (Ticket Granting Service) to KDC—ticket for service returned
3. Client can then use this TGS to access the resource or service that it needs

**Note:** Kerberos does not provide authorization to access resources—only authorization to the ‘system’. Celerra uses only Kerberos over UDP Port 88—support for Kerberos over TCP will be added with NAS 5.2

## **HOW WINDOWS CLIENTS LOCATE KDC’s:**

Clients discover DC’s via SRV records after querying DNS. Clients then resolve *ldap.tcp.dc.\_msdcs* SRV records to resolve the KDC Server to contact

**Kerberos consists of (3) sub protocols:** AP--TGS--CS { Application Exchange--Ticket-Granting Service--Client/Server Exchange} NTLM Legacy Authentication for NT and Windows 2000 Mixed Mode Domains--supported by 4.0.12.1

Key Distribution Center for Long-Term Key and Short-term Session Keys

KDC provides Authentication Services and Ticket-Granting Services

Supports UDP port 88 and TCP port 88

**Note :** Celerra only supports Kerberos over UDP—Kerberos over TCP will be supported at NAS 5.2]

## **KERBEROS UDP vs. TCP Packets:**

By default, Windows 2000 Clients use UDP datagrams to Port 88 of KDC Server. Because UDP size is limited to 2000 bytes, there is potential for Domain NETLOGON to fail [Event Log Error 5719]

## **PREVENTING USE OF UDP PACKETS FOR KERBEROS COMMUNICATIONS:**

**HKLM>System>CurrentControlSet>Control>Lsa>Kerberos>Parameters**

To create ‘Parameters’ Key; Edit>Add Value>REG\_DWORD Value=1

## **HOW KERBEROS AND TIME SERVICES INTERACT:**

--Kerberos Client presents Kerberos ticket to Server to prove identity on the network, and includes encrypted authenticator information that contains a secret key and timestamp that proves the presenter’s validity, and prevents replay attacks.

--Server decrypts authenticator and extracts Client’s clock time & checks for allowable time skew against Server time, and also checks that the time is not the same or earlier than another authenticator message, then returns modified authenticator to client

## **CLOCK SKEW:**

Servers check to make sure that client time is within an allowable skew limit [5 minutes by default, or as defined with GPO] before sending authenticator back. If greater than skew limit, return ‘Clock skew too great’ error. Clients can compensate for skew error by adjusting their time used to communicate back to the Server—up to 4 attempts to authenticate are allowed before Kerberos forces a Client to resync its clock, provided that the skew is within the lifetime of the Kerberos ticket.

## **COMMON WINDOWS 2000 KERBEROS ERRORS:**

### **0x6 (KRB\_ERR\_C\_PRINCIPAL\_UNKNOWN) "Client not found in Kerberos database"**

KDC could not translate client principal name into an AD account. AD replication problem--error generated by DC trying to auth.

### **0x7 (KRB\_ERR\_S\_PRINCIPAL\_UNKNOWN) "Server not found in Kerberos database"**

KDC could not translate server principal name into an AD account. AD replication problem--error generated by DC trying to auth.

### **0xE (KDC\_ERR\_ETYPE\_NOTSUPP) "KDC has no support for the encryption type"**

Client does not have proper encryption key; KDC (cross-realm trust) does not have proper encryption key; Also, if Administrator’s account has never been changed, then password change must occur for the MIT-compatible key to be available.

### **0x18 (KDC\_ERR\_PREAMTH\_FAILED) "Pre-authentication information was invalid"**

This indicates failure to obtain ticket, possibly due to the client providing the wrong password.

### **0x25 (KRB\_AP\_ERR\_SKEW) "Clock skew too great"**

#### **When there is a time discrepancy between client and server or client and KDC:**

2003-10-28 12:37:54: SMB: 3: SSXAK=c000006d origin=600 stat=d0000,-1765328347

2003-10-28 12:39:32: SMB: 3: SSXAK=c000006d origin=600 stat=d0000,-1765328347

From Kerberos web site, this means: 37 -1765328347 Clock skew too great

<http://www.net.berkeley.edu/kerberos/k5msgs.html>

"KDC has no support for the encryption type": Client lacks correct encryption key type

"Pre-authentication information was invalid": Client failed to obtain KDC ticket, possible due to wrong password

"Clock skew too great": Time or Timezone discrepancy client/server or client/KDC

## **USING SERVER KERBEROS COMMANDS TO VERIFY KERBEROS:**

**\$ server\_kerberos server\_2 -l**

Kerberos realms configuration:

realm name: NETWORK.LAN

KDC: util002.network.lan

admin server: util002.network.lan

default domain: network.lan

kpasswd server: util002.network.lan

## **\$ server\_kerberos server\_2 -keytab**

Dumping keytab file  
keytab file major version = 5, minor version 2  
realm: NETWORK.LAN, principal: host, host: FILE014,  
    principal type 1, key version: 1, encryption type 3  
realm: NETWORK.LAN, principal: host, host: DM2.NETWORK.LAN,  
    principal type 3, key version: 1, encryption type 3  
realm: NETWORK.LAN, principal: cifs, host: FILE014,  
    principal type 1, key version: 1, encryption type 3  
realm: NETWORK.LAN, principal: cifs, host: DM2.NETWORK.LAN,  
    principal type 3, key version: 1, encryption type 3  
realm: NETWORK.LAN, principal: user, host: dm2,  
    principal type 1, key version: 1, encryption type 3  
End of keytab entries.

## **USING KERBEROS COMMAND TO ADD SPECIFIC KERBEROS REALM TO DM:**

**\$server\_kerberos server\_2 -add**

**realm=T2DOM2.LOCAL,kdc=win2k.t2dom2.local, domain=t2dom2.local**

**Comment:** In certain situations, Data Mover may not be able to “Join” AD Domain because the Kerberos info is different than the AD Domain name. Use this command to specify the correct “kdc” domain suffix.

**Symptoms:** DM seeks krbtgt for certain domain suffix but receives “Server not found in Kerberos database” message.

## **MICROSOFT KERBEROS TROUBLESHOOTING TOOLS:**

**Kerbtray:** Displays ticket information, purges Tickets

**Klist:** CLI tool to display or purge Kerberos tickets and tgt

**Setspn:** List, add, delete SPNs objects

**W32tm:** Use to verify client time with DC, Timezone, etc.

**Kerberos Logging:** Add the following registry value

HKEY\_LOCAL\_MACHINE\SYSTEM\CurrentControlSet\Control\Lsa\Kerberos\Parameters

Registry Value: LogLevel

Value Type: REG\_DWORD

Value Data: 0x1

## **MICROSOFT SUPPORT TOOLS:** From Server CD>Support>Tools

**Adsedit.msc:** Useful GUI to observe and edit Active Directory OU's, Containers, and Objects

Domain NC → OU's

Configuration Container → CN's

Schema → AD Objects

**Ldifde Utility for AD Dumps :** Used to import or export all or specific portions of Active Directory into text file [LDAP Data Interchange Format]

**Ldifde -f c:\usersdump.txt -d CN=Users,DC=buzzsaw,DC=com** [dump Users container only]

**Ldifde -f domain.ldf -d "DC=buzzsaw,DC=com" -s 10.76.10.160** [dump entire domain partition]

**repadmin Utility for AD Replication:** Use to replicate AD Servers

**run: repadmin /syncall /e /d /P <adservername>**

## **NTP SERVICE (RFC1305):**

--Required for successfully implementing a CIFS 2000 Server in a WINDOWS 2000/2003 Active Directory Domain

--NTP Service Version 2 [Up to (4) NTP servers can be specified]

--Default clock skew of 5 minutes allowed to retain Kerberos ticket validity, or tighter tolerances if using GPO Policies

--W32Time is default SNTP Service (RFC 1769) on Windows 2000 based on UTC (Coordinated Universal Time) atomic time scale

--W32Time meets NTP requirements for W2k3/XP

--W32Time synchronizes every 45 minutes until synch successfully 3 times, then once every 8 hours thereafter

--Use an external Time Service if needing tighter time sync specs, or for PDC Emulator service at domain root, or child domains, etc.

## **SNTP:**

RFC 1769 was used by MS to implement W32Time for Kerberos V5 authentication and is called “loose synchronization”. SNTP protocol guarantees clocks within 20 secs of each other in the enterprise, or 2 secs between sites.

**Best Practice:** Use the AD Server as the NTP Server [Most AD's are configured as such by default]

**Note: Celerra uses SNTP for time synchronization within fractions of a second—as opposed to the NTP Implementation which is within milliseconds**

**Automatic NTP Detection:**

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
If not manually configured, Celerra datamovers will automatically detect & synchronize with up to (4) NTP sources  
Normal time corrections between local clock and NTP clock are done by “clock slewing”—time is gradually increased or decreased to bring back within synchronization. Normal slew rate is 10%, or about (6) minutes to correct an hour’s time difference

### **CELERRA NTP COMMANDS:**

```
$server_date server_x timesvc  
$server_date server_x timesvc start ntp 192.10.4.5  
$server_date server_x timesvc set ntp      [Immediately synchronizes with NTP server]  
$server_date server_x timesvc update ntp  [begins updating system time to NTP]  
$server_date server_x timesvc stats ntp  
$server_date server_x timesvc delete ntp
```

### **CELERRA DATABASE FILES: DNS/NIS/NTP**

**/nas/server/slot\_x/misc**

#### **DYNAMIC DNS: Default mode for Windows 2000**

--A FQDN must be defined on DM and added to "Enterprise Admin Group" in order to use LDAP for credentialing

--Namespace based on DNS--root Domain Controller is also root DNS server for domain

--DNS updates occur whenever CIFS is started and at regularly scheduled times

#### **WINDOWS 2000 DCLIST BUILD BEHAVIOR FOR CELERRA:**

When joined as a W2K Server, Celerra uses DNS exclusively to get the list of Domain Controllers to use & sorts the list based on LDAP ping response time, site information, and subnet information. Param is enabled by default—when the param is disabled, Celerra will use the list obtained from WINS to build its DC list:

```
# server_param server_2 -f cifs -info NTsec.getDCfromADServices -v
```

server\_2 :

```
name          = NTsec.getDCfromADServices  
facility_name = cifs  
default_value = 1  
current_value = 1  
configured_value =  
user_action    = restart Service  
change_effective = restart Service  
range         = (0,1)  
description   = Use DNS to get list of Domain Controllers
```

#### **DNS is a distributed database that provides name & service resolution services for datamovers:**

1. Mapping machine names to IP addresses    2. Mapping protocols    3. Mapping NT Domain names for client applications

4. Mapping IP Addresses to hostnames [Reverse Zone Lookups]

#### **DNS RULES:**

→DNS protocol runs over TCP or UDP Port 53[UDP is the default] and can point to (3) DNS Servers per DOMAIN for the datamover. DNS uses UDP by default, but if response is greater than 512 bytes, a TC bit is set and the Resolver knows to query again, this time using TCP.

→Default timeout is 60 secs if first DNS Server does not respond.

#### **Configure more than one DNS domain if the following conditions exist:**

--CIFS Services with DNS Domains that are in different forests

--CIFS Services with different DNS servers

```
$server_dns server_4 -protocol tcp t2dom2.local 192.10.3.2
```

--CIFS 4.0 Server will first try to resolve computernames by DNS, then NETBIOS Broadcast, then WINS [if configured]

--CIFS 4.0 Server also supports resolving Windows Service Names via DNS [such as LDAP, Kerberos, etc.]

--Dynamic DNS allows for automatic registration/un-registration of CIFS 4.0 Server—Secure Dynamic Updates are the default

**Note: 4.0.12.1 works when "Allow Dynamic Updates?" is set to "Yes"**

--With successful JOIN, the 'COMPNAME' is created by default in the "ou=Computers,ou=EMC Celerra" organizational unit (AD)

--CFS Servers retrieve list of supporting LDAP, Kerberos, & other services from the DNS Server

--Zones are stored in the AD database if using the “Integrate DNS with Active Directory” mode [not the default mode]

--Windows 2000 FQDN can only be 15 characters in length

#### **TURNING OFF DYNAMIC DNS FOR SERVER:**

Why do this? Data Mover automatically registers NetBIOS/COMPNAMES on all interfaces to all DNS Servers in its list--might be desirable to configure DNS Servers with A & PTR records manually, especially if DNS lookups are causing performance problems due to slow Name Resolution.

```
param dns updateMode=0  [Turns off dynamic DNS and introduces Static DNS behavior]
```

## **DNS-RELATED UTILITIES:**

```
C:>netdiag /v /l /test:dns [testing individual Servers]  
C:>dig <ip of NS> <resource record name> <type record>  
C:>dnslint  
C:>dnscmd 10.241.169.16 /zoneinfo mouse.com
```

## **MICROSOFT DNS:**

As long as DNS is BIND (Berkeley Internet Name Domain) 8.1.2 compliant, can be used with Windows 2000 (8.1.2 meets RFC2052 spec for Service Resource Records & RFC2136 for dynamic dns updates). ISC BIND 8.2.4 will run on NT 4.0 or Windows 2000. Microsoft DNS is based on BIND standards.

### **DNS BASICS:**

DNS is a distributed database similar to an inverted ‘tree’—the ‘root’ of DNS is at the top & is represented by the ‘null’ zero-length label “”, or more commonly by a dot “.” in Unix. Windows represents ‘root’ with a backslash \. In DNS, an ‘absolute name’ is synonymous to a FQDN. FQDN’s are read from left-to-right as Node→Subdomain→Parent Domain→TopLevel Domain→Root. Top-level or first-level domains are known as .com, .edu, .net, etc. Domains beneath top-level domains are called sub-domains or child domains.

## **TROUBLESHOOTING DNS ISSUES/DNS BEST PRACTICES:**

→Ensure that clients are pointing to correct DNS Server or they will not be able to use their Locator service to find SRV Service records that advertise DC services, such as Kerberos, LDAP, and Global Catalog—with the SRV records, clients will not be able to authenticate and operate within the domain. DC’s use their Netlogon service to register SRV records

→Default TTL for cached DNS entries on the DNS Server is one hour

→TCP/IP Properties, ensure DNS Resolver Suffix is correctly chosen for the intended purpose:

**Primary DNS Suffix:** Obtained from the TCP settings defined for the local host, appended to the local host’s ‘flat’ name

**Append Parent Suffix:** If the ‘Append primary and connection specific DNS suffixes’ option is checked, a DNS query will use first the Primary Suffix, then the Parent Suffix

### **EXAMPLE:**

ts2.emc.com →Primary suffix

emc.com →Parent suffix

**Interface Suffix:** With this setting checked, means that a DNS query will use first the Primary Suffix, then any unique suffixes that might be defined in the TCP settings [Static or DHCP-provided], then the Parent Suffix

**Search Table:** If the ‘Append these DNS suffixes’ is checked, the local DNS Resolver will not use the Primary, Connection-Specific, or Parent Suffixes, rather, will use whatever suffixes are defined in the table, in order from the top to bottom entry. If using this method, ensure that local FQDN is defined first.

→Use DNS Forwarding to reduce recursive DNS queries. During normal operation, if a client submits a request for a resource record from an outside domain, the DNS Server will assume responsibility for contacting a nameserver from the target domain to submit the query and provide the results—recursive queries. But, recursive queries can tie up Server resources, so using a Forwarder to conduct the recursive work for the Client is better. Even if using AD-integrated DNS, DNS forwarding information needs to be defined on the Forwarder as this is a local Registry entry, not an AD value.

→When configuring Zone Transfers between DNS Servers, allow transfers only to specified Name Servers, and ensure that Port 53 is not firewalled, as Zone Transfers use TCP ports 53 and ports above 1023

→When configuring AD-integrated zones, make sure all Secondary Zones are removed from DC’s & dns service is restarted

→Clients register their A & PTR records with Start of Authority (SOA) DNS Server (With AD-integrated DNS, any DC running DNS can update a Zone record)

→Improperly Delegated Child Zones can lead to ‘lame delegation’ and ‘replication’ issues

AD identifies Domains and DC’s in DNS via CNAME records correlated to GUID & FQDN

With delegation, Parent DNS Zone contains A Records and Names of Child DNS Servers—if the Child DNS Server is offline, then lame delegation and replication issues can occur. Check Event Logs for RPC connection errors and KCC problems. Win2k3 Servers use Stub DNS Zones to help avoid lame delegation issues.

## **CONFIGURATION CHECKING UTILITY: dnsllint**

C:>**dnsllint** is a utility that can test DNS configurations to see if they correctly support an AD domain, and also to test if DNS configuration meets standard practices for zones

→As a Best Practice, do not use your public DNS server as a Forwarder. Instead, install a Caching-Only DNS Server in the DMZ and use this as the Forwarder, and Disable Recursion for all Public DNS servers

→Best practice, if using BIND-style DNS Servers, do NOT allow dynamic updates. Only allow Dynamic Updates if using AD-integrated DNS Zones, then permit only SECURE updates

## **NAMESERVERS:**

NameServers are used to administer the DNS ‘namespace’ in subdomains called “zones”. NameServers are ‘authoritative’ for Single or Multiple Zones if they’ve been given ‘delegated’ authority. For example, a Parent Windows 2000 Domain called “gm.com”

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
[Windows 2000 Domains by default are DNS Domains] is authoritative for “gm.com” but might not be for “sub-domains” or ‘child domains’ such as “corp.gm.com”, which would have its own authoritative NameServer while the “gm.com” NameServer would contain pointers to it. If a NameServer is Authoritative for a Zone that it receives a Query for, it will either provide the data or return an error that the requested data does not exist. If not Authoritative, it would use Iterative Queries to other NameServers to obtain answer.

## **DNS ZONES:**

Zones contain all domain names for a Domain for which it is ‘authoritative’, represented by the SOA Record (Start of Authority). Every DNS Zone has at least one NameServer that is authoritative & one SOA record. Domain names and data that are located in subdomains that have been given their own ‘delegated authority’ (SOA) will not be found in the parent domain. NameServers load DNS Zones as either a Primary or Secondary Servers. A ‘Zone Transfer’ means that a Primary DNS Server provides a Secondary with dns records upon request.

>**ls t2dom2.local** [Conducting a Zone Transfer using nslookup]

**Forward Lookup Zones:** Contain “A” Address Records. A Records provide resolution for “names” to its corresponding IP Address.

**Reverse Lookup Zones:** Contain “PTR” Pointer Records. PTR Records provide resolution for “IP Addresses” to a ‘name’.

## **TYPICAL ZONE RECORDS:**

SOA Record—one per Zone to indicate that it is authoritative for its Zone

NS Record—one per Zone to identify the NameServer for the Zone

A & PTR Records

CNAME—Canonical Names or “aliases” –are aliases to the Address or Host Record.

## **PRIMARY vs. SECONDARY NAMESERVERS:**

Primary NameServers load the DNS Zone directly from local zone files. Secondary NameServers obtain their Zone records from the Primary via a “zone transfer” over the network. The purpose of Secondaries are for adding redundancy, spreading out the load, and also to have a DNS server as close to its Client systems as possible.

**Important Point:** Secondary NameServers are also authoritative for the Primary Zone.

## **DNS NAME RESOLUTION:**

Primary or Secondary NameServers provide data about the Zones for which they are authoritative. When resolving queries for data to which they are not “authoritative”, the process is called “name resolution”. In DNS, there are two major types of Queries from Client “Resolver” systems.

### **RECURSIVE DNS QUERY:**

Recursive queries instruct the NameServer to return an authoritative answer, or query Remote DNS Servers until it obtains an answer, or returns an error message. If the local DNS Server does not contain the authoritative data being requested, it will then most likely contact other DNS Servers using Iterative Queries until it finds the right answer for the Client.

### **ITERATIVE DNS QUERY:**

Local DNS Server will consult its own database first, then provide a list of NameServers to the Client ‘Resolver’ as a ‘referral.’

### **I18N SUPPORT:**

MS DNS supports all characters from UTF-8 Unicode character set

### **Native Windows 2000 Clients Startup Process:**

During startup & logon, native WIN2K clients use DNS to locate LDAP & Kerberos services, obtain address of DC, and registers its Hostname & IP Address with DNS.

### **DNS UPDATE MODES FOR WIN2K:**

Celerra works with "Secure Updates Only" [Default Mode]; 'Allow Dynamic Updates'--"Yes", or "No"[Insecure & No Updates].

**Note:** DM will automatically use Secure Updates if configured on DNS

### **DEFAULT DNS IS SECURE UPDATES:**

To configure Insecure or No Updates, use following params:

**param dns updateMode=0 [No Updates]**

**param dns updateMode=1 [Insecure Updates]**

**param dns updateMode=2 [Secure updates—default]**

## **REFRESHING DNS RECORDS ON AD SERVER:**

**Command Prompt: c:>ipconfig /refreshdns**

Admin Tools>DNS: Action>Update Server Data Files

## **QUERYING DNS RECORDS ON WINDOWS HOST—“Resolver”:**

**C:>ipconfig /displaydns                   c:>ipconfig /flushdns**

## **DATAMOVER DNS CACHE:**

**\$server\_config server\_8 -v “dns dump”**

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
**Note:** Outputs dump of datamover DNS cache. Useful for displaying DC A Records & SRV records for LDAP, KERBEROS, KPASSWD

**Note:** Standard Unix implementations use cache on the Server, not the Resolver Client system

## **SETTING UP DNS CLIENT ON CONTROL STATION:** Facilitates DNS testing of Windows environment

Step 1. #vi /etc/resolv.conf file and add DNS Domain

```
domain t2dom3.local  
search t2dom3.local  
nameserver 192.10.4.4
```

Step 2. Use #nslookup server\_x [verifies Server DNS name resolution and IP Address]

**Note:** Other tools are "dig"; dnswalk; etc

**Comment:** Other files to check—

```
# cat /etc/nsswitch.conf |grep hosts  
hosts: db files nisplus nis dns  
hosts: files nisplus nis dns [order of host resolution—i.e., /etc/hosts, nisplus, nis, dns]
```

## **VERIFYING HOST NAME RESOLUTION:**

```
#nslookup 192.168.51.22
```

```
Server: 192.168.26.10  
Address: 192.168.26.10#53  
22.51.168.192.in-addr.arpa name = .  
22.51.168.192.in-addr.arpa name = ogcovsrv01.ocfl.net.
```

```
# /usr/bin/host 192.168.51.22
```

```
22.51.168.192.in-addr.arpa domain name pointer .  
22.51.168.192.in-addr.arpa domain name pointer ogcovsrv01.ocfl.net.
```

## **COMMON WINDOWS 2000 DNS PROBLEMS:**

--Domain Controller is not pointing to itself for DNS resolution in its TCP/IP Properties>DNS Tab page

--A “root” [ . ] Zone exists in the Forward Lookup Zone [Having this set as a “Root” DNS Server will prevent Internet lookups]

**Solution:** Do not have a “. .” Zone unless only used for a closed network

--Ensure that Client Systems point to the DC DNS Server for DNS Name Resolution

--‘Disjointed Namespace’—when the DNS Domain Suffix differs from the Windows 2000 Domain Suffix

--Incorrectly configured Child Domains:

**Solution--Configuring the Parent DNS Domain with a Child DNS Domain:**

→Create Delegation Record on Parent DNS Server for Child DNS Server>Create Secondary Zone of Child DNS Server & configure to transfer from the Parent DNS Zone>Point Secondary Child DNS Server to itself only

## **WINDOWS 2003 ISSUE WITH CELERRA:**

If aliases are being used with W2K3, as set with SPNs, the KDC will include the alias name in the Kerberos ticket instead of the real compname, and the Data Mover will reject the client connection. W2K does not behave like this and is fine. A NAS workaround is built into 5.5.29.x, AR92558, by setting the following param, the name in Kerberos ticket does not need to match real server name, when checking Kerberos keys during authentication:

```
param cif's LanmanServer.disableNameChecking=1
```

## **DNS SRV RECORDS:**

**Intro:** Perhaps the single most important feature in making DNS & Active Directory function smoothly together is the registration of “SRV” Service Records. If the AD DNS Server points to its own IP Address in TCP/IP Properties for DNS, upon startup the “NETLOGON” Service will register SRV records and create the following (4) Records in the Forward Lookup Zone:

```
__msdcs  
__sites  
__tcp  
__udp
```

→Records contained within these folders are vital for Windows 2000 [CFS 2000 Servers] to register themselves and query AD!

## **LOCATION OF SRV RECORDS ON DC:**

c:\winnt\system32\config\netlogon.dns

## **VERIFYING REGISTRATION OF DNS SERVICE (SRV) RECORDS ON AD SERVER & HOST “RESOURCE RECORDS” (A & PTR) FOR CELERRA USING NSLOOKUP:**

**Comment:** There may be a need to ensure that Windows 2000 Services & Host Resource Records are properly registered with DNS. One tool for verifying these records is “nslookup”. Other tools can also be used, such as the Unix “dig” command.

**Change Update:** **NAS 4.2.6.x no longer requires that Celerra use the Reverse Lookup Zone “PTR” Record. Only the “A” Record for Hostname & IP in the Forward Lookup Zone is required.**

### **DNS RESOURCE RECORDS:**

Each host system on the network requires a minimum of (2) Resource Records for integration as a Active Directory client for Windows 2000 Domains.

#### **Forward Lookup Zones:**

**A=Address Records**—these records resolve Hosts by converting “Host Names-to-IP Addresses”

Every Host must have an “A” record for proper DNS registration and Name Resolution.

**nas.us.dg.com →**      **Name**      **Type**      **Data**  
                              cpc233120      A      10.240.16.113

#### **Reverse Lookup Zone:**

**PTR=Pointer Records**—these records resolve Hosts by converting “IP Addresses-to-Host Names”

**16.140.10.in-addr.arpa →** **Name**      **Type**      **Data**  
                              113      PTR      cpc233120.nas.us.dg.com

## **USING NSLOOKUP TO VERIFY IF AD SERVICES ARE REGISTERED:**

**#nslookup**

**>set type=SRV**

**> \_ldap.\_tcp.dc.\_msdcs.nas.us.dg.com**      [Shows DC List, IP Addresses, Port 389]

Server: cpc233120

Address: 10.240.16.113

\_ldap.\_tcp.dc.\_msdcs.nas.us.dg.com      SRV service location:

    priority      = 0  
    weight      = 100  
    port      = 389  
    svr hostname      = cpc233120.nas.us.dg.com

cpc233120.nas.us.dg.com internet address = 10.240.16.113

### **VERIFYING LDAP SERVICE IN DNS:**

**> \_ldap.\_tcp.nas.us.dg.com**      [Shows LDAP Servers]

Server: cpc233120

Address: 10.240.16.113

\_ldap.\_tcp.nas.us.dg.com      SRV service location:

    priority      = 0  
    weight      = 100  
    port      = 389  
    svr hostname      = cpc233120.nas.us.dg.com

cpc233120.nas.us.dg.com internet address = 10.240.16.113

### **VERIFYING KERBEROS SERVICE IN DNS:**

**> \_kerberos.\_udp.nas.us.dg.com**      [Shows list of Kerberos Services & Port 88]

Server: cpc233120

Address: 10.240.16.113

\_kerberos.\_udp.nas.us.dg.com      SRV service location:

    priority      = 0  
    weight      = 100  
    port      = 88  
    svr hostname      = cpc233120.nas.us.dg.com

cpc233120.nas.us.dg.com internet address = 10.240.16.113

### **VERIFYING KERBEROS PASSWORD SERVICE IN DNS:**

**> \_kpasswd.\_tcp.nas.us.dg.com**      [Shows list of Password Servers Port 464]

Server: cpc233120

Address: 10.240.16.113

\_kpasswd.\_tcp.nas.us.dg.com      SRV service location:

    priority      = 0

```
weight      = 100
port       = 464
svr hostname = cpc233120.nas.us.dg.com
cpc233120.nas.us.dg.com internet address = 10.240.16.113
```

**VERIFYING PTR RECORD OF DATAMOVER:** [Also called “Reverse Lookup Zone Record”]

```
#nslookup -sil
```

> **set type=PTR**

> **10.240.16.113**

```
Server: cpc233120
Address: 10.240.16.113
113.16.240.10.in-addr.arpa name = cpc233120
```

**VERIFYING DATAMOVER A RECORD IN DNS:** [Also known as Host Record for Forward Lookup Zone]

> **set type=a**

> **10.240.16.113 [or Hostname ‘cpc233120’]**

```
Server: cpc233120
Address: 10.240.16.113
Name: cpc233120
Address: 10.240.16.113
```

**USEFUL WINDOWS RESKIT TOOLS OR UTILITIES:**

**TROUBLESHOOTING ACTIVE DIRECTORY: Q247811**

**TROUBLESHOOTING, REPAIRING, TESTING “SRV” RECORDS IN DNS FOR AD SERVER:**

**Note:** Quickest way to rebuild an AD Server’s DNS SRV records is to stop its netlogon service, then dns, restart Netlogon & dns  
Use Microsoft WIN2K Resource Kit tools to troubleshoot and fix: c:>dcdiag.exe & netdiag.exe

**C:>netdiag.exe /fix** [This utility helps re-register the AD Server’s SRV Records in DNS]

**C:>dcdiag.exe /fix** [Utility to analyze state of domain controllers in a forest or enterprise and reports problems. Also used to help repair Domain Controller records, problems, etc.]

**C:>nlttest /dsgetdc:win2kdomain** [Verify that Domain Controller can be located for specific domain]

**C:>nlttest /dsgetdc:example.com /force**

**C:>nlttest /dbflag:0x2000ffff** [Turns on Debug Log in AD Server’s “Netlogon” directory>netlogon.log]

**C:>ldp.exe** [Use this utility to connect & bind to DC to verify LDAP connectivity]

**C:>netdom.exe /network.lan:network.lan** [Use to verify Trust relationships]

**C:>setspn.exe -l server\_x** [Lists Service Principal Names of AD Object]

**C:>setspn.exe -r server\_x** [Resets SPN to default; Use -a to add new SPN]

**#nslookup**

**>movername.umb.umbbank.com** [Use to verify that DataMover is registered correctly with DNS]

**>guid.\_msdcs.umbbank.com** [Use to verify that GUID SRV record for Parent Domain can be resolved]

**DCDIAG TOOL:**

Verifies connectivity, replication, topology, DS partitions, DS records, Inter-site health, External trusts only  
dcdiag /s:SERVER1 /c /v [Use these switches to verify whether critical DC Services are running on the AD Server]

**NETDOM UTILITY:**

Use this tool to verify Kerberos v5 Trusts

**Symptoms:**

Celerra Server cannot obtain SID after successful “Join”

**Server Log Indicates SRV Records Not Found:**

2002-07-11 20:16:56: KERBEROS: 3: krb5\_locate\_srv\_dns: DNS returned no KDC for service \_kerberos at realm DSUNET.COM

**NSLOOKUP Fails to Find SRV Records:**

**C:>nslookup**

> **set type=SRV**

> \_ldap.\_tcp.dc.\_msdcs.nas.us.dg.com [Shows DC List, IP Addresses, Port 389]

**Note:** In this example, no SRV record is returned for the LDAP service

**SETSPN [ServicePrincipalName] TOOL:**

Win2k Reskit tool to locate target principal name for a service—“Service Principal Names” of an AD Object

**C:>setspn -l pc10corp**

Registered ServicePrincipalNames for CN=PC10CORP,OU=Laptops,OU=US Clients,

DC=corp,DC=emc,DC=com:

HOST/ PC10CORP

HOST/ PC10CORP.corp.emc.com

**Note:** Details of ServicePrincipalName for computername “pc10corp”

**C:>setspn -r pc10corp**

**Note:** Resets default SPN for host “pc10corp”

**C:>setspn -a http/pc10corp pc10corp**

**Note:** Sets a new SPN for “pc10corp”

#### **NAS SYSTEM:**

C:\Program Files\Support Tools>setspn -l mview\_dm2

Registered ServicePrincipalNames for CN=mview\_dm2,OU=Computers,OU=EMC Celerra,DC=2k3,DC=pvt,DC=dns:

cifs/mview\_dm2.2k3.pvt.dns

cifs/mview\_dm2

host/mview\_dm2.2k3.pvt.dns

host/mview\_dm2

### **USING LDIFDE UTILITY TO VIEW ACTIVE DIRECTORY DATABASE OBJECTS:**

**C:>\Ldifde -f domain.ldf -d “DC=buzzsaw,DC=com: -s 10.7.10.160**

**Note:** Dumps the entire AD partition from the AD Server indicated to a text file

Ldifde -f c:\domaindump.txt -d DC=buzzsaw,DC=com →dumps domain

Ldifde -f c:\usersdump.txt -d CN=Users,DC=buzzsaw,DC=com →dumps Users container

#### **SUBINACL.EXE:**

Useful tool to display security information on files or services, change ownership on objects, replace security from one SID to another, migrate security information on objects, modify any part of ACL.

#### **SUBINACL SYNTAX:**

c:>subinac /noverbose /output=filename subinac command

**Example:** To backup NTFS ACLs on all files on c:\ drive

**c:>subinac /noverbose /output=c:\aclbackups.txt /file c:\\***

→ Subinac /noverbose & /output parameters allows for snapshot of ACLs on objects--the /playfile option lets you restore ACLs.

→/output param will report on ACLs and dump to file

→/playfile will restore ACLs from a previous backup

**c:>subinac /playfile c:\aclbackups.txt**

### **USING NLTEST UTILITY:**

**C:>nltest /DOMAIN\_TRUSTS**

List of domain trusts:

0: ASSDPS (NT 4) (Direct Inbound)

1: ZAPPA (NT 4) (Direct Inbound)

2: ASD (NT 4) (Direct Inbound)

3: PSOLAB2 psolab2.emc.com (NT 5) (Direct Inbound)

**C:>nltest /DCLIST:corp**

Get list of DCs in domain 'corp' from '\\CORPMAHO1'.

gen1dc1.corp.emc.com [DS] Site: CorpEUGENuernberg1

bebr1dc1.corp.emc.com [DS] Site: CorpEUBEBrussels1

uksy3dc1.corp.emc.com [DS] Site: CorpEUUKLondon1

**C:>nltest /SC\_QUERY:corp**

Flags: 30 HAS\_IP HAS\_TIMESERV

Trusted DC Name \\corpmawe1.corp.emc.com

Trusted DC Connection Status Status = 0 0x0 NERR\_Success

**C:>nltest /dcname:corp**

PDC for Domain corp is \\CORPMAWE2

The command completed successfully

### **TROUBLESHOOTING DIRECTORY SERVICE INFORMATION:**

**\$server\_config server\_8 -v “ds dump”**

**Note:** Useful for showing CIFS FQDN, DOMAIN GUID, FOREST name, SITE name, DOMAIN name, capabilities of Domain Controllers for each AD Server, CIFS Services information, etc.

### **CODE CHANGES AFFECTING CELERRA & ACTIVE DIRECTORY INTEGRATION:**

**NAS 4.0.12.1:**

- DM will use first DC, LDAP, or KDC server returned by DNS without further processing
- Required “Enterprise Admins” rights in multi-domain environments for Join

**NAS 4.0.16.100:**

- Uses DC on same subnet first, then first returned from DNS list
- Still required “Enterprise Admins” rights in multi-domain environments for Join

**NAS 4.0.18.0, 4.1.6.0, 4.2.4.1:**

- No longer requires “Enterprise Admins” account to Join multi-domain environments

**NAS 4.2.6.1/5.0.9.2:**

- Now locates DC Services in same manner as Windows 2000 Client
- No longer requires Reverse DNS Entry to validate the DM Host Record

## **CONFIGURING DATAMOVERS TO USE SPECIFIC AD SERVERS (DSFile):**

**Note:** The 'directoryservices.tmp' file represents the basic outline of the AD Servers & Services in use on the network, as seen by the DM. Edit the file as necessary, deleting AD Servers/Services that are not desired--retaining DC's and Services that are desired. This option was introduced with NAS 4.2.10.x. You would need to execute this file for each domain in multi-domain environments & build a list of preferred DC's in the directoryservices file.

1. Create the Directory Services File by executing the following command:

**\$ .server\_config server\_4 -v "ds domain=ts2.mouse.com makeDsFile"** [Use cmd to create DS File in .etc]

**Note:** Creates a file called “directoryservices.tmp” in the ./etc of the DM:

-rw-r--r-- 1 root bin 3140 Apr 5 10:07 directoryservices.tmp

1081174049: SMB: 4: The list of services is stored in file ./etc/directoryservices.tmp.

2. After editing the directoryservices.tmp file to reflect the AD Servers & Services desired, rename as “directoryservices”

3. Add following parameters to data mover slot:

**param ds useDSFile=1** [Enables the use of the 'Directory Services' file by the DM]

**param ds useDCLdapPing=0** [Disables default behavior of using DC's sorted by Site & LDAP Ping response times]

4. Reboot Server

**Note:** By using this file, DART will access only DC's for Services configured here, and in the order listed. For Services not listed here, DART will use DNS.

## **EXAMPLE OF AN ‘ABBREVIATED’ CONFIGURATION FILE: ./etc/directoryservices**

#Domain: alpa.com – Domain Controller: alpadc1

```
_ldap._tcp.dc._msdcs.alpa.com 0 100 389 alpadc1.alpa.com  
_kerberos._udp.alpa.com 0 100 88 alpadc1.alpa.com  
_kpasswd._udp.alpa.com 0 100 464 alpadc1.alpa.com  
_kpasswd._tcp.alpa.com 0 100 464 alpadc1.alpa.com
```

**Note:** Configure the above services for each Domain Controller to be used by Data Mover

**EXAMPLE ‘DIRECTORIEServices’ FILE**—Caution: File abridged—do not use as real file!

**\$ cat /nas/rootfs/slot\_4/.etc/directoryservices.tmp**

```
#  
# File ./etc/directoryservices.tmp  
#  
# This file is a temporary file created by the Dart.  
# It contains service descriptions for the services: ldap, Kerberos,  
# and kpasswd, which are provided by a Windows 2000 domain/realm.  
#  
# For each Windows 2000 Domain, modify the list of Domain Controller(s) to  
# keep the servers the Data Mover will access and are geographically close.  
#  
# For each Windows 2000 Domain Controller you can create manually other entries, one for each of the following services:  
# _ldap._tcp.dc._msdcs  
# _kerberos._udp  
# _kpasswd._udp  
# _kpasswd._tcp  
#  
# For the configured services, Dart will try to access only the Domain Controllers  
# configured and in the order the Domain Controllers have been configured.  
#
```

```
# When trying to access a Domain Controller for a service NOT configured,
# Dart will use regular DNS mechanisms to locate the Domain Controllers that provides
# the service.
#
# Each service is defined with the syntax below:
#
# <service_fqdn_path> <priority_value> <weight_value> <port> <domain_controller_FQDN>
#
# Where:
#   <service_name> : full DNS path of the service, in the form: service_name.server_fqdn
#           Example: "_kerberos._udp.mydomain.com"
#
#   <priority_value> : The priority of the server.
#           Dart attempt to contact the server with the lowest priority
#
#   <weight_value> : Used for load balancing. Use the same value for all the
#           Domain Controllers in a Windows domain.
#
#   <port> : TCP/UDP port where the service listens for connections (decimal)
#           "_ldap"   listens on port 389
#           "_kerberos" listens on port 88
#           "_kpasswd"  listens on port 464
#
#   <domain_controller_FQDN> : fqdn of the Domain Controller that provides the service
#
# All numeric values are decimal.
#
# Lines starting with "#" are considered as comment.
#
# NOTE: Dart 'param'
# =====
# To activate service lookup using this file,
#   - rename this file in /etc/directorieservices
#   - set the 'param' ds.useDSFile to 1
# In the file /nas/server/slot_x/param , add:
#   param ds useDSFile=1
#
# Reboot the Dart using the 'server_cpu' command
# =====
# Services for domain ts2.mouse.com
#
# Domain controller(s):1
_ldap._tcp.dc._msdcs.ts2.mouse.com 0 100 389 minnie.ts2.mouse.com
#
# Global catalog server(s):0
# Kerberos server(s) over UDP protocol:1
_kerberos._udp.ts2.mouse.com 0 100 88 minnie.ts2.mouse.com
#
# Kerberos server(s) over TCP:1
_kerberos._tcp.ts2.mouse.com 0 100 88 minnie.ts2.mouse.com
#
# Kerberos admin server(s) over TCP protocol:0
# Kerberos password change server(s) over UDP protocol:1
_kpasswd._udp.ts2.mouse.com 0 100 464 minnie.ts2.mouse.com
#
# Kerberos password change server(s) over TCP protocol:1
_kpasswd._tcp.ts2.mouse.com 0 100 464 minnie.ts2.mouse.com
# -----
```

## WINDOWS SMB PROTOCOL:

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
Resource-sharing protocol used in client-server operations to access files, printers, mail slots, named pipes, and API's.  
Current version is known as CIFS (Common Internet File System)--slightly different than NT LM 0.12—and is the native file sharing protocol for Windows 2000. Two Windows 2000 computers use NTLM 0.12, which allows exchanges to use Unicode for file names,, resource, and user names.

**SMB Commands fall into four categories:** Session Control; File Commands; Print Commands; Message Commands

Supports Two Levels of Security:

Share Level: password

User Level: files, user access and rights

#### **How does the Protocol Work?**

CIFS uses commands packaged via “redirection” requests to remote computers or the local computer. Connection establishment messages start & end “redirection” connections to remote resources. Namespace & File Manipulation messages are used to gain access to files and to read & write. Printer messages are used by redirection to send data to a print queue. Other messages send redirection requests to mail slots and named pipes.

#### **MICROSOFT RPC:**

- RPC is a protocol that permits a Client computer using a process to make a network (or local) request to use the processing resources of a Server. Thus, Client requests execution of instructions by another process located on a remote Server on the network.
- End Point Mapper listens on Port 135 for TCP/IP
- Client must RPC-bind to an interface before it can make procedure calls to Server
- RPC over SMB is also done

### **MICROSOFT NETWORK NAME RESOLUTION:**

#### **WINS & NT 4.0 Domains:**

Windows Internet Name Service & DNS uses TCP Port 134 for Name Resolution

#### **WINS/DNS for WIN2K Domain:**

WINS & DNS can still use TCP Port 134 for Name Resolution

SMB CIFS & NetBT can also travel directly over TCP/IP using TCP Port 445

#### **LDAP:**

- Clients submit LDAP requests to Servers—Servers can refer Client to other Servers for information
  - Celerra runs LDAP version 3.0 as an agent on DART to query & access Active Directory in Windows 2000 Domains
  - Windows 2000 supports LDAPv3 defined in RFC 2251 and LDAPv2 as defined in RFC 1777
  - Lightweight Directory Access Protocol allows clients to query, create, update, delete information stored in Active Directory.
  - LDAP Clients submit requests for operations based on entries for objects that are comprised of specific attributes. Entries have attributes consisting of a ‘type’ with one or more values [ASCII strings; URL’s; Public Keys]. Attribute values and distinguished names are defined by ISO 10646 character set.
  - LDAP protocol is carried over TCP or UDP
  - LDAP Versions 3 & 2 are used for Windows 2000
  - LDAP replaces ‘named pipes’ SMB communication with Domain Controllers
  - SASL (Simple Authentication and Security Layer) mechanisms can be used with LDAP to provide security services.
- Include DataMover in "Enterprise Admins" account  
Enable "Authenticated Users" to have read access to user account in all domains in forest that will be accessing datamover  
Distinguished Names; Universal Resource Locator name; Globally Unique Identifier

#### **LDAP over TCP:**

LDAP PDUs (Protocol Data Units) are mapped directly to TCP byte stream, usually listening on Port 389. AD also supports port 636 for LDAP Secure Sockets Layer communications. Global Catalog AD Servers listen on ports 3268 for LDAP & 3269 for LDAP SSL.

**LDAP Referral:** If a DC does not have the requested object during a query, will refer them to a different DC.

**RootDSE:** Consists of top portion of LDAP search tree containing logical namespace. Identifies domain, schema, and configuration directory partitions for a single DC and the forest root domain directory partition. Clients connect to the “rootDSE” when making LDAP inquiries. RootDSE publishes information about LDAP Servers, such as versions, SASL mechanisms, etc.

**Client→LDAP RootDSE→Server provides standard information in reply**

#### **How LDAP is Used During Windows 2000 Client Startup & Logon:**

Clients use LDAP locator process to locate DC to use for logon. LDAP is also used to find Group Policy Objects for computer or user. Also locates certificates if used.

#### **LDAP SERVICE OPERATIONS:**

BindRequest; BindResponse; UnbindRequest; SearchRequest; SearchResponse; ModifyRequest; ModifyResponse; AddRequest; AddResponse; DelRequest; DelResponse; ModifyRDNRequest; ModifyRDNResponse; CompareRequest; CompareResponse; AbandonRequest

→BindRequest is made by Client to Server

→BindResponse from Server

→LDAP Operation sent to Server

## **CELERRA CIFS SERVICES LIMITATIONS:**

- No physical limit to number of CIFS services that can be configured [just practical limit due to finite resources—old limit was 29]
  - Use multiple IP interface configurations for a single CIFS service for High Availability and Network Redundancy
  - Shares can be local to a specific Netbios name or Global across all netbios names for a specific Server
  - Each CIFS Service can participate in a separate NT Domain or Forest
  - Each CIFS Service can have its own DNS domain defined [up to (1) Primary & (2) Secondary DNS Servers]
- Note:** Use \$server\_dns command to add DNS Service & \$server\_cifs –add compname=, domain=, dns= to specify DNS domain  
--VLAN tagging can also be used to ensure security of packets when using CIFS Services in multiple NT Domain environments on the same DataMover

## **LOCALGROUPS SUPPORT:**

Celerra maintains a localgroups database similar to other Windows Servers

CIFS 4.0 Servers → The localgroups database can be managed by either User Manager or Computer Management Snap-in

CIFS 2000 Servers → Can use only Computer Management snap-in [also manages Shares & Event Viewer]

\\\172.19.3.22\compmgmt.msc

## **SPARSE FILE SUPPORT:**

Celerra frees unused space from 'sparse' files--these files allocate a logical size, but may only take up a small physical portion

## **NAMED STREAMS SUPPORT: aka, Alternate Data Streams (ADS), Multiple Data Streams, Named Data Streams, etc**

Every file uses a main, unnamed stream associated with it for 'data'. Alternate Data Streams are hidden files that are linked to the normal data stream file, and contains various information used by some applications and operating systems. For file systems that provide ADS capability, applications can write data at specific offsets within a stream as an alternate sequence of bytes, and then access the streams via their names. Thumbnail bitmaps and Summary Tab information on files with Windows NT/2000 are examples of files containing ADS. Name Streams (ADS) was introduced by Microsoft in the 1990's as a way to support Macintosh platforms as a file server. Macintosh files use a Data Fork and Resource Fork—the Resource Fork is used to store information about how the file was created, and by what application, allowing the correct application to subsequently reopen the file—whereas Microsoft uses file extensions to determine which application opens a file.

Celerra supports ADS with NAS 5.0 and higher. 'Streams' consist of data associated with a main file or directory. In FAT file systems, this data stream was called an "unnamed stream". NTFS introduced the concept of multiple data streams (MDS/ADS) within a file called 'named streams'. 'Named Stream' data is a hidden file—even Windows Explorer itself does not have the ability to see 'named streams'. To create an example of a 'named stream' within a file on an NTFS system, at the command prompt type the following:

**C:>echo hello > test:stream**

Opening "test" in notepad shows an empty file. Using **c:>more < test:stream** from prompt shows word "hello" inside the file.

### **Enabling Streams for NAS 5.0:**

**param cifs.multipleStreams=1**

Alternate Data Streams works on files with no spaces in name, but not with directories—param no longer found in NAS 5.5.

### **VERIFYING ADS NAMED STREAMS ON WINDOWS DATA:**

Download freeware utility called 'streams v1.53' from [www.sysinternals.com](http://www.sysinternals.com)

### **PROBLEMS WITH ADS:**

NAS 5.0.9.1 and 5.1.9.4 have problems when Clients try to create files with null strings, resulting in CIFS deadlock condition and loss of CIFS access to users. NAS 5.1.9.301 fixes problem in 5.1 family. A bug also exists in 5.1 in which NFS clients can corrupt a file system if they access files that contain ADStreams, causing panics. Issue first seen at Thales, 9660961, AR33615 and is fixed in 5.1.19.0.

### **Use the following param to turn off ADS and prevent FS corruption prior to 5.1.19.0:**

**param shadow stream=0**

### **SIDHistory:**

Migration capability from NT 4.0 Domains to WIN2K

## **WINDOWS PLATFORMS SUPPORTED BY CELERRA:**

NT 4.0 SP 4-6a → CIFS 2.1 and higher

Windows 2000 SP 1-3 → CIFS 2.2.35.4 and higher

Windows XP SP1 → CIFS 4.1.7 and higher

Windows 20003 Server (W2K3) → Not yet supported with any version NAS

**Note:** NAS 5.1M2 will have W2K3 Client Support; NAS 5.2 will include Domain Controller support

#### **Issues with 2003 Server:**

Default Domain Controller security policies require that clients use encryption and digital signatures via the secure channel socket between DC and Client, aka “SMB Signing”. DM does not yet support this mode and therefore cannot effectively “Join” the Domain, perform NETLOGON, or obtain SID [DART supports SMB Signing with NAS 5.2].

#### **Symptoms:**

Server Log →NETLOGON fails with “STATUS\_ACCESS\_DENIED” or “SMB: 3: SSXProcessReply: TreeXError=c0000022

Server\_Cifs output does not show SID

Users cannot map successfully to Celerra Shares

Network Traces show that Celerra cannot access the IPC\$ share on the 2003 DC's

## **BASIC STEPS IN CONFIGURING CELERRA AS A WINDOWS 2000 CIFS SERVER:**

- Step 1. Mount File System with relevant CIFS options [Locking Policies and Accesspolicies]
- Step 2. Start the Time Service on the DataMover [server\_date timesvc start ntp -i 0800:mn host] [Up to (4) Time Servers]
- Step 3. Configure DNS on the DataMover [server\_dns] **[Up to (3) DNS Servers can be specified per Data Mover]**
- Step 4. Enable I18N or ASCII Filtering [Note: Strongly recommended to enable I18N!]
- Step 5. Add CIFS Server to WIN2K Domain [server\_cifs -a compname=four0, domain=T2DOM2.LOCAL.COM, interface=ana1]
- Step 6. Join DataMover to Domain [server\_cifs -Join compname=four0, domain=T2DOM2.LOCAL.COM, admin=administrator]
- Step 7. Implement Usrmapper 3.0/NTMigrate/Manual Passwd/Group files
- Step 8. Start the CIFS Service [server\_setup]
- Step 9. Create CIFS Shares

**NTMigrate Note:** NTMigrate is supported for Windows 2000 Active Directory & produces same (3) output files that are then merged with a Unix Passwd/Group file or NIS database using "ntmiunix.pl" perlscript utility.

## **SETTING UP THE CIFS 2000 DM:**

### **I. Mount & Export File Systems:**

**Mount Options:** Mount Type; File Locking; Opportunistic Locks; File Change Notification; Access Checking Policies

**Mount Type:** rw [default] ro \$server\_mount server\_4 -o rw fs3 /mnt3

**File Locking:** nolock [default] wlock rwlock \$server\_mount server\_4 -o nolock fs3 /mnt3

**Opportunistic Locks:** nooplock [on by default] \$server\_mount server\_4 -o nooplock fs3 /mnt3

**File Change Notification:** on by default \$server\_mount server\_4 -o nonotify fs3 /mnt3

**Access Checking Policies:** NT UNIX SECURE NATIVE [default] \$server\_mount server\_4 -o accesspolicy=NT fs3 /mnt3

### **II. Configure NTP Time Service for CIFS Server:**

a.) **Start Time Service on DataMover:** Default clock delay is 60 minutes

\$server\_date server\_4 timesvc start ntp 172.24.80.11

\$server\_date server\_4 timesvc [Status of Time Service]

\$server\_date server\_4 timesvc stop | start | update ntp

\$server\_date server\_4 timesvc delete ntp

#### **UPDATING SERVER TIME MANUALLY:**

\$server\_date server\_4 0209041130 [Year, Mnth, Day, Time]

**Note:** In practical terms, best to use the NTP Service already running on one of the Windows 2000 Active Directory Servers!

**Important:** Ensure that Data Movers time is synchronized with Windows 2000. If not synchronized, CIFS Servers will not “Join” Active Directory properly or function properly after “Joining”—you would lose ability to “manage” the system from “Active Directory Users and Computers”. Authentication will not occur if the Time differential is greater > 5 minutes!!

### **III. Configure DNS Service for CIFS Server:** [Up to (3) DNS Servers can be listed]

**Note:** Windows 2000 and “Dynamic DNS” replaces “WINS”--when configured properly, this service will automatically register the CIFS Server in the WIN2K domain.

#### **NAS 4.0/4.1 REQUIRES REVERSE LOOKUP ZONE IN DNS:**

**CAUTION:** Make sure that the Windows 2000 Server’s DNS Service is configured with a “Reverse Lookup Zone”—the CIFS Server will not be able to register without this “zone” enabled. However, if Reverse Lookup Zone is used, it must match entries in Forward Lookup Zone or else the Join will fail—Celerra requires this because in a multi-protocol environment, this zone lookup is used to determine if we are communicating with the correct hosts. More specifically, we use an implementation of Kerberos used for mixed Unix & Windows environments—this MIT-based Kerberos obtains actual Host names by querying the Host’s IP Address in the Reverse Lookup Zone. In reality, if this Zone is not configured, the Kerberos Service may still obtain the correct name from the DNS query. NAS 4.0 & 4.1 may fail to obtain Kerberos ticket if Host names in Forward & Reverse Lookup Zones do not match. NAS 4.2.6.1 & higher no longer requires use of Reverse Lookup Zone for DM. See Microsoft Q174419 for info on setting up MS DNS.

Note: Reverse Lookup Zone is not enabled by default with Windows 2000 Server installations of DNS.

a.) **Setup the DNS Service on the DataMover:**

**\$server\_dns server\_4 -p tcp | udp t2dom2.local 192.10.3.2** [UDP is default protocol used]

Note: Recommendation is to use UDP protocol

**\$server\_dns server\_4 -o stop | start | flush**

**\$server\_dns server\_4 -delete t2dom3.local**

*Windows 2000 Domain:* t2dom2.local

*Default DNS Zone:* t2dom2.local

*Windows 2000 Active Directory Server FQDN:* win2ksrvdom2.t2dom2.local

b.) **Setting Up DNS on Windows 2000 Active Directory Server: REVERSE LOOKUP ZONE!**

1. Programs>Administrative Tools>DNS>View>Advanced>WIN2KSRVDOM2: Expand Folders

2. Reverse Lookup Zone: Rightclick and select "New Zone"

3. Welcome to the New Zone Wizard: Next

4. Zone Type: Select "Active Directory-integrated"

\*Active Directory-integrated

\*Standard Primary

\*Standard Secondary

5. \*Network ID: Enter: 192.10.3

6. You have successfully completed the New Zone wizard.

Name: 3.10.192.in-addr.arpa

Type: Active Directory Integrated Primary

Lookup Type: Reverse

FINISH [Shows up under "Reverse Lookup Zones" as "192.10.3.x Subnet"]

7. Highlight DNS Server & Stop/Restart DNS: WIN2KSRVDOM2>All Tasks: Stop | Start

8. FORWARD LOOKUP ZONES>t2dom2.local>Properties>General: Allow Dynamic Updates? Change to "Yes"

9. REVERSE LOOKUP ZONES>3.10.192.in-addr.arpa>Properties>General: Allow Dynamic Updates? Yes

**Example of Correct Entry for Datamover in Reverse Lookup Zone:**

Note: PTR Record for Datamover should look like the following entry:

DNS SERVER>Reverse Lookup Zones

100.168.192.in-addr.arpa 100 PTR dm2.dsunset.com

Rightclick entry for datamover>Properties>

Pointer (PTR)

Subnet:

100.168.192.in-addr.arpa

Host IP Number:

100

Host Name:

dm2.dsunset.com

**Important!** If AD Server is multi-homed, it most likely will have "A" Records for each IP Address in the Forward Lookup Zone. Make sure that it has only (1) PTR record in the Reverse Lookup Zone and reflects the IP Address of the Interface being used by the Datamover for Active Directory. When adding cifs server and 'joining' to domain, entries will show up immediately in DNS if successful. If a version other than Microsoft Windows 2000 DNS is being run on the network, ensure that it is Unix BIND 8.2.2 compliant to support Dynamic DNS. If Dynamic DNS is not being used, then add the CIFS Server manually to the DNS Servers by adding an "A" Record in the Forward Lookup Zone and a "PTR" Record in the Reverse Lookup Zone.

**Best Practice:** With NAS 4.2.6.1, no longer need the "PTR" record to function in AD

**IV. Enable I18N or SetUp the ASCII Filter Parameter:**

Note: If I18N is not used, then the following ascii parameter is defined in the DataMover "param" or Celerra Global "slot\_param" file

**Best Practice:** Engineering strongly recommends that I18N be configured by default on all new installations!

Why? Because all CIFS clients normally negotiate UNICODE on connection to File Servers. And, later 'conversions' tend to present a majority of the problems here.

a.) **Setting the ASCII Filter on a DataMover Basis:**

**#vi /nas/server/slot\_4/param**

**param shadow asciifilter=1**

b.) **Setting the ASCII Filter on the Whole Celerra:**

**#vi /nas/site/slot\_param**

**param shadow asciifilter=1**

Note: By default, ASCII filtering limits "Share" names to 12 characters

### c.) Enabling I18N on DataMovers:

1. **\$/nas/sbin/uc\_config -setup** [Normal→“Common Unicode translation subdirectory already exists.”]
2. **\$/nas/sbin/uc\_config -update** [Run this if the Customer has other Code Pages to load on the Servers]
3. **\$/nas/sbin/uc\_config -on -mover ALL | server\_2**

**Note:** After this command, Server\_2 will show "I18N Mode=UNICODE"

**Additional I18N Commands:** \$/nas/sbin/uc\_config -setup -i | -l

## V. Add CIFS Service Computer Name to Celerra Database:

- a.) **\$server\_cifs server\_4 -add compname=bubba, domain=t2dom2.local, interface=ana0**

**Note:** Adds the computer account to CIFS

Can only choose to hide the netbios name from Network Neighborhood [hidden=y]

Authentication=kerberos | all ["all" is default for both Kerberos & NTLM]

Netbios=legacy [use this option when you need to have a different netbios name than the FQDN]

#### Corresponding Server Log Entry:

2002-03-15 14:01:33: SMB: 4: CIFS Server TODD1[T2DOM2] created (0)

2002-03-15 14:01:33: SMB: 4: Full computer name todd1.t2dom2.local, Realm T2DOM2.LOCAL

2002-03-15 14:01:33: SMB: 4: Interface ana0 added into CIFS Server TODD1[T2DOM2]

2002-03-15 14:01:33: ADMIN: 4: Command succeeded: cifs add compname=TODD1 domain=T2DOM2.LOCAL interface=ana0

**Error Seen When DNS Not Setup Properly:** Note that Join command still works!

2002-03-14 14:41:25: SMB: 4: Dart account todd not registered with DNS.

2002-03-14 14:41:25: ADMIN: 4: Command succeeded: domjoin compname=todd domain=t2dom2.local admin=tmattha password=7?!?7 init

2002-03-14 14:41:25: LIB: 3: dns\_updateMessage: unable to get hostname of DNS server : Domain not exists

2002-03-14 14:41:25: LIB: 3: DNS update ERROR: Domain not exists while registering todd.t2dom2.local

## CHANGING DEFAULT COMMENT FIELD FOR COMPNANE:

Add Computer account to Create CIFS Service Using Following Syntax:

**\$ server\_cifs server\_6 -add compname=denver, domain=t2dom3.local, interface=ana1 -comment 'Look Out World'**

#### Server CIFS Output:

CIFS Server DENVER[T2DOM3]

Full computer name=denver.t2dom3.local realm=T2DOM3.LOCAL (domain not joined !)

Comment='Look Out World'

if=ana1 l=192.10.3.29 b=192.10.3.255 mac=0:0:d1:1f:61:2f

FQDN=denver.t2dom3.local (Updated to DNS)

## EXAMPLE OF USE OF OPTIONS FOR SERVER\_CIFS:

**\$server\_cifs server\_2 -add  
compname=cel2k, domain=ts2.emc.com, hidden=y, authentication=Kerberos, netbios=Celerra, interface=ace0, wins=192.168.3.2, dns=emc.com -comment 'EMC\_Celerra'**

**Note:** For no comment at all, leave a space between tic marks: -comment ‘ ’

## HOW TO REMOVE 'EMC-SNAS' FROM DRIVE MAPPINGS TO CELERRA ON XP CLIENTS:

**HKeyCurrentUser>Software>Microsoft>Windows>CurrentVersion>Explorer>ComputerDescriptions:**  
→edit and remove the 'EMC-SNAS' comment

## VI. Joining the CIFS Server to the Windows 2000 Domain:

- a.) **\$server\_cifs server\_4 -Join**

**compname=bubba, domain=t2dom2.local, admin=tmattha, DC=win2k.t2dom2.local**

**server\_4 : Enter Password:\*\*\*\***

**done [Use -U to unjoin a domain]**

**Note:** This command adds computer account to AD 'OU' & DNS Forward & Reverse Lookup Zones

#### JOIN PROCESS DEFINED:

1. **Join Command Issued→DART queries DNS Server for list of LDAP Services [DM→DNS→DM]**
2. **DART queries LDAP Server via UDP to obtain attributes of Domain Controllers [DM→LDAP→DM]**
3. **DART queries DNS for Kerberos Information [DM→DNS→DM]**
4. **DM obtains Session ticket with KDC [DM→KDC→DM]**
5. **DM uses LDAP via TCP to request add of Compname and Celerra 'OU' to AD [DM→LDAP→DM] & obtains SID**

#### Corresponding Server Log Entry:

2002-03-15 14:03:24: SMB: 4: Dart account todd1 not registered with DNS.

2002-03-15 14:03:24: ADMIN: 4: Command succeeded: domjoin compname=todd1 domain=t2dom2.local admin=tmattha password=7?!?7 init

2002-03-14 14:41:25: LIB: 3: dns\_updateMessage: unable to get hostname of DNS server : Domain not exists

2002-03-14 14:41:25: LIB: 3: DNS update ERROR: Domain not exists while registering todd.t2dom2.local

**SERVER LOG MESSAGES FOR UNSUCCESSFUL JOIN:**

2002-03-15 13:26:35: ADMIN: 3: Command failed: domjoin compname=todd domain=t2

dom2.local admin=tmatta password=7?!?7 unjoin

2002-03-15 13:49:01: KERNEL: 4: Time set

2002-03-15 13:52:30: SMB: 4: LGDB: Database started

2002-03-15 13:52:30: SMB: 4: ShareDB: DART database started

2002-03-15 13:52:30: ADMIN: 4: Command succeeded: cifs start

2002-03-15 13:52:44: KERBEROS: 3: krb5\_sendto\_kdc: unable to send message to any KDC in realm T2DOM2.LOCAL.

2002-03-15 13:52:44: KERBEROS: 4: krb5\_gss\_release\_cred: kg\_delete\_cred\_id failed: major GSS\_S\_CALL\_BAD\_STRUCTURE or GSS\_S\_NO\_CRED, minor G\_VALIDATE\_FAILED

2002-03-15 13:52:44: LDAP: 0: LdapGssAuthenticator::getCredentials: gss\_acquire\_cred\_ext failed: @ = 0x83d0f04, for principal user/todd@T2DOM2.LOCAL - GSS-API

major error: Miscellaneous failure

2002-03-15 13:52:44: LDAP: 0: LdapGssAuthenticator::getCredentials: gss\_acquire\_cred\_ext failed: @ = 0x83d0f04, for principal user/todd@T2DOM2.LOCAL - GSS-API

minor error: Preauthentication failed

2002-03-15 13:52:44: LDAP: 3: LdapBinder::bind: SASL\_PROTOCOL\_VIOLATION

2002-03-15 13:52:44: SMB: 4: Unable to connect to Active Directory server win2ksrvdom2.t2dom2.local, port 389

2002-03-15 13:52:44: SMB: 3: smb\_threadCtx::SSXAuth\_KERBEROS: Failed to obtainattributes for: user Administrator - realm T2DOM2.LOCAL - server todd@T2DOM2.LOCAL

**RESETTING COMPUTER PASSWORD ACCOUNT IN DOMAIN:**

-o reuse | resetserverpasswd [Used when migrating from Mixed to Native Mode using same computer account; Resets CIFS compname password and encryption keys]

**Example of Syntax:**

**\$server\_cifs server\_4 -Join compname=suncor, domain=network.lan, admin=migratory -o  
resetserverpasswd**

**Following Error Indicates a Password or Computer Account Desynchronization With Domain:**

2002-11-09 20:44:50: SMB: 3: SSXAK=c000006d origin=600 stat=d0000,-1765328246

**Comment:** Server does not register with DNS until CIFS started! If error messages persist, check DNS Forward & Reverse Lookup Zones and Allow Dynamic Updates setting!!

**PURPOSE OF ADDING CELERRA TO 'ENTERPRISE ADMINS' GROUP IN AD:**

By default, the Celerra CIFS 2000 Server is added to the "Enterprise Admins" group by default so that the datamover will have the capability to query any Domain Controller in the Forest so as to validate Users accessing Celerra shares. While not a requirement, this is our default behavior. Celerra no longer requires the "Enterprise Admins" group for NAS 4.x or 5.x

**Misc. Notes:**

--Join command creates a machine account in Active Directory & adds Server to an "Organizational Unit" called "ou=Computers,ou=EMC Celerra". Celerra becomes member of both "Domain Computers" and "Enterprise Admins" for T2DOM2.LOCAL/Users. Compname can be up to (63) UTF8 characters in length and represents the FQDN. 'Admin=' account is an Administrator with rights in the Domain Forests to add a member to 'Enterprise Admins' Group.

--Join command also immediately creates entries in DNS Forward and Reverse Lookup Zones

**USING JOIN COMMAND TO JOIN SPECIFIC "OU" CONTAINERS IN ACTIVE DIRECTORY:**

**\$server\_cifs server\_2 -Join  
compname=dm4, domain=network.lan, admin=migratory, ou="ou=energy:ou=suncor"**

**Note:** Last "ou" added by this command is actually the first "ou" that is created [i.e., OU=Suncor, followed by OU=Energy]

**\$server\_cifs server\_2 -Join  
compname=dm4, domain=network.lan, admin=migratory, ou="ou=energy,ou=suncor"**

**Note :** Alternative syntax for OU's

**VII. Setup UID/GID Mapping Service:**

**VIII. Start the CIFS Service:**

a.) #server setup server 4 -P cifs -o start

b.) Access the Netbios Name and Check for SID Generation in "server\_cifs" Output.

**SERVER LOG ENTRIES WHEN STARTING CIFS FOR FIRST TIME:**

2002-03-14 15:11:55: SMB: 3: LGDB: New Database created

2002-03-14 15:11:55: SMB: 4: LGDB: Database started

2002-03-14 15:11:55: SMB: 4: Registry initialization OK

2002-03-14 15:11:55: CFS: 3: Init new log file

2002-03-14 15:11:56: CFS: 3: attempt to set a zero-length filename

2002-03-14 15:11:56: SMB: 4: ShareDB: DART database started

2002-03-14 15:11:56: ADMIN: 4: Command succeeded: cifs start

2002-03-14 15:11:57: LIB: 3: dns\_updateMessage: unable to get hostname of DNS server : Domain not exists

**Example CIFS Output:**

DOMAIN T2DOM2  
DC=WIN2KSRVDOM2(192.10.3.2) ref=1 time=0 ms  
CIFS Server (Default) TODD[T2DOM2]  
Full computer name=todd.t2dom2.local realm=T2DOM2.LOCAL  
Comment='EMC-SNAS:T4.0.12.2'  
if=ana0 l=192.10.3.26 b=192.10.3.255 mac=0:0:d1:1f:83:1e  
FQDN=todd.t2dom2.local (Update of "A" record failed during update: Domain not exists)

**Note:** Above example reflects incorrect DNS configuration! Troubleshoot separately.

**Example of CIFS Output After Rebooting DataMover:**

CIFS Server (Default) TODD[T2DOM2]  
Full computer name=todd.t2dom2.local realm=T2DOM2.LOCAL  
Comment='EMC-SNAS:T4.0.12.2'  
if=ana0 l=192.10.3.26 b=192.10.3.255 mac=0:0:d1:1f:83:1e  
FQDN=todd.t2dom2.local (Retrying DNS update...)

**AFTER CHANGING FORWARD & REVERSE LOOKUP ZONES PROPERTIES TO:**

Allow Dynamic Updates: Changed from "Allow Dynamic Updates 'Secure Updates Only' to 'Yes'  
Stopped & restarted DNS, then CIFS and following DNS Update made to DNS & CIFS OUTPUT.

\$ server\_cifs server\_5

32 Cifs threads started  
Security mode = NT  
Max protocol = NT1  
I18N mode = ASCII  
Home Directory Shares DISABLED  
Usermapper[0] = [172.24.80.11] last access 0  
Enabled interfaces: (All interfaces are enabled)

DOMAIN T2DOM2  
SID=S-1-5-15-323e04be-36d67c9a-32eac016-ffffffffff  
DC=WIN2KSRVDOM2(192.10.3.2) ref=1 time=0 ms  
CIFS Server (Default) TODD[T2DOM2]  
Full computer name=todd.t2dom2.local realm=T2DOM2.LOCAL  
Comment='EMC-SNAS:T4.0.12.2'  
if=ana0 l=192.10.3.26 b=192.10.3.255 mac=0:0:d1:1f:83:1e  
FQDN=todd.t2dom2.local (Updated to DNS)

**Note:** If successful, DNS entries show up immediately in DNS REVERSE & FORWARD ZONES  
And, can now "manage" the account "todd1" in Active Directory

**IX. Create CIFS Shares:**

a.) **From CommandLine: Control Station**

\$server\_export server\_4 -P cifs -n share /mnt3

b.) **From WIN2K: "Computer Management" Interface**

Programs>Administrative Tools>Active Directory Users and Computers>T2DOM2.LOCAL>EMC Celerra>Computers  
>todd1: Rightclick netbios name and select "manage">Computer Management (TODD1.T2DOM2.LOCAL)>System Tools>  
**Shared Folders>Shares:** Rightclick & select "New">File Share>Folder to Share>C\$>mnt4 [Name Share & assign perms]

C\$ C:\  
CHECK\$ C:\  
cmdshare C:\todd  
IPC\$ \etc

**Shared Folders>Sessions**

administrator 192.10.3.2  
win2ksrvdom2\$ 192.10.3.2

**Shared Folders>Open Files**

\etc\PIPE\srvsvc administrator

**X. UNJOINING/DELETING CFS 2000 SERVER FROM WINDOWS 2000 DOMAIN:**

**Step 1. Unjoin the Server from the Domain:**

\$server\_cifs server\_4 -Unjoin compname=win2k, domain=t2dom2.local, admin=tmatta

Enter Password: xxxxxxxx

Removes from Active Directory and DNS

## **Step 2. Deleting the CIFS Service from the Data Mover:**

**\$server\_cifs server\_4 -delete compname=win2k [Deletes CIFS Service configuration]**

**Note:** Must have DNS service running on DataMover to do this

## **NOTE ABOUT "REALM" IN WINDOWS 2000:**

The word "realm" refers to the Kerberos Security space of Windows 2000 active directory piece. Therefore, any errors referencing "realm" would most likely be related to a problem with Kerberos security

## **JOIN PROCESS FOR WIN2K:**

1. FQDN request and response
2. Bind & query LDAP server for DC's
3. Kerberos session between DM & KDC
4. LDAP add request/response

**Note:** When the Netbios name is different than the prefix name for the FQDN, then the "netbios" parameter needs to be specified using the "server\_cifs -add" command

## **MOVING COMPNAME FROM ONE “OU” TO ANOTHER IN ACTIVE DIRECTORY:**

**NAS 4.2.9.0:** Moving CIFS “compnames” from default location, “EMC Celerra : Computers” to a different Organizational Unit [OU] Process of moving a compname using the Active Directory Users and Computers is very straightforward.

1. As Administrator, go to Start>Programs>Administrative Tools>Active Directory Users and Computers
2. Expand the “EMC Celerra>Computers” OU and highlight the Celerra compname to be “moved”.
3. Rightclick the compname and select “move” and select the appropriate OU container to move the computer account to.
4. Verify data and administrative access to the compname

## **CREATING CIFS “SHARE” FROM WINDOWS 2000 SERVER:**

Step 1. Computer Management>Action>Connect to another computer>Look in Entire Directory

Step 2. Highlight Celerra Machine Account: “win2k (user/win2k@t2dom2.local) T2DOM2.LOCAL/EMC Celerra/Computers”

Step 3. Expand ‘System Tools’>Shared Folders>Shares>rightclick and select “New”>File Share>Folder to Share>C\$>mnt4  
[C:\mnt4 = path for cifs file system]

- Step 4. Name the New Share and Assign Network Access Permissions
- \* All Users have Full Control [Default Value assigned]
  - Administrators have Full Control; Others Read Only
  - Administrators have Full Control; Others No Access
  - Customize Share and Folder Permissions
- “Finish”

## **HIDING COMPNAME FROM NETWORK NEIGHBORHOOD:**

Use hidden=y option with server\_cifs command

## **WINDOWS 9x LIMITATION ON RETRIEVING NETWORK SHARES INFO: Q202892**

By default, Windows 9x Clients are limited in the amount of info that can be retrieved, to 65,536 bytes. If the NT Server or Datamover delivers more than that, then not all shares will be displayed by Client. This is a LAN Manager 2.1 limitation.

### **Condition Represented by Following Irritating Server Log Message:**

"Max buf Size of Client too small"

## **WIN2K RESTRICTED LOGON ACCESS PROBLEM:**

NAS Code prior to 4.2--Client uses port 445 to connect to Celerra and we used IP Address of Client in place of machine name. Therefore, if access is restricted based on machine name of client, then connection will be refused by Server.

## **USING NFS TO COPY FILES TO A CIFS FILE SYSTEM:**

If Customer conducts a data migration from another Unix system to a CIFS filesystem, using strictly NFS CPIO or RSYNC operations, the “Ownership” of the directories copied will probably be the Unix Root User, indicated as “UNIX Uid=0x0” when examining folder permissions after the migration. Similarly, the default NT Permissions should be for “EVERYONE Full Control” Properties>Advanced>Owner **UNIX UID=0x0”**

If required, the Administrator can “Take Ownership” of files and re-permission directories as needed.

## **COMMON ERRORS WHEN CIFS SERVER ACCOUNT NOT ‘JOINED’ WITH WIN2K DOMAIN:**

Indicators: Missing DNS entry for Server/Server not ‘manageable’ from Active Directory Users and Computers Menu

**I. WINDOWS 2000 SERVER POPUP ERROR:**

Manage Computer>todd1>Shared Folders>Shares

"MMC The System encountered the following error while reading the list of shares: Error 5: access is denied"

**II. SERVER LOG ENTRIES:**

2002-03-22 13:06:02: SMB: 4: >DC=WIN2KSRVDOM2(192.10.3.2) R=2 T=0 ms S=0,1/-1

2002-03-22 13:06:02: KERBEROS: 3: krb5\_sendto\_kdc: unable to send message to any KDC in realm T2DOM2.LOCAL.

2002-03-22 13:06:02: KERBEROS: 4: krb5\_gss\_release\_cred: kg\_delete\_cred\_id failed: major GSS\_S\_CALL\_BAD\_STRUCTURE or GSS\_S\_NO\_CRED, minor G\_VALIDATE\_FAILED

2002-03-22 13:06:02: SMB: 3: KC\_GetServerCreds:gss\_acquire\_cred\_ext failed; maj Status=0xd0000, min=-1765328347

2002-03-22 13:06:02: SMB: 3: KC\_ComputeBlob Status=70000,0

2002-03-22 13:06:02: SMB: 3: DC\_GetBlob T2DOM2.LOCAL\todd1 WIN2KSRVDOM2\HOST=70000 -2045022973

**III. ACCESS FROM NT CLIENT [CORRESPONDING SERVER LOG ERROR]: [Line logged each time access to CIFS Server attempted]**

SMB:3:SSXAK=c000006d origin=600 stat=d0000, -1765328347

**IV. EXAMPLE OF SERVER "ALPHADM2" WITH BROKEN SECURITY TRUST:**

"Network path not found" from clients:

2002-03-28 16:43:31: SMB: 4: >DC=NTSRVDOM1(192.10.2.2) R=2 T=0 ms S=0,1/-1

2002-03-28 16:43:31: SMB: 3: NetLogon::buildSecureChannel=7 E=0xc000018b

2002-03-28 16:43:31: SMB: 3: NetLogon::buildSecureChannel=7 E=0xc000018b

2002-03-28 16:43:31: SMB: 3: NLogon\_SecureChannel not OK=7

2002-03-28 16:43:31: SMB: 3: No more valid DC (currDC=6ecbc04 6ecbc04/6ecbc04)

2002-03-28 16:43:31: SMB: 4: DC=\*SMBSERVER(192.10.2.2) R=2 S=15,c000018b/-1

2002-03-28 16:43:31: SMB: 3: No nextDC available, keep DC=6ecbc04

**Resolution:** Stopped CIFS, then UsrMapper, deleted Netbios Name & Readded back to Domain for Success

**RESETTING MACHINE ACCOUNT FROM ACTIVE DIRECTORY GUI:**

**Note:** There are situations whereby the "trust" account for the CIFS Server may need to be reset with the Windows Domain From Windows 2000 Active Directory Users and Computers Menu:

Manage Computer>rhombus: Rightclick, select "Reset Account"

**WINDOWS 2000 POPUP MESSAGE:**

"Active Directory Are you sure you want to reset this computer account? Yes  
Account rhombus was successfully reset. OK"

**LOSS OF CIFS ACCESS WHEN CELERRA TIME SERVICE BECOMES UNSYNCHRONIZED**

**WITH KERBEROS KDC SERVER:**

**Users have lost access to a CIFS Share and the Server Log is logging the following error(s):**

2002-04-08 11:01:32: SMB: 3: SSXAK=c000006d origin=600 stat=d0000,-1765328351

2002-04-08 11:01:32: SMB: 3: SSXAK=c000006d origin=600 stat=d0000,-1765328351

Clock skew too great

**Users See Following Errors from Windows Clients:**

a.) My Network Places>Computer>CIFS2000>"Enter Network Password: Incorrect Password or unknown username for: \CIFS2000"

b.) The Following Error is encountered when using the "Active Directory Users and Computers" interface to "manage" the CIFS2000 Netbios name: "Computer Management" "Computer \CIFS2000 cannot be managed. The network path was not found."

c.) Or, when trying to "manage" the CIFS2000 Server from within the "Active Directory Users and Computers" interface, the following errors are encountered: "Error 5: Access is denied"

**Note:** You may see several different messages, all resulting in an "Access Denied" message. Also, you may see certain elements of the CIFS2000 Netbios name circled with a red X around the subject, such as in "Local Users and Groups"; all these are indicators that the CIFS Server has lost the ability to login to the Windows 2000 Domain using Kerberos Security.

**Cause:**

By default, the Kerberos Key Distribution Center [KDC] Time Clock requires that Client Computer Systems Clocks on a Windows 2000 Active Directory Domain be synchronized to within a 5 minute variance. When Client Systems become unsynchronized, they can no longer successfully establish secure network communications using Kerberos Security and are effectively 'denied access' to services on the network until the Time Synchronization problem is corrected.

**Resolution:**

Resynchronize the DataMover's Time Service with that of the Customer's Network Time Service:

\$server\_date server\_5 timesvc stop | start

**TROUBLESHOOTING WIN2K ISSUES:**

1. Verify NTP Service and Time Synchronization
2. Encryption issues related to RC4 v. DES [see Q248808]
3. DNS Zones properly created [FWD & REVERSE] and non-authenticated updates enabled!

## **IMPROPER DNS SETUP=WIN2K CLIENT ERRORS:**

**"Reverse & Forward Lookup Zones" "Allow Dynamic Updates?": "Only Secure Updates"**

**Comment:** With "Only Secure Updates" set, you can add the CIFS Server and "Join" the Windows 2000 Domain, Create a "Share" from the commandline that is network accessible, Obtain a SID, but the appropriate DNS Entries will not be made and you will not be able to "Manage" the cifs server from "Active Directory Users & Computers"--if you try, you obtain a client error:

- a.) No entries in DNS Forward or Reverse Zones for CIFS server
- b.) Active Directory Users and Computers>Manage Computer>todd>Shared Folders>Shares:

## **WINDOWS 2000 POPUP ERROR:**

Microsoft Management Console

"The System encountered the following error while reading the list of shares:  
Error 5: access is denied"

Computer Management

"Computer <\\todd1.t2dom2.local> cannot be managed. The Network path was not found. Choose 'Connect to another Computer' from the Action menu to manage a different computer."

"WIN32: The RPC Server is unavailable"

## **SERVER LOG DNS ERRORS:**

1. CIFS Stopped or 2. REVERSE & FORWARD LOOKUP Zones not set properly; Allow Dynamic Updates? Yes!!!

2002-03-14 15:38:28: LIB: 3: DNS error: unable to get SOA to update todd.t2dom2.local entry

2002-03-14 15:38:28: LIB: 3: DNS update ERROR: Primary name server not found while registering todd.t2dom2.local

2002-03-14 15:41:16: LIB: 3: dns\_updateMessage: unable to get hostname of DNS server : Domain not exists

2002-03-14 15:41:16: LIB: 3: DNS update ERROR: Domain not exists while registering todd.t2dom2.local

2002-03-14 15:43:04: LIB: 3: DNS ERROR: no response from name server 192.10.3.2 after 2 seconds

## **WIN2K DNS PROPERTIES:**

DNS>WIN2KSRVDOM2>Forward Lookup Zones><t2dom2.local>><todd1>>Doubleclick netbios name for "Properties"

### **Host(A) Tab:**

Parent Domain: t2dom2.local

Host: todd1

IP Address: 192.10.3.26

Update associated pointer (PTR) record

Delete this record when it becomes stale

Record time stamp: 3/14/2002 5:00:00 PM

Time to Live (TTL): 0:0:20:0

### **Security Tab:**

Administrators (T2DOM2\Administrators)

DnsAdmins (T2DOM2\DsAdmins)

Domain Admins (T2DOM2\Domain Admins)

Enterprise Admins (T2DOM2\Enterprise Admins)

ENTERPRISE DOMAIN CONTROLLERS

Everyone

Pre-Windows 2000 Compatible Access

SYSTEM

Permissions: Full Control Read Write

Allow inheritable permissions from parent to propagate to this object

## **SYMPTOMS SEEN WHEN DNS NOT CORRECT:** 'Allow Dynamic Updates?' is set to "Only Secure Updates"

1. Cifs Service can be successfully added using \$server\_cifs server\_4 -add compname=todd1, domain=t2dom2.local, interface=ana0
2. FQDN is added to Active Directory: \$server\_cifs server\_4 -Join compname=todd1, domain=t2dom2.local, admin=tmatta
3. Name "todd1" is not added to DNS Forward or Reverse Zones; Server Logs give errors; Server\_Cifs output also indicates failure
4. Name "todd1" is added to Active Directory, but cannot be 'managed'; GUI Errors; Server\_Log errors, etc
5. Active Directory Users and Computers>todd1>Properties can be viewed but Shares cannot be created [Access Denied]
6. Shares can be created from CommandLine and then successfully mapped from Windows Clients!
7. SID is also obtained

### **Windows Client Errors:**

"Computer <\\todd1.t2dom2.local> Cannot be managed. The Network Path was not found. Choose 'Connect to another Computer' from the Action menu..."

## **WINDOWS 2000 DOMAINS:**

Each Domain contains its own "Active Directory". Multiple Domains comprise a "forest". The Global Catalog Service is the common denominator. Likewise, each Domain contains its own "DNS Zone" and trusts are automatic within the forest. Otherwise, create trusts between "forests" by configuring each Domain as a Secondary DNS Server for the other [New Zone Type>Standard Secondary]. Note, must have a working Global Catalog Service [first DC in the forest] for authentication/logons to work.

**SETTING WINDOWS 2000 USER RIGHTS/POLICIES:**

Programs>Administrative Tools>Domain Security Policy>Local Policies>User Rights Assignments | Security Options  
 >Domain Controller Security Policy>Local Policies>User Rights Assignments | Security Options  
 >Local Security Policy>Local Security Settings>Security Settings>Local Policies>User Rights Assignment>Security Options

**WINDOWS 2000 DNS INTERFACE:****Programs>Administrative Tools>DNS>WIN2KSRVDOM2>Properties**

| <u>Interfaces</u>                                                                                                                                                                                                                                                                                                                                                                                                                                      | <u>Forwarders</u>                                                                                                     | <u>Advanced</u>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | <u>Root Hints</u>                |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------|
| Select IP Addresses that will serve DNS requests.                                                                                                                                                                                                                                                                                                                                                                                                      | <input type="checkbox"/> Enable forwarders<br>Forwarders help Resolve any DNS queries not answered by this server.    | <b>Server Version Number:</b><br>5.049664<br><b>Server Options:</b><br><input type="checkbox"/> Disable recursion<br><input checked="" type="checkbox"/> BIND secondaries<br><input type="checkbox"/> Fail on load if bad zone data<br><input checked="" type="checkbox"/> Enable round robin<br><input checked="" type="checkbox"/> Enable netmask ordering<br><input type="checkbox"/> Secure cache against pollution<br><b>Name Checking:</b> Multibyte (UTF8)<br><b>Load Zone data on Startup:</b> From Active Directory and registry | m.root-servers.net. 202.12.27.33 |
| <b>Listen On:</b><br>*All IP Addresses<br><input type="checkbox"/> Only the following IP Addresses:<br>192.10.3.2                                                                                                                                                                                                                                                                                                                                      |                                                                                                                       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |                                  |
| <b>Logging</b>                                                                                                                                                                                                                                                                                                                                                                                                                                         | <b>Security</b>                                                                                                       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |                                  |
| <b>Debug Logging Options:</b> Administrators<br><input type="checkbox"/> Query<br><input type="checkbox"/> Notify<br><input type="checkbox"/> Update<br><input type="checkbox"/> Questions<br><input type="checkbox"/> Answers<br><input type="checkbox"/> Send<br><input type="checkbox"/> Receive<br><input type="checkbox"/> UDP<br><input type="checkbox"/> TCP<br><input type="checkbox"/> Full packets<br><input type="checkbox"/> Write through | Authenticated Users<br>Dns Admins<br>Domain Admins<br>Enterprise Admins<br>Pre_Windows 2000<br>SYSTEM<br>Advanced Tab |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |                                  |
|                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                                                                                                                       | <b>Monitoring</b><br>To verify the configuration of the servers, you can perform manual or automatic testing.<br><b>Select a test type:</b><br><input type="checkbox"/> A simple query against this DNS Server<br><input type="checkbox"/> A recursive query to other DNS Servers<br>Test Now                                                                                                                                                                                                                                             |                                  |
|                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                                                                                                                       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |                                  |
| <b>Log file location on Server:</b> %System Root%system32\dns\dns.log                                                                                                                                                                                                                                                                                                                                                                                  |                                                                                                                       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |                                  |

**Programs>Administrative Tools>DNS>Forward Lookup Zones>T2DOM2.LOCAL>Properties**

| <u>General</u>                      | <u>Start of Authority (SOA)</u>              | <u>Name Servers</u> | <u>WINS</u> | <u>Zone Transfers</u>                         | <u>Security</u>     |
|-------------------------------------|----------------------------------------------|---------------------|-------------|-----------------------------------------------|---------------------|
| Status: Running                     | Serial Number:                               | win2ksrvdom2.       | n/a         | <input type="checkbox"/> Allow Zone transfers | Administrators      |
| Type: Active Directory-Integrated   | 76                                           | t2dom2.local        |             |                                               | Authenticated Users |
| Data is stored in Active Directory. | Primary Server:<br>win2ksrvdom2.t2dom2.local | 192.10.3.2          |             | *Only to the following servers                | Dns Admins          |
| Allow Dynamic Updates?              | Responsible Person:<br>C.                    |                     |             | 192.10.3.2                                    | Domain Admins       |
| No   Yes   Only Secure Updates      |                                              |                     |             |                                               | Enterprise Admins   |
|                                     |                                              |                     |             |                                               | ENTERPRISE DOMAIN   |
|                                     |                                              |                     |             |                                               | Everyone            |
|                                     |                                              |                     |             |                                               | Pre-Windows 2000    |
|                                     |                                              |                     |             |                                               | SYSTEM              |

**Programs>Administrative Tools>DNS>Reverse Lookup Zones>3.10.192.in-addr.arpa>Properties**

| <u>General</u>                    | <u>Start of Authority (SOA)</u>              | <u>Name Servers</u> | <u>WINS-R</u> | <u>Zone Transfers</u>                         | <u>Security</u> |
|-----------------------------------|----------------------------------------------|---------------------|---------------|-----------------------------------------------|-----------------|
| Status: Running                   | Serial Number:                               | win2ksrvdom2.       | n/a           | <input type="checkbox"/> Allow Zone transfers | same as above   |
| Type: Active Directory-Integrated | 2                                            | t2dom2.local        |               |                                               |                 |
| Allow Dynamic Updates?            | Primary Server:<br>Win2ksrvdom2.t2dom2.local | 192.10.3.2          |               |                                               |                 |
| No   Yes   Only Secure Updates    | Responsible Person:<br>admin.t2dom2.local    |                     |               |                                               |                 |

**CIFS OUTPUT WITHOUT DNS CONFIGURED FOR DYNAMIC UPDATES:**

\$ server\_cifs server\_5  
CIFS Server (Default) CIFS1[T2DOM2]

Full computer name=todd.t2dom2.local realm=T2DOM2.LOCAL

Comment=EMC-SNAS:T4.0.12.2'

if=ana0 l=192.10.3.26 b=192.10.3.255 mac=0:0:d1:1f:83:1e

**FQDN=cifs1.t2dom2.local (Update of "A" record failed during update: Domain notexists)**

## **CIFS OUTPUT AFTER ALLOWING DYNAMIC UPDATES IN DNS FWD & REVERSE ZONES:**

Changed from "Allow Dynamic Updates 'Secure Updates Only' to 'Yes'

DOMAIN T2DOM2

SID=S-1-5-15-323e04be-36d67c9a-32eac016-ffffffff

DC=WIN2KSRVDOM2(192.10.3.2) ref=1 time=0 ms

CIFS Server (Default) CIFS1[T2DOM2]

Full computer name=todd.t2dom2.local realm=T2DOM2.LOCAL

Comment=EMC-SNAS:T4.0.12.2'

if=ana0 l=192.10.3.26 b=192.10.3.255 mac=0:0:d1:1f:83:1e

**FQDN=cifs1.t2dom2.local (Updated to DNS)**

## **DFS NAMESPACE (Replaces Distributed File System name) SUPPORT:**

--DFS's are Shared folders that can map to more than one root target, or Servers, which in turn can map to Links underneath the root

--Each Namespace is seen as a Shared folder, with subfolders underneath

--Celerra supports DFS as “leaf volume” only—DFS Server must be configured to see Celerra as a leaf volume [Native Windows supports ‘root volumes’]

--Celerra does not support File Replication service

--Since NAS 5.4, we only support Standalone DFS Root, which means there is only one root target

--Windows 2003 R2 changes from DFS to DFS Namespace, replaced DFS File System snap-in with DFS MMC Snap-in, and introduced DFS Replication. Windows 2008 added domain namespaces and other improvements.

**Note:** Celerra does not currently support DFS Replication, which replaced the File Replication Service (including NAS 6.0)

## **ACCESSING WIN2K OR NT SERVER FROM UNIX:**

Setup an "rshd" service on the Windows machine to enable a Unix Host to connect & use "rsh" and "rcp"

## **ENABLING RSH ON DATA MOVER:** /usr/bin

**# .server\_config server\_2 -v "rshd"**

1090516820: ADMIN: 4: Command succeeded: rshd

**Note:** To stop the RSH daemon on the DM, must reboot! No other way to turn off from CLI.

#rsh server\_x “pdc dump” or #rsh server\_x pdc dump

#rsh server\_x ls

## **ENABLE/DISABLE CIFS NETWORK INTERFACE:**

**\$server\_cifs server\_2 -Enable ana0 |Disable ana0**

**Note :** Use to associate a specific interface with the default CIFS configuration

## **CFS 4.0 32-BIT GID SUPPORT:**

**NFS support introduced with 2.2.35.4 +[new min NAS 2.2.46.x]**

Celerra supports 32-bit GID values (2 Billion--old values were 16-bit or 64k), but the actual value is still capped at a total of 64K GIDs per FS [0 – 2147483647]

16-Bit FileSystem Values range from 0 – 65,534

32-Bit FileSystem Values range from 0 – 2,147,483,647, but are still limited to a total of 65,534 GIDs per filesystem

**Note:** GID values are stored in the ushort of the inode, which is limited to 64k unique values, but are mapped differently for 32-bit so that the values can range from 0 – 2 billion (Using a 31-bit mapping scheme)

**Note:** Solaris, HP-UX, Netapps, AIX all support 32Bit GIDs or more

--Implement only during filesystem creation & not during mounting

--/etc/gid\_map file is created On-Disk at time of 32-bit file system creation [Superblock misc flag value FS\_GID\_32 0x20] to support translation between 32-bit values to 16-bit on-disk values

--Initial support only pertains to NFS file systems

--Linux 6.2 Control Station only supports 64k GID's [therefore you will not be able to see the 32-bit GID values]

--Linux 7.2 Control Station kernel now supports 2 Billion UID/GIDs

--Disable 'Quotas' prior to setting up 32-bit GID filesystem

--If NIS is being used, the GID can be resolved to a Group name

**Note:** A DataMover can mount and handle a 32-bit file system even if the 32-bit feature has not been enabled

## **NAS 5.3:**

32-bit GID support on by default

#nas\_fs -G [used to verify 32-bit fs]

**\$ nas\_fs -G fs06**

FS GID = 32BIT

### **32-BIT GID MAP:**

The gid\_map file is a way of mapping a 32 bit gid to a 16 bit gid and is unique for every file system. GIDs are only allocated sequentially on-disk, meaning that freed gid's have to be cleaned up from the database. The server\_mount option –o gid32Rebuild fs1 /fs1 frees up unused on-disk GIDs and puts a sentinel free value in unused slots in gid\_map file.

**# server\_mount server\_2 -o gid32Rebuild fs05 /fs\_repS**

**Note:** Each file system and rootfs that are 32bit will have the following gid\_map file. The gid\_map file is not visible in Versions 5.2 or 5.3, but is visible in 5.4:

```
-r--r--r-- 1 root bin 262144 Nov 9 16:45 gid_map
```

### **RECOVERING GID MAP FILE IF DAMAGED BY FSCK OR OTHERWISE:**

1. Mount the affected file system with the gid32Ignore option:

**\$ server\_mount server\_2 -o gid32Ignore home /home**

2. Verify the file system is mounted with the option:

```
$ server_mount server_2 | egrep home
```

```
home on /home uxf5,perm,rw,gid32Ignore
```

3. Check to see if the gid\_map file was moved to lost+found directory of the 'home' file system. The size of the gid\_map file is always 262144 bytes and the inode number of the file should be a low order number. If the gid\_map is found in the lost+found directory, or the gid\_map file needs to be obtained from a different file system\*, first copy the gid\_map to a temporary directory on the Control Station (e.g., /tmp), then copy the file back to the appropriate file system's hidden /etc directory using the server\_file command, as demonstrated in Step 4. The reason for this two-step process is that the /etc directory is hidden and only accessible via the server\_file command:

**Note:** The gid\_map is not universal and each file system will be unique—however, there are rare occasions where Eng. may propose restoring a gid\_map from one file system to another.

4. [root@laip1 lost+found]**# cp -ip gid\_map /tmp**

5. Copy the file to /etc (not /.etc) of the affected file system:

```
$ cd /tmp
```

**\$ server\_file server\_2 -put gid\_map /home\\Vetc/gid\_map**

```
server_2 : done
```

6. Umount the file system and remount without gid32Ignore option.

```
$ server_umount server_2 -p /home
```

```
server_2 : done
```

```
$ server_mount server_2 home /home
```

```
server_2 : done
```

7. Verify access and re-export for NFS and CIFS if required.

### **Example of what typical listing of the gid\_map file should be:**

```
$ ls -ail
```

```
7 -r--r--r-- 1 root bin 262144 Nov 9 16:45 gid_map
```

### **VERIFYING WHETHER FILE SYSTEM IS 16 or 32-BIT SUPPORTED?**

**\$server\_config server\_5 -v "file query /metro"**

```
1058875245: CFS: 4: fsRawSize = 4418808
```

```
1058875245: CFS: 4: inodes = 535294
```

```
1058875245: CFS: 4: kbytes = 4350792
```

```
1058875245: CFS: 4: used = 248
```

```
1058875245: CFS: 4: avail = 4350544
```

```
1058875245: CFS: 4: refCount = 13
```

**1058875245: CFS: 4: gid = 16**

```
1058875245: ADMIN: 4: Command succeeded: file query /metro
```

**Note:** Superblock contains a misc flag value of FS\_GID\_32 0x20 when 32-bit support is enabled

**\$ nas\_fs -G fs06**

FS GID = 32BIT

### **CIFS SUPPORT FOR 32-BIT GIDs:**

--Introduced with NAS 4.1

--Usrmapper Version 3 also supports

**Note:** NAS 5.2.9.6 still defaults to 16-bit file system support

## **32-BIT GID BACKUP SUPPORT:**

**NDMP:** Veritas & Legato NDMP supports 31-bit GID's for Backups [emctar; ustar; cpio; bcpio; sv4cpio; sv4crc; tar does not] --Server\_archive emctar also supports 31bit UID/GID Backups --Other Backup Formats: ustar=21bit; tar=18bit; cpio=15bit; bcpio=16bit; sv4cpio=21bit; sv4crc=21bit

## **MODIFY PARAM FILE TO ENABLE 32-BIT VALUES:**

```
param ufs gid32=0 [default value=16-bit GID's]
param ufs gid32=1 [Sets 32-bit values]
$ .server_config server_2 -v "param ufs"
ufs.gid32      0x00e0f734  0x00000000  0x00000000
```

## **CREATING A FILESYSTEM TO SUPPORT 32-BIT GID's:**

1. #vi /nas/server/slot\_13/param  
**param ufs gid32=1**
2. Reboot Server\_13
3. Create 32Bit FileSystem Using Server\_13:  
**#nas\_volume -n metav\_d91\_32bit -c d91,d92**  
**#nas\_fs -n 32bitfs01 -c metav\_d91\_32bit -o mover=server\_13**
4. Create mountpoint & export for NFS:  
**#server\_mountpoint server\_13 -c /32bitfs01**  
**#server\_mount server\_13 32bitfs01 /32bitfs01**  
**#server\_export server\_13 -o anon=0 /32bitfs01**

**Note:** Turn Off the param after the filesystem is created! [gid32=0]

5. Verify that FileSystem was created as 32-bit by checking Server Log:

### **SERVER LOG ENTRY:** UFS: 4: Creating GID MAPS

**Note:** NAS Version 5.1 and higher requires use of “mover=server\_x” in order to create 32-bit file system support—previous versions would default to Server\_2, but this is no longer the default behavior.

```
$server_config server_5 -v "file query /metro"
```

## **GID's GREATER THAN 65,535 ARE WRITTEN TO ACL's ON DISK WITH INCORRECT GID VALUES—HOW TO MIGRATE FROM 16-BIT TO 32-BIT FILE SYSTEMS:**

### **EXAMPLE OF FILE SYSTEM ISSUE WHEN USING UID OR GID VALUES > THAN 64k:**

If GID value of 75000 is used, unless 32-bit support is enabled for all file systems, the value will be truncated and written to disk as “9464”. 75,001 will be written as “9465”, etc. The meaning of this should be clear—actual GID values will be written to ACL records with the “new” truncated value, not the value that Usermapper has assigned. If a file system has encountered this problem, the following is one method for correcting:

1. Set following parameter on Data Mover that will be used to create 32-bit File System support:

```
/nas/server/slot_4/param
```

```
param ufs gid32=1
```

2. Reboot Server\_4 and create 32-bit File System:

```
#nas_fs -n newfs32 -c mtvfs32 -o mover=server_4
```

**Note:** Must specify the Server to be used for File System creation!

3. Migrate data from 16-bit file system to new 32-bit file system:

```
$server_archive server_4 -J -rw /oldfs /newfs32
```

**Note:** Server\_archive can only be used to migrate data located on the same Server. The above syntax preserves CIFS attributes. Sticky bit values and Quotas are also preserved. Data migration rate using server\_archive is only 16-18GB per hour.

4. Verify ACL's on 16-bit file system with known GID problem

```
# .server_config server_4 -v "acl if=fsn1 dump=\fs00620\0\GRPF-KPC-4-15-03"
```

Default values:

Interface 'elv0' Addr='0.0.0.0'

```
===== UNIX =====
```

USER 4196 GROUP **45007** mode=rwxr-xr-x

```
===== NT =====
```

Owner=USER 4196 UNIX UID=0x1064 ".:S-1-5-12-1-1064

Group=GROUP **110543** UNIX GID=0x1acf ".:S-1-5-12-2-1acf

**Note:** The NT Value for the Group in this example is 110,543—this is the value assigned by Usermapper service. Problem is that this value translates/truncates to “45007” when written to UNIX record on file system that has only 16-bit support enabled.

5. After a successful data migration from 16-bit to 32-bit file system, update the “on-disk” ACL records for all Users and Groups by

running Setacl--this will update the “UNIX value from “45007” to “110,543” in the above example:

**#server\_config server\_4 "cifs update /newfs32/testfolder force setacl if=ace0"**

6. Run ACL Dump again to validate the setacl repairs

## **RECOVERING 32-BIT GID MAP FILE IF FSCK REMOVES TO LOST+FOUND:**

**Note:** On rare occasions the gid\_map of a file system can become corrupted and require recovery—use following procedure

1. Remount file system with gid32Ignore option

**\$ server\_mount server\_2 -o gid32Ignore fs1 /fs1**

2. Locate gid\_map file from lost+found (if FSCK moved) or use same file from another 32-bit file system—gid\_map file size is always 262144 bytes

3. Copy gid\_map to /tmp directory

4. Copy gid\_map to /etc directory of the appropriate file system (NOT .etc)

**#server\_file server\_2 -put /tmp/gid\_map /fs1\etc\gid\_map**

5. Unmount file system and remount with normal mount options

## **NFS Protocol:**

Client-Server model where Server exports local directories for remote NFS Clients. NFS is a Session-layer protocol [RFC 1057], but with RPC used with NFS running over UDP or TCP. Data values are encoded using XDR. A Remote Procedure Call passes data to a function and awaits a reply [or value] to be returned. RPC can also conduct Host and UID authentication, and supports DES encryption. NFS is stateless in the sense that a server does not keep track of the dialogues it conducts with remote clients. NFS uses a number of Procedures, called synchronous, as the ‘server’ blocks additional requests until a valid reply is sent. Asynchronous requests generally lead to higher performance.

## **COMPONENTS USED WITHIN NFS:**

→RPC protocol RFC1057

→XDR protocol RFC1014

→NSM Network Status Manager for statd, crash & recovery, client reboot notification to Server, lock re-claim

→NLM Network Lock Manager for lockd, advisory locks only [NLMv1, 3 for NFSv2, NLMv4 for NFSv3]

→Port Mapper

→Mountd

→Xid, xid cache

## **NFS RETRANSMISSIONS:**

Duplicate XIDs are seen by the RPC layer and indicate that UDP packets have been lost and requires retransmission.

## **NFS NETWORK LAYERS:**

NFS v2 or v3 uses RPC Client-Server calls over TCP/UDP and IP. Portmapper service [RPCBind] registers applications. RPC Calls are assigned unique XID numbers.

## **NFS OPERATIONS:**

**NULL:** NULL does not perform real work, but is used to allow Server response testing and timing

**GETATTR:** Get File Attributes; retrieves attributes for file object by file handle [fhandle] from Server in response to a LOOKUP, CREATE, MKDIR, SYMLINK, MKNOD, READDIRPLUS

**SETATTR:** Used to change attributes of file system or directory object on Server; new attributes specified by sattr3 structure.

**LOOKUP:** Searches directory for specific name & returns file handle for file system object

**ACCESS:** Check Access Permission; Determines Client access rights for a File Object [ACCESS3\_READ--read from file or directory; ACCESS3\_LOOKUP--lookup name in directory; ACCESS3 MODIFY--Rewrite file date or modify directory; ACCESS3\_EXTEND--Write new data or add directory entries; ACCESS3\_DELETE--Delete directory entry; ACCESS3\_EXECUTE--Execute a file]; Client encodes permissions in bit mask, Server checks permissions & returns "NFS3\_OK" with enclosed bitmask permissions if o.k.

**READLINK:** Reads data from symbolic link as ASCII string.

**READ:** Reads data from file in 8-Kbyte blocks

**WRITE:** Writes data to file/or cache in NFS v3 in 8-Kbyte blocks

**CREATE:** Creates regular file

**MKDIR:** Create new subdirectory

**SYMLINK:** Creates symbolic link

|                     |                                                                                                                    |
|---------------------|--------------------------------------------------------------------------------------------------------------------|
| <u>MKNOD:</u>       | Create new special file for block, device, named pipes, or socket                                                  |
| <u>REMOVE:</u>      | Removes entry in directory                                                                                         |
| <u>RMDIR:</u>       | Removes a subdirectory                                                                                             |
| <u>RENAME:</u>      | Rename a file or directory                                                                                         |
| <u>LINK:</u>        | Creates hard link from file to link name in directory                                                              |
| <u>REaddir:</u>     | Retrieves variable number of entries in sequence from directory & returns name & file identifier (reads directory) |
| <u>REaddirplus:</u> | Extended read from directory for variable number of entries; enhanced REaddir (reads directory)                    |

FSSTAT: Retrieves dynamic/volatile file information

FSINFO: Retrieves non-dynamic or non-volatile information

PATHCONF: Retrieves POSIX information, pathconf information for a file or directory

COMMIT: Commit cached data to stable storage using WRITE procedure

**NFS runs over IP using UDP for NFS v2 and both UDP [port 2049] and TCP for NFS v3.**

Note: Must run NFS v3 if you are using CIFS! [Uses RPC over TCP]

CIFS runs only over TCP/IP [No UDP support here!]

DataMover can support both NFS & CIFS simultaneously

XDR—External Data Representation: A Presentation Layer protocol acting as RPC translator between disparate Hosts and is equivalent to the CIFS Unicode protocol.

ACCESS3\_READ: Read data from file or directory

ACCESS3\_LOOKUP: Look up name in directory

ACCESS3 MODIFY: Rewrite existing file data or directory entries

ACCESS3 EXTEND: Write new data or add directory entries

ACCESS3\_DELETE: Delete an existing directory entry

ACCESS3\_EXECUTE: Execute a file

## **UNNAMED PIPES:**

Way for processes to share data, usually in the context of the login Shell in use.

# ls | wc -l

Note: Example of unnamed pipe where the std-out of ls process is sent to std-in of wc process

## **NFS CALLS:**

NFS Packets that include RPC & NFS Headers, along with the IP fragment of the datagram. Biode's are Client NFS processes that communicate with Server nfsd processes (aka RPCBIND)—typical Unix systems run 8-16 biode processes. Threads spawn processes that conduct work. Modern NFS requests are typically multiplexed over a single TCP connection—note, if too many NFS requests are channeled over a single TCP stream, could create a performance issue—this is called a lack of I/O Concurrency. In otherwords, may need to implement multiple streams to improve performance.

## **NFS CLIENT-SERVER MODEL:**

Client always issues the Commands or Calls to the Server (NFS Commands (C))

Server always sends NFS Replies or Responses (R)

Clients always “write” to the Server or “read” from the Server

## **NFS PROCESSES:**

--“Biodes” are process threads that run on NFS Clients that send NFS Read or Write requests to the Server. Multiple Biodes allow for concurrent I/O requests. Each biod uses a different originating port number, but destination port number is always 2049. Normally 4-8 biodes running on a Client.

--Asynchronous Requests are when Client biod processes send multiple NFS requests prior to receiving reply from previous requests.

--Concurrent Requests are when multiple biod requests occur simultaneously.

--“Nfsds” are process threads running on the Server that handle NFS Client “biod” requests.

--Safe Asynchronous Writes occur when using NFS V3—Server replies with “Reply OK” to Client before writing to disk [Uses new COMMIT procedures].

--NFS Streaming is only supported when using TCP and Record Marking Standard markers for NFS. When NFS Streaming is used, multiple requests can be sent in a single packet, as well as multiple replies from the Server.

--NFS does not support the use of ‘nested mountpoints’ by design so that clients cannot see the same file across different mountpoints

## **ASYNCHRONOUS vs. SYNCHRONOUS NFS CALLS:**

Async requests, especially writes, are faster than synchronous because the client is not waiting for replies for every call, but making multiple write calls, which is more efficient and leads to greater throughput.

## **NFS MOUNTS:**

Mount is used to get the file handle of the directory being mounted. The “timeo” mount option specifies the NFS retransmission timer, defined in tenths of a second—often set as 11/10ths of a second for connectionless transports and 600/10ths for a connection-oriented transport. The “retrans” option sets the number of NFS retransmissions—default is 5 [when using TCP, this option does not normally have any effect]. The “rsize” and “wsize” options are used to set maximum data transfers for Read and Write requests, but Server negotiates with client during FSINFO call—maximum Celerra “wsize” is 64k. Mount protocol always uses UDP.

#### **Soft and Hard Mounts:**

These mount options determine how Clients react whenever the Server is overloaded or crashes. By default, all NFS filesystems are mounted Hard, meaning that Client RPC calls that timeout will retry indefinitely. The effect of this is to make the Server appear as a “local disk” to the Client. A side effect of Hard Mounting is that processes can “hang” or block other requests while waiting for an NFS RPC call to complete. Soft mounts allow NFS operations to fail and crashed Servers will return system calls with error messages. Do not use Soft Mounts whenever writing data or running executables from the Server [i.e., wouldn’t use soft mounts if filesystems are mounted RW]

**Hard Mount Error:** NFS server jassoc not responding, still trying

**Soft Mount Error:** NFS read for server jassoc: error 5 (RPC: Timed out)

### **NFS PERFORMANCE ISSUES:**

#### **TYPICAL NFS THROUHPUT ON 507 DATA MOVERS USING GbE:**

Single Client Reads: 15 MBps

Single Client Writes: 20 MBps

Multiple Clients Writing to DM: 35MBps

#### **TYPICAL NFS THROUHPUT ON 510 DATA MOVERS USING GbE:**

Single Client Writes: 35-40 MBps

Multiple Client Writes: 60 MBps

**Comment:** Data Mover DART is tuned to handle multiple Client requests more efficiently than a Single Client. So, best practice performance testing on DMs should be done using multiple Clients, as is more realistic of a network environment.

### **TESTING NFS WRITES TO NS600 FROM NFS RED HAT 8.0 CLIENTS:**

```
10.32.3.10:/fs1 /fs1 nfs      noauto,rw,hard,intr,proto=udp,rsize=32768,wsize=32768 0 0  [Using UDP mount on NS600]
10.32.3.10:/fs1 /fs1 nfs      noauto,rw,hard,intr,proto=tcp,rsize=32768,wsize=32768 0 0  [Using TCP mount on NS600]
dd if=/dev/zero of=/fs1/test/terminal1-1.dat bs=32k count=32000      [Crontab jobs from clients]
dd if=/dev/zero of=/fs1/test/terminal1-2.dat bs=32k count=32000
```

### **HOW TO CHANGE NUMBER OF NFS DAEMONS ON EACH DATAMOVER:**

**Background:** Sun Servers run 1024 nfsd daemons by default. Celerra DataMovers run 96 by default. The number of daemons can be changed but as a caution, this should be done carefully and in small increments to assess system impact

Step 1. Go to /nas/server/slot\_x/netd and vi

Step 2. **nfs start openfiles=15360 nfsd=96** [change nfsd=96 to nfsd=124, etc...]

### **CELLERRA NFS DAEMONS :**

Nfsd—manages the transfer of traffic across the network

Biod—manages the disk writes to FileSystems at the Block level

Lockd—processes file lock requests for NFS clients

Statd—maintains lock status and crash recovery [statd & NLM work to allow clients to reclaim previously held locks]

Mountd—manages mount characteristics of NFS filesystems

Portmapper [port 111]—Client requests Port for a transaction; Server provides port for communications; Occurs for every transaction

**Note:** Portmap provides a mapping service to network ports for services such as mountd, nfsd, etc.

### **NFS CLIENT DAEMONS:**

LOCKD & STATD daemons keep track of locks on NFS files

Client maintains list of FileSystems to mount on bootup, in the /etc/fstab file, via the /etc/rc.boot script

Alternatively, can manually mount the Fstab FileSystems: \$/usr/etc/mount -a -t nfs

### **NFS SERVER DAEMONS:**

RPC.MOUNTD daemon answers requests from Clients for NFS files and maintains /etc/rmtab file of FileSystems currently mounted NFSD daemon also processes Client requests

LOCKD & STATD daemons keep track of locks on NFS files

Exports FileSystems in the /etc/exports file, via the /etc/rc.local script [#exportfs]

Showmount -e -d shows exported filesystems [Solaris]

## **UNIX-BASED EXPORTS FOR NFS:**

# showmount -e [Shows how a server is exporting its file systems]  
# exportfs # exportfs -a

## **NFS VERSIONS:**

### **NFSv2 or NFSv3:**

Both have UDP/IP or TCP/IP support and can support 2000 Connections by default.

Both use Mount protocol, Status Monitor, and Network Lock Manager via RPC calls.

NFSv2 supports maximum data transfer size of 8k; supports UDP & TCP, but mostly uses UDP, & max file size of 2GB [NFSv2 uses 32-bit file sizes/offsets]

NFSv3 max data transfer size with UDP is 64k, supporting filesizes >than 2GB; NFS over TCP supports larger transfer sizes [Data transfers negotiated with FSINFO]. With NFSv3, safe asynchronous writes are supported using COMMIT. Clients have to wait for the Server to reply using NFSv2 before issuing more commands, with NFSv3, Client knows data is written to disk when “Reply OK” received to a COMMIT request. NFSv3 also returns file attributes with each call, reducing need for explicit “GETATTR” calls that NFSv2 requires [NFSv3 supports 64-bit filesizes/offsets]

### **NFSv4 Protocol:**

Through the use of “Open/Close Operations”, this version now makes NFS a “stateful protocol”, and implements File Locking & Delegation. NFSv4 introduces OPEN operation for file lookup, creation, and share reservation to become “stateful”.

Other Improvements: Enhanced access & performance across Internetworks; Strong security; Cross-platform interop.

NFS is still considered a “system” that is “Operating System” independent, using RPC procedures. NFS4 is now a single protocol with a well-defined port that uses RPC in “operations” v. “procedure calls”.

Mandatory Locking Introduced: Concept of “stateid” to block I/O operations by applications that record a lock on a file [similar to the way Windows locks files]

### **CELLERRA NFSv4:** (see Configuring NFS on EMC Celerra 300-004-152)

→ NFSv4 Clients can manage file system objects by ACLs when Celerra file systems use accesspolicy of MIXED or MIXED\_COMPAT

→ Celerra requires use of Unicode with UTF8 encoding

→ NFSv4 requires the use of UTF-8 encoded Usernames and Groupnames, as well as an NFSv4 domain

**Note:** File & Directory structure still maintains the UID/GID

→ Built on ONC RPC—RFC1831, uses XDR—RFC1832, TCP-Only

→ Supports CIFS, NFSv2-NFSv4, & FTP

→ NFSv4 RFC 3530, uses port 2049, connection-oriented only, i.e., TCP [Does not use UDP as NFSv2 or NFSv3 can]

→ NFSv4 uses 32-bit NFS clients (param nfsv4 32bitClient)

→ Use of Read/Write Client Delegation for caching data, metadata, & locking to improve network performance, is enabled by default on a per file system basis

**Note:** Celerra delegation levels are None (No file delegation granted); Read (Only Read delegation); and Read/Write (Read/Write delegation granted)

→ Support for Share reservations

→ Enhanced security as NFSv4 supports KerberosV5 user authentication

→ Offers better WAN performance than previous NFS protocols

→ Mount, NLM, Status, ACL protocols are removed

→ MKDIR & RMDIR replaced by CREATE & REMOVE

→ File locking now built into the NFS protocol, hence no more NLM

→ File System object names, owners, owner groups, & metadata are UTF-8 encoded

→ Mandatory attributes will be supp\_attr, type, lease\_time, others

→ Access Control Lists will be compatible with CIFS ACLs, requires MIXED mode for Client, requires synchronization between mode-bits and ACLs (Owner, Owner\_group, Everyone)

→ RPCSEC\_GSS layer built for AUTH\_SYS or KERBEROS v5 security via GSS-API

→ Access to file systems will be through a pseudo-root “/” construct instead of using mount protocol

→ NFSv4 adds mandatory locks, byte ranges and share reservations

→ NFSv4 Open/Close operations are Stateful, ensures atomicity of Share reservations, Exclusive creates, delegate authority to Clients

→ NFSv4 is stateful, required for proper locking and recovery from failures, but doesn't require constant connection—CIFS requires continuous TCP connection to maintain state, NFSv4 does not, will recover

→ Both mode bits for Unix and ACLs will be allowed for NFSv4

→ Set param nfsv4 domain=xxx & /nas/server/slot\_x/netd→hivers=4 and reboot Data Mover to enable NFSv4

→ Export a file system using “nfsv4only” option

- Configure name service (iPlanet, NIS, etc.) and configure NSSwitch
- Use Unicode and MIXED accesspolicy file systems in order to fully support NFSv4 ACLs
- Use NIS if using both CIFS & NFS to access file systems, NFSv4 uses user & group names for mappings, and has to then translate to UID/GIDs [using mapid daemon]

### **NFS VERSION 4 RESTRICTIONS/LIMITATIONS:**

- NFSv4 supported only for Server-side, not Client-side, hence CDMS migration from NFSv4 to Celerra is not supported, nor Celerra FileMover to NFSv4 Secondary storage
- SecMap caching cannot be used with this feature
- HighRoad MPFS is not supported
- NMFS file systems are not supported
- Only pseudo-roots supported at this time, not pseudo file systems
- Access Control will be through Windows SIDs
- Configuration support not provided by Celerra Manager
- Celerra does not yet support Volatile File Handles
- Blocking locks not supported
- No public key authentication yet [SPKM-3, Lipkey]

### **NFSv4 Clients:**

Solaris 10+ patches, AIX 5.3 ML-2 + patches, Linux, Hummingbird, HPUX

### **NFSV4 PARAMS & SETTINGS:**

**param nfsv4 domain=nfsv4.net** (use this param to set the NFSv4 domain)

/nas/server/slot\_2/netd file

**nfs start openfiles=240000 nfsd=256 hivers=4**

**Note:** Must specify the hivers=4 to set NFSv4—NFSv3 will remain the default

param dns updatePTRrecord →use when Windows is Kerberos environment

param cifs acl.sortAces →used to reorder ACEs per Windows Explorer model

\$server\_nfs server\_2 ALL | -v4 | -service -start | -stop | -client -list | -info | -release | -stats

**# server\_nfs server\_2 -v4**

server\_2 :

NFSv4 is not enabled

Restart system and use nfs option hivers=4

**\$ server\_nfs server\_2 -v4 -client -info index=0xef400000**

**\$ server\_nfs server\_2 -v4 -client -list**

**\$ server\_nfs server\_2 -secnfs -user -list** (Secure NFS stats)

**\$ server\_nfs server\_2 -v4 -secnfs -user -info handle=32**

**\$ server\_nfs server\_2 -v4 -service -start | -stop** (Cleans up NFS v4 state and recalls delegations)

**\$ server\_param server\_2 -facility nfsv4 -info domain -v**

**\$ server\_param server\_2 -facility nfsv4 -info 32bitClient -v** (return attribute values in 32-bit only)

**\$ server\_param server\_2 -facility nfsv4 -info leaseDuration -v** (duration server maintains client state)

**\$ server\_param server\_2 -facility nfsv4 -info recallTimeout -v** (timeout for recall delegations)

**\$ server\_param server\_2 -facility nfsv4 -info delegLeaseDuration -v**

**\$ server\_param server\_2 -facility dns -info updatePTRrecord -v** (use dynamic updates for PTR—is not by default)

**\$ server\_param server\_2 -facility cifs -info acl.sortAces -v** (sort ACEs in the ACL in the order Explorer expects)

**# server\_param server\_3 -facility nfsv4 -list**

| param_name    | facility | default | current | configured |
|---------------|----------|---------|---------|------------|
| recallTimeout | nfsv4    | 10      | 10      |            |
| vnodePercent  | nfsv4    | 80      | 80      |            |
| leaseDuration | nfsv4    | 40      | 40      |            |
| domain        | nfsv4    |         |         |            |
| 32bitClient   | nfsv4    | 1       | 1       |            |

### **Disabling Delegations:**

**\$ server\_mount -o nfsv4delegation=NONE**

**\$ server\_export server\_2 -P nfs -option nfsv4only /mnt** (restrict access from NFSv2 or 3 clients to NFSv4 only)

### **Problems:**

NFS v4 clients do not have a standard way of setting and interpreting ACLs

### **FILE SYSTEM DEFINITIONS:**

**File System Identifier [fsid]:** 128-bit unique per-server identifier of a single file name space.

**Regular File:** A simple byte stream.

**Filehandle:** Identifies a unique file on a per-server basis on a fs. NFSv2 & v3 had persistent filehandles. NFSv4 uses volatile filehandles to handle client caching operations.

## **IMPLEMENTING NFSv4 ON CELERRA:**

1. Enable a Name Mapping mechanism on the Celerra, using either local passwd/group files, NIS, or LDAP

**Note:** NFSv4 uses names in ACLs, not UID/GIDs

2. Edit /nas/server/slot\_x/netd and change “nfs start” line to read:

**nfs start openfiles=240000 nfsd=256 hivers=4**

3. Reboot Data Mover

4. Set the NFSv4 domain on the Data Mover:

**\$ server\_param server\_x -f nfsv4 -m domain -v europe.local**

5. Set the file system accesspolicy to mixed to support ACLs:

**\$ server\_mount server\_x -o accesspolicy=MIXED fs1 /fs1**

## **COMPARISON OF NFS VERSIONS:**

### **NFS Version 2   NFS Version 3   NFS Version 4**

|      |      |      |
|------|------|------|
| NULL | NULL | NULL |
|------|------|------|

#### **Compound Operations:**

COMPOUND/NVERIFY/VERIFY

#### **Open/Close Operations:**

OPEN/OPENATTR/OPEN\_CONFIRM/OPEN/DOWNGRADE/CLOSE

#### **Delegation Operations:**

DELEGPURGE/DELEGRETURN/SETCLIENTID/SETCLIENTID\_CONFIRM

#### **Client Callback Procedures for Delegation:**

CB\_NULL/CB\_COMPOUND/CB\_GETATTR/CB\_RECALL

#### **Locking Operations:**

LOCK/LOCKT/LOCKU/RENEW

#### **Filehandle Operations:**

PUTPUBFH/PUTROOTFH/GETFH/RESTOREFH/SAVEFH

#### **Security Operations:**

|        |                |
|--------|----------------|
| ACCESS | ACCESS/SECINFO |
|--------|----------------|

#### **Traditional File Operations:**

|          |             |                        |
|----------|-------------|------------------------|
| LOOKUP   | LOOKUP      | LOOKUP/LOOKUPP         |
| GETATTR  | GETATTR     | GETATTR                |
| SETATTR  | SETATTR     |                        |
| LINK     | LINK        | LINK                   |
| REaddir  | REaddir     |                        |
|          | REaddirplus |                        |
| Readlink | Readlink    | Readlink               |
| Create   | Create      | Create                 |
| Mkdir    | Mkdir/Mknod |                        |
| Remove   | Remove      | Remove                 |
| Rmdir    | Rmdir       |                        |
| Rename   | Rename      | Rename                 |
| Symlink  | Symlink     |                        |
| Read     | Read        | Read                   |
| Write    | Write       | Write                  |
|          | Commit      | Commit                 |
| Statfs   |             | FSSTAT/FSINFO/PATHCONF |

## **SERVER NFSSTAT:** NFS & RPC stats for NFSv2 & v3; Server cache statistics

**\$server\_nfsstat server\_4 [lists all statistics]**

-z {zeroes stats} -r [zeroes out RPC stats] -n [zeroes NFS stats] -s -c {server client}

**\$ server\_nfsstat server\_2 -s**

server\_2 :

Read/write size and alignment distribution (v3):

| size   | read | crossed | write  | crossed |
|--------|------|---------|--------|---------|
| 1 - 1  | 0    | 0       | 183    | 0       |
| 2 - 3  | 3    | 0       | 400586 | 0       |
| 4 - 7  | 39   | 0       | 12430  | 0       |
| 8 - 15 | 102  | 0       | 18140  | 0       |

|               |         |       |          |       |
|---------------|---------|-------|----------|-------|
| 16 - 31       | 100     | 0     | 4281     | 0     |
| 32 - 63       | 191     | 0     | 12464    | 0     |
| 64 - 127      | 330     | 0     | 94552    | 1125  |
| 128 - 255     | 538926  | 9992  | 751543   | 5791  |
| 256 - 511     | 490406  | 88200 | 10726464 | 6775  |
| 512 - 1023    | 3666    | 0     | 2078183  | 0     |
| 1024 - 2047   | 271057  | 15049 | 541734   | 9618  |
| 2048 - 4095   | 181237  | 0     | 1110809  | 0     |
| 4096 - 8191   | 5210239 | 0     | 1587799  | 32689 |
| 8192 - 16383  | 1552627 | 21820 | 407415   | 32985 |
| 16384 - 32767 | 79206   | 620   | 60851    | 5     |
| 32768 - 65535 | 3024210 | 981   | 3607389  | 1     |

**Clientside:** Large difference between “timeout” and “badxid” values may indicate client, server, network problem  
Use nfsstat –c from different clients to compare results

## **TROUBLESHOOTING NFS PROBLEMS:**

### **NFS & Client Mounting Process:**

1. Client initiates request to Server by contacting “portmap” daemon to discover service & port to communicate
2. Server returns list of services and ports
3. Client makes request to “mountd” daemon on Server to export a file system or directory
4. Server reads /etc/vfstab to verify availability and check permissions to directory and then returns a File Handle pointer to the file system or directory to the Client’s NFS kernel

### **NORMAL SEQUENCE OF EVENTS FOR MOUNTING FS NFS:**

1. PORTMAPPER/GETPORT/MOUNT [Mount executable tries to find mount Server]
2. MOUNT/NONE [Mount executable pings mount Server to verify that mount service is available]
3. MOUNT/MNT [Mount executable finds filehandle for the share]
4. PORTMAPPER/GETPORT/NFS [Mount executable finds NFS service]
5. NFS/NONE [Mount executable pings NFS server to verify service; after mounting, Mount executable notifies Client OS]
6. NFS/FSINFO [Client OS sends commands to Server querying for FileSystem info]

### **NFS CLIENT FILE ATTRIBUTE & READ BUFFER CACHING:**

Client caching is used in NFS for performance reasons, preventing some RPC calls to the Server. Operations such as ‘getattr’ and ‘ls’ collect metadata information and reuse it without always having to go back to the Server. File attributes, such as mod time & file size remain valid when read by a Client for 3 seconds, and if the attributes remain static, for 60 seconds before cache is flushed. Directory attributes typically remain valid for 30 seconds, up to 60 seconds before cache is flushed. Cache is also flushed whenever writes occur. Attribute changes created by write operations on the clientside are typically sent back to the server in a single setattr. Certain client operations, such as a chmod, are never cached on the client.

Cache consistency is used to ensure that the cached copy of file data is consistent with what is on the Server. The file’s modification time is used as a check to make sure that cached data is newer than the mod time—if mod time is newer, data is flushed.

Async threads typically perform read-ahead and write-behind operations for NFS clients. Data is read and uses cache consistency to keep track of changes, while writes are collected into buffers until a full page is collected or the buffer is filled, then written back to disk. With flush-on-close, NFS buffers are written to the Server. NFSv3 clients using write stable flag will force Server to write to disk using commit operation.

NFS file locking should be used in situations where overlapping Reads & Writes occur on the same file. File locking disables caching, meaning that writes are written back to the Server instead of local client cache.

#### **Attribute Caching can be disabled using following from Clientside:**

**# mount -o noac nfs\_server:/files/home /mnt**

**# mount -o actimeo=1 nfs\_server:/files/home /mnt** (use this to change attribute cache timeout on files & directories)

**Note:** Disabling attribute caching does not disable Read caching for NFS aysnc threads or VM system for consistency caching, but does require new getattr RPC call to Server for each check.

### **NFS CLOSE-to-OPEN CACHE CONSISTENCY:**

NFS Client applications opens files using GETATTR or ACCESS and write back any pending changes to Server when closing file, as well as provides opportunity for Server to report any write errors that may occur. This is done to provide Close-to-Open cache consistency. Weak Cache Consistency (WCC) is an NFSv3 feature that allows clients to check for attribute changes to a file before and after operations are made, though this method is not 100% reliable and can lead to stale data, though Linux no longer uses WCC. Linux uses NOAC mount option to achieve cache coherency by disallowing client file attribute caching, forcing the Client to check with the Server before changes are made.

## **NFS SERVER CACHING:**

- Inode cache, containing file attributes (speeds up NFS get and set operations)
- Directory Name Lookup Cache (DNLC)—caches directory entries to reduce pathname resolution—works at VFS layer
- Buffer cache used for data read from files—for NFSv2, writes are not cached, for NFSv3 writes are not cached unless the NFS stable flag is set to off

## **CELERRA NFS LOCKS:**

Celerra notifies Clients of locks that they previously had, after failover or failback. This is normal NFS Server behavior. Clients receive a 45sec grace period to reclaim locks. Implications here are that failover + grace period = about 75sec. Some applications may not respond to this length of outage. The 45sec “grace period” is not yet configurable.

## **UNCACHED WRITE MECHANISM:**

**\$ server\_mount server\_2 –option rw,uncached fs1 /fs1**

**Note:** Use this mount option to enhance performance for database applications with many connections to a large file for NFS.

**NFSSTAT:** Generally used for troubleshooting Networking/Performance Issues

**Badxid ~timeout:** Usually an indicator of a slow Server

**Badxid ~0:** Dropped packets between NFS Client & Server in Call or Reply

**Nfsstat -m:** Useful to determine NFS mount options

**Server Statistics:** **\$nfsstat -s** [displays server stats, # NFS RPC Calls Received and Rejected {badcalls}]

Calls=total RPC calls from clients

Badcalls=number of calls rejected by RPC layer [Might indicate overloaded network]

Nullrecv=number times RPC call not available

Badlen=number RPC packets truncated or damaged

Dupreqs=number duplicate RPC calls

Xdrccall=number RPC calls whose XDR header could not be decoded (External Data Representation)

writes>than 5% may indicate a problem

readlink>than 10% may indicate excessive symbolic links problem

getattr>than 40% indicates need to increase DNLC [Directory Name Lookup Cache--Speeds up name to inode mapping lookups]

\$vmstat -s for cache hit rate] and Inode Cache

**Client Statistics:** **\$nfsstat -c**

Calls=total calls made to NFS

Badcalls=number of calls rejected by RPC layer

Retrans=number calls retransmitted due to timeouts [>than 5% of total calls means requests aren't reaching server]

Badxid=number of times Server's reply was taking too long; Server Not Responding or other NFS Program Errors [If this value is approx. 0 may mean network is dropping requests [reduce rsize and wsize]; If value is large, then Server is taking too long

Timeouts=number times call timed-out while waiting for Server reply

Null>0 means Automounter retrying to mount too frequently

**NFS Statistics for Mounted FileSystems: \$nfsstat -m [Client mount status information]**

**\$nfsstat -m** [Useful for determining Server name and address, current NFS Read/Write sizes, retransmission count, hard/soft mount, and other mount status information of Client]

“NFS Server Not Responding”—Increase “timeo” parameter in /etc/vfstab by doubling the value to start with

Lookups>80; Reads>150; or Writes>250—Requests are taking too long to service [slow network or slow server]

**Note on NFSSTATs:** High Client “retrans” and “Timeouts” may indicate networking issues

## **USING NFSSTAT ON SOLARIS/LINUX CLIENTS:**

**#nfsstat -m** [Will show remote mounted filesystems, protocol in use, NFS Version 2 or 3, and Read/Write Transfer Size, such as 32k; It will also tell you what the NFS timeout values are for minor & major timeouts]

**#nfsstat -rc** [Will display BadXid and Timeout Stats for NFS—if these two values are equivalent, may indicate Server Not Responding problem, resulting in Clients sending duplicate NFS requests; If BadXid near zero but Timeout value is large, would indicate a network problem resulting in packets being dropped]

**#nfsstat -cz** [If gathering statistics, run this command first to zero out historical information]

## **Other NFSSTAT Switches:**

**#nfsstat -cnr** [Client NFS & RPC Information, such as NFS & RPC Calls made and rejected; -n NFS -r RPC]

**#nfsstat -snr** [Server NFS & RPC Information; -n NFS -r RPC]

**#nfsstat -z** [zeroes out statistics]



**SOFT MOUNTS:** Soft-mounts are typically used for Read-Only mounts—a key indicator of a soft-mounted filesystem would be an NFS “RPC Timeout”—client retries a number of times before giving up.

## **NFS TIMEOUT MESSAGES:**

SERVER NOT RESPONDING—Client sees major timeout error on hard-mounted systems

RPC TIMEOUT—Client will only retry the connection a number of times before timing out—soft-mounted systems

**Note:** Timeout Errors can be adjusted via the “timeo” option on the Client

## **INCREASING NFS PERFORMANCE:**

### **USING NFSv3 V. NFSv2:**

#### **Use of TCP v. UDP Transport Protocol:**

Since UDP is connectionless, this transport protocol works best over clean and pristine networks. For WAN networks, or busy networks with slower servers, TCP might be best because it will reduce number of retransmissions. TCP only has to retransmit the frame that was lost—UDP needs to retransmit the whole Datagram, which can be a large number of packets. Duplicate XIDs as seen by the RPC layer would indicate a need to retransmit UDP. Also with UDP, Server cannot slow down or ‘flow control’ the Client, meaning that datagrams will be discarded and all needed to be retransmitted. TCP can use congestion avoidance, slow start, and zero window advertisement to control flow, UDP does not. TCP can use fast retransmission at Layer 4 where UDP relies on Layer 7.

#### **Write Throughput:**

Increased because the number of write requests made to server are aggregated, then written to server cache as “safe asynchronous writes”, which can reduce number of overall disk I/O requests to a Server

#### **Read Throughput:**

Sequential ‘read’ operations on client can be enhanced by caching file data locally from Server

Reduced File Attribute Requests: NFSv3 Servers return file attributes for all operations—Clients cache this information and less likely to have to ‘re-request’ file attribute information as a result

#### **Increased NFS BufferSize:**

NFSv2 only supported NFS buffer sizes of 8192kb

NFSv3 supports larger data transfer sizes for “read/writes” of 32k—provides better “read/write” performance

#### **Reduced Directory Lookups:**

Server returns all File Names listings and File Attributes in one operation v. piecemeal method in NFSv2

## **NFS VERSION 2 ERRORS:**

NFSERR\_PERM=1, Not owner--Caller does not have the correct ownership to perform the requested operation.

NFSERR\_NOENT=2, No such file or directory--does not exist.

NFSERR\_IO=5, Some sort of hard error occurred when the operation was in progress--disk error, etc.

NFSERR\_NXIO=6, No such device or address.

NFSERR\_ACES=13, Permission denied. The caller does not have the correct permission to perform the requested operation.

NFSERR\_EXIST=17, File exists. The file specified already exists.

NFSERR\_NODEV=19, No such device.

NFSERR\_NOTDIR=20, Not a directory. The caller specified a non-directory in a directory operation.

NFSERR\_ISDIR=21, Is a directory. The caller specified a directory in a non-directory operation.

NFSERR\_FBIG=27, File too large. The operation caused a file to grow beyond the server's limit.

NFSERR\_NOSPC=28, No space left on device. The operation caused the server's file system to reach its limit.

NFSERR\_ROFS=30, No space left on device. The operation caused the server's file system to reach its limit.

NFSERR\_NAMETOOLONG=63, File name too long. The filename in an operation was too long.

NFSERR\_NOTEEMPTY=66, Directory not empty. Attempted to remove a directory that was not empty.

NFSERR\_DQUOT=69, Disk quota exceeded. The client's disk quota on the server has been exceeded.

NFSERR\_STALE=70 The fhandle given in the arguments is invalid--the file referred to by that file handle no longer exists, or access to it has been revoked. The fhandle is the file handle passed between the server and the client. All file operations are done using file handles to refer to a file or directory. The file handle can contain whatever information the server needs to distinguish an individual file. A filehandle that consists of 32 zero bytes is called the public filehandle. It is used by WebNFS clients to identify an associated public directory on the server.

NFSERR\_INVAL The fhandle given in the argument does not refer to a symbolic link.

NFSERR\_NOSPC No space left on device. The operation caused the server's file system to reach its limit.

**Note:** We no longer support WebNFS!

## **NFS VERSION 3 ERRORS:**

NFS3\_OK = 0 Indicates the call completed successfully.

NFS3ERR\_PERM = 1 Not owner. The caller does not have the correct ownership to perform the requested operation.

NFS3ERR\_NOENT = 2 No such file or directory. The file or directory name specified does not exist.

NFS3ERR\_IO = 5 I/O error--a hard error occurred when the operation was in progress. This could be a disk error, for example.  
NFS3ERR\_NXIO = 6 No such device or address.

NFS3ERR\_ACES = 13 Permission denied. The caller does not have the correct permission to perform the requested operation. Contrast this with NFS3ERR\_PERM, which restricts itself to owner permission failures.

NFS3ERR\_EXIST = 17 File exists. The file specified already exists.

NFS3ERR\_XDEV = 18 The caller attempted to do a cross-device hard link.

NFS3ERR\_NODEV = 19 No such device.

NFS3ERR\_NOTDIR = 20 Not a directory. The caller specified a non-directory in a directory operation.

NFS3ERR\_ISDIR = 21 Is a directory. The caller specified a directory in a non-directory operation.

NFS3ERR\_INVAL = 22 Invalid argument or unsupported argument for an operation. Two examples are attempting a READLINK on an object other than a symbolic link or attempting to SETATTR a time field on a server that does not support this operation.

NFS3ERR\_FBIG = 27 File too large. The operation would have caused a file to grow beyond the server's limit.

NFS3ERR\_NOSPC = 28 No space left on device. The operation would have caused the server's file system to exceed its limit.

NFS3ERR\_ROFS = 30 Read-only file system. A modifying operation was attempted on a read-only file system.

NFS3ERR\_MLINK = 31 Too many hard links.

NFS3ERR\_NAMETOOLONG = 63 The filename in an operation was too long.

NFS3ERR\_NOTEMLTY = 66 An attempt was made to remove a directory that was not empty.

NFS3ERR\_DQUOT = 69 Resource (quota) hard limit exceeded. The user's resource limit on the server has been exceeded.

NFS3ERR\_STALE = 70 Invalid file handle. The file handle given in the arguments was invalid. The file referred to by that file handle no longer exists or access to it has been revoked.

NFS3ERR\_REMOTE = 71 Too many levels of remote in path. The file handle given in the arguments referred to a file on a non-local file system on the server.

NFS3ERR\_BADHANDLE = 10001 Invalid NFS file handle. The file handle failed internal consistency checks.

NFS3ERR\_NOT\_SYNC = 10002 An update synchronisation mismatch was detected during a SETATTR operation.

NFS3ERR\_BAD\_COOKIE = 10003 A REaddir or REaddirplus cookie is stale.

NFS3ERR\_NOTSUPP = 10004 The operation is not supported.

NFS3ERR\_TOOSMALL = 10005 The buffer or request is too small.

NFS3ERR\_SERVERFAULT = 10006 An error occurred on the server, which does not map to any of the valid NFS Version 3 protocol error values. Clients based on an XPG system may choose to translate this to EIO.

NFS3ERR\_BADTYPE = 10007 An attempt was made to create an object of a type not supported by the server.

NFS3ERR\_JUKEBOX = 10008 The server initiated the request, but was not able to complete it in a timely fashion. The client should wait and then try the request with a new RPC transaction ID. For example, this error should be returned from a server that supports hierarchical storage and receives a request to process a file that has been migrated. In this case, the server should start the immigration process and respond to client with this error.

## **CELLERA NFS ERRORS AND CAUSES:**

2006-05-02 13:05:34: NFS: 3: checkExportedByHandle: nodeFromHandle failed with status 9 for client 10.203.3.53

**Note:** checkExportedByHandle message can occur for a number of reasons:

→Occurs when Quotas treeid conflict is on the mountpoint of the file system

→Can occur if path is no longer exported

→Can occur if file system is no longer mounted

2006-04-13 13:21:09: UFS: 3: foundNode(valid), with treeId 5 different than passed treeId 1,please umount & re-mount fs 237 from nfs clients

**Note:** Can occur if Quotas treeid conflict is on the filehandle and mountpoint is o.k.

## **CELLERA BACKUP SOLUTIONS:**

### **I. MANUAL/LOCAL TAPE BACKUPS:**

--backups based on "server\_archive" PAX [Portable Archive Interchange] and standard unix formats such as CPIO/TAR

--backups do not go over network, thus, no network traffic

--direct Symm to Celerra to Tape connections via SCSI or Fibre

--no Scheduling, Cataloging, or TLU support or Media control

--provides for multi-protocol support, backing up both CIFS & NFS attributes using "server\_mt" commands

--conducted as either a Full or Incremental "Logical" backup

--Fibre Channel tape support only for LUNS 0-7

--MD5 authentication for NDMP now supported with CFS 4.0

--CFS 4.0 supports Level 10 NDMP Incremental Backup

--Celerra supports (2) tape drives per controller for a total of (4) Tape Drives per DM

## **MULTI-PROTOCOL BACKUP SUPPORT:**

Celerra supports Backup & Restore of multiple protocols (CIFS & UNIX) using Server\_Archive or NDMP

### **NDMP PAX:**

TAR-like archive protocol based on standard UNIX tape formats, provides file-level backup and recovery [server\_archive]

### **TAR:**

PAX format traverses file tree in depth-first order

### **NDMP DUMP:**

Veritas-style PAX backup by inode order vs. tar tree-style backup (mixed width-first depth-first order)

### **NDMP CCB:**

High-speed raw image backup

## **SERVER ARCHIVE COMMANDS:**

#server\_archive commands using tar/cpio/emctar formats. Uses Dart PAX technology [Portable Archive Interchange]

**Note:** Default format used with server\_archive is “ustar”. Ustar is an extended tar interchange format from the -p1003.2 standard, with a default block size of 10,240 bytes.

**\$server\_archive server\_2 -w -e device -x emctar -tvN /fullpath**

[contents of tape]

**\$server\_archive server\_4 -v -e device**

[Write to tape]

**#server\_archive server\_2 -w -e tape1 -x emctar -tvN /home/user/file**

[restore archive]

**#server\_archive server\_2 -r -f /fullpath -e device**

**Note:** Server\_archive -v -e functions broken in NAS 5.2/5.3—requires server reboot

## **SERVER MT COMMAND: \$server\_mt**

Used to perform manual tape operations--rewinding, unloading, etc, & to verify status of communications to tape drive.

### **REWINDING TAPE DRIVE:**

**\$server\_mt server\_4 -f c1t4l0 rewind**

**\$server\_mt server\_4 -f tape2 rewind**

### **COMMUNICATING TO TAPE DRIVE:**

**\$server\_mt server\_4 -f c128t0l1 status**

Device name: /dev/c128t0l1

Product Id: ULTRIUM06242-XX

Block size: 0

Block number: 1

**Note:** Following output means that the tape device cannot be read—i.e., empty, not defined, not cabled, etc.

Device Not Ready

Device Not Ready

## **CAN THE SERVER MT COMMAND DETERMINE & SET TAPE DENSITIES ON MEDIA?**

Answer: NO. Some Unix O/S's allow for the querying and setting of tape compression attributes, aka "multi density support".

### **With Solaris the following (4) commands can determine density settings for a DLT/7000:**

mt -f /dev/rmt/0l status -->20GB Uncompressed

mt -f /dev/rmt/0m status -->20GB Compressed

mt -f /dev/rmt/0h status -->35GB Uncompressed

mt -f /dev/rmt/0u status -->35GB Compressed

**Note:** When using these commands, only one will return without error, indicating the compression density mode on that tape.

How do I force a specific mode, usually the highest performance mode available for the tape?

On Solaris, just append an "l", "m", "h", or "u" to the /dev/rmt/{n} format device name to specify tape density.

l=20GB m=20GB compressed ~40GB h=35GB u=35GB compress ~70GB [DLT7000 Example]

**Cause:** Celerra uses "auto-sense" mode and defaults to tape density being loaded—it cannot manipulate or force specific densities

**Fix:** The "server\_mt" command can only manipulate & detect tape drives for the Celerra. A workaround is to load all tapes onto a general purpose Unix or NT system and write a dummy record, selecting the proper density/compression pseudo-device alias name. Once these tapes are pre-formatted, then they can be unloaded, and assigned to the Celerra NDMP media pool.

## **CELERRA SERVER ARCHIVE:** [For more detail on server\_archive, use #info server\_archive]

--Server\_archive reads, writes, and lists archive files and directory hierarchies & supports a variety of different formats

--Cannot be used to backup or restore between Servers

--Server\_archive uses PAX (Portable Archive Interchange) & other standard UNIX tape formats—TAR, CPIO

--Server\_archive offers multiprotocol backups of both Unix & Windows attributes

--PAX is the subsystem that provides both NDMP & Server\_archive services for Celerra

**Note:** *Default format used with server\_archive is “ustar”.* Ustar is an extended tar interchange format from the -p1003.2 standard, with a default block size of 10,240 bytes.

## **BACKUP ARCHIVE FORMATS:**

**Ustar:** Unix Standard called 'extended tar interchange'—Celerra's Default format referenced in -p1003.2

Default blocksize is 10,240 bytes and max pathnames are 250 characters (150 char. for basename, 100 for filename)

**Emtar:** EMtar is the format of data written to disk or tape through PAX, and is not valid for Veritas or other vendor applications. Can be used to backup CIFS attributes. Not compatible with "ustar". Can "archive" files > than 8GB to an archive name. Pathnames are limited to 3070 characters. Not -p1003.2 compatible. Otherwise performs as Ustar.

**Caution:** Backing up Celerra filesystems to an NT box destroys UNIX file attributes

**Cpio:** Extended interchange format found in -p1003.2 standard. Default blocksize is 5120 bytes.

**Tar & PAX:** Tar is the old BSD format found in BSD4.3 and is what is used by Backup vendors. Default block size is 10240 bytes. Pathnames must be 100 characters or less. Only regular files, directories, & hard & soft links are archived. Backup methodology goes from bottom of directory tree to top—depth first order. Tar sends file details with full path names.

**Dump:** This is also the "server\_archive" format and traverses file system in mixed width-first, depth-first order. Dump sends file history in form of inodes and file names as a tree.

**NDMP:** NDMP Backups generate “emtar” data images or NDMP Dump formats, which can be read by ‘server\_archive’

## **TYPICAL SERVER ARCHIVE FORMAT RESTRICTIONS:**

--file pathname length

--file size

--link pathname length

--file type

## **ARCHIVING FILES TO TAPE:** [Tape direct-attached to DM]

**\$server\_archive server\_4 -w -f /dev/c1t4l0 -v /fs01** [Backs up files to tape device file]

**\$server\_archive server\_4 -f /dev/c1t4l0 -v** [Lists files from tape device file]

**\$server\_archive server\_4 -r -f /dev/c1t4l0 /fs01** [Restores file to tape device file]

**\$server\_mt server\_4 -f c1t4l0 rewind** [Rewinding tape]

**\$server\_mt server\_4 -f c1t4l0 status** [Tape drive status]

## **MIGRATING & ARCHIVING DATA TO TAPE:** [Using either Tar or Cpio formats]

**Note:** Useful for backing up files from network computer to tape, & then restoring to Celerra Filesystem

### **Archiving Using Relative Path:**

\$tar cvf /dev/rmt/0

### **Archiving Using Absolute Path:**

\$tar cvf /dev/rmt/0 /full\_pathname

### **Listing Archive Files from Tape:**

\$tar tvf /dev/rmt/0

### **Listing Archive Files from Tape Using Server Archive:**

\$server\_archive server\_2 -r -f c1t4l0 -v

### **Restoring From Tape Using Server Archive & Relative Path:**

\$server\_archive server\_2 -r -f c1t4l0 -s, ^ ., /path\_name\_new,

### **Restoring From Tape Using Server Archive & Path Beginning with Dir:**

\$server\_archive server\_2 -r -f c1t4l0 -s, ^ ., /path\_name\_new,

### **Restoring From Tape Using Server Archive & Absolute Path:**

\$server\_archive server\_2 -r -f c1t4l0 -s, ^ /path\_name\_new,

## **USING SERVER ARCHIVE TO LIST EMCTAR CONTENTS:**

**\$server\_archive server\_3 -f /fs01**

## **USING SERVER ARCHIVE TO BACKUP & RESTORE:**

### **Writing From Tape to FileSystem:**

**\$server\_archive server\_x -w -J -e tape1 /cifs/homedirs**

**Note:** This will backup UNIX/NT Names + Special NT Attributes + ACLs

**Archiving CIFS Attributes From One FileSystem to Another:**

**\$server\_archive server\_3 -J -rw /fs01 /emc** [-J switch preserves CIFS Attributes]

**Archiving NFS Attributes From One FS to Another:**

**\$server\_archive server\_3 -w /fs01 -f /emc** [Backs up /fs01 using emtar to /emc]

**Restoring EMCTAR Archive to Original Location:**

**\$server\_archive server\_3 -r -f /emc /fs01** [Restores archive to original location]

**Archiving From FileSystem to Tape:**

**\$server\_archive server\_3 -w -f /dev/c128t0l1 /fs\_d3** [Backs up /fs\_d3 using emtar to tape]

**USING "SERVER\_ARCHIVE" TO BACKUP & RESTORE FROM TAPE:**

**Conducting BackUp From File System to Tape Drive:**

1. Insert tape into drive and run status command

**\$server\_mt server\_4 -f c128t0l1 status**

2. Rewind Tape & run status command again [Need to see "Block number: 0"]

**\$server\_mt server\_4 -f c128t0l1 rewind**

3. Conduct Backup To Tape:

**\$server\_archive server\_4 -w -f /dev/c128t0l1 /fs\_d3**

4. Restoring from Tape to File System:

**Note:** Run steps 1 & 2 prior to conducting restore

**\$server\_archive server\_4 -r -f /dev/c128t0l1 /fs\_d3**

**CAUTION:** In practical terms, "server\_archive" is an "all or nothing" Restore, meaning that it is impractical to restore individual files if you have a large backup file. "Server\_archive" can only "restore" to the original location as well.

**USING "SERVER\_ARCHIVE" TO BACKUP & RESTORE FILES:**

**Backups:** \$server\_archive server\_2 -v -w -f /fleetmnt/archive /testfs [Archive Name and Source Folder, respectively]

**Restores:** \$server\_archive server\_2 -r -i -f /fleetmnt/archive /testfs

**Note:** Using the "i" interactive switch will prompt you for each file that you would like restored [i.e., impractical if you had thousands of files because you cannot specify an individual file by path]

**SERVER\_ARCHIVE SWITCHES:**

-r      Reads archive file from "archive name" and extracts specified files  
-w      Writes to "archive name" in standard input or format specified  
-r -w    Copies  
-0      Full referenced backup put into a reference file stored ./etc/BackupDates  
-e      archive\_name [Specifies archive name when it is a streamer]  
-f      archive\_file [Specifies archive name when it is a file]  
-i      interactive mode to specify files, etc.  
-k      will replace only "non-existing" files and will not overwrite existing files  
-I      Allows filenames to be translated into UTF-8  
-p      string [a, e, m, o, p ]  
-t      Resets access time to same time before server\_archive accessed files  
-J      Backs up, Restores, or Displays CIFS extended attributes [u w d]

**# info server\_archive** [for more complete detail of all switches]

**DATA MOVER RECORDS:**

Successful server\_archive operations, including NDMP tape operations which use the server\_archive function, records successful Full or Incremental Backups or Restores in the .etc directory of the DM: **./etc/BackupDates**

**\$ cat BackupDates |grep -i /fs00621/1/GRPF-HRS\_EnvelopeManagement**

|                                        |   |          |                     |
|----------------------------------------|---|----------|---------------------|
| /fs00621/1/GRPF-HRS_EnvelopeManagement | 0 | 3f2376f5 | 06:53:41 07/27/2003 |
| /fs00621/1/GRPF-HRS_EnvelopeManagement | 5 | 3f29fd6e | 05:41:02 08/01/2003 |

**II. NETWORK-BASED BACKUPS:**

--Generally speaking, NFS clients backup & restore Unix attributes, CIFS clients backup & restore CIFS attributes  
--tape devices typically attached to a Backup Server using standard backup products such as; EDM/Epoch, ADSM, Openvision, Cheyenne, Legato Networker, Veritas NetBackup, ARCServe  
--high network traffic load created using this method

--scheduling, cataloging, and TLU support provided  
--person performing backup must be member of Celerra's Backup Operator's Local Group  
--EDM involves backing up from Symm to DataMover and across network to EDM attached TLU  
  EMC Data Manager (EDM) 4.6/EDM 4.6 & EDM NT Client = NFS/CIFS backups, respectively

### **III. CELERRA BACKUP-to-DISK SOLUTIONS (LAN B2D or NDMP B2D):**

#### **CELERRA B2D:**

##### **Celerra offers two methods for backing up data to disk:**

NS700 LAN B2D → Requires configuration on Veritas or Legato Backup host to accomplish  
NDMP B2D → Requires configuration of VTLU, NAS 5.4, and NDMP Backups to disk-based storage  
Veritas NetBackup is LAN B2D qualified with v4.5+ and Celerra 5.4, set param cifs prealloc=6 for 512kb block writes  
Legato NetWorker is LAN B2D qualified with v7.0+

##### **FEATURES/RESTRICTIONS:**

VTLU configuration required for NDMP

Disks must be presented from Celerra

Server\_archive or server\_mt are not valid troubleshooting tools for VTLU configurations

VTLU requires dedicated file system

After backing up to disk with either LAN or NDMP B2D solution, data can be cloned to tape for offsite storage

##### **TROUBLESHOOTING:**

LAN B2D would require customer network investigation and traces

NDMP B2D would require configuration checks, file system layouts, etc.

### **IV. CELERRA NDMP PROTOCOL-BASED BACKUPS: NDMP=NAS 2.1.15.20 +**

##### **Intro:**

Network BackUp architecture based on a Client/Server model using NDMP protocol with backup software on NDMP Host Client, backing up using NDMP Server [Celerra]. NDMP Protocol is based on XDR encoded messages transmitted over TCP/IP. NDMP uses asynchronous XDR-encoded messages over bi-directional TCP/IP channels that are grouped together by NDMP Category or Interface (Such as SCSI, TAPE, DATA, etc). Actual Backup/Restore of Data is handled completely on the NDMP Server side, with Client Host Software providing control, etc

##### **BASIC NDMP BACKUP OPERATION:**

NDMP Client [NDMP Management SW] initiates Backups/Restores, indirectly accessing Tape devices via NDMP protocol interfaces. NDMP Client is not directly involved in any DATA movement--this is the NDMP Server's responsibility.

NDMP Client accesses tape via "pass through" commands.

NDMP Server does not require a special BackUp agent but uses the Media Agent provided by the NDMP Client instead and writes data to tape using "dump" or "tar" format.

**Note:** Celerra Data Movers serve as the "tape/media" Server, while NDMP Host serves as the "Console" providing Scheduling, Cataloging, and Tape Management functions.

##### **NDMP BackUp Flow:**

Data flows from Storage system to the DM to an attached Tape Library device, direct-attached, but not over network—only Backup SW control data for Scheduling, Cataloging, and TLU commands traverse network.

##### **CELERRA PAX SUPPORT:**

PAX is the Celerra subsystem that supports both NDMP and server\_archive services

PAX supports up to (4) concurrent archive streams

PAX code ensures that file names are the same during the RESTORE operation as in the original backup

**Note:** NAS 5.4.16.106 (see AR62002) will increase supported NDMP streams to 16 sessions.

##### **NDMP BEST PRACTICES:**

- Optimize file system layout when possible
- Use RAID-3 LUNs for large sequential IO for ATA drives
- Balance LUNs across SPs for file systems being backed up
- Limit IO from other applications while backing up or restoring file systems
- Ensure tape drives perform digital speed matching
- For DAR backups, ensure NMDP Client can handle history workload
- Use supported backup vendors and TLU solutions
- Enable CLARiiON read cache for prefetching
- Use jumbo Ethernet frames for 3-way backups

##### **ENVIRONMENTAL VARIABLES ASSOCIATED WITH NDMP & VENDOR SUPPORT:**

|                  |                                                                                                           |
|------------------|-----------------------------------------------------------------------------------------------------------|
| BASE_DATE        | →Clientside, works with dump_date, ignores LEVEL if set                                                   |
| DUMP_DATE        | →Serverside, works with base_date for incremental backups                                                 |
| DIRECT           | →Set to Yes for DAR                                                                                       |
| EMC_EDIR         | →use to exclude directories from backup, uses filter.numDirFilter (1-50)                                  |
| EMC_EFILE        | →use to exclude files or filetypes from backup, uses filter.numFileFilter (1-50)                          |
| EMC_OFFLINE_DATA | →Set to No by default to only backup offline attributes. Set to Y will backup attributes & data.          |
| FILESYSTEM       | →path to be backed up                                                                                     |
| HIST             | →Creates file history if set to y                                                                         |
| LEVEL            | →dump levels 0-10                                                                                         |
| OPTION           | →NT=NT attributes; LK=follow sym links; AT=preserve access times; MI/MC/MM=collision policy I18N restores |
| PREFIX           | →use in place of FILESYSTEM for backward compatibility                                                    |
| RECURSIVE        | →required for 5.5+, performs recursive restores when set to y                                             |
| SNAPSURE         | →uses SnapSure checkpoints for NDMP backups when set to y                                                 |
| TYPE             | →backup format tar, dump, or vbb                                                                          |
| VLC              | →Volume level copy to tape using RO file system; set to n by default                                      |

### **COMMENT ABOUT TLU (Tape Library Unit) SUPPORT:**

TLUs are not directly qualified by NAS, only Tape Drives are tested for support. However, the rule is that if the Backup Vendor software is listed in NSM as supported, if the Tape Drives and Bridges are listed as supported, and the TLU, Tape Drive, & Bridges are shown as supported by the Vendor software, then the entire configuration would be considered supported by NAS.

### **NAS 5.2 NDMP SUPPORT:**

#### **NASS:**

Thread that reads directories and file metadata and provides filenames to NASA thread via a Stat Buffer

#### **NASA:**

Archive thread that reads files and provides history information to NDMP Client

5.2 introduces multiple NASA threads for backups (not restores), see nThread parameter

5.3 introduces multiple nasA threads for Backups and for Restores

#### **NASW:**

Threads that write data to the tape unit

### **TROUBLESHOOTING NASW THREADS:**

#### **# .server\_config server\_2 -v "diag nasw"**

nasw00 No session found

nasw01 No session found

nasw02 No session found

nasw03 No session found

#### **# .server\_config server\_2 -v "diag tape"**

tape[0] No information found

tape[1] No information found

tape[2] No information found

tape[3] No information found

#### **USAGE STATEMENTS:**

#### **# .server\_config server\_2 -v "diag help"**

1138037924: PAX: 3: Usage: diag nasw

1138037924: PAX: 3: diag nasw=0..3

1138037924: PAX: 3: diag nasw=reset

1138037924: PAX: 3: diag tape

1138037924: PAX: 3: diag tape=0..3

1138037924: PAX: 3: diag tape=dev\_name blksz=kb(64) total=mb(1024) pattern=rand/(text)/zero

### **NDMP BACKUP PERFORMANCE CRITERIA/NDMP TROUBLESHOOTING:**

--File System layout is a consideration when evaluating NDMP performance throughput

--Raid Group structure is also important on Clariion systems (Raid5 4+1 & clar\_r5\_performance)

--If using (4) NDMP streams, overlapping file systems sharing same spindles would impact performance throughput

--Disk Read performance is paramount in importance and needs to be tuned for sequential access by backup applications [Using prefetch and more read cache are things that may help with Clariion systems]

--File names and metadata are often stored in separate locations, often affecting backups of many small files

--Use of DAR imposes more overhead on NDMP Clients and impede performance

--Use Server\_Archive dump when testing backup performance of Celerra

## \$ time server\_archive server\_2 -w -x emctar -J -b 60k -e /dev/cc32t0l0 /fs1

--PAX Backup performance statistics are available for tuning and troubleshooting

--MPD file systems yield higher performance throughputs

--Use TAR vs. DUMP file ordering whenever possible [default mode for Celerra]

--Using HIST with many very small files can cause performance overhead issues with job history archiving

--Typical Backup performance for NAS 5.2 would be 12MB/sec or higher [50-112GB/hour]

**Note:** Typical performance with NAS 5.6 is 40MB/sec

--Upgrade to NAS 5.3--new params and defaults are in place [clarion queue depth, FS traversing threads, etc]

## **NDMP PERFORMANCE COMPONENTS:**

Server, BackEnd, Client, TLU, Bridge, Drive, Switch, Cables

## **TROUBLESHOOTING 5.2 NDMP:**

### **\$server\_config server\_x -v “printstats pax full”**

**Note:** Outputs stats on NASA & NASW threads. NASA contains the backup/archive file reader threads whereas NASW concerns writing to tape. Get & Put Pools serve as buffers between NASA & NASW threads.

\*\*\*\*\* SUMMARY PAX STATS \*\*\*\*\*

#### ---- NASA STATS ----

cifs\_ea new count: 515079, delete count: 515077

\*\* nasa thid 0 \*\*

Backup root directory: /fs1

Total bytes processed: 15978103777

Total file processed: 515077

throughput: 5 MB/sec

average file size: 30KB

Total nasa wait nass count: 236012

Total nasa wait nass time: 2464059 msec

Total time since last reset: 2666 sec

Tape device name: /dev/c80t0l0

nasa01 isnot doing backup

nasa02 isnot doing backup

nasa03 isnot doing backup

#### ---- NASW STATS ----

nasw00 BACKUP (in progress)

Session Total Time: 00:44:26 (h:min:sec)

Session Idle Time: 00:37:17 (h:min:sec)

KB Tranferred: 15603588 Block Size: 64512 (63 KB)

Average Transfer Rate: 5 MB/Sec 20 GB/Hour

Average Burst Transfer: 35 MB/Sec 124 GB/Hour

Write Block Counters: 246225/1451 (List/Direct)

Point-in-Time (over the last 10 seconds):

Rate=2 MB/Sec Burst=55 MB/Sec Idle=948 msec/sec

Get Pool: 0 buffers Put Pool: 639 buffers

## **Get Pool:**

Buffer used for thread used to write data . Value >than 0 indicates that write thread has data available and is not waiting. 0 value indicates thread is waiting. Increase nThread value if this value is 0 and NASA wait counts are low.

## **nThread:**

Specifies number of backup threads used by NASA file reader for backup stream. Increasing value with small file sizes may enhance performance.

## **Put Pool:**

Buffer used for file reading threads. When this value is 0, tape unit is limiting the backup speed. Non-zero values indicate buffers available for reader threads.

**Note:** NDMP backups use a data buffer size of 60kb [Use server\_mt status command to verify tape buffer size]

## **paxStatBuff:**

Number of buffers available to the NASS thread. Increase this value if NASA wait counts or wait times are high.

## **nPrefetch:**

Number of buffers that NASA thread will prefetch

## **paxWriteBuff:**

Total number of buffers in GetPool and PutPool. Increase this value if GetPool buffers fall to zero. This buffer helps keep data streaming to tape.

## **TAPE DRIVE TIMEOUT PARAM:**

**param streamio timeout=180** →Used to increase read/write timeout value when a tape drive takes too long to respond to a single tape command. For NAS 5.5, default value hex 708, 1800 decimal, in secs. = ½ hour.

**NAS 5.5 PAX PARAMS:**

| Name                       | Current    | Default    |                                                |
|----------------------------|------------|------------|------------------------------------------------|
| PAX.allowVLCRestoreToUFS   | 0x00000000 | 0x00000000 |                                                |
| PAX.checkUtf8Filenames     | 0x00000001 | 0x00000001 |                                                |
| PAX.dump                   | 0x00000000 | 0x00000000 |                                                |
| PAX.filter.caseSensitive   | 0x00000001 | 0x00000001 |                                                |
| PAX.filter.dialect         | " "        |            | →Use when restoring ascii-mode tape to I18N DM |
| PAX.filter.numDirFilter    | 0x00000005 | 0x00000005 |                                                |
| PAX.filter.numFileFilter   | 0x00000005 | 0x00000005 |                                                |
| PAX.noFileStreams          | 0x00000000 | 0x00000000 |                                                |
| PAX.readWriteBlockSizeInKB | 0x00000040 | 0x00000040 |                                                |
| PAX.scanOnRestore          | 0x00000001 | 0x00000001 |                                                |
| PAX.writeToArch            | 0x00000001 | 0x00000001 |                                                |
| PAX.writeToTape            | 0x00000001 | 0x00000001 |                                                |

**param PAX.nFTSThreads** →hex & decimal 8; number of threads used for single backup job—increase to reduce NASA thread bottleneck

**param PAX.nPrefetch** →hex & decimal 8; amount data each backup thread prefetches from disk before finishing file read & writing to tape

**param PAX.nRestore** →hex & decimal 8; number restore threads per job. Increasing can help restore speed improve

**param PAX.nThread** →hex 040, decimal 64; number of backup threads used

**param PAX.paxReadBuff** →hex 040, decimal 64; →number of buffers between threads that read data from tape (NASA) & threads that restore to disk (NASW). Increasing can enhance tape streaming and restore speeds.

**param PAX.paxStatBuff** →hex 080, decimal 128; →number of buffers between threads that send & receive metadata (NASA) and read data from disk (NASS). Increasing can increase speed of metadata info.

**param PAX.paxWriteBuff** →hex 040, decimal 64; no. buffers between threads that read file data (NASA) & write data to tape (NASW)

**CELLERRA NDMP SUPPORT:**

--With NDMP Version 4, NDMP Client=Data Management Application, and NDMP Server=Data Service Provider

--Provides multi-protocol backups/restores [NT Attributes require use of HIST=Y and OPTIONS=NT

--Scheduling, cataloging, and TLU support for up to (8) Tape Drives, but only (4) streams at any given time

**Note:** Driver can see a maximum of 16 tape devices down any given FA port

--NDMP dump uses a control and backup data path

--Recommended that tar backup be used with Celerra vs. dump backup

--High Capacity--minimal network traffic

--Higher capacity with up to 4-simultaneous backup streams to each datamover

--Centralized control & media management, separation of Control and Data paths.

--Local backup using SCSI Tape Libraries and/or drives

--Support for NDMP Drive sharing through SANs only over Fibre

--NDMP does not support Archive Bit method, only Date/Time stamp method (i.e., only mtime is changed for backups)

--NDMP does not support use of wildcard characters, or exclude/include options

--3-Way Backups/Restores, but only with other Celerra DMs

--I18N support with default translation dialect of 8859-1

--DAR restores with HIST=Y and DIRECT=Y set

--NDMP Restores are not supposed to be intercepted by Virus Checking and scanned, however, AR45645 proves that this was happening, and causing customer issues. Set following param to disable VirusChecking on NDMP Restores:

**param pax scanOnRestore=0** [NAS 5.3, 5.1.26.x, 5.2.15.x]

**NAS 5.4:**

Celerra has capacity for (16) NDMP Jobs (still only 4 concurrent streams) if using (2) NDMP buffers per session, or (32) Jobs if using only a single NDMP buffer.

**HOMOGENEOUS 3-WAY NDMP BACKUP:**

--Data Server and Tape Server are on different Data Movers, with NDMP Host, and data is backed up over the Network

**HETEROGENEOUS 3-WAY NDMP BACKUP:**

--Data Server is on DM but Tape Server is on non-Celerra Server, with NDMP Host Server, and data backed up over Network

--Veritas NetBackup 4.5

**BASIC NDMP COMMUNICATION INTERFACES:**

**CONNECT:** Interface is responsible for establishing NDMP connection between Server & Client, authenticating Client, and negotiating version to use (Secure MD5 digest using special param file entry or clear-text). Uses NDMP well-known port #10000.  
**CONFIG:** Interface where Client discovers NDMP Server configuration and attributes (Tape Drives, Jukeboxes, File Systems, Database)

**SCSI:** This interface is used to pass SCSI CDB (Command Descriptor Blocks) commands to a SCSI device & receive status back. NDMP Client software interprets the status and data information received via SCSI commands. NDMP client uses SCSI to control locally-attached jukeboxes. Called a “Pass through” SCSI device driver.

**TAPE:** This interface supports control of Tape Devices, such as positioning of tape read/write operations for Backups/Restores, & header & trailer file writes. Backup software uses Interface to maintain Tape Labels, Tape Format, Tape positioning.

**DATA :** NDMP client initiates Backups/Restores with this interface. Generates Backup image & retrieves data from this image. Handles parameters and formatting. Performs the actual task of "streaming" NDMP data to and from Tape.

**MOVER:** This interface controls the reading/writing of data from or to a tape device. During Backups, this interface reads data from DATA Interface, buffers into tape records, and then writes data to tape device. During Restores, this interface reads data from Tape Device and writes data to DATA Interface.

## **NDMP CLIENT/SERVER MESSAGES:**

NOTIFY: Used by Server to obtain Client response

FILE HISTORY: File-by-file record during Backups

LOGGING: Informational & diagnostic data to a log file

## **3-WAY NDMP BACKUPS/RESTORES: NDMP v.2**

BackUp Server initiates backup job to datamover, which reads data from Symmetrix and streams data over the network to a datamover connected to the tape device, which then receives the data stream and writes to tape, all using NDMP. As opposed to direct-attached local backup, both File System & Control Data traverse the network.

**Note:** 3-Way NDMP Backups are only supported where the Celerra DM servers as NDMP Tape Server (direct-attached). Port 1000 is used for listening for NDMP requests, while ports 10001 – 1004 are used for 3-way Backups and Restores. Local Backups/Restores do not use these extra ports.

**Example:** An NDMP Sun “Client” serves as the Backup Station accessing the NDMP Server (DataMover), which pulls the filesystem data from the Symm and pushes to the Tape Device Station.

**Caution:** Ensure “params” are set properly for all Servers involved with NDMP Backups, not just ones attached to Tape Drives!

NDMPv1: Local Tape Backup

NDMPv2: Remote Data Stream Backup from one DM to another

## **NDMP SUPPORT:**

### **NDMP v4 SUPPORT:**

--Supported with NAS NAS 5.2

**param NDMP maxProtocol=4** [Default setting for 5.2. Set to 1, 2, or 3 if customer is using other NDMP Version]

**param NDMP.maxProtocolVersion** 0x00000004 0x00000004 [NAS 5.5 syntax change]

--NDMP to run as TCP sevice on DART, listening on port 10000

--NDMP Client and Dart to negotiate protocol version for each backup/restore

--PAX implementation to be changed in order to change File History handling to conform to version 4

### **NDMP V2 & V3:**

--Both NDMP & Server\_archive allows for simultaneous backup or restore of NT/Unix attributes due to Logical-level vs. Volume-level file backup

--Default DataMover port for NDMP=10000

--Shared tape drives only over a SAN environment [only Legato & IBM Tivoli offers this]

--Total of (4) Tape Drives per DM

--DM can host up to (4) simultaneous BackUp jobs

--Windows Backups require that User account be added to Celerra’s local “Backup Operator’s Group”

--NDMP uses Tar or Dump formats to write to tape

## **NAS 5.5 RESTORE ISSUE ON BACKUPS CREATED WITH PRE-NAS 5.5:** AR84993

In certain cases, the combination of the DMA settings and the pre-5.5 backups sets both environmental variables, Direct=N and Recursive=N, in which case the restore only restores the directory, and not the subdirectories and files. In order to circumvent this issue, turn on the following param on the DM to allow full recursive restore, new with NAS 5.5.25.1 +:

**param NDMP.forceRecursiveForNonDAR=1**

**DAR:** Direct Access Restore

**DDAR:** Directory Level DAR restore—Not yet supported by DART 5.3 or lower, though Legato NetWorker V7.1.1 and Dart 5.1.11.300 or higher will restore directory permissions successfully. Both DAR & DDAR are NetApp implementations. NAS 5.5 introduces DDAR support.

**DDS:** Dynamic Drive Sharing, supported NAS 5.1 and above. Set following param to 0 if using DDS with NDMP vendors.

### # server\_param server\_2 -facility NDMP -info scsiReserve

server\_2 :

```
name      = scsiReserve
facility_name = NDMP
default_value = 1
current_value = 1
configured_value =
user_action = reboot DataMover
change_effective = reboot DataMover
range     = (0,1)
description = Do scsi reserve for NDMP tape open
```

### # server\_param server\_2 -facility NDMP -modify scsiReserve -value 0

Note: With NAS 6.0, the scsiReserve parameter is set to 0 by default, i.e., do not do scsiReserve

## **NDMP VENDOR MATRIX:**

### **A TEMPO TIME NAVIGATOR (Quadratoc):**

Version 3.7 SP0 p946 3-way & DAR NAS 5.1; SP0 P1110 for NAS 5.2.10+

Version 3.7.0.3 SP1 p1270—NDMPv3 tar format, 3-Way & DAR—NAS 5.3 only

Version 3.7.0.4 SP2 p1929—NDMPv3, tar, 3-way, DAR—NAS 5.4.21.4

Version 4.0 SP2 p1929—NDMPv3, tar-3-way, DAR—NAS 5.4.21.4

Version 4.0.1—NDMPv4, tar, 3-way, DAR, NAS 5.5 for limited features only

Version 4.1—V4, tar, 3-way, DAR, no DDS, NAS 5.5 core, NDMP2D, Integrated Checkpoints

Version 4.2 SPO P3267--V4; dump/tar; 3-way; DAR; DDS; DDAR; NAS 5.6; NDMP2d; Filters; Int. Ckpts; NVB

### **BAK BONE NETVAULT 7.0/7.1:**

Version 7.1+ SPE or 7.1.1 NDMP APM v6.3 with PCHR8 or 6.301 NDMP v4, 3-way, DAR, DDS, NAS 5.3.14.2

Version 7.1.1, NDMP APM 6.302 NDMP v4 dump format, 3-way, DAR, DDS, NAS 5.4 only

Version 7.1.2, V4, dump, 3-way, DAR, DDS, NAS 5.5 Core, NDMP2D, DDAR, Integrated checkpoints

Version 7.4.2, V4, dump, 3-way, DAR, DDS, NAS 5.5 Core, NDMP2D, DDAR, Integrated checkpoints, & VBB

Note: Celerra supports NDMP v4 in NAS 5.2+

Version 7.4.5 APM 6.5, V4, Dump, 3-way, DAR, DDS, NAS 5.5

Version 8.0 NDMP APM 7.0.6—V4; dump; 3-way; DAR; DDS; NDMP2D; DDAR; Int. Ckpts; NVB

Version 8.2 NDMP APM 7.1.5—“

### **BRIDGEHEAD HYPERTAPE:**

Version v3.1.3 NMDP Agent (v)—NDMP v3 dump/tar format, 3-way, DAR, DDS, NAS 5.4.19.4 only

Version 3.1.5 NDMP Agent v, ay—V3; dump/tar; 3-way; DAR; DDS

### **CA BRIGHTSTOR ENTERPRISE BACKUP (BEB):** v10.5-b2290 patch QO44763 Win2k

Supports NDMP v3, 3-Way, & DAR on NAS 4.2 & 5.1

Brightstor Enterprise Backup V10.5-b2290 patch QO44763 W2K NDMP v3, 3-way & DAR, NAS 4.2 & 5.1

V10.5 SP1 patch QO53005 W2K NDMP v3, 3-way & DAR, NAS 5.2

V10.5 SP1 patch QO57428/QO54460 NDMP v3, dump format, 3-way & DAR, NAS 5.3 only

Brightstor ARCserve Backup V11.0/11.1 NDMP v3 dump format, 3-way & DAR, DDS (v11.1), NAS 5.1 (V11.0), 5.2, 5.3

Note: BEB not qualified for NAS 5.4. & is being obsoleted (BAB r.11.1 does work with 5.4.15.2)

### **CA BRIGHTSTOR ARCSERVE BACKUP (BAB):**

V11.0 NDMP v3 dump format 3-way & DAR; NAS 5.1-5.3

V11.1 includes DDS support; NAS 5.3 only

V11.1 patches QO63594, QO72174, QO72924, QO73743 NDMPv3 dump format, 3-way, DAR, DDS, NAS 5.4 only

V11.5 patch Q076318 V3, dump, 3-way, DAR, DDS—NAS 5.4.21.4 only

V11.5 SP2 patch Q084383 V4, dump, 3-way, DAR, DDS, Core, NDMP2D, DDAR, Filters, Integrated Checkpoints NAS 5.5 only

V11.5 SP4 patch RO02596, RO06497 (bb) V4, dump, 3-way, DAR, DDS, Core, NDMP2D, DDAR, Filters, Integrated Checkpoints

V12.0 SP2 with T31A531, same as above NAS 5.6

V12.5 RO07441, RO09048, T36A558, T31A530, same as above NAS 5.6

### **COMMVAULT GALAXY:**

Version 5.0 SP1 & 5.9, NDMP v4, 3-way & DAR, NAS 5.2 & 5.3

Version 5.9 SP1 NDMP v4 dump format, 3-way & DAR; NAS 5.4 only

V5.9 SP4 update 671, v4, dump, 3-way, DAR, DDS—NAS 5.5 for limited features only

V6.1 SP1, v4, dump, 3-way, DAR, DDS—NAS 5.5 for limited features only

V7.0, V4, dump, 3-way, DAR, DDS, NAS 5.5 core, NDMP2D, DDAR, Filters, Int.Checkpoints

Note: Disable SCSI-3 Reservations if in use with CommVault

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!

V7.0 SP4a—V4; dump; 3-way; DAR; DDS; NAS 5.6; NDMP2D; DDAR; Filters; Int.Ckpts; NVB; Tape Silvering

V8.0 SP1 updates 11638, 11823—Same as V7.0 SP4a

### **EDM:**

EMC Data Manager 4.6.1/5.0: NDMP v2; 3-Way; no DAR; no DDS; Dump format

EDM 4.6.1 = NDMP v2 with 3-way only for NAS 4.1.12.1, 4.2.9.0, 5.0

EDM 5.0 = NAS 4.2.18.2, 5.0 & 5.1; MPFS NOT supported

EDM v5.1 NDMP v2 dump format 3-way only, NAS 4.1.12.1, 4.2.18.2, 5.0, 5.1, 5.2, 5.3, 5.5 with EDM patch 9--PE510\_09

### **FUJITSU SIEMENS NETWORKER:**

V7.0 patch NSx70A001 NDMP v3, 3-way & DAR NAS 5.1

V7.1A00 patch NSx71A001 NDMP v3 dump/tar format, 3-way & DAR NAS 5.2 & 5.3

### **HP OMNIBACK II/HP DATA PROTECTOR:**

Omniback II 4.1: No 3-way, has DAR, NAS 2.2.53.4; NAS 4.0, 4.1, 4.2, 5.0

OpenView Storage Data Protector 5.0/5.1, NDMPv3 dump format, No 3-way, DAR; NAS 4.1 – 5.1 are all qualified

Version A.05.50 NDMPv4, dump, DAR--NAS 5.2, 5.3, 5.4, & limited features in 5.5

OpenView Storage Data Protector 6.0, v4, dump, DAR,--NAS 5.4 & limited features in 5.5, Core, NDMP2D, Filters, Int. Checkpoints

Data Protector version 5.5 is in qual process (HP DP 5.1 is not supported with NAS 5.2 or 5.3)

### **IBM TIVOLI TSM:**

Version 5.2 NDMP v3 dump format with DAR & DDS; NAS 5.1.9.4 – 5.4; No 3-way backups

IBM Tivoli Storage Manager Version 5.3 v3, dump, DAR, DDS—NAS 5.3, 5.4, and limited 5.5 features

V5.4 V4, dump, DAR, DDS, NAS 5.5 Core, NDMP2D, Int. Checkpoints

V5.5—V4; dump; DAR; DDS; Core 5.6; NDMP2D; Int.Ckpts

V6.1; V4, dump, no 3-way, DAR, DDS, CORE, NDMP2D, Integrated Checkpoints NAS 5.6

### **EMC NETWORKER with NDMP (Legato):**

V7.0 NDMP v3 dump/tar NAS 4.1 to 5.1

V7.1/7.1.1 NDMP v4 dump or tar format, 3-way, DAR, DDS, NAS 4.1 to 5.4

V7.2 NDMPv4, dump/tar format, 3-way, DAR, DDS, NAS 5.1 to 5.4

V7.3 NDMPv4, dump/tar, 3-way, DAR, DDS, NAS 5.2 – 5.4, & most features in 5.5

V7.4 V4 dump/tar, 3-way, DAR, DDS, NAS 5.5 core, NDMP2D, Filters, Int.Checkpoints, VBB

V7.4.1, same as 7.4 plus DDAR

V7.5—V4; dump/tar; 3-way; DAR; DDS; NAS 5.6; NDMP2D; Filters; Int. Ckpts; NVB; Tape Silvering

EMC NETWORKER Fast Start with NDMP:

V7.4 SP3—v4; dump/tar; 3-way; DAR; DDS; Core 5.6; NDMP2D; Filters; Int.Ckpts; NVB; Tape Silvering

V7.5, V4, dump, 3-way, DAR, DDS, CORE, NDMP2D, DDAR, Filters, Int.Ckpts, VBB, Silvering NAS 5.6

### **NEC WEBSAM NetWorker:**

V7.2.1 V4, dump/tar, 3-way, DAR, NAS 5.5 Core, NDMP2D

V7.4.1 V4, dump/tar, 3-way, DAR, DDS, NAS 5.5 all features (DDAR, Chkpnts, VBB, etc)

V7.5—V4; dump/tar; 3-way; DAR; DDS; Core 5.6; NDMP2D; DDAR; Filters; Int.Ckpts; NVB; Tape Silvering

### **ORACLE SECURE BACKUP:**

V10.1.0.2 (aa), v4, dump, 3-way, DAR, DDS—NAS 5.4.28.4 only

V10.1.0.3 (aa,ak) V4, dump, 3-way, DAR, DDS, NAS 5.5 Core & NDMP2D [No 5.6 support]

### **SYMANTEC BACKUP EXEC:**

v11.0.7170 V4, dump, 3-way, DAR, NAS 5.5 core, NDMP2D, DDAR, Int. Checkpoints

v12.0 (an)—V4; dump; 3-way; DAR; Core 5.6; NDMP2D at 5.6.43; DDAR; Int.Ckpts

v12.5 (an)—V4; dump; 3-way; DAR; DDS; Core 5.6; NDMP2D at 5.6.43; DDAR; Int.Ckpts

### **SYNCSORT BACKUP EXPRESS:**

v2.1.4D patch I-1634 for Solaris NDMP v3, 3-way & DAR, NAS 4.2.18.2

v2.1.5D NDMP v3 3-way & DAR, NAS 5.1

v2.1.5D NDMP v3 3-way, DAR, DDS NAS 5.2

v2.2.1 NDMP v3 3-way, DAR, DDS, NAS 5.3

v2.3.1 NDMPv3 dump/tar, 3-way, DAR, DDS, NAS 5.4 only

v2.35.2, v4, dump/tar, 3-way, DAR, DDS—NAS 5.5 only and all features [No 5.6 support]

### **SYMANTEC VERITAS NETBACKUP:**

Version 3.2 JO820443 & SO820446 NDMP v2 dump/tar 3-way

Version 3.4 JO820443; NAS 4.1 only

Version 4.5 MP1 NDMP v3 dump/tar 3-way; NAS 4.2.9.0, 5.0, 5.1

Version 4.5 FP3 DAR, NAS 5.1 – 5.3

Version 5.0 MP1/v5.1 NDMPv4 dump/tar format, 3-way & DAR; NAS 5.1 – 5.4

Version 6.0 v4, dump/tar format, 3-way, DAR, DDS; NAS 5.2 – 5.4, and most features in 5.5

## **Supported Tape Devices/Drives:**

**SCSI Protocol over single Fibre or SCSI Cable HVD Only** [LVD is NOT supported]

### **QUANTUM:**

DLT 4000, 7000, 8000 SCSI or SCSI/FC Bridge for HVD—NAS 5.0 to 5.5

SDLT 220 SCSI or SCSI/FC LVD—NAS 5.0-5.5

SDLT 320 or 600--SCSI or SCSI/FC LVD—NAS 5.1-5.5

LTO-3—supported SCSI & Fibre NAS 5.1-5.6

### **EMC CDL:**

**Note:** See EMC Support Matrix for the latest information on EMC CLARiiON Disk Library solutions. Default emulation is ATL-P3000 with (4) DLT-7000 drives for pre-version 2.x systems. Post 2.x systems have expanded the emulation types.

Latest release CDL 4000 Series Rev 3.0 for DL310, 710, 720, & 740 with Flare 19 patch 040 minimum, NAS 5.4 & 5.5

### **EXABYTE SERIES:**

85xx series & Mammoth-2: SCSI & FCSW/FCAL or SCSI/FC HVD Bridge or Native NAS 4.0 – 5.5

### **HP:**

LTO-1; SCSI 5.1 – 5.4; SCSI/FC with LVD

LTO-2; SCSI 5.1 – 5.5; FCSW/FCAL 5.2 – 5.5; SCSI/FC Bridge with LVD or Native

LTO-3; NAS 5.1-5.5 for SCSI; NAS 5.1-5.5 for FCSW, or SCSI/FC with LVD

SDLT 320; SCSI 5.1-5.5, SCSI/FC LVD

SDLT 600 5.1 – 5.6 Fibre

### **IBM:**

MagStar 3590 SCSI or FCSW/FCAL or SCSI/FC Bridge with HVD or Native

3592-J1A FCSW/FCAL or SCSI/FC Native

Ultrium LTO-1 SCSI 5.0 – 5.5, FCSW/FCAL 5.0-5.5, or SCSI/FC Bridge with HVD or Native

Ultrium LTO-2 SCSI 5.1 – 5.5, FCSW/FCAL 5.1 – 5.5 or SCSI/FC Bridge with LVD or Native

IBM 3588 LTO-3 Ultrium SCSI for NAS 5.1-5.5, FC support for NAS 5.1 – 5.5 LVD or Native

LTO-4, 5.1 – 5.6 Fibre

### **SONY AIT:**

AIT-2 SCSI or SCSI/FC with LVD—NAS 5.0 to 5.5

AIT-3 SCSI NAS 5.1 – 5.5 or SCSI/FC with LVD

S-AIT FCSW/FCAL NAS 5.0-5.6 for FC only

AIT-4, 5.1-5.6, SCSI/FC LVD

### **STK:**

9840A or 9940A all SCSI or FCSW/FCAL NAS codes, or SCSI/FC with HVD or Native

9840B or 9940B FCSW/FCAL or SCSI/FC Bridge Native only

9840C FC only, NAS 5.0-5.6

10000A FC, NAS 5.0-5.6

## **FC-to-SCSI BRIDGES SUPPORT MATRIX:**

Chaparral 2620: NS Series LVD

CrossRoads 4200, 4150, 4250: NS Series HVD

CrossRoads 4450, 6000, 8000, 10000: NS or CNS Series LVD or HVD

HP M2402 NS or CNS Series LVD

HP e1200/e2400 LVD; HVD

Pathlight 3000 or 4000 ADIC NS Series LVD

Pathlight 5000 ADIC NS or CNS Series LVD

Quantum ATL FC 210 NS or CNS Series HVD

Quantum ATL FC 230 NS or CNS Series LVD

Quantum ATL FC 310, 420, 470 NS or CNS Series LVD

Quantum ATL FC 1202 NS & CNS

Spectra Logic G2 F-QIP: NS or CNS Series LVD

StorageTek SN3250, 3300, 3400 LVD & HVD [3250 not tested LVD]

**Purpose :** Convert FC Protocol to SCSI protocol. Sharing of FA ports between DMs & other hosts not supported!

**Note:** Quantum FC470 in arbitrated loop mode should have HARD ALPA port configured with range DC - EF

## **I/O VALUES FOR FIBRE TAPE DRIVES:**

**param streamio timeout=180**

**Note:** Celerra default timeout is 180 secs (3 minutes), 0xb4 hex. Increase this value for IBM LTO2 Fiber Ultrium drives to 30 minutes or 1800 seconds, as their timeout is 17 minutes. Caution, this param is only for fibre tape drives. NAS 5.1.25.0, 5.2.13.0, & 5.3.5.0 contain the new “param streamio” param, see AR57118. NAS 5.5 raises default to 708hex, or 1800 secs. in decimal.

### **NDMP TLU SUPPORT FOR CELERRA:** 3 Ways to Connect to TLU Robot

1. DataMover Fibre Channel or SCSI connectivity: NDMP Client routes commands to TLU via DataMover
2. IP LAN: NDMP Client controls TLU Robot directly over Network via commands
3. NDMP Client Direct Connection: NDMP Client directly connected to TLU [No LAN or DataMover communication]

### **NDMP INCREMENTAL BACKUPS:**

Performs backups based on file C-Time, not Archive bits

### **NDMP PARAMETERS FOR DART:**

**param NDMP bufsz=128**

**param PAX nbuf=8** [x \* 4 = y]

**Note:** **bufsz** is the max buffer size in kb for tape\_read and tape\_write operations from NDMP Client. When reading output of “param NDMP”, current value shows up as 20000 Hex, or 131072 bytes, which is 128KB.

#### **Typical Server Log Error When Exceeded:**

2002-10-12 19:55:36: NDMP: 3: Read count exceeds bufsz

**nbuf** is the preallocated buffer used by Celerra DART for backing up to tape or restoring from tape

**Note:** nbuf values must also be set the same on a 3-Way Server as for the Server directly-attached to tape! Nbuf memory buffers are allocated by DM during bootup, thus, changing values requires reboot to put into effect.

**param NDMP ntape=2** [ x \* 2 = y ]

**param NDMP md5=0** [purpose of this is to encrypt the NDMP passwd]

**param NDMP dialect=** [Default param value for I18N]

**Note:** NAS 5.3 and higher no longer contains the nbuf param--it is no longer applicable. The default ntape value is set to 4, sufficient to handle up to (4) attached tape drives, which is the Celerra limit for concurrent operations.

### **NAS 5.6.40.3:** MD5 Encryption

This NAS version enables MD5 encryption as the default, meaning that current backups and restores will fail unless an NDMP useraccount and password is setup on the Data Mover using server\_user.

**# server\_param server\_2 -facility NDMP -info md5**

server\_2 :

```
name          = md5
facility_name = NDMP
default_value = 1
current_value = 1
configured_value =
user_action    = none
change_effective = immediate
range         = (0,1)
description   = Turns on MD5 authentication for NDMP
```

**# /nas/sbin/server\_user server\_2 -add -md5 -passwd ndmp**

Creating new user ndmp

User ID: 505 →Select a UID

Group ID: 505 →Select a GID

Home directory:

Changing password for user ndmp

New passwd:

Sorry, password must be 6 characters or more

New passwd:

Retype new passwd:

**Note:** Used “ndmpndmp” as the new password since it had to be a minimum of 6 characters

**Adds line to Server ./etc/passwd file with encrypted password entry:**

ndmp:klvaFqcJuuRP.:505:505:HRb7QIpm2kdnpmndnqiIBLwaP0l::ndmp\_md5

### **DATA MOVER AUXILIARY FIBRE CHANNEL PORT SETTINGS NAS 5.6:**

## # server\_param server\_2 -facility fcTach -info linx\_speed\_aux0 -v

```
server_2 :
name      = linx_speed_aux0
facility_name = fcTach
default_value = 0x00008000
current_value = 0x00008000
configured_value =
user_action = reboot DataMover
change_effective = reboot DataMover
range      = (0x00001000,0x00008000)
description = Set the link speed for Fibre Channel tape port AUX0 (default auto-negotiate)
detailed_description
param fcTach linx_speed_aux0=0x8000 sets the tape backup port AUX0 to auto-negotiate the link speed. param fcTach
linx_speed_aux0=0x4000 sets the tape backup port AUX0 link speed to 1Gbps. param fcTach linx_speed_aux0=0x2000 sets the tape
backup port AUX0 link speed to 2Gbps. param fcTach linx_speed_aux0=0x1000 sets the tape backup port AUX0 link speed to
4Gbps. Note: This parameter applies only to the Celerra NS Systems.
```

**NAS 5.6.41:**

## # .server\_config server\_2 -v "param fcTach"

| Name                   | Location   | Current    | Default    |
|------------------------|------------|------------|------------|
| fcTach.linx_speed_aux0 | 0x02647b30 | 0x00008000 | 0x00008000 |
| fcTach.linx_speed_aux1 | 0x02647b34 | 0x00008000 | 0x00008000 |
| fcTach.linx_speed_be0  | 0x02647b28 | 0x00008000 | 0x00008000 |
| fcTach.linx_speed_be1  | 0x02647b2c | 0x00008000 | 0x00008000 |

→Aux ports default to Auto-sense, 0x8000

**AUX PORT LINK SPEED SETTINGS:**

**0x4000=1Gb (decimal 16384)**

**0x2000=2Gb (decimal 8192)**

**0x1000=4Gb (decimal 4096)**

**0x8000=auto (decimal 32768) →Code default since 5.3.15.3 & 5.4.12**

**CONFIGURING AUX FIBRE PORT SETTINGS ON NS DATA MOVERS (NAS 5.1-5.4):**

**Note:** Auxiliary port settings apply to NS Series Celerra--NS500, NS600, NS700, or X-Blade data movers. No aux ports on 510, 514 DM's, etc.

**param fcTach enable\_fabric\_aux0=0** →Indicates FC Arbitrated Loop, default Celerra setting for NAS 5.3 & 5.4

**param fcTach enable\_fabric\_aux0=1** →Set param value to 1 to hard-code support for FC-Switched Fabric

**fcTach.linx\_speed\_aux0 0x00008000** →Indicates auto-negotiate, default Celerra setting for NAS 5.3.15.3 & 5.4.12

**AUXILIARY PORT PARAMETERS FOR NAS 5.4.12:**

## # .server\_config server\_2 -v "param fcTach"

| Name                      | Location   | Current    | Default    |
|---------------------------|------------|------------|------------|
| fcTach.enable_fabric_aux0 | 0x017b98e4 | 0x00000000 | 0x00000000 |
| fcTach.linx_speed_aux0    | 0x017b98d0 | 0x00008000 | 0x00008000 |

**Note:** 5.4 removes references to aux1. Only aux0 link speed and topology can be specified, and defaults to auto-negotiate for speed & auto-sense for topology.

**AUXILIARY PORT PARAMETERS FOR NAS 5.5.79.0 (fabric is auto-sensed):**

| Name                   | Location   | Current    | Default    |
|------------------------|------------|------------|------------|
| fcTach.linx_speed_aux0 | 0x01fd4110 | 0x00008000 | 0x00008000 |
| fcTach.linx_speed_aux1 | 0x01fd4114 | 0x00008000 | 0x00008000 |

**AUX PORT INFO FOR NS500 SYSTEMS :**

The AUX port is an extension of the backend disk loop and can ONLY run at 2GB, not 1GB. AR61224 should address this issue in NAS 5.4. Port can be hard-coded to 2Gb/sec or left to auto-negotiate (0x8000).

**NAS 5.4 FCTACH PARAMS:**

## \$ server\_param server\_2 -f fcTach -i -all

```
server_2 :
name      = linx_speed_aux0
facility_name = fcTach
default_value = 0x00008000
current_value = 0x00008000
```

```

configured_value      =
user_action          = reboot DataMover
change_effective     = reboot DataMover
range                = (0x00002000,0x00008000)
description          = Set the link speed for AUX0 (default 1GBit)
name                 = enable_fabric_aux0
facility_name         = fcTach
default_value         = 0
current_value         = 0
configured_value     =
user_action          = reboot DataMover
change_effective     = reboot DataMover
range                = (0,1)
description          = Enables fabric support on port AUX0

```

**CHANGING PARAM SETTINGS USING SERVER PARAM :**

**\$ server\_param server\_2 -f fcTach -i -all**

```

server_2 :
name           = linx_speed_aux0
facility_name   = fcTach
default_value   = 0x00008000
current_value   = 0x00008000
configured_value =
name           = enable_fabric_aux0
facility_name   = fcTach
default_value   = 0
current_value   = 0
configured_value =

```

**\$ server\_param server\_2 -f fcTach -m enable\_fabric\_aux0 -v 1** →Example of change to FC Switched fabric

server\_2 : done

Warning 4126: server\_2 : You must reboot server\_2 for enable\_fabric\_aux0 changes  
to take effect.

**# server\_param server\_2 -f fcTach -m linx\_speed\_aux0 -v 0x4000** →Example of change to 2GB speed

server\_2 : done

**\$ server\_param server\_2 -f fcTach -i -all**

```

server_2 :
name           = linx_speed_aux0
facility_name   = fcTach
default_value   = 0x00008000
current_value   = 0x00008000
configured_value =
user_action     = reboot DataMover
change_effective = reboot DataMover
range          = (0x00002000,0x00008000)
description    = Set the link speed for AUX0 (default 1GBit)
name           = enable_fabric_aux0
facility_name   = fcTach
default_value   = 0
current_value   = 0
configured_value = 1
user_action     = reboot DataMover
change_effective = reboot DataMover
range          = (0,1)
description    = Enables fabric support on port AUX0

```

**PARAMS FOR I18N SUPPORT:**

**param NDMP convDialect=8859-1** [Default value; Change to convDialect=UTF8 to restore from non-ASCII backups]

**param NDMP dialect=UTF8** [Set this value when DM operates in I18N mode with UTF8 client backups/restores]

**Note:** Apply the above params in situations where UTF8 is being used and conversions are failing, as seen in the logs:

2005-07-03 21:06:36: NDMP: 6: NDMP client dialect is set (clientType=0xc4df4004) although NDMP.dialect is NOT

### **NDMP TUNING PARAMETERS (NAS 5.2 & 5.3+):**

**Example:** DM using two tape drives and buffer size of 128--recommended tuning settings

**param PAX paxStatBuff=1024** [Buffers between threads that send & receive metadata and Read data from disk]

**Defaults:** 128 for 507DM & 1024 for 510DM; Values range from 1-2048

**param PAX paxWriteBuff=640** [Buffers between threads that read file data and write data to tape]

**Defaults:** 64 for 507DM & 640 for 510DM; Values range from 1-2048

**param PAX nPrefetch=32** [Number blocks each backup thread prefetches from disk before finished file read and write to tape]

**Defaults:** 8 for 507DM & 32 for 510DM; Values range from 1-128

**param PAX nThread=64** [Number of backup threads]

**Defaults:** 16 threads for 507DM & 64 for 510; Values range from 1-256

**param PAX paxReadBuff** [Buffers between threads that read data from tape and restore to disk]

**Defaults:** 4 for 507DM & 16 for 510DM; Values ranges from 1-64

**Note:** Some PAX default values have been lowered from earlier NAS versions due to possibility to panic for out of memory

### **DETERMINING VALUES FOR 5.4 PARAMS USING SERVER PARAM:**

**\$ server\_param server\_2 -facility NDMP -info dialect -v** (with description)

server\_2 :

|                  |                            |
|------------------|----------------------------|
| name             | = dialect                  |
| facility_name    | = NDMP                     |
| default_value    | = "                        |
| current_value    | = "                        |
| configured_value | =                          |
| user_action      | = none                     |
| change_effective | = immediate                |
| range            | = '*'                      |
| description      | = Set the dialect for NDMP |

### **NAS 5.4 Values:**

|                          |            |            |            |                                                     |
|--------------------------|------------|------------|------------|-----------------------------------------------------|
| NDMP.bufsz               | 0x018f8d24 | 0x00000080 | 0x00000080 |                                                     |
| NDMP.convDialect         | 0x018f8ec4 | '8859-1'   | '8859-1'   | →Use when restoring ascii-mode tape to I18N DM      |
| NDMP.dialect             | 0x018f8ec0 | " "        | " "        | →Default “ “(null), or UTF8, use to match NDMP host |
| NDMP.maxProtocolVersion  | 0x00f06474 | 0x00000004 | 0x00000004 |                                                     |
| NDMP.md5                 | 0x018f8ec8 | 0x00000000 | 0x00000000 |                                                     |
| NDMP.scsiReserve         | 0x00f06478 | 0x00000001 | 0x00000001 |                                                     |
| NDMP.v4OldTapeCompatible | 0x00f0647c | 0x00000001 | 0x00000001 |                                                     |

### **SETTING NDMP VERSION:**

**param NDMP maxProtocolVersion=3** [Default version=4]

### **SETTING UP NDMP BACKUPS WITH LEGATO NETWORKER 7.2 WITH VTLUs:**

1. Create NDMP account

#/nas/sbin/server\_user server\_x –add –md5 –passwd ndmp

2. WebUI, create VTLU server using default values

3. Create Storage for VTLU: Rightclick new VTLU>properties>Storage Tab>New>Select VTLU file system created previously, Select number of tapes, use \*Slot for tape location, and LGTO for Barcode Prefix

4. VTLU properties>Drives: enter device names in Backup application

5. Create Auto-changer from Legato CLI:

C:\Program Files\Legato\lsrc\bin>jbconfig (Configure an Autodetected NDMP SCSI Jukebox)

**Note:** Enter Jukebox type 1, Tape Server IP address, User ndmp & password, Jukebox device name, NetWorker auto-cleaning, drive intended for NDMP--yes

6. Label Tapes Using Legato CLI: >nsrcjb –L –S <tape slot #>

7. Enter IP Address and Server name of DM in Windows>System32>drivers>etc>hosts file

8. Launch NetWorker: Programs>Legato NetWorker>NetWorker Administrator>NetWorker Groups>Manage Groups>rightclick groups Create>assign name

9. NetWorker Servers>rightclick on Windows system name>connect to this Server>Client Operations>Manage Clients>rightclick, Create>General Tab: Enter Server name, Save Set path to file system, Group name. Preferences Tab: Server name as alias, Storage

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
nodes, and Clone storage nodes, NDMP- Yes; Remote Tab: Remote access: “@”; Remote user: ndmp; Password: ndmpndmp; Backup command: nsrndump\_save –T tar; Application information: HIST=y, UPDATE=y, DIRECT=y

10. Create c:>bootstrap folder on Legato Windows Server
11. Create Boostrap from Legato GUI: Media Management>Devices>rightclick Create: Name: c:\bootstrap; Media type: file;
12. Rightclick Bootstrap device>Operations>Select Label
13. Perform manual backup: Networker Groups>Manage Groups>rightclick Group>Start  
→Click on Monitor tab to monitor progress  
→Rightclick Group>Details to see details of backup  
→Check Server Log

### **NDMP COMPATIBILITY ISSUE V3 vs. V4 ISSUE:**

Omniback 4.2 and Data Protector 5.0 have incompatibility between NDMP versions 3 & 4 in which restores of all appended backups will fail due to a tape positioning issue. Restores may complain about filemark problems:

“server\_archive: Cannot identify format. Searching...

ATTENTION! server\_archive archive volume change required...”

#### **NAS 4.2.18 + & 5.1.15 + Fix:**

1. Logon to CS
2. Add following to server param file: **param NDMP v4OldTapeCompatible=0**
3. Reboot

**Note:** Setting this param should allow restores of NDMP v3 Backups to work

#### **NAS 4.2.17.2 or 5.1.9.4 and Lower:**

1. Set following param in the “omnirc” file on the Omniback Client Host: **OB2NDMPFMV4=1**

**Note:** See NAS 5.1.15.3 Release Notes for more details. Also note that if you upgrade from 4.2.17.2/5.1.9.4, then this fix needs to be disabled and the previous fix applied

**TROUBLESHOOTING NDMP:** NDMP Messages categorized as SCSI, CONFIG, TAPE messages

### **INCREASING CELERRA NDMP LOGGING LEVELS FOR TROUBLESHOOTING:**

#### **TURNING NDMP LOGGING ON:**

**\$server\_config server\_2 “logsys set severity NDMP=LOG\_DBG2”**

**\$server\_config server\_2 “logsys set severity PAX=LOG\_DBG2”**

#### **TURNING NDMP LOGGING OFF:**

**\$server\_config server\_2 “logsys set severity NDMP=LOG\_PRINTF”**

**\$server\_config server\_2 “logsys set severity PAX=LOG\_PRINTF”**

**Note:** If implementing any 'DEBUG' levels of NDMP logging, it is highly recommended that you setup the following command to pipe the Server Log into a file so that all debug information is capture. Otherwise the Server Log wraps too quickly.

**\$nohup server\_log server\_3 -f -s > s3log.log & [\$jobs #fg %1 ctrl + c]**

**Server Log:** RPC, NDMP, SCSI/FC Errors [RPC=backup utility errors; NDMP=protocol errors; SCSI/FC=device errors]

**BackUp Software Logs:** [Depends on Platform being used for BackUps]

**Note:** NDMP only supports a single data stream!

NDMP does not require specific "client" software on DataMover--uses NDMP protocol

**\$server\_config server\_2 “printstats pax”**

**Note:** ‘getpool’ means waiting on backend while ‘putpool’ is waiting on tape drive

### **CREATING AN NDMP SERVER LOG CAPTURE SCRIPT:**

**1. # mkdir /nas/var/emc/ndmpleg**

**2. # vi ndmp\_capture.log and modify script to collect log from correct DM, chmod 755 to make executable:**

===== script template used to capture the server log. Pls modify as need=====

**#!/bin/sh**

**DM=server\_2**

**#**

**NAS\_DB=/nas**

**export NAS\_DB**

**STIME=`date '+%y%m%d%H%M'**

**\$NAS\_DB/bin/server\_log \$DM -s -a > /nas/var/emc/ndmpleg/log.\$DM.\$STIME**

**gzip /nas/var/emc/ndmpleg/log.\$DM.\$STIME**

**3. Create cronjob to run log capture script at 5 minute intervals:**

**#crontab -e**

**0-59/5 \* \* \* \* /nas/var/emc/ndmpleg/ndmp\_capture.log**

## **NDMP BACKUP LEVELS: LEVEL 0 and LEVELS 1-9**

### **LEGATO:**

Level 0 is a full backup of all files and directories into a “SaveSet” and is considered the lowest NDMP Backup level  
 Levels 1 – 9 are considered ‘differential’ and only backs up files that have changed since last lower level backup  
 i.e., Level 0 is taken, Level 1 will then backup all files that have changed since Level 0. Level 2 will backup files that have changed since Level 1, and so on.

## **NDMP INCREMENTAL BACKUPS/RESTORES:**

Supported Veritas 3.2/3.4 & Legato NetWorker 6.1 using BackUp Level 10

1. Verify NDMP Daemon is running on the Server: `$.server_config server_x -v "ndmp"`
2. Verify NDMP passwd account has been setup correctly: Test by FTP'ing to the DataMover [Resolved by NIS or local Passwd file]
3. Verify Devices seen by the DataMover [Tape & Jukebox]: Record device numbers for future reference  
`--server_devconfig server_x -p -s -n` [Probe verifies the SCSI devices on the bus]  
`--server_devconfig server_x -c -s -n`  
`--server_devconfig server_x -l -s -n`
4. Review NDMP Server Logs from either the NT or SUN Solaris Backup Server [Legato, Veritas, etc]

### **5. Review/Change Logging Levels on Celerra Server:**

`$server_config server_2 "logsys set severity NDMP=LOG_DEBUG"`

`$server_config server_2 "logsys set severity NDMP=LOG_DBG2"`

`$server_config server_2 "logsys set severity PAX=LOG_DBG2"`

[Higher logging level--use this]

[New logging parameter]

**Verify Logging On:** `$server_log server_2 -s -f ltail`

6. Conduct Testing/Operations to capture Verbose Server Logs

7. Turn off Celerra Server Verbose Logging:

`$server_config server_2 "logsys set severity NDMP=LOG_PRINTF"`

`$server_config server_2 "logsys set severity PAX=LOG_PRINTF"`

8. Verify NDMP & PAX Param Settings:

**Note:** Necessary to avoid Memory Fragmentation for NDMP/PAX Backup/Restores on DM's with Tape Drives Attached:

/nas/server/slot\_x/param

`param NDMP bufsz=128` [default: o.k. for most clients except Legato NetWorker!]

`param PAX nbuf=8 <number_for_pax>` [(2) tape drives \* 4 = 8]

`param NDMP ntape=2 <actual number of tape drives in use>`

**Warning:** *The above "params" are for NAS Code 2.2.46.0 and 3.1.3.0 and higher with (2) tape drives attached!*

**Note:** NAS 5.3 and higher no longer contains the nbuf param--it is no longer applicable. The default ntape value is set to 4, sufficient to handle up to (4) attached tape drives, which is the Celerra limit for concurrent operations.

## **NDMP/PAX PARAMS ON DATAMOVER:**

`# .server_config titanic -v "param NDMP"`

| <i>Name</i>   | <i>Location</i> | <i>Value</i>                                                            |
|---------------|-----------------|-------------------------------------------------------------------------|
| NDMP.dialect  |                 | 0x00b061f8 'latin1'                                                     |
| NDMP.ntdump   |                 | 0x00b061f0 0x00000001 / 1                                               |
| NDMP.md5      |                 | 0x0afe274 0x00000000 / 0                                                |
| NDMP.bufsz    |                 | 0x0afbe44 0x00020000 / 131072 [Actual value is 131,072 bytes, or 128kb] |
| NDMP.ntape    |                 | 0x01048a20 0x00000000 / 0                                               |
| NDMP.prefixNT |                 | 0x00afbcc8 '/.NT.'                                                      |

`# .server_config titanic -v "param PAX"`

| <i>Name</i> | <i>Location</i> | <i>Value</i>              |
|-------------|-----------------|---------------------------|
| PAX.nbuf    |                 | 0x01048558 0x00000000 / 0 |

## **DEFAULT PAX PARAMS NAS 5.2.x & 5.3:**

**param PAX paxStatBuff=1024**

**Note:** paxStatBuff specifies number of buffers between threads that send and receive metadata (NASA) and those that read data (NASS). Increasing this param value increases speed of metadata work. Values are 1-1024. Old default for 507 DM 128. 510 DM default is 1024 with latest 5.2 and new 5.3 code.

**param PAX paxWriteBuff=640**

**Note:** Number buffers between threads that read data (NASA) and write data to tape (NASW). Tape streaming can be enhanced with higher values. 507 Value=64. New 510 values=640.

**param PAX nPrefetch=32**

Note: Number blocks backup thread prefetches from disk before completing file read and write to tape. 507 value=8 510 value=32

**param PAX nThread=64**

Note: Number of backup threads. 507 value=16 510 value=64

**\$ server\_param server\_2 -f PAX -l**

server\_2 :

| name         | facility | default | current | configured |
|--------------|----------|---------|---------|------------|
| paxWriteBuff | PAX      | 64      | 64      |            |
| paxStatBuff  | PAX      | 128     | 128     |            |
| nFTSThreads  | PAX      | 8       | 8       |            |
| nThread      | PAX      | 64      | 64      |            |
| nPrefetch    | PAX      | 8       | 8       |            |
| nRestore     | PAX      | 16      | 16      |            |

Note: The higher default settings for PAX led to memory panics as buffers were gobbling up Server memory. New PAX defaults are reflected by NAS 5.3.12 output above

**# server\_pax server\_2 -s -v**

server\_2 :

\*\*\*\*\* SUMMARY PAX STATS \*\*\*\*\*

---- NASS STATS ----

\*\* nass thid 0 \*\*

Total file processed: 81

throughput: 0 files/sec

Total nass wait nasa count: 78

Total nass wait nasa time: 142791 msec

Total time since last reset: 1107 sec

fts\_build time: 1 sec

getstatpool: 0 buffers putstatpool: 126 buffers

nass01 is not doing backup

nass02 is not doing backup

nass03 is not doing backup

---- NASA STATS ----

\*\* nasa thid 0 is running backup with dump format \*\*

Backup root directory: /PST

Total bytes processed: 26829497451

Total file processed: 81

throughput: 23 MB/sec

average file size: 323465KB

Total nasa wait nass count: 4

Total nasa wait nass time: 157 msec

Total time since last reset: 1107 sec

Tape device name: c9000t011

dir or 0 size file processed: 3

1 -- 8KB size file processed: 0

8KB+1 -- 16KB size file processed: 0

16KB+1 -- 32KB size file processed: 0

32KB+1 -- 64KB size file processed: 0

64KB+1 -- 1MB size file processed: 2

1MB+1 -- 32MB size file processed: 7

32MB+1 -- 1GB size file processed: 64

1G more size file processed: 5

fs /PST size is: 48563675136 Bytes

Estimated time remain is 901 sec

nasa01 is not doing backup/restore

nasa02 is not doing backup/restore

nasa03 is not doing backup/restore

---- NASW STATS ----

nasw00 BACKUP (in progress)

Session Total Time: 00:18:27 (h:min:sec)

Session Idle Time: 00:11:30 (h:min:sec)

KB Tranferred: 26200680 Block Size: 61440 (60 KB)

Average Transfer Rate: 23 MB/Sec 81 GB/Hour

Average Burst Transfer: 61 MB/Sec 215 GB/Hour

\_Point-in-Time\_ (over the last 10 seconds):

Rate=23 MB/Sec Burst=93 MB/Sec Idle=780 msec/sec

Get Pool: 0 buffers Put Pool: 63 buffers

nasw01 No session found

nasw02 No session found

nasw03 No session found

**\$ .server\_config server\_2 -v "printstats pax full"**

### **5.3 ENHANCEMENTS:**

--multithreading restore process  
--multithreaded directory traversal, allows for better performance with small files and fragmented file systems  
--improves large file handling by allowing multiple NASA threads to work on the single file  
--NASA reads data and writes to file, uses larger data buffer pool  
--More threads and buffers will help on small file restores

**Note:** NAS 5.3 and higher no longer contains the nbuf param--it is no longer applicable. The default ntape value is set to 4, sufficient to handle up to (4) attached tape drives, which is the Celerra limit for concurrent operations.

NASA → Writes file header, reads file data, writes to buffers (paxStatBuff)

NASS → Traverses file system, provides metadata for each directory/file

NASW → Gets data from buffer pool (paxWriteBuff), writes them to tape, sends back to remote mover

**param PAX nRestore=16** [Default value for number of restore threads, with range from 1 – 256]

**param PAX paxReadBuff=64** [Default value for number of tape read-ahead buffers, with range from 1-512]

**param PAX nFTSThreads=8** [Default value for number of FTS threads, ranges from 1-256]

**\$server\_pax server\_x -stats -verbose** [PAX Statistics]

NASS Stats→stats on file system traversal, provides meta data for each directory/file

NASA Stats→reads file data and writes to buffer with NASW

NASW Stats→gets data from buffer pool with NASA and writes to tape or sends to remote mover

**\$server\_pax server\_x -stats -reset**

### **SETTING I18N PARAMETERS FOR NDMP:**

**Note:** Set following parameter when Data Mover has I18N enabled, for all NAS versions that support I18N. This will ensure that the DM will use UTF-8 translation instead of the default “Latin 1”

**param NDMP dialect=**

**param NDMP convDialect=8859-1** [Default value; Change to convDialect=UTF8 to restore from non-ASCII backups]

**param NDMP dialect=UTF8** [Set this value when DM operates in I18N mode with UTF8 client backups/restores]

### **INVALID UTF CHARACTERS MAY RESULT IN BACKUP/RESTORE FAILURE (NAS 4.2):**

#### **Legato NetWorker NDMP Service Log:**

server\_archive: Can't convert filename using the specified NDMP.dialect, or invalid utf8 filename:  
0x2f41302f6d6f6e746167652f68647574696c7334312f737461362f3fea3456789abcd4d45432e75736572.

#### **Resolution:**

**param NDMP.dialect=null**

**Note:** New default codepage with NAS 5.2 is null to address I18N issues.

### **SERVER LOG ERRORS REGARDING INVALID UTF8 NAMES DURING NDMP BACKUPS:**

2005-05-05 10:34:55: NDMP: 3: Can't convert filename using the specified NDMP.dialect, or invalid utf8 filename:  
0x2f766f6c302f706c69752f434354656d706c6174652f49432f496d61676520436166822063616e63656c6c6174696f6e5f72657175657  
3745f72656365697665645f6b6565705f646f6d61696e2e646f63.

2005-05-05 10:34:55: NDMP: 3: inode\_2137907 is using for filename

**Note:** Point of above example is that Celerra NDMP backups no longer fail when running into UTF8 names that cannot be translated. We log the error in the log, then instead of passing the file name history back to the NDMP Client, we pass the inode number of the file/folder.

### **OTHER NDMP PARAM VARIATIONS:**

**param NDMP dialect=<dialect string>**

**Note:** Dialect String is the dialect used by the NDMP Client. Celerra default without this param change is Latin1.

Restores From DM With ASCII Enabled:

**param NDMP convDialect=<dialect string>**

**Note:** Dialect string is the dialect used by the ASCII Mode DM. Default value is Latin1.

**param NDMP convDialect=null**

**Note:** ‘Dialect=null’ actually defaults to “cp437.txt” code page

### **RESTORING ASCII BACKUP TO UNICODE: NAS 5.0.11.4**

**param NDMP convDialect=null** [Restores in UTF-8 format]

**param NDMP convDialect=big5.txt** [Restores in alternate code page format]

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
**Note:** By default, Servers restore using Latin-1 (8859-1.txt). Changing to Dialect=null will stop translations from occurring.

### **I18N ISSUE WITH NDMP DIALECT SETTINGS:**

NDMP Service Log: server\_archive: Can't convert filename by the specified NDMP.dialect

NDMP Service Log: server\_archive: Backup continues with environment variable HIST disabled

**Comment:** Customer may see these errors if there is an NDMP dialect error or mismatch in the locale translations used by a Client and that configured for NDMP. In other words, a filename may contain a character that the default NDMP.dialect setting cannot translate properly.

### **NDMP BACKUPS WITH I18N USING TAR FORMAT:**

If files are found that cannot be translated, catalog information is not created, error logged, but backup continues. Mis-translated files can only be restored via Full Restore of a directory.

### **NDMP BACKUPS WITH I18N USING DUMP FORMAT:**

If files are found that cannot be translated, catalog information is not created for any files after that point, error logged, but backup continues. All data from point of first mis-translated file can only be Restored via full restore of directory.

### **INCREASE NDMP LOGGING IN SERVER LOG:**

**\$server\_config server\_x "logsys set severity NDMP=LOG\_DBG2"**

NDMP:5: HALT reason:NDMP\_DATA\_HALT\_SUCCESSFUL; <TAPE\_CLOSE>

--Keep NDMP Tape pools separate from Open Host formats!

--Must conduct restores to same host for NDMP--but can change the directory destination

--DataMovers support (4) datastreams [four separate tape drives using dual scsi cards--two drives handled per each SCSI channel].

But NDMP does not support multiple data stream backups yet!!!!

--"Cannot communicate with Server" --might mean that wrong NDMP passwd was used

**NDMP & I18N SUPPORT:** You cannot restore from tape from a system that was backed up as non-Unicode to a new file system created using Unicode [i.e., I18N is now turned on]--it will fail with this message in Server Log:

2002-06-01 12:34:11: NDMP: 3: Pattern Not Matched: /chainlink/ismp004 (not restored)

**Note:** New default codepage with NAS 5.2 is null to address I18N issues.

### **PATTERN NOT MATCHED:**

2003-05-28 15:39:13: NDMP: 3: Pattern Not Matched: /fs-dm2-data01/data01/Dept/Minimed/Public/ERP\_Imp/SECURITY PROFILES (not restored)

**Comment:** Generally see this error if the wrong Restore path was typed in or 'case' was not exact. Also see this error in 2.2 code for Veritas if a linked null file was backed up—the problem is that restore fails for a linked file that is null. Other probable causes for the "Pattern Not Matched" error are incorrect Restore path and 'case' sensitivity issues with the path defined for the Restore.

**DAR RESTORE ISSUE:** Also can lead to 'Pattern not matched' entry in 5.2. Single file DAR restores o.k. Memory leak and panic.

### **NDMP BACKUP or RESTORE ISSUE:**

NAS < 4.2.10.0 & 5.0.12.0

**Server Log Message: NDMP: 3: TAPE\_OPEN c0t110 fails with stat 16**

Celerra Bug that causes Tape Hang/Busy condition in Veritas and Legato

#### **Status 16 Error Defined:**

"stat 16" error returns are an EBUSY response which means another application has the device open

#### **Cause of Issue:**

The issue is a resource deadlock between two threads during an incremental backup which spans tape volumes and there are no more tapes available. PAX times out and begins to clean up resources. Buffer threads get into a deadlock situation.

#### **Workaround/Resolution:**

Reboot or failover/failback the data mover and make the new media available.

Issue resolved with NAS patch 4.2.20.0 & 5.1.17.0

#### **FULL RESTORES FAILED:**

NAS 5.2 full restores failed—fixed NAS 5.2.11 and higher

#### **NDMP RESTORE COMPLETES BUT NOT ALL FILES RESTORED:**

**Note:** For NDMP restores that encounter name collisions, the default policy is to NOT restore a file with a name collision

**MIGNORE:** Default, do not restore file when name collision exists [**OPTION=MI**]

**MDELETE:** Restore file upon name collision and delete existing file on target [**OPTION=MD**]

**MMANGLE:** When DOS names the same, will rename the restored file but leave original file alone [**OPTION=MM**]

## **DLm4080/DLm4020 (EMC Disk Library for mainframe) Version 1.2**

- Enterprise class tape-on-disk appliance solution for IBM mainframes, using virtual tape emulation (disk storage arrays replace tape libraries)
- A benefit of DLm is that it offloads the Tape Library function and processing from the Mainframe
- DLm connects to Mainframe using Escon or Ficon connections
- Uses SATA drives and RAID 6 protection (12+2)
- DLm4020 is a single rack system based on NS20 Integrated, with up to 47.5TB capacity (100TB compressed)
- DLm4080 can be from 1-7 racks, based on NS80 Integrated, with up to 570TB capacity using IDRC=Force option (1.5PB compressed)
- Can replicate between either solution
- Both NAS platforms ship with dual Control Stations and extra switch to connect to, but is in “cold standby”
- Reduce unknown CallHomes by configuring ConnectHome with following:  
**# /nas/sbin/nas\_connecthome –modify –site\_id “DLm4020”**

### **Components:**

- DLm Controller, or VTEC (Virtual Tape Emulation Controller) running version 2.0 (ACP hardware—DLm2-ACP)
- NAS front-end
- Storage back-end

### **DLm MODELS:**

- DLm4080 Version 1.0.0 NAS 5.5.33.2 Flare 26
- DLm4080 Version 1.1.0 NAS 5.6.37.6 & 5.6.38.2 Flare 26
- DLm4080 Version 1.2.0 NAS 5.6.40.3 Flare 26
- DLm4020 Version 1.2.0 NAS 5.6.40.3 Flare 26

### **DLm SUPPORT:**

MFSW supports VTE & Mainframe portion of the product, NAS the NS80/NS20

### **VTEC (Virtual Tape Emulation Controller) Components:**

- VTEC sits between IBM mainframe and EMC Storage array, providing IBM 3494 virtual tape library emulation for IBM 3480/3490/3590 tape drives, and connects using ESCON, FICON SM, or FICON MM
- Access Control Point (ACP) [provides configuration, management, & maintenance of VTE]
- Gigabit Ethernet switches [use two switches for path redundancy]
- LVD SCSI to HVD SCSI converter
- Virtual Tape Engines (VTE) emulate Tapes [IBM tape drive emulation, directs data to and from Celerra and its attached array]
- NS80 Integrateds or NS20 Integrated (NAS Support)

### **DLm120 Version 2.0:**

- Replaces NS20 configuration with NS-120
- Ships with (2) ACPS, (1-2) VTEs, FICON or ESCON Channel adapters, (2) GbE Ethernet switches, Modem, (2) NS-120 Blades with dual Control Stations (one CS is in ‘cold standby’, not cabled or powered on), CX4-120 array with dual SPS and (up to 5) DAEs, all SATA
- Support up to 47.5TB

### **DLm960 Version 2.0:**

- Replaces NS80 high-end with NS-960
- Ships with (2) ACPS, (1-6) VTEs, (2) 1GbE Ethernet & (2) 10GbE Ethernet switch, (2-6) NS-960 Blades, dual Control Stations cabled and ready, CX4-960 array, with up to (12) DAEs, all SATA except boot DAE, which uses (6) FC drives for vault
- Can ship with (2) NS-960's?
- NS-960 Blade configuration will consist of (1) Tomahawk FC card in slot\_0, Firestorm cards in slot\_1 & slot\_2, and Thunderbolt cards in slot\_4 and Annex slot\_5 [DLm960-DM-8C3 blade type]
- AUX configuration consists of (4) FC IO Modules
- Supports up to 1.2PB

## **DATA MIGRATION METHODS:**

EMCopy for NT; CPIO for Unix [cp -r; cp -pr; cpio; tar]

**Data Migration Considerations:** Using CPIO tools

Incremental if “live” systems v. static copy

Higher GID's than allowed would be a big problem

Having Negative UID/GID's a problem

Symbolic Links that go Across or Up are a problem

Deleted Files on Source are a problem—would be migrated over to target

Negative TimeStamps

Multiprotocol Environments [Permissions & ACL's; even tape Backups don't do it all]

Owners—Permissions—ACL's—LocalGroups—Share ACL's

CIFS—Robocopy/Scopy

**Possible Solution:** Use Lgdup, Sharedup, and EMCopy to migrate a File System.

## **PS MIGRATION TOOLS:**

CDMS, Rsync, EMCopy, Scopy, Robocopy, Secure Copy, Aelita, FastLane, NetIQ

Scopy--Migrate files between DC's; does not migrate "localgroups"

Robocopy--Migrates data & permissions; does not migrate "localgroups"

Secure Copy--GUI-based migration of data, shares, perms, ownership, and localgroups

Rsync--Useful for migration of NFS data from SOURCE to TARGET IP4700 using "rsh" & "ssh" communications. Rsync 2.5.6 is a fast incremental file transfer utility. Note that if copying sparse files, must use the -S flag or else the migration will convert files to dense.

Aelita: Consolidation Wizard 5.63--comprehensive migration tool

XfrPerms: Windows tool backing up SD of files/directories and can reapply to same or migrated directory tree

XfrShares: Windows tool backing up shares and perms. Do not need source up to apply to target system

rmtshare: MS tool to remotely manage Shares from command prompt [ \\server\\sharename=drive:path]

MapSID: tool to find and manage orphaned SIDs

## **QUEST FASTLANE CONSOLIDATOR:** Version 4.1

Win2k SP3, IE 5.5, PIII 400Mhz, 512MB RAM; 350MG space

Migration tool that can preserve cross-domain file, folder, share, & local groups; Maintains file ownership, last access date, W2K inheritance, File filters, migration reports, User Profiles, Storage analysis prior to migration

Best Practices for using Consolidator with Celerra: [www.quest.com/fastlane/consolidator](http://www.quest.com/fastlane/consolidator)

## **MIGRATING TO CELERRA USING ROBOCOPY XP010 WINDOWS 2003 VERSION:**

1. Use LGDUP to migrate local groups from Source to Target
2. Run initial migration: c:>`robocopy.exe \\source \\destination /NP /V /B /MIR /FFT /COPY:DAT /LOG:logname`
3. Repeat step 2 for incremental copies. Stop access to Source and run a final time to complete data migration.
4. Run following to copy security, ownership, and auditing information:  
c:>`robocopy.exe \\source \\destination /NP /V /B /MIR /IS /FFT /COPY:SOU /LOG :logname`

**Note :** User account for migration should have Backup Operator privileges

## **HOW TO COPY FILESYSTEM VIA NFS STREAM AND THEN COPY OVER NT ATTRIBUTES:**

Step 1. From Remote Unix Host, mount to Celerra NFS Exports:

FS\_A = Source filesystem

FS\_B = Destination filesystem

Step 2. From UNIX host, cd to the mounted directory of FS\_A & run following command:

**`tar cf - . | (cd <<full path of directory mounted of FS_B>>; tar xfBp -)`**

[this will stream the data from FS\_A to filesystem FS\_B & retain all UNIX rwx and sticky bit permissions]

Step 3. When NFS Stream operation is complete, conduct following from an NT Server:

--Map the two filesystems as CIFS shares from the NT server

--Run following from NT Server command prompt:

**`c:>emcopy <<FS_A dir full path>> <<FS_B dir full path>> *.* /s /sefix`**

[this will copy the NT permissions from the directory on FS\_A to the same named files on FS\_B]

## **ROBOCOPY NTRESKIT TOOL:** Robust File Copy for Windows NT v.1.96

**Intro:** Robocopy is an NT Resource Kit tool similar to "xcopy" or "scopy", only better. For basic copying of data from an NT Server to the Celerra, this requires very little preparation. The only caveat is that without running another tool such as EMCopy's "LGDUP", "robocopy" does not copy "ownership" information over for Celerra:

**`C:>robocopy.exe \\win2kbde\ robocopy \\celerra\migrate /E /SEC /SECFIX >robo.txt`**

### **Robocopy Switches:**

/mir [combination of /E to copy directories and /Purge to purge files on Source after copying

/SEC [copy security information—requires volumes be NTFS]

/SECFIX [fixes security information on files & directories]

/mir [mirrors Data & NTFS attributes from Source to Target]

/S : copy Subdirectories, but not empty ones

/E : copy subdirectories, including Empty ones

/NP : No Progress - don't display % copied

/XO : eXclude | Older files

/R:n : number of Retries on failed copies: default is 1

/W:n : Wait time between retries: default is 30 seconds

/Z : copy files in restartable mode

**ROBOCOPY XP010:** Windows 2003 Resource Kit

## **CELERRA CDMS VERSION 2 (CDMSV2): CIFS & NFS**

**Note:** CDMS v1 was first version and introduced with NFS support only in NAS 2.2

Introduces CIFS support using 4.2.10+, 5.0.11+, 507DM+, running on NT or WIN2K Platforms; Adds NFSv2 support (keeps NFSv3 support) with both UDP & TCP support; Map source file system root to a 'directory'; Intended to fully support NT/Windows 2000 Domains; Does not require port mapper to support migration; Uses standard RPC client generated by "drpcgen"; Can now consolidate multiple source file system types into one local file system; Migrates all CIFS file attributes

**Note: NAS 5.1M introduces support that allows Customers to conduct their own CDMS migrations**

### **CDMS MIGRATION EVALUATION TOOLS FOR CIFS/NFS:**

dircount.pl /mntpoint → Script to verify tree structure depth and number of dirs & files

diskUsage.pl -m → Script to validate amount of space needed on target for the migration

### **CDMS CONNECTION THREADS:**

→NAS 5.2 only supports a single connection thread for the migration process

→NAS 5.3 supports multiple connection threads to different connection paths

**Note:** Important to know that even when errors are seen in the migration process, the MGFS Threads do not stop, even if the WebUI says that they have failed. This is one of the main reasons for not kicking off a verify until you know that the MGFS thread is completed.

### **CDMS MIGRATION PROCESS:**

CIFS CDMS migration does not impact Users and is done synchronously for directories [entire directory migrated & offline inodes created] and asynchronously for files [data is pulled over per Client requests on reads—when whole file is read, the offline inode is brought online]. A key point of our target “mgfs” filesystem on the Celerra is that we have added a new inode type [offline inode] to support this type of migration. Another important point is that inodes are read block-by-block, not individually. Cross-domain CDFS migrations are possible but are very complicated and should engage EMC PS support.

### **HOW DO MIGRATION DIRECTORIES AND FILES GO ONLINE?**

When running the CDMS migration from Source to Target, the root directory on the Target is the first to be brought “online”, then the parent directories, but each directory is done linearly across and recursively down the entire subdirectory structure before moving to the next parent-level directory. In other words, the directory itself is marked online immediately, though until the process completes, many of the sub-directories are not yet online. Files themselves are migrated piecemeal as Clients request to read various portions of the file—if the last byte of the file is read by a Client, then a 2<sup>nd</sup> CDMS thread ensures that any other parts of the file are read, then migrates and marks the file header as “online”. Files might not be properly migrated if they contain incompatible special characters or linkages that the CDMS process cannot resolve—this is why the use of the dirprivW.pl tool is important and is run on the Source file system prior to migration—files should be identified and corrected prior to the migration. The term “offline inodes” refers to the state of directories and files that are in the migration process.

### **POST-MIGRATION RULES & THINGS TO CHECK PRIOR TO CONVERTING FROM MGFS:**

Rule 1: Do not convert from mgfs file system before reviewing AND correcting any CDMS migration log errors!

Rule 2: Do not run the verify command before reviewing AND correcting any CDMS migration log errors!

**Note:** If you run the verify command and I/O is still occurring for the migration process, the verify will stop and restart from the beginning of the file system—essentially, you will bog down the Server CPU and never get anywhere. For blocks that were successfully checked by the verify, if the process has to start from the beginning again, it will cover the checked blocks very quickly—however, it DOES have to start from the beginning each time.

Rule 3: Check Server Log for evidence of offline inodes and other errors

### **WHAT TO DO AFTER A MIGRATION:**

#### **I. ALWAYS CHECK & RESOLVE ERRORS FROM MIGERR LOG:**

→Temp files, such as .xls, .tmp, may be recorded in migErr log, but these are normal and will have been migrated by the User, even if the migration thread did not read

→Certain system files & process can lock files, preventing migration—in many cases this will be normal and they can be deleted on the Celerra, but customer should make these decisions

→Other files or directories may be listed in migErr log, but when checking for the files on the Celerra, if the files are no longer present, it generally means that the files have been deleted since the migration thread was run—as in reading a file, the act of deleting files also puts the inode “online”, so this is normal and expected & no further action required

→Remaining files or directories in migErr log are seen on the Celerra—try to read the contents of the files to see if the files are good—if so and there are many such files, run the migration thread again

→If there are files that cannot be read on the Celerra (and are referenced as errors in the migErr log), it’s possible that the source files are corrupt, or there could be access problems on the source, or the ‘Backup Operator’ rights were not properly assigned to the migration user account

→As a last step, try reading files from the Source server—if they cannot be read then they can be safely deleted. If they can be read, then try migrating via another method, such as EMCopy or Robocopy

→Network problems can cause files to be dropped or skipped during the migration [process can only ‘pause’ for a finite time before continuing], resulting in offline inodes

→Only run the ‘getOnline’ procedure as a last resort and only after all the above measures have been taken, as the ‘getOnline’ procedure does not ‘fix’ problem inodes. For example, if there is a directory with only partially migrated files, the rest of the files will not be migrated and will be missing. Make sure that all ‘offline’ inodes are understood before running this:

**#/nas/tools/cdms/specialCmd getOnline server\_2 /fs01**

## **II. RUN VERIFY AFTER ENSURING MIGRATION IS COMPLETED & MIGERRs CHECKED:**

**#server\_cdms server\_2 –verify fs01**

### **TACTICS FOR RESOLVING OFFLINE INODES:**

1. Determine if the offline inode can be seen on the Celerra file system. If you cannot see the offline inode on the Celerra, then there is a good chance that it was deleted on the Source side during the migration.

**Note:** Deleted inodes are actually placed “online” before they are deleted for CDMS

2. If cannot be seen on Celerra, check to see if it can be seen on the Source and handle accordingly. Determine if file is corrupt on the Source side, which could prevent migration.

→In most cases, the first step after reviewing the Logs would be to start the MGFS threads again to force the migration to bring inodes online

**#server\_cdms server\_2 –start mgfs -path /conn -log /path/subpath -Force**

**Note:** For CIFS CDMS migrations, use the –Force flag only for “restarting” a CIFS thread for the migration, or else the CDMS thread will think that a directory marked ‘online’ does not need checking for its subdirectories and files! This command does not alter inodes. NFS migrations do not experience this limitation. Deleting an offline inode could also be an option at customer discretion. This command is documented in NAS 5.4.

3. Running the /nas/tools/cdms/specialCmd can be used to bring an offline inode online—specify the complete path to the inode  
4. Cat the offline inode file directly to /dev/null and this should also place the inode online

### **SERVER CDMS COMMAND OPTIONS:**

**\$ server\_cdms**

```
-connect <mgfs> -type {nfsv2|nfsv3} -path <localpath> -source <srcName>:<srcPath> [-option <options>]
-connect <mgfs> -type cifs -path <localpath> -netbios <netbios> -source \\<srcServer>[.<domain>]\<srcShare>[\<srcPath>]
-disconnect <mgfs> {-path <localpath>|-path <cid>|-all}
-verify <mgfs> [-path <localpath>|-cid>]
-Convert <mgfs> [Caution—use this command only as a last resort and after all else has been checked]
-start <mgfs> -path <localpath> -log <logpath> [-include <include_path>] [-exclude <exclude_path>]
-halt <mgfs> -path <localpath>
-info [<mgfs>] [-state {START|STOP|ON_GOING|ERROR|SUCCEED|FAIL}]
```

### **OFFLINE INODE PROBLEM WITH CDMS MIGRATION:**

#### **PROBLEM INDICATORS:**

**# server\_cdms server\_2 -i fs01**

```
server_2 :
fs01:
path  = /temppath
cid   = 0
type  = cifs
source = \\source1.rcb.ford.com\u1$ netbios= RCB9010201.NA1.FORD.COM
admin  = na1.ford.com\$kcardwel
threads:
path  = /temppath
state = FAIL
log   = /cdms_logs
```

#### **SERVER LOG:**

**2004-12-30 14:54:14: MGFS: 3: node 547058 is offline, abort**

2004-12-30 18:35:58: MGFS: 3: node 171621 is offline, abort

1. Use calculator to convert offline inode number from Server Log to true inode number if using /nasmcd/quota/slot\_2 to locate and fix the offline inode. If you nfs mount the file system from the Control Station, then the inode to find and fix is the same number as listed in the Server Log:

-->Open calc, enter inode number, 547058, then xOR file system id, 28, and result is 547070. This will be the inode number to use when conducting find and fix to the offline inode.

2. Locate file on file system:

**\$ find /fs01 -inum 547070**

547070 -rw-r--r-- 1 1000943 3000002 553 Aug 4 08:24 img\_presentation.gif

3. Check Migration logs to identify problems

4. Run MGFS Thread against file system path or subdirectory path to try and convert again

5. Optionally, bring inode online by running cat offline\_file /dev/null from NFS, or use following script tool:

**#/nas/tools/cdms/specialCmd getOnline server\_2 /fs01/path/specific\_file**

6. Run Verify to Check Status only after verifying that MGFS Threads are completed and no I/O is occurring:

**# server\_cdms server\_x –verify fs01**

**# server\_cdms server\_2 -i fs01**

**# Server\_log server\_2**

2004-12-31 10:59:16: MGFS: 4: Checking

2004-12-31 10:59:16: MGFS: 4: in-processing: fsid=28 0% done

2004-12-31 11:01:13: MGFS: 4: in-processing: fsid=28 10% done

2004-12-31 11:03:06: MGFS: 4: in-processing: fsid=28 20% done

2004-12-31 11:16:11: MGFS: 4: in-processing: fsid=28 90% done

2004-12-31 11:18:03: MGFS: 4: in-processing: fsid=28 100% done

2004-12-31 11:18:03: MGFS: 4: remove cid 0 in namedb succeed

2004-12-31 11:18:03: ADMIN: 4: Command succeeded: mgfs action=verify fsid=28

## **CDMS MIGRATION NOTES:**

--“mgfs” file system creates “offline migration inodes” which contains the actual data during the migration

--An “mgfs” file system can be extended during the Migration--will not to restart the ./migrate.pl script

--“migrate.pl” script should be started during periods of low I/O--works by reading last byte of each file in File System

--File System may be larger after migration due to our 8k block size requirement

**Note:** After the script is run to “read” contents of every file that has been migrated, the offline inodes are marked “online” and the data is written to disk.

--recommendation would be to export Source File System to Celerra as “root=” so that there will be no problems migrating files that are owned by Root

--If using 32-bit GIDs for file systems, set param on DM prior to migration [param ufs gid32=1]

--CDMS does not support sparse files during migrations, will result in dense files

## **CDMS FOR NFS V2:**

### **SUPPORT:**

--I18N supported, but must configure prior to migration

--NFS V2 & V3 supported, for UDP & TCP protocols

--Can combine multiple source filesystems into single UxFS filesystem

--Export aliasing permits consolidated filesystems to be reexported under original names

--Simultaneous migration of multiple filesystems

--Migrates Inode Information: Size, Access & Modification Times, UID, GID, Privilege & Mode bits, Link count

--Migrate from NT 4.0 Domain or Windows 2000 Domain, but not both at same time

--migration supported without using Portmapper

--Symbolic and Hard links are migrated and preserved

### **LIMITATIONS/REQUIREMENTS:**

--100MB Network, with recommended limit of (3) migration streams in parallel

--Files that cannot be opened by NFS cannot be migrated

--Filesystem Quota information is not migrated

--Timestamps with attribute changes are not migrated [chmod attributes, etc]

--NFS V2 supports max. 2GB Filesizes

--Compressed files migrated uncompressed

--When migrating from NetApp Source using Network Appliance Snapshot, must use additional software on Celerra side

--local mountpoints can cause dircount.pl script to fail

**Cautionary Note:** Files that cannot be opened by the migrating Data Mover (for instance, files in use by the system or other locks), will not be migrated!

## **CDMS V2 FOR CIFS SUPPORT:**

--Migration of CIFS attributes: ACLs, DOS attributes, Audit bits

## **CDMS V2 CIFS LIMITATIONS:**

- ADS named streams are not supported for migration
- I18N must be configured on the Target Celerra Server in order for CIFS CDMS migration to work, as Unicode is used in process!
- Migration of DOS 8.3 names not guaranteed due to name mangling differences between Windows and Celerra
- No cross-domain migrations allowed (NT → WIN2K & not vice versa—however, in order to migrate from an NT Domain to a Windows 2000 Domain, must have the normal Trust prerequisites in place and have run ADMT migration with SID History)
- Migration of Quotas Information is not supported—recreate Quotas after the migration
- Sparse Files are not supported—migrated sparse files will become dense files
- Compressed data will be migrated uncompressed
- Single Domain Migrations only
- Does not support EFS (Encrypted File System) migrations for Windows 2000—decrypt files first. Nas 4.2 & 5.0 will allow these files to be migrated, but they will be useless. Nas 5.1 will not allow the migration of encrypted files.
- Migrations are done for NFS or CIFS attributes, but cannot migrate over both attributes at one time!
- DFS is not supported for CDMS as there is no WIN32API built in to support this yet
- Windows has a directory path length limitation of 255 characters [Identify these problem paths first, then run separate migrations at different levels to accomplish complete migration of all data from the Source]
- Source file server should be set to RO once share migration begins
- XP systems cannot be a migration Source
- Source NT Server only supported for migrations within an NT 4.0 security domain, not W2K or W2K3
- Local Groups and CIFS Shares cannot be migrated by CDMS alone [use lgdup and sharedup utilities]
- Local Users can be migrated if the Local Users feature is configured on the CIFS Server, and running NAS 5.4 or higher

**Note:** Previously, Local User account ACLs were dropped during migration. Invalid owner or primary group SIDs were replaced by SID of account used in migration

## **CIFS CLIENT TIMEOUT VALUE:**

# server\_param server\_2 -facility cifs -info cifsclient.timeout -v

```
server_2 :  
name      = cifsclient.timeout  
facility_name = cifs  
default_value = 20000  
current_value = 20000  
configured_value =  
user_action = none  
change_effective = immediate  
range      = (0,4294967295)  
description = CIFS client timeout used in CDMS (default:20000)  
detailed_description
```

Specifies the amount of time in milliseconds that is required for a response to a cifs client request from one Data Mover to another cifs server (other DM, regular MS server or other cifs server). Mainly used by CDMS. Value is defined in milliseconds.

## **CDMSv2 SUPPORT:**

- Timefinder/fS
- HighRoad MPFS
- NT/Windows 2000
- SRDF
- MGFS filesystem can be extended during a migration
- Migrating up to 1024 filesystems into one can be done [but migrate only 4-5 at a time]

**Note:** What this really means is that you can have up to 1024 concurrent connections to the Source Server's local directory structure, implying that you can migrate multiple filesystems, directories, etc.

## **BASIC METHODOLOGY FOR CIFS CDMS MIGRATIONS:**

1. LGDUP from Source to Target Server to copy over local group database
2. Setup, mount, and export the MGFS file system & log file system on the Target
3. Create the CDMS connection, which then copies over the inode structure of Source fs to Target fs
4. Conduct Sharedup to copy over Share information
5. Start the actual CDMS migration thread(s)
6. Monitor migration using nas\_fs -s
7. Inspect logs after completion to verify migration of fs
8. Convert file system to MGFS

## **CONDUCTING CIFS CDMS MIGRATION:**

### **1. Migrate LocalGroups database using lgdup, from Source to Target systems**

### **2. Create MGFS File System on Destination using GUI or CLI:**

2005-12-12 12:04:16.692 db:201:29385:S: nas\_fs -name speced -type mgfs -create size=30720M pool=clar\_r5\_performance storage=SINGLE -option slice=y,mover=server\_2

### **3. Mount MGFS file system Read/Write:**

2005-12-12 12:04:28.641 server\_2:201:29485:S: server\_mountpoint server\_2 -create /speced

2005-12-12 12:04:31.789 server\_2:201:29543:S: server\_mount server\_2 -option rw speced /speced

### **4. Start CDMS Service on Destination CIFS Server:**

**\$server\_setup server\_2 -P cdms -o start** (default 32 threads)

### **5. Create CIFS CDMS Connection between Source and Target (always run from Target Side):**

2005-12-12 12:06:43.703 server\_2:0:30570:S: server\_cdms server\_2 -connect speced -type CIFS -path /speced -netbios dpsnas01 -source \\dpsnas01\_t.dpsk12.org\c\$\speced -admin dpsk12.org\cdms

**Note:** From CLI, actual syntax requires the use of ‘ ‘ for UNC and Admin paths, as well as password

**# server\_cdms server\_2 -connect speced -type cifs -path /speced -netbios dpsnas01 -source**

**'\\dpsnas01\_t.dpsk12.org\c\$\speced' -admin 'dpsk12.org\cdms'**

server\_2 : Enter Password:\*\*\*\*\*

### **6. Create Share on Destination Server for MGFS file system called “speced”**

### **7. Start CDMS Migration Thread:**

2005-12-12 12:10:26.945 server\_2:0:32238:S: server\_cdms server\_2 -start speced -path /speced -log /mgfslog

### **8. Verify Migration Progress:**

**# server\_cdms server\_2**

server\_2 :

CDMS enabled with 32 threads.

speced:

path = /speced  
cid = 0  
type = CIFS  
source = \\dpsnas01\_t.dpsk12.org\c\$\speced  
netbios= DPSNAS01.DPSK12.ORG  
admin = dpsk12.org\cdms

threads:

path = /speced  
state = ON\_GOING  
log = /mgfslog  
cid = NONE  
recall = FALSE

**# nas\_fs -i speced**

id = 32  
name = speced  
acl = 0  
in\_use = True  
type = mgfs  
worm = off  
volume = v118  
pool = clar\_r5\_performance  
member\_of = root\_avm\_fs\_group\_3  
rw\_servers= server\_2  
ro\_servers=  
rw\_vdms =  
ro\_vdms =  
status =

|           | Total KBytes | Used KBytes | Total inodes | Used inodes |
|-----------|--------------|-------------|--------------|-------------|
| Dart cid: | 30977264     | 18675872    | 3778558      | 107350      |

remote 0: 1978496 184624 0 0

type=CIFS path=speced cifs=\\dpsnas01\_t.dpsk12.org\c\$\speced localServer=DPSNAS01.DPSK12.ORG account=dpsk12.org\cdms passwd=\*

**# nas\_fs -s speced**

```
total = 30251 avail = 11527 used = 18723 ( 61% ) (sizes in MB) ( blockcount = 62914560 )
```

```
volume: total = 30720 (sizes in MB) ( blockcount = 62914560 )
```

**# .server\_config server\_2 -v "mgfs action=query fsid=32"**

```
1134416714: MGFS: 4: Total KBytes Used KBytes Total inodes Used inodes
1134416714: MGFS: 4: Dart cid: 30977264 9590128 3778558 62495
1134416715: MGFS: 4: -----
1134416715: MGFS: 4: remote 0: 1978496 184624 0 0
1134416715: MGFS: 4: type=CIFS path=speced cifs=\dpsnas01_t.dpsk12.org\c$\speced localServer=DPSNAS01.DPSK12.ORG
account=dpsk12.org\cdms passwd=*
1134416715: MGFS: 4: -----
1134416715: ADMIN: 4: Command succeeded: mgfs action=query fsid=32
# server_export server_2 |grep -i speced
share "vol2" "/speced/speced/vol2" netbios=DPSNAS01 maxusr=4294967295 umask=22 comment="Student Services Share"
# cat /nas/server/slot_3/param
param mgfs logPathPrefix=/mgfslog/mgfs
# cat /nas/server/slot_3/netd
migThreadPool action=startPool
```

**NEW SERVER CDMS COMMANDS USED WITH NAS 5.2:**

**CREATING MGFS FILE SYSTEM:**

```
#nas_volume -n vol_mgfs1 -c d30
#nas_fs -n mgfs1 -t mgfs -c vol_mgfs1
```

**MOUNTING & EXPORTING MGFS FILE SYSTEM:**

```
#server_mountpoint server_x -c /mgfs1
#server_mount server_x mgfs1 /mgfs1
```

```
#share -F nfs -o ro,root=10.241.169.20/0 /fs1 [Exporting Source for DM with root access privileges]
```

**SETTING UP CDMS SERVER THREAD POOL:**

```
#server_setup server_x -P cdms -o start=50
```

**CONNECTING TO CDMS EXPORT FOR NFS:**

```
#server_cdms server_x -connect mgfs1 -type nfsv3 -path /cdmsexport -source
10.241.169.30:/cdmsexport -o useRootCred=true proto=TCP
```

**Note:** The ‘useRootCred’ option ensures that data reads on source are UID=0, GID=1, instead of file owner’s UID/GID, especially in environments with mixed permissions on files and directories. Also export Source file system to Data Mover as root.

**CONNECTING TO CDMS SHARE FOR CIFS:**

```
#server_cdms server_x -connect mgfs1 -type cifs -path /cdmsshare -netbios cdmsserver -source
\\source.domain\cdmsshare -admin Administrator
```

**EXAMPLE CIFS CONNECT COMMAND:**

```
# server_cdms server_2 -connect fs303 -type cifs -path /Groups -source "\\\xbwsdata1\groups" -netbios
bwssdata1 -admin ex88889
```

```
server_2 : Enter Password:*****
```

```
done
```

**Note:** When specifying the path “/Groups”, this is the path and share that actually exists on the Source side and does not exist on the Target side by foldername or sharename—there should be no CIFS shares in use when conducting the migration to the Data Mover. Furthermore, the CDMS ‘connection’ and ‘migration start’ commands can be run from CLI or WebUI GUI.

**VERIFYING CDMS CONNECTION:**

```
# server_cdms server_x -verify mgfs1 -path /cdmsshare
```

**VERIFYING DATA MOVER CDMS CONNECTION TABLE:**

```
$ .server_config server_5 -v "queryCid fsid=547 cid=0 range=20"
```

```
1124861497: MGFS: 4: ----- cid information -----
```

```
1124861497: MGFS: 4: cid=0 fsid=547 type=CIFS path=/studappdata1 cifs=\studentspc8.students.intranet.epfl.ch\AppData\
localServer=STUDENTSPC6.STUDENTS.INTRANET.EPFL.CH account=students.intranet.epfl.ch\rootb wins=128.178.50.44
passwd=*
```

```
1124861497: MGFS: 4: ----- end cid information -----
```

1124861497: ADMIN: 4: Command succeeded: queryCid fsid=547 cid=0 range=20

### **VERIFYING DM CREDENTIALS TO REMOTE HOST:**

**#.server\_config server\_5 -v "mgfs action=query fsid=547"**

```
1124860820: MGFS: 4:          Total KBytes  Used KBytes  Total inodes  Used inodes
1124860820: MGFS: 4:  Dart  cid:  206515184    6104288    25190398    208377
1124860820: KERBEROS: 7: send_as_request
1124860820: KERBEROS: 7: _krb5_use_dns: use dns = yes
1124860820: KERBEROS: 7: krb5_locate_srv_dns: for service _kerberos at realm STUDENTS.INTRANET.EPFL.CH, # KDC: 2
1124860820: KERBEROS: 7: krb5_locate_srv_dns: name studentsdc1.students.intranet.epfl.ch
1124860820: KERBEROS: 7: krb5_locate_srv_dns: prior. 0
1124860820: KERBEROS: 7: krb5_locate_srv_dns: port 88
1124860820: KERBEROS: 7: krb5_locate_srv_dns: IP @ 128.178.50.116
```

### **IN MEMORY CID MAP vs. ONDISK MAP:**

**\$ .server\_config server\_2 -v "queryCid action=onDiskCidQuery fsid=1503"**

```
1263504004: MGFS: 6: ----- cid information -----
1263504004: MGFS: 6: ----- end cid information -----
```

1263504004: ADMIN: 6: Command succeeded: queryCid action=onDiskCidQuery fsid=1503

**\$ .server\_config server\_2 -v "queryCid action=incoreCidQuery fsid=1503"**

```
1263504032: MGFS: 6: ----- cid information -----
1263504032: MGFS: 6: Last Cid Available is 1022
1263504032: MGFS: 6: cidMap: (0) state=2 rpolicy=15 wpolicy=1 http=http://tlsm389a.cw01.contiwan.com cgi=
1263504032: MGFS: 6: cidMap: (1023) dart://RDE
1263504032: MGFS: 6: ----- end cid information -----
1263504032: ADMIN: 6: Command succeeded: queryCid action=incoreCidQuery fsid=1503
```

### **VERIFYING DATA MOVER CLIENT CONNECTIONS TO SOURCE:**

**\$ .server\_config server\_2 -v "cifsclient audit"**

```
1134574620: SMB: 4:
Server:CLUSFS001NASTST.us.cly (10.55.9.92) Client:CLUSFS001NASTST.US.CLY (10.55.9.92)
1134574620: SMB: 4: NativeOS: EMC-SNAS:T5.3.20.1
1134574620: SMB: 4: NativeLanman: NT1
1134574620: SMB: 4: Status: Connected
1134574620: SMB: 4: Auth mode: kerberos
1134574620: SMB: 4: Capabilities: 8000e3fd
1134574620: SMB: 4: Encrypted password: Yes
1134574620: SMB: 4: Protocol: NT LM 0.12
1134574620: SMB: 4: Refcount: 3
1134574620: SMB: 4: List of sessions:
1134574620: SMB: 4: * Name: us.cly\hsadmin
1134574620: SMB: 4: uid: 3f
1134574620: SMB: 4: Vc: 0
1134574620: SMB: 4: RefCount: 3
1134574620: SMB: 4: List of connections::
1134574620: SMB: 4: * Name:zummy_archive
1134574620: SMB: 4: tid:: 3f
1134574620: SMB: 4: RefCount: 2
1134574620: SMB: 4: Service: (NULL)
1134574620: SMB: 4: Mounted by: us.cly\hsadmin
```

### **DISCONNECTING FROM CDMS SHARE:**

**# server\_cdms server\_x –disconnect mgfs1–path /cdmsshare**

**Note:** Do not issue this command unless absolutely required as the entire migration will have to be done over. Rebooting Data Mover should allow a halted migration to continue.

### **STARTING CDMS MIGRATION:**

**# server\_cdms server\_x –start mgfs1 –path /cdmsexport –log /mgfs1/logs.txt**

### **EXAMPLE STARTING CDMS CIFS MIGRATION:**

**2004-06-26 12:37:59.954 server\_2:201:28663:S: server\_cdms server\_2 -start fs303**

### **-path /Groups -log /fs00**

**Note:** Do not have to use a special param to create a log file, simply point to a file system mountpoint and the default log name is as follows: **/fs00 → migLog\_fs303\_Groups**

### **HALTING CDMS MIGRATION:**

**# server\_cdms server\_x –halt mgfs1 –path /cdmsexport**

### **OBSERVING CONNECTION AND THREAD STATUS OF CDMS MIGRATION:**

**# server\_cdms server\_x –info | mgfs1 | -info –state failed | -info –option format=parseable | format=raw**

**Note:** -state switch displays threads for state info requested. Parseable is in comma-separated fields, raw is human readable form.

### **PROBLEM WITH ALL CDMS THREADS IN USE:**

#### **Server Log:**

2007-04-07 22:14:01: MGFS: 4: migThread: All migration threads are in use

2007-04-07 22:14:01: ADMIN: 3: Command failed: migThread action=start fsid=107 migpath=/h/fs01\_ns80g logpath=/miglog1 recall=FALSE

→For a variety of reasons (deleting MGFS filesystem with thread in use; forced converts then threads are hung or incomplete, connection is broken while threads are active, etc.), there are situations where all CDMS migration threads can show up as in use, preventing further migrations. Perform the following steps to remove the “stuck” CDMS threads:

- 1.) Halt all running migration threads (if any) using “server\_cdms server\_2 –halt”, but leave the CDMS connections as is
- 2.) Run “server\_setup server\_2 -P cdms -o delete” to clean out the “stuck” CDMS threads, but not to remove the CDMS data table, or affect the CDMS connections—cleans out the state and data information in the table used to track threads, but not the table itself
- 3.) Restart the migration threads per the normal method

### **MIGRATION STATES:**

start → thread has started stop →thread has been stopped by user

ongoing →thread is running without errors

error →thread is running with errors—check migration log and server log for error messages

succeed →thread is completed with user-specified directory migration

fail →thread is completed but with errors

### **FORCING MGFS FS TO UXFS TO COMPLETE MIGRATION:**

**# server\_cdms server\_x –Convert mgfs1**

**Caution:** This should never be used except as a last resort—ensure integrity of migrated filesystem before using this

### **EXAMPLE OF CDMS SERVER OUTPUT:**

**# server\_cdms server\_5**

```
server_5 :  
CDMS enabled with 50 threads.
```

```
test2:  
path = /test5  
cid = 1  
type = nfs  
source = 10.64.133.11:/export/spare/toby/TEST3  
options= useRootCred=True
```

### **EXAMPLE OF MIGRATION STATUS WITH ERROR CONDITION:**

**# server\_cdms server\_2 -info fs303**

```
server_2 :  
fs303:  
path = /Groups  
cid = 0  
type = cifs  
source = \\brwsdata1.PHARMA.AVENTIS.COM\groups\  
netbios= BRWSDATA1.PHARMA.AVENTIS.COM  
admin = ex88889  
threads:  
    path = /Groups  
    state = ERROR  
    log = /fs00
```

**Note:** Corrective action for a migration that has stopped would be to restart migration using –start command

### **MIGRATION STATE ON GOING=ACTIVE:**

**# server\_cdms server\_2 -i fs303**

```
server_2 :  
fs303:  
path = /Groups  
cid = 0  
type = cifs
```

```
source = \\xbrwsdata1.PHARMA.AVENTIS.COM\groups\  
netbios= BRWSDATA1.PHARMA.AVENTIS.COM  
admin = ex88889  
threads:  
path   = /Groups  
state  = ON_GOING  
log    = /fs00
```

## **VERIFYING CDMS MIGRATION STATUS/PROGRESS:**

# **server\_cdms server\_2 -i fs303**

# **nas\_fs -s fs303**

total = 1008374 avail = 684787 used = 323587 ( 32% ) (sizes in MB)

volume: total = 1024000 (sizes in MB)

## **LISTING OFFLINE INODES:**

# **.server\_config server\_2 -v "param mgfs listAllOfflineInodes"**

## **CDMS TOOLS: APPS & TOOLS CD**

### **NFS:**

ch\_group.pl—change GID numbers on a tree against GID template, used for consolidating file systems with duplicate GIDs

dircount.pl—tool to size the directory tree

dirprivU.pl—verifies file privileges & access and what files can be read so migration will work

diskUsage.pl—estimates storage space required on Target Celerra FS

migrateU.pl—script for Unix donor systems

**Note:** Recommendation is to use Perl 5.6 or 5.8

### **CIFS:**

migrateW.pl—Windows migration script

dirprivW.pl—checks that files can be Read by Windows for successful migration (checks privileges & access)

dircount.pl—use this tool to traverse directory tree & output map of filesystem into CSV file, import to Excel to sort by size, etc.

diskusage.pl—use this tool to determine how much space will be required on the Target filesystem—takes Shadow file into account

connBuilder.pl—Pre-migration tool, output from sharedup.exe to map top level shares for connection cmd's, builds exclude file list

backupWrapper.exe—Gives Backup Operator privileges to running process so scripts can read exclusively owned files for CIFS migrations

Pototype.PM—perl module required by Windows for running all scripts, used in CIFS migrations only

### **LOGFILES:**

/mgfs1.migLog\_mgfs1 and migErr\_mgfs1

## **NATIVE TOOLS ON CELERRA: /nas/tools/cdms**

disconnect.pl—script to remove unwanted connections

setParam.pl—

snapshotHandler.pl—marks inodes online, used especially for NetApps systems that use ‘snap shot’ feature

#specialCmd getOnline server\_x fsid cid [brings offline directories online; CID=Common Identifier]

#./specialCmd disconnect server\_2 27 <cid> [removes connection without migrating data]

## **DOS 8.3 NAME MANGLING:**

Celerra handles 8.3 names exactly as does Windows. Whenever large numbers of like-named files are copied, Windows or Celerra will not retain exact 8.3 shortened name but will dynamically assign new 8.3 names based on the order of the copy operation.

## **SUPPORTED CDMS V2 NFS PLATFORMS:**

|                                       |                    |
|---------------------------------------|--------------------|
| NetApp Data ONTAP 5.3.4, 5.3.6, 6.1.1 | NFS v2/v3; 4.2.x + |
| AUSPEX 1.9.2                          | NFS v2/v3; 4.2.x + |
| Sun Solaris 2.6/2.7/2.8               | NFS v2/v3; 4.2.x + |
| IP4700 R2.0, R2.1                     | NFS v2/v3; 4.2.x + |
| Celerra 1.2.57, 2.2, 4.1              | NFS v2/v3; 4.2.x + |
| HP UX 11.0                            | NFS v2/v3; 4.2.x + |
| AIX 4.3.3                             | NFS v2/v3; 4.2.x + |

## **SUMMARY OF MIGRATION STEPS:**

1. Create MGFS File System on Celerra
2. Mount MGFS File System
3. Connect to Source File System
4. Export the MGFS File System
5. Query, Verify, Convert MGFS File System after migration, permanently unmount and remount from Celerra

## **CDMS V.2 CIFS MIGRATION:**

- Scenario:** Migrating FileSystem from NT 4.0 Server to CIFS NT 4.0 Data Mover  
--Not recommended for WAN connections  
--Not recommended for many very large files  
--Does not support migration from NT to Win2k domains [Just NT to NT or Win2k to Win2k]  
--Biggest benefit of CDMS is that it allows for migration online with less downtime  
--Multiple sources can now be migrated to a single filesystem

### **1. Create Migration User Account on NT DC:**

- Create User called “cdms\_migrator”, add to Domain Admins and BackUp Operators Global Groups  
--Set Primary Group membership to Domain Admins  
--Remove cdms\_migrator account from Domain Users Group  
--Add User to Target & Source Servers Administrators and BackUp Operators local groups  
--Grant User the Generate Security Audits & Manage Auditing & Security Logs rights on Source & Target Servers  
**Note:** During migration, assign “Deny Access to this computer from the Network” to Domain Users group & grant “Access to this computer from the Network” for the migration User account.

### **2. Conduct complete Tape Backup of Source Server’s Data:**

### **3. Setting Up NT Server With CDMS Migration Tools:**

- Download and install Perl 5.6.1 or later from [www.activestate.com](http://www.activestate.com)  
--Download and install special Win32 API Library for Perl from [www.roth.net/perl/packages](http://www.roth.net/perl/packages)  
c:>ppm  
PPM>install Win32 – API – Prototype

### **4. Copy CDMS Tools from CD-ROM to NT Server [c:\Copy Tools CD\CDMS\cifs>]**

### **5. Run LGDUP Tool to migrate Local Groups from Source to Target Servers:**

c:>lgdup -l logfile.log \\nt1 \\ntdm3

### **6. Run following CDMS Perlscript Tools to Verify & Validate FileSystem to be Migrated:**

**>backupWrapper.exe dirprivW.pl**

This script attempts to read one byte from every file, ensures that directory path names are less than 256 characters, and determines access rights to Share based on the User Account you are logged in with. Script also checks to see if encrypted bits are set on files for Windows 2000 Servers—cannot migrate these files.

**>backupWrapper.exe dircount.pl j:\\**

This script also traverses directory tree path and determines number of path levels deep and number of names involved, and is capable of outputting information into CSV file, which can be imported into Excel, then sorted for path lengths, file sizes, etc.

**>backupWrapper.exe diskusage.pl -m -h**

Script that looks at Source disk usage and translates into Celerra’s 8k block size requirement to estimate Target disk usage

### **7. CREATE LOG FILESYSTEM:**

```
#nas_slice -n disk3 -c d3 11000 1
#nas_volume -n log_vol -c disk3
#nas_fs -n log -c log_vol
#server_mountpoint server_2 -c /mgfslog
#server_mount server_2 /mgfslog
#vi /nas/server/slot_2/param
param mgfs logpathprefix=/mgfslog/mgfs
#server_export server_2 -o anon=0 /mgfslog
Map or Mount to log file system to view logs
```

### **8. Create MGFS FileSystem:**

```
#nas_volume -name str1 -create -Stripe 32768 d3,d4,d5,d6
#nas_volume -name -mtv1 -create -Meta str1
#nas_fs -name mgfs1 -t mgfs -create mtv1 -o mover=server_3
#nas_fs -i mgfs1 [Verify that filesystem is mgfs]
```

**Note:** NAS Versions prior to 5.1 require CDMS license and invokes CallHome

### **9. Mount MGFS FileSystem:** [Front End Mount]

```
#server_mountpoint server_3 -c /mgfs1
#server_mount server_3 mgfs1 /mgfs1
#nas_fs -i mgfs1 [Note change after mounting the filesystem]
```

### **10. Create CDMS Connection to Source:**

```
#server_mount server_3 -o connect,type=CIFS,path=/f$,cifs=\\\\nt1.cdmstn\\f$\\,netbios=ntdm3.cdmstn,account=cdmsnt\\cdmsntb
mgfs1
```

Password: \*\*\*\*\*

**Note:** path=/f\$ is Sharepoint on Source NT Server; \\nt1.cdmsnt\\ f\$\\ is Source NT Server Netbios & Domain name; netbios=ntdm3.cdmsnt is Netbios name for Target Server\_3 & NT Domain name; account=cdmsnt\\ is NT Domain Name; cdmsntb is User Account used in migration—must be a member of Domain Admins and BackUp Operators Groups.

## **11. Create CIFS Share for the MGFS FileSystem:**

```
#server_export server_3 -P cifs -n mgfs1 /mgfs1
```

## **12. Migrate Shares from Source to Target Servers:**

```
c:>sharedup \\nt1 \\ntdm3 f: /SD /P:server_3\mgfs1 /LOG:mgfs1.log
```

## **13. Map Drive From Source NT Server to Target DataMover Sharepoint:**

```
Start>Run: net use * \\ntdm3\mgfs1
```

**Note:** In this example, the drive mapping created drive J: on the NT Server, against which the Migration script will be run

## **14. Start the Migration Script:**

```
c:\Copy Tools CD\CDMS\cifs>backupWrapper.exe migrateW.pl J:\
```

## **15. Observe Progress of Migration:**

```
$server_log server_3 -s -f
```

```
$nas_fs -i mgfs1
```

**Note:** Touch data on folders in migration path for f\$ to help migration along

## **16. Run Verify Command at end of Migration:**

```
#server_mount server_3 -o verify mgfs1
```

**Note:** This will run through entire filesystem to verify the migration & output results to screen

## **17. Convert MGFS FileSystem to UXFS:**

```
#server_mount server_3 -o convert mgfs1
```

```
#nas_fs -rename mgfs1 uxfs1
```

```
#server_umount server_3 -p /mgfs1
```

```
#server_mount server_3 uxfs1 /mgfs1
```

## **RESOLVING INODE ISSUES DISCOVERED WITH VERIFY COMMAND:**

**Note:** CDMS migration cannot complete without the resolution of all unknown inodes. Verify command will indicate which inodes are a problem.

1. \$server\_export server\_x -o anon=0 /mgfs1

2. #mount 192.10.3.32:/mgfs1 /mnt1

3. #find /mgfs1 -inum 187 -print or \$ls -ailR /mgfs1 |grep 187 #find . -inum 14 -print [Prints out path & file for inode]

**Resolution:** Delete inode or force online using specialCmd getOnline server\_x /mgfs1

## **#/nas/tools/cdms/specialCmd getOnline server\_2 /path/file**

**Note:** Use \* for each space if there are spaces within names in the path and enclose entire path in quotations “ “

**[specialCmd getOnline server\\_2 '/fs05/01/Data/Warehouseusers/!!20050120\\_SE\\_EC\\_MIGRATION\W534MGR3/Tony\'s'](#)**

**Note:** Example of various escape slashes and single quotes required to make command execute with \$ & ‘ marks in path

4. Verify effect of command in Server Log

## **TWO METHODS FOR MIGRATING CIFS DATA FROM SOURCE:**

**PREREQUISITES:** sharedup.exe operation & connbuilder.pl utility

## **CONNECTION BUILDER UTILITY:**

Connbuilder.pl utility's purpose is to migrate only those directories of a Source Server that had Shares that you wanted. In otherwords, any folders [such as Program Files, applications, etc] that you did not want to migrate over, you would not have Shared.

Step 1. Use sharedup.exe tool with the /FO switch to obtain a share map of the Source filesystem and output to a file

Step 2. Use connbuilder.pl [must be using a system with Excel & Word programs] to run against sharedup.exe file output to produce template files to support either an Administrative or Top-Level Share migration.

## **ADMINISTRATIVE SHARE METHOD:**

Everything below the Administrative mountpoint is migrated. Data not needed can be ‘excluded’ using exclude parameters in the migrateW.pl, but resulting filesystem “structure” still exists.

## **TOP LEVEL SHARE METHOD:**

Everything underneath the “toplevel” share is migrated, but nothing above.

## **CDMS V.1 MIGRATION:**

|                     |             |                    |
|---------------------|-------------|--------------------|
| Source: Unix System | 192.1.4.250 | Netapps NFS Server |
|---------------------|-------------|--------------------|

|                  |            |                    |
|------------------|------------|--------------------|
| Target: Server_3 | 192.1.4.21 | Celerra NFS Server |
|------------------|------------|--------------------|

|                      |             |                                                                          |
|----------------------|-------------|--------------------------------------------------------------------------|
| Client: Unix Solaris | 192.1.4.200 | [Requires Perl 5.0+ to run migrate.pl script, diskusage.pl script, etc.] |
|----------------------|-------------|--------------------------------------------------------------------------|

**CDMS NFS MIGRATION STEPS:**

**Note:** Target Server requires “root” access to everything on Source because migration requires the Reading of every file!

Step 1. Sun Solaris Client: Must have Perl 5.005\_3+ installed. Load "migrate.pl" & "diskUsage.pl" scripts to Client.

**Note:** Ensure perl is located in correct path! [#which perl]

Step 2. Mount Source NFS Server's File System to SUN Solaris Client:

**#mount 192.1.4.250:/vol/test /mnt2**

a. Determine Disk Usage Requirements for Migration: **./diskUsage.pl -m | 1 | /mnt2 | 8 | no**

Step 3. Create MGFS File System on Target Celerra NFS Server:

a. **\$nas\_volume -n mv\_mgfs1 -M -c d11** [creates metavolume for mgfs fs]

b. Run **#nas\_license -I** to obtain Password Key: 3a8dedc0

b. **#nas\_fs -n mgfs1 -t mgfs -c mv\_mgfs1** [Creates mgfs File System]

**Note:** During creation of MGFS FS, will prompt for password key obtained above

c. Create mountpoint for MGFS File System on Celerra:

**\$server\_mountpoint server\_3 -c /mgfs1**

d. Mount MGFS FileSystem on Server\_3:

**\$server\_mount server\_3 mgfs1 /mgfs1**

Step 4. Create MGFS Logging File System on Target Celerra:

a. Create standard ufs file system for MGFS Logging:

**\$nas\_fs -n mgfs1log -c mtvmgfs1log**

b. Create mountpoint for Log File System:

**\$server\_mountpoint server\_3 -c /mgfs1log**

c. Mount MGFS Log FS:

**\$server\_mount server\_3 mgfs1log /mgfs1log**

d. Edit Param File on DataMover so as to log to this new File System:

**\$vi /nas/server/slot\_3/param**

**param mgfs logPathPrefix=/mgfs1log/mgfs1** [MGFS FS/Log Path]

e. Reboot DataMover

f. Export Log File System on DataMover:

**\$server\_export server\_3 /mgfs1log**

g. Mount to Target File System from Source Server:

**\$server\_mount server\_3 -o connect,type=NFSV3,path=/source,nfs=10.100.50.14:/mgfs1**

Step 5. Unmount Source NFS FileSystem and remount as "RO" for Host Access during Migration

Step 6. Mount Source NFS File System to Server\_3's "MGFS" File System as ReadWrite:

**\$server\_mount server\_3 -p -o nfs=192.1.4.250:/vol/test rw mgfs1 /mgfs1**

Step 7. Export Source NFS File System on Server\_3's MGFS Mountpoint:

**\$server\_export server\_3 -o anon=0 /mgfs1**

**\$server\_export server\_3 -P nfs -n mgfs1 -o anon=0 /mgfs1**

Step 8. From Sun Client remote mount Server\_3's MGFS File System:

**#mount 192.1.4.21:/mgfs1 /server\_3**

Step 9. From Sun Client, begin data migration from Source Server to Target Celerra Server\_3:

a. Have Users mount "mgfs1"on Server\_3 and access data--this will "pull" data over automatically

b. Use Migration Tool Script:

**#export/home/user2/migrate.pl -m /server\_3** [Use of -m switch recommended]

**./migrateU.pl /remote\_mnt**

c. Monitor Migration Status from Celerra: **#nas\_fs -i mgfs1** [From Client: #df -k]

**\$server\_mount server\_3 -o verify mgfs1**

Step 10. After Migration, convert MGFS File System to "UXFS":

**\$server\_mount server\_3 -F -o convert mgfs1 /mgfs1**

Step 11. Permanently unmount "mgfs1" from Server\_3, rename if desired, and remount on Server\_3

**SERVER LOG ENTRIES DURING MIGRATION:**

MGFS: 4: in-processing: fsid=23 40% done

**MONITORING CIFS MIGRATION:**

**\$nas\_fs -i mgfs1**

**\$server\_mount server\_3 -o verify mgfs1**

[Monitor completion of migration via ‘Used Kbytes’ value]

```
$server_mount server_3 -o convert mgfs1      [Converting MGFS1 to UxFS]
$server_config server_3 "mgfs getOnline=/mgfsMountPoint/parentDir1/parentDir2/offlineDir"
```

## **TROUBLESHOOTING CDMS:**

### **NFS Data Migration Summary: Auspex machine to a Celerra**

Step 1. Run disk\_usage.awk & verify that no existing user links remain

Step 2. Copy data from one File System to another and then compare results [using what copy command/tools?]

### **NFS Post Migration Process:**

Step 1. Mount Target & Destination FileSystems from Client [/source /target]

Step 2. #cd source #(date ; find . -print -depth | cpio -pdmv /target/. ; date) > /tmp/move/log.original

Step 3. #cd destination #find . -print -depth > /tmp/move/log.target

Step 4. Vi Original Log: #vi /tmp/log.source

Step 5. Check File System Consistency: #diff /tmp/move/log.source /tmp/move/log.target

Step 6. #df -g /source /target

Step 7. Run disk\_usage.awk on new Celerra to confirm Files & Directories identical

## **KNOWN CDMS MIGRATION ISSUE:**

Primus emc65794 indicates potential for file corruption during migration if >than 4GB in size. Contact TS2 for patch.

## **USING CELERRA EMCOPY MIGRATION TOOLS:**

**Note:** Tools are not publicly posted and are available from the Apps & Tools CD

### **Migrating From NT Server to Celerra:**

**Introduction:** Main objective is to migrate "data" and "NT Attributes" from an existing NT Server to the CIFS Server, while retaining as many original ACL and ACE Directory/File Entries, including file "ownership" permissions, as possible.

**Note:** EMCopy does not transfer files like a 'backup'--that is, as a data "stream", though recent versions of EMCopy are supposed to be equivalent to a 'robocopy' in speed. Therefore, there is a lot of overhead as each file has a Start & Received signal associated with it--up to 20ms delay for each file copied. A better solution for large transfers might be to use a full Tape Backup, and Incrementals, to do the bulk of the migration & then run EMCopy to cleanup. Unlike robocopy, Emcopy does not list and sort directory files first before copying, so there is a slight risk that any Long Filenames that are named with DOS 8.3 names (filename~1.txt) could be overwritten by a file name with an 8.3 name that is the same.

**Caution:** In order to properly copy over NT User Rights, Ownership, & Permissions, the Celerra must have complete UID/GID-mapping information, or Usrmapper running, in order to copy over files & directories completely.

## **BASIC EMCOPY METHODOLOGY:**

### **Source Windows System to Target Celerra CIFS Server**

- I. Select User Account for EMCOPY migration with Domain Admin Privileges
- II. Assign User to 'Localgroups' called 'Administrators' on SOURCE & TARGET Servers
- I. Assign User Rights to this User on both the SOURCE & TARGET Servers
- II. Use LGDUP.EXE to Mirror the "Localgroups" attributes from SOURCE to TARGET
- III. Use EMCOPY.EXE to Mirror both the "Data" & "NT Attributes" from SOURCE to TARGET
- IV. Use SHAREDUP.EXE to replicate NT Shares/Share ACL's from SOURCE to TARGET

### **I. USE A USER ACCOUNT WITH DOMAIN ADMIN PRIVILEGES FOR EMCOPY:**

**Note:** To run any EMCopy Tools, you must select a User Account to be used to conduct the "migration" that has Domain Admins privileges. In the following example, User "newby" is a member of Domain Admins.

### **II. ASSIGN USER ACCOUNT TO LOCAL GROUPS "ADMINISTRATORS" ON SOURCE NT & TARGET CELERRA SERVERS:**

- a.) As Administrator, open "User Manager for Domains" >User>Select Domain> \\SOURCENTSERVER
- b.) Doubleclick on "localgroups" called "Administrators" and add User "newby"
- c.) Repeat step (a) for Celerra: "User Manager for Domains" >User>Select Domain> \\TARGETCELERRA
- d.) Doubleclick on "localgroups" called "Administrators" and add User "newby"

### **Note: If multiple NT Domains are used, add following Global Groups from each Domain:**

"Domain Admins" from each accessing Domain to Celerra's Local Group—"Administrators"

"Domain Users" from each accessing Domain to Celerra's Local Group—"Users"

### **III. ASSIGN USER RIGHTS TO USER ACCOUNT ON SOURCE & TARGET SERVERS:**

#### **SETTING USER RIGHTS FOR NT SERVER & CIFS NT 4.0 DM:**

a.) As Administrator, open "User Manager for Domains" >User>Select Domain> [\\\$SOURCENSERVER](#)

b.) Select Policies>User Rights>Show Advanced User Rights>

c.) For each of the following "rights", highlight the "right" and then "Add" the User "newby":

**Summary of User Rights Required:**

"[Backup Files & Directories](#)"

"[Generate Security Audits](#)"

"[Manage Auditing & Security Logs](#)"

"[Restore Files & Directories](#)"

d.) Repeat Steps (a) thru (c) on [\\\$TARGETCELLERRA](#)

e.) Stop and Start the CIFS Service on the Target Celerra Server [if allowed to do so]

f.) Log Off your NT System and then Log On as the User "newby"

**SETTING USER RIGHTS FOR CIFS 2000 TARGET DM:**

a.) Grant following privileges or rights to the "migration\_user" account on the Data Mover using Celerra Management snapin:

b.) Start>Programs>Administrative Tools>Celerra Management>EMC Celerra Management>Data Mover Management>

c.) Rightclick and select 'Connect to Data Mover'

d.) Add the appropriate compname from the dropdown box

e.) Expand 'Data Mover Security Settings'>User Rights Assignment: Add following rights to "migration\_user"

[→Backup Files & Directories](#)

[→Restore Files & Directories](#)

[→Generate Security Audits](#)

[→Manage Auditing and Security Log](#)

**SETTING RIGHTS ON WINDOWS 2000 SOURCE SERVER:**

Start>Programs>Administrative Tools>Domain Controller Security Policy

Expand "Security Settings">Local Policies>User Rights Assignment

"[Backup Files and directories](#)" >Add>User and group names>select User>Add>o.k., o.k. & repeat for other rights

"[Generate Security Audits](#)"

"[Manage Auditing & Security Logs](#)"

"[Restore Files & Directories](#)"

**IV. RUN LGDUP.EXE: Current Version 1.0.10 (NAS 5.4.19.4)**

**PURPOSE:**

Replicates "Local Groups" database information from SOURCE to TARGET SERVERS. Also replicates User and Group privileges to DataMover. Without this, DataMover would not be able to access Domain-based local Groups.

**PREREQUISITES:**

--Member of either 'Administrators' or 'Account Operators' Local Groups on Source & Target Servers

--Member must have "Generate Security Audits" & "Manage Auditing & Security Log" Privileges within the NT Domain(s)

**LGDUP LIMITATIONS:**

--LGDUP does not migrate "Unknown Accounts" or Local User Accounts such as "local" Administrator account.

**LGDUP SWITCHES:**

**Note:** By default, LGDUP will merge the SOURCE to TARGET localgroups database

**-r** --Use the "-r" switch if you would rather "replace" the TARGET localgroups database with the SOURCE

**Caution:** Versions prior to 1.0.6 "erased" the localgroups database on TARGET!!

**-p** --Use "-p" switch if you are copying localgroups databases from multiple SOURCE servers & want to maintain each SOURCE Server's privileges on the TARGET—adds SOURCE Server name prefix to TARGET database

**-v** --Use "-v" switch for Verbose logging details and redirect output into a text file

**-s** --Use "-s" switch if you do not wish to set or add any member of localgroups on first resolve error—LGDUP will stop

**-l** --Use "-s" switch to redirect output to a new log file

**-l + logfile.name** --Use this switch to append output to existing log file

**-nopriv** --Use "nopriv" switch if you do not wish to duplicate local groups privilege settings

**-u** --Allows for migration of local users (5.4 code)

**Note:** Encoding status of 0 means LGDUP operation was successful

**EXAMPLE OF LGDUP COMMAND PROMPT SYNTAX:**

**FROM SOURCE NT SERVER → TO TARGET CELERRA**

**C:\emctools>lgdup.exe -v \\ntwk1dom1 \\suncor > suncorlgdup.txt**

**C:>lgdup.exe -r -p -v \\source \\target**

**TROUBLESHOOTING:**

--Redirect LGDUP output to file. Check for errors after running to verify results. If following error obtained, User probably does not have appropriate Administrative membership or User Rights to conduct the "migration"

**Error: you do not have the "Generate security audits" privilege on this server.**

*Please set the required privilege(s).*

*You must logon again to get the new privilege(s).*

## **V. RUN EMCOPY.EXE: Current Version 03.01 (NAS 5.6.38.2)**

### **PURPOSE:**

Duplicates Directory Tree from SOURCE to TARGET Servers, preserving Windows NTFS/Ownership ACL's & ACE's on TARGET. Copies both "Data", Windows Ownership/DACL?SACL information, and time stamps.

### **EMCOPY SLOWER THAN ROBOCOPY?**

Robocopy and Emcopy are similar tools used to copy data and NT attributes from a Source to a Target Server. Robocopy, however, caches the target directory & filesystem handle information in a local cache file while performing the copy operation, whereas EMCopy has to open each file and do a GetFileInformationByHandle request across the network to the target side. Next version of EMCopy will cache directory and file attributes so that it more closely emulates Robocopy behavior and will therefore be much faster.

### **PREREQUISITES:**

- Member must have "Backup files and directories" on Source and "Restore files and directories" privileges on Source/Target
- Member of either the 'Administrators' or 'Account Operators' Local Group on Source & Target

**Note:** Must use the Celerra Data Mover Security Management Console for Windows 2000, when setting Celerra User Rights

--Must run "LGDUP.EXE" first or Security attributes will not copy over correctly

--Use following params if Invalid or Orphaned SIDs exist on the Source Server

**param cifs acl.mappingErrorAction=3** [Save ACE & SID if resolved during 1<sup>st</sup> lookup, but don't try again if fails]

**param cifs acl.retryAuthSid=600**

**param cifs acl.FailOnSDRestoreError=0**

### **EMCOPY SWITCHES:**

**/nosec** Disables the copying of NT Security properties for Files/Directories [Takes priority over /o /a /lg /i switches]

**Note:** By default, DACL's are replicated on TARGET—this switch disables copying of DACLs

**/o** Copies ownership information of files & directories--otherwise defaults to account used in Migration, when original owner SID cannot be copied over

**/a** Enables copying of auditing info but "Manage Auditing" privilege must be granted to User account doing copy (SACL)

**Note:** User account requires "Manage audits and security log" Privilege

**/secfix** Forces update to NT security on destination folders & files when they already exist (i.e., updates security)

**Note:** If switch not used, NT Security will only be copied to newly created files & folders during the copy operation

**/lg** Copies local groups security entries—[Use LGDUP first]—local group entries are ignored without this switch

**/lu** Copy local user security entries—ignored without this switch [Added in version 2.17]

**Note:** If SOURCE & TARGET localgroups are not identical, then EMCOPY will stop

**/i** Ignores security entries with local users & suppresses errors

**/create** Create 0 file length rather than copy data

**/s** Copies all subdirectories and files in the tree

**/lev:n** Sets depth of copy operation for scanned subdirectories

n=1 Copies only files in specified source directory

n=2 Copies first level of subdirectories with files

n=3 Copies 2<sup>nd</sup> level of subdirectories with files, [n=4, etc.]

**/d** Copies only those files where LAST modification timestamp is newer than those on the TARGET

[Use this switch if you were running the EMCopy a 2<sup>nd</sup> time, like an incremental copy]

**/sd** Preserves security of Source files--does not copy file if there are Security Descriptor errors, erases Target files with errors

**/l** Evaluates the number of files to copy without conducting any copying [Use with /secfix switch to test security]

**Note:** When used with /secfix switch, will compare NT Security on Source & Target files

**/c** Allows process to continue after retries, otherwise will stop

**/z** Restarts from point of failure

**/r:n** Sets the number of retries that EMCopy will resume after encountering failures—default=100 retries

**/w:n** Set the number of seconds between retries—default=30 secs

**/log:d:\logpath** Creates new file & redirects console messages to the file

**/log+:d:\logfile** Appends new messages to existing logfile, as in incremental migrations

**/purge** Removes files & directories from Destination tree which no longer exist on the Source

**Note:** The 'purge' switch in 2.0.9 removes all contents at Target, acting not as an incremental but a full copy again

**/stream** Switch added version 2.22 to copy ADS (alternate data streams) for files and directories—not copied if switch not used

**/q** Disables file printing as standard output—added version 2.05

**/f** Prints path of SOURCE & DESTINATION files [default is to print only Source Path—version 2.05]

**/u** Switch that forces logfile to be written in Unicode characters (version 2.30)

**/nocase** Creates all files & directories in lower case—added to emcopy 2.06

**/nd:domain\_name** Allows translation of SD using new domain name instead of original domain name

**/preserveSIDh** Preserves SIDs from previous domain [Default is to replace obsolete SIDs with current domain SIDs—version 2.11]

**/xd dir <dir...>** Exclude directories with specified names, paths, or wildcards [Added emcopy 2.16]

**/xf file <file...>** Exclude files with specified names, paths, or wildcards

**/backupsd** Backup security properties without copying file contents—extra files are removed at TARGET as in /purge switch

**Note:** Takes precedence over /nosec; /o; /lg; /i; /create; /purge; /l; /nocase; /d switches

**/restoresd** Restores security properties without copying data

**Note:** When using either backupsd or restoresd switches, must have param set on DM if you wish to have security information restored even if errors are encountered. The /restoresd & /backupsd was introduced with EMCopy 2.07 to allow for the backing up of all Security Descriptors from a Source tree without copying any data. EMCopy creates 0 length files and duplicates the SD without localgroup SID translations and local SID suppression. Restore switch allows SD's (Security Descriptors) to be restored to a destination tree.

**/preserverSIDh** Forces preservation of historical SID info during SD translation

**param cifs acl.FailOnSDRestoreError=0**

## EMCOPY SYNTAX EXAMPLE:

**FROM SOURCE NT SERVER → TO TARGET CELERRA**

**C:\migration>emcopy.exe c:\migration \\suncor\users\ /o /a /lg /s /secfix /d > emcopy.txt**

**Note:** Above command copies the contents of "c:\migration" from the SOURCE NT Server "NTWK1DOM1" to the TARGET CELERRA Server "SUNCOR". File Ownership and ACL properties are mirrored from the Source to the Target.

## EMCOPY SECURITY DESCRIPTOR BACKUPS[ACL Backups]:

### SAMPLE COMMAND SET:

**C:\emcopy x:\ c:\acl /s /lg /o /backupsd**

**Note:** Above command backs up Security attributes on a share called "dave\_test", as represented by the drive mapping letter "x" from the Windows system, and backs up to local "c:\acl" folder on the Windows system. Look at Readme.txt file for other syntaxes available and logging options for debugging.

**C:\emcopy c:\acl x:\ /s /lg /o /secfix /restoresd**

**Note:** Above command will restore Security attributes to the mapped "x" drive for a share called "dave\_test". Also note that in order to run EMCOPY, you will need to configure the following (4) User rights on both the Windows system (Computer Management) and the Celerra (Celerra CIFS Management tool>Security>User Rights Assignment), for the Domain Admin User account being used. You may also need to add this User account explicitly to the Compname's localgroup called Administrators on both systems.

"Backup Files & Directories"

"Generate Security Audits"

"Manage Auditing & Security Logs"

"Restore Files & Directories"

## TROUBLESHOOTING EMCOPY:

### Error When Local Groups Not Migrated Properly or Wrong EMCopy Version Used:

\hsbc\hsbcshare\: Security Descriptor error: Unable to set SD: Error 1337: The security ID structure is invalid.

An ERROR occurred on \hsbc\hsbcshare\

### Error (1359): An internal error occurred

→Above error occurs with Win2k3 migrations where user account names contain illegal characters:

### ILLEGAL WINDOWS USERNAME CHARACTERS:

/ \ “ [ ] : | < > + = ; ? \* , [Nor can names end with period]

**Server Log:** lsa.acct ==\UNIX GID=0x1 (Unsupported character in account name)

User names limited to 20 characters while Group names limited to 256 characters

**Fix:** Use emcopy switch /preservesidh option

### Encoding Status Error Codes:

|   |                                                                                   |
|---|-----------------------------------------------------------------------------------|
| 0 | No error                                                                          |
| 1 | Mismatch error—directory not created because file with same name exists on TARGET |
| 2 | Security error—one or more SD's not applied                                       |
| 4 | Data Copy error—one or more files not duplicated                                  |
| 8 | Critical error—no files or directories copied                                     |

**SHARE LEVEL SECURITY:** Must use /nosec option for Share-Level Security on Celerra

## **VI. SHAREDUP.EXE: Current Version 01.09 (NAS 5.4.14)**

### **PURPOSE:**

Replicating NT “Shares” & “Share ACL’s” from SOURCE to TARGET; replicates only data shares, not admin shares

### **PREREQUISITES:**

--Member of Administrators Local Group on Source & Target [No special privileges or User Rights required]

--Target folder path & directories must already exist

### **SHAREDUP.EXE SWITCHES:**

**ALL** New ‘ALL’ value allows all Source Shares to be migrated over to the Target

**/P** Target Server rootpath, followed by mountpoint to CIFS file system [physical path]: \mnt39 represents File System Mount Point—Unix FileSystem name is “FS39”, Shared as “Suncor”, and Mountpoint is “/mnt39”]

**/r** Replaces target shares if they already exist

**/sd** Duplicates Security Descriptors on all Shares. All "localgroups" on SOURCE must match TARGET Celerra--this is why "LGDUP" is run first when conducting an EMCOPY or SHAREDUP migration

**/lu** Enables copying of local user security entries, else local user info is ignored

**/nd:domain\_name** SIDs translated to new domain based on user/group name from source

**/PREFIX** Adds prefix to shares on Target Server with Source Server’s name

**/FO:shareoutput** Creates list of shares to duplicate on Target Server

**/FO+:shareoutput** Concatenates several drive letters for SOURCE server to TARGET server

**/FI:inputfile** Uses specified file as list of Shares to create on Target

**/LOG:path** Creates log file

**/LOG:+path** Appends to log file

### **EXAMPLE OF SHAREDUP COMMAND PROMPT SYNTAX:**

**FROM SOURCE NT SERVER → TO TARGET CELERRA**

**C:\emctools>sharedup.exe \\ntwk1dom1 \\suncor c: /P /mnt39 /PREFIX /r /sd**

### **Exit Codes:**

- |   |                                            |
|---|--------------------------------------------|
| 0 | No error                                   |
| 1 | Syntax error                               |
| 2 | Server name error or Server Not Responding |
| 3 | Duplication error                          |

### **ALTERNATE NTRESKIT TOOL: PERMCOPY.EXE**

## **VII. EMCACL: Current Version 01.08 (NAS 5.4.14.3)**

### **PURPOSE:**

Used to display or edit ACLs or to change Ownership of files or directories

### **PREREQUISITES:**

--Assign account used to run EMCACL the Backup Files and Directories and Restore Files and Directories privilege

--Member of Domain Users

--Enclose User/Group Account names in Quotations if there are spaces within the names [“Domain Admins\account”]

--Can use wildcards to specify more than one file or directory

--Can combine Access Rights

### **EMCACL.EXE SWITCHES:**

**/T** When used by itself, will scan a directory’s ACLs and display without changing [c:\>emcacl stuff /T]

**Caution:** If used with /O; /G; /R; /P; /D switches, will change ACLs

**/E** Used to Edit an existing ACL without replacing—adds ACEs to the ACL list without deleting

**/C** Continue if errors encountered

**/O** Changes owner to specified User

**/G user;perm F; T spec P** Grants specified Access Rights to Users

**Note:** “perm” applies to files—if “spec” is not included, will also apply to directories; “spec” option will apply only to directories [Values for “spec” & “perm” are the same]

R=Read (equiv. E+X); C=Change (write); F=FC (R+C+P+O+D)

P=Change Perms; O=Take Ownership; X=Execute; E=Read; W=Write; D=Delete (all are Special Access rights)

T Flag will apply if used before the “spec” option. If T used after “;T”, then ACE inheritance is disabled. If T Flag is present, new files will not inherit User privileges of directory itself. If T flag is omitted, new files will inherit ACEs of parent Directory.

**/R** Revokes specified User’s right

**/P user:perm F; T spec** Replaces specified User's access rights—perm is same as /G

**/D user** Denies specified User access rights

**/Y** Replaces User's Access rights without confirmation

**/Q** Prints only Errors & Summary

**EXAMPLE OF EMCACL COMMAND PROMPT SYNTAX:**

**C:\>emcacl.exe \\wkdom4\stuff\\*.\* /T**

**Note:** Will display ACL's for mountpoint

**PRINTING ACLs FOR GIVEN PATH (Not the Sharename):**

**C:\>emcacl.exe \\wkdom4\stuff\\*.\* /T**

**PRINTING ACLs for GIVEN MAPPED DRIVE :**

**C:\emcacl>emcacl.exe z:\\*.\* /T**

**CHANGING OWNERSHIP RECURSIVELY TO ADMINISTRATOR :**

**C:\emcacl>emcacl.exe z:\emcacl /E /O administrator /G administrator:F /Y**

**Exit Codes:**

0 No error

1 Syntax Error

2 Given path not found

3 ACL change failure

OI: Object Inherit ACE (file creation)

CI: Container Inherit ACE (directory creation)

NP: No Propagate Inherit ACE

IO: Inherit Only ACE (used only for inheritance, not access checking)

IA: Inherited ACE

**FSTOOLBOX.EXE:** Version 01.02 NAS 5.5

Tool to remove or change ownership on files or directories for a specific owner, as well as managing Quotas; needs to be run as local administrator—specify User name or SID

**C:>fstoolbox.exe z: info |list |enumallfiles |removeallfiles |removequota |removefilesandquota |changeowner |moveusertree**

Info→Print quota setup of a file system

List→lists all user quotas or specific User if specified

Enumallfiles→list all files/dirs owned by particular user

Removeallfiles →remove all files owned by Username or SID

Removequota →remove quota entry for particular User

Removefilesandquota →remove files & directories of the User and then removes quota entry

Changeowner →change file/dir owner

Moveusertree →move files owned by User in a tree structure

**EMCABE:**

Used to manage access-based enumeration feature for Shares, version 1.00 NAS 5.5

/e →Enables ABE on specified shares

/d →Disables ABE on specified shares

/g →Gets ABE state of share/shares on target server

/t →Target server name

/a →run cmd against all Shares on Target server

/s →specify name of share to run command

**Note:** Must be run against mapped drive letter

**XCACLS:** Command Line Reskit Tool to view and modify NT Access Control Lists that are accessible from Windows Explorer

/E Edit ACL instead of replacing

/G Grants access to User/Group specified

/P Replaces access rights for User or Group specified

**USING EMCACLS UTILITY TO REPERMISSION COMPLETE FILE SYSTEM:**

1. Start with clean slate by running “cifs update /mnt resetacl” to make all files & folders Everyone FC

2. Run recursive #chown -rv administrators to make Administrator owner of all files & folders

3. Run emcacls command to grant Domain Admins, System, Administrator FC permission

4. Run emcacls again to remove Everyone FC group and to assign other permissions as required

## **CELERRA NETWORKING:**

**Intro:** When troubleshooting Celerra Networking issues, aside from the usual customer variables, key areas will be:

1). Ethernet IP configuration, Broadcast Address, subnet mask , gateway address--for both Control Stations and DataMovers.

**Example:** Celerra will not be accessible from NT in a Windows environment unless a valid Subnet Mask and Broadcast Address are used [for Netbios Name Resolution & communication]

2). Verify both Physical-layer and Application-layer connectivity to Celerra using Ping, FTP, TRACERT, etc.

3). Verify Switch/Router setup, such as Duplexing and Speed settings.

4). Verify Name Resolution, WINS, Hosts files are setup, etc.

5). Verify appropriate Celerra services are configured and running—such as CIFS, etc.

## **ETHERNET FRAMES:**

802.3 Frames are always less than 1500 bytes

### **PHYSICAL LAYER 1:**

Wiring & devices used to connect Hosts to a network to transmit and receive data; ability to detect signaling errors on cable

**MEDIA:** Layer 1 physical medium on wire sends data in Bits

**DATALINK:** Layer 2 Protocol using MAC Addresses to deliver Frames; min. frame size 60 bytes, max 1514 bytes

**IP:** NetworkLayer 3 Protocol using IP Addresses delivering Packets

IP is connectionless datagram service designed to interconnect packet-switched networks [Each packet handled as individual entity]

Due to size of UDP segments, IP must fragment for delivery and reassembles UDP Datagrams on receiving end via ID & Offset

UDP is limited to maximum datagram size of 64k

TTL of IP = 255

**MTU = Max. Transmission Unit** = 1500 bytes for Ethernet [Obtain true MTU setting on Interface from server\_ifconfig -a]

**Note:** Windows clients use PMTU Discovery method (Path Maximum Transmission Unit Discovery)

**MSS = Max. Segment Size** = MTU – IP & TCP Header sizes = 1460 bytes [maximum size of data within a TCP segment as negotiated during SYN—largest segment that TCP will send to other end]. Default T\_maxseg for Celerra is 536 bytes.

**UDP:** Layer 4 User Datagram Protocol that provides connectionless delivery of datagrams using IP. Any retransmissions required are generated at Application Layer and must resend the whole datagram.

**TCP:** Layer 4 Transport Protocol using Ports & Sockets delivering segments in both directions in a virtual full duplexing circuit. TCP uses timers when sending data and waiting for Acknowledgements and uses Receiver Flow Control to maintain the optimal data stream. For missing or dropped packets, TCP performs retransmission of only the missing data.

## **TCP/IP 3-WAY HANDSHAKE—“SESSION SETUP”:**

1. SYN seq=x to Server

2. SYN ACK=x+1 Seq=y from Server

3. ACK Seq=y+1 from Client completes session setup

## **CLOSING TCP SESSIONS:**

1. Client sends FIN seq=x to Server

2. Server ACKs with seq=x+1

3. Server performs ‘half-close’ FIN seq=y

4. Client ACKs with seq=y+1 to perform “complete close”

**Purpose:** TCP is a Byte-stream protocol that requires Byte-Stream synchronization of sequence numbers for each host

## **Options Field:**

Negotiated during SYN

--In event of multiple segments lost in a single Window, SACK allows for Sender to specify packets needed to be retransmitted

--Advertised Receive Window also negotiated here

## **TCP FLAGS:** URG, ACK, PSH, RST, SYN,FIN

**Note:** PSH is the “pushbit” and is set by Sender to indicate that no more data remains in its buffer—requires immediate ACK

SYN—synchronize sequence numbers

FIN—sender is finished sending data

RST—reset the connection when segment arrives that appears incorrect for the socket connection in place

PSH—push data to receiving process immediately

-- --no flag is set

## **LISTEN STATE:**

A TCP endpoint that Server uses when new connection requests are received

## **ESTABLISHED STATE:**

TCP module of kernel creates new end points after accepting an incoming connection requests→from LISTEN to ESTABLISHED

**2MSL WAIT STATE (TIME\_WAIT state):**

Max. segment lifetime (MSL) that any segment can exist on the wire before expiring. Common implementations are that upon a TCP close, the TIME\_WAIT status of a connection must remain open for twice the MSL, while not allowing the socket connection to be reused.

**FIN\_WAIT\_2 STATE:**

Time while waiting for other end to acknowledge our FIN, or half-close.

**ACK:**

Acks are typically triggered for every other full sized frame sent

Acks are also triggered when Receiving end gets a frame out of order

**DEFAULT MTU SETTINGS:**

Gigabit Ethernet: 1500 bytes

Fast Ethernet: 1500 bytes

FDDI: 4500 bytes

ATM: 1500 bytes

**SETTING LARGE PACKET 9000MTU SUPPORT ON GIGE:**

**\$server\_ifconfig server\_2 ace0 lcge0 | fxg0 mtu=9000**

**MAXIMUM TRANSMISSION UNIT—MTU:**

Each media type has a maximum frame size that can be transmitted, but in turn, each NIC Interface may have a different MTU value. Link Layer protocol discovers what the common MTU denominator is for TCP for each TCP session between two hosts. The MTU is negotiated during the 3-Way Handshake in setting up a TCP session between two hosts and is negotiated in the SYN packet. Every “route” used typically will have a unique MTU. Most modern switches & routers support 1500 byte frames without fragmenting.

**FIGURING OUT TRUE MTU SETTING FOR NETWORK DEVICES:**

**# server\_netstat server\_2 -i**

| Name           | Mtu  | Ibytes     | Ierror | Obytes    | Oerror           | PhysAddr         |
|----------------|------|------------|--------|-----------|------------------|------------------|
| mge0           | 1500 | 25621621   | 0      | 139978212 | 0                | 0:60:16:f:df:55  |
| mge1           | 1500 | 1112423355 | 0      | 196511688 | 0                | 0:60:16:f:df:57  |
| ---abridged--- |      |            |        |           |                  |                  |
| cge-3-0        | 9000 | 2784149    | 0      | 975247    | 0                | 0:60:48:1b:58:b2 |
| ---abridged--- |      |            |        |           |                  |                  |
| fxg-1-0        | 9000 | 0          | 0      | 0         | 0:60:16:1d:d5:14 |                  |
| fxg-2-0        | 9000 | 0          | 0      | 0         | 0:60:16:1d:d5:17 |                  |

**Note:** The Server Netstat command displays the default physical NIC MTU values (Ethernet Layer), and most Celerra production NICs default to 9000 MTU support—internal NICs use 1500 MTU.

**# server\_ifconfig server\_2 -a**

```
server_2 :
fxg1 protocol=IP device=fxg-1-0
    inet=192.1.4.3 netmask=255.255.255.0 broadcast=192.1.4.255
        UP, ethernet, mtu=1500, vlan=0, macaddr=0:60:16:1d:d5:14
192-1-4-1 protocol=IP device=cge-3-0
    inet=192.1.4.1 netmask=255.255.255.0 broadcast=192.1.4.255
        UP, ethernet, mtu=1500, vlan=0, macaddr=0:60:48:1b:58:b2
```

**Note:** Server Ifconfig displays the MTU setting at the IP level, which is where we generally control the MTU values. Since we default to 1500 MTU, we manually change the device to 9000 MTU Large Packet support.

**# server\_ifconfig server\_2 fxg1 mtu=9000**

server\_2 : done

**CELLERRA TCP ADVERTISED WINDOW SIZE:**

**Definition:** TCP Window Size is the number of data packets that the Sender is allowed to have outstanding with having received an ack from the Receiver. TCP uses a ‘sliding window’ to adjust this Window Size value dynamically during the Session.

Parameter for Windows Clients that determines size of TCP ‘receive’ window, which is the number of bytes that can be received before the Sender must acknowledge. Sender advertises its “receive” buffer size window with every segment of data sent. Sender cannot send more data once full number of bytes is sent that represents the “Receive Window”.

**Default TCP Window Size for Celerra is 65kbytes, max 256kbytes [“param tcp”=tcp.doRFC1323=1]**

**Note:** This option is related to what is called “Windows Scaling” and allows for the Celerra to supply Clients that request “Receive Window” sizes > than 65kb with the larger data window.

→Advertised Window is flow control imposed by Receiver

## **CONGESTION WINDOW & SLOW START:**

Segments are incrementally injected into the network, up to the Window Size advertised by receiver. However, with more complex networks with multiple routers and slower links, there is a need to have a “slow start” mechanism that calculates how quickly to add segments to the network based on the rate of acknowledgments received. This is achieved by using a “Congestion Window” whereby after each ACK, the cwnd is increased by one segment.

→Congestion Window is flow control imposed by Sender

## **HOW THE TCP ADVERTISED WINDOW WORKS:**

--Sender must wait for ACK once full the number of bytes sent equals the Receive Window & cannot send more until ACK

--Normally, TCP ACKs occur for every other full sized segment sent

--Missing frames also automatically cause ACKs for each subsequent frame until the system is back to a normal state

## **KeepAliveTime:**

Determines how often TCP verifies idle connections to determine whether to maintain or breakdown. Keep Alive is not enabled by default. A KeepAlive packet is sent to host connections and if acknowledged, will keep connection alive.

## **TCP IP STACK (NAS 5.4 Enhancements):**

SACK—Selective Acknowledgement—param tcp do\_sack

FACK—Forward Acknowledgement

NEWRENO—param tcp do\_newreno (New Reno algorithm)

**Note:** 3 Algorithms designed to help recovery from lost TCP segment traffic

## **MAC OS (Apple) McIntosh Performance Issues with NAS 5.4:**

Macs do not handle the delayedACK feature well with CIFS and this can cause bursts in traffic followed by long pauses, and overall performance degradation.

### **Solutions:**

**param tcp.maxburst=0** (turn off maxburst from default value of 4—NAS 5.4.17.0 will have this disabled by default)

**# /sbin/sysctl -w net.inet.tcp.delayed\_ack=0** (turn off delayed ack on Mac Client)—to make change persistent, edit the /etc/sysctl.conf and add the string “net.inet.tcp.delayed\_ack = 0” to the end of the file as a new line.

## **CHANGING DM MTU SETTINGS:**

**#server\_ifconfig server\_2 ana0 mtu=1500**

## **Celerra CIFS Interface Rules:**

--multiple NT Domains can be supported by one DataMover, but is not recommended

--Can have multiple Netbios names for a NIC if IP's or Domains are different

--Multiple NIC Ports can have the same name, however, only if different IP's are used within same Domain

--Different NIC Ports can have different Netbios names on Same or Different Domain with different IP's

--All Shares go out over all Interfaces and Domains by default when exported from CLI, to specific Comname when using MMC

**SUPPORTED CELERRA INTERFACES:** ana0=Ethernet fpa0=FDDI fa20=ATM ace0=Gigabit Ethernet

**Maximum of 32 Interface Configurations per DataMover** [2 for Internal network; 1 for loopback; 29 for IP configurations]

## **INTERNAL IP ADDRESSES FOR LINUX CONTROL STATION:** NAS 2.2.34.5

**emcnas\_i0 192.168.1.100**

**emcnas\_i1 192.168.2.100**

## **STOPPING & STARTING INTERNAL CONTROL STATION INTERFACES:**

**# /sbin/ifconfig down eth0**

**# /sbin/ifconfig up eth0**

## **INTERFACE CONFIGURATIONS:** \$server\_ifconfig server\_2 ana0 | ace0 | fsn0

--Shows IP, Mask, Broadcast, whether device is UP or DOWN, MTU Size, and VLAN info

**Caution:** Interface variable is the “name” of the device in lefthand column of server\_ifconfig -a command, not the logical name!

## **CHECKING INTERFACE SETTINGS/STATS using “.server\_config”:**

**\$server\_config server\_x -v “ace ana0 clearstat”** [Use this to clear the server\_netstat -i Statistics counters]

```
$server_config server_2 -v "ana ana0 stat"
$server_config server_2 -v "ace ace0 stat"
$server_config server_2 -v "fsn fsn0 stat"
$server_config server_x -v "bcm cgeX stat" |showmac |intrstat |timerstat [Stats for Broadcom NIC]
$ .server_config server_2 -v "bcm"
```

Usage: bcm <device> <command> [parameters]

commands and parameters are:

- up | start - Enable interface
- down | stop - Disable interface
- clearstat|clearstats - Clear all statistics counters
- showmac - prints current and native mac addresses
- setnative|mac - set current mac addr to native mac addr
- stat - prints default statistics
- ifstat - prints interface statistics (MIB-II[8,9])
- intrstat - interrupt statistics
- timerstat - Timer statistics

```
# .server_config server_2 -v "bcm cge0 showmac"
```

Current Mac Address -> 8:0:1b:42:48:65

Native Mac Address -> 8:0:1b:42:48:65

1074543292: ADMIN: 4: Command succeeded: bcm cge0 showmac

```
$ .server_config server_X -v "bcm cge0 stats"
$ .server_config server_X -v "bcm cge0 macstat"
$ .server_config server_X -v "bcm cge0 intrstat"
$ .server_config server_X -v "tcp stat"
$ .server_config server_X -v "ip allstat"
$ .server_config server_X -v "trunk trk0 stats"
```

## **TRUNKED OR FSN CONFIGURATIONS PRESENT SINGLE MAC ADDRESS BASED ON ACE0:**

```
# server_ifconfig server_3 -a
```

server\_3 :

```
fsn protocol=IP device=fsn0
    inet=10.73.3.24 netmask=255.255.255.0 broadcast=10.73.3.255
        UP, ethernet, mtu=1500, vlan=0, macaddr=0:60:cf:20:26:71
```

```
# .server_config server_3 -v "ace ace0 showmac"
```

Current Mac Address -> 0:60:cf:20:26:71

1112965793: DRIVERS: 4: ace0: Using Local MAC address 0:60:cf:20:26:71

Native Mac Address -> 0:60:cf:20:26:71

1112965793: ADMIN: 4: Command succeeded: ace ace0 showmac

```
# .server_config server_3 -v "ace ace1 showmac"
```

Current Mac Address -> 0:60:cf:20:26:71

1112965801: DRIVERS: 4: ace1: Using Local MAC address 0:60:cf:21:2e:2b

Native Mac Address -> 0:60:cf:21:2e:2b

```
# server_log server_3 -a -s |grep -i mac
```

2004-06-02 17:07:57: DRIVERS: 4: ace0: Using Local MAC address 0:60:cf:20:26:71

2004-06-02 17:07:57: DRIVERS: 4: ace1: Using Local MAC address 0:60:cf:21:2e:2b

2005-03-21 01:08:30: ADMIN: 4: Command succeeded: ifconfig fsn protocol=IP dev

ice=fsn0 local=10.73.3.24 netmask=255.255.255.0 broadcast=10.73.3.255 mtu=1500 vlan=0 mac=0:60:cf:20:26:71

## **TROUBLESHOOTING PCI DEVICES ON CELERRA:**

```
# .server_config server_2 -v "showpci" |grep -v no
```

PCI DEVICES:

| VendID | DevID | Bus | IRQ | Dev# | Slot# | Func | Class | Driver | Name                   | OR | 2nd Bus |
|--------|-------|-----|-----|------|-------|------|-------|--------|------------------------|----|---------|
| 8086   | 1229  | 0   | 16  | 1    | 0 (0) | 0    | 20000 | yes    | fxp0                   |    |         |
| 15bc   | 100   | 2   | 22  | 4    | 0 (0) | 1    | c0400 | yes    | fcp-0:5006016010600278 |    |         |
| 15bc   | 100   | 2   | 21  | 4    | 0 (0) | 0    | c0400 | yes    | fcp-1:5006016110600278 |    |         |
| 15bc   | 100   | 1   | 18  | 4    | 0 (0) | 1    | c0400 | yes    | fcp-2:5006016210600278 |    |         |
| 15bc   | 100   | 1   | 20  | 6    | 0 (0) | 1    | c0400 | yes    | fcp-3:5006016310600278 |    |         |
| 14e4   | 1645  | 3   | 23  | 4    | 0 (0) | 0    | 20000 | yes    | cge0                   |    |         |
| 14e4   | 1645  | 3   | 25  | 6    | 0 (0) | 0    | 20000 | yes    | cge1                   |    |         |

```
14e4 1645 3 24 5 0 (0) 0 20000 yes cge2
14e4 1645 4 26 4 0 (0) 0 20000 yes cge3
14e4 1645 4 27 5 0 (0) 0 20000 yes cge4
14e4 1645 4 28 6 0 (0) 0 20000 yes cge5
```

### **CLARIION COPPER CGE INTERFACES—BROADCOM ETHERNET CONTROLLER:**

```
$server_config server_2 -v "bcm cge0 stop"
$server_config server_2 -v "bcm cge0 start"
$server_config server_2 -v "bcm cge0 ifstat"
$server_config server_2 -v "bcm cge0 stat"

Gigabit Ether cge0 --> LINK is UP, NIC is STARTED
Speed - 1000 Mbps, Duplex - full
PktForTx TxPktSnt TxThres High TxThresh Low Rx Pkts Rx Dlvr
000004b1 000004b1 0 0001e62b 0001e081
-----Tx Dropped----- -----Rx Dropped-----
RingFull No Descr Not Data Too Big NoProtoC MsgbProt MsgbData
00000000 00000000 00000000 00000000 0000005aa 00000000
Bad CRC Collisic LinkLost PhyDecode OddNibbl MacAbort ShortPkt
00000000 00000000 00000000 00000000 00000000 00000000
NoResrc GiantFrm
00000000 00000000
```

### **GOLDEN EAGLE INTEL CGE INTERFACES:**

```
$server_config server_2 -v "ians cge0 stat"
$server_config server_2 -v "ians cge0 ifstat"
$server_config server_2 -v "ifconfig"
```

#### Devices:

```
fxp0 dmtu=1500, dmac=8:0:1b:43:54:ab
cge0 dmtu=9000, dmac=8:0:1b:42:48:65
loop dmtu=65536, dmac=0:0:0:0:0:0
```

Interfaces:(4)

```
el30 on fxp0 l=192.168.101.2 n=255.255.255.0 b=192.168.101.255 DNIF UP
    mtu=1500, dmtu=1500, vhid=0, mac=8:0:1b:43:54:ab dmac=8:0:1b:43:54:ab
el31 on fxp0 l=192.168.102.2 n=255.255.255.0 b=192.168.102.255 DNIF UP
    mtu=1500, dmtu=1500, vhid=0, mac=8:0:1b:43:54:ab dmac=8:0:1b:43:54:ab
192.168.1.180 on cge0 l=192.168.1.180 n=255.255.255.0 b=192.168.1.255 DNIF UP
    mtu=1500, dmtu=9000, vhid=0, mac=8:0:1b:42:48:65 dmac=8:0:1b:42:48:65
loop on loop l=127.0.0.1 n=255.0.0.0 b=127.255.255.255 UP
    mtu=32768, dmtu=65536, vhid=0, mac=0:0:0:0:0:0 dmac=0:0:0:0:0:0
```

**Note:** This command also is useful for determining whether the Server is negotiating Full, Auto, or Half Duplex, Speed, etc!

**10/100 Interface:** Tells whether Auto/Manual, Speed and Duplex Settings [Half, Full, 10/100Mb/s]

**Gigabit Interface:** Tells you whether interface “LINK is UP, NIC is STARTED”

**Virtual Interface:** Tells you whether Virtual Interface is UP, and which NIC Port is ACTIVE--Also lists NIC's configured.

### **DETERMINING DM SPEED & DUPLEX SETTINGS:**

```
$server_config server_2 -v "ana ana0 stat" [Shows status of link]
$server_log server_2 > slog2 [Negotiation status reflected after a BootUp too]
```

### **SPEED & DUPLEX SETTINGS ON GOLDEN EAGLE:**

```
# .server_config server_2 -v "ians cge0 ifstat"
```

Intel Gigabit Controller

| IfType   | IfMTU  | IfSpeed   | IfOper    |        |
|----------|--------|-----------|-----------|--------|
| ethernet | 9018   | 1 Gb/s    | Up        |        |
| AutoNeg  | Duplex | RXFlowCtl | TXFlowCtl | Media  |
| disabled | full   | disabled  | disabled  | Copper |

**Note:** Above output indicates speed and duplex as FULL DUPLEX.

### **SERVER LOG WILL ALSO SHOW SPEED SETTINGS:**

```
# server_log server_3 -s |grep -i speed
```

```
2005-02-26 14:19:25: DRIVERS: 4: cge4 : Link Status : UP, Speed : 1000, Duplex: FULL
2005-02-26 14:19:25: DRIVERS: 4: cge0 : Link Status : UP, Speed : 1000, Duplex: FULL
```

### **TROUBLESHOOTING COMMANDS:**

```
$server_config server_2 -v "printstats rpc" [RPC Statistics]
```

**\$server\_config server\_2 -v "ace ace0 stat"** [Gigabit Interface]

**\$server\_config server\_2 -v "ace ace0 updown"** [Gigabit Interface]

**\$server\_config server\_2 -v "logsys set severity IP=LOG\_DEBUG"** [Debug cmd for IP]

### **CLEARING STATISTICAL COUNTERS ON INTERFACES FOR IP, TCP, ICMP, UDP:**

**\$server\_config server\_2 -v "printstats tcpcstat reset"**

**\$server\_config server\_2 -v "printstats ipstat reset"**

**\$server\_config server\_2 -v "printstats udpstat reset"**

**\$server\_config server\_2 -v "printstats icmpstat reset"**

### **TUNABLE UDP/TCP TIME-TO-LIVE PARAMETERS:**

Default udp.ttl and tcp.ttl values are 0x40, while MAXTTL value is 255

**\$server\_config server\_x -v "param udp ttl=0x80"**

**\$server\_config server\_x -v "param tcp ttl=0x80"**

**# .server\_config server\_3 -v "param fulldescription udp ttl"** [add'tl syntax required to output param info NAS 5.5]

udp.ttl 0x01b74710 0x00000040 0x00000040

### **PLACING INTERFACE UP or DOWN/DELETING:**

**\$server\_ifconfig server\_4 ana0 up [down|trk0,fsn0]**

**\$server\_ifconfig server\_2 -d ana0**

**Note:** This also is the way to make changes to the Interface effective, rather than rebooting the whole server!

### **FAST ETHERNET SETTINGS:**

1. Setting IP/Mask/Broadcast:

**\$server\_ifconfig server\_2 -c -D ana0 -n ana0 -p IP 193.1.21.172 255.255.255.0 193.1.21.255**

2. Setting Speed and Duplex Settings:

**\$server\_sysconfig server\_2 -pci ana0 -o speed=10|100|auto,duplex=half|full|auto**

**# server\_sysconfig server\_2 -pci cge0 -o speed=auto,duplex=full**

server\_2 : done

**# server\_sysconfig server\_2 -pci cge0**

server\_2 :

On Board:

Broadcom Gigabit Ethernet Controller

0: cge0 IRQ: 23

speed=1000 duplex=full txflowctrl=disable rxflowctrl=disable

### **VERIFY DM SPEED/DUPLEX/LOAD BALANCING SETTINGS (MAC, IP, TCP):**

1. Server Log contains correct speed/duplex condition

**2. \$ server\_ifconfig server\_2 ana3**

server\_2 :

ana3 protocol=IP device=ana3

inet=10.64.25.72 netmask=255.255.255.0 broadcast=10.64.25.255

UP, ethernet, mtu=1500, vlan=0, macaddr=0:0:d1:20:50:4c

**3. \$ server\_sysconfig server\_2 -v -i fsn01**

server\_2 :

\*\*\* FSN fsn01: Link is Up \*\*\*

active=trunk01 standby=trunk01 trunk02

**\$ server\_sysconfig server\_2 -v -i trunk02**

server\_2 :

\*\*\* Trunk trunk02: Link is Up \*\*\*

**\*\*\* Trunk trunk02: Statistical Load Balancing is IP \*\*\***

Device Link Duplex Speed

---

ana4 Up Full 100 Mbs

ana5 Up Full 100 Mbs

**Note:** Output from these commands may be incorrect—see Server Log!

**\$ server\_sysconfig server\_2 -v -i trk001**

server\_2 :

\*\*\* Trunk trk001: Link is Up \*\*\*

\*\*\* Trunk trk001: Statistical Load Balancing is TCP \*\*\*

**4. \$ .server\_config server\_2 -v "ana ana2 stat"**

```
ana2 : AUTO : *** NO LINK *** : IRQ 5 : IOBASE 0x3400
Tx Pkts Tx Bytes TxPcktQu TxByteQu TxPcktEr TxByteEr
00000000 00000000 00000000 00000000 00000000 00000000
```

**5. \$ .server\_config server\_5 -v "trunk trk0 stat"**

\*\*\* Trunk trk0: Link is Up \*\*\*

| Device | InPackets | InErrors | OutPackets | OutErrors | Link | Speed   |
|--------|-----------|----------|------------|-----------|------|---------|
| ana2   | 606744    | 1        | 942759     | 0         | Up   | 100 Mbs |
| ana3   | 1179469   | 1        | 1032832    | 0         | Up   | 100 Mbs |
| ana6   | 942740    | 1        | 672547     | 0         | Up   | 100 Mbs |
| ana7   | 1109979   | 1        | 1718295    | 0         | Up   | 100 Mbs |

1113462577: ADMIN: 4: Command succeeded: trunk trk0 stat

**\$ .server\_config server\_2 -v "param trunk"**

| Name              | Location   | Current | Default |
|-------------------|------------|---------|---------|
| trunk.LoadBalance | 0x01356f20 | 'tcp'   | "       |

**\$ .server\_config server\_3 -v "trunk trk0 showcfg"**

1139001535: DRIVERS: 4: \*\*\* Trunk trk0: Link is Up \*\*\*

1139001535: DRIVERS: 4: \*\*\* Trunk trk0: Timeout is Short \*\*\*

1139001535: DRIVERS: 4: \*\*\* Trunk trk0: Statistical Load Balancing is IP \*\*\*

1139001535: DRIVERS: 4: Device Local Grp Remote Grp Link LACP Duplex Speed

| 1139001535: DRIVERS: 4:      | cge0  | 10000 | 37381 | Up | Up   | Full     | 1000 Mbs |
|------------------------------|-------|-------|-------|----|------|----------|----------|
| 1139001535: DRIVERS: 4: cge1 | 10000 | 37381 | Up    | Up | Full | 1000 Mbs |          |
| 1139001535: DRIVERS: 4: cge3 | 10000 | 37381 | Up    | Up | Full | 1000 Mbs |          |
| 1139001535: DRIVERS: 4: cge4 | 10000 | 37381 | Up    | Up | Full | 1000 Mbs |          |

**FAST ETHERNET QUAD CARD PERFORMANCE CHARACTERISTICS:** Adaptec Quad FE for 507 DM**Theoretical maximum output is 100Mbps=12.5MB/second****Actual Performance:** READ TRANSFER RATES = 7MB/second [Adaptec Hardware limitation]

WRITE TRANSFER RATES = 11.7MB/second

**Typical Performance:** 5MB/sec**CASE STUDY: COPYING DATA FROM ONE DIR TO ANOTHER ON SAME FILESYSTEM:**

**Situation:** Slow performance while performing copy of data from one subdirectory to another on the same Celerra file system. Customer calculates throughput of 4.375MB/sec while copying 500GB from one part of the filesystem tree to another when using Windows 2000 Client Explorer interface [rightclick copy, rightclick paste]. Network Traces show that DM SRT [Server Response Time] is very good--1.5ms, while Windows 2000 Client CDT [Packet Data Unit] and CNRT [Client time to next Read request] times of 5.1ms & 6ms, respectively, were poor.

In actuality, throughput operation on the Data Mover had a combined "read" and "write" rate of 8MB/sec. The limiting factor in this scenario was the Windows 2000 client, which [as designed] conducts non-overlapping synchronous 64k Reads from and Writes to, the DM. The overall effect is somewhat analogous to half-duplexing and limits the theoretical throughput to a maximum of 12.5MB/sec. If the Windows 2000 client used overlapping asynchronous IO operations, then the combined theoretical maximum rate would be closer to 25MB/sec.

The combined throughput of 8MB/sec, out of a total theoretical maximum of 12.5MB/sec, reflects a healthy and normal copy operation rate on the Data Mover. While using a more powerful Windows 2000 Client would increase the actual overall read/write operation throughput, a better solution for the customer would be to perform a "cut & paste" operation, which does not have to copy actual data, just changes the metadata pointers to reflect the change in directory location of the data.

**OPTIMAL ETHERNET SETTING FOR CELERRA FE:**

Hard-code the Link Partner settings to 100 Full for Celerra Interface and Switch Port.

**FAST ETHERNET SPECS:**

**100BASE-T** →Original Fast Ethernet spec 100Mbits/sec based on old Ethernet standard (IEEE 802.3 ) using CSMA/CD (Carrier Sense Multiple Access/Collision Detection), IEEE 802.3u is the updated spec for 100Mbits/s

**100BASE-TX** →2 prs of cat5 wire, performs at 100Mbits/s half-duplex mode or 200Mbits/s Full Duplex, distance 100 Meters 4 bits/transmission sent at 25MHz clock speed to yield 100Mbits/s

**100BASE-T4** →4 prs twisted pair wire, purpose to link two 100 meter segments together on the LAN

**100BASE-FX** →100Mbits/s Fast Ethernet over Fibre Optic cables half-duplex (400meters) or Full Duplex mode (2km), two multi-mode optical fiber strands for receive and transmit

**Changing Ethernet Max Received Packets-Processed:** Change when more packet processing per interrupt is required

**param ana rxburst=512** [Default value=hex 100, decimal 256]

## **FAST ETHERNET PORT LIGHTS:**

Righthand light for each ana port: Green Light indicates network link and running at FULL DUPLEX 100TX SPEED  
Amber Light indicates network link but running at HALF DUPLEX 10BASE-T SPEED

Lefthand light for each ana port: Amber light indicates network traffic [port is passing data]

## **DM ETHERNET INITIALIZATION:**

When links are being brought up, link pulses are suspended to ensure that links are severed.

When link pulses are restarted, Switch senses and initiates link from its end.

DM then broadcasts ARP request for itself in order to update MAC forwarding tables of Switches

## **DISABLING/ENABLING RIP PROTOCOL ON DM:**

**\$server\_rip server\_2 noripin ana0 | ripin ana0**

**\$server\_rip server\_2** [Displays RIP Status]

**Note:** Introduced with NAS 4.0.12.1

## **USING TRACEROUTE ON CS:**

**# /usr/sbin/traceroute**

## **USING TRACERT ON DATA MOVER (NAS 5.6):**

**\$ .server\_config server\_2 -v "tracert ip=128.222.121.1"**

\$ server\_log server\_2 -a -s ltail

HOP:1, DST:128.222.121.1, Router:10.8.126.1

HOP:2, DST:128.222.121.1, Router:10.5.191.202

HOP:3, DST:128.222.121.1, Router:57.225.20.57

HOP:4, DST:128.222.121.1, Router:57.215.105.188

HOP:5, DST:128.222.121.1, Router:57.215.105.182

HOP:6, DST:128.222.121.1, Router:10.6.193.233

## **SWITCH BASICS:**

--Celerra should work with all switches [FE or GbE] that are IEEE compliant

--Switches should guarantee that packets from single TCP connection always go through same link in each direction if using Celerra Trunking [link aggregation]

--Celerra understands ARP, RIP, 802.3ad, 802.1q

--Each port on a switch is its own Collision Domain [Dedicated vs. Shared Network that Hubs offer]

--Switches operate in Full Duplex mode

--switches are multiport ‘bridges’ that make use of ‘learn & forward’ tables

## **SPEED & DUPLEX MISMATCHES: [Settings of Host and Switch Port are key]**

### **DUPLEX MISMATCH SYMPTOMS:**

--On side with Full Duplex set, may see Runt, Alignment errors, CRC errors, but no collisions

--On half-duplex side will see High Collisions & Errors

**Note:** Netstat command can help in determining collisions. Most important concept in Duplex mismatches is that the “link partners” must be set the same [Host interface to Switch interface]. Settings between Connection Partners is not as meaningful.

**HUBS:** Hubs always check for Collision Detection since by definition they are always half-duplex. Hubs can never negotiate with a switch—it does not have the intelligence built-in. Since Hubs cannot autonegotiate with the switch, if the Switch cannot set itself to half duplex by default, then the link may never come up. Newer Cisco switches tend to default to 100 Full.

**Note:** You can still change the switch setting manually to half-duplex.

### **LINK PARTNERS:**

Physical NIC connection between Host and Switch—this is where Duplexing matters, between the Host and the Switch, not between one Host and another Host on the other side of a switch. Link pulses, or link light comes on when each Interface is linking together.

### **CONNECTION PARTNERS:**

Logical connection between two Host systems on different networks [more along TCP Socket connection]

### **BRIDGES:**

- Bridges bring two different Broadcast Domains together
- Uses table of MAC Addresses
- bridges forward Broadcasts to all ports if it hasn't learned the specific MAC Address [Learn & Forward]
- Spanning Tree attempts to detect Broadcast loops

### **Route Table:** Celerra uses "rip" protocol by default

Temporarily Flushing Route Table: \$server\_route server\_2 -f [good until next reboot]

Permanently Flushing Route Table: \$server\_route server\_2 -DeleteAll

Removing an Entry from Table: \$server\_route server\_2 -d guppy 192.10.2.34

**Note:** server\_route -list can only output 64 lines. Use “amgr showroute” to output all entries.

### **CELERRA GIGABIT ETHERNET:** Stnd SC Cables: FDDI 62.5 micron or FC-AL 50 micron

**Alteon 1000Base-SX Gigabit Ethernet Rules:** (2) ports/Datamover. + (1) Quad Eth. Or FDDI card. Jumbo Frames, mtu=9000bytes 505DM+

1. **\$ server\_ifconfig server\_4 -c -D ace0 -n ace0 -p IP 193.1.21.122 255.255.255.0 193.1.21.255**
2. **# server\_sysconfig server\_2 -pci ace0 -o linkneg=disable** [disables autonegotiation]

### **COPPER ETHERNET vs. FIBER ETHERNET:**

Copper Ethernet usually configures Port Negotiation for Catalyst Cisco Switches as part of the port speed settings:

speed:auto | duplex:auto

speed:1000 | duplex:full

For Cisco versions of CatOS, Fiber Ethernet (1000BaseX) still uses Speed & Duplex settings, but Port Negotiation is handled via separate protocol and configuration. If Port or Gigabit Negotiation is Enabled (#show port negotiation), Celerra ACE ports will not work unless Linkneg is Disabled.

For Cisco versions of Native IOS or CatIOS, Port Negotiation would be enabled or disabled when configuring port speeds. If speed is hard-coded, i.e. Full, then Port Negotiation is Disabled. Again, Celerra ACE ports would not work correctly unless Linkneg were Disabled. If port speed is Auto, Port Negotiation is Enabled, and Celerra default is also Linkneg Enabled.

**# show port**

**# show port negotiation**

**# set port negotiation**

**# show tech**

**# show etherchannel load-balance** [Port load balancing scheme in use]

### **CELERRA GbE SETTINGS & LINK NEGOTIATION:**

**Note:** Following case example is based on Alteon 1000Base-SX Interface on Celerra, and the Cisco 6509 GbE Switch. Newer GbE switches that use fibre ports vs. copper will have different criteria.

#### **CISCO 6509 & ALTEON 1000Base-SX NIC:**

Speed settings are automatic and are set to Full Duplex by Default. In some cases, you may need to change Celerra “linkneg”, or autonegotiation, to “disable” and set the Switch accordingly, per the following example:

Step 1. Disable Gigabit AutoNegotiation: **\$ server\_sysconfig server\_2 -pci ace0 -o linkneg=disable**

Step 2. Catalyst 6509 GB Switch:

- a. Set port flowcontrol 9/3 send off
- b. Set port negotiation 9/3 disable
- c. Set spantree portfast 9/3 enable

**Explanation:** Problem can be that the NIC & Switch will have trouble negotiating the link because of conflicting flowcontrol settings. Setting the flowcontrol on the switch to off, then disabling negotiation on both sides, “hard-codes” a mutually acceptable link configuration on each side. Also, use the “portfast” command on Switch if the server is an end node, as it is desirable to bypass the

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
full spanning-tree calculations on ports connected to end nodes. DO NOT use this command on ports connected to other network devices, because that can create a bridging loop, which can degrade an Ethernet network.

## **CELERRA & GIGABIT SWITCH RULES:**

Always Full Duplex with effective bandwidth of 2GBPS on each Switch port

Port Negotiation—all Ports on switch enabled by Default

Duplex-----Full Duplex by default [CANNOT be changed for Gigabit] {two stations Tx and Rx simultaneously}

Flow Control-----Set to Off for Rx and Desired for Tx

Jumbo Frames-----Disabled by default

### **Enabling Jumbo Frames on a Port:**

#set port jumbo 2/1 enable [disable]

#show port jumbo

**Switch Management:** Telnet or SNMP or SSH using the In-Band sc0 Interface or Out-of-Band s10 Interface

In-Band sc0 Interface: Connected to switch fabric and used with STP, CDP, VLAN's—tool normally used

    Switch IP Routing Table: Used to forward traffic originating on the switch, not for forwarding traffic sent by devices  
    Connected to the switch

Out-of-Band s10 Interface [Serial Line Internet Protocol—SLIP]: Not connected to switch fabric

### **Setting Switch VLAN, IP address, and MASK:**

#set interface sc0 5 172.20.52.124/255.255.255.248

#set interface sc0 up [down]

**Note:** IP Address and Subnet Mask required. Broadcast address optional. Gateway needed for inter-networking.

## **TROUBLESHOOTING GB SWITCH PORTS:**

|                                                     |                                                            |
|-----------------------------------------------------|------------------------------------------------------------|
| #ping -s 12.20.5.3 800 5                            | [ping host with packet size of 800 bytes with (5) packets] |
| #traceroute 12.20.5.3                               | [Layer 3 IP Trace between network Hosts]                   |
| #l2trace 00-01-22-33-44-55 10-22-33-44-55-66 detail | [Layer 2 MAC address trace]                                |
| #show interface                                     | [Shows interface details]                                  |
| #show ip route                                      | [IP route table]                                           |
| #show module                                        | [Shows Switch Ports and Blades]                            |
| #show port 2/1                                      | [Port(s) Status, such as Speed and Duplex settings]        |
| #show port capabilities 2/1                         | [Port settings, etc]                                       |
| #show port negotiation                              | [Gigabit or Port Negotiation for switch ports]             |
| #set port negotiation                               |                                                            |
| #show tech                                          | [Useful output on switch specs]                            |
| #show interfaces counters errors                    | #show interfaces capabilities                              |
| #show interfaces counters etherchannel              | #show etherchannel load-balance                            |
| #show interfaces flowcontrol                        | #show interfaces switchport                                |
| #show interfaces trunk                              | #show mac-address-table                                    |

## **BASIC INFORMATION WHEN WORKING ISSUE INVOLVING CELERRA GbE PORTS :**

→Are Celerra interface ports built into Trunks, FSNs, or both ? [Relevant for all Ethernet 10/100/1000 Ethernet ports]

→Is network backbone built on Copper or Fiber GbE ?

→What is the link negotiation on the Switch ports that Celerra is connected to ?

→What is the Switch Port Speed Settings & Duplexing for ports connected to Celerra ? [Relevant for all Ethernet issues]

→If EtherChannel port aggregation is used, are Switch Ports enabled for EtherChannel (Cisco) ? [Relevant for all Ethernet]

→If Trunking is in use, verify type of Statistical Load Balancing in use—NAS 5.4 provides new default based on IP [MAC, IP, TCP]

→Is Flow Control Enabled or Disabled on Switch ports and data mover interfaces ?

→CatIOS or CatOS Version of the Switch that we are connected to ? [Relevant for all Ethernet issues]

→Network Topology would be ideal (though frequently not available) [Relevant for all Ethernet issues]

→Switch chassis model & switch module model numbers for ports to which DM is connected to [Relevant for all Ethernet issues]

→If LACP aggregation is configured on the DM, are the switch ports also configured ? [Relevant for all Ethernet issues]

→Switch port state and port error statistics on the ports to which the DM are connected to [Relevant for all Ethernet issues]

→Switch error log (if available) [Relevant for all Ethernet issues]

→Switch port VLAN configuration ?

## **GBIC [Gigabit Interface Converter] MODULES :**

I/O device that plugs into GB Ethernet Switch port supporting UTP Copper or Fibre Optic media. A GBIC is a transceiver that converts serial electrical current to optical and optical to digital electrical current.

**GIGABIT PORT NEGOTIATION:**

- Not used for negotiating port speeds
- Used for flow-control parameters and duplex information
- Port negotiation set to Enable by default for all ports
- Ports on both ends of a link must have same setting!!!! {Link may NOT come up if both ends are set differently}

**AutoNegotiation States and Resulting Link Status Table:**

| <i>Local Port Neg. State</i> | <i>Remote Port Neg. State</i> | <i>Local Link Status</i> | <i>Remote Link Status</i> |
|------------------------------|-------------------------------|--------------------------|---------------------------|
| Off                          | Off                           | Up                       | Up                        |
| On                           | On                            | Up                       | Up                        |
| Off                          | On                            | Up                       | Down                      |
| On                           | Off                           | Down                     | Up                        |

**ENABLING/DISABLING SWITCH PORT NEGOTIATION:**

# set port negotiation 2/1 enable [disable]

# show port negotiation 2/1

**GIGABIT FLOW-CONTROL:** Flow-control for a Gigabit port is implemented when the port's "receive" buffer becomes full. The port transmits a "pause" to remote ports in order to delay inbound traffic for Xamount of time.

**Flow Control Table:**

|                 |                                                                   |
|-----------------|-------------------------------------------------------------------|
| Receive on      | Port uses flow-control if dictated by neighboring ports           |
| Receive desired | Port uses flow-control if neighbor ports uses, otherwise does not |
| Receive off     | Does not use flow-control at all                                  |
| Send on         | Port sends flow-control frames to neighboring ports               |
| Send desired    | Port sends flow-control fames to neighbor ports only if requested |
| Send off        | Port does not use flow-control to send to neighboring ports       |

**SETTING SWITCH FLOW CONTROL:**

# set port flowcontrol 2/1 send on [receive on, etc.]

# show port flowcontrol [To verify settings]

**SPANNING TREE PROTOCOL:**

**Purpose:** Since there should only ever be (1) active path between (2) nodes on an Ethernet network, the Spanning Tree Protocol {STP} is used to prevent loops in multi-path configurations by automatically calculating the best path. Also used for path redundancy. A problem with Spanning Tree is that new connections are blocked for 45 sec. while deciding if the connection will cause a bridge loop. Fast Access is an option that will prevent this Spanning Tree Blocking.

**Basic Operation:** STP Works at each Port by Blocking>Listening>Learning>Forwarding>Disabled

**SPANNING TREE PROTOCOL: Port Transition States**

- I. Initialization to Blocking
- II. Blocking to Listening
- III. Listening to Learning [or Disabled] [This state not active if port uses Frame Forwarding for PortFast]
- IV. Learning to Forwarding [or Disabled]
- V. Forwarding to Disabled

**Note:** Use "PortFast" mode only on Ports that are directly connected to workstations so as to go directly to the "Forwarding" state, which bypasses the "Listening" & "Learning" states. The use of PortFast helps to overcome the problem where new connections are blocked for 45 secs. by the STP.

**ENABLING PORTFAST FOR SPANNING TREE PROTOCOL:**

# set spantree portfast 2/1 enable [disable]

# show spantree portfast 2/1

**Troubleshooting Switches:**

- |                         |                               |                                             |
|-------------------------|-------------------------------|---------------------------------------------|
| --switch configuration  | --switch blades configuration | --are there VLANs configured                |
| --Jumbo packets set?    | --Flow Control on?            | --Is there a diagram available?             |
| --switch software/patch | --switch model number         | --Network traces--port spanning of switches |

**SETTING PAUSE FRAME FLOW CONTROL ON CELERRA GbE INTERFACES:**

With NAS 4.2 and higher, gigabit interfaces can now be enabled for "Pause Frame Flow Control"

**Example:** Switch buffers can overflow, resulting in dropped packets to or from Celerra using GbE

**Cause:** While not generally likely in modern switches, some switches may be susceptible to having their buffers fill, overflow, and drop packets due to traffic bursts from the Celerra using Gigabit Ethernet.

**Fix:** Celerra can be configured to conform to 802.3 GbE Standard in which a feature called "Pause Frame Flow Control [RxTx]" can be enabled on the DM GbE interface. Implementing the following settings means that a capable switch can stop traffic flow from a Host if its buffers are getting full, then allow a resumption of traffic [similar to an X-on, X-off operation]:

### **SETTING PAUSE FRAME FLOW CONTROL ON CELERRA GbE INTERFACES:**

**1. \$server\_sysconfig server\_7 -pci ace0 -o txflowctl=enable | disable**

**2. \$server\_sysconfig server\_7 -pci ace0 -o rxflowctl=enable | disable**

### **PAUSE FRAME FLOW CONTROL:**

NAS 4.2 adds this feature where the switch can stop the flow of data if its buffers are full, and restart client when ready [X-on -off]

#### **Celerra Pause Frame Flow Control works under the following conditions:**

1. Autonegotiation is turned on: linkneg=enable
2. Pause Flow Control Set using Sysconfig: txflctl=desired; rxflctl=desired
3. Switch port also set to autonegotiate
4. Only supported on full duplex GbE links

### **WHEN WOULD PAUSE FRAME FLOW CONTROL BE USED FOR DM?:**

If inbound packets are being received and dropped by DM. Prior to implementing, check outbound flow control RingFull Errors on DM: \$.server\_config server\_x v “ace ace0 stat”. If autonegotiation is turned off on the DM & Switch port, pause frame send and receive can be enabled via the “txflowctl” & “rxflowctl” parameters.

**Note:** Pause Frame Flow Control support was added for Copper Gigabit interfaces on the DM with NAS 5.2.10.0.

### **RINGFULL ERRORS ON DM ETHERNET CARDS:**

RingFull errors result when outbound frames are being dropped from the DM interface, often due to oversubscribed hardware buffers during heavy load conditions or during GETTATTR storms from Solaris clients. Engineering has provided a software buffer to help handle these conditions, but the patch is not yet developed for BCM Copper Gigabit devices yet:

**param maxsecringsize=4000**

### **INTEL (ians) & BROADCOM (bcm) COPPER GbE INTERFACES:**

Both devices can experience a problem with ringfull errors on their transmit rings. We added secondary transmit ring for BCM devices—this is also being reviewed for IANS.

### **GIGABIT ETHERNET & AUTONEGOTIATION:**

**Term “autonegotiate” means that a device will detect Speed and then try to find out what the Duplex setting is at the other end**

1. Duplexing—usually always set to Full
2. Pause Frame Flow Control [RxTx]
3. Remote Fault—ability of link partner to signal status even if its receive link is not operational

### **CELERRA BEST PRACTICE WITH GIGE AND AUTONEGOTIATION:**

Celerra performs autonegotiation by default if Switch is enabled—in otherwords, if Switch supports Auto, then set Celerra to Auto  
If autonegotiation is not supported on the Switch, then it must be disabled [-o linkneg=disable]

**\$ server\_sysconfig server\_2 -pci ace0 -o linkneg=disable | enable** [flow control is also on]

### **GIGABIT ETHERNET NEGOTIATION RULE FOR CELERRA:**

→GbE on Celerra should never have a setting that is different from the corresponding Switch Port. Also, potential performance issues if Clientside is not set correctly from Interface to Switch:

**Full—Full or Auto—Auto**

### **GIGABIT INTERFACE TROUBLESHOOTING/KNOWN ISSUES:**

Symptom: Gigabit interface loses connectivity

**\$server\_config server\_x -v “ace ace0 clearstat”** [Clears the stats—server\_netstat server\_x -i]

**\$server\_config server\_x -v “ace ace0 stat”**

**\$server\_config server\_x -v “ace ace0 ringstat”**

Resolution: Delete Gigabit interface and rebuild using server\_ifconfig command

### **XBLADE 65 TEMPEST MODULE NETERION XFRAME 10GbE OPTICAL NIC:**

**\$ .server\_config server\_x -v “xena fxg0 stat”** [Viewing port statistics on the 10GbE interface]

**\$ .server\_config server\_x -v “xena fxg0 macstats”**  
**\$ .server\_config server\_x -v “xena fxg-1-0 stop”**  
**\$ .server\_config server\_x -v “xena fxg-1-0 start”**

### **10GbE NETERION FAILOVER ISSUES (Trunking, LACP, 10GbE to 1GbE, etc) AR152623:**

→A recent issue in 5.6.44 & 5.6.45 code exists when using Jumbo frames with 10GbE

#### **Suggested Workarounds:**

1. Set the following param

**param xena maxRxJumboRefill=192**

**\$ .server\_config server\_2 “xena fxg0 maxjumborefill 192”**

2. Or, if the customer is not using Jumbo Frames, then remove the following param from the /nas/site/slot\_param file

**param xena maxsdu=9000**

### **CELERRA MULTIHOMEING CHARACTERISTICS:**

- I. Multihome V1.0: Ability to have multiple interfaces in different subnets, but each interface supports only one subnet
- II. Multihome V1.1: Multiple interfaces on same subnet
- III. Multihome V1.2: Multiple NIC ports on one DM
- IV. Multihome V2.1: Trunking device; Single IP for multiple interfaces; Bandwidth aggregation; statistical load balancing; Link Failover; Supports Fast and Gigabit Ethernet.

### **CELERRA 10GbE ETHERNET NETWORKING:**

Single port Neterion 10GbE Ethernet module: NS80, NSX, NS-960 (Firestorm), NS-G8

Dual port Neterion 10GbE Ethernet module, + dual port Copper 1GbE: NX4, NS-G2, NS-120, NS-480 (Mojito)

#### **Media Specification for 10Gbps Ethernet:**

→Based on the IEEE 802.3ae standard, which allows for 10Gbps Ethernet transmission over short wavelength Multi-mode fiber (MMF), using core diameters of 50 or 62.5 microns

→10GBASE-SR/SW 10G Ethernet topology (SR=short range; SW=short wavelength)

→Full duplex and optical fiber only

→Firestorm and Mojito 10GbE cards use Finisar SFP+ Transceivers, FTLX8571D3BCL-E5, EMC Part # 019-078-041, which is a Class 1 VCSEL Laser Product (Vertical Cavity Surface Emitting Laser—low cost & power) that is RoHS-6 compliant, lead-free, and operates at 850nm wavelength as a 10Gb/s Datacom SFP+ Transceiver, delivering serialized data over multi-mode fiber (MMF) single wavelength @10.3125Gbit/s, modal bandwidths from 160MHz/km – 2000MHz/km, and distances from 26meters to 300meters, and is specifically designed for use in 10GBASE-SR/SW networks

→The SFP+ uses a duplex LC fiber optic connector (1.25mm), basically a miniature version of the SC connector (2.5mm)

#### **OM2 Shortware Optical Cable Media:**

8Gbps 50 meters

4 Gbps 100 meters

2 Gbps 300 meters

1 Gbps 500 meters

#### **OM3 Shortware Optical Cable Media:**

8Gbps 150 meters

4 Gbps 380 meters

2 Gbps 500 meters

1 Gbps 860 meters

### **CELERRA HIGH AVAILABILITY: (Ethernet Channels; Link Aggregation; FailSafe Networking)**

#### **CELERRA TRUNKING vs. CISCO:**

→Celerra trunking is considered link or port aggregation

→Cisco trunking is more along the line of VLAN tagging or multiplexing on a single logical port

→Celerra trunking consists of 2 or more ports logically trunked (aggregated for load distribution & statistical load balancing) so that the primary interface (ace0, cge0, etc) presents a single MAC address to the Switch, regardless of which actual physical interface the traffic may have left on. All traffic for individual TCP conversations (or IP or MAC, if designated), will travel from the same physical link.

→EtherChannel is called “port-channel” with Cisco’s CatIOS

#### **CELERRA NETWORK FAULT-TOLERANCE DEFINED:**

Both Ethernet Channels and Link Aggregation are methods for logically grouping physical ports together into fault-tolerant connections. Data Mover would have (2) or more ports connected to a switch configured for either Ethernet Channels or Link

Aggregation (trunking). Failure of (1) physical port in the group means that the Switch would automatically failover and route all traffic over remaining ports in configuration—this is not FSN, but normal switch failover from one port to another in an ethernet channel configuration. Failure of the entire ethernet channel or link aggregation channel would entail failing over to the next available connection.

### **BASIC TRUNKING/FSN RULES:**

#### **CELERRA PRESENTS SINGLE MAC ADDRESS FOR FSNs & TRUNKs FROM ACE0/CGE0:**

- For Trunking (EtherChannel), Switch ports must be enabled for EtherChannel. For Celerra, the act of creating a trunk between 2 or more interfaces invokes Trunking protocol and the MAC address of the primary NIC port (ace0/cge0) is presented to the switch.
- For Trunking, by definition, a set of Trunk interfaces must be plugged into the same Switch—cannot be split to 2 switches
- Trunking protocol must be enabled both on DART & the Switch Ports plugged into
- For FSNs, one port must be plugged into one switch while the other is plugged into a different, redundant path switch
- FSN feature is governed by DART, failing over from one FSN to another when the Active FSN fails

**Comments:** ‘MAC Address flapping’ occurs when a switch continually has to re-learn the same MAC address as it arrives on different ports (as would occur from a set of trunked DM interfaces), when the switch ports themselves are not configured for trunking. Some switch model types disable ports periodically when ‘mac address flapping’ occurs.

### **CELERRA FAILSAFE NETWORKS DEFINED:**

FSN's are fully redundant network connections consisting of (2) or more sets of connections—(1) connection is Active and the Other is in Standby, taking over if the Active segment fails. This type of network redundancy is monitored by the DataMover and upon failure, the DM will failover to next available connection in the FSN group. This feature is monitored and managed solely by DART and not the Switch. Each Link Aggregation or Ethernet Channel is considered a single logical connection & only one is active at a time. If implementing Ethernet Channels/Trunking and Link Aggregation with FSN, be aware that FSN failover would only occur if all devices failed in the Channel or Aggregation—therefore performance could degrade. In otherwords, implement FSN independent of Channel or Link Aggregation.

If using FSN's with a Primary & Secondary connection, the DM will automatically failover to secondary upon loss of primary, and then fail back to Primary if the latter is available. Best practices state that FSN's should not be configured with a primary so that the DM will switch to the next available connection and continue to use the connection until it too fails.

**# server\_netstat server\_3 -i**

| Name  | Mtu  | Ibytes     | Ierror | Obytes     | Oerror | PhysAddr        |
|-------|------|------------|--------|------------|--------|-----------------|
| ***** |      |            |        |            |        |                 |
| ana0  | 1500 | 2353619332 | 0      | 3449320128 | 0      | 0:0:d1:1f:ae:99 |
| ana1  | 1500 | 23206933   | 0      | 0          | 0      | 0:0:d1:1f:ae:9a |
| ana2  | 1500 | 409281831  | 0      | 3156524163 | 0      | 0:0:d1:1f:ae:9b |
| ana3  | 1500 | 44781179   | 0      | 0          | 0      | 0:0:d1:1f:ae:9c |
| ana4  | 1500 | 35610160   | 0      | 0          | 0      | 0:0:d1:1e:6a:f4 |
| ana5  | 1500 | 1915113172 | 0      | 2744764655 | 0      | 0:0:d1:1e:6a:f5 |
| ana6  | 1500 | 24914132   | 0      | 0          | 0      | 0:0:d1:1e:6a:f6 |
| ana7  | 1500 | 1844263097 | 0      | 722474104  | 0      | 0:0:d1:1e:6a:f7 |

**Note:** Above example shows pure FSN setup only, where primary interface for each of the (4) fsn's is handling the traffic, while the standby interface is only showing inbound bytes due to network broadcasts, etc., and no outbound bytes.

### **I. CELERRA ETHERCHANNEL (PORT AGGREGATION, aka CISCO ETHERCHANNEL):**

**Supported in NAS 2.1 +**

**Purpose:** Provides High Availability by using multiple physical links or ‘channels’.

- All Ports in the Ethernet Channel share the same IP and MAC addresses and are connected to the same switch
- EtherChannel can combine 2, 4, or 8 ports (8 links/channel) and can use comb. of Fast Ethernet and/or Gigabit, i.e., different speeds
- Provides for statistically based data distribution, not load balancing

**Note:** Inbound trunk traffic is distributed across receiving MAC addresses based on an algorithm called “statistical multiplexing”, which is done using the last two-four bits of the MAC address. In juxtaposition, ‘Etherchannel’ provides for Bandwidth aggregation, Statistical load balancing, and Link failover. Ethernet Channels group multiple physical ports into a single logical port. Network Switches are responsible for distributing traffic across ports in the Channel and for redirecting traffic to remaining ports upon a port failure (failover). Ethernet Channels can consist of 2, 4, or 8 ports and is similar to Link Aggregation.

**Note:** Statistical data distribution may increase combined throughput, but not single client throughput. We no longer refer to “Ethernet Channel or Trunking” for Celerra—EtherChannel.

**CELERRA ETHERCHANNEL DEFINED:** Mode used on Cisco switches to support EMC “trunking” [uses ISL protocol]. ‘EtherChannel’ is a proprietary Cisco protocol that allows multiple physical interfaces to be combined into a single logical channel, a single MAC address, and one or more IP Addresses. Celerra’s implementation of EtherChannel will use the MAC Address of the first interface specified in the Trunk Group. EtherChannel provides for more physical ‘links’ for clients to access the DM,

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
thereby allowing better throughput as well as higher link availability [However, the caution is always given that this is not “Link Aggregation” of throughput, as is real “EtherChannel Trunking”].

**ETHERCHANNEL AVAILABILITY:** Both ethernet channel "link failover" and Statistical Data Distribution are handled by the "switch", not datamover. In otherwords, on the DM itself, all EtherChannel ports are active and potentially in use.

**I. Link Failover:** If an active ‘link’ or ethernet port fails on DM, switch reroutes traffic to remaining “links” in channel.

**II. Statistical Data Distribution:** Switch controls distribution of packets based on source-destination MAC addresses and XOR bits.

**Note:** If a router were used, the MAC address of router will be used & will be same link [No Statistical distribution]

## **ETHERCHANNEL SUPPORT:**

--Celerra supports both Fast Ethernet and GbE interfaces at the same time

--Celerra uses 802.1Q VLAN Trunking protocol

**10/100 Fast Ethernet Support:** **2 or 4 ports on Cisco 5000; 2, 4, or 8 ports on Cisco 6000**

**Gigabit Ethernet Support:** **2 ports only**

**Key Concept:** Switches will transfer traffic to specific NIC ports in a channel based on XOR of predefined number of bits in MAC address of client and MAC address of Celerra channel [1,2,3 XOR bits for 2 NICs, 4 NICs, 8 NICs respectively]

**Celerra EtherChannel Trunking v. “Ethernet Trunking”:** True Ethernet Trunking uses the 802.1Q protocol and is known as “aggregate trunking”—it can combine multiple Ethernet Ports, such as four 100MB Ethernet Ports, to create an aggregated output of 400MB. While overall throughput may increase for Celerra Trunking, we DO NOT support it—we employ this mainly to provide “link” redundancy.

## **ETHERCHANNEL DEFINITIONS:**

--Ties multiple physical interfaces into logical channels that use a single MAC address to (1) or more IP Addresses

--Uses the MAC address of the first interface in the Celerra Trunk Group

--EtherChannel trunks are setup manually on Celerra

--Can combine 10/100/1000 Ethernet interfaces

--IP4700 uses PagP (Port Aggregation Protocol)

--A (2) Trunk Port sends data down one or the other pipe based on the last bit of the originating MAC

--4-Trunk Ports use the last 2 bits to determine path & 8 Trunk Ports uses 3 bits

Ethernet trunking is a point-to-point link between one or more switch{es} ports.

Trunks carry traffic for multiple VLANs over a single link, with the purpose of extending VLAN’s across the network.

**Multiple Ethernet Ports share a single IP address through one switch.** These port devices are combined into “Virtual Devices” that appear as a single “link.” Celerra Trunking can combine 2, 4, or 8 NIC ports on a DataMover into a single Channel—these ports would all “share” the same IP & MAC addresses. All “links” in a channel connect to the same switch

**Link:** A “link” is synonymous with an ethernet Port on the DataMover. I.e., all 2, 4, or 8 port combinations are “links” that combine under a single IP address known as a single “channel”.

**Channel:** A “channel” is synonymous with a single DataMover trunk. A channel is a virtual device that combines multiple physical devices into a single link.

**Rules:** NIC Ports all must be the same type!

Switch determines both Definition and Utilization of Ports.

Can configure 2, 4, or 8 NIC Ports on a Single DataMover [Sharing a single IP & MAC Address through one switch]

Over a Single Switch only.

## **FACTSHEET ON CELERRA ETHERCHANNEL:**

1. If switch supports Trunking, will conduct **statistical load balancing** based on hashing MAC addresses of Client--Server

2. If one port fails, other port will assume load, providing **High Availability—Link Failover**

3. **Single IP Address over multiple network devices** [i.e., ports share a Single address through one switch—virtual devices]

**Note:** Does not provide any increase in throughput for single clients but may for multiple clients

4. Compatible with Cisco “EtherChannel”

5. Can Trunk 2, 4, or 8 ports using mixed Duplexing and Speeds

6. 2 or 4 Ports on Celerra 10/100 Ethernet supported for Cisco 5000 or Bay Networks 450

7. 8 Ports on Celerra 10/100 Ethernet supported for Cisco 6000

8. 2 Ports for Gigabit ethernet on Alteon

9. A “trk” represents a ‘channel’ device [default=trk0]. Devices in a channel can be of different speeds! [not so with FSN]

10. Maximum of (8) Fast Ethernet ports can be trunked [ana0 – ana7]

11. Maximum of (2) Gigabit Ethernet ports can be trunked [ace0 – ace1]

12. Only one NIC port can be a Standby for another

13. Only one Port can be active at a time

14. Channel link on DM and Switch must be set the same

15. Spanning Tree must be enabled on switch in order to prevent bridge loop multicast/broadcast storms

## **SETTING UP CELERRA ETHERCHANNEL:**

1. Create Trunk Device: **\$server\_sysconfig server\_4 -v -n trk0 -c trk -o "device=ace0,ace1"**  
8-Port Trunk Device: **\$server\_sysconfig server\_4 -v -n trk2 -c trk -o "device=ana0,ana1,ana2,ana3,ana4,ana5,ana6,ana7"**
2. Set Trunk IP: **\$server\_ifconfig server\_4 -c -D trk0 -n trk0 -p IP 193.1.21.124 255.255.255.0 193.1.21.255**  
8-Port Trunk IP: **\$server\_ifconfig server\_4 -c -D trk1 -n trk1 -p IP 193.1.21.125 255.255.255.0 193.1.21.255**

**Note:** You can trunk 2, 4, or 8 ports across single or multiple Network Interfaces

## **TROUBLESHOOTING ETHERCHANNEL:**

1. Determine which devices are assigned to the Virtual Trunk: **\$server\_sysconfig server\_x -v**
2. Verify Duplex & Speed Settings: **\$server\_sysconfig server\_x -pci**
3. Check for Connectivity/Activity: **\$server\_netstat server\_x -i | -a | -s | -r**
4. Check ARP Table to verify Networking: **\$server\_arp server\_x -a**
5. Check Routing Table to verify Gateway, Routes, etc: **\$server\_route server\_x -l**
6. Verify IP, Mask, and Broadcast are Correct: **\$server\_ifconfig server\_x -a**
7. Check server logs—if Link goes Up and Down = bad port: **\$server\_log server\_x legrep trk**
8. Use Ping to verify/troubleshoot physical layer connectivity: **\$server\_ping server-x -send xxxx.xxxx.xxxx.xxxx**

**Note:** Windows systems, use c:>ping -n 20000 -l 65500 -w 0 10.241.169.24 [Max buffer size is 64k, 65500]

**Linux CS, use #ping -c 20000 -f -q -s 65507 10.241.169.25** [IP address of CS]. Use server\_ping -s command to hostname to see how fast the network is performing & whether there are timeouts.

9. Determine which method of ‘Statistical Load Balancing’ is in use [MAC Address, TCP, IP] and then make sure that the Switch ports are also configured to support the same thing

**\$ server\_sysconfig server\_2 -v -i trk001**

server\_2 :

\*\*\* Trunk trk001: Link is Up \*\*\*

\*\*\* Trunk trk001: **Statistical Load Balancing is TCP** \*\*\*

## **REMOVING AN ETHERCHANNEL GROUP FROM CELERRA:**

1. Delete IP Configuration of Trunk Group first using **\$server\_ifconfig server\_x -d trk0**
2. Delete the Virtual Interface using **\$server\_sysconfig server\_x -v -d trk0**

## **TROUBLESHOOTING NETWORKS USING .SERVER CONFIG COMMANDS:**

**\$server\_config server\_x -v "amgr showarp"**

**\$server\_config server\_x -v "amgr showroute"** → Outputs complete Routing Table

**# .server\_config server\_2 -v "amgr stat"**

UseHash is 0

Received:

amgr\_output = 13474

Transmitted:

amgr\_noARPOutput = 86

amgr\_beastOutput = 1438

amgr\_ipOutput = 11400

Drops:

amgr\_noMem = 0

amgr\_noReply = 36

amgr\_pendingOverflow = 55

amgr\_timerDrop = 349

amgr\_noIfp = 0

amgr\_releaseEntry = 0

amgr\_noRoute = 2

Informational:

amgr\_nonlocalSrc = 0

amgr\_orphans = 0

amgr\_addrAlloc = 1493

amgr\_addrFree = 1486

amgr\_arpBadVlan = 0

**\$server\_config server\_x -v "ace ace0 stat"**

**\$ .server\_config server\_x -v "printstats ipstat"** [IP Statistics and IP Reflect Statistics]

**\$ .server\_config server\_2 -v "tcp stat" "ip stat" "amgr stat"**

**\$ .server\_config server\_2 -v "ip allstat"** (good summary of IP statistics)

**\$ .server\_config server\_x -v "ace ace0 macstat"**

**\$ .server\_config server\_x -v "ace ace0 ifstat"**

**# .server\_config server\_3 -v "trunk trk0 stats"**

\*\*\* Trunk trk0: Link is Up \*\*\*

\*\*\* Trunk trk0: Timeout is Short \*\*\*

| Device | InPackets | InErrors | OutPackets | OutErrors | Link | LACP | Dupl | Speed |
|--------|-----------|----------|------------|-----------|------|------|------|-------|
|--------|-----------|----------|------------|-----------|------|------|------|-------|

|      |           |   |           |   |    |    |      |          |
|------|-----------|---|-----------|---|----|----|------|----------|
| cge0 | 110325486 | 0 | 133328911 | 0 | Up | Up | Full | 1000 Mbs |
| cge1 | 76109700  | 0 | 125650319 | 0 | Up | Up | Full | 1000 Mbs |
| cge3 | 49139801  | 0 | 136290971 | 0 | Up | Up | Full | 1000 Mbs |
| cge4 | 66039540  | 0 | 118871057 | 0 | Up | Up | Full | 1000 Mbs |

**# .server\_config server\_3 -v "trunk trk0 state"**

| Device | Receive | Periodic | Mux |
|--------|---------|----------|-----|
|--------|---------|----------|-----|

|      |         |         |            |
|------|---------|---------|------------|
| cge0 | CURRENT | SLOWPER | DISTRIBUTI |
| cge1 | CURRENT | SLOWPER | DISTRIBUTI |
| cge3 | CURRENT | SLOWPER | DISTRIBUTI |
| cge4 | CURRENT | SLOWPER | DISTRIBUTI |

**# .server\_config server\_3 -v "trunk trk0 showcfg"**

1139001535: DRIVERS: 4: \*\*\* Trunk trk0: Link is Up \*\*\*

1139001535: DRIVERS: 4: \*\*\* Trunk trk0: Timeout is Short \*\*\*

1139001535: DRIVERS: 4: \*\*\* Trunk trk0: Statistical Load Balancing is IP \*\*\*

1139001535: DRIVERS: 4: Device Local Grp Remote Grp Link LACP Duplex Speed

| 1139001535: DRIVERS: 4:      | Device | Local Grp | Remote Grp | Link | LACP | Duplex | Speed |
|------------------------------|--------|-----------|------------|------|------|--------|-------|
| 1139001535: DRIVERS: 4: cge0 | 10000  | 37381     | Up         | Up   | Full | 1000   | Mbs   |
| 1139001535: DRIVERS: 4: cge1 | 10000  | 37381     | Up         | Up   | Full | 1000   | Mbs   |
| 1139001535: DRIVERS: 4: cge3 | 10000  | 37381     | Up         | Up   | Full | 1000   | Mbs   |
| 1139001535: DRIVERS: 4: cge4 | 10000  | 37381     | Up         | Up   | Full | 1000   | Mbs   |

**# server\_sysconfig server\_3 -v -info trk0**

server\_3 :

\*\*\* Trunk trk0: Link is Up \*\*\*

\*\*\* Trunk trk0: Timeout is Short \*\*\*

\*\*\* Trunk trk0: Statistical Load Balancing is IP \*\*\*

Device Local Grp Remote Grp Link LACP Duplex Speed

| Device | Local Grp | Remote Grp | Link | LACP | Duplex | Speed    |
|--------|-----------|------------|------|------|--------|----------|
| cge0   | 10000     | 37381      | Up   | Up   | Full   | 1000 Mbs |
| cge1   | 10000     | 37381      | Up   | Up   | Full   | 1000 Mbs |
| cge3   | 10000     | 37381      | Up   | Up   | Full   | 1000 Mbs |
| cge4   | 10000     | 37381      | Up   | Up   | Full   | 1000 Mbs |

**# .server\_config server\_3 -v "trunk trk0 lacpinfo"**

| Device | Local SysId | Loc Port | Remote SysId | Rem Port | Link | LACP | Dplx | Astate | Pstate |
|--------|-------------|----------|--------------|----------|------|------|------|--------|--------|
|--------|-------------|----------|--------------|----------|------|------|------|--------|--------|

|      |                  |       |                  |       |    |    |      |    |    |
|------|------------------|-------|------------------|-------|----|----|------|----|----|
| cge0 | 0:60:16: 4:3f:41 | 10000 | 0:d0: 3:73:28: 0 | 37381 | Up | Up | Full | 3f | 3c |
| cge1 | 0:60:16: 4:3f:41 | 10000 | 0:d0: 3:73:28: 0 | 37381 | Up | Up | Full | 3f | 3c |
| cge3 | 0:60:16: 4:3f:41 | 10000 | 0:d0: 3:73:28: 0 | 37381 | Up | Up | Full | 3f | 3c |
| cge4 | 0:60:16: 4:3f:41 | 10000 | 0:d0: 3:73:28: 0 | 37381 | Up | Up | Full | 3f | 3c |

**# .server\_config server\_3 -v "trunk trk0 lacpstats"**

| Device | InLACP | OutLACP | BadLACP | InMarke | OutMark | BadMark |
|--------|--------|---------|---------|---------|---------|---------|
|--------|--------|---------|---------|---------|---------|---------|

|      |       |      |   |   |   |   |
|------|-------|------|---|---|---|---|
| cge0 | 51981 | 1806 | 0 | 0 | 0 | 0 |
| cge1 | 51980 | 1807 | 0 | 0 | 0 | 0 |
| cge3 | 51980 | 1807 | 0 | 0 | 0 | 0 |
| cge4 | 51980 | 1807 | 0 | 0 | 0 | 0 |

**\$ .server\_config server\_x -v "fsn fsn0 stat"**

**\$ .server\_config server\_5 -v "tcp connstats"**

1132578385: TCP: 4: TCP Connection Statistics:

```
1132578385: TCP: 4: Port 0: 0
1132578385: TCP: 4: CIFS : 175
1132578385: TCP: 4: NFS : 0
1132578385: TCP: 4: ISCSI : 0
1132578385: TCP: 4: FTP CTRL : 0
1132578385: TCP: 4: FTP DATA : 0
1132578385: TCP: 4: HTTP : 0
1132578385: TCP: 4: OTHERS : 36
```

**\$ .server\_config server\_2 -v "tcp stat"**

TCP stats :

```
connections initiated 381
connections accepted 16506
connections established 16887
connections dropped 0
embryonic connections dropped 0
conn. closed (includes drops) 16877
segs where we tried to get rtt 2537138
times we succeeded 2628288
delayed acks sent 11351
-----output abridged-----
```

### **CELLERRA ETHERCHANNEL STATISTICS:** Use “nfsstat” for traffic statistics

Use \$server\_netstat server\_x -i to view which interfaces are active for In and Out Bytes

### **CELLERRA LINK AGGREGATION CONTROL PROTOCOL: LACP**

**New with NAS 5.0:** 802.3ad Link Aggregation Control Protocol. Tested with Cisco and Foundry switches, but Foundry's implementation does not work well.

#### **PURPOSE:**

Similar to Ethernet Channels (Trunking) in that groups of physical ports are combined into a single logical interface to provide for enhanced performance or enhanced availability. Link Aggregation works with Switches that support IEEE 802.3ad Link Aggregation standards for two or more ports. A major difference from Ethernet Channels is that all ports in the Channel must have same Speed & Duplex settings & have the same type of NIC cards. Switch is responsible for failing over traffic to all unaffected ports. Similar to Cisco's PAgP EtherChannel protocol.

#### **LINK AGGREGATION ATTRIBUTES:**

- Combines multiple Ethernet Links or Ports into a single Virtual Device trunk, and does not need to be in multiples of two [could be 2, 3, 4, 5, 6, 7, etc. links combined]
- Number of links per trunk are not limited
- Unlike EtherChannel, however, LACP does not support different speeds among the devices in the Trunk
- Provides for Higher Availability [One link goes down, others still active—may result in loss of bandwidth, but not loss of service]
- Port Distribution: Links are determined by Source & Destination MAC Addresses
- Provides for better Link Control between DM and Switch
  - [LACPDU Frames are transmitted over each link in the Trunk to monitor]
- DM Ethernet devices must be set to FULL Duplex as Links operate at FULL Duplex
- No ATM or FDDI with LACP
- All Ports should be set at the same Speed, Full Duplex, and of same NIC type
- Maximum of (12) physical Ethernet devices can be combined into a single LACP Trunk [Only Physical Ports can be used]
- LACP protocol can automatically identify links operating correctly and take improperly configured links offline
- Link aggregations provide more overall bandwidth, but any single client only communicates through one port and is limited to the bandwidth of that port

#### **SETTING UP AN LACP TRUNK:** Very similar to EtherChannel Trunking Setup

Step 1. Create Virtual LACP Trunk:

**\$ server\_sysconfig server\_6 -v -n trk0 -c trk -o “device=ana2,ana1,ana0,ana3 protocol=lACP”**

Step 2. Add IP to Device: **\$server\_ifconfig server\_6 -c -D trk0 -n trk0 -p IP 192.10.3.28 255.255.255.0 192.10.3.255**

#### **DELETING AN LACP TRUNK DEVICE:**

**\$ server\_sysconfig server\_6 -v -d -Force trk0** [Deletes both IP & Trunk]

#### **CREATING LACP TRUNKS WITH FSN NETWORKS:**

1. **\$ server\_sysconfig server\_6 -v -n trk0 -c trk -o “device=ana0,ana1,ana2,ana3 protocol=lACP”**

2. **\$ server\_sysconfig server\_6 -v -n trk1 -c trk -o "device=ana4,ana5,ana6,ana7 protocol=lacp"**
3. **\$ server\_sysconfig server\_6 -v -n fsn0 -c fsn -o "device=trk0,trk1"**
4. **\$ server\_ifconfig server\_6 -c -D fsn0 -n fsn0 -p IP 192.10.3.28 255.255.255.0 192.10.3.255**

### **VERIFYING LACP VIRTUAL DEVICES:**

```
$ server_sysconfig server_6 -v  
$ server_netstat server_6 -i  
$ .server_config server_6 -v "fsn fsn0 ifstat"  
$ .server_config server_6 -v "fsn fsn0 stat"  
$ .server_config server_6 -v "fsn fsn0 clearstat"
```

**FROM CISCO SWITCH:** Only Supported on Catalyst OS Version 7.1 and above

```
# show port lacp-channel [LACP Statistics]  
# show lacp-channel traffic
```

### **CISCO CATALYST 6000 SWITCH:**

```
#set channelprotocol lacp 9  
#show channelprotocol  
#set port lacp-channel 3/3-16 [Sets Ports to an Admin Key]  
#set port lacp-channel 3/3-16 mode active  
#show port lacp-channel
```

### **LACP LIMITATIONS:**

--Only supports full duplex Ethernet Links  
--All links must be same speed setting

### **ENHANCING LACP AND TRUNKING STATISTICAL LOAD BALANCING:**

**param trunk LoadBalance=tcp**

**param trunk LoadBalance=ip**

**param trunk LoadBalance=mac**

**Note:** Default is statistical load balancing based on source IPv4 and Destination IP addresses—binds to IP of NIC card (ip). MAC value binds to mac address of NIC cards connecting to Clients, useful if large number of NICs with sequential MAC values, not similar values, where load is distributed based on ranges of MAC values. TCP filtering is similar to IP but also takes ‘port’ into account.

**\$ .server\_config server\_2 -v "param trunk"**

| Name              | Location   | Current | Default |
|-------------------|------------|---------|---------|
| trunk.LoadBalance | 0x01356f20 | 'tcp'   | "       |

**\$ server\_sysconfig server\_2 -v -i trk001**

server\_2 :

\*\*\* Trunk trk001: Link is Up \*\*\*

\*\*\* Trunk trk001: Statistical Load Balancing is TCP \*\*\*

Device Link Duplex Speed

**Note:** Prior to NAS 5.4, which now uses IP as the default value for statistical load balancing, mac address was used

### **LACP TROUBLESHOOTING COMMANDS:**

**# server\_sysconfig server\_3 -v -info trk0**

**# .server\_config server\_3 -v "trunk trk0 lacpinfo"**

**# .server\_config server\_3 -v "trunk trk0 lacpstats"**

**# .server\_config server\_3 -v "trunk trk0 showcfg"**

## **II. CELERRA FAILSAFE NETWORKING [FSN]: Failover is DataMover dependent!**

FSN = 2.2.6+ NAS code

### **Purpose:**

FailSafe Networks are redundant network connections that failover from an Active to a Standby. FSNs are monitored and failed over by the Data Mover alone, and not the Switch. FSNs can consist of a group of Ports, Ethernet Channels, or Link Aggregation connections—Data Mover supports up to 2-8 “connections”. FSNs typically retain full bandwidth performance after failover. All connections in the FSN device share a single MAC address and a set of common IP Addresses. Additionally, the underlying Ethernet devices can be of different speeds. By default, failover is automatic, but fallback is manual.

**Note:** Unlike trunking, where all ports are still “active”, FSN has only a single active port at a time, with one in standby

## **FSN ATTRIBUTES:**

- Composed of multiple NIC ports on same DataMover as a Virtual Device using one IP and MAC Address
- Fails over to a Standby device if the primary Failsafe network device fails
- Fails over if the Switch fails on the Failsafe Network Device
- Each 'channel' must be identical in an FSN group!
- Can use FSN in Multiple Switch, VLAN, or Trunking configurations
- Can use a combination of NIC Ports together--e.g., Fast Ethernet or Gigabit Ethernet
- Only one port in an fsn network is active at a time

## **Typical FSN Configurations:**

Datamover NIC ports connected to different physical switches in support of Failsafe Switch Failover  
NIC ports combined as a virtual primary device, failing over to Standby device when necessary

## **Typical FSN Failover:**

- Primary port fails--the operational Standby port will take over
- When active communications link between Switch & DM fails--will failover to an available link within the fsn group
- If primary device is recovered, the communications link is restored automatically
- If no 'primary' device was specified in an fsn setup, then the original device does not restore automatically

## **IMPORTANT NOTE:**

**FSN Failover is monitored and conducted by DART, not the network switch! We failover from an fsn on one switch to a different fsn on an alternate switch.**

### **Restores:**

- If Primary device is restored, will take over from Standby port

### **FailOver Between Different IP Channels in a Switch for the same DataMover:**

"Automounter" feature can be used so that a timeout value will force failover to another IP channel if the first one becomes too busy.

## **TESTING GIGABIT FAILOVER BETWEEN PRIMARY [ace0] & STANDBY [ace1]:**

Step 1. Testing "FSN" failover: Remove sc connector from "ace0"--assumption is that it is the "primary" interface

Step 2. Standby port [ace1] will take over immediately with no noticeable pause in activity

Step 3. When getting ready to plug "ace0" back in, insert the top "TX" connection first, wait 5 seconds to allow time for the switch to bring that port back up, then insert the "RX" connector--this should allow for a seamless failback. Otherwise, there could be a 3-4 second delay as the switch reactivates the port

## **RESULTS OF MISMATCHED INTERFACES ON DM FAILOVER:**

**Configuration:** DMs have FSNO configured (Active Port=ACE0 Stndby=ACE1); Primary DMs have linkneg=disabled and corresponding Switch ports disabled for autonegotiation. Standby Server has switch ports set for autonegotiation.  
Failover of Primary to Standby Server is this situation will not allow ACE0 to come up due to switch port mismatch—instead, ACE1 will come up.

## **LAYER 3 FAILOVER [NAS 5.6]:**

- IP Addresses are assigned to Virtual devices and not to physical devices
- Data Mover uses Routing Tables to determine proper return route
- IPReflect feature must be disabled if using Layer 3 failover
- Data Mover uses RIPv2 to advertise to Routers for all interfaces configured to a Virtual device
- Available RPQ only, requires Eng. Key to enable

### **Example of how Layer 3 failover is used for High Availability:**

Data Mover A in one location uses 192.168.50.10 IP on a virtual device with RIP cost of 0

Data Mover B in other location uses 192.168.50.10 IP on a virtual device with RIP cost of 15

Scenario→Data Mover A goes down, Routers update tables and send packets to Data Mover B

### **# server\_param server\_2 -facility vdevice -list**

```
server_2 :  
param_name           facility default   current   configured  
key                 vdevice    'none'   'none'
```

### **ENABLING LAYER 3 FAILOVER:**

```
# server_param server_2 -facility vdevice -modify key -value <eng_key>  
# server_param server_2 -facility ip -modify reflect -value 0  
# server_sysconfig server_2 -virtual -name ha_vdevice -create ha [failover device]  
# server_ifconfig server_2 -create -Device ha_vdevice -name ha_interface -p IP 192.168.50.10 [failover interface]
```

# server\_rip server\_2 –modify ha\_out\_enabled y [Enable RIP advertisement on DM]

# server\_rip server\_2 –append ha\_out cge1 [Globally accessible gateway interface]

### DELETING LAYER 3 FAILOVER:

# server\_ifconfig server\_2 –delete ha\_interface

# server\_sysconfig server\_2 –virtual –delete ha\_vdevice

### STOPPING RIP ADVERTISEMENTS:

# server\_rip server\_2 –remove –ha\_out cge1

# server\_rip server\_2 –modify –ha\_out\_enabled n

### LISTING VIRTUAL TRUNK/FSN DEVICES:

# server\_sysconfig server\_2 -v

server\_2 :

Virtual devices:

fsn1 active=cge3 primary=cge3 standby=cge4

fsn0 active=cge0 primary=cge0 standby=cge1

fsn failsafe nic devices : fsn0 fsn1

trk trunking devices :

# .server\_config server\_2 -v "fsn fsn0 stat"

\*\*\* fsn0 is Up, active link is cge0 \*\*\*

Transmit drops = 10, Ad. requested = 58, Ad. sent = 57

Link ups = 9, Link downs = 7

fsn0 ups = 4, fsn0 downs = 3

Device InPackets InErrors OutPackets OutErrors Link Speed

-----

cge0 1463786 0 1185782 0 Up 10000000000

cge1 75 15 113 0 Up 1000000000

1122381376: ADMIN: 4: Command succeeded: fsn fsn0 stat

# server\_sysconfig server\_2 -v -i fsn0

server\_2 :

\*\*\* FSN fsn0: Link is Up \*\*\*

active=cge0 primary=cge0 standby=cge1

### Verifying IP Configuration for Virtual Devices: \$server\_ifconfig server\_2 fsn0

Shows IP, Mask, Broadcast, whether device is UP or DOWN, MTU Size, and VLAN info

Caution: fsn0 represents the “name” of the virtual interface, as given in lefthand column of server\_ifconfig –a command

### CONFIGURING FAILSAFE NETWORKING: [FSN]

1. Create Failsafe Configuration:

\$server\_sysconfig server\_4 -v -n fsn0 -c fsn -o “primary=ana0 device=ana1”

2. Create IP Configuration:

\$server\_ifconfig server\_4 -c -D fsn0 -n fsn0 -p IP 193.1.21.126 255.255.255.0 193.1.21.255

Note: Must use quotations around the “primary=” & “device=” statements!!

Specifying Primary & Failover FSN Device: -o “primary=ana0”litrk0|fsn0 “device=ana1”litrk1|fsn1” [device=failover Standby]

Specifying Failover List of FSN Devices: -o “device=trk0,trk1” [trk1 represents failover device for trk0]

### CONFIGURING TRUNKING WITH FAILSAFE NETWORK:

1. Trunk Group: \$server\_sysconfig server\_4 -v -n trk0 -c trk -o “device=ana0,ana1”

2. FailSafe Group: \$server\_sysconfig server\_4 -v -n fsn1 -c fsn -o “primary=trk0 device=ana2,ana3”

3. IP Configuration: \$server\_ifconfig server\_4 -c -D fsn1 -n fsn1 -p IP 193.1.21.127 255.255.255.0 193.1.21.255

### CONFIGURING MULTIPLE TRUNKS WITH FAILSAFE NETWORKS:

1. Trunk Group One: \$server\_sysconfig server\_4 -v -n trk0 -c trk -o device=ana0,ana1,ana4,ana5

2. Trunk Group Two: \$server\_sysconfig server\_4 -v -n trk1 -c trk -o device=ana2,ana3,ana6,ana7

3. FailSafe Group: \$server\_sysconfig server\_4 -v -n fsn2 -c fsn -o device=trk0,trk1

4. Reboot server!

Configuration Output: fsn2 active=trk0 standby=trk1  
trk0 devices=ana0,ana1,ana4,ana5

trk1 devices=ana2,ana3,ana6,ana7  
Fsn failsafe device: fsn2 Trk trunking device: trk0 trk1



--VLANs are best used for routing a single logical network over a single VLAN—however, with caution, multiple logical networks can be configured over the same VLAN

--Data Movers will not allow two different physical interfaces to be on the same subnet if the VLAN ID is different

**VLAN TRUNKING:** Celerra uses “VLAN Trunking”, which consists of a grouping of Ports on a switch. VLAN Trunking is better defined as a point-to-point link between two switches that can pass unique VLAN numbers over a single or multiple links. We support “trunking” in a VLAN environment without reference to a specific “protocol”. The two basic VLAN Trunking protocols are 802.1Q (Celerra) and Cisco’s proprietary ISL protocol.

**Note:** All ports in a trunk are “active” & capable of passing traffic

**Trunking Protocols:** ISL and 802.1Q

Inter-Switch Link [ISL]—Cisco trunking encapsulation protocol [Catalyst 5000+ switches]

IEEE 802.1Q—industry standard trunking encapsulation protocol [Catalyst 4000+ switches]

**UNSUPPORTED:**

*Nortel Switches do not support Celerra EtherChannel Trunking*

## **USING VLANS WITH DATA MOVERS:**

Customer might want to multi-home the DM so that clients can access from different logical networks without having to go through a Router. Would do this by setting up a Virtual Interface with different IP addresses, one for each logical network.

## **TWO WAYS TO IMPLEMENT VLANS FOR CELERRA:**

1.) Configure VLAN on switch port with VLAN identifier and connect Server to the port.

**Note:** Typical client configuration. File Server is not aware that it is part of the VLAN, and no special configuration required. The VLAN ID is set to 0.

2.) Configure Switch ports for VLAN tags and configure DM to process VLAN tags on packets. This enables DM to connect to multiple VLANs through a single physical interface, also useful when configuring Standby DMs for failover if the failover server is on a different VLAN. Using this method, both Switch Ports and Celerra Servers are configured to used VLAN tags on packets.

## **ADVANTAGES OF VLANS:**

--ease of Administration

--Reduces Broadcast traffic by creating separate VLANs or Broadcast Domains

--Better than Routing by reducing Broadcast traffic & enabling better efficiency in network traffic

### **Disadvantages:**

--Fewer standards: LANE; 802.10; 802.1q; 802.1p

### **Misc:**

--Setup VLAN trunks between switches to allow traffic to pass between VLAN groups

--Use a router to route traffic between VLANs

--VLANS share ports on multiple switches using trunking

--Grouping ports under one IP Address

--Ports in a VLAN share Broadcast traffic

--DM can have multiple VLANs per port or single VLAN, spanning multiple IPs and Subnets

--Support for up to 14 DataMover's per VLAN

--Requires Layer 3 Switches

## **VARIOUS TYPES OF VLAN IMPLEMENTATIONS:**

**PORT-BASED VLANS:** Easy setup; only (1) virtual LAN per port; Static assignments

**MAC-BASED VLANS:** Difficult setup; Easy to maintain

**LAYER-3 VLANS:** Based on IP Subnets

**PROTOCOL-BASED VLANS:** Protocol type of frames determine membership

**MISC VLANS:** Policy-Based; Multicast; Authenticated Users

## **CISCO VLAN:**

IOS Virtual LAN Services

IEEE 802.1Q

IEEE 802.10 Interoperable LAN/MAN Security Standard (SILS)

--allows LAN traffic to carry a VLAN identified

--MAC layer frame with 802.10 header, called PDU using SDE (Security Data Exchange)

## **SOME CISCO DEFINITIONS OF VLAN:**

→VLAN is a shared broadcast domain within a switched network that is logically segmented into distinct groups of users

→End stations may be in separate physical LANs, but grouped together in a VLAN segment

- VLAN ports can be grouped on a single switch or across multiple switches per VLAN group
- Communication between (1) VLAN and another is done across a Router
- VLANs are normally numbered from 2-1000 (1 is reserved for the Management VLAN)
- Configuring VLAN on single switch is easy, just configure
- Configuring VLAN across multiple switches requires VLAN Trunking Protocol (VTP), and requires certain Trunk mode and Trunk encapsulation mode match (the manner in which frames are tagged for the trunk)
- Frames travelling over a trunk are tagged to identify which VLAN the frames belong to

### **FOUR TYPES OF TRUNKING ENCAPSULATIONS:**

- 1) Inter-Switch Link Protocol (ISL) -- Cisco proprietary trunking protocol.
- 2) IEEE 802.1Q (dot1q) -- Industry standard trunking protocol.
- 3) LAN Emulation (LANE) -- Used for trunking VLANs over ATM links.
- 4) IEEE 802.10 (dot10q) -- Cisco proprietary method for transporting VLAN information inside standard FDDI frames.

### **CONFIGURING VLAN ON SINGLE CISCO IOS SWITCH:**

```
SwitchA# vlan database
SwitchA(vlan)# vlan 2 name vlan2
SwitchA(vlan)# exit
SwitchA# configure terminal
SwitchA(config)# interface fastethernet 0/1
SwitchA(config-if)# switchport mode access
SwitchA(config-if)# switchport access vlan 2
SwitchA(config-if)# end
```

### **SETTING UP TRUNKING BETWEEN TWO CISCO IOS SWITCHES:**

1. SwitchA(config)# interface fastethernet 0/1  
SwitchA(config-if)# switchport mode trunk  
SwitchA(config-if)# switchport trunk encapsulation dot1q
2. Repeat commands on SwitchB.

**Note:** Sets up fast Ethernet interface 0/1 to be a trunk port using dot1q encapsulation. It is important to note that once you change one side of a connection to trunk mode, communication between the two switches will be lost until the other side is configured for the same mode/encapsulation. If you are trying to set up trunking remotely, always change the far side of a connection first. The port is currently passing information for all VLANs (1-1005).

3. To limit which VLANs will be allowed to pass information on the port:

```
SwitchA(config)# interface fastethernet 0/1
SwitchA(config-if)# switchport trunk allowed vlan remove 1-1005
SwitchA(config-if)# switchport trunk allowed vlan add 1-3
```

4. Repeat commands on SwitchB.

5. This removes the default of all VLANs, and adds back support for VLANs 1-3.

### **VLAN SWITCH/ROUTER TROUBLESHOOTING:**

1. Ports assigned to wrong VLAN—i.e., same IP Address each VLAN  
**Note:** VLANs are setup on the SWITCH
2. Running wrong protocol, as set on the “Router”—ISL v. dot1Q {802.1Q}—must be all .1Q or ISL.
3. Check to see if VLAN trunk is “active”
4. Cross Connection Issue: To route between VLAN’s, must be on a different subnet and must use a Router  
**Note:** (5) Modes of Trunking for Fast Ethernet/Gigabit—default mode is AUTO [on—off—desirable--auto—nonegotiate]  
Commands: sh config; sh vlan; sh port; sh trunk [commands to verify VLAN setup on a Switch]

### **CELLERRA VLAN TAGGING:**

--VLAN Tagging can be implemented over a single “link” or multiple “links” such as in EtherChannel.

--Feature using 802.1Q protocol over GigabitEthernet to support VLAN tagging of packets.

--Supported with NAS 4.2

--VLAN Tagging for network devices is only supported for GbE ace and cge ports using 802.1Q

--‘Frame Tagging’ using 802.1Q; 22-byte VLAN ID in header of every packet

--Another IEEE standard is 802.3q for VLAN frame Tagging

--802.1Q only supported with Gigabit Ethernet

--VLAN Tagging is useful in situations where you have multiple VLANs

--VLAN Tagging is controlled on the Switch

**Note:** NS600 changes this with NAS 5.1 code—supports 10/100/1000

### **VLAN TAGGING & NIC DEVICE DRIVER SUPPORT:**

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
VLAN Tagging support is a function of device drivers, not speed. Hence, both Alteon (ACE) & Broadcom (CGE) can support VLAN Tagging for 10/100/1000 Base T, while LSI (ANA) drivers do not.

## **CONFIGURING VLAN TAGGING:**

Step 1. Create Gigabit Interface:

```
$server_ifconfig server_2 -c -D ace0 -n Group1 -p IP 192.50.13.24 255.255.255.0 192.50.13.255
```

Step 2. Set Interface to Use VLAN: \$server\_ifconfig server\_2 Group1 vlan=100 [Value can be between 1-4095]

```
Deleting VLAN: $server_ifconfig server_2 Group1 vlan=0 [Vlan=0 means no VLAN ID set]
```

Step 3. Verifying VLAN on an Interface: \$server\_ifconfig server\_2 Group1

**Note:** Do not use “vlan=1” as an ID, as many switches reserve this for their native VLAN

## **CREATE MULTIPLE VLANs USING SINGLE PHYSICAL INTERFACE & 802.1Q TAGGING:**

Step 1. Create Virtual Interfaces:

```
$server_ifconfig server_6 -c -D ace0 -n acev1 -p IP 192.10.4.28 255.255.255.0 192.10.4.254
```

```
$server_ifconfig server_6 -c -D ace0 -n acev2 -p IP 192.10.4.29 255.255.255.0 192.10.4.254
```

Step 2. Assign VLANs:

```
$server_ifconfig server_6 acev1 vlan=acev1
```

```
$server_ifconfig server_6 acev2 vlan=acev2
```

## **CONFIGURING FSN WITH 802.1Q TAGGING:**

Step 1. Create Virtual FSN:

```
$server_sysconfig server_6 -v -n fsn0 -c fsn -o device=ace,ace1
```

```
$server_ifconfig server_6 -c -D fsn0 -n acev1 -p IP 192.10.4.29 255.255.255.0 192.10.4.254
```

```
$server_ifconfig server_6 acev1 vlan=10
```

## **CONFIGURING TRUNK WITH 802.1Q TAGGING:**

Step 1. Create Trunk:

```
$server_sysconfig server_6 -v -n trk0 -c trk -o device=ace0,ace1
```

```
$server_ifconfig server_6 -c -D trk0 -n acev1 -p IP 192.10.4.28 255.255.255.0 192.10.4.255
```

```
$server_ifconfig server_6 acev1 vlan=10
```

```
$server_ifconfig server_6 acev1
```

## **CONFIGURING VLAN TAGGING ON CISCO SWITCHES: 802.1Q, Cisco 4000, 5000, 6000**

1. #set vlan 100 3/10 [Assigns switch port to the VLAN]
2. #show vlan 100 or #show port 3/10 [Use to verify VLAN membership, speed, & duplex settings]
3. #set trunk 2/8 on dot1q [Setting the Trunk on the switch]
4. #show trunk 2/8 [Verify trunk configuration]
5. #set vlan mapping dot1q 100 [Set valid VLAN range for 802.1Q: 1-4095 value; For isl\_vlan, values are 1-1005]
6. #show vlan mapping [Verify mapping]

## **ISSUE WITH OUT OF SEQUENCE VLAN PACKETS AND TOE DRIVERS:**

# .server\_config server\_2 -v "param ip vlanFilter"

ip.vlanFilter INT 0x0136b2f8 1 1 (0,4294967295) FALSE REBOOT 'NA'

**Note:** VlanFilter is set by default to drop packets that are received with unexpected VLAN id. Disabling this param does not affect DM id tagging of outbound VLAN packets

## **TROUBLESHOOTING TCP:**

### **Max Number Client TCP Connections to the DataMover:**

Max number of connections/concurrent sessions per Datamover using 2.2.35.4+: 20,000/3000

Max number of concurrent TCP connections prior to 2.2.35.4: 2000

## **TCP PROTOCOL:**

- Connection-oriented byte-stream service using IP network layer.
- Applications generate data into segments, which are transmitted by TCP
- TCP keeps track of segments with timers and acknowledgements to ensure delivery
- TCP segments are also transmitted as IP segments
- TCP is full duplex to the Application layer

## **I. CELERRA TCP/UDP RETRANSMISSIONS (NFS or CIFS):**

Layer 4 retransmissions at the TCP layer. Basic rule of thumb is to have .01% or lower retransmit rates.

**\$ .server\_config server\_2 -v “tcp stat”** [Includes TCP retransmits & retransmit timeouts--RTOs]

**\$ server\_netstat server\_2 -s -p tcp** [Does not contain the RTO information]

**\$ .server\_config server\_2 -v “printstats tcpstat reset”**

#### **FAST RETRANSMISSIONS(fastRTO):**

Fast retranmissions are triggered whenever (4) consecutive duplicate acknowledgements are received. The sending host retransmits data after the 3<sup>rd</sup> duplicate ACK. i.e., whenever the Server receives (3) or more consecutive ACKS that are out of order, it will request a retransmission. Thereafter, will ACK every segment received for a period of time. Fast retranmissions generally occur when a single packet is missing from the Receiver's advertised TCP window & have a minimal impact on performance. FastRTO = 1 slow clock tick, 500ms. FastRTO is a TCP parameter that would apply to CIFS, NFS, TCP, or IP Replication activities.

#### **RETRANSMISSION TIMEOUTS (RTO's):**

RTO's occur when there is multiple packet loss within Receiver's TCP window, occurring when sending hosts retransmission timer expires. Celerra RTO is (3) slow clock ticks, or (1.5) seconds [most NT clients wait much less, say 200ms]. RTO's have a very negative impact on performance. NAS 5.4 will reduce RTO's with implementation of NewReno and SACK RFC's.

**Comment:** TCP runs on a slow and fast clock. Problem can occur on networks where because of the slow RTO mechanism, the Celerra will seem to take a long time before retransmitting because of lost packets. Usually, this all happens on dirty networks.

Microsoft's implementation of RTO is to resend the missing packets without waiting for RTO Timer to expire. Usually, the Server should wait for 3 consecutive out of order packets before doing a retransmit: Microsoft can adjust this to between 1-3 consecutive packets.

--Celerra's Slow Clock is based on a retransmission timer (RTO) of 500ms (by default the retransmission timer is set to (3) slow clock ticks, or 1.5ms total)

--Celerra's Fast Clock is based on a retransmission timer of (1) slow clock tick, or 500ms (Windows' Hosts=200ms)

#### **Celerra Fast Retransit Mechanism: NAS 2.2.53.5 +**

/nas/site/slot\_param [Changing this value will trigger fast RTO whenever 3 or more consecutive ACKs received out of order & changes slow retansmit clock to 200ms]

**param tcp fastRTO=1**

**Note:** Setting this parameter may have the effect of increasing response times for the Celerra when experiencing network packet loss--however, it serves to mask network problems and may actually degrade the overall network with increased retransmission traffic. Celerra doesn't yet use SACK (Selective Acknowledgements).

#### **CELLERRA RPC RETRANSMISSIONS (NFS over UDP):**

Layer 5 retranmissions at the RPC layer using NFS over UDP. Typical Linux RPC timeouts are now 35ms. Goal is to have no greater than .01% retranmissions. Entire Request or Reply is retransmitted. If Client performing Reads, client must retransmit Request and DM must retransmit the Reply data. If Client performing Writes, client must retransmit Request data write and DM must retransmit the Reply.

**# server\_nfsstat server\_2 -rpc**

server\_2 :

Server rpc:

| ncalls     | nBadRpcData | nDuplicates | nResends | nBadAuths |
|------------|-------------|-------------|----------|-----------|
| 3762239871 | 0           | 27447775    | 149909   | 0         |

nResends indicate number of duplicate RPC requests (XIDs) to the DM and retransmitted responses (i.e., RPC retranmissions)

nDuplicates indicates number of duplicate client retransmission requests while DM still working on request—could indicate slow responses or locking problem

**# nfsstat -c** [Use this on the Client to examine RPC retranmissions]

retrans

6

#### **CELLERRA PACKET LOSS:**

**# server\_netstat server\_2 -i**

Ierror or Oerror columns could indicate packet loss due to corruption

#### **Windows Client:**

C:\>netstat -e

Look at Discards & Errors for Client packet loss corruption

#### **Unix Client:**

# netstat -i

Look for RX-ERR and TX-ERR for evidence of packet loss due to corruption

**Note:** Causes of packet loss can be Duplex mismatch (10/100Mb), Faulty cable, Faulty switch port, Faulty DM or Client NIC

#### **Cisco Switch:**

#sh tech

Look at Duplex settings, for Runts, Giants, and Collisions

### **CELERRA PACKET LOSS DUE TO CONGESTION:**

1. Look for ringfull errors on the DM NIC (output ring buffer overflows and packets are dropped)

```
$ .server_config server_2 -v "ians cge0 stat"
```

```
$ .server_config server_2 -v "bcm cge0 stat"
```

-----Secondary Send Ring-----

DescCnt MaxDescCnt SecRingSize **OutDrops** →This is the key field to examine for ringfull issues

```
0x00000 0x00000000 0x000000800 0x0022bf71
```

2. For excessive DM retransmissions, look at output buffers on Client's switch port

3. For excessive Client retransmissions, look for packet drops in receive buffer of client's switch port or send buffer of DM's switch port

#### **Corrective Actions:**

Upgrade to NAS 5.4 to take advantage of SACK, NewReno, and FACK to reduce RTO's

Mount NFS over TCP vs. UDP

Escalate OutDrops ringfull errors to EE

#### **Try setting following param if DM is retransmitting:**

```
param tcp fastRTO=1
```

Lower “rsize” by 4k decrements on mount option for NFS Clients if DM is retransmitting until retransmits are reduced

For TCP Windows Size for Windows clients, and DM is retransmitting, set Windows size to 32k, then lower in decrements of 4k

#### **Client retransmitting with NFS:**

Reduce mount option wsizze by decrements of 4k. Check using netstat –s for TCP and nfsstat –c for UDP.

## **II. TCP ACKNOWLEDGMENTS & WINDOWS 98/95 CLIENTS:**

Celerra no longer sends an immediate acknowledgment [TCP ACK] for frames received with TCP Push Bit set, causing a 150ms delay before ACK's are returned. By default, this TCP Push Bit was set for previous versions of NAS [2.1 & 2.2]. With introduction of NAS 4.x, Celerra no longer has this feature set by default and must be enabled in the /nas/site/slot\_param file:

**param tcp ackpush=1**

**Note:** Setting this bit should increase copy performance when using Windows 95/98 Clients to Celerra Shares

#### **DEFINITION OF TCP PUSHBIT: [PSH flag]**

When Client sets this Flag [PSH], the Server will respond with an ACK immediately. The Push Bit Flag is usually set whenever the Client has sent all data in his buffer out onto the wire. Push Bit is an efficiency measure by which the Sender is telling Receiver not to wait for anymore data and to offload data up the protocol stack. Sun Solaris has an old bug that if the Server does not immediately reply with an ACK after sending data with PSH Flag set to 1, then it will wait a long time before sending again, by which time the Celerra may timeout. Setting the Push Bit (by Sender) may indicate that its system is running slow due to load.

## **III. INCREASING TCP STREAMS/PCB CACHE VALUES FOR MORE CONCURRENT TCP CONNECTIONS:**

Some Customer sites have very high numbers of Users, which can overwhelm the default "tcp maxStreams" and "tcp pcbCacheSize" values. This can create a situation where Users may begin to have problems connecting to the Celerra, and/or being able to effectively use the connection once established.

**Tcp maxStreams:** Default is now 64k concurrent TCP dialogues taking place at once [Reads/Writes/Logons].

**Tcp pcbCacheSize:** Default buffer size 401 [Other values: 5077; 8209; 9973]

### **HOW TO OBSERVE OPEN TCP STREAMS ON A DATAMOVER:**

Method 1:      \$server\_netstat server\_2 | wc -l                [Outputs number of active connections]

Method 2:      \$.server\_config server\_2 -v "tcp stat"        [Subtract 'Connections Closed' from 'Connections Established']

Method 3:      \$.server\_config server\_2 -v "param tcp"        [Viewing Open TCP Streams on a DataMover—"tcp.maxStreams"]

### **OBSERVING PROTOCOL STATISTICS ON DATA MOVER (retransmissions, etc.):**

**\$ server\_netstat server\_2 -s** [Use to verify DM TCP retransmission rate]

ip:

\*\*\*

```
219026 total packets received-----abridged
```

icmp:

\*\*\*\*

```
echo reply: 174-----abridged
```

tcp:

\*\*\*\*

```
298483227 packets sent
```

**1402 data packets retransmitted** →TCP retransmits

0 resets  
273079526 packets received  
108 connection requests  
36 connections lingered  
udp:  
\*\*\*\*  
0 incomplete headers  
32211 bad ports  
56722904 input packets delivered  
56679958 packets sent

### **PARAMETER TWEAKING:**

**Example:** Changing parameters to allow for increased "fs" memory caching for users pertaining to the "NFS" parameter called "openfiles". Allowing for 16,000 simultaneous User TCP connections and increased "openfiles" capability at two files each:

#### **TCP Protocol Parameters for Celerra:**

##### **/nas/site/slot\_param**

param tcp maxStreams=16384 [Add this value to slot\_param]

#### **VALUES FOR MAXSTREAMS AND OTHER PARAMS—TO BE SET CONCURRENTLY:**

**param tcppcbCacheSize=16993** Hex Value=0x00004261 Decimal Value=16993 (param no longer used)

**param tcppcbMaxCacheSize=39839** Hex Value= 0x00009b9f Decimal Value=39839 (param no longer used)

**param tcp.maxStreams=16384** Hex Value=0x00004000 Decimal Value=16384 (Default now 65535)

#### **NFS Protocol Parameters for Celerra:**

##### **/nas/server/slot\_x/netd**

nfs start openfiles=15360 nfsd=96 [default Celerra values]

#### **UFS/CIFS File System Parameters for Celerra:**

##### **/nas/server/slot\_x/file**

**file initialize nodes=65536 dnlc=262144** [default Celerra values; Change to nodes=49152 in this example]

### **DATAMOVER SPEED/DUPLEX SETTINGS:**

**#server\_sysconfig server\_2 -pci ana0 -o speed=100** {10|100|auto}

**Note:** Eventhough Celerra autonegotiates speed settings, it is recommended that the settings be manually set!

**Deleting Speed & Duplexing Manually:** vi /nas/server/slot\_x/sysconfig {sysconfigtab} Delete speed & duplexing entries

**CHANGING DATAMOVER DUPLEX/SPEED:** Orange Light on NIC = Half Duplex [10MB]; Green Light = Full Duplex [100MB]

**#server\_sysconfig server\_2 -pci ana0 -o duplex=half {full|half|auto}**

Or, directly edit /nas/server/slot\_2/sysconfig and sysconfigtab files [Remove references to Speed and Duplexing]

**Note:** Eventhough Celerra autonegotiates duplex settings, it is recommended that the settings be hard-coded

### **AUTO NEGOTIATION AND ETHERNET:**

Autonegotiation is a system that takes control of the cable when NIC devices are being initialized, and uses 16-bit word FLP (Fast Link Pulse) to detect modes of other devices on the other end of the wire (Link Partner), while advertising its own capabilities. What this does is allow devices to detect and choose the best connection type to use between nodes. Control of cable is relinquished after the negotiation is completed.

#### **APPLICABLE ETHERNET TYPES:**

10-BaseT; 10BaseT Full Duplex; 100BaseTX; 100BaseTX Full Duplex; 100BaseT4

#### **AUTO SENSE:**

Older technology that is passive only, and devices do not send FLP to detect or communicate with Link Partners

**CELERRA PACKET REFLECT (aka IP Reflect):** Enabled by default with 4.0.12.1 [Supported with Linux 2.2 NAS]

→Packet Reflect operates at TCP layer and ensures that inbound traffic to a specific port is replied to from the same port

→Packet Reflect does not allow routing of packets across interfaces (Routing Tables are not used when Packet Reflect is enabled)

→Packet Reflect is on by default

→Incoming & Outgoing ports are always the same

→When disabled, outgoing traffic will be determined by IP Routing Tables

→Routing Tables are used for all data mover initiated outbound traffic, such as DNS resolution requests, etc.

**Note:** In most cases, Network packets will go out via the same NIC port that they were received on because the Source & Destination MAC Addresses and VLAN ID are saved and passed up the protocol stack to the application layer. In the case of TCP, MAC Addresses of every packet are passed up to TCP, saved, and then sent back down to the same MAC address. See exception below.

#### **DISABLING IP REFLECT:**

\$server\_config server\_x -v “param ip” [Verify current settings]

**param ip reflect=0**

\$vi /nas/site/slot\_param **param ip reflect=0**

#### **BENEFITS OF USING PACKET REFLECT:**

--supports PVLANS on Celerra [replies will go out same VLAN as received if VLAN Tagged]

--Better network security

--speeds up router failover, example HSRP environments

--Especially useful in environments where Customers want multiple Routes defined and want to turn RIP off on the DMs—used in place of multiple default Routes

--Yields better performance by cutting down on number of Route & ARP lookups

--Works with Clients having single IP Address and Multiple MACs

#### **How IP Reflect Works:**

UDP→Source & Target MAC Address & VLAN ID are passed up to Application Layer. Application Layer saves and sends back down for transmission out original port [good example of application would be NFS].

TCP→Source & Target MAC Addresses & VLAN ID are saved in TCP connection data structure [which is then used when sending packets back out from the same TCP connection]

#### **USING PACKET REFLECT:**

/nas/server/slot\_2/param param ip reflect=0 [disables packet reflect]

**/nas/server/slot\_2/param param ip reflect=1** [default value and enables packet reflect feature]

ip.reflect 0x00fd6184 0x00000001 0x00000001

#### **PACKET REFLECT & NIC TEAMING:**

In general, the use of the IP Reflect feature & NIC Teaming at the same time is not encouraged and may lead to inconsistent communication on the Data Mover, especially if being used on AD or DNS Servers. What can happen when IP Reflect is configured is that the Data Mover will not use local ARP tables and will instead transmit to the NIC-teamed server using the underlying physical MAC addresses and not the virtual MAC address for the team. Disabling IP Reflect takes care of this issue.

In general, the use of NIC Teaming on network Servers where the Data Mover co-exists will work, but in certain cases NIC Teaming could be incompatible with the use of IP Reflect, and IP Reflect would have to be disabled. There is no bug here, it's just an interop issue between two different feature sets. On the Microsoft side, they pretty much say that NIC Teaming is a function of the various hardware components involved: Servers, NICs, Switches

If MS is engaged for a case that involves NIC Teaming, their pat requirement would be to disable the teaming to see if this resolved the issue. If it did resolve the issue, they would say go talk to the hardware vendors involved, etc. In general MS recommends that you only use NIC Teaming for externally presented NICs that do not transmit traffic to private intranets, go through firewalls, or are routed to NAT devices. In otherwords, NIC Teams should be used for a single network presentation.

For the Celerra, NIC Teaming is something that we have been doing for a long time now, more commonly known as EtherChannel or Link Aggregation. From my understanding, NIC Teaming can provide for Fault Tolerance (High Availability), and maybe Load Balancing, depending on the vendor implementation.

#### **SITUATION WHERE IP REFLECT WILL NOT ALWAYS BE SENT OUT SAME “PHSYICAL” PORT AS RECEIVED:**

With Celerra EtherChannel and Packet Reflect, if the DM is using two ana ports configured in an EtherChannel “Trk0” where both interfaces share a common IP address, in certain situations, packets received via a switch will enter the “logical” network on “TRK0” via either ana0 or ana4. In this situation, Packet Reflect only knows how to reply via the “logical” port TRK0. An algorithm on the data mover determines which “physical port” to send reply out, meaning that it could go out either ana0 or ana4 for any given received packet. All that IP Reflect can do in this instance is to ensure that the packet goes out the same logical “network”, which in this example is TRK0. So, IP Reflect cannot guarantee reply via same physical port, just logical port.

**Note:** Both ports in an EtherChannel configuration are active

**Rule:** IP Reflect can only reply using the “logically” defined trunk, TRK0 in the above example, not physical port.

#### **TROUBLESHOOTING TCP/IP CHECKLIST:**

1. ARP [ARP IP address to MAC address resolution at physical layer]
2. Hostnames [nodenames]
3. Ipconfig [Windows IP configuration, WINS, DNS, DHCP information]
4. Nbtstat -c [shows contents of netbios cache] -n [shows locally registered names] -S [Netbios sessions]

5. Netdiag /fix [diagnose network problems and DNS issues] /v /l /DCAccountEnum
6. Netstat -a -s -p -r [Network protocol statistics & TCP connections]
7. NSLookup [Useful in verifying DNS name resolution on a network & Win2k DNS SRV resource records]
8. PathPing [combination Ping & Tracert tool]
9. Ping/Route/Tracert [default ICMP echo request TTL is 32 hops]

### **PDUs/SDUs—Protocol Data Units & Service Data Units:**

Protocol Data Units operate at each Layer in the OSI model, and are used to encapsulate data and protocol headers for transmission to the next layer. So, PDUs from one Layer become SDU's when received at the next layer, and so on throughout the process.

### **DATA ENCAPSULATION DOWN THE OSI LAYERS:**

→TCP receives SDU's from upper layer protocols, adds its headers to encapsulate and create Layer 4 PDU's called "segments" that are passed down to the IP Layer, which receives the segments as SDUs, which in turn are encapsulated into Layer 3 PDUs called IP packets or IP Datagrams, and passed to Layer 2 Ethernet, which receives Datagrams as SDUs, and in turn are encapsulated into Layer 2 PDUs called "Ethernet frames", which are passed to Layer 1 for transmission as electronic bits over the physical media.

### **DATA ENCAPSULATION UP THE OSI STACK:**

→Basically, the process of encapsulation is reversed. The Ethernet device inspects the Layer 2 PDU Ethernet frames, removes the Layer 2 SDU IP Datagram, and passes to IP as Layer 3 PDU. IP Layer 3 removes the Layer 3 SDU TCP Segment and passes to TCP Layer 4 as a Layer 4 PDU. TCP continues to process up the protocol stack.

### **APPLICATION LAYER—Layer 5:**

Protocols such as NFS, FTP, DNS, SNMP, Telnet communicate with Layer 4 through use of Port numbers  
Data Containers in this layer are called 'messages'

### **TRANSPORT LAYER—Layer 4:**

TCP & UDP protocols—data containers in this layer are referred to as "datagrams" or "segments"

→TCP is connection-oriented, full duplex, contains error checking, sequencing of packets, uses flow control and acknowledgements, and contains packet recovery services for retransmission of lost packets

→UDP is connectionless, unreliable, non-guaranteed transport, used especially in broadcast or multicast deliveries

→TCP encapsulates data from higher-level protocols into "segments", which are then passed down as TCP PDU's (Protocol Data Units), and passed to the IP protocol, Layer 3.

### **TROUBLESHOOTING UDP TRANSPORT PROTOCOL:**

#### **UDP PROTOCOL:**

Protocols that use UDP are TFTP, SNMP, NFS, DNS

NFS→RPC→UDP→IP

--UDP is a datagram transport protocol that outputs a single UDP datagram per transmission, as received from Applications via the IP Layer. Depending on MTU, etc., IP datagrams may need to be fragmented into separate packets. This allows for the out of order delivery of fragments—problem arises when one of the fragments is lost, the whole datagram must then be retransmitted. UDP header is always 64 bits.

--RPC Calls on behalf of NFS use Transaction IDs (XID) and is registered with Portmapper process on Server

--Much different than TCP—UDP has no concept of lost packets and retransmissions

--Applications enforce lost packets, and when this happens UDP must retransmit the whole datagram, which could be many IP Packets, unlike TCP which performs timeout and retransmissions of individual packets

--NFS typically uses UDP to mount remote filesystems

--RPC (NFS) keeps track of timeout values and if they expire, will require that UDP retransmit

**Note:** Unlike TCP, which uses ACKs to indicate packet loss, UDP uses silence

#### **UDP REASSEMBLY QUEUE:**

UDP uses unique IDs and Offset values to reassemble inbound IP fragments from its 'reassembly queues'—by default, the Queue is set to expire after 60 secs, meaning that an IP Datagram would have to be completely retransmitted.

#### **TIPS WHEN USING UDP:**

**Comment:** On clean networks, UDP has the potential to deliver more payload with less overhead than TCP. However, from experience, any dropped 'packets' on the network can significantly slowdown delivery of data when using UDP since the complete "datagram" must be retransmitted if any single "packet" is dropped or lost. TCP, on the otherhand, only needs to re-transmit a single lost "packet" and can therefore handle dirty networks by providing better overall throughput.

--Because of the above statement, it is usually always recommended to "mount" NFS File Systems for TCP rather than UDP.

--But, if NFS over UDP is required, it may be better to set an rsize/wsize of 8k or lower for the NFS mounts

### **IP NETWORK LAYER 3→IP PROTOCOL:**

- IP protocol receives PDU's from TCP, and encapsulates them as Layer 3 SDU's (Service Data Units) called “packets or datagrams”
- IP provides routing of interconnected packet-switched networks
- IP provides for unreliable, connectionless datagram delivery service
- IP provides for data transfer via datagrams, addressing via 32-bit IP Addresses, routing, & fragmentation of datagrams into packets
- All IP fragments carry same ID as original packet
- ICMP, ARP, DHCP also work at this layer
- IP uses protocol field to determine which application protocol to route datagrams to [e.g., 6=TCP; 17=UDP, etc]
- IP datagrams are used for both TCP and UDP

#### **CHARACTERISTICS OF IP PROTOCOL:**

- Packets are treated independently between sender and receiver
- Packet delivery is not guaranteed
- Lost or corrupted packets are not recovered, except by other layers

#### **ICMP PROTOCOL:**

Announces network errors, network congestion, assists in network troubleshooting with ping echo, announces timeouts

### **DATALINK ETHERNET LAYER 2:**

Data containers at this layer encapsulate packets into “Ethernet frames”, turned into bit streams at the physical layer

### **PHYSICAL LAYER 1:**

Receives Ethernet Frames from Layer 2 and converts into bits for transmission on the wire

#### **NFS NETWORK LAYERS:**

NFS v2 or v3 uses RPC Client-Server calls over TCP/UDP, IP, Ethernet

#### **CIFS NETWORK LAYERS:**

RPC, Pipe Protocol, SMB, NetBIOS, TCP, IP, Ethernet

#### **NETWORK LAYERS:**

|                  |     |                                                                                                                                                                           |
|------------------|-----|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Layer 7 NFS      | PID | Application Layer [Unix NFS, Mount, NIS; Windows WINS, Login, SMB]                                                                                                        |
| Layer 6          |     | Presentation Layer [XDR for NFS—External Data Representation Standard]                                                                                                    |
| Layer 5 RPC      | XID | [Client & Server requests are tied together by unique transaction XID numbers]                                                                                            |
| Layer 4 TCP      |     | Transport Layer--Port Sequence Numbers and Acknowledgments                                                                                                                |
| Layer 3 IP       |     | Network Layer--IP Addresses—IP protocol ‘interconnects’ packet-switched networks                                                                                          |
| Layer 2 DLC      |     | Data Link Layer--MAC Addresses & Frames—Ethernet transmits & receives data packets, decodes packets; detects errors; method behind physical delivery of packets to a Host |
| Layer 1 Physical |     | Media & bits [This is where bad cables & mismatched duplexing manifest themselves]<br>NICs use MDI pinout scheme, while Switches & Routers use MDIX pinouts               |

### **ROUTERS, SWITCHES, HUBS:**

**Hubs:** By definition Hubs run at Half-Duplex—if you have a Hub involved with Celerra, must force the DM's to talk Half Duplex. Why? Because a switch will automatically negotiate at Full Duplex to the Hub and pass along to Celerra at Full Duplex—over time, the switch buffers will overflow and cause poor performance, possibly failure. Half-duplex interfaces transmit on a transmit pair while listening to receive traffic on a receive pair. Receive pair is where Collision Detection and Avoidance takes place. Full Duplex does not need to do this.

**Switches:** Layer 2 device that builds tables based on MAC addresses and port received from [Configure VLAN's on switches]

- show config: Displays switch configuration
- show vlan: Displays VLAN assignments—confirm that switch ports 3/1-16 are in correct vlans
- show port: Verify Port connection status; Duplex negotiation; speed negotiation
- show trunk: Displays trunking info; a)Mode set to On b)Encapsulation c)verify vlans memberships in trunk

**Basic Troubleshooting:** Check Port Light Negotiation and VLAN configurations

**Cisco 2900 Layer 3 Switch:** >enable [type “enable” at prompt to log into Privileged mode on a switch or router]

**Looking at settings:** sh config; sh int fast6

Run from Privileged Mode: Config t>interface fast6>speed auto [duplex auto] or speed full [duplex full] ctrl + Z to end session

#### **DART & SWITCH PROTOCOLS:**

DART understands only (3) Switch protocols; Ethernet Trunking; ARP; RIP

### **FULL/HALF DUPLEXING; AUTOSENSING/AUTONEGOTIATION:** #server\_sysconfig server\_2

Rule: Celerra must match settings of Switches!! [Celerra to Switch; Switch to Host; Switch to Switch—ALL must be same!!]

Switch              DataMover

|                 |                 |
|-----------------|-----------------|
| Auto            | Auto            |
| Fixed 10/100    | Fixed 10/100    |
| Fixed Half/Full | Fixed Half/Full |

**Performance Issues Related to Hubs, Switches, Celerras:** Speed and Duplex settings; or an Overloaded device

### **CISCO SWITCH COMMANDS:**

|                                                                    |                                                                                      |
|--------------------------------------------------------------------|--------------------------------------------------------------------------------------|
| #set port channel 5/17 mode on                                     | [Sets EtherChannel Mode On for the Switch]                                           |
| #show port capabilities 5/17                                       | [Shows capabilities of the port indicated on the switch]                             |
| #show port channel                                                 | Channel Mode On=Cisco Etherchannel=Celerra “trunking” ADMIN GROUP 12 Channel ID 777] |
| <b>Note:</b> This is our basic Celerra Trunking mode of operation! |                                                                                      |
| #show port 5/17                                                    | [VLAN 2 and other information displayed]                                             |

### **QUALIFIED SWITCHES FOR USE WITH CELERRA:** Others can be used with Product Qualifier Statement

**Ethernet 10/100 Base T:** RJ45 connectors and manual configuration of Speed and Duplexing recommended

Cisco Catalyst 5000/5500/6500; Alantec; FORE Powerhub 7000; Cabletron MMAC; SMC TigerSwitch 100; Bay Networks Baystack Hub; 3COM Linkswitch 3000; LANPlex 2500; Linkbuilder FMS100

**Gigabit Ethernet:** SC Connectors 62.5 or 50 micron cables with Jumbo Packet support

Alteon; Cisco 6500; Foundry BigIron 4000; HP ProCurve 9304

**FDDI Switches:** Uses SC connectors and is Half Duplex

DEC GIGAswitch FDDI; DECconcentrator 900/500MX; Synoptics 2914; Cableton MMAC Plus; 3COM Fiber Concentrator; Xylan Omniswitch; Cisco 5000/5500; Fibronics GigaHub

**ATM Switches:** SC connectors with LANE V1.0 emulation

Fore Systems ASX200/1000; Cisco LightStream 1010; Bay Network Centillion 100

### **CLIENT OPERATING SYSTEM SUPPORT:**

Compaq Digital UNIX Tru64; HP-UX; IBM AIX; LINUX; Silicon Graphics IRIX; SUN Solaris; Windows NT/2000

### **Routers:**

Layer 3 device designed to deliver packets from Layer 3 to Layer 2 MAC Addresses

RIP Protocol=Distance Vector      OPSF Protocol=Link State      Hybrid Protocol=IGRP [Cisco]

Router Access Lists: Standard, Extended, Deny, and Filters—all based on Protocols, Port Numbers, and/or IP Addresses.

**#show run** → Verify Encapsulation; Vlan assignments; IP Address & Mask; Access-Group assignments; That no IP redirects or Shutdowns listed

**#show route** → Confirm destination network has a defined route

**#show access-list 101** → Examine access-lists as necessary

**#show ip route; show init; show encaps** → Trunking protocol in use]; show config; show port 4;

**Note:** Trunking Protocols are set up on Routers, not Switches! VLANS are set up on Switches, not Routers!

**Basic Troubleshooting:** Check Access Lists; Routing Tables; Encapsulation protocol

**Blackhole Router:** Troubleshoot with ping command: #ping -f -i 50 [MTU size—vary size until no reply]

**Cisco 5500:** >enable>set port speed 4/1 auto    >set port duplex 4/1 full

### **TROUBLESHOOTING CELERRA NETWORKING:**

**Server\_route Command:** Routing Table entries. -f to flush entries -l to list -d to delete a route -D to delete all entries -a to add \$server\_route server\_4 -flush      \$server\_route server\_4 -list      \$server\_route server\_4 -add 193.1.21.200 193.1.21.254

**Adding DataMover Default Gateway Command:**

**#server\_route server\_2 -add default 10.127.15.238**

**Viewing a Routing Table:** #server\_route server\_2 -l

### **CELERRA ROUTING TABLES/DEFAULT GATEWAY:**

Ifconfig command adds IP entry for interface and takes the assigned IP address and netmask assigned to determine the Destination Network address seen in the Route Table

Default Routes can be assigned as Static entries and are always checked last in the Route Table before routing a packet  
packets are delivered in Ethernet network via MAC hardware addresses

### **Dynamic Routing:**

RIP Protocol runs by default on the Celerra

Disable RIP in order to build Static Routing Tables using the “server\_route -add” command

### **ICMP REDIRECTS:**

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
If RIP is disabled on DM, & Routing Tables are Static, ICMP Redirects have the capability of changing Routing Tables on the Receiving end & causing problems.

## **VERIFYING DM ICMP STATISTICS & ROUTE REDIRECTS:**

**\$server\_config server\_x -v “icmp stat”**

```
Source quench      0/0  
Redirect route     0/0  
Echo request      143654/0
```

-----abridged-----

**SERVER NETSTAT UTILITY:** Network Statistics for IP, ICMP, TCP, UDP [also, use to verify NIC transmitting!]

-a lists TCP & UDP sockets      -i state of interfaces      -r Routing Table      -s Protocol stats      -p tcpludplip

**Note:** Input errors could indicate network cabling problems, bad CRC checksum errors from failing NIC's, etc.

Output errors may indicate host problem or flooded network. Sustained Errors >than 5% would indicate a problem.

```
$server_netstat server_4 -a          [active TCP/UDP connections]      $server_netstat server_4 -s [stats for all protocols]
```

```
$server_netstat server_4 -s -p tcp    [udp/ip protocol stats]      $server_netstat server_4 -i [interface info]
```

```
$server_netstat server_4 -r          [routing table statistics]
```

## **DISPLAYING NETWORK STATISTICS FOR IPv4:**

**# server\_netstat server\_2 -A inet -a**

```
Proto Local Address                    Foreign Address        (state)  
*****  
tcp 128.221.252.2.mac            128.221.252.100.990 ESTABLISHED  
tcp *.portmapper                    *.*                    LISTEN  
tcp *.mount                        *.*                    LISTEN  
tcp *.nfs                         *.*                    LISTEN  
tcp *.Network-Block-Service        *.*                    LISTEN  
tcp *.Replication-Services        *.*                    LISTEN  
tcp *.Replication-Services        *.*                    LISTEN  
tcp *.Replication-Services        *.*                    LISTEN  
tcp *.5085                        *.*                    LISTEN  
tcp *.7777                        *.*                    LISTEN  
tcp *.RCP                        *.*                    LISTEN  
tcp *.NDMP                        *.*                    LISTEN  
tcp *.usermapper                 *.*                    LISTEN  
tcp *.xattrp                      *.*                    LISTEN  
tcp *.rquotad                    *.*                    LISTEN  
tcp *.statd                      *.*                    LISTEN  
tcp *.lockd                      *.*                    LISTEN  
tcp *.SNMP                        *.*                    LISTEN  
tcp *.pax                        *.*                    LISTEN  
tcp *.mac                        *.*                    LISTEN  
tcp 128.221.252.2.3081          *.*                    LISTEN  
tcp 128.221.252.2.5082          *.*                    LISTEN  
tcp 128.221.253.2.3081          *.*                    LISTEN  
tcp 128.221.253.2.5082          *.*                    LISTEN  
Proto Local Address  
*****  
udp *.SNMP  
udp *.Routing-Information-Protocol  
udp *.4646  
udp *.lockd-forward  
udp *.pax  
udp *.RFA  
udp *.mac  
udp *.56790
```

## **ANALYZING NETSTAT OUTPUT:**

**Network Collision Rate:** Divide Number of Output Collisions by Number of Output Packets [greater than 10% is a problem]

**Input Packet Error Rate:** Divide Number of Input Errors by Number of Input Packets [> than 25% could indicate host problem]

**Statistic Commands:** #server\_netstat server\_2 -a -i -r -s [tcp/udp/lip stats] #server\_netstat server\_2 -s [stats on all protocols]  
#server\_nfsstat ALL -rpc -nfs #server\_sysstat ALL {Key stats on DataMover activity} #server\_nfsstat ALL -z -n

**Netstat:** Should have less than 10% errors on a NIC or else could indicate a problem!

**Netstat -e** command will show network statistics, especially if there are a lot of collisions [packet loss]

### Linux Interface Statistics: #netstat -i

### OBSERVING PORTS AND SERVICES CONNECTED ON CONTROL STATION:

\$ netstat -alp

### ADDRESS RESOLUTION PROTOCOL—ARP:

**Purpose:** To resolve a node's Layer 3 IP Address to its Layer 2 Data Link MAC, hardware, or Network Interface Address.

--ARP sends ARP Request message to request MAC address for a forwarding IP address via MAC-level broadcast

--Receives ARP Reply

--Default ARP Cache TTL for Windows 2000 is 2 minutes, with a max lifetime of 10 minutes

**Gratuitous ARP:** Method for detecting duplicate IP addresses. Node sends an ARP Request for own IP address—if no ARP reply is received, then assumption is made that its IP address is o.k.

**Note:** If another Host shares same IP address, then ARP cache might have MAC address of other computer, or vice versa, leading to ARP resolution problems.

**Proxy ARP:** Answering ARP Requests on behalf of another node, commonly used in Routing & Remote Access services.

### CELLERRA ARP ISSUE:

NAS 5.1.15.3 fixes gratuitous ARP issue by providing for a new ARP engine that will issue gratuitous arps every 10 seconds for up to 2 minutes. Purpose is to allow time for multiple switches to be properly updated with new ARP information from Celerra.

### ARP PARAMETERS:

**param ip ifAdvertiseDuration=119** [Defines how long DART sends advertisements for ARP--default is 119 secs in length.  
If set to 0, DART sends only a single ARP message when the Interface is started—previous code behavior]

**param ip ifAdvertiseFrequency=10** [Defines how frequently advertisements are sent—default=10 secs. Gratuitous arp is advertised every 1-10 secs for the 2 minute window]

**ARP TABLES:** #server\_arp; arp -a; arp -d 10.127.15.229; arp -s 10.127.15.227 macaddress or arp -d\* [delete all]  
Arp maps layer 3 IP addresses to Layer 2 MAC addresses [Reverse ARP maps Hostname to IP Address]

**Default ARP Cache Value:** 10 minutes most systems, 5 minutes for DM

\$server\_arp server\_4 193.1.21.210 [Arp cache from IP Address] \$server\_arp server\_4 -a [Arp cache locally]

\$server\_arp server\_4 -s 193.1.21.210 macaddress [adding entry to ARP] \$server\_arp server\_4 -d 193.1.21.210 [deleting entry]

**Server arp Command:** -a Displays ARP tables {IP to MAC address resolution} -d delete entry -s set static entry

**Note:** Default ARP Cache timeout = 5 minutes. View ARP Cache to see entries for other systems.

# server\_arp server\_13 -all

server\_13 :

192.168.2.100 at 0:4:76:30:c:6a

192.168.1.100 at 0:1:2:ee:74:1f

10.10.28.254 at 0:0:c:7:ac:1c

### DUPLICATE IP ADDRESS ISSUE:

**Symptom:** Customer adds new network interface to Celerra and receives “Network path not found” error

**Cause:** Duplicate IP Address

### Troubleshooting Duplicate IP for Celerra:

→Ping from Client to DM IP

→Do arp server\_x and note MAC address of DM

→Do server\_ifconfig on DM and compare MAC addresses [If different, then there is a duplicate IP address on network]

**Network Commands:** #server\_arp #server\_dns #server\_ifconfig #server\_nis #server\_ping #server\_route ALL -l [lists out the DM routing table] #server\_snmp Network Connections, Info, & Ports Listening: #server\_netstat server\_x -a [-i]

#server\_ifconfig server\_2 ana0 [parameters for specific interface] #server\_ifconfig server\_2 ana0 up [enable/disable interface]

#server\_netstat server\_2 -i [interface info] #server\_netstat server\_2 -r [server routing tables]

### Server Sysconfig Command: Displays Interface Information about DataMovers

Device name, slot, port, IRQ, Processor, RAM, Motherbd, Bus speed, etc

-pci : Display or configure NIC's for duplex speed -o {speed=10|100 default} duplex=full|half default

/nas/bin/server\_sysconfig ALL -P [shows server configurations]

## **DISABLING RIP PROTOCOL ON DM:**

#vi /nas/server/slot\_2/netd [comment out #routed]

## **NETWORK ANALYZERS:**

Solaris Snoop Program; NT Network Monitor; Network Associates SnifferPro 3.5; Ethereal Program

**Sniffer Placement on Network:** Consideration needs to be given as to where the desired traffic capture should be—between Client & Switch; Data Mover and Switch, etc. Consideration also needs to be given as to taking the trace by spanning switch ports or doing the trace in-line.

## **CONFIGURING A SWITCH TO SPAN A PORT:**

1. #set span 1/1 1/3 [All traffic from port 1/1 will be sent to 1/3]

## **NETWORK BROADCASTS:**

Two Types: Layer 2 and Layer 3

LAYER 2 BROADCASTS: Conducted by applications such as ARP and system discover

LAYER 3 BROADCASTS: Conducted by NetBIOS, etc. [Most common]

Most networks should have no more than 50-100 Broadcast packets per second to be healthy

## **IPv4 CLASSES:** 32-bit values, up to 4 billion addresses, Layer 3 in OSI model

→interfaces with Layer 4 [TCP & UDP] via the Protocol ID field in the packet header

→Autoconfiguration through use of DHCP

→Handles fragmentation

→QOS and Security functions not really useful or used

→Use of broadcasts [ARP, RIP, etc.]

→ICMPv4 ping for Error and Informational messaging

## **IPv4 Enhancements:**

→Classless addressing [CIDR—Classless Interdomain Routing]

**Note:** Network ID indicated by leftmost high order bits, usually written like /24, /30, etc., requires use of subnetting

→NAT

## **RFC 1918 Establishes Private Network Addresses:**

10.0.0.0 – 10.255.255.255

172.16.0.0 – 172.31.255.255

192.168.0.0 – 192.168.255.255

**Note:** These private address spaces are NOT maintained by Internet Routers—hence “private”

| Class | Default Mask  | Network Range     | # Host Bits | #Network Bits        | #Networks | #Hosts per Network |
|-------|---------------|-------------------|-------------|----------------------|-----------|--------------------|
| A     | 255.0.0.0     | 1-126.255.255     | 24 bits     | 8 bits               | 126       | 16,777,214         |
| B     | 255.255.0.0   | 128.0-191.255.255 | 16 bits     | 16 bits              | 65,000    | 65,534             |
| C     | 255.255.255.0 | 192.0-223.255.255 | 8 bits      | 24 bits              | 2,097,152 | 254                |
| D     |               | 225.0-239         |             | [Multicast networks] |           |                    |

## **IP Subnetting Refresher:**

n<sup>n</sup> - 2

### **I. How to Find Subnet Mask Value for a Given Number of Required Subnets**

Example: 172.20.30.40 Given IP Address Problem: Need to support 6 subnets

Class B: 255.255.0.0

**Note:** To provision a Class B network for 6 subnets, requires (3) bits from the Host octet because you count the Value from l-to-r; 2    4    8    16    32    64    128 ....to obtain (8), the minimum number of bits required for (6) subnets.

Subnet Mask= 11111111.11111111.11100000.00000000

Or 128 + 64 + 32 for the (3) bits = 255.255.224.0

### **II. How to Calculate Broadcast Address: Subnet Mask: 11111111.11111111.11100000.00000000**

Note: Take value of 3<sup>rd</sup> bit, which is (32), and that becomes your network boundaries;

172.20.0.0 – 172.20.31.255                      Broadcast Address = 172.20.31.255

172.20.32.0 – 172.20.63.255                      Broadcast Address = 172.20.63.255

172.20.64.0 – 172.20.95.255, etc.                Broadcast Address = 172.20.95.255, etc.

## **Internal IP Address Ranges:**

\* 10.0.0.0 - 10.255.255.255  
\* 172.16.0.0 - 172.31.255.255  
\* 192.168.0.0 - 192.168.255.255

## **IPv6 STANDARD:**

### **IPv6 NETWORKING FEATURES/BENEFITS:**

- Larger public & private address space
- Automatic addressing easier
- Support improved for extensions and options
- Uses QOS in header for Traffic Class and Flow Label [New, allows for prioritization & filtering of traffic using RSVP—Resource Reservation Protocol]  
**Note:** Flow Labeling is one-way directional data transfer and is identified using Source & Destination addresses, and Flow Label. Flow Labeling is really just a special handling indicator for Routers. DiffServ for internal network priority schemes [DSCP—Differentiated Services Code Point]. DiffServ is class-based approach, while IntServ is flow-based approach to QOS
- Authentication & Privacy capabilities
- No more broadcasts as in IPv4—instead, uses multicasting and anycast addressing [1:1:many]
- Security built-in with IPSEC
- Use of scopes based on topology or geography for communications [Link local, subnets, Global]
- Subnetting is done using CIDR rules [Classless Inter-Domain Registration] with network prefix notation, NOT subnet masks
- IPv6 supports Stateful DHCP and Stateless (no DHCP) address configuration

**Note:** With Stateless configuration, Hosts auto configure IP addresses on local link and globally unique prefixes advertised by local routers

## **IPv6 HEADER:**

- (8) fields fixed at 40 bytes, using Version (4-bits), Traffic Class (8-bits), Flow Label (20-bits), Payload Length (16-bits), Next Header (8-bits), Hop Limit (8-bits), Source Address (128-bits), Destination Address (128-bits)
- Note:** Traffic Class of 0 means best-effort. Anything else means routers treat with preference. Next Header identifies the next header after the IPv6 header—replaces Protocol field. Hop Limit max. is 255 and packets discarded if decremented to zero—replaces TTL

## **IPv6 PACKET TERMINOLOGY:**

Frame consists of Layer 2 (Ethernet), Layer 3 (IPv6), Layer 4 (TCP/UDP), Layer 7 (Applications), and Data fields

Packet/Datagram consists only of Layer 3 – Data fields

Segment consists of Layer 4 – Data fields

Link MTU is Maximum Transmission Unit, packet size in octets, that can be conveyed over a link

Path MTU is the minimum MTU to all links between Source & Destination systems

## **IPv6 EXTENSION HEADERS:**

Optional Layer 3 headers placed between IPv6 and Upper Layer headers in a packet, multiples of (8) octets (bytes)

Packets can contain multiple extension headers

Extension Headers are used for Hop-by-Hop, Routing, Fragment, Destination, Mobility, Authentication, & ESP options

### **Hop-by-hop Extension Header:**

Option used to inform Routers of special QOS processing or MLD messages, or Jumbo payload attached

Indicated to Router by Next Header=0 flag in IPv6 Header

Immediately follows the IPv6 Header, and is examined by all Routers

### **Routing Extension Header:**

Indicated by value of 43 in Next Header of preceding header

Used as Loose Source Route option when one or more intermediate nodes are desired for routing

Lists addresses to be visited

### **Fragment Extension Header:**

IPv6 Routers are not allowed to fragment packets, and IPv6 Header cannot be fragmented

Source system can fragment large packets using the Fragment Header option, multiples of (8) bytes

Receiving system reassembles fragments in correct order at Layer 3

### **Destination Extension Header:**

Processed on destination system at upper-layer, though each Router along the way inspects

## **IPv4 HEADER:**

- Could be variable in length, 20-60 bytes, using (14) fields: Version, IHL, TOS, Total Length, ID, O, DF, MF, Fragment Offset, TTL, Protocol, Hdr Checksum, Source Address, Destination Address

## **CONFIGURING IPv6 ON WINDOWS CLIENTS:**

XP requires use of CLI utility called netsh c:>netsh interface ipv6 install [Installs the IPv6 stack]

Vista offers a GUI for configuration

## **IPv6 ADDRESS:** 128 bits in length

→ An IPv6 Address is 128 bits, represented by 16 octets (aka, hex pairs) totaling 32 hex characters [see RFC 2460]. Visually, the address is segmented into (8) fields of 4 Hex characters separated by colons, and shortened per accepted abbreviations.

→ Colons are used to separate groups of 16 bits, totaling (8) sections

**2001:030b:0000:0000:0004:6c2a:008d**

**2001:30b::4:6c2a:8d** → Example of acceptable abbreviation

### **IPv6 Addressing Abbreviation Rules:**

\* Successive fields of all zeroes can be indicated by double colon ::, which can be used at beginning, middle, or end of address, but only used once within the address

\*\* Beginning zeroes of any field can be dropped—they are implicitly known

\*\*\* Never drop trailing zeroes from any field

→ IPv6 addresses are based on Interfaces, not Nodes

→ All IPv6 addresses have a specific scope built-in, such as Link-local, Enterprise, and Global, based on topology or geography

→ IPv6 addresses are written in Hex, with each hex character representing 4-bits

### **CONFIGURING IPv6 NODES:**

#### **Stateless:**

→ autoconfig with Router providing the network prefix & subnet via ICMPv6 Router Advertisements, with host creating Host ID from MAC address

→ Router might also reply to use DHCP for other configuration information (such as DNS, Domain Name, etc.)

→ Hosts use ICMPv6 Router Solicitation messages to talk to Routers [sends to FF02::2, hop limit of 255]

→ Routers use ICMPv6 Router Advertisement messages to talk to Hosts

#### **Stateful:**

→ Autoconfig with DHCP Server providing the network prefix for hosts & other required information

**Manual** configuration also possible

→ IPv6 Header has been simplified in the PDU packet [Protocol Data Unit] and now fixed at 40 bytes vs. variable length

**Note 1:** IPv4 header fields IHL, Identification, Fragment offset, Header checksum, Options, all removed

**Note 2:** TTL is now Hop Limit, TOS is Traffic Class, Total Length is Payload Length, Protocol is Next Header

→ IPv6 uses Extension Headers instead of Header Options, which are dynamic, used as needed, processed as needed

### **INTERFACE IDENTIFIERS:**

→ Minimum of two Interface IDs are configured with this method:

1. EUI-64 ID 2. Temporary interface ID, which is a hash of the EUI-64 ID

**Note:** EUI 64-bit Interface ID is created from the interface MAC address, and substitutes a “02” hex as the Universal/Local bit at the beginning of the identifier, followed by Vendor ID in 4 hex numbers, pads the middle with a standard FF-FE, and ends with last 6 hex numbers.

#### **EXAMPLE:**

48-bit MAC Address: 00-03-47-28-45-8F

Converted to EUI 64-bit Identifier: 02-03-47-FF-FE-28-45-8F

→ All interfaces have at least one Link-local unicast address, and usually other addresses as well

### **STANDARD NET PREFIX AND COMPOSITION OF AN IPv6 ADDRESS:**

2000::/3

#### **FULL IPv6 ADDRESS as assigned to ISPs:**

2001:0DB8:0000:0000:0000:0000:0000/48

#### **First 48 High Order Bits Represent Network Prefix:**

2001:db8::/48

#### **Next 16 Bits Represent Subnet ID:**

0000

#### **Last 64 Bits Represent Interface ID:**

0000:0000:0000:0000

### **IPv6 ADDRESS CLASSES:**

Unicast 1:1 → Unique IPv6 interface node

Multicast 1:many → Group of IPv6 interfaces, members of a multicast group

Anycast 1:1 of many → Address assigned to multiple interfaces; Packet sent to anycast address usually goes to nearest interface

### **MANAGING MULTICAST ADDRESSES:**

→ Routers use Multicast Listener Discovery (MLD) to add, advertise, and prune entries

→ MLD replaces the IPv4 functionality of IGMP

→ Uses ICMPv6 as the transport protocol

→ Routers can query a link to see who's present using FF02::1 [Closest thing to IPv4 Broadcast 255.255.255.255]

→ Host Listeners can register with Router

→ Host Listeners should inform Router when leaving the group [FF02::2]

→ Solicited node multicast used in Neighbor Discovery process when only a single Host discovery is desired

## **ICMPv6:**

Used to return Error messages—represented by bits 0-127

Used to return Informational messages—represented by bits 128-255 in Echo Request and Echo Reply messages

Neighbor Discovery protocol replaces IPv4 ARP

An 8-bit ICMPv6 Header is indicated by 58 (0x3A), and follows the IPv6 Header [0-127; 128-255 values]

DAD (Duplicate Address Detection) uses Neighbor Solicitation to ensure that node address is not duplicated

## **DEFINED ICMPv6 ERROR MESSAGE TYPES:**

Destination Unreachable, Type 1 Code 1-6

Packet too big, Type 2 Code 0 [Used as part of the Path MTU discovery process between Interfaces]

Time exceeded, Type 3 Code 0 [Hop Limit Exceeded] or 1 [Packet reassembly time exceeded, from Host]

Parameter problem, Type 4 Code 0-2

## **NEIGHBOR DISCOVERY PROTOCOL:**

→ ND uses ICMPv6 with values 133-137

→ ND is used by IPv6 nodes (hosts & routers) to discover each other's Link-layer addresses of neighbors using multicast

→ ND used by Hosts to locate Routers

→ Used by Hosts to track reachable and non-reachable neighbors, and changed link-layer addresses

Type 133 Router Solicitation

Type 134 Router Advertisement → default timer 200secs

Type 135 Neighbor Solicitation → FF02:0:0:0:0:1:FF00::/104

Type 136 Neighbor Advertisement

Type 137 Redirect [from Router]

## **ND CACHE TYPES:**

### **Destination Cache:**

→ Entries regarding recent destination interfaces

→ Stored on Source hosts as cache maps of destination IP addresses, along with next-hop neighbor

→ Updated by Router Redirect messages

→ Stores Path MTU information for destination addresses

→ Stores link-local addresses of destinations

### **Neighbor Cache:**

→ Entries regarding neighbors, with flag to indicate Router or Host

→ Stores on-link unicast IP addresses and link-layer information (MAC)

→ Information on reachability

# show ipv6 neighbors

c:\>netsh interface ipv6 show destinationcache | neighbors | route

## **IPv6 ADDRESS TYPES:**

### **UNICAST:**

→ Global Unicast Addresses are manually set, and are the addresses used to route over the Internet. 64bits define the Network portion of the address, while 64bits define the Interface portion of the network. Currently, ISPs are only using prefix 2000::/3 addressing for Internet routing

→ A single Unicast Global address identifies only a single Interface

→ A single Interface can have multiple Unicast Global addresses

→ Global scope addresses are internet routable whereas Link-local scope addresses are IDs for the local link/network only

→ Global Unicast addresses usually consist of a Temporary address with same /64 prefix, with a default life of 24hrs, and an Interface ID that is hashed (i.e., unique) from the EUI-64 portion of the MAC address

→ Global unicast addresses are equivalent to public IPv4 addresses

→ Unicast is an identifier for a single interface and is point-to-point

→ Usually, an interface will have multiple unicast addresses (Global & Local unicast, and multicast addresses)

→ Local Unicast addresses (aka, private addresses inside a firewall) are unique, use a prefix of FD00::/8, and are not used for routing over the Internet

→ Link Local addresses are automatically configured by DART when using a Global Unicast or Unique local Unicast address

→ Link Local addresses only apply to the local Ethernet broadcast domain

→ Link-local Unicast addresses should never be used as a Source address if the Destination is a Global Unicast address

→ Link local interfaces are used for Neighbor Discovery, Router Discovery, and Duplicate Address Detection (DAD)

→ Link local addresses always begin with FE80::/10 or 64

### **MULTICAST:** Multicast---FF00::/8

→ Identifier for a set of different interfaces that are combined into a multicast group

→ Packets sent to multicast are delivered to all interfaces that are members of the address [FF02::2 all Routers on this link]

→ Multicast groups can be local, enterprise wide, or Internet

→ Many scopes possible

→Multicast addresses should not be used as a Source address

→Multicast addresses identified by first 8-bits, FF's, in network prefix portion of address, followed by 8-bits to define flags and scope

#### **4-bit Scope Field for Multicast Addresses:**

| Binary | Hex | Scope              |                  |
|--------|-----|--------------------|------------------|
| 0001   | 1   | Interface-local    | [FF01::1]        |
| 0010   | 2   | Link-local         | [FF02::1]        |
| 0100   | 4   | Admin-local        |                  |
| 0101   | 5   | Site-local         | [FF05::2, et al] |
| 1000   | 8   | Organization-local |                  |
| 1110   | E   | Global scope       |                  |

#### **ANYCAST:**

→Identifier for set of interfaces

→Packet is sent to only one interface in a group of different address nodes/interfaces

→Can be Global or Link-local scope addresses

→Address created from 64-bit network prefix and 57-bits of host ID portion, with 7-bits of host ID reserved for anycast ID

#### **IPv6 ADDRESS TYPES:**

**Link-local Unicast—FE80::/10** [1111111010 is the leftmost 10 bits]

**Link-local Multicast—FF02::/16**

**Note:** By definition, Link-local addresses are used only on local link and are not Routable

Local Unicast—Private internal networks—FD00::/8

Dual-stack Unicast--::FFFF:a.b.c.d [IPv4 over IPv6 for dual-stacked systems]

Site-local Unicast—FEC0::/10

**Multicast---FF00::/8** [All 1's for leftmost 8 bits, 4-bits for flags, 4-bits for scope]

FF02::2 →Link-local all-routers multicast [1<sup>st</sup> 8-bits indicate multicast, next 4-bits=flags, next 4-bits=Link-local]

**Global Unicast—2000::/3** (001) [2001::/16]

**Loopback---1/128** [all zeroes and a single bit on far right]

**Unspecified---1/128** [all zeroes]

#### **USING IPv6 ADDRESS WITH BROWSER:**

**Note:** Key is to enclose the IPv6 address within brackets [ ]

http://[2001:0db8:0:cd30:0200:abff:fe2e:2345]:8080/page.html

#### **MIGRATION/CO-EXISTENCE STRATEGIES FOR IPv6 & IPv4 NETWORKS:**

Creating Dual stacks on Hosts and Routers is one way to allow both IPv4 & IPv6 to co-exist (requires RIPng or OSPFv2/v3)

--Applications must support dual-stack

--IPv6 DNS Resolver call is getaddrinfo() vs. IPv4 gethostbyname()

Using Static or Automatic Tunneling methods is another way to bridge IPv6 networks that span IPv4

#### **MANUAL IPv6 TUNNELING (Point-to-Point):**

→point-to-point and unidirectional

→IPv6 packets are encapsulated and carried over the IPv4 network

→ideal for connecting two IPv6 networks together over the IPv4 internet

→Border Routers are dual-stacked

→Uses OSPFv3, RIPng (RIP for IPv6 distance vector routing protocol), or EIGRPv6 protocols for Routers

→Requires default IPv4 gateway on Hosts

#### **6to4 AUTOMATIC TUNNELING:**

→point-to-multipoint

→Automatic and dynamic tunneling of IPv6 packets over IPv4 networks, using 6to4 or ISATAP addresses

→most common technique is to use 6to4 tunnels [Net prefix 2002::/16]

→Hosts must be dual-stacked and have global IPv4 address to use 6to4 tunnel

→tunnels automatically created

#### **Example 6to4 Address:**

2002:AE13:4AOC:4:0:0:0:C7 (Dest. Router IPv4 address embedded: 174.19.74.12)

**Note:** Number of other tunneling protocols: GRE, ISATAP, IPv6IP, NAT, TEREDO, etc. NAT servers cannot handle IPv6.

#### **OSPFv3 (Open Shortest Path First) ROUTING PROTOCOL:**

→Updated version of OSPF that supports IPv6

→Protocol used in large or small networks based on link state algorithms, fast convergence, uses metrics to establish least-cost path

→Enabled on an interface-by-interface basis

#### **Enabling OSPFv3 on Router:**

```
# config t
```

```
# interface fastethernet 0/1
```

```
# ipv6 ospf 1 area 0
```

## **OTHER FACETS OF IPv6:**

### **DHCPv6 Servers:**

→Supplies all configuration parameters to Nodes except network prefixes for STATELESS configuration, requires Router [O flag set in Router Advertisement]

→Supplies all configuration parameters for STATEFUL, does not require use of a Router [M flag set in Router Advertisement]

### **DNS:**

IPv6 entries are stored as AAAA records using an IPv6 compliant BIND

Linux adds /etc/named.conf entry:

```
listen-on-v6 {any};
```

### **MOBILE IPV6:**

Supports an internet connection for mobile users when transiting from one IPv6 network to another

Router supports through use of MN entries (Mobile Node) onto foreign networks when traveling

Home Agent refers to mobile user when on the Home network

## **WINDOWS XP IPv6 NETWORK SHELL UTILITY:** netsh.exe

### **Configuring IPv6 on XP:**

```
C:\>netsh interface ipv6 install  
C:\>netsh interface ipv6 show interface  
C:\>netsh interface ipv6 show address  
C:\>ping <router_ipv4_addr>  
C:\>ping <router_ipv6_addr>  
C:\>ping ::1 (loopback) | ff02::1 (all hosts) | ff02::2 (all routers)  
C:\>nets hint ipv6 add dns "P-IV Laptop LAN #1" 2001:abcd:0:1001::6 [adding host IPv6 ID to DNS]
```

## **CISCO ROUTER COMMANDS IPv6:**

### **Enabling IPv6 on Router, then Interfaces:**

```
# ipv6 unicast-routing  
# interface fastethernet 0/0  
# ipv6 enable  
#ctrl z  
# show ipv6 interface [to verify—creates several multicast group addresses and a link-local address]  
# write memory [saves to startup config]
```

Configuring an IPv6 Network on the Router:

```
# config terminal  
# interface fastethernet 0/0  
# ipv6 address 2001:abcd:0:6::/64 eui-64  
# ctrl z  
# write mem  
# show ipv6 interface fastethernet 0/0
```

### **Configuring Manual Tunnel between IPv6 Sites:**

```
# config t  
# ipv6 unicast-routing  
# interface fastethernet 0/1  
# no shutdown  
# ip address 10.1.6.254 255.255.0.0  
# ipv6 address 2001:abcd:0:1001::6/64  
# ipv6 enable  
# ipv6 nd dad attempts 2  
# ctrl z  
# show run # write mem  
# config t  
# router rip [enabling RIP]  
# network 192.168.6.0 # network 10.0.0.0 # exit  
# interface fa0/0  
# ipv6 rip process1 enable [Enabling RIPng]  
# interface fa0/1  
# ipv6 rip process1 enable  
# ctrl z # write mem # show run  
# config t  
# interface tunnel 100  
# ipv6 address 2001:abcd:0:106::1 [Tunnel endpoint on one side] # ipv6 enable # ipv6 rip process1 enable  
# ipv6 address 2001:abcd:0:106::2 [Tunnel endpoint on other side] # ipv6 enable # ipv6 rip process1 enable
```

```
# tunnel source 10.1.6.254 # tunnel destination 10.2.16.254 [Configuring endpoints & associating with tunnel]
```

```
# tunnel mode ipv6ip # ctrl z #write mem # show run # show ip route # show interface tunnel100
```

## **CELERRA COMMANDS:**

```
# .server_config server_2 -v "ifconfig showall"
```

-----output abridged-----

Interfaces:(1)

loop6 on loop address= ::1/128 DNIF UP

```
mtu=32768, dmtu=32768, vlid=0, mac=0:0:0:0:0:0 dmac=0:0:0:0:0:0 if_index=4
```

## **CELERRA NAME SERVICES [HOSTS, NETGROUPS, NIS, DNS, WINS]:**

### **NIS SERVICE:**

**Intro:** NIS is usually used when using UNIX Security & User-Level Authentication [NT Domain Controllers are not used]. NIS is a network lookup service that provides for hostname-to-IP translation or IP-to-hostname, netgroup memberships, etc.

However, NIS is sometimes configured in “mixed” environments where NT Security is also used. In these cases, an NT client accesses the Celerra via Netbios name [as registered by WINS], which in turn interrogates the Domain Controller with the Security Access Token information of the User for validation and access rights. When using NIS, the DataMover points to the NIS service, which must have the NT User Names and Domain listed in its Map files—this information can be maintained on the NIS service by using NT Migrate [or manually by adding the proper entries whenever new NT Users are added to the Domain], or pushed out to the DataMover for better reliability and redundancy in case the NIS server becomes unavailable.

**Caution:** NIS+ is not supported by Celerra!

**Note:** When using following param entry, Usernames will be resolved to first matching name in NIS passwd file without regard to NT Domain extension—in otherwords, it would no longer be required to add the “username.domain” syntax.

```
/nas/server/slot_x/param
```

```
param cifs resolver=1
```

**NIS LIMITATION:** CIFS User Migration Tool, UNIX User Management Tool, & UNIX Users & Groups property pages might not be able to return Users or Groups with names longer than (17) characters

### **NIS SERVICE CONFIGURATION FILES: Network Information Service [NIS]**

--Functions as the Unix Nameserver & Security Server; files are located in **/var/yp** directory [ypserv ypbind]

--Master copy of **Hosts**, **Passwd**, **Group**, and **Netgroup** files

--Ypserv—Server daemon servicing client requests: client lookups; map maintenance; internal NIS calls

--Ypbind—Client daemon making requests and server uses to reply to clients.

**CELERRA NETGROUPS (NIS):** A netgroup file is a NIS database map file that is used for creating sets of Users and Hosts that help administratively define access to specific Hosts or NIS Domains. A netgroup file’s format consists of a set of triples or three ‘fields’ within ( ): setname (hostname, username, domainname). If an entry within the ( ) is left blank, that field becomes a wildcard. If an entry contains a dash (-), then the field has no value. One suggested method for building netgroup files is to use a netgroup to define “users” and a netgroup to define “hosts”:

#### **EXAMPLE 1:**

gate-users: ( ,isaac, ), ( ,moby, ), etc..... [Defines Users for Netgroup “gate-users”]

gate-hosts: (taco, ,), (bell, ,), etc..... [Defines Hosts for Netgroup “gate-hosts”]

#### **EXAMPLE 2:**

```
/etc/netgroup
```

root-users (,user1,), (,user2,), (,user3,)

trusted-machines (machine1,,), (machine2,,),

**Purpose:** When exporting NFS filesystems, customer may want to use NetGroup files to control access. Default Celerra behavior is to allow any unix client to mount a directory. You could restrict this behavior by using NetGroup files to restrict Users or Hosts.

Netgroups can be defined by local Datamover **/.etc/netgroup** files, or using NIS.

1. Defined in the **/.etc/netgroup** file on the DataMover

2. Or defined using NIS

**Note:** Recommended that 'Hosts' not be members of more than (1) netgroup!

--Netgroups only works when NIS is implemented, and combines password & host files to control User access via password file or Host access to NFS filesystems via host file.

--You can use Netgroups in the Passwd file by adding following at end of /etc/passwd: +@root-users

--Workaround for NIS+; Push the Netgroup Tables from NIS+ Server to NIS Host. Then use Cron job to send to Control Station, then put out onto the **/.etc** directory

--Netgroup line length limitation is 16,383, see AR37564 [5.1.20.2 & 5.2.6.0]—previous limitation was 1024 characters per line

## **NFS EXPORT RESOLUTION PRECEDENCE:**

- Subnet entries take more precedence than netgroup entries and overrule
- Host entries take precedence over Subnet entries

## **NETGROUP EXAMPLE FOR NFS:**

**\$server\_export server\_2 -Protocol nfs -ignore -option root=test\_test1 /fs12**

**Note:** If using DNS with FQDN, then netgroups must also use the FQDN

## **CONFIGURING THE NETGROUPS FILE:**

--Each line of Netgroup File consists of a Groupname followed by (3) fields that define:

**groupname (hostname, username, domainname)**

## **EXAMPLE:**

jupiter (, , galaxy) →Represents the group called “Jupiter” in the NIS domain called “Galaxy” that includes all Hosts and Users  
planets (-, ,galaxy) →Represents the group called “Planets” in NIS domain called “Galaxy’ that includes all Users but no Hosts

**Note:** Celerra does not support netgroup queries in NIS+ compatibility mode—must use local netgroup file on DM.

## **ORDER OF CELERRA HOST NAME RESOLUTION:**

→ Celerra searches its ./etc/hosts, then ./etc/netgroup files first, then NIS, then DNS until a match is obtained

**Note:** Configure local ./etc/hosts file with IP Addresses and Host names; Configure ./etc/netgroups file; Configure DM to run NIS & DNS Services

### **Example of Hostname resolution from ./etc/hosts file:**

**# .server\_config server\_2 -v "gethostbyname loner"**

1211898331: LIB: 7: Querying LOCAL host file by name.

1211898331: LIB: 6:

1211898331: LIB: 6: Name: loner

1211898331: LIB: 6:

1211898331: LIB: 6: Aliases:

1211898331: LIB: 6: No aliases

1211898331: LIB: 6: IPv4 Address(es):

1211898331: LIB: 6: 192.1.4.222

1211898331: LIB: 6:

1211898331: LIB: 6: IPv4/IPv6 Socket(s):

1211898331: LIB: 6: ::ffff:192.1.4.222 AF\_INET

**Note:** NAS 5.5 only returns a single IP address for a given Hostname, whereas NAS 5.6 correctly displays all known IP addresses for a given Host. See AR117706 for more details—NAS 5.5.33 and higher contains the fix.

## **CONFIGURING NIS CLIENT ON LINUX CONTROL STATION:**

Step 1. #linuxconf [provide NIS Domainname & server IP address]

Step 2. #/sbin/chkconfig ypbndt on [sets up ypbndt to start on next reboot of CS]

Step 3. #/sbin/chkconfig --list ypbndt [verifies your ypbndt setup]

Step 4. Reboot Control Station #reboot

## **CONFIGURING DATA MOVER AS NIS CLIENT:**

Step 1. Copy Control Station Hosts file to DataMover: \$server\_file server\_2 -p hosts hosts

Step 2. Configure DataMover to use NIS: **\$server\_nis server\_2 nsgprod 172.24.80.100,172.24.80.101** [NIS IP Addresses]

**Note:** Be aware that data mover uses names when mappings are made using NIS or local passwd/group files, not SIDs. Also know that all NIS entries must be in lowercase as Celerra forces all Windows names to lowercase before making NIS ‘yp match’ call.

## **EXAMPLE OF PROBLEM NAME:**

**\$server\_config server\_2 -v "yp match nas-nis 13434 group.bygid"**

1121950202: NETLIB: 4: Research=20Admins.umc-users:x:13434:

## **NIS SERVICE LIMITATIONS:**

Only a single NIS domain can be used, with up to 10 NIS Servers specified

## **WINDOWS SERVICES FOR UNIX 3.0:**

Be aware that NIS can be run as a service on Windows 2000/2003 systems—Celerra Servers can be NIS clients

## **NIS COMMANDS:**

**\$server\_nis server\_2 nsgprod 172.24.80.100,172.24.80.101** [IP Addresses of NIS servers]

```
$server_nis server_2 135.112.5.11,135.112.5.12 [Configures DataMover to use NIS]
$server_nis server_x      [Displays NIS configuration]
$server_nis server_x -d   [deletes NIS Services]
$server_nis server_x hostname | ip      [queries NIS for Hostname or IP Address]
$server_nis server_x -status           [Displays status of NIS Servers]
# ypwhich
# ypmatch reduced2@mouse.com.mouse passwd [Using Linux CS to conduct lookup of User in NIS]
reduced2@mouse.com.mouse::52007:110:S-1-5-15-42f831d9-66417ccd-28a68b82-4ca:/usr/S-1-5-15-42f831d9-66417ccd-28a68b82-4ca:/bin/sh
# ypmatch admins.mouse group          [Using Linux CS to conduct lookup of Group in NIS]
admins.mouse:*:2000:S-1-5-15-42f831d9-66417ccd-28a68b82-200:
# ypcat hosts
```

### **USING YP MATCH TO RESOLVE USERS, GROUPS, HOSTS BY DART:**

**Note:** You can use YP calls from data mover to query NIS status, to find users by name, groups by gid, and hosts by name

#### **Typical map names are:**

```
group.byname
group.bygid
passwd.byname
passwd.byuid
hosts.byname
hosts.byaddr
netgroup
netgroup.byhost
netgroup.byuser
```

```
# .server_config server_2 -v "yp status"
```

```
1103244764: NETLIB: 4: NIS default domain: atd.gmeds.com
1103244764: NETLIB: 4: NIS server 143.242.8.152
1103244764: NETLIB: 4: NIS server 143.242.8.153
```

```
# .server_config server_2 -v "yp match atd.gmeds.com tzn2y3 passwdbyname"
```

```
1103244796: NETLIB: 4: tzn2y3:RDxsnL1a4.ymk:35176:5000:Rusk, Clinton (NO ALIAS):/home/tzn2y3:/bin/csh
```

```
$ .server_config server_2 -v "yp match mynisdomain 5003 group.bygid"
```

```
1103173920: NETLIB: 4: managers::5003:
```

```
1103173920: ADMIN: 4: Command succeeded: yp match mynisdomain 5003 group.bygid
```

```
$ .server_config server_2 -v "yp match mynisdomain corp002 hostsbyname"
```

```
1103174936: NETLIB: 4: 10.16.10.13 CORP002 corp002
```

```
1103174936: ADMIN: 4: Command succeeded: yp match mynisdomain corp002 hostsbyname
```

```
# .server_config server_2 -v "yp match 171dc domain=20admins.eng groupbyname"
```

```
1129229947: NETLIB: 4: domain=20admins.eng: *:25001:
```

**Note:** Must use fully qualified groupname when resolving groups by name if CIFS resolver=0 is set, including the “=20” for a space between names

#### **DO NOT USE UPPERCASE CHARACTERS IN NIS MAPS:**

**Note:** Celerra forces all Windows names to lowercase before making NIS call, meaning that any uppercase names would fail because NIS is case sensitive and would not be able to provide a match. Also know when to use the cifs resolver param. For NIS entries that contain name.fqdn, you would keep default cifs resolver=0 param.

```
$ .server_config server_2 -v "yp match nas-nis 13434 group.bygid"
```

```
1121950202: NETLIB: 4: Research=20Admins.umc-users:x:13434:
```

```
$ .server_config server_2 -v "lsarpc user='research admins'"
```

```
1121950720: SMB: 5: Unix group 'research=20admins' unknown
```

```
1121950720: SMB: 5: dom='UMC-USERS' (2d8e3ce4)
```

```
1121950720: SMB: 5: ExtractSIDs:Usr='UMC-USERS\research admins' RID=347a U=4 D=0 UID=65534 T=3 (4)
```

```
1121950720: SMB: 5: MsError sendLookupNames=6 NTStatus=0
```

Finding User SID from Name=0

Interface 'elv0' Address=128.206.8.41

UserName0='UMC-USERS\research admins' (0) use=4 nameType=0

S-1-5-15-bfc2566-26bd6a90-49c7643a-347a research admins

UNIX ID=65534 Type=3

**Note:** Lsarp call does retrieve SID from DC but fails to map, so assigns the default ‘Nobody’ group gid or 65534

## **CONFIGURING/VERIFYING DNS/NIS ON DM:**

```
#ypwhich [NIS setup on CS] #more /etc/resolv.conf [Nameservers for client] #more /nas/server/slot_x/misc [DM DNS info]
#/usr/sbin/nslookup hostname {IP Address} #server_dns ALL #server_nis ALL #server_arp ALL -a
```

## **SETTING UP NIS ON 7.2 LINUX CONTROL STATION: NAS 4.1 and above**

Edit the following file to allow the Control Station to resolve NIS Hostnames:

1. **#vi /etc/nsswitch.conf**

hosts: files nisplus dns

2. **Change above line to read:**

hosts: files nis dns

## **SETTING UP LINUX AS NIS CLIENT WITH AUTOMOUNTER & PAM SECURITY:**

1. Define NIS Domain and Server in /etc/sysconfig/network file
2. Define NIS Domain and Server in /etc/yp.conf
3. Uncomment “nis or yp” line in /etc/nsswitch.conf
4. Run #authconfig to configure system as NIS client, specifying NIS Domain and NIS Server IP
5. #/sbin/service sshd restart [Restart Sshd so as to use NIS]
6. Configure AutoMounter so as to define NIS Users’ Home Directory on NIS Server:  
#vi /etc/auto.master → /home/guests /etc/auto.guests --timeout=60  
#vi /etc/auto.guests → \* -rw,soft,intr 192.168.1.20 :/home/guests/&  
#chkconfig autofs –runlevel 345 on #/sbin/service autofs restart

## **TROUBLESHOOTING NIS:**

### **Example of NIS Service on DataMover:**

\$ server\_nis server\_2

**server\_2 : yp domain=peoplesoft.com server=216.131.193.79**

### **Using YPWHICH Command to Resolve NIS Server Name:**

**\$ ypwhich 216.131.193.79**

it-sun66

### **Test for NIS Resolution of the DataMover and ‘Client’:**

1. Success would show test1 as resolving with the NIS server—[or use IP address]
2. Failure might indicate that the DataMover’s Unix name is not listed in the NIS Server’s /etc/hosts file

## **MORE NIS TROUBLESHOOTING:**

1. Ping by Hostname from DataMover
2. Gather output from NIS Server: #ypcat -k hosts #ypcat -k netgroup
3. If ‘automounter’ is used, can we see it?
4. Obtain Snoop traces between NFS Client and DataMover; between NIS Server and DataMover
5. For multiple NIS Servers, do map files update properly? /var/yp/domainname /etc/nsswitch.conf
6. \$more /nas/server/slot\_2/misc [Obtain NIS & DNS configuration on DataMover]
7. Check ARP tables

**Note:** NIS runs in Broadcast Mode by default—be careful when going across subnets!

## **USING NIS TO RESOLVE NT USERS FOR ACCESS TO CIFS SHARES ON DATAMOVER:**

**Note:** Common in environments that are predominantly Unix in nature, but require some CIFS User access

### **REQUIREMENTS:**

#### **1. NT DOMAINS REQUIRE GID MAPPING ENTRY IN GROUP FILE ON DM’s:**

\$ more group

**na01:\*:11992:** [NT Domain called “na01” with GID of 11992]

nj7460rd01:\*:11993:

domain=20admins.nj7460rd01:\*:2001:

domain=20users.nj7460RD01:\*:2002:

domain=20users.na01:\*:2003:

#### **2. NIS PASSWORD FILE ON NIS SERVER [or locally on Datamover]:**

kola:x:35012:4000:Ojebuoboh,Kolapo=kola:/home/kola:/bin/ksh

**kola.na01:x:35012:11992:Ojebuoboh,Kolapo=kola::**

**Note:** This is correct entry for NT User “kola.na01”

### **3. DM AS NIS CLIENT:**

**\$ server\_nis server\_4**

server\_4 : yp domain=wayne\_manor server=135.17.47.100

**\$ server\_nis server\_4 135.17.47.100**

server\_4 :

joker-cc = 135.17.47.100

### **4. ADD PARAM & REBOOT DM's:**

**param cifs useUnixGid=1**

**Note:** For Windows User access, will apply the GID from the local passwd file or NIS database as the default Windows group to be used, instead of the default Windows group, “Domain Users”

**YP Domain:** wayne\_manor 135.17.47.100 [Server name is ‘Joker’]

**NIS Password File Location:** /var/yp/etc/passwd

Kola:x:35012:4000:Objebuoboh:,Kolapo:/home/kola:/bin/ksh

Kola.na01:x:35012:11992:Objebuoboh,Kolapo=kola::

**Comment:** Second password entry for User “Kola” defines the NT Domain entry for the User whereas the first line defines Unix

**NIS Netgroup File:** /var/yp/netgroup [view netgroup: allft2k—itself contains a list of Users & machines]

## **ADDING USER TO LOCALGROUPS DB WITHOUT CORRESPONDING NIS PASSWORD ENTRY:**

**“A new member could not be added to a local group because the member has the wrong account type”**

**Note:** This would be equivalent to not having a password UID/GID match to the User Name

## **TROUBLESHOOTING CELERRA NAME SERVICES:**

--#more /nas/server/slot\_x/misc --Contains DNS/NIS/NTP info for DM's

--#more /etc/resolv.conf --Shows the DNS server info if DNS is configured on the CS.

--Use “traceroute” tool & “ping” to examine routes and responses

--Use “nslookup” on Control Station to verify DNS resolution

--Use “dig” to query nameservers

--Use “host” command to resolve hostnames or IP Addresses in DNS

--#dnsdomainname --shows DNS Domain

--#ypdomainname --shows NIS Domain

--#ypwhich --shows NIS Server

--#ypcat --shows NIS database

## **TROUBLESHOOTING DNS CONFIGURATION FILES:**

/etc/resolv.conf [File used to setup DNS client on Control Station]

/etc/host.conf [File used to indicate the order of resolution for DNS, NIS, Hosts file, etc]

/etc/nsswitch.conf [hosts: files nisplus dns →Order of resolution is Hosts File, NIS, then DNS]

/usr/bin/host 172.24.80.10 [Command to see if DNS will resolve Host by IP or machinename, etc.]

## **WINDOWS 2000 DNS SERVICES:**

--Integrated with Active Directory

--Uses Secure Dynamic updates

--Aging & scavenging feature

--Unicode character support for UTF-8

--Commandline tool: dnscmd.exe

## **DNS UPDATES WITH CIFS:**

→Prior to NAS 5.6 code, DNS updates were performed via the default interface

→With NAS 5.6, DNS updates are performed on the interface configured for the CIFS server

**param dns.bindUpdInterface**

**Note:** The above param added with NAS 5.6.46 to support NAS 5.5 DNS bind method (AR137299). Problem happens when a CIFS server is on a different physical network, and cannot reach the DNS server with the CIFS interface because the route to the DNS is defined on a different interface via the default route. New param is set to 0 by default and will now allow Dynamic DNS updates to occur via the default route interface on the Data Mover, as opposed to the CIFS comppname interface, which may be on a different network.

## **CONFIGURING DNS CLIENT ON DM:**

**#server\_dns server\_2 -p tcp {udp} domain\_name 192.1.2.7** [IP Address DNS server]

## **TEXTBOOK WAY TO UPDATE DNS CLIENT CONFIGURATION ON DM:**

- 1. \$ server\_dns server\_2 -o stop**
- 2. \$ server\_dns server\_2 -o flush**
- 3. \$ server\_dns server\_2 -p tcp win2kemc.com 10.62.25.115,10.64.25.114**
- 4. \$ server\_dns server\_2 -o start**

**Note:** The abbreviated method would be just to issue Step 3 and then verify that DNS is still updating o.k. DNS starts with UDP protocol by default, use “-p tcp” to define TCP protocol.

**Server dns Command:** Display & setup DNS for client DataMover \$server\_dns server\_4 -p tcp comics 193.1.21.200

-p tcpludp {default} start DNS client using tcp or udp protocol Domain\_name -d delete client from DNS

**Verify DNS Settings:** \$more /etc/resolv.conf \$more /nas/server/slot\_4/misc \$/usr/sbin/nslookup

## **Configuring a Sun Machine as a DNS Client:**

Step 1 #cd /etc and vi nsswitch.conf [add the following to the Hosts Line: “Hosts: files [NOTFOUND=continue] dns”]

Step 2 #vi resolv.conf file and add following: “domain <dns domain name>” “nameserver <IP of nameserver>”

Step 3 #ps -ef |grep inetd [record PID number and #kill -HUP 407]

**Comment on NIS and DNS:** Up to (10) NIS servers can be defined [1 Primary, 9 Slaves] for a single NIS domain, and (3) DNS servers per Domain [1 Primary and 2 Secondaries]—no limit to DNS Domains that can be defined.

**NIS+ NOT supported for Celerra!!** Local Netgroup file is supported by NIS.

## **MICROSOFT WINS:** Windows Internet Name Service

**Purpose:** Distributed database for registering and querying netbios computer names-to-IP addresses

## **CONFIGURING WINS CLIENT ON DM:**

**# server\_cifs server\_2 -a wins=10.127.15.252**

Netbios name is limited to 15 characters.

## **Default Celerra Name Resolution Process:**

1.) Local Hosts File      2.) NIS server, if used [DataMover client]    3.) DNS Server, if used [DataMover client]

## **NETBIOS NAME RESOLUTION:**

Uses UNC paths or Browsing

Resolves via: Cache/WINS/Broadcast/LMHosts/Hosts/DNS

## **BROWSE ISSUES ON MS NETWORKS:**

--Netmask incorrect

--Broadcast address incorrect

--Ensure that NT Server is setup as Browse Master: HKLM>CurrentControlSet>Services>Browser>Parameters>MaintainServerList

## **NTRESKIT TOOLS:**

--browstat.exe [CLI program]

--browmon.exe [GUI Program]

## **NBTSTAT UTILITY:**

**Purpose:** Useful for determining NetBIOS Name Table Information on both Local & Remote Hosts

nbtstat -n [Local Host NetBIOS registrations]

nbtstat -a [Remote Host NetBIOS Name Table by Hostname]

nbtstat -A [Remote Host NetBIOS Name Table by IP]

nbtstat -r [Name statistics resolved by Broadcast & WINS locally]

nbtstat -c [Netbios cache on local or remote system]

nbtstat -S [Sessions by IP]

nbtstat -s [Sessions by Name]

nbtstat -RR [Sends name release to WINS and Refreshes]

nbtstat -R [Purges & reloads remote cache table]

**EXAMPLE:** c:>nbtstat -a corpusdc1 [-A 128.221.13.132]

NetBIOS Remote Machine Name Table

| Name         | Type        | Status     |
|--------------|-------------|------------|
| CORPUSDC1    | <00> UNIQUE | Registered |
| CORP         | <00> GROUP  | Registered |
| CORP         | <1C> GROUP  | Registered |
| CORPUSDC1    | <20> UNIQUE | Registered |
| CORP         | <1E> GROUP  | Registered |
| CORPUSDC1    | <03> UNIQUE | Registered |
| CORP         | <1D> UNIQUE | Registered |
| .._MSBROWSE_ | <01> GROUP  | Registered |
| CORP         | <1B> UNIQUE | Registered |

MAC Address = 00-B0-D0-B0-93-5E

**WINS-Enabled Client LookUps:** Once IP address obtained, client communicates directly with ‘other’ computer

- 1.) Client checks cache
- 2.) Client directs query to WINS server (h-node mode, if configured with WINS IP address)
- 3.) Client b-node Broadcasts query on local subnet—if target on same subnet, will return its IP address
- 4.) Local LMHosts file checked

**Note:** If “Enable DNS for Windows Resolution” is configured on client, additional steps are used to resolve:

- 5.) Hosts file checked
- 6.) DNS Server queried

**Non-WINS Client LookUps:**

- 1.) b-node broadcast name query on local subnet—if target on same subnet, will return its IP address
- 2.) Local LMHosts file checked

**Host Name Resolution:** DNS & NIS

Used for Internet names; Unix Hosts, TCP protocols such as Ping, Telnet, FTP

**Note:** Ping & FTP use Hosts file by default

**USING NSLOOKUP TO TROUBLESHOOT DNS ISSUES:**

**\$ /usr/sbin/nslookup**

**Note:** NSLOOKUP uses UDP queries by default, but will use TCP if necessary

**START OF AUTHORITY:**

>set type=soa [Start of Authority]  
>t2dom2.local [Returns information about the dns domain]

**NAME SERVERS:**

**>set type=ns** [querying for Name Servers]  
>t2dom2.local [Name Servers for specific domain]  
>server win2dom2.t2dom2.local [Changing DNS lookup to a specific Name Server]  
>ls -t t2dom2.local [lists out subdomains in a domain]

**QUERYING A RECORDS:**

**>set type=a**

>cpc1223 [queries for “A” Address Record for Host name-to-ip address resolution]  
\$/usr/sbin/nslookup cssc1223 [Returns Address Record for Host “cssc1223”]

**QUERYING PTR RECORDS:**

**>set type=ptr**

>192.10.3.2 [Queries for “PTR” Pointer Record for Host ip-to-name resolution]  
\$/usr/sbin/nslookup 192.10.3.2 [Returns address record for cssc1223.us.dg.com]

**Note:** Querying by IP automatically checks Host Name against PTR Record

**QUERYING REMOTE HOST NAME:**

\$/usr/sbin/nslookup [ftp.emc.com](http://ftp.emc.com) [If resolution returned, means that your DNS server is querying DNS Root servers]

**Comment:** Together, querying for A, PTR, and Remote Host names will validate whether local DNS System is functioning

**Aliases:** Additional names beyond the A Record name [CNAME or canonical name]

**Example:** Multi-homed systems might want to have (1) Alias name to point to all addresses

**(3) Default Zones for Every DNS Server that are SOA:**

0.in-addr.arpa 127.in-addr.arpa 255.in-addr.arpa

**CELLERRA & SYMMETRIX: Bin Files & Volume Management**

**Symmetrix IntegratedCachedDiskArray [ICDA]:** ICDA is the “Technical” name for the Symmetrix

Data Volumes for Storage [SA Ports]; BCV's for EMC Timefinder/FS;

**Max. 240 Target & LUN Addresses used on a Symm:** All vol.'s addressed by SCSI Target ID [0 – F except 7] and LUN [0 – 7]

Maximum Number of Logical Volumes seen by a single SA or FA: 554

Largest Single MetaVolume Size: 3.825TB

**Note:** SCSI ID 7 is reserved for the DataMover controller

**Gatekeeper Devices:** Normally, one configured per Symm-Celerra, as an AS400 Gatekeeper device, located on the last logical volume, 7 cylinders in size [3.5MB] and addressed as FF. The Celerra Gatekeeper interfaces with Symm API calls using SCSI commands for applications such as Timefinder, SRDF, or Celerra Monitor. Without a GateKeeper device, NAS will not install on Celerra! Additionally, a 2<sup>nd</sup> GateKeeper would need to be configured for the “Volume Logix” [VCM] application [16 cyl.]

**Max. Number SA/FA Ports for Celerra:** SA's=32 SCSI ports      FA's=14 ports

**Note:** If using multiple Symmetrixes with Celerra [Outside of an SRDF configuration], then only (1) of the Symmetrixes can have the (6) Celerra Hyper Volumes configured. I.E, only one Symm can host a Celerra's Boot Devices. First two Symmetrix Volumes, d1, d2, are reserved for the Symm O/S and Logs

## **CYLINDER SIZES FOR CELERRA CONTROL LUNS:**

### **PRE-DMX3:**

LUN 0 & 1 = 24788 cylinders for 11GB LUNs

LUNs 2-4 = 4432

Gatekeeper = 6, set to AS400

### **DMX3/DMX4 (Pre 5.6):**

LUN 0 & 1 = 12394

LUNs 2-4 = 2216

Gatekeeper = 2

### **POST 5.6:**

LUNs 0 & 1 = 12394

LUNs 2-4 = 2216

LUN 5 varies: 2GB=2216; 32GB=34956; 64GB=2@34956

## **SYMMETRIX STORAGE SYSTEMS:**

### **DMX SERIES—Symmetrix Direct Matrix**

**DMX800, DMX1000, DMX2000 (aka Symm 6), replaces Symm 8000 series**

→Enginuity microcode

→RAID 0, 1, or 10 software RAID

→Direct Matrix architecture, using up to 128 direct dedicated data paths

→2Gb/s FC backend speeds

### **DMX800:** Jan 2003

→8-120 2Gb/s drives, 4-64GB memory, raw capacity of 58 TB, high-end entry point (73, 146, 300GB drives)

→Rack-mountable version of Symm 6.0

### **DMX1000:**

→DMX1000 is a single-bay system from 64-144 drives, raw capacity of 59TB & 128GB memory, 16-48 Host FC ports or ESCON, and up to 24 FICON, GbE, or iSCSI Host connections

→Up to 12 director/cache slots

→Symm 6 technology

### **DMX2000:**

→Dual bay version of the DMX1000, 24 director/cache slots

### **DMX3000:** Sep 2003

→Enginuity v5670, 3-bay system, up to 576 drives total, 84TB

→Really an extension of the DMX2000 platform by adding a bay

### **DMX-2 SERIES:** Feb 2004, both DMX-1 & DMX-2 run Enginuity 5671

→Extends capabilities of DMX platform for iSCSI, faster processors, etc. for DMX1000/2000/3000 series, Enginuity 5670

### **DMX-3 SERIES:** DMX3500/4500 GA 31 Aug 2005 Enginuity 5771

→24-card slot for channel, disk directors, cache & XCM boards

→Enginuity 5772 released March 2007 for enhanced security, cache partitioning, workload controls, 4Gb/s FC connectivity, RAID 6 support, and up to 128 RDF groups per array

→OpenHosts supported via iSCSI & FC, with mainframe hosts supported via FICON (GbE & ESCON later release)

→Front-end directors are not compatible with DMX-2 Series

→Each Bay will support 240 drives, with a max system of 960 drives

→Max total of 2400 disk drives, 12 Channel directors, 8 disk directors, FC, iSCSI, GbE, FICON, ESCON host connectivity, 256GB memory

→Cache bandwidth 32GB/sec and new 1.3GHz Power PC processors for Disk & Channel directors

→128 hosts per Fibre port; 256 hosts per iSCSI port

→Jan 2006, support for 1024TB, (1) Petabyte, up to 2400 drives, LC-FC (low cost) drives; Native GigE support

**DMX-3 950:** Oct 2006, extension of the DMX-3 series, Enginuity 5771

→Runs Enginuity 5771 code, replacement for DMX800 for OpenSystems hosts

→Latest high-end Symmetrix platform with central bay of 120 disks, with one additional bay of 240 disks, up to 1PB data

→Up to 64 FC hosts or 48 FICON/GbE or iSCSI Host connections

→64GB memory, 32-360 disks total

→Bandwidth twice that of DMX-2 Series, at 128GB/s

→12 Channel directors & 8 Disk directors

→73GB or 146GB 15k RPM FC drives, 146GB or 300GB 10k RPM FC drives, 500GB 7200RPM FC drives

**DMX-4 1500-4500:** Highend Symmetrix family July 2007 GA, Enginuity 5772

→System bay & up to (8) storage bays, 2-8 DAs, up to 12 Channel Directors [total 16 combined], 2, 4, 6, or 8 global memory directors

→Supports from 96 to 1920 disk drives with total capacity of nearly 1 PB

→(4) dual 1.3GHz PowerPC processors per director

→Enginuity 5772 software, which provides most of the performance gains

→4Gbps end-to-end architecture

→Supports data protection methods:

RAID 1, 10, 5, 6, SRDF, Dynamic & Permanent sparing

→Supports disk drives:

73, 146, 300, & 500GB FC drives; 750GB SATA (Serial Advanced Technology Attachment) II drives

**Logical Volume Structure:**

Each logical volume has n cylinders

Each cylinder consists of (15) tracks/heads

Each track contains 128 blocks of 512bytes each

To calculate volume size: n \* 15 \* 128 \* 512

**DMX-4 950:** August 2007

→Touts itself as first high-end array to support 4Gbps fibre speed, 750GB capacity SATA II disk drives

**Note:** Serial Advanced Technology Attachment Drives (SATA II) have a number of improvements, the biggest being a doubling in speed from 1.5 to 3Gbps

→Total of 180TB capacity, 6-slot, supporting 2-4 DA directors, with up to 360 drives when daisy-chained to additional Storage Bays

**SYMMETRIX TIGON (V-MAX):**

GA April 2009, new virtual matrix storage platform using Enginuity 5874, 96-2400 drives up to 2PB, system can have up to 1TB memory, supports 128 FC ports, 64 GbE ports, or 64 iSCSI connections. Drive options include SATA, FC, and EFD.

Celerra tolerance for V-MAX with 5.6.45. With Solutions Enabler 7.0 and VMax, device masking is done using ACLX, for Auto-Provisioning volumes from storage pools, Enginuity 5874 and later, and using SMC or ECC. You would need to set the ACLX flag on front-end Director ports to support auto provisioning of Symmetrix devices. You do not need to manually address the ALCX db volume with the OE address.

**TOTAL NUMBER SUPPORTED DEVICES PER FA/SA PROCESSOR PORT:**

904 for SymmWin 5266/5267

2048 for SymmWin 5566/5567/5568/5669/5670

4096 for SymmWin 5771

**OLDER SYMM MODELS**

**Symmetrix Product Family Nomenclature:**

3xxx = Open Systems Hosts      3300/5300 = Half Bay

5xxx = Mainframe Hosts      3400/5400 = Full Bay

8xxx = Symm 5      3700/5700 = 3-Bay

Symm 4

Symm 4.8

Symm 5

Symm 8230 & 8000 Series (Uses Enginuity O/S and Mosaic hardware architecture)

**CELERRA FA PORT SHARING:**

FA Port Sharing is supported by Celerra with Symm 8000 & DMX series—see E-Lab Navigator and EMC Support Matrix

ARB director flag still required for all Celerra FA ports to prevent host reset being propagated back to DM & for FC connectivity

**Different Types of Symmetrix “Contention”:**

1. Disk Adapter Contention
2. SCSI Contention
3. Physical Disk Contention
4. RAID-S Contention

**QUERYING FOR SYMM SERIAL #, MICROCODE LEVEL, SYSTEM & DISK INFORMATION:**

**\$/nas/bin/nas\_symm -l** [Outputs ID and NAME of Symmetrix]

```
id          acl      name
000283600226 0      000283600226
```

**\$/nas/bin/nas\_symm -i -all**

**nas\_symm -i Output:**

**Note:** Use of nas\_symm -i across Multiple Symmetrixes is a good way to test whether the Gatekeeper devices are properly setup. That is, you will not receive output for a Symm if the Gatekeeper is not in place or addressed correctly to a DataMover

**\$/nas/bin/nas\_symm -i 000283600226** [Outputs specifics on Symm Hardware, Software, Celerra Configuration, etc.]

```
id          = 000283600226
name        = 000283600226
ident       = Symm48
model       = 3630
microcode_version = 5266
microcode_version_num = 1492AA01
microcode_date   = 07242000
microcode_patch_level = 22
microcode_patch_date = 07242000
symmetrix_pwron_time = 980174655 == Mon Jan 22 14:44:15 GMT 2001
db_sync_time    = 985272913 == Thu Mar 22 14:55:13 GMT 2001
db_sync_bcv_time = 985272913 == Thu Mar 22 14:55:13 GMT 2001
db_sync_rdf_time = 985272913 == Thu Mar 22 14:55:13 GMT 2001
last_ipi_time   = 985208404 == Wed Mar 21 21:00:04 GMT 2001
last_fast_ipl_time = 985208404 == Wed Mar 21 21:00:04 GMT 2001
API_version     = V4.1-131-4.0
cache_size      = 2048
cache_slot_count = 53417
max_wr_pend_slots = 42775
max_da_wr_pend_slots = 21387
max_dev_wr_pend_slots = 3612
permacache_slot_count = 7042
num_disks       = 32
num_symdevs     = 144
num_pdevs       = 2
sddf_configuration = Enabled
config_checksum  = 0x0884da
num_powerpath_devs = 0
config_crc       = 0x06fbec98c
Physical Devices
/nas/dev/c0t15i15s2
/nas/dev/c0t15i15s3
```

**Director Table**

| type | num | slot | ident  | stat | scsi  | vols | ports | p0_stat | p1_stat | p2_stat | p3_stat |
|------|-----|------|--------|------|-------|------|-------|---------|---------|---------|---------|
| DA   | 1   | 1    | DA-1A  | On   | Wide  | 61   | 2     | On      | Off     | NA      | NA      |
| DA   | 2   | 2    | DA-2A  | On   | Wide  | 60   | 2     | On      | Off     | NA      | NA      |
| R2   | 3   | 3    | RA-3A  | On   | NA    | 0    | 1     | On      | NA      | NA      | NA      |
| SA   | 14  | 14   | SA-14A | On   | Ultra | 0    | 2     | On      | On      | NA      | NA      |
| FA   | 15  | 15   | FA-15A | On   | NA    | 0    | 1     | On      | NA      | NA      | NA      |
| SA   | 16  | 16   | SA-16A | On   | NA    | 0    | 2     | On      | On      | NA      | NA      |
| DA   | 17  | 1    | DA-1B  | On   | Wide  | 60   | 2     | On      | Off     | NA      | NA      |
| DA   | 18  | 2    | DA-2B  | On   | Wide  | 61   | 2     | On      | Off     | NA      | NA      |
| R2   | 19  | 3    | RA-3B  | On   | NA    | 0    | 1     | On      | NA      | NA      | NA      |
| SA   | 30  | 14   | SA-14B | On   | Ultra | 0    | 2     | On      | On      | NA      | NA      |
| FA   | 31  | 15   | FA-15B | On   | NA    | 0    | 1     | On      | NA      | NA      | NA      |
| SA   | 32  | 16   | SA-16B | On   | NA    | 0    | 2     | On      | On      | NA      | NA      |

**HOW TO DETERMINE SYMAPI VERSION ON CELERRA:**

**#/nas/symcli/bin/symcli**

Symmetrix Command Line Interface (SYMCLI) Version V4.3.1.0

**QUERYING DETAILS OF SYMMETRIX DEVICE:**

# /nas/symcli/bin/symdev show 00AD

Symmetrix ID: 000184503008

Device Physical Name : Not Visible

Device Symmetrix Name : 00AD

Device Serial ID : N/A

Symmetrix ID : 000184503008

Attached Snap TGT Device : N/A

Vendor ID : EMC

Product ID : SYMMETRIX

Product Revision : 5567

Device Emulation Type : CELERRA\_FBA

Device Defined Label Type: N/A

Device Defined Label : N/A

Device Sub System Id : 0x0140

Device Block Size : 512

Device Capacity

{

Cylinders : 18590

Tracks : 278850

512-byte Blocks : 17846400

MegaBytes : 8714

KiloBytes : 8923200

Device Configuration : BCV

Device is WORM Enabled : No

Device is WORM Protected : No

Dynamic Spare Invoked : No

Dynamic RDF Capability : None

Device Service State : Failed

Device Status : Not Ready (NR)

Device SA Status : Ready (RW)

Front Director Paths (2):

{

-----  
POWERPATH DIRECTOR PORT

| PdevName    | Type | Num | Type | Num | Sts | VBUS | TID | LUN |
|-------------|------|-----|------|-----|-----|------|-----|-----|
| Not Visible | N/A  | 03A | FA   | 0   | RW  | 000  | 00  | 06F |
| Not Visible | N/A  | 14B | FA   | 0   | RW  | 000  | 00  | 06F |

}

Mirror Set Type : [Data,N/A,N/A,N/A]

Mirror Set DA Status : [NR,N/A,N/A,N/A]

Mirror Set Inv. Tracks : [0,0,0,0]

Back End Disk Director Information

{

Hyper Type : Data

Hyper Status : Not Ready (NR)

Disk [Director, Interface, TID] : [16A, C, 1]

Disk Director Volume Number : 16

Hyper Number : 8

Disk Capacity : 70007m

}

BCV Pair Information

{

Standard (STD) Device Symmetrix Name : 00F6

Standard (STD) Device Serial ID : Not Visible

Standard (STD) Device Group Name : Not/Grouped

BCV Device Symmetrix Name : 00AD

BCV Device Serial ID : Not Visible

BCV Device Associated Group Name : Not/Associated

BCV Device Status : Not Ready (NR BCV)

State of Pair ( STD <=> BCV ) : Split

Time of Last BCV Action : Thu Nov 21 23:02:58 2002

State of BCV Mirrors : N/A

BCV State Flags : (inc) (CantRevSpl)

Percent Split : 100%

Number of Inv. Tracks for STD Device : 0

Number of Inv. Tracks for BCV Device : 1625

**DETAILS OF SYMMETRIX DEVICE GROUPS:**

#/nas/symcli/bin/symdg list

DEVICE GROUPS

| Name | Type    | Valid | Symmetrix ID | Devices | GK's | BCV's |
|------|---------|-------|--------------|---------|------|-------|
| 1REG | REGULAR | Yes   | 000184503008 | 143     | 0    | 110   |

**#/nas/symcli/bin/symdg show 1REG**

Group Name: 1REG

|                                                |   |                          |
|------------------------------------------------|---|--------------------------|
| Group Type                                     | : | REGULAR                  |
| Valid                                          | : | Yes                      |
| Symmetrix ID                                   | : | 000184503008             |
| Group Creation Time                            | : | Tue Dec 31 17:41:35 2002 |
| Vendor ID                                      | : | Celerra CS               |
| Application ID                                 | : |                          |
| Number of STD Devices in Group                 | : | 143                      |
| Number of Associated GK's                      | : | 0                        |
| Number of Locally-associated BCV's             | : | 110                      |
| Number of Remotely-associated BCV's (STD RDF): | : | 0                        |
| Number of Remotely-associated BCV's (BCV RDF): | : | 0                        |
| Standard (STD) Devices (143):                  | : | {                        |

---

| LdevName | PdevName | Sym  | Cap                       |
|----------|----------|------|---------------------------|
|          |          | Dev  | Att. Sts (MB)             |
| d1       | N/A      | 0000 | RW 4153                   |
| d2       | N/A      | 0001 | RW 4153                   |
| d145     | N/A      | 004F | RW 8714 [output abridged] |

**QUERYING FOR GATEKEEPER DEVICES:****#/nas/symcli/bin/symgate list**

Symmetrix ID: 000184503008

| Device Name | Directors | Device |
|-------------|-----------|--------|
|             |           |        |

---

| Physical             | Sym               | SA :P DA :IT Config | Cap        | Attribute | Sts (MB)               |
|----------------------|-------------------|---------------------|------------|-----------|------------------------|
| /nas/dev/c112t0l15s5 | 0007 03A:0 15B:D2 | 2-Way Mir           | N/Asstd GK | RW        | 3 [Celerra GateKeeper] |

**UNDEFINING & REDEFINING A GATEKEEPER DEVICE:****\$ /nas/symcli/bin/symgate undefine pd /nas/dev/c1t15l15s5****\$ /nas/symcli/bin/symgate define pd /nas/dev/c1t15l15s8****QUERYING FOR SYMMETRIX DISK TYPE:****#/nas/symcli/bin/syndisk list**

Symmetrix ID : 000184503008

| Disk Selected | : 56         |     |     |        |      |        |       |      |
|---------------|--------------|-----|-----|--------|------|--------|-------|------|
|               | Capacity(MB) |     |     |        |      |        |       |      |
| Ident         | Symb         | Int | TID | Vendor | Type | Hypers | Total | Free |

---

|       |     |   |   |         |         |   |       |     |
|-------|-----|---|---|---------|---------|---|-------|-----|
| DA-1A | 01A | C | 0 | SEAGATE | CHET_73 | 9 | 70007 | 287 |
| DA-1A | 01A | C | 1 | SEAGATE | CHET_73 | 8 | 70007 | 294 |

**LISTING OF EVENTS ON SYMMETRIX:****#/nas/symcli/bin/symevent list****MISCELLANEOUS SYMMETRIX QUERIES:****\$/nas/sbin/hostinq -h [help menu for switches] -et [for emulation and type]****#/nas/sbin/hostinq -celerra | -sid | -showvol | -clarion | -btl | -ckd****\$/nas/bin/nas\_symm -l [List of Symmetrixes]****\$/nas/bin/nas\_symm -i "serial number"****QUERYING SYMMETRIX DATABASE FOR CELERRA:****APPENDING PATH TO ENVIRONMENT VARIABLES FOR SYMCLI COMMANDS:****# echo PATH=\$PATH:/nas/symcli/bin****PATH=/usr/local/bin:/bin:/usr/bin:/usr/X11R6/bin:/home/nasadmin/bin:/nas/bin:/nas/symcli/bin****# export PATH=\$PATH:/nas/symcli/bin****SYMCLI COMMANDS:**

**#/nas/symcli/bin/symgate -sid 000185700809 list** [GateKeepers]

**#/nas/symcli/bin/symcfg verify -sid 09**

Wed Sep 7 15:55:33 CDT 2005

The Symmetrix configuration and the database file are NOT in sync

**#/nas/symcli/devs** [all devices]

**#/nas/symcli/bin/symdev list -v**

**#/nas/symcli/bin/symdev list** [Lists Symm devices, Mirrors, BCV's, Sizes]

**#/nas/symcli/bin/symdev show 00AD** [Very detailed status of Symmetrix Device]

**#/nas/symcli/bin/symdev -sid 0107 show** 030A [Use to interrogate specific symms for devices, especially over SRDF links]

**#/nas/symcli/bin/symcli**

**#/nas/symcli/bin/symcli -env** [Lists out Symcli Variables]

**#/nas/symcli/bin/symcli -def** [Lists out defined Symcli variables]

**#/nas/symcli/bin/symcfg discover** [Creates the Symapi Database]

**#/nas/symcli/bin/symcfg -v list** [Displays Symmetrix configuration info]

**#/nas/symcli/bin/symcfg -db** [Symapi version]

**#/nas/symcli/bin/symcfg list -services** [Displays network services available]

**#/nas/symcli/bin/symcfg list -lock** [Displays locks held on Symmetrix]

**#/nas/symcli/bin/symcfg list -meta**

**#/nas/symcli/bin/symcfg -lockn 15 list** [Displays info on a lock found with list -lock command]

**#/nas/symcli/bin/symcfg release** [Basic command to release symm locks]

**#/nas/symcli/bin/symdev -lock list** [See what symm device has lock]

**#/nas/symcli/bin/symdg -lock show devicegroupname** [Locks for SRDF devices]

**#/nas/symcli/bin/symdg -lock 9 release devicegroupname -force**

**#/nas/symcli/bin/symdg release devicegroupname**

**#/nas/symcli/bin/syminq** [SCSI inquiry to all devices]

**#/nas/symcli/bin/symdg list** [List of Device Groups--would show GateKeepers and BCV's]

**#/nas/symcli/bin/symdev list** [Displays list of Symmetrix devices]

**#/nas/symcli/bin/symbev list** [List of BCV devices]

**#/nas/symcli/bin/sympd list** [List of host physical devices]

**#/nas/symcli/bin/symhostfs list** [File System information]

**#/nas/symcli/bin/symrdf -g <group name> query** [State of RDF mirroring]

**#/nas/symcli/bin/symrdf -g <group name> -i <nn> -c <nn> verify** [Verifies synchronization state of RDF]

**#/nas/symcli/bin/symcfg -sid 000185700870 -FA 04B -v list** [or -FA ALL -v list]

**Note:** Provides Director Status, WWN, & Flags set on each Director

**#/nas/symcli/bin/symstat -i 5 -c 9996 -type PORT -dir ALL**

**Note:** Use command to obtain I/O stats on SCSI or Fibre Directors. I/O/sec of 3000 is extremely busy.

## SETTING UP RDF LINKS:

**#/nas/sbin/nas\_rdf -init** [Run on R1 side, then R2 side]

**\$nas\_server -a -i** [verify datamover relationships]

## SYMMETRIX APPLICATIONS:

**#/nas/symcli/bin/symcfg list -applications**

Wed Sep 7 16:20:26 CDT 2005

Symmetrix ID : 000187721009

| Host | Application | Node Name   | IP Address    | ID                   | Vendor ID | Version             | Attr   |
|------|-------------|-------------|---------------|----------------------|-----------|---------------------|--------|
|      |             | HK187721009 | 192.168.137.1 | EMC STP<br>OPTIMIZER | EMC Corp  | 0.0.0.0<br>4.1.70.1 | -<br>- |
| cs0  |             |             | 10.224.64.61  | SYMCLI               | EMC Corp  | 5.4.0.5<br>5.1.1.1  | -<br>- |
|      | Celerra CS  |             |               | EMC Corp             |           |                     |        |

```
cs1      10.224.64.62  Celerra CS    EMC Corp     5.1.1.1  -
dal01man0001 10.224.64.4  SYMCLI      EMC Corp     5.5.0.0  -
          ECC          EMC Corp     5.2.0.1  -
```

## **VIEWING HOSTS CONNECTED TO SYMMETRIX:**

**#/nas/symcli/bin/symcfg list -connections**

Wed Sep 7 16:13:06 CDT 2005

Symmetrix ID : 000187721009

| Symmetrix | Host |
|-----------|------|
|-----------|------|

| Director | Port | Node Name | IP Address | HW Type | OS Name | OS Revision |
|----------|------|-----------|------------|---------|---------|-------------|
|----------|------|-----------|------------|---------|---------|-------------|

|       |     |              |               |       |          |            |
|-------|-----|--------------|---------------|-------|----------|------------|
| DF-1A | N/A | HK187721009  | 192.168.137.1 | INTEL | WinNT-SP | 5.0.2195   |
| FA-9A | 0   | dal01man0001 | 10.224.64.4   | sun4u | SunOS    | 5.8        |
| FA-7C | 0   | cs0          | 10.224.64.61  | i686  | CelerraL | 2.4.9-34.5 |
|       | 0   | cs1          | 10.224.64.62  | i686  | CelerraL | 2.4.9-34.5 |

**#/nas/symcli/bin/symcfg list -connections -sorthost**

Wed Sep 7 16:18:27 CDT 2005

| Host | Symmetrix |
|------|-----------|
|------|-----------|

| Node Name | IP Address | OS Name | OS Revision | ID | Director | Port |
|-----------|------------|---------|-------------|----|----------|------|
|-----------|------------|---------|-------------|----|----------|------|

|              |               |          |            |              |        |   |
|--------------|---------------|----------|------------|--------------|--------|---|
| HK187721009  | 192.168.137.1 | WinNT-SP | 5.0.2195   | 000187721009 | DF-1A  | 0 |
| dal01man0001 | 10.224.64.4   | SunOS    | 5.8        | 000187721009 | FA-9A  | 0 |
|              |               |          |            | 000187721055 | FA-9A  | 1 |
| cs0          | 10.224.64.61  | CelerraL | 2.4.9-34.5 | 000187721009 | FA-7C  | 0 |
|              |               |          |            | 000187721009 | FA-8C  | 0 |
|              |               |          |            | 000187721009 | FA-7D  | 0 |
|              |               |          |            | 000187721009 | FA-8D  | 0 |
|              |               |          |            | 000187721009 | FA-9D  | 0 |
|              |               |          |            | 000187721009 | FA-10D | 0 |

**#/nas/symcli/bin/symcfg list -lock**

**#/nas/symcli/bin/symcfg list -meta**

## **USING SOLUTION ENABLER v6.0.3 to PRESENT SYMM DEVICES (LUN Masking) FOR WINDOWS OR CELERRA HOSTS:**

**Note:** Solution Enabler uses the symmask API to perform the work

1. Load Symmetrix BIN file for the new configuration & set FA bits: PP/UWN/EAN/C; SC-3 for heterogeneous; ARB for Celerra
2. Use Solution Enabler v6.0.3 to run symmask command

### **For Windows Hosts:**

```
symmask -sid 0178 -wwn 10000000C928CACF -dir 04b -p 0 add devs 005CA:005D1
symmask -sid 0178 -wwn 10000000C926D762 -dir 13b -p 0 add devs 005CA:005D1
```

### **For Celerra Hosts:**

```
symmask -sid 0178 -wwn 5006016039a004f5 -dir 04b -p 0 add devs 0005:00F7,05B9:05BA -celerra
```

3. Host Action after performing LUN Masking:

### **For Windows Hosts:**

Reboot Windows system

### **For Celerra Hosts:**

Perform NAS install

## **VERIFYING CELERRA HBA LOGINS TO SYMMETRIX FA's:**

**\$ .server\_config server\_2 -v "fcp bind show"**

\*\*\* Persistent Binding Table \*\*\*

Chain 0000: WWN 5006048acc193b58 HBA 0 FA-09da

**Note:** Run “fcp show” to see which DM WWN’s are associated with each FA

**\$ ./symmask list logins -sid 941 -dir 9d -p 0** (FA-09da = 9d –port 0)

Symmetrix ID : 000187720941

Director Identification : FA-9D

Director Port : 0

| Identifier       | Type  | Node Name        | User-generated Logged On |        |     | Fabric |
|------------------|-------|------------------|--------------------------|--------|-----|--------|
|                  |       |                  | Port Name                | FCID   | In  |        |
| 10000000c939ed49 | Fibre | NULL             | NULL                     | 613d13 | Yes | Yes    |
| 10000000c93d3dac | Fibre | NULL             | NULL                     | 614113 | Yes | Yes    |
| 10000000c9483d88 | Fibre | 10000000c9483d88 | 10000000c9483d88         | 614913 | No  | Yes    |

**./symmask list logins -sid 941 -dir 9a -p 1** (port 1 on FA 9a)**VERIFYING REMOTE CSHOSTS:**

\$cat /nas/site/cshosts

**Finding Celerra & Symmetrix Information for SCO Installations: [/nas/symcli/bin]**

|                               |                                                                        |
|-------------------------------|------------------------------------------------------------------------|
| \$/nas/symcli/bin/symcfg list | [Lists out Symmetrix ID, Model, basic Microcode level, device summary] |
| \$/nas/symcli/bin/symcfg -h   | [Listing of switches available]                                        |
| \$./symcfg list -v            | [Verbose listing of Symmetrix characteristics]                         |
| \$./symcli -v                 | [Prints out summary of each command available with SYMCLI]             |
| \$/nas/server/slot_x/scsidevs |                                                                        |
| \$/nas/bin/nas_symm -l        | \$/nas/bin/nas_symm -i "serial number"                                 |

**HOW TO CHECK TO SEE IF SYMAPI DATABASE COMPONENTS ARE LOCKED:**

# /nas/symcli/bin/symcfg -semaphores list [or semaphores -list]

| S Y M A P I   S E M A P H O R E S |            |          |      |      |                                 |
|-----------------------------------|------------|----------|------|------|---------------------------------|
| Lock Proc.                        |            |          |      |      |                                 |
| ID                                | Key        | State    | Type | Wait | Lock Full Path Specification    |
| 24                                | 0x45601565 | Unlocked | GK   | 0    | /nas/dev/c0t15l15s2             |
| 22                                | 0x45601b7e | Unlocked | DB   | 0    | /var/sympapi/db/sympapi_db.bin  |
| 104                               | 0x4560155a | Unlocked | FILE | 0    | /var/sympapi/config/sympapislck |

  

| ----- Semaphore Arrays ----- |       |        |       |       |        |
|------------------------------|-------|--------|-------|-------|--------|
| key                          | semid | owner  | perms | nsems | status |
| 0x00000000                   | 0     | apache | 600   | 1     |        |
| 0x450540e8                   | 32769 | root   | 666   | 1     |        |
| 0x45054154                   | 65538 | root   | 666   | 3     |        |

-----output abridged-----

**SYMMETRIX INLINE COMMANDS:**

Logging In brings you to the Default Director [e.g., DA-15a\* = Default Director for INLINES]

\$ sign commands are “global” commands , Used to indicate the end of a command

|                  |                                                                                                                  |
|------------------|------------------------------------------------------------------------------------------------------------------|
| <b>\$RS</b>      | [Returns you to Default Director]; \$Rx [Enables you to change Director—e.g., \$R2A]                             |
| <b>E0,</b>       | [Online/Offline director status—Displays status of all Dir’s {0A/0B v. 00}, Memory Cards, Cables connected {X} ] |
| E7,              | [DA Host&Disk subsystems—Listing of all Primary and Secondary devices controlled by the DA logged into]          |
| T#               | are SCSI target numbers for that Director                                                                        |
| ‘Shadow devices’ | are listed as grayed-out secondary devices for its partner DA                                                    |
| D2,tt            | [tt=T# number of device]—Physical disk info of the logical device                                                |
| <b>9B,A</b>      | [Displays Error & Event Codes logged at each Director—errors logged on all Directors]                            |
| 99,A             | [Displays Error Codes logged at each Director]                                                                   |
| E1,,E,A          | [Fatal errors on all Directors]                                                                                  |
| E6               | [Provides errors logged by the indicated Director]                                                               |
| <b>A7,D</b>      | [Location of all Drives, Directors, and SCSI—Information on all Logical Devices configured on the SYMM]          |
| A7,              | [Displays all devices with anything irregular in status--Not Ready; Write Disabled; Write Pending, Invalids]     |
| A4,              | [Write Pending in Cache]                                                                                         |
| A7,FE            | [Physical device information such as SCSI code & Splits]                                                         |
| A7,’LGND’        | [Legend of Abbreviations used in the various output columns]                                                     |
| A7,’EMUL’        | [Displays Emulation & Cylinder Size for each SYMM Volume; 3390, FBA, etc]                                        |
| <b>A7,C</b>      | [Status of logical vol., Invalid Tracks, Not Ready, Write Disabled, Write & Format Pending, etc—all Directors]   |
| A7,DC            | [BCV status]                                                                                                     |
| E7,D,A           | [EEPROM & IMPL revision levels for all directors—Verifying running code on System]                               |

A7,AAAA [Listing of A7 Inlines commands]  
 E7,AAAA [Listing of E7 Inlines commands]  
 A7,FD [device table as defined by IMPL file]  
 E1,,AB,AB [memory config]  
 A7,B [all SA's & assoc.devices]  
 A7,'GRUP' [RAID-S Groups—3 data & 1 parity per group]  
 A1,B [Target & LUN addresses]  
 A2,BA,ADD [I/O on all SCSI Channel Directors]  
 A2,CA [Host activity & I/O on Channel Directors]  
 E7,CF [Basic SYMMETRIX Configuration Information]  
 A7,DC [Shows the Standard Volumes and the corresponding BCV volumes]  
 8D,,,,'GDAT' [Shows history of BCV establishes and splits]  
**FD,CE** [FAST IML of a Single Director]  
**\$F0,CE,0** [Bringing single director back online after Fast IML]  
 How to Stop a Running Script? Esc key and/or Menu>Sympl>Halt script

#### **INLINES to Run Whenever First Logging Into a Symmetrix and Before Logging Out:** ‘Symmetrix Keys to Success’

9B,A Look for new Errors & Events by Director & Time [eg, 0472=Environmental RealTime errors]  
 E0, Verify that all Directors are alive and all Memory available. Verify cable connections.  
 A7,C Verify Ready status of drives. Check for Invalids on valid mirrors.  
 A3 Verify Cache slots are o.k.—look for Errors listed in RED  
 A4 Verify that ‘Invalid Cache Flags Count’ is 0  
 E1,,AB,AB Look for fragmented memory in Error column  
 E1,,DB Determine if any Memory is disabled—look for values ‘other than 0’ in “mem enable table”  
     Each ‘ADDR SLOT MAP’ horizontal location = 256MB of RAM  
 E1,,E,A Look for “Director Fatal Errors” in the “EXCEPTION\_COUNTERS” section  
 9B,A Verify that there are no Director errors  
 E7,D,A Verifies Microcode on each Director’s “Control Store” [RUNNING CODE], to EEPROM, giving “CODE OK!” msg  
     [Code is first loaded on Service Processor, then on each Director using Procedures>Code Load Procedures]  
 A1,DA/A1,CA/ A1,F3 Look for any RED Values indicating errors  
 E7, Run from RA director--verify that SRDF Flags are set properly [S column, red NR=link down]  
 \$A0, Clears all Test parameters that may have been used

#### **LINUX-CELERRA VOLUMES—OLD STYLE:**

##### **[6 regular volumes, 1 GateKeeper 7cyl. volume]: 1-6=4430 Cylinders or 2GB:**

|                                                      |                                                                                                                                                         |
|------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------|
| Volume 00=Hyper1 CS0 only; Target 0/LUN 0            | 4430 Cylinders; 2 GB; SCO O/S; CS0 Boot partition                                                                                                       |
| Volume 01=Hyper2 CS1 only; Target 0/LUN 0            | 4430 Cylinders; 2 GB; SCO O/S; CS1 Boot partition                                                                                                       |
| Volume 02=Hyper3 CS0&1 as Target 0/LUN 1; DM's T0/L0 | Shared NAS Work Partition—DOS & DART; DM boot & rootfs; Dump; map areas--CS's Read & Write;<br>CS Logs. Sys. DB Partition: Celerra Config. & Swap files |
| Volume 03=Hyper4 CS0&1 as Target 0/LUN 2             | All DataMovers [File System Transactions UxFS]                                                                                                          |
| Volume 04=Hyper5 DM's as Target 0/LUN 1              | Additional logs for Panics, Event Logging, etc.                                                                                                         |
| Volume 05=Hyper6 CS0&1, Target 0/LUN 3               | 7Cylinders; TID FF; AS/400; SRDF, TimeFinder, C.Monitor                                                                                                 |
| Volume 07=optional Gatekeeper Volume; TID FF         |                                                                                                                                                         |

**Note:** All (7) Volumes are usually created on one physical drive

#### **LINUX CELERRA VOLUMES—OLD STYLE:**

Device 00 Linux CS0 Mapped to SCSI Port 16A, Processor A, Port A; Target 0, LUN 0  
 Device 01 DOS Boot Volume for DataMovers Mapped to SCSI Port 16A, Processor A, Port A; CS0=01; DataMovers=00  
 Device 02 Linux Log Files for Failover Mapped to 16A, Processor A, Port A; CS0 02  
 Device 01 DOS Boot Volume for DataMovers Mapped to SCSI Port 15A, Processor A, Port A; CS1=01; DataMovers=00  
 Device 02 Linux Log Files for Failover Mapped to SCSI Port 15A, Processor A, Port A; CS0=02  
 Device 03 Linux CS1 Mapped to SCSI Port 15A, Processor A, Port A; CS1=00  
 Device 04 Data Mover Log Files Mapped to all SCSI Ports except 15A& 16 A, Processor A, Port A; DataMovers=01  
 Device 05 TBD Mapped to SCSI 15a & 16A, Processor A, Port A; CS0 & CS1=03  
 Device 06 DataVolumes Mapped to all SCSI Ports except 15A & 16A, Proc. A, Port A; DataMovers=10 and up numerically

#### **CELERRA BIN FILES--After NAS 2.2.25.5+:**

VCM GateKeeper TID = OE for all ports

**Note:** This is a change—CS now sees all volumes that DM sees

Regular GateKeeper TID = OF for all ports

Data Volumes should be addressed as TID 10+ for all ports

Control Volumes for DM & CS: Two Volumes--8860 cyl, TID 00 & 01

Volume 00 & 01: All DM's and CS's. 8860 cylinders. TID 00 & 01

Volume 02, 03, 04, 05. 4430 cylinders. TID 02 – 05.

## **HOW TO DETERMINE IF AS400 BIT IS SET/ON FOR CELERRA GATEKEEPERS:**

BIN File>Configuration>Edit Volumes>vol "02E">Go to Heading: "Flags 2. As400 Gate Keeper"

Gate Keeper bit is "ON" if "GK" is capitalized but is off if "gk" is lowercase

### **GateKeeper Device [AS400 Volume]:**

*Looking at the System Impl Volumes Map: configuration>VolumesMap>Channel Addresses tab*

| Volume | Emulation | Size | SA,5A | SA,5B | SA,12A | SA,12B |    |
|--------|-----------|------|-------|-------|--------|--------|----|
| 000    | FBA       | 4430 | 00    | --    | --     | --     | -- |
| 001    | FBA       | 4430 | --    | --    | --     | --     | -- |
| 002    | FBA       | 4430 | 01    | --    | 00     | 00     | -- |
| 003    | FBA       | 4430 | 02    | --    | --     | --     | -- |
| 004    | FBA       | 4430 | 03    | --    | --     | --     | -- |
| 005    | FBA       | 4430 | --    | --    | 01     | --     | 01 |

**Purpose of GateKeeper Devices:** Allows Celerra to execute SCSI commands in support of TimeFinder & Celerra Monitor.

## **CELERRA TIMEFINDER/FS [Snapshot]:**

Symm 4.0+ and 5265+ code; 5266 code introduces "InstantSplit"

New min. microcode levels for Timefinder & SRDF is 5567.24.13.S

**Chronological TimeFinder Process:** Establish, Synchronize, Freeze, Split, Thaw

### **TIMEFINDER PURPOSE:**

→Creating point-in-time copy (snapshot) of Celerra file system onto dedicated BCV devices from R1 devices

**Note:** i.e., fs\_timefinder creates copy of fs or fs group that can be mirrored with original pfs

→Timefinder/FS copies are used for Backups, Replication of data for live testing, mirrored file system updating

**Note:** Copies can be mounted and made accessible over the network

### **TIMEFINDER/FS REQUIREMENTS:**

--Use only one filesystem per STD volume for Timefinder/FS

--Requires symm backend and STD volumes (R1 or R2 devices)

--Do not use nas\_slice command when creating file systems

--Snapshots occur only with FileSystems from STD to BCV which contain Metas

--Only a single Snapshot can take place at a time

--Can only restore to a single FileSystem

--Avoid slicing when using AVM by specifying slice=n when creating file systems

--BCV's are used as mirror devices for STD device

--BCV volumes must be the same size as the volumes used by the production file system

--Whenever changing device type from BCV to STD, delete using nas\_disk -d -p, then vtoc volume, then readd using devconfig -c

**NAS 2.1 & 2.2 Commands:** nas\_fs -S; nas\_fs -M on {off}; nas\_fs -R [NAS 2.1.x; 2.2.x]

**NAS 4.0.12.1 Commands:** fs\_timefinder -S; fs\_timefinder -M on | off |refresh; fs\_timefinder -R | fs\_group\_name

**Note:** NAS 4.0.12.1 introduces support for grouping of multiple filesystems for TimeFinder/FS operations

### **DEVICE REPLACEMENT FOR FAILED BCV DRIVES:**

--Replace disk for failed BCV device

--PSE Lab needs to do full establish for that hyper only (clear GDAT tables)

--Other BCVs associated with failed device will also need to be established and split using fs\_timefinder --refresh, which recreates diskmark on the replaced device

### **TIMEFINDER/FS OPERATIONS:**

--automatically "splits" after the snapshot copy is complete (fs\_timefinder -S)

--Incremental copies are made using --mirror on or refresh switches

--Split by using the --mirror off command

--the --restore command will copy the BCV to STD device (no fs restore, just entire volume, which should only consist of one fs)

--BCV's are automatically selected when --Snap is run, but can manually create BCV volumes on same structure as STD

--Snapshot is synonymous with a ‘full establish’ (fs mirrored) and split (fs becomes ufs again)  
--With the mirror on command, the BCV becomes “mirrors” type (fs01\_snap2 -M on)  
--With the mirror off, the BCV becomes “ufs” type (use #fs\_timefinder -i fs01\_snap1 to query state of volume)  
--Later NAS codes do not lock the GK during the entire refresh or establish operation (only at start & end of sequence)  
--When using the –Restore –Force command, the BCV must be ufs and “unmounted” (cannot be in mirrored on state)  
--When fs\_timefinder operations are in progress, the STD & BCV devices are added to a “device group” which becomes part of the symapi\_db.bin, and then deleted from the symapi\_db.bin after the devices are mirrored off or split (1REG are normal BCV devices & 1R1\_1 are where BCV’s are group as RDF BCV devices)  
--#/nas/symcli/bin/symmir –g 1REG or 1R1\_1 query shows BCVs as fs type=1 when split, or fs type=6 when mirrored  
--#/nas/symcli/bin/symdev show <device> will show the state of the STD==>BCV pair, with “synchronized” meaning “mirrored”  
--#/nas/symcli/bin/symdev –bcv list will show NR for ‘Not Ready’ for hosts, meaning it is in an “established” state (also, symbcv dev list)  
--After a mirror on establish, the BCV device will move into the M3 position on the paired STD device  
--After a split mirror off, the BCV M3 position will be deleted from the paired STD  
--SDDF tables keep track of changes between STD & BCVs  
**Note:** Please note that at the end of each refresh operation, when the BCV device is being split from the standard device, it will have its former diskmark replayed from the Camdisk file and reassigned to itself. A diskerr:mismatch:tid error can be seen in a Server devconfig –probe if the diskmarks have not been correctly replayed—See AR69029  
tid/lun= 0/12 type= disk sz= 172631 val= 75 info= 56705532F560 diskerr= mismatch:tid  
1128101311: STORAGE: 3: IsMarkerValid(): BadDiskMark vname:54 disk\_id:54 marker:6  
1128101311: STORAGE: 3: 1: Error: Path:c32t4l3 for volume:54 not validated (Path/DiskMark inconsistent)

## **FS TIMEFINDER TROUBLESHOOTING:**

[/nas/log/symapi.log](#) is a good place to see Incremental\_Establish, pairings, Succeeded, SPLIT operations, etc.

Other important logs: [cmd\\_log](#), [cmd\\_log.err](#), [nas\\_log.al.trace](#)

**# export SYMAPI\_DEBUG=-1;export SYMAPI\_DEBUG\_FILENAME=debug\_sep05.log** (debug logging)

## **NAS UPGRADES:**

--All BCVs must be split prior to upgrades, otherwise the devices will not be visible to Celerra for the setup\_slot discovery

**Note:** BCVs can only be seen via probe when “mirrored off” or “ufs”. Look for any volumes in the filesys file as type 6

**# grep ^6 /nas/volume/filesys**

## **GROUPING FILESYSTEMS TOGETHER FOR TIMEFINDER OPERATIONS: NAS 4.0.12.1**

### **CREATING GROUP FILE SYSTEMS:**

**\$ fs\_group -n grp1 -c fs1 fs2 fs3**

**Results:** grp1 [Type 100 FileSystem]

### **SNAPPING GROUP FS:**

**\$ fs\_timefinder grp1 -S**

**Results:** grp1\_snap1

### **RESTORING GROUP FS:**

**\$ /nas/sbin/rootfs\_timefinder grp1\_snap1 -Restore -Force**

### **REFRESHING GROUP FS:**

\$nas\_fs grp1\_snap1 -Mirror refresh

**\$ fs\_timefinder grp1\_snap1 -Mirror refresh** [alternative command]

**Verify:** \$fs\_group -i grp1\_snap1

### **OTHER COMMANDS:**

fs\_group -shrink; fs\_group -xtend; fs\_group -delete; fs\_group -l; fs\_group -i

## **COMMAND OUTPUT EXAMPLES:**

**# fs\_group -l**

```
id    name      acl in_use type member_of fs_set
24   KF2_nggf_all 0  n   100      18,19,20,21,22,23
```

**# fs\_group -i KF2\_nggf\_all**

```
id      = 24
name    = KF2_nggf_all
acl     = 0
in_use  = False
type    = group
```

**fs\_set = nggf\_fs01,nggf\_fs02,nggf\_fs03,nggf\_fs04,nggf\_fs05,nggf\_fs06**

rw\_servers=

ro\_servers=

symm\_devs = 000183700172-00E,000183700172-013,000183700172-018,000183700172-01D,

## **GROUPING FILESYSTEMS TOGETHER FOR TIMEFINDER OPERATIONS: NAS 4.2**

--One filesystem or fs\_group can belong to multiple fs\_groups

--Can add new FS to fs\_group and then conduct Snap Refresh to update snap volume

## **TIMEFINDER/FS METAS NEED TO BE IDENTICAL ON BOTH STD & BCV VOLUMES:**

**Problem:** Timefinder/FS fails to complete Snap with screen error: Error 2231: Ecap1\_fs02 : invalid duplicate [Ecap1\_fs02 = BCV Meta Volume]. The problem is that all production filesystems were built from regular metas, such as d11, d12, etc., except the filesystem in this case, which was built from stripe volumes and then a meta. Since the corresponding BCV Meta was not built on the same stripe volume layout, the TimeFinder/FS command failed with the "invalid duplicate" message.

**Solution:** Deleted the BCV Meta & recreated using the same Striping scheme as used to build the PFS that was being snapped.

## **RDF CHAND PROCESS:**

1. Verifying Process: #ps -ef |grep chan [Should be running on R1 & R2 side]

2. Starting Process: #/nas/sbin/rdf\_chand -c rdf\_channel

**Note:** RDF\_Chand process is required for TimeFinder functions, not nas\_rdf init, activate, or restore commands

## **SRDF TIMEFINDER/FS FARCOPY & NEARCOPY SUPPORT: NAS 2.2.49.2/4.0.12.1 [Uses diff. commands]**

**SRDF FARCOPY ++:** GA with 2.2, 4.1, 4.2

## **CREATING A FILE SYSTEM COPY USING TIMEFINDER (4.0):**

Step 1. For CIFS File Systems, it is recommended that you first update the production File System CIFS attributes

**\$server\_cifs server\_2 -update /home mindirsize=0**

Step 2. Create the File System Copy: **\$fs\_timefinder fs01 -S**

**Note:** Once a timefinder command is issued, do not stop the process! The default BCV volumes are used unless otherwise specified, such as in this example: \$fs\_timefinder fs01 -S -volume mtv1

## **MIRRORING A FILE SYSTEM USING TIMEFINDER(4.0): Synchronizing orig. file system with timefinder copy**

Step 1. Creating the Mirror: **\$fs\_timefinder fs01 -M on | off | refresh** [automatically conducts mirror on, mirror off]

**Note:** Normally, the mirrored file system copy should be "unmounted", but use the following command if this cannot be done:

**\$fs\_timefinder fs01 -M on -Force**

In case of mirroring problems, contact Symmetrix PSE Lab to complete the establish/split operation manually. Basically, PSE can conduct a complete Establish to sync up all volumes, then we can Split the mirrorfs using our commands.

## **PROBLEM WHERE DISK DRIVE FAILS & HOT SPARE IS INVOKED:**

1. Problem may manifest itself with "smbcv" list showing invalids for bcv meta heads used to make up bcv's of file system (s)

2. Replace disk drive

3. Establish file systems

4. Split bcv's from production file systems

## **RESTORING A FILE SYSTEM FROM A TIMEFINDER COPY:**

**\$fs\_timefinder fs01 -Restore -Force**

## **REFRESHING A FILE SYSTEM COPY(4.0):**

**\$fs\_timefinder fs01 -M refresh**

## **RUNNING REFRESH ON R2 SIDE R2BCVs:**

**\$fs\_timefinder fs01 -M refresh -o protect=no**

## **RUNNING REFRESH ON R1 SIDE:**

**\$fs\_timefinder fs01 -M refresh -o**

## **REMOTE TIMEFINDER/FS--NEAR COPY & FAR COPY:**

Purpose is to provide remote copies of data for Backups, Application Testing, Data Warehousing

## **REMOTE TIMEFINDER/FS (FARCOPY):**

Far Copy uses R1BCV devices on local Active Celerra, snaps a copy of the local file system onto one of these R1BCV's [command must specify the R1BCV's as opposed to regular BCV's or this whole process will not work!], then uses the RDF link to Import the snapped File System across the link to the Remote Celerra's R2BCV's. From there, the file system can be snapped onto a local R2 Celerra BCV Volume and exported for use in a RO or RW mode. This snapped filesystem can then be routinely "refreshed" or synchronized with the original file system on the R1 side, and restored from Remote to Local side.

## **FAR COPY/EXTENDED DISTANCE REMOTE TIMEFINDER/FS/R-2COPY: 2.2/4.x**

### **Farcopy:**

- Requires SRDF adaptive copy write-pending mode (extended distance) configuration between local and remote symmetrix
- Local R1 side may have local BCVs and R1BCVs
- FS snapped from local STD to Local R1BCV's
- FS imported from R1BCV to R2BCV on Destination side
- Imported FS is snapped to local STD volumes on Destination side
- ACTIVE/PASSIVE ASYNCHRONOUS MODE
- NOT A DISASTER RECOVERY SOLUTION
- Supports Semi-Synchronous (Journal1) mode—1<sup>st</sup> write to local client acknowledged, but 2<sup>nd</sup> write not acknowledged until Symm replies to 1<sup>st</sup> write
- Supports Adaptive Copy mode—writes acknowledged to local client & cached until written to Symm storage

## **SETTING UP FARCOPY & ESTABLISHING REMOTE FILE SYSTEMs (NAS 5.5):**

1. Have BIN files applied to each Symmetrix to create local R1BCV's on Active Side & R2BCV's on Passive Side
2. Establish RDF Link between Symms & IP Data Network between Control Stations
3. Initialize Remote Celerra from Local side:

**# /nas/sbin/nas\_rdf –init Celerra\_remote 10.241.169.51**

**Note:** Create “rdfadmin” account and give same password as name. Select passphrase when prompted. Verify visibility of the Remote Celerra using nas\_cel –list. Should see Local and Remote Celerras listed.

4. Initialize Local Celerra from Remote side:

**# /nas/sbin/nas\_rdf –init Celerra\_local 10.241.169.50**

**Note:** Create “rdfadmin” account and passphrase, using same passwords are in step 3. Run # nas\_cel –list—should see both systems in the output.

5. Create local snapshot of production file system:

**\$ fs\_timefinder fs1 –Snapshot –option disktype=R1BCV**

**Note:** This automatically creates the mirrored relationship between R1BCVs and R2BCVs—note that FarCopy requires snapping to local R1BCVs and not local BCVs.

6. Refresh the R2BCV with changes from R1BCV & to resume communication between device pairs over RDF:

**\$ fs\_timefinder fs1\_snap1 –Mirror refresh**

7. Import the refreshed snapshot of the production file system to the destination R2BCVs from the Remote Celerra:

**\$ cel\_fs Celerra\_local –Import fs1\_snap1**

**Note:** Use \$ nas\_disk -l to see devices that have been imported to Remote Celerra as “R2BCV rootdxxx”, etc

8. Create a local snapshot of the imported file system on the Remote Celerra:

**\$ fs\_timefinder fs1\_snap1 -Snapshot**

**Note:** This will snap a copy of the file system located on the R2BCVs to the local STD devices

9. Mount and export the file system as needed

**Note:** File system will be mounted RO by default, but you can specify to mount RW

10. Refreshing changes to original Source file system and importing changes to Remote Celerra:

**\$ fs\_timefinder fs1\_snap1 –Mirror refresh**

**\$ fs\_rdf fs1\_snap1 –Mirror refresh** (imports updates from Source R1BCVs to Destination R2BCVs)

11. Update the Destination File System (located on STD volumes) from R2BCVs:

**\$ fs\_timefinder fs1\_snap1 –Mirror refresh**

**Note:** At this point, the updated changes in the source file system snap are transferred to the destination snap filesystem

## **RESTORING FROM REMOTE SNAPSHOT TO LOCAL R1BCV SNAPSHOT:**

Unmount and unexport remote snapshot file system

**\$ /nas/sbin/rootfs\_timefinder fs1\_snap1 –Restore** (Restoring Remote R2BCV from snapshot file system)

**\$ /nas/sbin/rootfs\_rdf fs1\_snap1 –Restore** (Restoring remot R2BCV copy to local R1BCV)

**FAR COPY BIN FILE ARRANGEMENT:****R1 ACTIVE SOURCE SYMMETRIX:**

Data Volumes are seen as R1BCV's locally and on R2 Side: 00c0→2DF (RA Group A)

R1Control Volumes are Mirrored: 015E→14E (RA Group A)

R2Control Volumes are Mirrored: 043F→423 (RA Group B)

**TROUBLESHOOTING FARCOPY OPERATIONS:**

**Note:** Lun0 should be in synchronous mode for rdf\_chand process

1. Conduct nas\_fs -l to see if any filesystems are in a status 6, which is mirrored, when they should be split

**id inuse type acl volume name server**

**27 n 6 0 186 uslist22\_BCV\_ftp**

2. Run **nas\_fs -i uslist22\_BCV\_ftp**

**# /nas/bin/nas\_fs -i uslist22\_BCV\_ftp**

id = 27

name = uslist22\_BCV\_ftp

acl = 0

in\_use = False

**type = mirrorfs [Means filesystem is mirrored and not split]**

volume = uslist22\_mtv01\_BCV

rw\_servers=

ro\_servers=

backup\_of = fs01 Thu May 15 05:32:49 CDT 2003

remainder = NOT ACCESSIBLE

symm\_devs = 000185700111-02F8,000185700111-02F9,000185700111-02FA,000185700111-0

3. Run List of BCV Devices:

**# /nas/symcli/bin/symbcv list**

|                    |                      |                      |             |                       |
|--------------------|----------------------|----------------------|-------------|-----------------------|
| <b>Not Visible</b> | <b>0314 R1 (M) -</b> | <b>1 Not Visible</b> | <b>02FC</b> | <b>0 Invalid</b>      |
| <b>Not Visible</b> | <b>0315 R1 (m)</b>   | <b>- Not Visible</b> | <b>02FD</b> | <b>- Synchronized</b> |
| <b>Not Visible</b> | <b>0316 R1 (m)</b>   | <b>- Not Visible</b> | <b>02FE</b> | <b>- Synchronized</b> |
| <b>Not Visible</b> | <b>0317 R1 (m)</b>   | <b>- Not Visible</b> | <b>02FF</b> | <b>- Split</b>        |

4. Try to mirror off the offending filesystem:

**# /nas/bin/nas\_fs -M off uslist22\_BCV\_ftp**

operation in progress (not interruptable)...Error 2201: unable to acquire lock(s), try later

5. Look for Devices that are Locked—status “9”:

**# /nas/symcli/bin/symdev -lock list**

Symmetrix ID: 000185700188

S Y M M E T R I X D E V I C E L O C K S

Device Name Device

| Sym  | Physical    | Lock Number | Lock Holder ID | Seconds Held | Device Flags | Attribute   |
|------|-------------|-------------|----------------|--------------|--------------|-------------|
| 02FC | Not Visible | 9           | 0x0E0236EB     | 1362759      | None         | Grp'd (M)   |
| 02FD | Not Visible | 9           | 0x0E0236EB     | 1362759      | None         | N/Grp'd (m) |
| 02FE | Not Visible | 9           | 0x0E0236EB     | 1362759      | None         | N/Grp'd (m) |
| 02FF | Not Visible | 9           | 0x0E0236EB     | 1362759      | None         | N/Grp'd (m) |

6. Release the locks on STD & BCV volumes by using the following:

**# /nas/symcli/bin/symdev -sid 0188 -lock 9 -RANGE 1<sup>st</sup>#:2<sup>nd</sup># -noprompt release**

[Alternatively, release one volume at a time: #/nas/symcli/bin/symdev -sid 1118 -lock 9 release ]

7. Contact PSE Lab to establish all devices in same Meta groups making up filesystems

8. Once synchronized, /symbcv list should show all devices for PFS and Snap filesystems in synchronized state

9. Snap Volume should now mirror off:

**# nas\_fs -M off uslist22\_BCV\_ftp**

**Verifying RDF Chand Process is Running on Both Sides:**

**# ps -ef |grep rdf\_chand**

root 1296 1124 0 Mar16 ? 01:42:02 /nas/sbin/rdf\_chand -c rdf\_chann

**EXAMPLE OF REGULAR & R2BCV'S ON R2 SIDE:**

```
133 n 13943 000185700967-35F BCV rootd133 1,2,3,4,5
134 n 13943 000185700967-360 BCV rootd134 1,2,3,4,5
139 y 13943 000185700967-2DF R2BCV rootd39
140 y 13943 000185700967-2E0 R2BCV rootd40
```

**LISTING OF RDF GROUPS & ASSOCIATED SYMMS:****\$ /nas/symcli/bin/symdg list****DEVICE GROUPS**

Num of Num of

| Name      | Type    | Valid | Symmetrix ID | Devices | GK's | BCV's |
|-----------|---------|-------|--------------|---------|------|-------|
| 1R1_2     | RDF1    | Yes   | 000185700967 | 6       | 0    | 0     |
| 1REG      | REGULAR | Yes   | 000185700967 | 110     | 0    | 22    |
| cs_500    | RDF2    | Yes   | 000185700967 | 1       | 0    | 0     |
| 1R2_500_1 | RDF2    | Yes   | 000185700967 | 5       | 0    | 108   |

**\$ /nas/symcli/bin/symcfg list****SYMMETRIX**

Mcode Cache Num Phys Num Symm

| SymmID       | Attachment | Model | Version | Size (MB) | Devices | Devices |
|--------------|------------|-------|---------|-----------|---------|---------|
| 000185700967 | Local      | 8830  | 5567    | 32760     | 10      | 1065    |
| 000184500694 | Remote     | 8530  | 5567    | 8184      | 0       | 204     |
| 000184702190 | Remote     | 8730  | 5567    | 32760     | 0       | 1093    |

**# /nas/bin/nas\_symm -l****# /nas/sbin/rootnas\_symm -l:**

id acl name

**000183700172 0 000183700172****000184701179 0 000184701179****# /nas/bin/nas\_cel -l**

id name owner mount\_dev channel net\_path CMU

|   |          |     |           |                                      |
|---|----------|-----|-----------|--------------------------------------|
| 0 | padmesc6 | 0   | 10.28.4.4 | 00018570455406B5                     |
| 1 | padmesc1 | 500 | /dev/ndj1 | /dev/ndg 10.3.4.166 000187720368066B |
| 2 | padmcsa1 | 501 |           | 172.18.4.166 0001874306850041        |

**# /nas/sbin/rootnas\_cel -l:****nas/rdf # more cshosts:****1:cnt1:500:57.8.109.135:/dev/sdj1:/dev/sdg:****CSHOSTS FILE & NS702G ATTACHED TO SYMMETRIX:****# cat /nbsnas/site/cshosts**

1:padmesc1:500:10.3.4.166:/dev/ndj1:/dev/ndg:000187720368066B:

2:padmcsa1:501:172.18.4.166:::0001874306850041:

0:padmesc6:0:10.28.4.4:::00018570455406B5:

**#/nas/symcli/bin/symcfg list :****SYMMETRIX**

Mcode Cache Num Phys Num Symm

| SymmID       | Attachment | Model | Version | Size (MB) | Devices | Devices |
|--------------|------------|-------|---------|-----------|---------|---------|
| 000183700172 | Local      | 3930  | 5267    | 12280     | 16      | 606     |
| 000184701179 | Remote     | 8730  | 5567    | 16376     | 0       | 866     |

**# /nas/symcli/bin/symrdf -g 1R1\_2 query****Device Group (DG) Name: 1R1\_2**

DG's Type : RDF1

DG's Symmetrix ID : 000183700172

Source (R1) View Target (R2) View M O D E S

| ST       | LI       | ST     | M             |
|----------|----------|--------|---------------|
| Standard | A        | N      | O             |
| Logical  | T R1 Inv | R2 Inv | K             |
| Device   | E Tracks | Tracks | S Dev         |
|          | E Tracks | Tracks | E Tracks      |
|          |          |        | Dom ACP STATE |

d1 006 RW 0 0 RW 00A WD 0 0 SYN DIS OFF Synchronized

```
d2 007 RW 0 0 RW 00B WD 0 0 SYN DIS OFF Synchronized  
csroot 008 RW 0 0 RW 00C WD 0 0 SYN DIS OFF Synchronized  
cs009 009 RW 0 0 RW 00D WD 0 0 SYN DIS OFF Synchronized  
cs00A 00A RW 0 0 RW 00E WD 0 0 SYN DIS OFF Synchronized  
cs00B 00B RW 0 0 RW 00F WD 0 0 SYN DIS OFF Synchronized
```

## **REMOTE TIMEFINDER/FS--NEARCOPY CAMPUS MODE (60km limit):**

- NearCopy Supports Campus Mode (Writes are acknowledged to Client only after remote Symm acknowledges as written to disk)
- Active/Passive SRDF/S with Synchronous operation for DR purposes
- Bin files same as for regular Active/Passive SRDF

### **Basic Operation:**

- Step 1. R1 Celerra, discover SRDF-linked Celerras: \$nas\_cel -list
  - Step 2. From R2 Celerra, choose the R1 File System to be imported: \$cel\_fs R1Celerra -list [choose fs01\_orig]
  - Step 3. R2 Celerra, Import File System from R1 Celerra to R2STDs: \$cel\_fs R1Celerra -Import fs01\_orig
  - Step 4. Create Local TimeFinder File System Copy to BCVs: \$fs\_timefinder fs01\_orig -S
  - Step 5. Mount and export the Remote Copy on the R2 Celerra
- Note:** By default timefinder copies are mounted as "Read Only"--use -F switch to force RW mount

## **SETTING UP NEARCOPY & ESTABLISHING REMOTE FILE SYSTEMs (NAS 5.5):**

1. Setup SRDF/S in Synchronous Mode Campus configuration for DR
2. Initialize Active/Passive Celerra configuration from Source:

### **\$ /nas/sbin/nas\_rdf -init**

**Note:** At prompt, create the RDF standby Servers from the Remote side that will be used for failover

3. Run RDF Initialization again to establish RDF relationship between Source & Remote Celerras:

### **\$ /nas/sbin/nas\_rdf -init Celerra\_remote 10.241.169.51**

**Note:** At prompt, create "rdfadmin" account & password, and Passphrase (this step not used with NAS 5.3)

4. Initialize Remote Celerra:

### **\$ /nas/sbin/nas\_rdf -init**

**Note:** At prompt, create "rdfadmin" account & Passphrase as in previous step, and create rdf\_standby relationships

5. Verify configuration by running \$ nas\_cel -list on both sides (both sides should see two Celerras) & symdg list:

|           |         |     |           |                                     |
|-----------|---------|-----|-----------|-------------------------------------|
| 1R1_1     | RDF1    | Yes | <symm_ID> | →Represents R1 volumes on Remote CS |
| 1REG      | Regular |     |           | →Local non-SRDF volumes             |
| 1CS_500   | RDF2    |     |           | →Source CS R2 volumes               |
| 1R2_500_1 | RDF2    |     |           | →R2 data volumes on Remote CS       |

**Note:** Remote side nasadmin account should not be able to see or manage the RDF standby servers

6. List SRDF-linked Celerras from Remote side

### **\$ nas\_cel -list**

7. List Source Celerra file systems from Remote side:

### **\$ cel\_fs Celerra\_local -list**

8. List & Identify Remote file system in order to Import Production File System from Source:

### **\$ nas\_fs -list**

9. Import PFS from Source to Remote Celerra & Verify:

**Note:** Imports to R2STD devices on Remote Celerra

### **\$ cel\_fs Celerra\_local -Import fs1**

### **\$ nas\_fs -list**

### **\$ nas\_fs -info fs1**

10. Create Snapshot of imported file system onto BCVs:

### **\$ fs\_timefinder fs1 -Snapshot**

11. Refresh the newly created BCV Snapshot:

### **\$ fs\_timefinder fs1\_snap1 -Mirror refresh**

12. Mount and export snap file system on Remote Celerra

## **RESTORING NEARCOPY SNAP FILE SYSTEM TO R2STD & THEN PFS:**

Unexport and unmount snap file system

### **\$ /nas/sbin/rootfs\_timefinder fs1\_snap1 -Restore**

**Note:** SRDF R1 & R2 device pairs are automatically suspended after conducting a Restore. Verify using the following commands:

### **\$ /nas/symcli/bin/symdg list**

|           |      |     |          |    |   |    |   |
|-----------|------|-----|----------|----|---|----|---|
| 1R2_500_1 | RDF2 | Yes | <sym_ID> | 55 | 0 | 64 | 0 |
|-----------|------|-----|----------|----|---|----|---|

**\$ /nas/symcli/bin/symrdf -g 1R2\_500\_1 query** (look for suspended devices)

Unmount PFS on Source

**\$ /nas/sbin/rootfs\_rdf fs1 –Restore**

**Note:** Restores file system from R2STD

## **OPERATIONS PERFORMED FOR SNAP OR REFRESH WITH ‘NEARCOPY’:**

- Establishes BCV against R2 volumes on R-2 side
- Monitors operation until all BCV’s are in synchronized state
- Freezes the R-1 Source filesystem and ceases I/O to the file system and temporarily unmounts fs
- Proceeds to split the BCV volumes on R-2 side
- Thaws the R-1 Source file system, remounts, and resumes I/O

### **SYMMETRIX Microcode Requirements:**

5265 supports only the last STD-BCV pair for establishes and restores

5266 supports the last (8) STD-BCV pairs for establishes and restores

### **SYMMETRIX BCV’s:**

Located in Drive Map on Disk Adaptor Tab: BCV’s show up in teal coloring

**INLINES:** A7,D [will show BCV volumes under Type column with blue “B”]

A7,DC [shows standard volumes with BCV’s established against it]

8D,,,’GDAT’ [History of BCV’s—Splits show up as “I” in the ‘DC ACT’ column]

E7, [displays devices for a specific director]

**After snapshot:** BCV volume shows up in red as “NR BCV” and a corresponding green “B” shows up for the standard device. BCV then shows up as the M3 device for the Standard

### **TimeFinder Logs/Troubleshooting:**

Verify TimeFinder scripts on Celerra: \$crontab -l [/usr/local/BCVscript: BCVerrorlog/BCVlist/BCVlog/BCVmaint

Logged Information: /nas/log/symapi.log /nas/bin/server\_log /nas/log/cmd\_log.err /nas/log/cmd\_log

### **Example of BCV Error in Cmd log.err:**

5-22-2001 13:50:54 db:201:nas\_fs -S c1dm1fs50 -n c1dm1fs50\_snap2: [“Volume not available”—no more BCV volumes available to do a snap!!]

### **Examples of Symapi.Log Activity:**

5/22/2001 22:00 Celerra CS STARTING a BCV ‘INCREMENTAL\_ESTABLISH’ operation for 1 [STD-BCV] pair:[10f-3C2]

22:00:39.449 The BCV ‘INCREMENTAL\_ESTABLISH’ operation SUCCEEDED.

22:46:11.205 STARTING a BCV ‘SPLIT’ operation for 8 [STD-BCV] Pairs [123-2B5, etc.]

### **TimeFinder F/S and Symm Meta Volumes are operational in the EMCNAS versions below.**

2.1.12; 1.2.50; 2.2.15.4 + [i.e., cannot use SYMM metavolumes in NAS versions lower than these]

## **USING CELERRA TIMEFINDER/FS FOR SNAPSHOTTS, MIRRORING, & CLONING FS:**

**I. Introduction:** Symmetrix Logical volumes [d3, d4, etc.] are typically visible by all DataMovers!

Only (1) FileSystem should be on a STD or BCV

If a FileSystem spans (3) symm volumes, then the BCV also needs to span (3) symm volumes

Symm 5266 code can do “InstantSplits”, which occur in background

**II. Create Original File System:** #time nas\_fs -n ufs8 -c mtv8

**III. Create BCV Snapshot File System:** #time nas\_fs -S ufs8 [Uses first series of rootfs devices in list=ufs8\_snap1]

BCV Types: 1=slice 2=volume 3=meta 4=bcv

**Note:** After all BCV volumes have been used once for snaps, you have to delete the old ones before you can conduct more snaps!

**IV. Conducting Mirror:** Permanently unmount the Snap FS to be mirrored!!

\$server\_umount server\_2 -p ufs8\_snap1 \$time nas\_fs -M on ufs8\_snap1 [create mirror—Type=mirrorfs]

**V. Mount Snapshot FS ReadWrite:** \$server\_mount server\_2 -F -o rw,nolock,accesspolicy=NATIVE ufs8\_snap1 /ufs8\_snap1

**VI. Unmount Snapshot File System and Delete:** \$server\_umount server\_2 -p ufs8\_snap1 \$nas\_fs -d ufs8\_snap1

**VII. Cloning BCV:** \$/nas/sbin/rootnas volume -C mtv8 disktype=bcv

**VIII. Restoring a Snapshot File System:** Unmount STD volume first, then BCV, then restore

#server\_umount server\_2 -p ufs8 #server\_umount server\_2 -p ufs8\_snap1 #/nas/sbin/rootnas\_fs -R ufs8\_snap1

**CREATING SNAPSHOT FILESYSTEM:** #nas\_fs -S fs2 [STD volumes are either mirrored or SnapShot onto BCV volumes]

**Note:** Tracks on STD to BCV are mirrored until 2 seconds of synchronization remain, then fs is frozen [unmounted], and client access is broken. BCV then splits off and STD File System unthaws, or is remounted and made available. Might need to create a BCV Volume first in order to save a Snapshot to.

Step 1: \$nas\_volume -n bcv1 -c -M roottd15,roottd16,roottd17,roottd18

Step 2: \$nas\_fs -S globix1 -v bcv1

### **EXAMPLE OF SNAPSHOT FILE SYSTEM IN PROGRESS...:**

```
# nas_fs -i fs-1_snap1
id      = 67
name    = fs-1_snap1
acl     = 0
in_use  = False
type   = mirrorfs [mirrorfs designation indicates that Mirroring is On]
volume  = v431
rw_servers=
ro_servers=
backup_of = fs-1 Tue Aug 27 22:03:37 EDT 2002
remainder = 9091 MB (3%)
symm_devs = 000185700438-29B,000185700438-29F,000185700438-2A3,000185700438-2A7,
000185700438-2AB,000185700438-2AF,000185700438-49B
disks   = roottd3,roottd4,roottd5,roottd6,roottd7,roottd8,roottd131
```

### **MIRROR COMPLETED:**

```
# nas_fs -i fs-1_snap1
id      = 67
name    = fs-1_snap1
acl     = 0
in_use  = False
type   = mirrorfs
volume  = v431
rw_servers=
ro_servers=
backup_of = fs-1 Tue Aug 27 22:03:37 EDT 2002
remainder = 0 MB (0%)
symm_devs = 000185700438-29B,000185700438-29F,000185700438-2A3,000185700438-2A7,
000185700438-2AB,000185700438-2AF,000185700438-49B
disks   = roottd3,roottd4,roottd5,roottd6,roottd7,roottd8,roottd131
```

**Note:** Next step would be to break Mirror & Split off the BCV Volume—then reverts to Type=uxfs [#nas\_fs fs-1\_snap1 -M off]

**Mirroring:** To “mirror” an existing BCV, the BCV must be unmounted. Then issue the **nas\_fs -M on** command.

**Mounting Snapshot FS:** Snapshots are made onto BCV Metavolumes [BCV’s are created manually using “rootnas\_volume” command]. ‘Mount’ a BCV volume READ-ONLY, but mounting “R-W” destroys the BCV-STD relationship.

**#server\_mount server\_2 -perm -o ro snap\_fs2 /mntpoint** FS permissions and inode numbers are maintained in the BCV copy.

### **FORCING SNAPSHOT FILESYSTEM TO MOUNT “RW”:**

**\$server\_mount server\_2 -F -o rw fs8\_snap1 /fs8\_snap1**

**Multiple Snapshots:** Code 5265 limited to only last incremental BCV for “Restores”  
Code 5266 can “Restore” up to 8 previous incremental “Snapshots”.

**Restoring a Snapshot File System:** #rootnas\_fs -R ufs1\_snap1

**BCV to STD Volume:** [BCV filesystem volume to be restored to STD]

STD volumes are mirrored on the symmetrix as BCV volumes and can be used for Testing or Backup purposes.

Timefinder is oriented towards File Systems: #nas\_disk -l [will show root\_d3, root\_d4 for BCV volumes]

**Deleting a Snapshot File System:** \$nas\_fs -d snap\_fs2

**#nas\_fs -i ufs1** [details about a FileSystem]

**Snapshot Fails:** Snapshots may fail if new drives or DataMovers are added [SYMAPI\_C\_DEVICE/GATE error]

1. Run \$/nas/symcli/bin/symcfg discover [Updates the Symmetrix device database]

\$/nas/bin/server\_devconfig server\_2 -p -s -a [Lists devices found]

\$/nas/bin/server\_devconfig server\_2 -c -s -a [Adds any new devices found]

Reboot DataMover and repeat for each Datamover affected

**Note:** If this doesn’t work, may need to failover & retry

### **TIMEFINDER/FS FEATURES**

**vs.**

### **SNAPSURE CHECKPOINT**

- |                                        |                                    |
|----------------------------------------|------------------------------------|
| --Full Mirrors                         | --Increments                       |
| --Made on BCV's                        | --Made on STD's                    |
| --Detached copies can be mounted RW    | --Export SavVol to DataMover as RO |
| --Full Restores ONLY                   | --File Level Restores              |
| --System CPU handles the load          | --DataMover CPU handles load       |
| --Grouping of FileSystems with NAS 4.0 |                                    |

**SNAPSURE:** Introduced with 2.2+ Code

Does not use BCV's as in TimeFinder.

## **CELERRA METAVOLUMES—CLONING VOLUMES:**

--Metas are only used for Open System Hosts

--Metas can contain up to 255 devices

--Volumes Map in SYMMWIN shows Meta Group, Meta Position, and Meta Stripe [no=not striped; 2=striped]

Q—Why do we require a MetaVolume for FileSystems in Celerra? A—allows for On-Line extension of FileSystem; and, allows large grouping of physical disks into a single Logical Volume. Also, uses less SCSI addresses.

A metavolume is arranged in addressable logical blocks, from Logical Block Address 0-N

### **CELERRA VOLUME MANAGEMENT:**

Managing Disks and/or Volumes attached to Celerra [Slice; Stripe; Meta; BCVs]

### **CELERRA DISK TYPES:**

BCV=Business Continuance Volumes for Timefinder/FS operations

STD=Standard Volume used for ordinary file systems

R1STD=Standard volumes that have been mirrored

R1BCV=BCV volumes that have been mirrored

### **CELERRA VOLUME CLONING:**

**\$nas\_volume -C volume\_name -o disktype=svol:dvol [STD|BCV]      Code 2.0.28.3+/4.0.x**

**\$nas\_volume -Clone d2 disktype=BCV root\_d2:disk15**

**\$nas\_volume -Clone stv1 d3:d8,d4:d9,d5:d10,d6:d11**

Nas\_volume -Clone makes an exact copy of a volume of either a BCV or STD volume

svol:dvol →represents Source and Destination volumes, respectively

Can also restore a file system copy to the original location.

## **CELERRA DISASTER RECOVERY SOLUTIONS vs. REMOTE BACKUP SOLUTIONS:**

### **DISASTER RECOVERY:**

Celerra only supports disaster recovery with synchronous SRDF where there is 100% synchronization of data between production and remote sites. If “point in time” captures of data meets a customer’s criteria for “disaster recovery”, then TimeFinder FarCopy can also be used as a DR solution.

### **REMOTE BACKUP SOLUTIONS [Non-DR solutions]:**

TimeFinder NearCopy, FarCopy, Celerra Replicator

### **CELERRA SRDF:**

#### **MAINTAINING HEALTHY SRDF ENVIRONMENT & CONDUCTING FAILOVER TESTING:**

See primus emc131411 for information on maintaining and testing SRDF environments. Check the link to download the latest chk\_srdf\_groups.sh script, and run on Source & Destination Celerras prior to running failover. The script performs two basic and fundamental integrity tests:

I. Checks to ensure that all RDF devices match between the Source and Destination sides using /nas/symcli/bin/symdg. Source side RDF1 devices should add up to same number of RDF2 devices on the target side in the basic device groups. Target RDF1 devices should add up to the same number of RDF2 devices on the Source side. If they do not, one suggested fix would be to run #nas\_rdf –init on Source and Destination.

II. Verifies that all devices in the RDF groups are “synchronized”, without which a failover will likely not succeed.

**Symm Minimum Code:** 5567.24.13.S

Symmetrix Remote Data Facility supports disaster recovery when used in synchronous mode. Celerra SRDF writes data to local and remote volumes synchronously. Local & Remote Storage systems must be within 60Km or 37 miles for synchronous operation.

**Modes:** *Active-Active or Active-Passive* configurations using synchronous mode campus—37.5 miles limit!

--Synchronous Mode only for Disaster Recovery situations

--Semi-synch or asynchronous mode for 'data' volumes only using R2 Copy

--Cannot have dual Control Stations in use for Celerra SRDF failover, restore [Except with new SW Patch—contact TS2]

**Note:** Dual Control Station support added with NAS 4.2.5.1 and higher

Provides for "remote mirror copy of data" between Celerras at different locations using SRDF Communications via a "Remote Adapte/RA" [R-1 → R-2] using a T3 or Escon WAN link. Celerra "SRDF" requires Symm 4.0/4.8 minimum running 5265 microcode. Only 'supported' when use Celerra CommandLine "NAS-to-SYMM" API. The top RA port [top port] is used for RA1--RA2 comms. Celerra SRDF supports only "journal 0" between symmetrixes. Logical volumes on primary R1 side are mirrored on the R2 side. RLD –Remote Link Director—RA1 and RA2, provide SRDF link between R-1, R-2 sides, respectively.

**Note:** Normally have 240 ids per SCSI port; SRDF 50-50 split reduces this to an effective total of 120SCSI ids.

## **PURPOSE OF CELERRA SRDF:**

--Provides for a Disaster Recovery solution [Active-Active or Active-Passive]

--Provides for remote backup of data [Farcopy & Nearcopy Timefinder/FS implementations]

--Provides for data center migration ability & use of Symmetrix Data Migration Services [Manual process, heavy planning]

--Provides for recovery from planned outages

## **HOW CELERRA SRDF WORKS:**

→NAS commands invoke SYMCLI commands on Control Station to accomplish tasks (/nas/sbin/nas\_rdf)

→Only single Source SYMM supported for Celerra SRDF

→Failover and Restores are invoked manually

→Synchronous modes keeps R1 & R2 devices in sync [Target notifies Source Host of 'write completion']

→Active-Active & Active-Passive can use multiple RLD's (Remote Link Directors)

→Only supports primary Control Stations

→BCV's on local R1 [and on remote R2 devices with RPQ]

→SCSI or Fibre Datamovers over SRDF Fibre Channel or ESCON

→Gatekeeper devices are not part of the SRDF devices

→Semi-Synchronous Mode not supported for "disaster recovery" [journal 1 mode---first write request acknowledged]

→"Farpoint" supports one outstanding I/O from different volumes—also not for 'disaster recovery' operations

## **SETTING DEBUG WHEN TROUBLESHOOTING SRDF:**

**#export SYMAPI\_DEBUG=-1**

**#export SYMAPI\_DEBUG\_FILENAME=/home/nasadmin/symapi\_debug.log**

Exit the shell to revert back from debug logging

**Note:** Don't forget to export symapi\_debug to a file. In certain situations, with debug turned on, symcfg or symcli commands may fail and output debug info to terminal.

## **SRDF CONFIGURATIONS:**

### **ACTIVE-PASSIVE SRDF (Synchronous Mode):**

(1) Production Celerra & Storage System and (1) Backup Celerra & Storage System

**Note:** Active side will have (6) R1 and (1) R2 device for Control Volumes

--DR Solution

--Campus 40 miles ESCON/Fibre

--Supplemented with NearCopy for Backups

### **ACTIVE-ACTIVE SRDF (Synchronous Mode):**

(2) Production Celerra & Storage Systems, each Backing up the Other [Each Symm has R1 & R2 Volumes]

**Note:** Each side should have (6) R1 devices and (4) R2 devices for Control Volumes

--DR Solution

--Campus 40 miles ESCON/Fibre

--Supplemented with NearCopy for Backups

### **NEARCOPY (Synchronous Mode):**

--Campus distance 40 miles

--Not a DR solution but can be supplemented with SRDF AA or AP

--Backup Solution RO

### **FARCOPY (Not Synchronous—Adaptive Copy WP Mode):**

--Adaptive Copy, Write-pending

**Note:** DOS partition, LUN 00, should be setup in synchronous mode vs. Adaptive mode

--No Geographic limitation for distance

--Not a DR solution

--Backup solution RO

### **FARCOPY ++ (Not Synchronous):**

--Adaptive Copy

--No Geographic limitation for distance

--DR Solution

--Backup Solution RO

### **SRDF/A (SNOW—Symmetrix Native Ordered Writes)—Asynchronous Mode:** NAS 5.4 and Microcode 5670 +

→ Asynchronous mode (Uni-directional) SRDF that is not a DR solution (i.e., data will be lost)

→ Provides for consistent point-in-time copy to Target side (R2) that is only slightly behind R1 side

→ Requires DMX hardware with 5670 code

→ SRDF/A defined at RDF Group level & all SRDF/A volumes will be in async mode

→ SNOW supports all existing SRDF topologies & Hosts

→ Uses (2) COVD (Cache Only Virtual Devices) devices for every SRDF/A volume on the R2 side only

→ SRDF/A over IP requires Qualifier

→ Both Failover & Failback are manually invoked

→ Identical network configurations required, one-to-one failover only

→ Target Celerra can have its own local data movers

**Purpose:** Extended distance solution that does not have performance issues—designed to replace FarCopy

#### **Features:**

--Hosts can read data from R2 side in the event the local source mirror fails

--Supports any host type

--Failover and failback

--Supports other SRDF technologies

#### **Setup:**

--Uses a single SRDF group per Symm pair in the bin file--“Reserve SNOW Resources”

**Note:** All changes are done at the group level and not individual devices.

--Introduces concept of COVD (Cache Only Virtual Devices) on R2 side—(2) COVD's are set for every R2 device

--Writes are sent across link in the correct order (doubly linked lists)

**SRDF/A Cache split into (3) groups/cycles:** Capture, Transmit, Restore

→ N—Cache cycle that captures new writes on the Source R1 side for Host writes

→ N-1—Cache cycle that transmits the dataset chunks of data from R1 to R2 (via the established cycle time)

→ N-2—Cache cycle that takes transmitted dataset chunk from R1 and makes consistent point-in-time dataset on R2 side

**Note:** Cycle Switch occurs once minimum cycle team (30 secs) has expired and when the N-2 has restored all its data to R2 side and the R-1 has completed its transfer. Once this criteria is met, the next cycle time begins.

#### **Three SRDF/A States:**

Both N and N-2 are Active states

N-2 is an Inactive State

Not-Ready State is set if SRDF/A if links down, R2 devices Not Ready, etc.

#### **Dataset Cache Model:**

R1 Active Side contains dataset changes in N cache

R1 Active Side contains dataset transmission changes in N-1 cache, which also spans to R2 Inactive Side as N-1

R2 Inactive Side contains dataset changes that are written to R2 side and are consistent, called N-2

#### **Tolerance Mode:**

Consistency disabled: Tolerance = ON

Consistency enabled: Tolerance = OFF

With Tolerance OFF, SRDF/A will become inactive if SRDF Links are lost or any R2 device becomes Not Ready

With Tolerance ON, SRDF/A will remain active with R2 disk failures, making only single device inconsistent. Tolerates short-term SRDF link interruptions (Link Limbo—10 secs by default), but will go inactive if links remain down. Can replace R2 devices.

#### **Failover:**

--Failover and Failback are manually invoked

--Data Movers are 1:1 from Source to Target side, and both sides can also have local DM's with local File Systems

--R2 RDF Standby Data Mover must have identical network configuration as Source, as well as access to R2 data volumes

--Graceful failover would retain all changes made in N-1 and N-2 sides, but not N, while ungraceful failover would lose contents of N-1 and N.

--place CS0 lun 002 out of SRDF group A from R1 side and into group B for R2 Control Volumes

--Max throttle is set to 60 secs, max time that SNOW will remain active once WP ceiling is reached, provided cache can handle it

#### **SETTING UP SRDF/A:**

1. Run /nas/sbin/nas\_rdf –init on Source, then Target side to setup SRDF/A, creating RDFADMIN account and password, as well as passphrase, when prompted

## **#nas\_cel -list**

2. Set async mode on Target side:      **#/nas/symcli/bin/symrdf -g 1R2\_500\_11 set mode async**
3. Set Consistency Mode Tolerance Off:    **#/nas/symcli/bin/symrdf -g 1R2\_500\_11 enable**
4. Query state of SRDF/A at any time using: **#/nas/symcli/bin/symrdf -g 1R2\_500\_11 query -rdfa**

## **LISTING DEVICE GROUPS:**

**#/nas/symcli/bin/symrdf list**

**#/nas/symcli/bin/symrdf -g 1R1\_1 query** [Querying R1 Device Group]

## **ACTIVATING SRDF/A FAILOVER (Target Site) or RESTORE:**

1. Real disaster scenario, shutdown Source data movers and Control Station (for testing, leave Source side running and let script shutdown)

2. Run Activate on Target Side:

**#/nas/sbin/nas\_rdf -activate**

3. Verify State after Failover:

**#/nas/symcli/bin/symrdf -g 1R2\_500\_1 query**

**Note:** Look for “Failed Over” in output.

**#df -h**

**Note:** Also run df -h on target Control Station and verify that **/net/500/nas** directory is present—this is the /dev/sd1 partition that becomes the active NAS database on the failed over site.

**#nas\_disk -l**

**Note:** All disks should reflect R2STD

**#nas\_server -l**

**Note:** All Servers should reflect acl of 2000 and not be in a faulted state

## **4. SRDF/a Restore to Source Side:**

- a) Making Consistent Snaps before doing Restore:

**# fs\_rdf <filesystem name> -Link suspend**

- b) Perform the import and snap on the target side

- c) Resume SRDF link:

**# fs\_rdf <filesystem name> -Link resume**

- d) Do not power-up Source CS or DMs

- e) **#/nas/sbin/nas\_rdf -restore** [on target side]

**Note:** Debug log kept in /tmp/nas\_rdf.debug.log on Target side. RDF links set to Synchronous mode for the Restore operation.

5. Query R2 Device Group after Restore:

**#/nas/symcli/bin/symrdf -g 1R2\_500\_1 query -rdfa**

6. Verify Disks after Failback:

**#nas\_disk -l**

**Note:** If devices show up as STD, run nas\_diskmark -m -a to make them revert to R1STD

## **SRDF/A BIN FILES:**

→RDF mode SRDF, Reserve SNOW Resources to Yes

→CS0 must not be in SNOW Volume group on each side, R1 & R2, ALL disks displayed on SOURCE side as R1STD, but only (6) Control Volumes displayed on Target Side Symmetrix, again as R1STD devices

→Volume map for R2 is normal SRDF Active-Passive Volume map

→(2) COVDs created for every device in SNOW group

## **SNOW INLINES:**

Session: 8D,,,SNOW

Status of Session: 8D,00,,SNOW

Verbose Session: 8D,00,192,SNOW

SNOW Help: F0,CE,SNOW,HELP

A7,RASN →Check SRDF links

E0, →Displays cable connections

## **Activating SNOW via InLines—always start with Target Side first:**

FO,CE,SNOW,ACTV,0 (Target)

FO,CE,SNOW,ACTV,0 (Source)

## **VERIFYING TIMEFINDER SCRIPTS:**

**/var/spool/cron**

**# cat rdfadmin**

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!

```
# DO NOT EDIT THIS FILE - edit the master and reinstall.  
# (rfadmin installed on Thu Jun 3 19:24:51 2004)  
# (Cron version -- $Id: crontab.c,v 2.13 1994/01/17 03:20:37 vixie Exp $)  
# DO NOT EDIT THIS FILE - edit the master and reinstall.  
# (rfadmin installed on Thu Jun 3 19:24:12 2004)  
# (Cron version -- $Id: crontab.c,v 2.13 1994/01/17 03:20:37 vixie Exp $)  
# DO NOT EDIT THIS FILE - edit the master and reinstall.  
# (/tmp/crontab.21337 installed on Thu Jun 3 17:22:43 2004)  
# (Cron version -- $Id: crontab.c,v 2.13 1994/01/17 03:20:37 vixie Exp $)  
25 19 * * 1-5 /home/nasadmin/snapfiles/snap.main5 >/home/nasadmin/snapfiles/snap.log 2>&1
```

## **SRDF TIMEFINDER/FS FARCOPY & NEARCOPY SUPPORT:** NAS 2.2.49.2/4.0.12.1 [Uses diff. commands]

Nearcopy uses Synchronous Campus Mode to create RO copy of FS that can then be exported on Remote File server. Farcopy can use Semi-Synchronous Mode for CS Volumes and Adaptive Copy modes for data devices except for Fibre target & luns of 00, 06, and SCSI target of 10 (Also known as R2 Copy). FarCopy makes a copy of PFS on R1 side to a local R1 BCV. The R1 BCV is then copied to the R2 BCV via SRDF. The R2 BCV is then copied to a STD volume that can then be exported RO.

## **CELLERRA SRDF CONFIGURATIONS:**

- I. SRDF ACTIVE--ACTIVE: Synchronous; Campus to 60km; Escon/Fibre; Disaster Recovery; Use with Near Copy
- II. SRDF ACTIVE--PASSIVE: Synchronous; Campus to 60km; Escon/Fibre; Disaster Recovery; Use with Near Copy
- III. SRDF NEARCOPY: Synchronous; Campus to 60km; Escon/Fibre; No Disaster Recovery; RO Backups
- IV. SRDF FARCOPY: Asynchronous; Adaptive; No limits; No Disaster Recovery; RO Backups
- V. SRDF FARCOPY RPQ: Asynchronous; Adaptive; No limits; Disaster Recovery; RO Backups

## **CELLERRA SRDF PREREQUISITES:** Symm 4.0+, 5265.48.30 and NAS 2.1.15.16

## **SRDF DISASTER RECOVERY ACTIVE-ACTIVE & ACTIVE-PASSIVE SUPPORT:** NAS 2.1.15.6

- Solutions apply to entire Celerra-to-Celerra Failover--no individual DataMover failover supported
- R1 & R2 volumes are synchronized in 'real-time'
- Only Synchronous Mode supported--means that every write request on R1 side must be acknowledged on R2 side
- For performance reasons, use of FC or Escon Link Repeaters is not recommended
- Both Modes require BI-DIRECTIONAL SRDF on Symmetrixes
- R1 devices must have R2 devices of equal size, emulation type
- Filesystems built on RDF Datamovers must be built on RDF volumes
- Network topologies between the two Celerras must be the same for 'failover' [try to use same Subnets of each side]
- R1 CS0 must be able to communicate over network to R2 CS0
- Failover from SCSI-to-Fibre & vice versa supported
- DataMover hardware can be different
- If using External Usrmapper, consider the impact of 'failover' operations
- Must maintain 1:1 relationship for R1 DataMovers and R2 Standby DataMovers
- Dual Control Stations on each Side not supported--cannot failover using dual control stations, will cause NAS DB corruption

## **STEP I. SRDF INITIALIZATION:**

**#/nas/sbin/nas\_rdf -init** [Initializes SRDF between two Celerra-Symmetrix pairs; configures Control Station to use SRDF; Identifies the R2 and R1 Celerras and the R1 to R2 volume mappings; Prompts for creation of remote admin account for R2 side, "rfadmin"; Prompts for definition of Standby DataMovers to be used for the R2 side]

## **STEP II. SRDF ACTIVATION:**

Conducting failover from Primary to Backup Celerra:

- a) Halt the Primary Celerra
- b) Log into the Backup or R2 Celerra using the RFADMIN account

**#/nas/sbin/nas\_rdf -activate** [Places R1 volumes in RO mode, and R2 in RW mode]

- c) Standby DataMover assumes IP & MAC Address of Primary, as well as all File Systems, Exports, & Shares

## **STEP III. SRDF RESTORE:**

Log into the R2 'Backup' Celerra and issue **#/nas/sbin/nas\_rdf -restore**

Data on R2's are copied to R1 and then mirrored; Primary DataMovers reboot and then assume IP address & file systems from R2

**Note:** NAS 5.5.30.x resolves an issue whereby restores would fail unless the target CS was rebooted first

## **RUNNING NAS RDF -INIT ON R2 SIDE OF ACTIVE PASSIVE SRDF FAILS:**

**# /nas/sbin/nas\_rdf -init**

Discover local storage devices ...

done

Contact cnt1 ... is alive

Please enter the passphrase for RDF site :

Passphrase:

Retype passphrase:

CRITICAL FAULT:

Unable to mount /dev/sdj1 at /nas/rdf/500

**Resolution:**

1. Rename R2 /nas/site/cshosts file
2. Run # /usr/sbin/userdel -r rdfadmin [This will delete all references to the original RDFADMIN account on R2 Control Station]
3. Run #/nas/sbin/nas\_rdf -init [This prompts for new RDFADMIN account and password, Passphrase password, and re-establishment of relationships between the R2 STANDBY data movers and the R1 production data movers. Successful completion of the -init will also create a new CSHOSTS file on the R2 Control Station].

**OTHER TROUBLESHOOTING TIPS FOR –INIT FAILURES:**

- Ensure that /.etc/hosts file is up-to-date.
- Verify /nas/site/cshosts or /nas/sys/cshosts file information--possible that this file may need to be renamed to allow a fresh –init. ----
- The -init could fail if the HTTP daemon is not running on R1 and R2 Control Stations, as the nas\_cel command needs to establish a secure connection between the two Control Stations in order to setup RDF communication.
- Passphrase also needs to be the same between the two sides.

**MAIN DIFFERENCE BETWEEN ACTIVE-PASSIVE AND ACTIVE-ACTIVE SRDF:**

Active-Active has R1 and R2's on each side that are active at all times while Active-Passive R2 side cannot be used for production as its dataMovers are dedicated only as Standbys

**HOW CAN I DETERMINE IF A CONFIGURATION IS ACTIVE-ACTIVE OR ACTIVE-PASSIVE?**

1. BIN Files
2. Or, on each Celerra, look at /nas/site/cshosts file--it should only contain an entry for R1 CS on the R2 side

**ACTIVE-PASSIVE & ACTIVE-ACTIVE SRDF:**

**I. CELERRA ACTIVE-PASSIVE SRDF SUPPORT(R-1 to R-2 Celerras):**

**CONFIGURATION REQUIREMENTS:**

- Must have (2) Symms and (2) Celerras
- Zone Control Station only to single FA zone and single Symm
- Either Active-Passive or Active-Active configurations require bi-directional RDF to operate
- Active side contains only R-1 RDF volumes
- Target Passive side will contain only R-2 RDF volumes
- Failover is for the total box, not individual Servers
- After failover is initiated, any local DM's will not be operational
- R-1 DataMovers need to be configured to failover to R-2 Standby DataMovers when RDF failover is invoked. Prior to failing back, all DataMovers/CS0 on R-1 side must be shut down before initiating failback command.
- SRDF Standby Target DM's must have same Ethernet & network configuration as primary Source DM's [GbE, FE, etc]
- Control Stations must have a network path to each other
- SRDF Standby must see all data volumes outlined in R2 Symmetrix BIN file
- Primary Datamover is paired to only one SRDF Standby in a 1 : 1 relationship

**Note:** Make sure to pair the Primary & Standby RDF Servers with same slot number

**REQUIRED BIN CONFIGURATION ACTIVE-PASSIVE:**

- R2 side must be able to access Control Luns of R1 side
- R1 side must be able to access boot lun 00 on the target side (/nas/dos)
- Source Bin file sees devices as RA2/RF LUNs 06, 07, 08, 09, 0A, 0B for Target side
- Target Bin file sees Source Luns as RA2/FS Luns 00, 01, 02, 03, 04, 05
- Source R-1 data volumes should be mapped to same device numbering scheme on Target, but will be R-2 volumes

**Flag Settings on R2 Devices on Symms:**

RDF R2 Not Ready if Invalid=no

RDF R2 Not Ready=no

**/nas/site/cshosts** [This file should be empty on R-1 CS0; On R-2 CS0 should have entry for R-1 IP Address--created by nas\_rdf -init]

**CONTROL STATION CONTROL VOLUMES:**

**FROM PRIMARY R1 BIN FILE:**

R1 CS0 & R1 Datamovers must always have Target & LUNs of 00, 01, 02, 03, 04, 05

R2 CS0 must always have Target & LUNs of 10, 11, 12, 13 if SCSI and 6, 7, 8, 9 if FIBRE

**Note:** R2 CS1 never gets mapped at all!!

**SETTING UP ACTIVE-PASSIVE (R-1/R-2) SRDF:**

- Step 1. Initializing SRDF on SOURCE: Login to R-1 Celerra as nasadmin, su to root and run **#/nas/sbin/nas\_rdf -init**
- Step 2. At prompts, associate R-1 datamovers with standby R-2 datamovers by slot # [recommend using slot-for-slot matching]
- Step 3. Initializing SRDF on TARGET: Login to Remote R-2 Celerra as nasadmin, su root and run **#/nas/sbin/nas\_rdf -init**  
**Note:** Local R-2 Storage is discovered and R-1 Control Station contacted
- Step 4. TARGET CS: When prompted, enter the "RDFADMIN" Username & Password to create this special RDF account  
**Note:** At this point, 'discovery' of remote R-1 devices occurs
- Step 5. TARGET CS: Assign Standby Servers to Primary Servers using "ID" numbers  
**Note:** R-1 and R-2 DataMovers must have a 1 : 1 relationship
- Step 6. TARGET CS: Verify creation of R-2 Standby Datamovers: **#nas\_server -l** [Logged in as RDFADMIN]

## **WHAT DOES THE "nas rdf -init" COMMAND DO?**

- Run first on R1 side, then R2 side
- Establishes the relationship between R1 & R2 devices & RDF Standbys
- Discovers local dos, root, nas devices (/dev/sda1; /dev/sdc3; /dev/sde1), using **nas\_disk -l**
- Creates cshosts file on R2 side
- Discovers remote dos, root, nas devices (/dev/sdg1; /dev/sdi3; /dev/sdj1)
- Mounts R2 remote root device & reads /etc/sysconfig/network & /etc/hosts files for remote CS0 hostname & IP address
- Pings R2 CS0 to verify network connectivity
- Reads /nas/sys/callhome.config to obtain remote site name
- Determines if CS can be a failover target
- Uses server\_devconfig server\_2 -p -s -a to locate diskmarked devices with "value=2 (local root\_ldisk, remote root\_ldisk)
- Creates RDFADMIN User account when run on R2 side to manage R2 Standby Datamovers (Can only 'see' these datamovers using this User account)
- Sets up communication path & devices for CS using nas\_rdf chand process (root\_rdf\_volume; nas\_cel -i)
- Adds Symm DEV002 to special group "cs\_500" (/dev/sdc; /dev/sdi) so that these devices will not be included in the failover & fallback operation
- Creates the "1R2\_500\_1" RDF Group of local R2's & remote R1's. Sets up Datamover relationships and ACL of 2000 for RDF Standby Datamovers

## **TROUBLESHOOTING TIP:**

If any Celerra devices have been added or reassigned after the initial setup of RDF, re-run **#/nas/sbin/nas\_rdf -init** on R-1 & then R-2

## **ACTIVE-PASSIVE FAILOVER—SRDF ACTIVATE:**

- Step 1. **#/nas/sbin/nas\_rdf -activate** (Target Side)

**Note:** Login as rdfadmin account, su to root, do failover [su - rdfadmin ]. Activate makes source SA's write-disabled; SRDF Link Suspends; Target SA's write-enabled; NAS DB fsck'ed; SRDF DM's Activated

- Step 2. Verify communications between R1 & R2 Control Stations [telnet, ping, etc]

- Step 4. Ensure that R-1 /nas/site/slot\_param file is copied over & is the same on the R-2 Control Station

- Step 5. Verify the status of RDF Groups: **\$/nas/symcli/bin/symdg list**

### **R-1 RDF DEVICE GROUPS:**

| Name         | Type        | Valid      | Symmetrix ID        | Devices    | GK's     | BCV's    |                                                |
|--------------|-------------|------------|---------------------|------------|----------|----------|------------------------------------------------|
| <b>1R1_1</b> | <b>RDF1</b> | <b>Yes</b> | <b>000285500353</b> | <b>198</b> | <b>0</b> | <b>0</b> | [Only 1 RDF Group on R1 Side--R1 Data Devices] |

### **R2 RDF DEVICE GROUPS:**

**\$/nas/symcli/bin/symdg list**

| Name             | Type        | Valid      | Symmetrix ID        | Devices    | GK's     | BCV's    |                                       |
|------------------|-------------|------------|---------------------|------------|----------|----------|---------------------------------------|
| <b>1R1_1</b>     | <b>RDF1</b> | <b>Yes</b> | <b>000285500186</b> | <b>6</b>   | <b>0</b> | <b>0</b> | [R1 Control Volumes]                  |
| <b>cs_500</b>    | <b>RDF2</b> | <b>Yes</b> | <b>000285500186</b> | <b>1</b>   | <b>0</b> | <b>0</b> | [R1 Control Volume to be failed over] |
| <b>1R2_500_1</b> | <b>RDF2</b> | <b>Yes</b> | <b>000285500186</b> | <b>197</b> | <b>0</b> | <b>0</b> | [R1 Data Volumes to be failed over]   |

- Step 6. Verify the Synchronization/Tracks Invalid Status of RDF Groups on R-2 Celerra:

**\$/nas/symcli/bin/symrdf -g 1R2\_500\_1 query** [No Invalid Tracks & All Devices Synchronized]

Target (R2) View      Source (R1) View      M O D E S

| ST<br>Standard<br>Logical<br>Device | A<br>T<br>E | LI<br>R1 Inv<br>Tracks | N<br>R2 Inv<br>Tracks | A<br>K<br>S | T<br>Dev<br>E | M                |                  |                        | RDF Pair<br>STATE |
|-------------------------------------|-------------|------------------------|-----------------------|-------------|---------------|------------------|------------------|------------------------|-------------------|
|                                     |             |                        |                       |             |               | R1 Inv<br>Tracks | R2 Inv<br>Tracks | o<br>d<br>e<br>Dom ACp |                   |
| DEV001                              | 000         | WD                     | 0                     | 0           | RW            | 000              | RW               | 0 0 SYN DIS OFF        | Synchronized      |
| DEV002                              | 001         | WD                     | 0                     | 0           | RW            | 001              | RW               | 0 0 SYN DIS OFF        | Synchronized      |
| <hr/> -----output abridged-----     |             |                        |                       |             |               |                  |                  |                        |                   |
| DEV197                              | 005         | WD                     | 0                     | 0           | RW            | 005              | RW               | 0 0 SYN DIS OFF        | Synchronized      |

```
Total ----- -----
Track(s) 0 0 0 0
MB(s) 0.0 0.0 0.0 0.0
```

Step 7. Verify Status of R-1 & R-2 DataMovers:

#### **R-1 DATAMOVERS:**

```
# nas_server -l
id type acl slot groupID state name
1 1 1000 2 0 server_2
2 1 1000 3 0 server_3
3 4 1000 4 0 server_4
```

# nas\_server -i -a [Output is abridged to show 'RDFstandby' properties only]

RDFstandby= slot=2

RDFstandby= slot=3

RDFstandby= slot=4

#### **R-2 DATAMOVERS:**

```
$ nas_server -l
id type acl slot groupID state name
1 4 2000 2 0 server_2
2 4 2000 3 0 server_3
3 4 2000 4 0 server_4
```

**Note:** Must log in as RDFADMIN account to 'see' RDF Datamovers!

Step 8. Verify accuracy of "cshosts" file on R-2 Control Station:

\$ cat /nas/rdf/cshosts [R2 Side Only--created by the #nas\_rdf -init command]

1:celerracs0-md:500:10.252.24.164:/dev/sdj1:/dev/sdg: [IP Address of R1 Control Station]

**Note:** File should not be populated on R-1 Control Station for Active/Passive--file located /nas/site later codes

Step 9. Permanently unmount any Checkpoint or BCV Snapshot FileSystems on R-1 Side prior to failing over!

Step 10. Conduct Disaster Recovery Failover by 'halting' all R1 DataMovers & CS0: \$server\_cpu ALL -h now #/sbin/init 0

Step 11. Login to R-2 Celerra as "rdfadmin", su root, and initiate failover command: #/nas/sbin/nas\_rdf -activate

### **ACTIVE PASSIVE CONFIGURATION FROM R2 SIDE ON NS SYSTEM:**

#### **# nas\_cel -l**

```
id name owner mount_dev channel net_path CMU
0 ncsp2l 0 192.168.21.13 00028789037301F9 →R2 Side
1 NCSP1L 500 /dev/ndj1 /dev/ndg 172.27.2.13 00028789037102D3 →R1 Side
```

#### **# cat /nas/site/cshosts**

1:NCSP1L:500:172.27.2.13:/dev/ndj1:/dev/ndg:00028789037102D3:

0:ncsp2l:0:192.168.21.13:::00028789037301F9:

#### **# nas\_disk -l**

```
id inuse sizeMB storageID-devID type name servers
1 y 11499 000287890373-01F9 R1STD root_disk 1,2
2 y 11499 000287890373-01FA R1STD root_ldisk 1,2
3 y 2076 000287890373-01FB R1STD d3 1,2
4 y 2076 000287890373-01FC R1STD d4 1,2
5 y 2076 000287890373-01FD R1STD d5 1,2
6 y 2076 000287890373-01FE R1STD d6 1,2
```

#### **# df -h**

| Filesystem       | Size        | Used        | Avail       | Use%       | Mounted on                                                                       |
|------------------|-------------|-------------|-------------|------------|----------------------------------------------------------------------------------|
| /dev/hda3        | 2.0G        | 960M        | 953M        | 51%        | /                                                                                |
| /dev/hda1        | 30M         | 4.1M        | 24M         | 15%        | /boot                                                                            |
| none             | 251M        | 0           | 250M        | 0%         | /dev/shm                                                                         |
| /dev/nde1        | 1.7G        | 479M        | 1.1G        | 29%        | /nbsnas                                                                          |
| /dev/hda5        | 2.0G        | 401M        | 1.4G        | 21%        | /nas                                                                             |
| <b>/dev/ndj1</b> | <b>1.7G</b> | <b>801M</b> | <b>885M</b> | <b>48%</b> | <b>/net/500/nas</b> →When failed over to R2 side, this is where the NASDB exists |
| /dev/nda1        | 133M        | 39M         | 95M         | 29%        | /nas/dos                                                                         |
| /dev/ndf1        | 1.7G        | 78M         | 1.5G        | 5%         | /nas/var                                                                         |

### **WHAT DOES THE "nas\_rdf -activate" COMMAND DO?**

→Runs on R2 side to initiate failover

→Command fails over R1 to R2 devices, the NAS\_DB, and R1 to R2 Standby datamovers

→Depending on Customer activity, failover may take some time to complete--Follow screen prompts

→When successful, the R-1 Volumes on Primary become Read-Only, SRDF link is suspended, and R-2 Volumes become Read-Write  
→RDF Standby DataMovers are then activated--R-2 merges track tables & devices from R-1 & restores SRDF Link  
→R-2 DataMovers assume IP Address, MAC Address, File Systems, and Export Tables of R-1 DataMovers  
→Specifically, this script discovers local dos, root, nas devices (/dev/sda1; /dev/sdc3; /dev/sde1), Reads /nas/rdf/cshosts file & then identifies remote /dev/sdj1 as /nas partition--mounts these as RO to its local /nas/rdf/500 mountpoint  
→Reads remote /etc/sysconfig/network & /etc/hosts files & pings R1 CS. Issues 'symrdf -g 1R2\_500\_1 -noprompt failover -all'  
→Unmounts & remounts "/nas" directory & then mounts /dev/sdg1 "/nas/dos" directory from R1 side  
→Updates Gatekeeper information on local symmetrix, copies NAS\_DB over, then initiates datamover failover using "server\_standby -activate rdf"

## **RESTORING/FAILING BACK ACTIVE-PASSIVE SRDF--R-2 to R-1 SIDE—SRDF RESTORE:**

Step 1. Restore power to R-1 Symmetrix and bring up SRDF link

**Note:** Make sure that no changes have been made on R1 side, such as adding devices, running rdf\_init, etc.

Step 2. Have PSE Lab conduct HealthCheck of R-1 & R-2 Symmetrixes

Step 3. Best chance of successful “restore” is to have the tracks between R2 and R1 devices as close to be synchronized as possible. Run the following command on R2 logged in as rdfadmin & root in a while true loop to keep the RDF devices synchronized:

**#/nas/symcli/bin/symrdf -g 1R2\_500\_1 -noprompt -all -until 500 update**

**Note:** srdf –restore script also calls on the –until 500 update function, but doing this first lessens R2 to R1 Track Table Merge operations.

Step 4. Only when sync’ed should you begin the restore:

**#/nas/sbin/nas\_rdf -restore**

**Note:** Alternatively, specify a higher track number before merging track tables, such as “-restore –until 1500 update”, etc.

### **RESTORE SCRIPT ACTIONS:**

Initiate command, checks for /dev/sda1 (dos), /dev/sdc3 (root), & /dev/sde1 (nas) on remote R1 side, defaults to 500 tracks for sync purposes, reads local /nas/site/cshosts file, reads remote /nas/sys/callhome.config file, prompts user if the remote side is ready for Storage restoration, reads remote /etc/sysconfig/network & /etc/hosts file, pings remote CS0 (to make sure that it is NOT yet running—will shut it down if it replies), loops through #symrdf -g 1R2\_500\_1 -noprompt -all -until 500 update, and then when the tracks reach 500 or lower, will ask if the remote side is ready for Network restoration (Power on the R1 Celerra and allow it to come up. Or, answer “no” to exit script and run “restore” later. Script then fails back Servers with server\_standby -restore rdf, umounts and fsck’s remote /nas partition (/dev/sdj1), verifies that all tracks are updated: #/nas/symcli/bin/symrdf -g 1R2\_500\_1 -noprompt -all -until 500 update , reboots all Data Movers on R1 side (nas\_mcd -e rdfrestore), verifies that all tracks are updated, then fails back devices using: #/nas/symcli/bin/symrdf -g 1R2\_501\_1 -all -noprompt fallback. Script loops to keep devices synchronized, pings remote CS0, runs #nas\_mcd -e <r1cs0>. Script then reboots any panicked servers, verifies all volumes synchronized, suspends RDF link (#symrdf -g 1R1\_1 suspend), runs #server\_standby ALL -restore rdf, resumes RDF link (#symrdf -g 1R1\_1), reboots all DM’s on R1 side again, reboots CS0 on R1 side, NAS\_DB data copied from R1 to R2 side, volumes are synchronized, changed RW on R1, ready on link, and set to WD on R2 side.

### **OTHER THINGS UPDATED:**

R1 /nas/volume/disks updated with R1 serial number, camdisk files on restored Servers updated, local hardware config files updated, R1 data movers activated, R2 side data movers rebooted as RDF Standbys, etc.

Step 5. If the restore operation fails due to network connectivity issues, continue with the following steps on the R-1 Celerra:

a) Login as nasadmin on R-1 Side and execute; **#/nasmcd/sbin/nas\_rdf -restore**

Verifies all devices synchronized, starts restore of R1 CS, then copies back NAS\_DB & verifies that 1R1\_1 RDF group is write-enabled. Completes failback of datamovers & starts Box Monitor on R1.

### **TROUBLESHOOTING:**

→Investigate the /tmp/nas\_rdf.debug.log files to determine where failure or hang is occurring (script failure, this log is NOT there)  
→Output of symcli commands found in /nas/log/symapi.log

## **DISCUSSION: SRDF FAILOVER/FAILBACK IN REAL DISASTER SCENARIO**

1. Initiate failover command on R2 side
2. But because R1 side is down hard, its R1 devices will not be changed from RW to RO.
3. Failing Back: When ready to failback, power up R1 Symmetrix, bring up RDF Links, then write disable the R1 devices (/nas/symcli/bin/symrdf -g 1R2\_500\_1 write\_disable r1 -force). Check for R1 Invalids using \$/nas/symcli/bin/symrdf -g 1R2\_500\_1 query. If invalids exist, then force update using \$/nas/symcli/bin/symrdf -g 1R2\_500\_1 update -force.
4. Once completed, log in to R2 side as "rdfadmin" user and issue #/nas/sbin/nas\_rdf -restore

## **NAS UPGRADES & BEST PRACTICES BEFORE MAKING SRDF LINKS NOT READY:**

**Note:** In certain situations, such as NAS Upgrades, the SRDF links should be taken offline.

1. Synchronize all R1 to R2 devices using Symmetrix Inlines
2. Break SRDF Links

3. Conduct NAS Upgrade
4. Restore SRDF Links

## **CELERRA SRDF BEST PRACTICES:**

- In BIN file, map all RDF devices to all channels
- Use same Target & LUN ID's on both R1 & R2 sides
- Maximize the use of LUNs for performance reasons (more LUNs in use the better)
- Minimize distance between sites
- Add multiple RDF links between Sites for increased performance (3 or more RDF links)
- Use Fibre v. Escon and do not use "link extenders"
- Though depends on Customer environment, use only Single Striped Meta Volume for each Celerra FileSystem (8k -32k stripe)
- Try to have same IP Subnets for both Sites
- Take into account network speed/duplexing settings, location of Usrmapper, WINS, DC servers, Viruschecker, NDMP Servers, etc.
- Use dedicated 1:1 R1 to R2 datamovers only, as well as RDF-only FileSystems, & use same slot numbering scheme
- Always use (1) & only (1) Control Station per Celerra [do not yet support dual control station environments]
- After any & all RDF Volume changes, run #/nas/sbin/nas\_rdf -init on R1, then R2 side

## **REMOVING DATA MOVERS FROM SRDF RELATIONSHIP:**

1. #nas/sbin/nas\_rdf -init [Answer "none" to the respective pairing questions on both R1 & R2 sides]
2. #nas\_server -l & #nas\_server -i -all [Verify ACL & Ownership info of DMs]
3. #nas\_server -acl 1000 <mover> [Restores ACL to correct value for NAS and owner=nasadmin]
4. #nas\_server -info -all [Verify DM type is correct—issue server\_setup to change if necessary]
5. If required removed the rdfadmin user --#**/usr/sbin/userdel -r xxxx**

## **SRDF MIGRATIONS & MULTICAST CAPABILITY:** Multicast SRDF from one Symm to many

Prior to NAS 4.x, SRDF sys calls were issued by SIL [Symm Interconnect Layer], which cannot interpret  
RDF\_MULTICAST\_ENABLE or Concurrent RDF bits]

With NAS 4.x and above, we use Sym API to make sys calls, which can recognize multicast bits

--cannot do Symm SRDF Migration from Symm to NS Clariion systems

## **MAXIMIZING SRDF PERFORMANCE:**

- Present more luns per RDF filesystem & stripe (Rule is 1 I/O per lun when using Synchronous mode)
- Using more than 3 RDF Links for better throughput (Example: Single Fibre link @500 I/O's/sec.=4MB/sec throughput)
- Use more Clients to spread out writing to filesystems (More Clients mean more data streams writing in parallel)

## **TROUBLESHOOTING SRDF FAILOVER/FAILBACK ISSUES:**

- R1 & R2 Control Stations must be able to talk to each other over network [Not critical for failover, but is for failback]
- Rerun #nas\_rdf -init on Source CS if network settings or control station names have changed
- Primary datamover must have only (1) dedicated standby datamover [don't need to be the same hardware]
- IP subnets should be the same but scripts can work around this
- File Systems must be built on RDF devices, not Standard devices
- Failover can occur Fibre to SCSI or vice versa
- If RDF devices have been added, run \$server\_devconfig -p -s -a to update Celerra database, followed by nas\_rdf -init  
(Problem might manifest itself when a particular server panics & will not recover)
- Ensure that R1 & R2 /nas/site/slot\_param files are the same
- Look at Server Logs
- Look for dumps in /nas/var/dump
- Verify hostnames in /nas/server/slot\_x/hostname
- Repopulate the contents of /nas/server/slot\_x/param on R1 side during Restore [does not copy over automatically]
- Examine file system structure [recommendations are for 16 or 32k stripe depths on metavolumes]
- Verify whether BCV or Checkpoint file systems are mounted
- Look for non-failed over/back devices (\$/nas/symcli/bin/symrdf -g 1R2\_500\_1 query)
- Ensure Customer is not creating heavy network traffic during Failover/Failback operations (will take very long time)
- NDMP Tape drives on R1 side may panic on a restore (Fixed in NAS 2.2.39.2)
- For any IP addressing or Hostname changes, re-run #/nas/sbin/nas\_rdf -init on both sides

## **SRDF FAILOVER:**

### **WHAT TO DO IF BOTH SOURCE & TARGET SITES HAVE VOLUMES WRITE-ENABLED:**

**Note:** Situation can occur if Source symmetrix has not ‘write-disabled’ the device group—this is called a “split device group”



**#/nas/symcli/bin/symrdf -g 1R2\_500\_1 query** [From R2 Side]

Device Group (DG) Name : 1R2\_500\_1  
DG's Type : RDF2  
DG's Symmetrix ID : 000183700172  
Target (R2) View      Source (R1) View    MODES

| ST             | LI    | ST       |               |      |       |          |               |                    |
|----------------|-------|----------|---------------|------|-------|----------|---------------|--------------------|
| Standard       | A     | N        | A             |      |       |          |               |                    |
| Logical Device | T Dev | R1 Inv E | R2 Inv Tracks | K S  | T Dev | R1 Inv E | R2 Inv Tracks | RDF MDA Pair STATE |
| DEV001         | 0000  | WD       | 0             | 0 RW | 0002  | RW       | 0             | 0 S.. Synchronized |
| DEV002         | 0001  | WD       | 0             | 0 RW | 0003  | RW       | 0             | 0 S.. Synchronized |
| DEV003         | 000E  | WD       | 0             | 0 RW | 0010  | RW       | 0             | 0 S.. Synchronized |

```
# /nas/symcli/bin/symrdf -g 1R1 1 query [From R1 Side]
```

Device Group (DG) Name : 1R1\_1  
DG's Type : RDF1  
DG's Symmetrix ID : 000184701179  
Source (R1) View Target (R2) View MODES

|                | ST    | LI   | ST         |               |       |       |              |                    |
|----------------|-------|------|------------|---------------|-------|-------|--------------|--------------------|
| Standard       | A     | N    | A          |               |       |       |              |                    |
| Logical Device | T Dev | R1 E | Inv Tracks | K             | T Dev | R1 E  | Inv Tracks   | RDF Pair MDA STATE |
|                |       |      |            |               |       |       |              |                    |
| d1             | 0002  | RW   | 0          | 0 RW 0000 WD  | 0     | 0 S.. | Synchronized |                    |
| d2             | 0003  | RW   | 0          | 0 RW 0001 WD  | 0     | 0 S.. | Synchronized |                    |
| d3             | 0010  | RW   | 0          | 0 RW 000E WD  | 0     | 0 S.. | Synchronized |                    |
| d4             | 0015  | RW   | 0          | 0 RW 0013 WD  | 0     | 0 S.. | Synchronized |                    |
| d5             | 001A  | RW   | 0          | 0 RW 0018 WD  | 0     | 0 S.. | Synchronized |                    |
| d6             | 001F  | RW   | 0          | 0 RW 0001D WD | 0     | 0 S.. | Synchronized |                    |
| d7             | 0024  | RW   | 0          | 0 RW 0022 WD  | 0     | 0 S.. | Synchronized |                    |

## **MISSING CSHOSTS FILE:**

“nas\_rdf –init” command prompts for creation of RDF Logon Account when one existed previously: Ensure that the /nas/site/cshosts is present—an incomplete or missing cshosts file will cause the “nas\_rdf” script to call for creation of a new RDFAdmin account.

## **SRDF NAS UPGRADE ISSUE:**

There is a possible Active—Passive SRDF Issue when upgrading from NAS 4.2 to 5.1.18 in that a nas\_rdf –init needs to be run on the R1 Control Station, then on the R2 Control Station. However, it may error out on the R2. If so, rename /nas/site/cshosts file and run nas\_rdf –init and re-establish the RDFADMIN account, the Data Mover relationships, and the PASSPHRASE account—the –init should then complete successfully.

### Symptoms:

#/nas/sbin/nas\_rdf -init command issued and hangs on R2 side without producing error

#/nas/sbin/nas\_rdf -init command issued and fails on R1 side with following error because it could not ping the R2 Control Station: "Unable to contact node nas2 at 10.0.248.22"

## **CONTROL STATION ISSUES:**

- Loss of RDF Link communication
- R1/R2 Control Stations missing "hostname=" entry in the "/etc/sysconfig/network" file [each CS checks the other's hostname here]
- R1/R2 Control Stations missing IP Address & Hostname entry for each other in "/etc/hosts" file [Checked by each CS during -init]
- IP connectivity interrupted or not functioning, causing either R1 or R2 Control Station to not be able to "ping" the other
- R1 & R2 local symapi database is inconsistent and requires updating prior to running the #/nas/sbin/nas\_rdf -init command
- Control Station hardware failed, requiring replacement
- R2 script emulates an R1 configuration if it cannot detect the correct Control Volumes for an R2 configuration—check BIN files

## **Possible Remedies:**

--Verify R1/R2 Control Stations' /etc/hosts file to ensure that each other is referenced here by IP Address and Hostname  
--Verify R1/R2 Control Stations' /etc/sysconfig/network file to ensure that Hostname entry is correct

--Verify IP connectivity between the R1/R2 Control Stations & troubleshoot accordingly

--Verify RDF Link is operational between Sites

--Update R1 & R2 Control Stations' local "syma

a.) #mv /nas/symapi/db/symapi\_db.bin /nas/symapi/db/symapi\_db.old [R1 & R2 side]

- b.) #/nas/sbin/nas\_rdf -localinit [R1 & R2 side]

- Note:** Generally, the -localinit update of the symap

problems, a good practice would be to start over by running the -localinit update. The local symapi database on each side must be intact before any RDF operations between the two sites can work.

--Update R1/R2 SRDF Configurations by running from R1 side first, then R2:

- a.) R1 Control Station: #/nas/sbin/nas\_rdf -init [If no errors encountered, proceed with next step]
- b.) R2 Control Station: #/nas/sbin/nas\_rdf -init

--Verify Status of RDF Devices by running the following on R2 side:

- a.) #/nas/symcli/bin/symdg list [Obtain list of RDF Groups to query for 'Synchronization' and 'Invalid Tracks' status]
- b.) #/nas/symcli/bin/symrdf -g 1R2\_500\_1 query

**Normal Entries show RDF Devices are Synchronized and have no Invalid Tracks:**

DEV001 002 WD 0 0 RW 002 RW 0 0 SYN DIS OFF Synchronized

DEV002 004 WD 0 0 RW 004 RW 0 0 SYN DIS OFF Synchronized

**Entries might normally indicate a problem, but in this example, are "direct-attached" devices used by another Host:**

08C 11D R1:1 ?? RW NR SYN DIS OFF 0 0 RW WD Suspended

0DA 0C7 R2:1 NR WD RW SYN DIS OFF 0 0 NR NR Invalid

0DB 0C8 R2:1 NR WD RW SYN DIS OFF 0 0 NR NR Invalid

## **II. CELERRA ACTIVE-ACTIVE SRDF:** Min. Microcode: 5265.48.30 NAS Code: 2.1.15.16

Bi-Directional SRDF--both sides configured with R-1& R-2 DataMovers & Symm volumes. Only one side of SRDF volume pair can be write-enabled at a time. Synchronous Mode ONLY [Journal 0]. Each side has an Active and Passive component [R-1 & R-2]. An SRDF Standby server on Primary side must be matched on the Failover side in a 1-to1 relationship. That is, a Primary datamover must have an RDF Standby designated, and vice versa.

### **DATAMOVER PREREQUISITES:**

--SRDF Standby must have same network configuration as primary DM

--SRDF Standby must have same data volumes outlined in R2 Symmetrix BIN file

--Primary Datamover is paired to only one SRDF Standby in a 1 : 1 relationship [Pair Primary & Standby with same slot number]

### **BIN FILE PREREQUISITES:**

RDF R2 Not Ready if Invalid—set flag to ‘no’

RDF R2 Not Ready—set flag to ‘no’

Primary R1 Control Station maps to R2 Control Station volumes [scsi=10-13; fibre=06-09]

### **SETTING UP ACTIVE-ACTIVE SRDF--R-1 to R-1:**

- Step 1. **Login to local Celerra** as nasadmin and execute: #/nas/sbin/nas\_rdf -init  
[Disks are discovered locally for SRDF and the Remote Celerra/Symmetrix is contacted]
- Step 2. Create new user account to administer local R-2 Standby Servers for the Remote Celerra  
New Login: rdfadmin-localname  
New password: \*\*\*\*\* Re-enter password: \*\*\*\*\*  
[Discovery process then begins for Remote Celerra from Local]
- Step 3. Setup local server(s) as RDF Standby for Remote by specifying "ID" number, not "Slot" number of datamover
- Step 4. Then associate Primary local datamovers as Standby(s) for Remote by entering "Slot" number when prompted! Log out.  
**Note:** After set-up on Local Celerra, run \$nas\_server -l --note that the RDF Standby Server designated in Step (3) is not listed! It is accessible only when logged in as the special "rdfadmin-local" account created in Step (2).
- Step 5. **Login to Remote Celerra:** #/nas/sbin/nas\_rdf -init  
[Local disk discovery occurs for SRDF volumes & remote Celerra is contacted]
- Step 6. "Please create a new login account to manage RDF site 'remote'"  
New login: rdfadmin-remotename  
New password: \*\*\*\*\* Re-enter password: \*\*\*\*\*  
[Discovery of remote Celerra's volumes takes place]
- Step 7. Setup local datamover as Standby for the remote Celerra by "ID" number  
**Note:** Ensure that you match this "ID" with the "Slot #" of the Primary Server designated on the Remote Celerra!!
- Step 8. Associate a Primary datamover with a Standby on the "remote" Celerra by "slot" number [the two slot #'s should match]

### **ACTIVE-ACTIVE SRDF FAILOVER:**

- Step 1. Login to Celerra at site experiencing failure and power-off the Celerra
- Step 2. Login to remote Celerra as "rdfadmin-remotename" and activate failover: #/nas/sbin/nas\_rdf -activate  
[At prompt, continue failover and answer "Y"; R1 volumes on failed site become RO, SRDF link suspends, and R2 volumes on failover symmetrix to RW]
- Step 3. RDF Failover continues and Standby Server(s) are activated [device track tables from R1 and R2 are merged and SRDF link brought back online]

### **RESTORING ACTIVE-ACTIVE SRDF:**

- Step 1. Power up Symmetrix on site to be restored & ensure SRDF link is operational
- Step 2. Login to Remote Symmetrix site running as failed over Celerra [Use "rdfadmin-remote" account]

- Step 3. #/nas/sbin/nas\_rdf -restore [SRDF link suspended, track tables merged, link restored]
- Step 4. When prompted, power up the Celerra on the site to be restored, and continue with Remote Site restoration by saying "Y" to restore. Remote Celerra Datamovers are rebooted and give up control of their identity and volumes back to original site. R2 devices on Remote Site become RO once again while R1 volumes on Local Site are set RW]
- Step 5. If restore "fails" continue with following steps at Local Site being restored:
  - a) Login as nasadmin and execute command #/nasmcd/sbin/nas\_rdf -restore

**BIN Files:** "RDF R2 Not Ready if Invalid" and "RDF R2 Not Ready" flags set to "NO"

**Creating SRDF Standby DataMover on R2 for R1:** \$server\_standby server\_4 -create rdf=6 [Server\_6 on R2 is standby unit]

### **TROUBLESHOOTING SRDF:** Primary source: /nas/symapi/log/symapi.log

|                                            |                                                                                   |
|--------------------------------------------|-----------------------------------------------------------------------------------|
| #/nas/site/cshosts                         | [File shows Other Celerra if Active-Active; If blank, should mean Active-Passive] |
| #/nas/symcli/bin/symcfg list               | [Local & Remote RDF]                                                              |
| #/nas/symcli/bin/syndg list                | [shows RDF Device Groups]                                                         |
| #/nas/symcli/bin/symrdf -g 1R1_1 query     | [Device details of an R-1RDF Device Group]                                        |
| #/nas/symcli/bin/symrdf -g 1R2_500_1 query | [R-2 RDF Device Group]                                                            |
| #/nas/symcli/bin/symcfg list -v  more      | [Provides Switched RDF Environment information]                                   |

### **ACTIVE-ACTIVE SRDF & RDF CHAND:**

'Unable to connect to host'—\$nas\_fs -M off or \$fs\_timefinder -M off commands fail

Rdf\_chand requires that both R1 & R2 Control Station processes communicate via root\_rdf\_channel gatekeeper slice—if this slice volume does not exist on each side, then Mirror off command will fail.

### **SRDF LOGS:**

#/nas/log/nas\_rdf.log \*New with NAS 5.1:

#/nas/log/symapi.log

**Note:** Use \$nas\_server -a -i command to see RDF relationship of DataMovers

### **EXAMPLE SRDF MIGRATION: Symm to Symm Only!**

1. Stop any Timefinder and Checkpoint operations & verify that Crontab jobs are commented out
2. Break Standby DM relationships with Primaries
3. Halt DMs
4. Halt CSs
5. Power off Celerra
6. Synch up R1 & R2 sides and verify no invalid tracks
7. Break SRDF Link
8. Load new bin on R2 Symm to change to STD volumes
9. Apply new Fibre Channel Zone sets
10. Remove DMs and CS1 from slots
11. Bring up CS0 & verify operation
12. Check nas\_disk list, IP connectivity, and Usrmapper service
13. Bring up CS1 & test failover of CS
14. Modify /nas/volume/symm and /nas/volume/disks files with new Symm serial # and Symm device ID number
15. Move /nas/dev files to backup location
16. Move /nas/server/slot\*/scsidevs files to backup location  
[#cd /nas/server; tar -cvpBf - slot\_\* /scsidevs | (cd /home/nasadmin/backup /server\* ; tar -xf -)]
17. Copy /nas/server/slot\*/camdisk files to new location
18. Move /nas/symapi/db/symapi\_db.bin file to new location
19. Insert Standby server first and run setup\_slot

**Note:** Compare devconfig output from previous copy to new devconfig [server\_devconfig server\_4 -p -s -a ltee /tmp/cmd-out/devcfg-psa\_s4.post diff /tmp/cmd-out/devcfg-psa\_s4.pre /tmp/cmd-out/devcfg-psa\_s4.post]

20. Verify disk listings are correct, IP connectivity of DMs, etc  
[#nas\_disk -l | tee /tmp/cmd-out/n\_disk-l.post diff /tmp/cmd-out/n\_disk-l.pre /tmp/cmd-out/n\_disk-l.post]
21. Run server\_devconfig ALL -c -s -a
22. Verify mounts, exports from previous copies to present output
23. Repeat verification process for other DMs
24. Define and test DM failover

## **CLARIION CORE CURRICULUM:**

### **COMPARISON OF CX SERIES:**

CX200→200MB/sec; 15 hosts/array; 256 LUNs; 4 FC-AL or 2 FC-SW host ports; 1GB cache; 30 drives Fibre or ATA

CX400→680MB/sec; 64 hosts/array; 512 LUNs; 4 host ports; 2GB cache; 60 drives

CX600→1300MB/sec; 128 hosts/array; 1024 LUNs; 8 host ports; 8GB cache; 240 drives

### **CLARIION CX600 STORAGE SERIES:**

High-speed, redundant arrays using Fibre Channel topology and fibre channel disk drives

#### **CX600 CONFIGURATION:**

CX600 series uses SPE (Storage Processor Enclosures)

Min. configuration requires the CX600 SPE, the DAE-O/S [first 5 drives contain dedicated space for O/S, 15 Drive DAE, SPS Module

Each SP contains (4) fibre ports for direct-connect to Hosts—uses PowerPath sw on Host for failover

#### **CX400/200 CONFIGURATION:**

This series uses the ‘DPE’ (Disk Processor Enclosure), 1<sup>st</sup> five drives with dedicated O/S space, SPS Module, and Standard DAE

Each SP on CX400 contains (2) Fibre ports for direct-connect, while CX200 uses only a single port for each SP

**Note:** DPE’s are built into the LCC cards for the CX400/200

### **CLARIION RAID TYPES & HOT SPARE:**

#### **RAID-0:**

Data distributed across 3-16 disks for performance, no fault tolerance—disk failure = data lost (no parity)

Used for performance as I/O is spread across many channels and controllers

#### **RAID-1:**

Data Mirrored to a second disk

#### **RAID-1/0:**

Combines mirroring and striping for performance, using 2-16 disks, used for high I/O systems, mirrored raid 0

**Note:** See emc183839. RAID 1/0 will be supported in the 5.6.39.x release. RAID 1/0 devices are mirrored, similar to RAID 1, but can be in groups of 2, 4, 6, 8, 10, 12, 14, or 16 disks. Celerra will not diskmark Raid type 1/0 devices and skips them:

# nas\_diskmark -m -a

Discovering storage (may take several minutes)

Warning:

17716810659: server\_2 c0t3l3 skipping unmarked disk with health check error,

APM00073801838 stor\_dev=0x0033, RAID10(2), doesn't match any storage profile

# server\_devconfig server\_2 -p -s -a|grep -v no

chain= 0, scsi-0

tid/lun= 3/3 type= disk sz= 20479 val= -99 info= DGC RAID 10 03263300330033NI diskerr= unmarked

#### **RAID-3:**

Use 5 or 9 disks; uses striping across volumes for perf.; parity is always on the same disk

Not used with CX Clariion series, only legacy FC4700

#### **RAID-5:**

Striping and parity striped across all drives (20% vol. space used for parity); 3-16 disks; Good on Reads but slight perf. hit on Writes  
Striping uses XOR algorithms

**Note:** Celerra/Clariion configurations support only RAID 5 & RAID 1

#### **HOT SPARE:**

A Hot Spare is normally required for every 30 disks and can be used with RAID 1, 3, 5, etc.

Typically flare will invoke the Hot Spare with a data drive failure & data is rebuilt to the Hot Spare

### **BASIC CLARIION SETUP OF DISK DRIVES:**

1. Create Raid Groups
2. Create and Bind Luns from Raid Groups
3. Map and assign Luns to Hosts and default SP owner

### **CLARIION FIBRE CHANNEL:**

--Clariion Storage Systems are configured to work with two main Fibre Channel Topologies—‘Direct-Connect’ Arbitrated Loop or Switched Fabric connected switches.

--Copper fiber is used primarily for short distances (up to 30M) for internal cabling in the array, while fibre optics are used for longer distances (SWL up to 500M & LWL up to 10k) and Host connectivity to the array [HSSDC High-Speed Serial Data Cables, & LC Lucent Connectors, respectively]

**Note:** All fibre switch ports on Clariion or Symmetrix fabric-connected ports must be set to same speed as DM ports. Default port speeds are 2GB. Mixing of speeds not supported. Auto negotiation is not supported.

### **CLARIION CACHE FUNCTIONALITY:**

- Read Cache is independent on each SP
- Write Cache is mirrored to peer SP
- Primary purpose of cache is to prevent ‘data loss’

- All Reads and Writes are completed in cache
- SPS modules exist primarily to ensure that write cache can be properly ‘destaged’ to disks in the event of a power outage
- Vaults are sections on the first five DAE-OS disks that cache destages to in emergency situations
- Read Cache uses ‘prefetch’ and LRU algorithms, first meaning that next ‘read’ will most likely be from next sequence of data on disk, and LRU means that oldest unused cache is overwritten with newer cache to enable better access to ‘current’ data
- All write cache mirrored between SP’s via the CMI peer bus [acknowledgements to Hosts must wait for this operation]
- When write cache is disabled, Host writes to directly to destination drives and acknowledgement is sent back to complete
- Low Watermark cache flush default is 40%--anything below this value is considered ‘idle flushing’
- Max of (4) threads can be assigned to flush cash if reaching the High Watermark [Forced Flushing can occur with write cache full]
- Cache is enabled at the Array level, but can also be enabled/disabled at the LUN level

### **WRITE CACHE DISABLED:**

- Requires that both SPs be operational
- Requires at least one fully charged SPS battery
- During FLARE upgrades
- Faulted SPS pwr supply
- Faulted vault drive in Encl 0

### **CLARIION STORAGE PROCESSORS (CX600):**

- Personality Module consisting of 4-ports for Host connectivity and (2) SP’s
- BE0 & BE1 ports to connect to the disk arrays & extend to additional DAE’s
- Aux 0 & 1 ports are NOT used in CX models
- Serial connection for system configuration
- Connection to SPS module
- Cat 5 Ethernet management port (Navisphere)
- Green and Amber indicator lights
- (2) power supply modules
- (3) Fan Modules containing (6) fans, hot-swappable, but requiring at least (5) fans to keep array up [2 minute shutdown with less]

### **CLARIION STANDY POWER SUPPY MODULE:**

- Input power switch from AC source
- Output receptables for SPA and DAE-OS
- RJ11 jack with db9 male connector to connect to CX600 SPE for monitoring

### **CLARIION DISK DRIVE SPECS:**

FC Drives 10K RPM (73, 146, & 300GB models); 140IO/s 2-8KB random requests; 10MB/s bandwith for sequential 128KB requests  
FC Drives 15K RPM (73 & 146GB); 180IO/s random requests; 12MB/s bandwidth for sequential requests  
PATA Drives 5400RPM 320GB; 50IO/s random; 7MB/s sequential  
SATA (Serial Advanced Technology Attachment) Drives 7200RPM (250 & 500GB); 60IO/s random; 8MB/s sequential  
LCFC Drives 7200RPM (500GB); 80IO/s random; 8MB/s sequential

**Note:** Drive capacity ratings can be misleading, as a 36GB drive is 36,000,000,000 bytes in base 10 used by mfg, but only 33GB in computer systems using Base 2

### **CLARIION DISK ARRAY ENCLOSURE (DAE):**

- 0-15 disk enclosures, with each disk dual-ported to LCC-A and LCC-B
- Contains SPB and SPA power supplies
- Contains SPB and SPA Link Control Card modules [LCC’s]
- Up to 15 disks per DAE, each drive shows amber fault light or green for normal
- DAE-OS Vault drives 0-3 required for XP O/S bootup, drives 0-4 required for enabling write cache
- Total of (8) DAE’s per bus loop (120 drives), for a maximum of 240 drives

### **CLARIION ATA DAE2 DISK ENCLOSURE:**

- Maxtor 250GB Tomcat Disk Drives, 5400 RPM, data written in 512 byte blocks
- Requires that first enclosure OS drives be RAID 5 Fibre Channel (true for all Clariion arrays)
- Requires Navisphere 6.4 to manage

### **FIRST FIVE ARRAY DISK DRIVES:**

- CX600 = DAE2-OS; CX400/300 = DPE2-OS
- First 6GB of each of the five drives is reserved for system environment
  - Flare database is triple mirrored across 28MB/disk on disks 0-2
  - PSM is triple mirrored across 0-2 and keeps SPA & SPB synchronized
  - Drives 0-4 contain 2176MB/disk vault area to destage write cache
  - SPA’s OS is mirrored on Disk 0 & 2 while SPB is mirrored on Disks 1 & 3

### **CLARIION DISK PROCESSOR ENCLOSURE CX400/200:**

- Contains SP’s where LCC cards are in the CX600
- SPA is directly connected to first 15 drives via backplane

--Min. configuration DPE + SPS

## **CLARIION RAID GROUPS & LUNS:**

Raid Groups are created to pool disks together and are carved up into usable Logical Units for Host Systems through the use of LUNs  
Luns are striped across the RAID Group and therefore must be of the same RAID type

**Note:** Prior to Flare 12, could only have 32 Luns per RG. Can now have 128 luns/RG.

## **RAID GROUP EXPANSIONS:**

--Raid Groups can be expanded with active LUNs attached

--Raid Groups can add drives within the limits of the Raid Type [max of 16 drives per Raid 5 RG, etc.]

--In the case of multiple LUNs, newly added disks do not change LUN size--free space is pooled at bottom of RG for expansion

--In case of single LUN Raid Groups, new disks added will increase the size of the LUN [optional]

--In fragmented RG's, consolidate free space in RG's by using defrag if a single large LUN is desired

## **LUN EXPANSIONS:**

### **METALUNS:**

Purpose of metaluns are to increase spindle count for increased performance of LUNs, and to be used for LUN migration to change LUN type

--Creating new LUNs and adding to a 'base' LUN; concatenated or striped metaluns

--With CX600, can add 16 components to a metaLUN, which in turn can have 32 Flare Luns per component

--CX400 can add 8 components to a metaLUN

### **Concatenating:**

--creates metaLUN components for increasing LUN size

--either protected or unprotected

--any size scheme except all FC or all ATA drives

### **Striping:**

--requires use of same Disk drives, same Raid scheme, and same sizes as base LUN

--restripes data across all space once new luns are added

## **CLARiiON LUN MIGRATION**

### **CLARiiON CX-to-CX3 CONVERSIONS (aka Data-in-Place Upgrade)**

→Clariion offers a migration path from CX to CX3 series arrays with FC drives—see February 2007 White Paper “Upgrading to CLARiiON CX3 UltraScale Series Storage Systems”

→Submit CCA for LUN Migrations that involve Celerra

→Upgrading from CX series results in 2Gb/s loop speeds (Purchase new CX3 array to get 4Gb/s capability)

→CAP2 Conversion Readiness tool is used to determine if User or Private LUNs need to be relocated prior to the hardware Upgrade

→Requires PS engagement

→CLARiiON Procedure Generator for “LUN Migration for CX3-Series Conversions” provides link in instructions on what to do if NAS is attached to a CLARiiON array: Celerra NAS – CLARiiON Backend Upgrades 5.5 Technical Note

→Backend upgrade conversion with attached Celerra is only supported on Gateway models

→ConversionPrep NDU utility required on array [UtilityPartition NDU, ConversionImage NDU pkg]

### **NAS Tech Note Recommendations:**

--CLARiiON LUN migration must be used to execute migration of NAS LUNs

--NAS Control LUNs must be migrated to Fibre Channel drives

--NAS Control LUNs must be moved to a single RAID group, with Source & Target LUN sizes identical by Block count

--NAS Data LUNs should be moved to a RAID group not already containing NAS LUNs

--NAS Data LUNs should (really needs to say “must”) be migrated to target disk technology compatible with the source

**Note:** Must migrate to same Raid Type, same Raid Protection—e.g., r5 (4+1) to r5 (4+1), same drive types

### **IF CAP2 TOOL REQUIRES LUN MIGRATIONS:**

1. Create Destination LUNS (destination LUNs must have same block count as source LUNs)

2. Use Navisphere to Migrate source to destination LUN and select migration rate of High

3. Do not continue with migration if a message indicates that target LUN is larger in size (because this will extend the LUN, which isn't allowed for Celerra)

4. Wait for LUN migrations to complete before continuing with CX to CX3 Upgrade procedure

### **EE Guidance on LUN Migrations on Celerra:**

--Do not use CLARiiON LUN extension to increase LUN size for an existing NAS LUN (see emc157859)—use nas\_fs -xtend

--Migrate LUNs where source & destination blocks are identical

--Migrate LUNs onto same RAID type (R5 vs. R3 or also means 4+1 to 4+1?) and physical disks

--Migrate LUNs to same Owner SP as source

## **WHY ARE LUN EXPANSIONS BAD FOR CELERRA LUNS?**

Diskmarks are written 128 sectors (512 bytes/sector) from the end of a LUN when first diskmarked by Celerra. Also, the first 8k block (sectors 0-15, 8192 bytes) is moved to the 2<sup>nd</sup> to the last 8k block of a LUN, meaning that the 1<sup>st</sup> and last 8k blocks are free for

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
MPFS client signatures. Both the diskmark and the 2<sup>nd</sup> to last 8k block are addressed from the END of the LUN, not from the beginning of the LUN. Celerra will lose diskmark if lun is expanded, and also data when first block of lun is used.

#### **512 bytes per Sector**

#### **16 Sectors per 8K block**

#### **Diskmark located at 8<sup>th</sup> 8K block from the END of LUN**

### **CLARIION SOFTWARE FEATURES:**

#### **Access Logix→Allows heterogeneous hosts to be attached to an array via “Storage Groups”**

→Provides Lun masking to Hosts for Shared arrays (Array-based SW that is NOT enabled by default)

**Note:** Access Logix used with NS600G/700G & CFS-14/CX600/CX700 on Shared SANs. Access Logix is NOT used with captive backends such as NS600/700. Access Logix Base Code contains “1” in 3<sup>rd</sup> field [2.05.1.60.5.xxx].

#### **To enable Access Logix:**

#### **Navisphere Manager>Properties>Storage Access>Select “Access Control Enabled”**

**SnapView Version 2.0**----→Point-in-time online virtual copy of LUNs--up to 8 per lun; Full synchronous copies called ‘clones’  
SnapView Snapshots and BCV’s

**MirrorView**→Remote DR solution over FC or T3 lines using synchronous mirroring—Source LUNs can have up to (2) mirrors to other Storage systems which can participate with Host writes.

**SANCopy**→Used to migrate data at a block level between LUNs, within an array or to other arrays, Symmetrix systems, and misc. supported 3<sup>rd</sup> party systems. Requires installation of a software enabler on the array. Uses the SAN to copy the data and does not use a Host. Considered the optimal method for migrating data within a CLARiiON system. Supports iSCSI beginning with Flare 26. Launched via Navisphere Wizard or from Navicli.

**PowerPath**—Host resident software delivers intelligent I/O path management, I/O balancing, and automatic failover capability

**Non-Disruptive Upgrade**--NDU upgrades of CX series are non-disruptive if data has path to both SPs & Hosts are using PowerPath.

### **CLARIION ARRAY INITIALIZATION & SETUP:**

#### **SETTING UP A NEW CLARIION ARRAY:**

1. Cable array for direct attached arbitrated loop Hosts or for Fabric Connected Hosts
2. Install Navi CLI, Navi Host Agent, Windows UIs, Windows Management Server, & Java Runtime Environment software on Windows 2000 Server or Laptop [d:>Host Software>W2K>Service Toolkit>ServiceT.exe]
3. Initialize array by powering on
4. Connect serial cable from Windows 2000 Server to SPA and create DialUp Networking PPP Connection:  
Network and Dial-up Connections>Make New Connection>Enter 3-digit area code if prompted>Network Connection Wizard>\*Connect directly to another computer>\*Guest>Communications cable between two computers (COM1)>\*Only for myself>type in name for the connection>Enter Username: clarion\_passwd: clarion!>General Tab: COM1>Configure>Set speed to 115200 and o.k.>Networking Tab: Uncheck all boxes except TCP/IP>Click on ‘connect’ using ‘clarion’ and password ‘clarion’!
5. Once PPP Session is launched and you are logged into the Array, open a Web Browser to connect to SPA:  
Default IP Address to configure Clariion Array using Web Browser: 192.168.1.1/setup
6. Set the desired IP address for SPA:  
IP Address: 192.168.2.124  
Hostname: SPA\_124  
Subnet Mask: 255.255.255.0  
Gateway: 192.168.2.1  
Peer IP Address: 192.168.2.125 [IP of SPB]
7. Click on ‘Apply Settings’ and SPA will reboot—connect to SPB while SPA is rebooting and configure IP Address for it  
**Note:** Enter Peer IP Address of SPA and ‘Apply Settings’ to reboot SPB
8. Reconnect serial cable to SPA, connect via PPP session, open browser to 192.168.1.1/setup:  
“Restart Management Server”  
**Note:** This will allow SPA to discover IP Address for SPB and syncs the PSM database. From this point forward, you no longer need the serial connection to SPA or the PPP session as you will use the Web Browser directly to the IP Address of either SPA or SPB.
9. Establish Global Security and Domain Name: c:>Program Files\EMC\ManagementUI\6.5.0.1.79\WebContent>start.html
10. Select ‘LAN Connection’ for ‘Select Connection Type’ and enter IP Address of SPA  
**Note:** Answer Yes to initialize global security and ‘Create Global Administrator’ Username and Password to be used on system
11. Click on ‘File’>Setup Domain>Select Master and enter ‘Master Node IP Address’ of SPA
12. Expand the ‘Enterprise Storage 1’ tree for the displayed CX600 system
13. Rightclick on SPA and update ‘Network’ information, set ‘Fibre Speed’, Add Privileged Users to ‘Agent’ tab, etc.  
**Note:** Repeat for SPB
14. To change Storage Domain Name, go to ‘File’>Setup Domain>Configure Domain>Change>add new domain name
15. To add Storage Domain Users, go to ‘Tools’>Security>User Management>Add>add users as required
16. For all new installations, one of the next steps will be to ‘Commit’ the Base Software package

Rightclick array>Properties>Software>Base>rightclick and select “commit”

**Comments:** First array configured is a ‘Domain Master’. User accounts have three roles, ‘Administrator’, ‘Manager’, & ‘Monitor’

17. Add or Upgrade SW packages on the array as required:

Highlight Array>Rightclick and select ‘Software Operations’>Software Installation Wizard>Browse CD-ROM for software and add: Management UI 6.5.0.3.3

**Note:** After installation of the Management UI software, you will be able to connect to the Management Server display directly from a Web Browser to the IP Address of SPA without using the ‘Start.html’ page

### **EXAMPLE OF SOFTWARE PACKAGES ON ARRAY:**

|                   |                                         |
|-------------------|-----------------------------------------|
| Base              | 02.05.0.60.5.008                        |
| ManagementServer  | 6.5.0.3.10                              |
| Navisphere        | 6.5.0.3.7                               |
| ManagementUI      | 6.5.0.3.3                               |
| Base              | 02.05.1.60.4.004 [Access Logix Base SW] |
| SnapViewUI        | 6.5.0.2.55                              |
| SnapView          | 02_20_01                                |
| SnapCloneProvider | 6.5.0.2.55                              |

**Note:** After new ‘base’ software is installed and then ‘committed’, the original base will be deleted from this list

### **INSTALLING UTILITY PARTITION AND RECOVERY IMAGE:**

1. Install Utility Partition Image and Recovery Image software using Software Installation

### **INVOKING ARRAY RECOVERY PROCEDURE:**

1. Connect null modem cable to serial port on SPA
2. Connect with hyperterminal session [9600; 8 bits; none; 1; Hardware]
3. Reseat SPA to reboot, during Extended POST, the following characters begin scrolling across the screen. At this point, type “ctrl + c” keys (or Esc key if that does not work), which allows POST to complete, and generates a “...Storage System Failure message, then enter the password indicated in the next step  
AabcdeBCDabEabcd
4. Then enter password on screen: **DB\_key**
5. Enter (3) for DDBS Service Sub-Menu at “Diagnostic Menu”
6. Select (2) Utility Partition Boot
7. After Utility Partition boots, press Enter key and select “Install Images” in Toolkit Menu
8. At ‘Select Images’ display, enter number corresponding to NDU Recovery Image installed for system
9. Install image on SPA only
10. Select (1) Reset Controller from “Diagnostic Menu”

**Note:** This will boot SPA from new image and requires (3) reboots to complete install

### **POWERPATH SOFTWARE:**

Host-based software providing Host-to-Storage I/O path management/failover and maximized performance, found predominantly on fibre channel. Support for multiple paths to logical devices so as to enable failover in the event of hardware failure. Provides dynamic load-balancing of I/O requests to logical devices across the available paths.

1. Install from CD-ROM by running ‘emcpp.w2000.3.0.5.ga.exe’
2. Enter License Key when prompted: B4P9-DB4Q-LF6W-QOSA-ML9O-VRL4
3. Reconfigure Windows Host HBA drivers for PowerPath using Emulex Driver Utility

**CURRENT VERSION:** <http://Powerlink.emc.com>

PowerPath v4.4.0 for Windows, March 2005—support for Symm Microcode 5x71 & Clariion Flare 17

PowerPath v4.4.0 for Solaris, March 2005—support for Solaris 10 & end of service life for PowerPath Volume Manager for Solaris, AIX (PPVM)—must backup data from PPVM volumes and destroy PPVM volume groups before installing this version.

### **PowerPath 4.5:**

Latest version supports EMC Symmetrix, EMC CLARiiON, Hitachi Data Systems (HDS) Lightning, HP XP, IBM Enterprise Storage Server (Shark), and HPQ storage systems. Supports Solaris, HP-UX, AIX, Linux, & Windows Servers.

### **USING ACCESS LOGIX:**

1. Install base Access Logix software
2. Open Navisphere Manager>Properties>Storage Access>Access Control Enabled
3. Create RAID Group; Bind LUNs to RAID Group; Create Storage Groups; Assign Hosts to Storage Group

### **CLARIION HOST CONNECTIVITY/TROUBLESHOOTING HOST CONNECTIVITY:**

→Fibre Channel Zoning must be correct

→Failover SW must be in place and active

→Storage Groups must be defined if Access Logix and LUN Masking used between LUNs and Hosts

→HBA Configuration setup: Persistent Binding; Initiator configuration, Failover mode, etc.; HBA registered manually or via Navi-Agent. Host HBA must log into the SP to be able to be Registered

#### **TROUBLESHOOTING TOOLS:**

- navicli commands
- HEAT grabs for Host configurations
- Switch Zoning, Port stats, logs

#### **USING NAVICLI TO VERIFY/CHANGE HBA & SP CONNECTIVITY STATUS:**

**`$ /nas/sbin/navicli -h 142.9.150.201 port -list | port -list -sp -all`**

**Note:** The port -list command provides detailed information on each Celerra HBA and what SP ports it is connected to, and the StorageGroup information. This command only works with SAN attached NSG products. Integrated NS platforms do not use StorageGroups or AccessLogix. This command also returns status of all the SP ports. In the following example, Celerra is only connected to Ports 0 & 1 on each SP, but each SP has a total of (4) ports for the NFS701G model. This command also has other options, such as -diagnose -host; -removeHBA -host; register -list

#### **Examples:**

##### **LIST OF DATA MOVER HBA's-to-SP PORTS:**

**`# /nas/sbin/navicli -h 192.1.4.220 port -list -hba`**

##### **DISCONNECTING HOSTS FROM STORAGEGROUP:**

**`# /nas/sbin/navicli -h 192.1.4.220 storagegroup -disconnecthost -host emcnas_i0_dm2_p1 -gname`**

**Celerra\_emcnas\_i0**

##### **REMOVING HOST INITIATOR RECORDS FROM SP:**

**`# /nas/sbin/navicli -h 192.1.4.220 port -removeHBA -host emcnas_i0_dm2_p1`**

Remove the following initiator records for: emcnas\_i0\_dm2\_p1

50:06:01:60:C1:E0:53:35:50:06:01:61:41:E0:53:35 SP B-2

Only logged out but registered initiator records will be removed. Do you want to continue (y/n)?

##### **REMOVING HBA INITIATOR RECORDS BY HBAUID FROM SP:**

**`# /nas/sbin/navisecli -h <sp_IP> port -removeHBA -hbauid 10:10:10:10:10:10:10:10:10:10:10:44:55:66:77:88`**

(y/n)?

##### **REMOVING ALL HBA RECORDS:**

**`# /nas/sbin/navicli -h 10.241.168.179 port -removehba -all`**

Remove the following initiator records for all attached hosts:

50:06:01:60:C1:E0:D9:54:50:06:01:60:41:E0:D9:54 SP A-0

50:06:01:60:C1:E0:D9:54:50:06:01:61:41:E0:D9:54 SP B-0

50:06:01:60:C1:E0:D9:54:50:06:01:68:41:E0:D9:54 SP A-1

50:06:01:60:C1:E0:D9:54:50:06:01:69:41:E0:D9:54 SP B-1

Only logged out but registered initiator records will be removed. Do you want to continue (y/n)? y

**Note:** If command fails because the Data Movers are logged-in, shutdown the Celerra or reboot the blades

##### **CREATING INITIATOR RECORDS FOR SPECIFIC PORTS IN STORAGEGROUP:**

**`# /nas/sbin/navisecli -h <sp_IP> storagegroup -setpath -hbauid 10:10:10:10:10:10:10:10:10:10:10:44:55:66:77:90 -sp b -spport 2`**

##### **ADDING & REGISTERING DATA MOVER HBA's TO SP's AND STORAGEGROUP:**

**`# /nas/sbin/navicli -h 192.1.4.221 storagegroup -setpath -gname Celerra_emcnas_i0 -hbauid`**

**50:06:01:60:C1:E0:53:35:50:06:01:61:41:E0:53:35 -sp b -spport 2 -ip 128.221.252.2\* -host**

**emcnas\_i0\_dm2\_p1 -failovermode -arraycommarray 0 -unitserialnumber array**

\***Note:** Data Mover 2 would use 128.221.252.2 and 128.221.253.2 for its internal IP address, while DM3 would use 252.3 and 253.3

##### **Information about each HBA:**

HBA UID: 50:06:01:60:90:60:2F:8D:50:06:01:60:10:60:2F:8D

Server Name: CS-NS\_DM2\_BE0

Server IP Address: 192.168.1.2

HBA Model Description: NS701G

HBA Vendor Description: Celerra

##### **Information about each port of this HBA:**

SP Name: SP A

SP Port ID: 0

HBA Devicename: N/A

Trusted: NO

Logged In: YES

Source ID: 7014419

Defined: YES

Initiator Type: 3

**StorageGroup Name:** C-NS701G

**SP Name:** SP B

**SP Port ID:** 1

HBA Devicename: N/A

Trusted: NO

**Logged In:** YES

**Source ID:** 7014419

**Defined:** YES

Initiator Type: 3

**StorageGroup Name:** C-NS701G

**Information about each HBA:**

HBA UID: 50:06:01:60:90:60:2F:8D:50:06:01:61:10:60:2F:8D

**Server Name:** C\_NS\_DM2\_BE1

Server IP Address: 192.168.2.2

HBA Model Description: NS701G

HBA Vendor Description: Celerra

HBA Device Driver Name: N/A

**Information about each port of this HBA:**

**SP Name:** SP A

**SP Port ID:** 1

Logged In: YES

Source ID: 7079955

Defined: YES

Initiator Type: 3

**StorageGroup Name:** C-NS701G

**SP Name:** SP B

**SP Port ID:** 0

Logged In: YES

Source ID: 7079955

Defined: YES

Initiator Type: 3

**StorageGroup Name:** C-NS701G

**Information about each SPPORT:**

**SP Name:** SP A

**SP Port ID:** 1

SP UID: 50:06:01:60:B0:60:09:BD:50:06:01:61:30:60:09:BD

Link Status: Up

Port Status: Online

Switch Present: YES

Switch UID: 10:00:08:00:88:03:36:00:20:05:08:00:88:03:36:00

SP Source ID: 7079187

**SP Name:** SP A

**SP Port ID:** 0 -----abridged-----

**\$ .server\_config server\_2 -v "fcp bind show"**

\*\*\* Persistent Binding Table \*\*\*

Chain 0000: WWN 50060160306009bd HBA 0 SP-a0 Bound

Chain 0016: WWN 50060169306009bd HBA 0 SP-b1 Bound

Chain 0032: WWN 50060161306009bd HBA 1 SP-a1 Bound

Chain 0048: WWN 50060168306009bd HBA 1 SP-b0 Bound

**Note:** Same system, but from Celerra Bind Table perspective

**\$ .server\_config server\_2 -v "fcp show"**

FCP ONLINE HBA 0: S\_ID 6b0813 WWN: 5006016010602f8d DX2 →WWN of Celerra HBA 0

FCP scsi-0: HBA 0: D\_ID 6b0413 SP-a0: 50060160306009bd Class 3 →WWN of SPA Port 0

FCP scsi-16: HBA 0: D\_ID 6b0513 SP-b1: 50060169306009bd Class 3 →WWN of SPB Port 1

FCP ONLINE HBA 1: S\_ID 6c0813 WWN: 5006016110602f8d DX2 →WWN of Celerra HBA 1

FCP scsi-32: HBA 1: D\_ID 6c0513 SP-a1: 50060161306009bd Class 3 →WWN of SPA Port 1

FCP scsi-48: HBA 1: D\_ID 6c0413 SP-b0: 50060168306009bd Class 3 →WWN of SPB Port 0

FCP OFFLINE HBA 2: ALPA 000001 WWN: 5006016210602f8d DX2

**\$ /nas/sbin/navicli -h 142.9.150.201 storagegroup -list**

**Storage Group Name: C-NS701G**

Storage Group UID: 6C:2D:F8:9E:22:F5:D8:11:A8:3F:08:00:1B:43:7A:AF

**HBA/SP Pairs:**

| HBA UID                                         | SP Name | SPPort |
|-------------------------------------------------|---------|--------|
| 50:06:01:60:90:60:2F:8D:50:06:01:60:10:60:2F:8D | SP B    | 1      |
| 50:06:01:60:90:60:2F:8D:50:06:01:60:10:60:2F:8D | SP A    | 0      |
| 50:06:01:60:90:60:2F:8D:50:06:01:61:10:60:2F:8D | SP A    | 1      |
| 50:06:01:60:90:60:2F:8D:50:06:01:61:10:60:2F:8D | SP B    | 0      |

**HLU/ALU Pairs:**

| HLU Number | ALU Number |
|------------|------------|
| 0          | 0          |
| 1          | 1          |
| 2          | 2          |
| 3          | 3          |
| 4          | 4          |
| 5          | 5          |
| 16         | 20         |
| 17         | 21         |
| 18         | 30         |
| 19         | 31         |
| 20         | 40         |

Shareable: YES

**Note:** Can also –addhlu, -connecthost, -create, -destroy, -enable, -removehlu, -sethost, -setpath, and other options from CLI**CREATING STORAGEGROUP FROM CLI:**

# /nas/sbin/navicli -h 10.241.168.57 storagegroup –create –gname Celerra\_new

# /nas/sbin/navicli -h 10.241.168.57 storagegroup –shareable –gname Celerra\_new yes

**REGISTERING DM WWN's WITH SP PORTS & HOSTNAMES IN STORAGEGROUP:**

# /nas/sbin/navicli -h 10.241.168.57 storagegroup -setpath -o -gname Celerra\_new -hbauid

50:06:01:60:00:60:03:7f:50:06:01:60:00:60:03:7f -sp a -spport 0 -type 3 -failovermode 0 -arraycommpath 0

**Note:** Data Mover WWN registered to SPA Port 0**CREATING RAID GROUP AND BINDING LUNs:**

# /nas/sbin/navicli -h 10.241.168.57 createrg 10 0\_1\_0 0\_1\_1 0\_1\_2 0\_1\_3 0\_1\_4 (RG 10 on drives indicated)

# /nas/sbin/navicli -h 10.241.168.57 bind r5 16 -rg 10 -rc 1 -wc 1 -aa 0 -sp a -sq gb -cap 100 elsz 128

**ADDING LUNS TO STORAGEGROUP:**

# /nas/sbin/navicli -h 10.241.168.57 storagegroup –addhlu –gname Celerra\_new –hlu 16 –alu 16

**CLARIION WORLD WIDE NAMES:**Resultant SPA port 0 WWPN – 50:06:01:**60**:00:60:01:b2Resultant SPA port 1 WWPN – 50:06:01:**61**:00:60:01:b2Resultant SPA port 2 WWPN – 50:06:01:**62**:00:60:01:b2Resultant SPA port 3 WWPN – 50:06:01:**63**:00:60:01:b2Resultant SPB port 0 WWPN – 50:06:01:**68**:00:60:01:b2Resultant SPB port 1 WWPN – 50:06:01:**69**:00:60:01:b2Resultant SPB port 2 WWPN – 50:06:01:**6a**:00:60:01:b2Resultant SPB port 3 WWPN – 50:06:01:**6b**:00:60:01:b2**Note:** SPA's ports are represented in blue (ports 0-3), while SPB's ports are in red (8, 9, a, b)

BITS 127 – 124 [Represent WWNN IEEE type]

BITS 123 – 100 [Represent WWNN ID assigned by IEEE]

BITS 99 – 96 [Represent WWNN padded to 0]

BITS 95 – 64 [WWNN bitwise ‘OR’ of 0x80000000 and the 32 bit WWN from Seed prom]

BITS 63 – 60 [WWPN IEEE type]

BITS 59 – 36 [Company ID number set by IEEE]

BITS 35 – 32 [Differentiates among multiple ports with same Node Name]

BITS 31 – 0 [32-bit WWN seed from prom on midplane]

**INDICATOR LIGHTS ON SYSTEM BOOT:**

BIOS Test: Green LED On; Amber LED blinks once every 4 seconds

POST Test: Green LED On; Amber LED blinks once every second

NT Boot & Driver Load: Green LED On; Amber LED blinks 4 times per second

Successful Boot: Green LED On; Amber Off

### **INTERNAL DISPLAY LIGHTS:**

(4) Green LEDs in a Row 0 1 0 0 [Green + Yellow LED row decodes to number error code 41]

(4) Yellow LEDs in a Row 0 0 0 1

**Decimal:** 8 7 6 5 4 3 2 1 0

**Binary:** 1000 0111 0110 0101 0100 11 10 1 0

### **HYPER TERMINAL SETTINGS:**

9600 baud; 8 data bits; no parity; 1 stop bit; no flow control

### **NAVISPHERE MANAGER INDICATORS:**

T [Means component or system is in transition] F [Means component(s) have failed] X [System inaccessible] ? [Unsupported]

### **CLARIION ARRAY EVENT/ERROR MESSAGES:**

**FIBRE DISK MIRROR DRIVE EVENTS:** Events all begin with 7124

7 1 2 4 **4** 0 0 5 [Translates into a warning: ‘Attempt to reconstruct primary after a verify error.’]

0=Informational 4=Warning 8=Error C=Critical Error

### **CLARIION SNAPVIEW:**

Main purpose is to provide online backups using point in time “snapshots” or full copies called “clones” of LUNs

Requires SnapView NDU, Clone Provider NDU, CLI Provider NDU, and SnapView UI for Navisphere

#### **SNAPSHOTS:**

Point-in-time virtual copy of LUN, requiring about 10-20% space

SnapShot Cache resides on the disk drive’s ‘private lun’ while SnapShot Lun resides in SP memory

SSLUN [SnapShot Source LUN] and SCLUN [SnapShot Copy LUN, stored in SP RAM] and one Snapshot Cache area per SSLUN

**Conducting a SnapShot:** Up to 8 concurrent snapshot sessions per LUN

1. Source LUN is ‘stabilized’ and put into ‘hot backup mode’, db is offline
2. Snapshot Session started
3. Source Lun restarted, taken out of ‘hot backup mode’
4. Snapshot is activated
5. Backup software backs up snapshot [mount the snapshot]
6. Snapshot deactivated & Snapshot Session stopped

#### **Limitations:**

--min. of 2 cache luns [private lun area] per Source LUN, total of 100 per Clariion array

--default snapshot chunks are in 128 64k blocks

--max of 8 Snapshots per source LUN, 300 per Clariion

--max of 8 Snapview Sessions per Source LUN, 100 per Clariion

#### **SnapClone:**

Full point-in-time copy of a LUN, equal in size to Source LUN

Max 8 active Clones per source LUN, 50 total per Clariion

## **FIBRE CHANNEL:**

### **(2) categories of fibre:**

**Multimode:** 125-400micron core allowing light source multiple paths down fiber—hence ‘multimode’—shorter distances than Single Mode; Cable color is orange

**Single Mode:** 2-8micron core, spans very long distances; cable color is yellow

**Light Source:** Multimode Fibre commonly uses LEDs for its lightsource while Single Mode uses Lasers.

**Transmission Methods:** TDM (Time Division Multiplexing) & FDM (Frequency Division Multiplexing). Combinations of the two are called WDM (Wavelength Division Multiplexing)

Fibre signals are pulsed as 0 & 1 bits onto copper wire or as “on & off” pulses of light via LED or Laser light sources

Fibre can travel long distances through the use of “regeneration” techniques using OEO (Optical/Electrical/Optical) or FA (Fibre Amplifiers).

#### **Fibre Channel Connectors:**

FC—connectors are screwed on

SC—connectors are plugged into sockets

LC—connectors are smaller but similar to the SC

ST—connectors are placed into sockets and turned ¼” to lock

MT—connectors are the smallest

#### **Fibre Optic Distances:**

SWL Multimode = 500M LWL Single Mode = 10K

#### **Copper Cable Distances :**

Both twinaxial and HSSDC support distances up to 30M

### **CELRRA FIBRE CHANNEL:**

→Supports switched fabric Class 3 service

→Receiving Host is always responsible for flow control with fibre channel

→FCP Class 3 FC does not perform retransmissions—relies on timeouts, but then must resend entire I/O stream

**Fibre Channel:** Serial interface, 100-200mb/sec, Copper Cable distance is 5-20m; Fiber Cable distance is 150m – 10km. Arbitrated Loop can have 126 devices, Switched Fabric 16 million. Supports SCSI, TCP/IP, VI & IPI, HIPPI

→Fibre channel is full duplex block-oriented high-speed serial network protocol 1-10Gbps [Celerra operates at 100MB/sec].

→Encoding schemes called 8b/10b control data flow & transfers using SOF (Start of Frame) & EOF (End of Frame). Frames can be up to 2148 bytes. The "fabric" of a Fibre Channel topology consists of (1) or more Switches connect to one or more port through F\_Port & N\_Port nodes.

**NODES:** Nodes represent equipment or devices such as HBA's, or SP's with built-in Fibre Channel Interfaces

**PORTS:** Ports represent the cable connections on the Node or Device

**Note:** For Fibre Channel Control Stations and DataMovers--Access up to 256 SCSI addresses per SCSI channel

### **TWO TYPES OF FIBRE CHANNEL TOPOLOGIES :**

#### **ARBITRATED LOOP (FC-AL) :**

**Note:** Also known as 'Direct Connect' or 'Point-to-Point', similar to 'token-ring' Hub technology in that only one device can be communicating to the Storage Processor at a time, therefore sharing the 'loop'.

→ALPA's are unique 2-digit numeric identifiers in hex for each port on an Arbitrated Loop topology, ranging from 0xef to 0x01, lowest to highest priority, respectively.

→Loop ID's are a decimal representation of the ALPA, 0-125 [lowest to highest]

→Loop Initialization Process [LIP] is the automatic login and arbitration for unique ID and for speed with the loop for new or disconnecting devices

**Note:** Loop ID's and ALPA's help maintain communication on the 'loop' via 'arbitration'.

→Daisy-chain of two or more devices using a FC Hub with one Port controlling the loop.

→Only one conversation on the 'loop' at a time.

→500 Meters, 2Gb/sec, 126 devices per loop, but Clariion imposes a limitation of 120 devices

→Limited to a single O/S type.

**Note:** PBC's [Port Bypass Circuit] allow for the protection of the loop while devices are accessing or leaving, or failing

→LCCs [Link Control Cards—FC-AL interface] are the actual built-in hub on each Clariion Storage Processor used in the 'loop'

**Note :** If one Host goes down in loop, if using Fibre Channel Hub architecture and Loop Initialization Process, the faulty or missing Host is updated in the ALPA tables [Arbitrated Loop Physical Address tables]

#### **SWITCHED FABRIC (FC-SW) :**

'Fabric' which consists of one or more dynamic switches connected via ISL (InterSwitch Link)

500 Meters, 2GB/sec, 2-to-24th ports per fabric

Multiple conversations can occur simultaneously over the SW Network

Multiple O/S can participate on Network by using Zoning to isolate ports

More hosts allowed

Any port can have conversation with another port

Limited to 16 switches total per fabric

FC-4 Type=SCSI Persistent Binding

Class 3=HBA/driver solution

### **GENERAL FIBRE CHANNEL FABRIC RULES:**

--all nodes should be within 3 hops of each other

--max of 16 switches in a single fabric

--mult. Equal costs paths recommended between all switches or (2) ISLs

--min. of (2) ISLs between two direct communicating switches

--Brocade "trunking" to consolidate bandwidth is recommended

--Principal core switch should be at the logical center of the fabric

### **WELL KNOWN SWITCH ADDRESSES:**

Broadcast: 0xFFFFFFF

Fabric Login Server: 0xFFFFFE

Name Server: 0xFFFFFC [All Nodes login and register with Name Server]

## **CLARIION HBA/ARRAY DEVICE DRIVERS FOR EMULEX:**

Clariion Arbitrated Loop=direct connect via arbitrated loop hub

Clariion Fabric for switched fabrics

## **EMC SUPPORTED TOPOLOGIES :**

Arbitrated Loop for Direct Connections

Switched Fabric

## **FIBRE CHANNEL OSI MODEL :**

Physical Layer—FC-0

Datalink Layer—FC-1 to FC-3

Transport/Networking Layer—ULP, FC-4

FC-4 for SCSI allows SCSI Initiator and Target to communicate over Fibre Channel using ;

FCP\_CMND, FCP\_XFER\_RDY, FCP\_DATA, FCIP\_RSP

## **FIBRE CHANNEL PORTS:**

Many types of Fibre Channel Ports, but in “Switched Fabric” [i.e., non-arbitrated loop!], we only use (3) types:

### **N\_Port:**

N\_Ports are hardware endpoints [Celerra HBA or Symmetrix FA] that connects to the FC topology. N\_Ports transmit and receive Fibre Channel frames in Switched Fabric environments.

### **F\_Port:**

N\_Ports connect to the FC Switch Ports, called F\_Ports, bringing that connection into the “Fabric”[Celerra or Symm to Switch]

### **E\_Port:**

Expansion port connecting two switches together—called ‘cascading’ switches [ISL’s—Inter-Switch-Links].

Total of 14 FA Ports per Symmetrix [7 FA’s, each having two ports]

**Login Process:** N\_Port host logs into fabric on F-Port switch

### **NL\_Port:**

Loop port used in ‘arbitrated loop’

### **FL\_Port:**

Used to connect fabric to Public Loop

## **FIBRE CHANNEL TERMINOLOGY:**

WWN—Unique 64-bit address broadcast by HBA on startup [Aliases are names given to WWNames]

Hops—Number of ISL’s a frame needs from time entering to exiting the F\_Port fabric

ISL—InterSwitch Link connects two E\_Ports on two different switches [aka “Cascading”]

Fan-In—Number of Host ports zoned into a single SYMM Fibre Channel FA[Director] Port

Fan-Out—Number of SYMM FA Ports consolidated into a single Host.

Zoning—Grouping together of a bunch of WWN’s or Aliases [ZoneSet—group of Zones]. Main purpose is to isolate each HBA from others in the fabric, called “Single HBA/Initiator Zoning.”

Hard Zoning—consists of WWN and physical Switch Port numbers—not the recommended approach

Soft Zoning—Recommended method, using WWNs and aliases instead of physical switch port numbers

Zone Configs—a collection of Zones to be activated simultaneously—can only have a single Active Zone Set at any given time

Full Mesh Topology—All switches interconnected with only a single hop between any two nodes [Dual ISLs recommended]

Core/Edge Topology—More common switch fabric, where “Core” switches are located in the center of the fabric in a “Mesh” configuration, while “Edge” switches are located at the edge of the fabric with single ISL

## **CONTROLLING ACCESS TO VOLUMES:**

Use VolumeLogix to manage what Hosts see what Volumes on the BackEnd by mapping HBA to FA to Volumes [LUN Masking].

## **THREE TYPES OF FC ZONING :**

Port Zoning—‘Hard Zoning’ of ports together at the Switch—not recommended

WWPN Zoning—more flexible Zoning allows for movement of physical cables from port-to-port [EMC zoning method]

Mixed Zoning—Combination of Port-to-Port and WWN Zoning

## **Fibre Channel Celerra Implementation:**

Fibre Channel connection to Symmetrix using SCSI protocol-over-Fibre-Channel via Class 3 Switched Fabric.

## **TARGET & LUN SUPPORT ON CELERRA:**

SYMM=max of 2048 devices

SCSI=240 LUNS

FIBRE=256 LUNS SYMM 5 [Note: This is a Symmetrix limitation. Celerra could otherwise support up to 4096 LUNs]

FIBRE=128 LUNS SYMM 4

Single Topology per Host

Multiple HBA's per Host of same type

### **FIBRE CHANNEL TOPOLOGIES:**

Public Loop: Shared bandwidth, connects to switch using FL\_Port

Private Loop : Shared bandwidth, not connected to a switch—isolated loop.

Point-to-Point Fabric : Dedicated bandwidth connection—N\_port to F\_port on switch

**SWITCHED FABRIC Class 3 [FC-SW]:** FC is a serial data transfer interface operating over copper and/or optical fiber at speeds up to 100MB/sec. Networking and I/O protocols [such as SCSI] are mapped to FC constructs, encapsulated, then transported within Fibre Channel frames. Supports SCSI, TCP/IP, and ESCON and Hosts : HP, Sun, SGI, NT, Sequent, IBM AIX, Celerra. Class 3 is SCSI 3 protocol [Upper Level Protocol—ULP] encapsulated over Fibre Channel without guaranteed delivery—no ACK. SCSI over Fibre is good for 500 Meters at 2Gb/sec. FC SW provides for dedicated bandwidth and virtual paths between ports.

### **DataMover Fibre Channel Specs: 506, 507, 510**

**EMULEX LP8000:** Emulex LP8000DC/N1; Two Ports Full Duplex; Multimode Cable with SC Connectors; FC5M-50M/FC10M-50M cables ; Supports 506/507/510 DMs running 2.1.24.4/2.2.15.4 NAS and SYMM 5265.43.26

**Note:** Emulex LP8000 only supports 1GB speed. Known problem with excessive emulex driver ‘jitter’ on the top port of the card, causing connectivity issues with McData 140M switches—for this issue, use only bottom port for Golden Eagle Control Stations.

**EMULEX LP9000:** 2GB Fibre HBA's for 510 DM's; IP Flag off by default; Current Bios levels 1.63a2 BIOS & 3.90a7 firmware

**EMULEX LP9002:** 2GB HBA's for 510 DMs ; Requires SC-LC or LC-LC cables ; LP9002 Board 250-735-900 for straight DM Host ; Board #250-734-902 for Tape Drive Support

**AGILENT HBA:** NS600, 700, & NSX Series

### **Control Station Fibre Channel Specs:**

**QLOGIC HBA:** Qlogic QLA-2212 Two Ports Full Duplex Multimode Cable with SC Connectors

**Prerequisites:** Linux O/S and BIN file to support FC-SW [CS0 should see all Data Volumes & GateKeepers]

**BIN File Flags:** PP, UWN, SCSI T & ARB bits set; SCSI C Flag cleared

When connected, orange or green light should illuminate—if not, reverse the orientation

**Note:** FC Control Stations were susceptible to rebooting after Zone changes due to a 30 timeout value. Value increased to 90 seconds with NAS 4.2.17.0, same timeout value as for FC Data Movers.

### **BOOTING FROM CLARIION ARRAYS:**

All CFS/CNS14 Control Stations should be zoned to a single HBA port only, using HBA0, to SPA. Fibre drivers do not support dual paths. Redundancy pathing is provided if CS1 is used, which is zoned to SPB.

### **Replacing Fibre Control Station:**

1. Replace hardware & find new WWN for CS0
2. Remove old hardware's WWN
3. Create replacement Zone on FC Switch for new Control Station
4. After Logon: #/nasmed/sbin/setup\_slot -init cs [#/nasmed getreason]
5. Failing Over CS0 Primary to CS1 from the Console of CS1 : **#/nasmed/sbin/cs\_standby -takeover**

### **FIBRE CHANNEL SWITCH SUPPORT MATRIX:**

#### **SWITCH TYPE:**

EMC Connectrix DS-8B or 16B

EMC Connectrix DS-16B2

EMC Connectrix ED-12000B, DS-16M

EMC Connectrix DS-16M2

EMC Connectrix DS-24M2

EMC Connectrix DS32B2, DS-32M

EMC Connectrix DS-32M2

EMC Connectrix ED-64

EMC Connectrix ED-64M

EMC Connectrix ED-140M, 1032 v 2.0

EMC Connectrix ED-48000B enterprise director (32-256 full duplex 4Gbit/sec ports)

EMC Connectrix DS-220B SAN switch (8-16 4Gbit/sec ports)

BROCADE SilkWorm 2400, 2800, 3800, 3900

BROCADE 12000

BROCADE DS-24M2, 32M2

#### **FIRMWARE LEVELS:**

2.5.1.b or later

3.0.2a or later

V2.00.00

V4.01.00 or later

V2.00.00

04.00.00

2.00.00 or later

4.01.00 or later

FUJITSU-SIEMENS PSFS-B161

MCDATA ED-5000, 6064

MCDATA ES-3016, 3032

MCDATA Intrepid 6140

MCDATA Spheron 4500

CISCO MDS 9216, 9509 (Multilayer Fabric Switches, 2Gbps and 1/2/4 or 10Gbps switches, respectively)

**Comments:**

All switches support Fan-Out Ratio of 6 :1 and Fan-In Ratio of 1 :6/Volume Logix 2.2.1/256 LUNS

Up to 16 switches can be cascaded [merged] together in a single ‘Fabric’

DS-8B & 16B are being discontinued

BROCADE=DS; MCDATA=ED

**Warning:** Celerra DM or CS cannot boot from port 0 on Brocade 3900 or 12000—BIOS issue [Fixed with later code]

**BROCADE DS-16B vs. McDATA ED-1032 SWITCHES :**

--Telnet or use GUI to manage Brocade (McData is GUI only)

--Default Zoning is to allow all ports to talk to all nodes (McData has all Zoning disabled by default)

--Brocade uses NL, G, U, or QL Ports (McData does not)

--McData Switches do not work with Arbitrated Loop topologies, only Switched Fabric

--McData Enterprise Director switches can generally perform ‘firmware’ upgrades while Online, Brocade DS cannot

**CISCO 9216/9509 FIBRE CHANNEL SWITCH ISSUES WITH CELERRA:**

NAS 5.2 installs or existing Celerras may experience a loss of FC connectivity, especially after firmware upgrade (Cisco 9120 sw 2.0(3)). Cisco switches have two admin modes: FX & AUTO. Switch ports may need to be set to FX emulation in order to work on firmware levels 1.2.1a and higher. See emc107196.

**CISCO SWITCH COMMANDS:**

# show tech-support (Fabric Manager map, write this to a file)

# show flogi (view HBA’s or SP’s logged in)

# show fcns (nameserver database and statistics)

# show interface brief (summary of interface and ports, as well as connected devices)

**PORT EMULATION SHOWS AUTO:**

#show interface brief

fc1/3 1 auto on notConnected swl -- --

**CHANGING PORT EMULATION TO FX:**

MDS9509# config t

MDS9509(config)# interface fc1/3

MDS9509(config-if)# switchport mode FX

MDS9509(config-if)# do show interface brief

---

Interface Vsan Admin Admin Status FCOT Oper Oper Port

Mode Trunk Mode Speed Channel

Mode (Gbps)

---

fc1/1 1 auto on down swl -- --

fc1/2 1 SD -- up swl SD 1 --

**fc1/3 1 FX** -- notConnected swl -- --

**FIBER MEDIA :**

**Single-Mode Fiber :**

Longest distances, uses 9 micrometer cable and LongWave Laser (LWL), 10km

**Multi-Mode Fiber :**

50 or 62.5 micron cables using ShortWave Laser (SWL), 500 or 300 meters, respectively

**Note :** Lucent ‘LC’ Connectors support either Fiber Cable Type

**FIBRE CABLE LENGTHS FOR CELERRA:**

3, 5, & 10 meters

**FIBRE MEDIA SPECS:**

|                          | <b>LongWaveLaser</b> | <b>ShortWaveLaser</b> | <b>ShortWaveLaser</b> |
|--------------------------|----------------------|-----------------------|-----------------------|
| <b>Transmission rate</b> | <b>9 micron</b>      | <b>50 micron</b>      | <b>62.5 micron</b>    |
| 100MB/s                  | 10,000 meters        | 500 meters            | 300 meters            |
| 200MB/s                  | 2,100 meters         | 300 meters            | 150 meters            |
| 400MB/s                  | 2,100 meters         | 175 meters            | 90 meters             |

## **FIBRE CHANNEL FRAMES :**

Total frame is 2148 bytes, with 2112 bytes of data

Header consists of S-ID (Source ID), D-ID (Destination ID), & Frame Type (08=SCSI over FCP Type)

Address field=3 bytes, consisting of Domain ID, Logical Port Number, and AL\_PA Loop ID (00=Brocade ; 13=McData)

McData : Domain ID's from 1-31 range

Brocade : Domain ID's from 1-239

## **PORT ADDRESS USAGE BY TOPOLOGY:**

Arbitrated Loop Private : Does not use Domain ID or Logical Port Number, just AL\_PA

Arbitrated Loop Public : Uses all three (Domain ID, Port Number, & AL\_PA)

Switched Fabric : Uses only Domain ID & Port Number

## **RESERVED PORT ADDRESSES IN FC :**

FFFFFE—Fabric Login Server

FFFFFC—Directory Name Server

## **WORLD WIDE NAMES (WWN) :**

128-bit Unique Identifier in a Loop or Fabric [64 bits identifies Node—WWNN; next 64 bits identifies Port—WWPN]

### **WWPN:**

When referring to WWN's, we most commonly mean the World Wide Port Name, which is a unique 64-bit value

Recommended Zoning using WWPN be used (as opposed to the WWNN, which can change)

**Two types of WWN's:** WWPortNames and WWNodeNames

**WWN SUN HOSTS :** Grep for WWPN in /var/adm/messages for EMULEX & /qla2200 for Qlogic

**SYMMETRIX :** Inlines E1,,0 and E1,,1 (See WWN Hi and Lo)

NT>Programs>Emulex Configuration Tools>Doubleclick LP7000

### **Composition of WWN :**

1st 4-bits identify format

Next 24-bits identify company or OUI

Last 36-bits reserved for company use

### **OUI's :**

006069 = Brocade

00<sup>E</sup>002 = Crossroads

006048 = EMC

0000C9 = EMULEX

## **CELLERRA BLADE WWN's—8 Byte field: Mfg ID; HBA Port; WWN Seed:**

# .server\_config server\_2 -v "fcip show"

FCP ONLINE HBA 0: ALPA 000001 WWN: 500601603b200c8c QE4

FCP scsi-0: HBA 0: ALPA 0000ef SP-a00: 500601603b200c4d Class 3

FCP ONLINE HBA 1: ALPA 000001 WWN: 500601613b200c8c QE4

FCP scsi-16: HBA 1: ALPA 0000ef SP-b00: 500601683b200c4d Class 3

# server\_log server\_2 -s |grep -i wwn

2009-03-23 12:47:45: FCP: 6: ONLINE HBA 0: ALPA 000001 WWN: 500601603b200c8c →Blade WWN HBA0

2009-03-23 12:47:50: FCP: 6: ONLINE HBA 1: ALPA 000001 WWN: 500601613b200c8c →Blade WWN HBA1

**500601603b200c8c**

→500601 represents Mfg ID

→60 represents HBA port, in this case HBA 0

→3b200c8c represents the WWN Seed, from enclosure\_0 resume prom

## **SWITCHED FABRIC LOGIN PROCESS :**

I. FLOGI—Fabric Login Table—each Host or Disk requires an FC ID. After HBA Link or Loop is initialized, Node Port transmits FLOGI frame to Well-Known Fabric Port "FFFFFE"—receives 'ACC' response if successful from Name Server. If a device logs in successfully, it will display in the FLOGI table.

### **Information Registered with FFFFFC Directory Service :**

S\_ID—Port Identifier ; PN—WWPort Name of N\_Port ; Class of Service—Class 3 ; FC-4 Types—SCSI-3 ; Port Type—N\_Port

II. PLOGI—Port Login. Node conducts 'N\_Port' login by transmitting PLOGI frame to destination N\_Port node—receives 'ACC' frame if successful. Process whereby a Port registers with Name Server and obtains attributes of other hosts.

### **Typical Port Login Exchange between Nodes :**

Host Address—S\_ID ; Frame Size—Buffer Size ; Flow Control & Version—TOVs ; Port Name--WWPN

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
III. PRLI—Process Login. SCSI 3 requires a process login so that SCSI protocol can BIND with FC protocol—Applications.

#### **MAPPING OF SCSI-3 to FCP :**

FCP\_CMND—Send SCSI Command service

FCP\_XFER\_RDY—transports offset & requests byte count objects of Send Data-In and Receive Data-Out Services

FCP\_DATA—Transports data object of Send Data-In and Receive Data-Out protocol

FCP\_RSP—Send Command Complete

**RSCN:** Registered State Change Notification is a FC Service that informs hosts about changes in the fabric (SCR), but each Hosts much query the Name Server to obtain the new information.

#### **TROUBLESHOOTING INLINES :**

A7,B or E7 : Check Volumes seen on FA Port

Run INQ query tool from Host

8F : Shows FA Status, Bit Settings, & Login History (Use command to determine if logged into switch)

E0 : Run to see if FA port sees light (not login)

FC,AAAA : Fibre Channel commands

#### **ISL Fibre Channel Limitations for Celerra:**

Celerra HBA's must be physically connected to the same switch as the Symmetrix FA's. This is a known Celerra limitation since the Celerra boots from devices on the Symmetrix and this is not presently negotiated successfully across the Interswitch Link [ISL] between the two fibre channel switches. The Celerra FA's cannot boot from Symmetrix if connected to multiple Fibre Channel Switches that are connected together using ISL. Up to (4) ISL's can be integrated together.

--If using ISL, use only in-family switches (i.e., brocade to brocade, etc)

--Zone switches using only WWN, not Port Zoning, which does not work well with ISL's

--Use only single switch-to-switch ISL link—multiple switches not supported

**Note:** The term “cascading” is commonly used to denote that an ISL is in use. The term “trunking” is generally applied to Brocade’s method of using multiple ISL’s as a single high-bandwidth trunk.

#### **Symptoms:**

--Celerra HBA's have valid Binding Tables and are logged into the fabric and Symmetrix FA's, yet upon booting up, cannot see Symmetrix devices. Performing setup\_slots does not correct this condition.

--Celerra might actually boot up to a "Contacted 5" state, yet in actuality, root filesystems are not mounted. When running the \$server\_devconfig -probe command, no Symmetrix devices are visible.

--Celerra will boot up successfully from Symmetrix if the Symmetrix FA is connected to the "same" Fibre Switch as the Celerra

--Celerra will not boot up or setup\_slot when the Celerra FBA is connected to (1) Switch and the Symmetrix FA is connected to a (2nd) ISL-linked Switch

#### **Fibre Channel Adapter/Switch Adapter :**

Connectrix SCB Spider—DP3-SCB1/DP3-SCQ1 [Microcode 5567.38.22; SYMAPI 4.3.1; NAS 2.2.49.2 +]

4 Port FA—DP3-FCD4/DP3-SCD44

**Latest 4-Port FA for Symmetrix :** DP3-FCD42G [2GB FA with minimum 5567 microcode]

#### **IBM LTO ULTRIUM TAPE DRIVE:**

**Problem:** \$server\_mt server\_3 -f tape1 status

Cannot format the device

**Cause:** I/O Check condition is occurring and sense information cannot be obtained from server\_mt command

**Fix:** Change ‘Host Type’ on Pathlight from ‘Generic’ to ‘autosense/NT’

#### **CELLERRA TIMEOUT PROBLEM:**

IBM Ultrium LTO2 Fibre Tape drives have timeout value of 17 minutes while Celerra has a timeout of 3 minutes. This can cause backup failures if we timeout too soon while waiting for tape device to reply to commands. NAS 5.1.20.4012 & higher will increase our timeout to (30) minutes.

**Symm Inlines :** FC,SWIT,A ; FC,FABR,A ; E1,,0 ; E1,,1 ; 8F,,,0 ; 8F,,,1

#### **Fibre Channel BIN File Director Flag Settings :**

PP, UWN, SCSI T & ARB, VCM [if using VL] are SELECTED; A, NP, TP, GVSA, VCM, C2S, and SCSI C are CLEARED

**BIN File :** FC (3) digits represent Virtual Bus, Virtual TID, and Virtual LUN. SCSI (2) digits represent Target ID and LUN.

#### **FAN-IN & FAN-OUT TOPOLOGY:**

Fan-In typically implies that there are (6) DataMover FBA's zoned with (2) Symmetrix FA's

A 32-port fibre switch can therefore support (24) datamover links to (4) Symmetrix links

Fan-Out implies that a single datamover can be zoned to see up to (6) symmetrix devices [3 symm's @2FA's each]

### **Fibre Channel DataMovers:**

Use only FC connection to Symmetrix!! SCSI Adapter is for Tape Drive support!!

HA—Each DM Fibre Port should be connected to the Fabric via a distinct data path, and then to different blades on the Fibre Switch—also, connections between Switch and SYMM FA should be on different switch blades. In otherwords, both ports should see same volumes but via different FA port and from different FA cards!

Identifying DataMover WWNames in Server Log: **\$server\_log server-x -a | grep WWN**

New 2GB Emulex LP9002 HBA's for 510 Datamovers will show WWN using **\$server\_sysconfig server\_x -pci**

### **Fibre Channel Zoning for DataMovers:**

--Single HBA Zoning—Single DM HBA port is zoned to one or more FA ports (i.e., multiple Symm FA's)—conversely, an FA may reside in multiple Zones

**Note:** One good reason for this is that it cuts down amount of time required for resets after Registered State Change Notifications (RSCN) & only ports within same Zone will have to log back into the Fabric. It also most closely resembles the 'single initiator' mode of SCSI.

--Both DM HBA Ports should see same Volumes but map via different FA Port and FA card on SYMM

--Keep each DM HBA port in separate Zone for load balancing

--Each DM FC HBA Port should be connected to Switches with low-numbered Ports (& each to different switches for HA)

--SYMM FA Ports should be connected to Switches with high-numbered Ports

--DM HBA Ports need to be zoned at switch for multiple SYMMS

-- Standby DataMover should be connected to a different Symmetrix FA Port

--DataMover & Control Station can share the same FA Port, but this is not recommended!!

### **How to Zone New Datamover FC HBA's Using Connectrix Switches:**

Step 1. If using Connectrix Switches, setup Port-to-Port Zoning for Datamovers to be upgraded using Connectrix Manager

Step 2. After Hardware install and setup\_slot, reconfigure Connectrix Zones for WWN

**Note:** This can all be done "on-line" and without any interruptions for a Connectrix Switch but not for a Brocade!!

### **FibreChannel Celerra Configuration:**

--Control Station must use Target & LUN 0/0. If a previous configuration of Volume Logix [aka, VCM] on the Symmetrix has used TID 0/0, then the VolumeLogix must be moved!

--Maximum of 14FA ports per Symmetrix [7 FA's with 2 ports each]

### **FibreChannel Failover on DM:**

--Upper port used for booting

--If either port is disconnected, other port will assume all traffic but could take up to 2 minutes for this transition to occur

--If both ports are disconnected, DMF will occur within 2 minutes

--Should never disconnect Fibre Cables from DMs online

**Comment:** Current NAS behavior is as follows: Loss of a single path will cause HBA to failover to alternate controller and port. Loss of both paths will initiate data mover failover to Standby server.

### **MIGRATING DATAMOVER FIBER CABLES FROM ONE SWITCH TO ANOTHER:**

1.) Match cables by name to respective Switch Ports

2.) Match WWNs on Celerra HBA's to respective Switch ports

3.) Create new Zonesets for new Switch Ports and HBA Ports

4.) Use Standby Server as test and move FC Cable to new Switch port

5.) Confirm connectivity to Symm down new path by using **server\_devconfig -p -s -a**

6.) Failover Primary Servers to Standby Server first, then move Cable paths on each Primary Slot

### **MOVING HBA1 FROM ONE FA TO ANOTHER :**

**Intro:** Moving Fibre HBA1 from Even FA to an Odd FA

1. Create new zoneset and activate

2. Servers need to reboot in order to update binding tables for new FA WWN

3. Do **server\_devconfig -probe** to ensure that DM can see symm devices down both FA's

4. Conduct **server\_devconfig -create** in order to update database files

5. Reboot server to allow kernel to register new volume information

### **REMOVING FA FROM DM AND ADDING A NEW FA:**

1. Document current configuration:

#/nas/sbin/log\_config →Monthly dial home script that produces output of Celerra configuration and information

#.server\_config ALL -v "fcp show" →Shows which channels are online

#.server\_config ALL -v "fcp bind show" →Shows current binding table  
#nas\_volume -i -a →Shows current information on Controll, Target, and LUNs  
#server\_devconfig -p -s -a & -l -s -a →Devices in NAS database & verifying Access Logix use  
#/nas/symcli/bin/symbcv list →Split any bcv's off first as seen from nas\_disk -l  
2. Stop CIFS on all DM's  
3. Permanently unmount all file systems on DM's [Save copy of mount file first]  
4. Clear binding tables: \$.server\_config ALL -v "fcp bind clear"  
5. Confirm by running: \$.server\_config ALL -v "fcp bind show"  
6. Halt all DM's: \$server\_cpu server\_x -h now  
7. Activate new zoneset for new FA  
8. Reboot data movers and confirm that new FA is being seen by running "fcp bind show"  
9. Create new chains in Celerra camdisk database by running \$server\_devconfig server\_x -c -s -a  
10. After verifying that all devices are visible down both paths, remount all File Systems and start CIFS  
11. Update SYMAPI database: #mv /nas/symapi/db/symapi\_db.bin & #/nas/sbin/nas\_rcf -localinit & #/nas/sbin/nas\_rdf -init

## **CELERRA WWN's:**

**\$ /nas/sbin/setup\_backend/zone -s 2 showdm** (command to collect WWN's for DMs)

Checking for Data Mover 2.

Collecting Data Mover WWNs...Done.

DM List:

DM Port:20 WWN:50:06:01:60:90:60:16:f8:50:06:01:60:10:60:16:f8  
DM Port:21 WWN:50:06:01:60:90:60:16:f8:50:06:01:61:10:60:16:f8  
DM Port:22 WWN:50:06:01:60:90:60:16:f8:50:06:01:62:10:60:16:f8  
DM Port:23 WWN:50:06:01:60:90:60:16:f8:50:06:01:63:10:60:16:f8

## **EXAMPLE OF 'PROBLEM' BINDING TABLE:**

**# .server\_config server\_2.faulted.server\_3 -v "fcp bind show"**

\*\*\* Persistent Binding Table \*\*\*

Chain 0002: WWN 50060482c0946954 HBA 0 FA-05ba Bind Pending →Problem—not bound to Fabric—no path to backend  
Chain 0018: WWN 50060482c094695b HBA 1 FA-12ba Bind Pending →Problem—not bound to Fabric—no path to backend  
Chain 0034: WWN 50060482c4698712 HBA 0 FA-03ba Bound →Only real path to backend, though Chain 0034 is not normal  
Chain 0050: WWN 50060482c469871d HBA 1 FA-14ba Bind Pending →Problem—not bound to Fabric—no path to backend  
\*\*\* Dynamic Binding Table \*\*\*

Chain 0002: WWN 0000000000000000 HBA 0 ID 0 Inx 00:81 Pid 0002 S\_ID 000000 Non  
Chain 0018: WWN 0000000000000000 HBA 1 ID 1 Inx 01:81 Pid 0018 S\_ID 000000 Non  
Chain 0034: WWN 50060482c4698712 HBA 0 ID 0 Inx 02:00 Pid 0034 S\_ID 691013 Sys  
Chain 0050: WWN 50060482c469871d HBA 1 ID 1 Inx 03:00 Pid 0050 S\_ID 6a1013 Sys

**Note:** Run "fcp show" to display WWNs for Data Mover HBAs

## **DEVCONFIG PROBE SHOWS SINGLE PATH ONLY:**

**chain= 34, scsi-34**

```
symm_id= 000185706012 celerra_id= 0001857060120563
tid/lun= 0/0 type= disk sz= 4153 val= 1 info= 556812563300
tid/lun= 0/1 type= disk sz= 4153 val= 2 info= 556812565300
tid/lun= 0/2 type= disk sz= 2076 val= -99 info= 556812567300 diskerr= unmarked
tid/lun= 0/3 type= disk sz= 2076 val= -99 info= 556812568300 diskerr= unmarked
tid/lun= 0/4 type= disk sz= 2076 val= -99 info= 556812569300 diskerr= unmarked
tid/lun= 0/5 type= disk sz= 2076 val= -99 info= 55681256A300 diskerr= unmarked
tid/lun= 0/15 type= disk sz= 3 val= -99 info= 55681256B300 diskerr= unmarked
tid/lun= 1/0 type= disk sz= 34856 val= 3 info= 556812533300
tid/lun= 1/1 type= disk sz= 34856 val= 4 info= 556812537300
```

**Note:** Problem here is that only a single path has true connectivity to the backend, as seen by Chain 0034 "Bound". Chains 0002, 0018 & 0050 show "Bind Pending" and no I/O can traverse any of these Chains. The other problem is that this is Fibre Channel, and the normal Chains that should be bound for dual FBA's are Chains 02 & 18, not 34 or higher. This Server's Binding Tables require immediate repair as there is no failover to the Backend, only a single path.

## **REBUILDING BINDING TABLES**

## **CLEARING BINDING TABLES & UPDATING CELERRA CAMDISK DATABASE:**

**Warning!** The following procedure should not be conducted without TS2 consult. Special care must be taken on systems with multiple storage backends, which may require temporary suspension of zones for the secondary storage systems, and other additional steps. Clearing the Binding Tables is potentially a very destructive action that could result in the entire loss of data on any file system mounted to the Server and should not be done on the “production” server’s slot—the primary means of clearing binding tables should be done by failing over the Server to its standby, then performing the binding table work on the unused slot.

### 1. Run “fcp bind show” on Production Server to document, then failover to Standby Server

### 2. Clear the Binding Table on the “unused” non-production faulted slot:

```
# .server_config server_2.faulted.server_3 -v "fcp bind clear 2"
```

**Note:** If you absolutely have to run on production slot and there are multiple backends, permanently unmount the production file systems before using this procedure. This action clears the persistent bind table.

### 3. Verify above action:

```
# .server_config server_2.faulted.server_3 -v "fcp bind show"
```

\*\*\* Persistent Binding Table \*\*\*

\*\*\* Dynamic Binding Table \*\*\*

```
Chain 0002: WWN 0000000000000000 HBA 0 ID 0 Inx 00:81 Pid 0002 S_ID 000000 Non
Chain 0018: WWN 0000000000000000 HBA 1 ID 1 Inx 01:81 Pid 0018 S_ID 000000 Non
Chain 0034: WWN 50060482c4698712 HBA 0 ID 0 Inx 02:00 Pid 0034 S_ID 691013 Sys
Chain 0050: WWN 50060482c469871d HBA 1 ID 1 Inx 03:00 Pid 0050 S_ID 6a1013 Sys
```

### 4. Conduct Setup Slot on Unused Slot:A

```
# /nas/sbin/setup_slot -init 2
```

Initializing server in slot 2 as server\_2.faulted.server\_3

Discover disks attached to server in slot 2 ...

server\_2.faulted.server\_3 : done

server\_2.faulted.server\_3 :

Warning 4017: server\_2.faulted.server\_3 : is out\_of\_service, use server\_standby  
to fix

Error 4008: server\_2.faulted.server\_3 : is restricted dart

**Note:** Do not be alarmed by the error messages—allow the setup\_slot to complete

### 5. Verify Binding Tables were rebuilt:

```
# .server_config server_2.faulted.server_3 -v "fcp bind show"
```

\*\*\* Persistent Binding Table \*\*\*

```
Chain 0002: WWN 50060482c4698712 HBA 0 FA-03ba Bound →Chain now correctly bound
```

```
Chain 0018: WWN 50060482c469871d HBA 1 FA-14ba Bound →Chain now correctly bound
```

\*\*\* Dynamic Binding Table \*\*\*

```
Chain 0002: WWN 50060482c4698712 HBA 0 ID 0 Inx 00:00 Pid 0002 S_ID 691013 Sys
```

```
Chain 0018: WWN 50060482c469871d HBA 1 ID 1 Inx 01:00 Pid 0018 S_ID 6a1013 Sys
```

### 6. Conduct Devconfig Probe:

```
# server_devconfig server_2.faulted.server_3 -p -s -a
```

server\_2.faulted.server\_3 :

SCSI devices :

chain= 0, scsi-0 : no devices on chain

chain= 1, scsi-1 : no devices on chain

**chain= 2, scsi-2**

symm\_id= 000185706012 celerra\_id= 0001857060120563

tid/lun= 0/0 type= disk sz= 4153 val= 1 info= 556812563190

tid/lun= 0/1 type= disk sz= 4153 val= 2 info= 556812565190

-----abridged-----

**chain= 18, scsi-18**

symm\_id= 000185706012 celerra\_id= 0001857060120563

tid/lun= 0/0 type= disk sz= 4153 val= 1 info= 556812563300

tid/lun= 0/1 type= disk sz= 4153 val= 2 info= 556812565300

### 7. Conduct Devconfig –create to update Camdisk File (if required):

```
# cat camdisk
```

1:c34t0l0+556812563190+,c50t0l0+556812563300+:

2:c34t0l1+556812565190+,c50t0l1+556812565300+:

3:c34t1l0+556812533190+,c50t1l0+556812533300+:

**Note:** Camdisk may not always be completely updated after setup\_slot!

```
# server_devconfig server_2.faulted.server_3 -c -s -a
```

server\_2.faulted.server\_3 : done

**#cat camdisk**

```
1:c2t0l0+556812563190+,c18t0l0+556812563300+:  
2:c2t0l1+556812565190+,c18t0l1+556812565300+:
```

**8. Reboot Data Mover to refresh memory with correct device list, and verify Binding Tables and Probe again**

**9. Fail production Server back to home slot**

**ALTERNATE PROCEDURE TO CORRECT TRESPASSED LUNS AND UPDATE CAMDISK FILES:**

**Situation:** Use this procedure if the binding tables do not need to be rebuilt—correcting for trespassed luns. Situation is that camdisk files do not show all (4) paths to backend, though when probe is conducted, all (4) paths are visible. Previous attempts to trespass luns back to Owner SP have failed.

**Step 1 Verify Trespassed Luns:**

**# /nas/sbin/navicli -h 10.241.168.57 getlun -tresspass -status**

**Note:** Best way to see if there are trespassed luns—will show all trespassed luns, regardless of which SP used to run cmd

**# /nas/sbin/navicli -h 10.241.168.57 getlun 5 legrep "Default|Current"** (checking status on single lun)

Current owner: SP A

Default Owner: SP A

**#/nas/sbin/navicli -h <ip of spa> getlun -default -owner**

**Note:** Try to fail luns back to correct SP Owner—if that fails, then proceed with step 3

**Step 2 Verify Chains using Probe for each Data Mover:**

**#server\_devconfig server\_x -p -s -a**

**Step 3 Temporarily Unmount all Production File Systems and Verify:**

**#server\_umount ALL -a**

**#server\_mount ALL**

**Step 4 Trespass each Lun to Respective SP Owner:**

**# /nas/sbin/navicli -h <ip of spx> trespass lun <lun#>**

**Step 5 Confirm that LUNs remain on Owner:**

**# /nas/sbin/navicli -h <ip of spa> getlun -default -owner**

**# /nas/sbin/navicli -h 10.241.168.57 getlun -tresspass -status**

**Step 6 Conduct Devconfig -create to Update Camdisk:**

**#server\_devconfig server\_x -c -s -a**

**Step 7 Verify Camdisk files and Probe matches:**

**#ls -la /nas/server/slot\_x/camdisk\***

**#server\_devconfig server\_x -p -s -a**

**Step 8 Remount all File Systems:**

**#server\_mount ALL -a**

**USING TS2 NAVILUN14.PL BACKEND CHECK SCRIPT (SangDon Shin):**

→Checks status of Backend luns

→Detects any trespassed luns and can fail them back if desired

→Displays Celerra HBA login status from SP perspective

→Displays fs disk & RG information and status of physical disks

→Checks validity of HBA-to-Storage Group relationship

**CLEARING & REBINDING FCP TABLE:**

**# .server\_config server\_2 -v "fcp bind show"**

**# .server\_config server\_2 -v "fcp bind clear"** →clears persistent bind table

**# .server\_config server\_2 -v "fcp bind rebind"** →Recreates persistent bind table

**Note:** Reboot Data Mover

**# .server\_config server\_2 -v "fcp show"**

**# .server\_config server\_2 -v "fcp bind show"**

**# .server\_config server\_2 -v "camshowconfig" (Luns 0-6 on Chains 00)**

**Other Commands:**

**# .server\_config server\_2 -v "fcp bind showbackup 2"**

**# .server\_config server\_2 -v "fcp bind restore"**

**# .server\_config server\_2 -v "fcp rediscover=Y"** [where Y = HBA0 or HBA1]

```
# .server_config server_2 -v "fcportreset=2" [forcing HBA2 to log back into fabric]
# .server_config server_2 -v "fc init"
# .server_config server_2 -v "fc rescans"
```

### **AVAILABLE FIBER CHANNEL COMMANDS IN NAS 5.5:**

**Caution:** These are not trivial options—do not experiment on customer systems!

```
# .server_config server_2 -v "help fc"
```

**Usage:**

'fc' options are: bind ..., flags, locate, nsshow, portreset=n, rediscover=n rescan, reset, show, status=n, topology, version  
 'fc bind' options are: clear=n, read, rebind, restore=n, show showbackup=n, write

**Description:**

Commands for 'fc' operations:  
 fc bind <cmd> ..... Further fibre channel binding commands  
 fc flags ..... Show online flags info  
 fc locate ..... Show ScsiBus and port info  
 fc nsshow ..... Show nameserver info  
 fc portreset=n ..... Reset fibre port n  
 fc rediscover=n ..... Force fabric discovery process on port n Bounces the link, but does not reset the port  
 fc rescan ..... Force a rescan of all LUNS  
 fc reset ..... Reset all fibre ports  
 fc show ..... Show fibre info  
 fc status=n ..... Show link status for port n  
 fc topology ..... Show fabric topology info  
 fc version ..... Show firmware, driver and BIOS version

**Commands for 'fc bind' operations:**

fc bind clear=n ..... Clear the binding table in slot n  
 fc bind read ..... Read the binding table  
 fc bind rebind ..... Force the binding thread to run  
 fc bind restore=n ..... Restore the binding table in slot n  
 fc bind show ..... Show binding table info  
 fc bind showbackup=n .. Show Backup binding table info in slot n  
 fc bind write ..... Write the binding table

### **VIEWING FIBRE CHANNEL STATISTICS ON CELERRA HBA's:**

1. Set FCP Stats on HBA's:

```
$ .server_config server_5 -v "printstats fc reset"
```

FCP stats Reset and Enabled

2. Review statistics:

```
$ .server_config server_5 -v "printstats fc"
```

Total I/O Cmds: +0%-----25%-----50%-----75%-----100%+ Total 10

|                                               |    |
|-----------------------------------------------|----|
| FCP HBA 0  XXXXXXXXXXXXXXXXXXXXXXXXXXXX  100% | 10 |
| FCP HBA 1   0% 0                              |    |
| +-----+                                       |    |

```
$ .server_config server_5 -v "printstats fc full"
```

Total I/O Cmds: +0%-----25%-----50%-----75%-----100%+ Total 15

|                                               |    |
|-----------------------------------------------|----|
| FCP HBA 0  XXXXXXXXXXXXXXXXXXXXXXXXXXXX  100% | 15 |
| FCP HBA 1   0% 0                              |    |

# Read Cmds: +0%-----25%-----50%-----75%-----100%+ Total 0

|                  |  |
|------------------|--|
| FCP HBA 0   0% 0 |  |
| FCP HBA 1   0% 0 |  |

# Writes Cmds: +0%-----25%-----50%-----75%-----100%+ Total 15

|                                               |    |
|-----------------------------------------------|----|
| FCP HBA 0  XXXXXXXXXXXXXXXXXXXXXXXXXXXX  100% | 15 |
| FCP HBA 1   0% 0                              |    |

Total # Blocks: +0%-----25%-----50%-----75%-----100%+ Total 240

|                                               |     |
|-----------------------------------------------|-----|
| FCP HBA 0  XXXXXXXXXXXXXXXXXXXXXXXXXXXX  100% | 240 |
| FCP HBA 1   0% 0                              |     |

# Read Blocks: +0%-----25%-----50%-----75%-----100%+ Total 0

|                  |  |
|------------------|--|
| FCP HBA 0   0% 0 |  |
| FCP HBA 1   0% 0 |  |

```
# Write Blocks: +0%-----25%-----50%-----75%-----100%+ Total 240
  FCP HBA 0 |XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX| 100% 240
  FCP HBA 1 |                               | 0% 0
  +-----+
  +-----+
```

**Note:** Available 5.4 only

### **MOVING DM FROM ONE SWITCH TO ANOTHER:**

1. Clear persistent binding table on Standby Server
2. Failover Production Server to Standby
2. Clear binding tables on original Server slot [faulted]
3. Halt faulted DM and move Switch connections
4. Activate new Zone set on switches
5. Reboot faulted DM & confirm binding tables
6. Run server\_devconfig on faulted mover
7. Failback server

### **CAUSES FOR DATA MOVER CHANNELS TO GO OFFLINE, REQUIRING PORT RESET:**

--reboot of FC Switch [Brocade firmware upgrade requires reboot of switch]  
--removal of Zone to tape library may not stop probing of devices until reset is done  
--zoning to remove a path, then adding back may require port reset  
--firmware upgrades on Connectrix  
--cable was out during DM boot

### **HOW TO CONDUCT FCP PORT RESET ON DATA MOVER HBA's:**

1. Verify HBA Paths that should be available:

**# .server\_config server\_2 -v "fc bind show"**

\*\*\* Persistent Binding Table \*\*\*

Chain 0000: WWN 50060167006007c0 HBA 0 SP-a7 Bound

Chain 0016: WWN 5006016f006007c0 HBA 1 SP-b7 Bound

2. Verify IO counts by running the following command several times:

**# .server\_config server\_2 -v "printstats scsi"**

Controller: IO-pending Max-IO IO-total Idle(ms) Busy(ms) Busy(%)

|  | 0:  | 17 | 217611 | 342359404 | 614656 | 0% |
|--|-----|----|--------|-----------|--------|----|
|  | 16: | 16 | 100016 | 342931080 | 38781  | 0% |

3. Issue Port Reset on affected HBA:

**# .server\_config server\_2 -v "fc portreset=0"**

1142623373: DRIVERS: 4: FCDMLT 0 [2.4.1] Scsi Port Bus Reset

1142623373: DRIVERS: 4: FCDMLT 0 [2.4.1] TPM Notify: st=0xa000002, flg=0x200, cmd=0x1

1142623373: DRIVERS: 4: FCDMLT 0 [2.4.1] Auto Neg Speed : In sync fmstat=0x80000000

1142623373: DRIVERS: 4: FCDMLT 0 [2.4.1] Auto Neg Speed : Current Speed: 2Gbps

4. Verify IO Total increases on affected path:

### **VIEWING PERSISTENT BINDING TABLE INFORMATION:**

**#/nas/sbin/workpart -r**

**NAS 5.3.18 Version:**

**# /nas/sbin/workpart -r**

Read Work Partition layout from LBA 0x43000, 528384 bytes.

After read:

Dump of Work Partition Structure (size 528384 bytes):

workpart\_magic = **0xfacebeaf** →Unique signature for NAS 5.3

**NAS Version 5.4.20:**

**# /nas/sbin/workpart -r**

Read Work Partition layout from LBA 0x43000, 528384 bytes.

After read:

Dump of Work Partition Structure (size 528384 bytes):

workpart\_magic = **0xbeaf0001** →Unique signature for NAS 5.4

### **DETERMINING IF WORKPART HAS BEEN UPGRADED TO NAS 5.4/5.5 OR HIGHER:**

**# /nas/sbin/workpart -c** (this check command is only avail. With 5.4+)

Newer Version Workpart Layout Found!

134

# /nas/sbin/rootnas\_volume -s root\_rdf\_channel (systems prior to 5.4 are 1 MB in size)

total = 2 avail = 2 used = 0 ( 0% ) (sizes in MB)

## **VolumeLogix:**

**Set IMPL flag to A080 if using Volume Logix with Celerra** (Normally set to 000 if no Celerra present).

**Purpose:** Once turned on, Volume Logix allows access to only those volumes behind FA ports that are configured to be accessed by a Host. Volume Logix is a combination of Symm Microcode and host-based software (NT or Sun) used to control Volume access in Fibre Channel environments. VolumeLogix issues a SCSI “not ready” to Hosts that are not “assigned” or granted “access” to volumes in the VCM database. WWN’s of each Host HBA is put into a Table on the Symm FA Port. VCMDB is constructed from WWN’s and volumes behind the FA’s.

### **SYMMETRIX:**

Symm 4, microcode 5265.49.31, 16-cylinder Gatekeeper which stores the VCMDB

Volume Logix flag can be turned on each individual FA port

**BIN File Settings:** Set Director port bits: T, ARB, UWN, PP, and VCM if using VolumeLogix

### **IMPL Flag=A080**

Use “SO” Host Emulation if using VolumeLogix in Mixed Host environment!

Use “S” Host Emulation if using only Celerra Hosts

Set VCM GateKeeper volume Target & Lun “FE” [NAS Versions 2.2.25.5 & Lower ONLY]

Set VCM GateKeeper volume Target & Lun “OE” [NAS Versions 2.2.25.5 & Higher]

**Note:** DataMovers and Control Station can share same FA port if using Volume Logix

**Server Devconfig Output:** Look for “SZ=7” for VCM Gatekeeper [7MB=16 cylinders]

**Fibre Channel Symmetrix Flags for BIN File Setup of Celerra:** FBA Host=SERVER\_CELERRA\*; FBA IMPL Flags=A000; FBA Multi-Access = Yes; PP=Selected [Point-to-Point enabled]; UWN=Selected; SCSI Flags “T” and “ARB”=Selected; SCSI Flag “C”=Cleared; VCM=Selected if using Volume Logix; A, H, V, NP, TP, GVSA, C2S Flags=Cleared; [\*FBA Host=Mixed if a Sun Host and Celerra will be connected to the Symmetrix]

### **Fibre Channel Media & HBA Adapters:**

Emulex LP-8000DC HBA’s [Dual ports per adapter; Full Duplex 100MB/sec; Multimode cabling using SC Connectors]

Fiber optic cable; 50/125mm and 62.5mm, with distances of 500 & 300 meters between ports

**Note:** Yellow or Green light should activate when connecting HBA port [N\_Port] to Switch port [E\_Port]

### **Fibre Channel Switches: “Class 3 Switching”**

Brocade SilkWorm 2800 [aka, Connectrix DS-16B]; 16 Ports; 16 ZoneSets [Each ZoneSet can have up to 256 Zones]

### **Connectrix ED-1032 Cabinet:**

Houses up to (2) ED-1032 Switches; Version 2.0 for each switch=32 ports; 64 ZoneSets

[256 Zones/set]; Version 1.0 has capacity of 16 zone sets for 31 zones per set

**Note:** Connect Celerra or Host connections to low-numbered ports and Storage-side connections to high-numbered ports  
Connectrix Switch has two WWN’s—Port WWN=Connectrix Port; Attached Port WWN=Attached Celerra or Symm WWN  
McData ES-5000

### **Fibre Channel Port Zoning:**

1. Use Symmetrix Volume Management [bin file map] to control FA ports; setup special FC flags on Symm
2. Use Connectrix or DS-16B switches to configure initial Connectrix Zoning
3. Symm should have VCM 16-cyl Gatekeeper Device configured for “Volume Logix” volume manager--must be visible to all FA ports Volume Logix would be used on a Unix Client to manage the Zoning; also can do Zoning from Switch
4. Symm FA Config: WWNaming; Point-to-Point for SW
5. Volume Logix: Software used to control access of Fibre Channel fabric date & Symm FA port devices. All FA ports should see Volume Logix gatekeeper

### **Fibre Channel Zoning Rules:**

Zone by WWN of Celerra HBA or WWN or Symmetrix FA Port

Each Celerra HBA Port should reside within its own Zone, that is, no other HBA’s in the same zone

A Symmetrix FA Port can reside in Multiple Zones, making it available to Multiple DataMovers

Both Ports on a Celerra HBA must see same Symmetrix volumes but be connected to different Symmetrix FA Ports for HA Standby DM HBA Boot Port should be on different FA Port than Primary DM HBA Boot Port

Fan-In Ratio: Up to (6) HBA Ports per Symm FA Port allowed; i.e., 6 HBA Adapters per Zone, different DataMovers

## **FIBRE CHANNEL TOPOLOGY ON GATEWAY:**

**\$ .server\_config server\_2 -v "fc topology"**

FCP HBA 0: F\_PORT D\_ID FFFFFC Node 100000051e0419e4 Port 100000051e0419e4 →D\_ID WWNs for Clariion

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!

FCP HBA 0: N\_PORT S\_ID 890900 Node 50060160b9a00e09 Port 5006016039a00e09 →S\_ID WWNs for Celerra  
FCP HBA 0: scsi-000 D\_ID 8a1c00 Node 5006016090600510 Port 5006016210600510  
FCP HBA 0: scsi-016 D\_ID 8a1b00 Node 5006016090600510 Port 5006016a10600510  
FCP HBA 1: F\_PORT D\_ID FFFF FC Node 100000051e0419e4 Port 100000051e0419e4  
FCP HBA 1: N\_PORT S\_ID 890a00 Node 50060160b9a00e09 Port 5006016139a00e09  
FCP HBA 1: scsi-032 D\_ID 8a1900 Node 5006016090600510 Port 5006016b10600510  
FCP HBA 1: scsi-048 D\_ID 8a1a00 Node 5006016090600510 Port 5006016310600510  
FCP HBA 2: OFFLINE ALPA 000001 Node 50060160b9a00e09 Port 5006016239a00e09  
FCP HBA 3: OFFLINE ALPA 000001 Node 50060160b9a00e09 Port 5006016339a00e09

### **FIBRE CHANNEL TOPOLOGY ON INTEGRATED:**

**# .server\_config server\_2 -v "fc topology"**

FCP HBA 0: FC-AL ALPA 000001 Node 50060160c1e05335 Port 5006016041e05335  
FCP HBA 0: scsi-000 ALPA 0000ef Node 50060160c1e043a3 Port 5006016241e043a3  
FCP HBA 1: FC-AL ALPA 000001 Node 50060160c1e05335 Port 5006016141e05335  
FCP HBA 1: scsi-016 ALPA 0000ef Node 50060160c1e043a3 Port 5006016a41e043a3  
FCP HBA 2: OFFLINE ALPA 000001 Node 50060160c1e05335 Port 5006016241e05335  
FCP HBA 3: OFFLINE ALPA 000001 Node 50060160c1e05335 Port 5006016341e05335

### **VIEWING DM BINDING TABLES:**

**\$ .server\_config server\_6 -v "fc bind show"**

\*\*\* Persistent Binding Table \*\*\*

Chain 0002: WWN 50060482bfd1fb0c HBA 0 FA-13a

Chain 0018: WWN 5006048000000000 HBA 1 FA-01a

\*\*\* Dynamic Binding Table \*\*\*

Chain 0002: WWN 50060482bfd1fb0c HBA 0 ID 0 Inx 00:00 Pid 0002 S\_ID 211413 Sys

Chain 0018: WWN 0000000000000000 HBA 1 ID 1 Inx 01:81 Pid 0018 S\_ID 000000 Non

FCP Base Chain: 2, Dump Chain: 2, Dump Slot: 6

Adapter Chain Offset 0:2:2 1:18:18

Existing CRC: 8a86935d, Actual: 8a86935d, CRC Matchs

**Commentary:** Above example shows an inconsistent Binding Table--evidenced by multiple zeroes. In this example, it may be that someone rezoned the Connectrix for Slot\_6, or disconnected the Fibre Port, etc. Any zone changes made to the Celerra that changes Port Channel assignments are reflected immediately in the Binding Table. Datamovers conduct "Persistent & Dynamic Bindings".

### **INTERPRETING OUTPUT OF "fc bind show":**

S\_ID refers to the Fibre Zone & Switch ID: 01=Zone 00=Brocade [011f00 Sys]

"Sys" means that this port is connected to a LUN0 (usually system or boot device), while "Non" would indicate a non-system volume

**# cat slog3.jan28 |grep -i wwn**

2005-01-27 08:50:11: CAM: 4: FCP ONLINE HBA 0: S\_ID 690913 WWN: 10000000c9394790 →Data Mover WWN HBA 0

2005-01-27 08:50:17: CAM: 4: FCP ONLINE HBA 1: S\_ID 610913 WWN: 10000000c9394791 →Data Mover WWN HBA 1

**Note:** The "13" represents a McData switch

2005-05-10 16:30:06: FCP: 4: ONLINE HBA 0: ALPA 000001 WWN: 5006016810601550 (direct-connect)

2005-05-10 16:30:10: FCP: 4: ONLINE HBA 1: ALPA 000001 WWN: 5006016910601550 (direct-connect)

### **RECOMMENDED CONNECTRIX ZONING WITH MULTIPLE SWITCHES:**

--All Boot ports should be connected to one switch

--All non-boot ports should be connected to the other switch

**LINUX CONTROL STATION:** Linux Fibre Channel does only "Dynamic Binding Table" updates. Also, Slot\_2's GateKeeper is the default gatekeeper that the Control Station uses. If slot\_2 is not working, will go to next GateKeeper on list for /nas/dev

### **DETERMINING WWN FOR BOTH CELERRA & SYMMETRIX HBA:**

**1. .server\_config server\_3 -v "fc show"**

FCP ONLINE HBA 0: ALPA 000001 WWN: 5006016810601550 DX2 [DataMover WWN S\_ID]

FCP scsi-0: HBA 0: ALPA 0000ef SP-a1: 500601611060172f Class 3 [Symmetrix WWN D\_ID]

FCP ONLINE HBA 1: ALPA 000001 WWN: 5006016910601550 DX2 [DataMover WWN S\_ID]

FCP scsi-16: HBA 1: ALPA 0000ef SP-b1: 500601691060172f Class 3 [Symmetrix WWN D\_ID]

FCP ONLINE HBA 0: S\_ID 011600 WWN: 10000000c9234711 LP8000

FCP scsi-2: HBA 0: D\_ID 011f00 FA-13b: 50060482bfd1fad Class 3

FCP ONLINE HBA 1: S\_ID 021600 WWN: 10000000c923470e LP8000

FCP scsi-18: HBA 1: D\_ID 021000 FA-14a: 50060482bfd1fad Class 3

1015516486: ADMIN: 4: Command succeeded: fc show

**Note:** This command will also indicate whether HBA is ONLINE or OFFLINE

ONLINE means that the HBA is logged into the Switched FC Fabric & has a physical connection

## 2. \$ server\_log server\_2 -a -s |grep -i wwn

```
2005-12-23 21:19:05: FCP: 4: ONLINE HBA 0: S_ID 614113 WWN: 10000000c92e6cb1 (DM HBA0 WWN)  
2005-12-23 21:19:12: FCP: 4: ONLINE HBA 1: S_ID 613213 WWN: 10000000c92e6cb2
```

## **CORRECTING INCONSISTENT BINDING TABLES FOR CELERRA FIBRE HBA's:**

**Caution:** Do not run fcp bind clear or init if multiple backends are involved & never without TS2 advice!

Step 1. Run the "fcp bind show" on the Server that is being rezoned or moved from one FA Port to another & record the info

Step 2. \$ls /nas/dev [Record GateKeeper info; shows all GK's in use by SLOT #--pay attention to the Chain being moved or rezoned--this will be removed later]

```
c18t0l15s2 c18t0l15s5 c2t0l15s3 c2t0l15s5 c18t0l15s6  
c18t0l15s3 c2t0l15s2 c2t0l15s4 c2t0l15s6
```

Step 3. Conduct the actual Zone changes on the Connectrix, etc.

Step 4. Reboot DataMover\_6 [in this example]

Step 5. Verify Binding Table ["fcp bind show"]

Step 6. Probe to create new devices for Slot\_6: \$server\_devconfig server\_6 -p -s -a | -c -s -a

Step 7. #cd /nas/dev & remove the GateKeeper for Slot\_6: **c18t0l15s6**

```
$more /nas/server/slot_6/scsidevs [1:gk21:0:c2t0l15:55665100A290:000184502251:]
```

Step 8. Update the /nas/sympapi/db database by running: #/nas/sbin/nas\_rdf -localinit

## **HOW TO DETERMINE IF MULTIPLE SYMMETRICES ARE ATTACHED:**

**/nas/bin/nas\_symm -l      or      \$nas\_disk -l**

### **Miscellaneous Fibre Channel Notes:**

Point-to-point Serial Data Transfer Interface using Optical Fiber or Copper Wire @100MB/sec.

**Support:** Switched Fabric=Switches (No Arbitrated Loop support!!) FC Class 3 service, connectionless, using N, F, and E\_Ports

**Note:** Because our FC method does not support arbitrated loop, we plug into a Connectrix or DS-16b Switch from a Datamover, which then goes to the Symmetrix. Cannot plug a Datamover Fibre channel connection directly into the SYMM.

### **Tracing Fibre Channel Path from DataMover to Symmetrix:**

| Name     | ControllerID | DMHBA  | DMWWN      | FAPort | FAWWN  |
|----------|--------------|--------|------------|--------|--------|
| Server_2 | >c2          | >hba-0 | >.....cd35 | >3a    | >7b42> |
| Server_2 | >c18         | >hba-1 | >.....cd7a | >14b   | >7b5d> |

### **FIBRE CHANNEL SOFTWARE SUPPORT:**

**SYMM:** 5265.43.26

**DATA MOVERS:** NAS 2.1.15.6+ and 506+ DataMovers

**CONTROL STATIONS:** NAS 2.2.25.5+

Dual, Full Duplex Emulex LP-8000DC dual HBA Ports [100Mbps throughput] using Multimode cable with SC connectors

Fan Out ratio of 6:1 [6 DM ports to 1 Symm FA port]

**Target & LUN Support:** Total of 540 targets can be visible on a Symmetrix, but only 128 from a single FA port.

**Extended LUN Support [FFF]:** DM Vol.Map Bits increase to support 4k targets=FFF, as opposed to FF for SCSI.

**Zoning on Brocade [DS-16B] Switch:** >switchShow; >cfgClear; >cfgCreate "celerra"; >aliCreate "cst9" [aliases for DM, and for Symm]; >zoneCreate "cstdm9"; >cfgEnable "celerra" >cfgShow [displays zoning information]

**Brocade DS-16B Switch Rules:** Max of (4) may be combined into one fabric; Max of (1) hop between any two switches; Max of (4) ISL's between two switches; Min of (2) ISL's between switches; FanOut 6:1; Max 0-8 E\_Ports per single switch [ie, switch to switch]

### **Fibre Channel HBA Adapter:**

Celerra 'discovers' FC devices on startup--N\_Ports query the Name Server for list of logged-in fabric devices, then registers them.

A Fibre Channel DataMover can address up to 4096 SCSI Addresses

## **CELERRA BOOT PORT FOR FIBRE FBA ADAPTER:**

*Top port of the Emulex is the Boot Port, Port A*

*Bottom port is the Data port, Port B*

**Pre 2.1.15.20 Code: FC Boot Port:** If Port A fails, DataMover must failover to a Standby unit!!! [If you unplug Port A, same thing]

Code 2.1.24.4: Both Top Port [C0] and Bottom Port [C1] are used as Boot and Data Ports!

**Capacity [Fan-out] v. Consolidation [Fan-in]:** Fan-Out means that a single Celerra can access multiple Symmetrix devices.

Fan-In means that a single Symmetrix FA port is shared by up to (6) DataMover connections as a 'shared' port; 6:1 ratio

## **LOCATING WWN FOR FIBRE CHANNEL CS:**

### **#dmesg |grep qla0**

scsi-qla0-adapter-node=200100e08b24dc0e;  
scsi-qla0-adapter-port=210100e08b24dc0e;

scsi-qla0-target-0=50060482c094decc;

**Note:** Above output is for Que Logix Adapter

## **CONNECTRIX--Fibre Channel Switches [EDM-1032]:**

### **ENTERPRISE DIRECTOR CABINETS:**

EC-1100: Houses up to (2) ED-1032 Switches; 10BaseT Ethernet Hub for Management; ZipDrive BackUps; MultiTech Modem

EC-1200: (2) ED-1032 or (12) DS-32/16M switches

ED-1230: Up to (3) ED-140M Directors & Connectrix SP

--Switched Fabric can use all (32) Connectrix ports in Full-Duplex mode, 100MB/sec, for aggregate bandwidth of 3.2 GB/second!

--4-ports/card GSM/GXX; 9 micron Single Mode 20km GXX Port Card; 50 micron Multimode 500m; 62.5 micron Multimode 325m

### **ENTERPRISE CONNECTRIX MANAGER (ECM):**

--Latest Connectrix Mgr SW: 7.0.x

--Latest McData Firmware: 5.0.1

--Runs on Laptop Service Processor but can be accessed remotely from Connectrix Client on Windows

--Most Management should be done via this interface, not Embedded Web Manager

--Normal operation is to have an IP configured for the ECM Laptop and for each Switch in the Cabinet

--Manage up to 48 Switches

## **CONFIGURING THE BROCADE DS-16B2 SWITCH:**

1. Connect to serial port from Windows Host with Hyperterminal session:

Setup HyperTerm as COM1 with 9600; 8; 1; None; None

2. At prompt, login: switch2a:admin>admin password

3. Set IP Address of Switch: admin>ipAddrSet

4. Disable Switch: admin>switchDisable

5. Configure Domain ID for Switch: admin>configure [Set Domain ID]

6. Verify license: admin>licenseShow

7. Reboot switch: admin>fastboot

8. Open Web Browser and enter IP Address of switch

9. Create Blocking Zone on Switch prior to plugging Hosts into Switch

Zone Administration>Create Zone called ‘Blocking Zone’ & add any port to the zone

Click on ‘Config Settings’ and create ‘Blocking Config’ & add ‘Blocking Zone’ to config

10. Actions>‘Save Config’ and then ‘Enable Config’

11. Create Aliases for each port and move respective port into the Alias

12. Create Zones and add alias members to proper zones

13. Create Config and add Zones to config [Save and then Enable Config]

## **SETTING UP NEW ZONES & ZONE SET:**

1. EMC Connectrix Manager>Fabrics>configure>Zoning Library [Contains all Zones created across all Switches managed by ECM]

>Zones>File>New>Zone Name: A1-Sun1-1pfco

2. Add HBA WWN of Host HBA and FA WWN of Symmetrix

--Repeat Steps above to create a Zone for each HBA-to-FA path

3. ZoneSets>File>New>Lab1>Add the Zones created previously to the new ZoneSet

**Note:** Create maximum of 1024 zones per zone set

### **ED-140M:**

140-port 2Gbps Switch; Online firmware upgrades; up to (3) ED-140M Directors per ED-1230 Cabinet

CTP Card: Control Processor Card contains System Service Processor and Embedded Port subsystems. SSP runs O/S, Applications, Communicates with Ports, Controls Serial & Ethernet Management Ports, Stores Config & Parameters in NVRAM. IML Button on the CTP, when held for 3 seconds, reloads firmware on Directory and resets both CTP cards without powering down or affecting operational Fibre Links—however, network connections are dropped & clients will have to log back into Fabric.

Port Map: 0-127

### **ED-64M ENTERPRISE FIBRE CHANNEL DIRECTOR:**

--64-ports, 4-ports per card; Uses Multimode 50 or 62.5 micron cables using Full Duplex 100MB/second speeds per port

--IML—Resets the CTP’s (Control Processors of the Switch)

### **DS-16/32M DEPARTMENTAL DIRECTORS:**

--16 or 32 Ports, Multimode 50 or 62.5 micron, Full Duplex 100MB/port

--Small Form Factor Optics (SFFO): Short & Long Wave cables using LC Connectors

--SFFO Card has green LED's that indicate Logins, yellow LED's that indicate traffic

## **DS-8B/DS-16B DEPARTMENTAL DIRECTORS:**

Setup using “Blocking Zone” as these switches see all ports by default [Create Zone, then add zone to a Config]

Commands: switchshow; supportshow; version; configure; switchenable; switchdisable; zoneshow; zoneadd; cfgadd; cfgcreate; cfgsave

### **MERGING:**

When “merging” or “cascading” these switches, must have different Domain ID’s per switch, but same BB Credit & same RA & ED TOV. Also must be at firmware 2.3 or higher.

1. Disable (1) of the Switches after ensuring that they have different Domain ID’s
2. Disable Zone Config on ‘disabled’ switch
3. Run ISL cable between two switches
4. Ensure that Both Switches have same Config
5. Enable Offline Switch
6. Add zones from Offline Switch to Fabric
7. Save Zone Config
8. Enable Zone Config
9. Save Zone Config again

## **DS-16B-2: Brocade Introduces ISL Trunking**

16-port switch with 2Gbps autonegotiate speed; supports ISL Trunking [allows up to (4) ISL’s to merge into single logical link with aggregate throughput of 8Gbps; trunkshow]

## **UPGRADING CONNECTRIX MANAGER AND SWITCH FIRMWARE WITH CELERRA:**

Latest ECM SW: 7.x

Latest McData Firmware: 5.x

**Note:** Ordinarily, switch upgrade activities should not impact Celerra, that is, if they are properly dual-pathed to different fabrics and you upgrade firmware in one fabric at a time. However, for the Control Station, we use only a single path, so shutdown the CS prior to upgrading firmware on switch to which it is connected. Also, conduct ‘Healthcheck’ of Data Movers after Connectrix work is completed just to make sure that we are still logged into the fabric and seeing devices properly, etc. Best method for verifying is to watch appropriate ports on the switch for I/O. However, in the event that I/O stops, conduct the following procedure on the appropriate HBA port:

## **FIBRE SWITCH INTEROPERABILITY:**

**Comment:** Interop means that you can “merge” or “cascade” different switch types into a single “Fabric” ED-1032, ED-64M, DS-32M, DS-16M, DS-8B, DS-16B—all can be used together

### **Prerequisites:**

1. Requires use of ECM Manager 4.1 or ESN Manager 1.1 [Being replaced by SAN Manager—ESN is EOL]
2. Max of (12) DS per Fabric
3. Supports only WWPN Zoning
4. Supports Switched Fabric only
5. Solaris 2.6, 7, 8 Hosts require JNI FC64-1063-EMC (SBus) HBA running 2.5.9 EMC driver  
NT/2000 Hosts require Emulex LP8000-EMC 1.27a3/1.27a5 drivers  
Linux 6.2 requires QLogic QLA2200/66 or QLA2200F/66 running 2.10.10 driver
6. DS-8B/16B requires 2.3.0 firmware if connected to ED-1032  
ED-1032 requires 3.2; ED-64M requires 1.2

**Fan-In v. Fan-Out:** Usually refers to the ratio of ports going “into” or “out of” the Symmetrix.

**Fan-In or Capacity Expansion:** (1) Server HBA can be zoned via Connectrix to go into multiple Symmetrix FA ports

**Fan-Out or Consolidation:** (6) Datamover Server HBA ports can be zoned via Connectrix to go into (1) shared Symmetrix FA port

**Current Firmware:** ED-1032 = 6.00.00.45; ED-64M, DS-32M, DS-16M = 1.2.0

### **Connectrix Logs:**

Audit Log—Configuration Changes      Event Log—Port/Hardware Problems      Hardware Log—hardware

Link Incident Log—Loss-of-Signal[Synchronization]—Occurs if cable unplugged from an attached node

## **SUPPORTED FIBRE CHANNEL SWITCHES: “Class 3 Switching”**

Brocade Silkworm 2800 [aka, Connectrix DS-16B]; 16 Ports; 16 ZoneSets [Each ZoneSet can have up to 256 Zones]

Connectrix ED-1032 Cabinet; Houses up to (2) ED-1032 Switches; Each switch=32 ports; 64 ZoneSets [256 Zones/set]

**Note:** Connect Celerra or Host connections to low-numbered ports and Storage-side connections to high-numbered ports

Connectrix Switch has two WWN’s—Port WWN=Connectrix Port; Attached Port WWN=Attached Celerra or Symm WWN

**Connectrix Enterprise Directors:** ED-64M & ED-1032 [Normally used as backbone switches in the fabric environment]

**Connectrix Departmental Directors:** DS-32M; DS-16M; DS-16B; DS-8B [Used as 'edge' switches]

**Operating Mode:** McData Fabric Mode; Open Fabric Mode; Non-Interop Mode; Interop Mode

**Connectrix Configuration:** Obtained from “ED-1032 Configuration Report” for SW1 & SW2 [16 ports per switch, totaling 32 ports]  
Shows Ports 0-31 and Nodenames attached.

**TroubleShooting Connectrix Setup:**

- Verify that unique Domain ID is used
- Verify that multiple switches are setup for Inter-Op Mode
- Verify Zone Set is active
- If ISL used, note Celerra limitations here [Celerra cannot boot up properly across an ISL link!]

**WWN—World Wide Name:** Attached N-Port node’s manufacturer HBA name or nickname. FC-4 Type=SCSI for FC-SW.  
64-bit number [10:00:08:00:88:44:50:ef]

**LINK RESET ERROR IN SERVER LOG:** 2003-06-17 07:53:55: CAM: 4: FCP HBA 0 OFFLINE: Link Reset

**POSSIBLE CAUSES FOR "Link Reset" ERROR:**

1. Bad fibre cable on HBA to switch, or from Switch to FA port--reseat or replace cables as required
2. Bad GBIC card on the Connectrix Switch--reseat or replace GBIC card as required
3. Partially completed Zoneset activation--reactivate Zoneset for affected HBA if above hardware checks out
4. Bad FA Port: Symm Inlines: 863E shows error on FA, probable cause is bad cable; 8F,,, shows Link Status and WWN Logon info; If 8F,,, doesn’t display WWN Status but has no 863E error, then switch may have a bad GBIC

**FIBRE CHANNEL PORT OPERATIONAL STATES & LED LIGHTS:**

No Green Light: No cable [no signal], plugged into port. Or, switch port blocked.

No Green Light with Yellow Triangle: Offline, Link Incident, Link Reset, or Not operational.

No Green Light with Red/Yellow Blinking Diamond: Port failure.

No Green Light, Blinking Amber, & Yellow Triangle: Testing or Beaconing in progress.

Green Light On: On Line and device communicating [blinking green when activity occurring!!]

Green Light On & Yellow Triangle: Invalid attachment or configuration.

Green Light, Blinking Amber, & Yellow Triangle: Testing or Beaconing in progress.

Flashing Green, but no flashing Amber—adapter has completed POST, but no valid link to FC or driver may not have loaded

Solid Green, flashing Amber—Post completed, link up on FC, and driver loaded

**Two Types of Switched Fabric Fibre Channel Zoning:**

1. Port-to-Port Zoning—Zones by physical switch port numbers [useful when setting up a new Fibre host & you don’t have WWN]
2. WWPN Zoning—recommended method, with zones defined by WWPortN’s of nodes and switch ports

**Note:** Specifically, **Single-HBA WWN Zoning** is recommended v. Port-to-Port Zoning. Celerra needs this to log into a single port during bootup or discovery, otherwise may try logging into all ports that it sees.

**ZONE RULES FOR FIBRE:**

--Zoning is fabric-wide, meaning that is applied to the whole fabric

--Zones can span switches

--Only a single Zone Set can be active on a switch at a time

**TROUBLESHOOTING GREEN vs. BLUE CONNECTIVITY STATUS OF HOST INITIATORS VIA NAVISPHERE:**

**Intro:** There are occasions where the “fcp bind show” will actually show an HBA with a persistent and dynamic port connection, but devconfig –probe will not be able to see the devices down the path. Navisphere port list command also shows that the path to the specific SP port is up. Navisphere might show the WWN connectivity record as a “blue” icon, when normal connectivity is reflected by a “green” icon.

**Resolution:**

There could be potential hardware port issues anywhere along the path, from HBA to Switch port to SP port. In one case, the “fcp bind show” did not trigger the path to reconnect to storage, nor did the “fcp rediscover=0”, nor did the nas\_storage –check –all.

1.) Block and unblock the Switch port should resolve the connectivity issue

2.) Log into **Navisphere>Storage Groups>Celerra\_nyip2>Properties>Hosts>Advanced>find the ‘Host Connection Path’ in question and if the box is unchecked**, check the box and apply—this should also force the path to refresh to the missing Chain or Storage path

**CELERRA SNAPSURE [Celerra Checkpoint]:**

Introduced with NAS 2.2--“nas\_fs” commands

NAS 4.0 and above allows refresh without client unmounting Checkpoint filesystem, introducing new “fs\_ckpt” commands  
NAS 4.2 [SnapSure V2] and above provides “restore”, automatic extensions when HWM is reached, single volume for all Checkpoints, and a “.ckpt” NFS entry in PFS that points to top of all snapshot directories.

NAS Upgrades from 4.2.5.1 or higher no longer requires deletion of existing Checkpoints!

## **CHECKPOINT FEATURES:**

- Checkpoint writes the first modification to any PFS block since the Checkpoint was created, to the SavVol
- NAS 4.2, Checkpoint SavVol will be same size as PFS if Checkpoint is less than 5GB
- SnapSure uses only STD volumes, not BCV's—it is not a “copy” of the PFS!
- Cannot FSCK a checkpoint filesystem
- Can extend a Production File System while a Checkpoint is established
- Checkpoint SavVol auto-extends when HWM [High Water Mark] is reached [NAS 4.2 uses pool of (8) disks]
- Can only mount Checkpoints on same DataMover as the PFS [Production File System]
- Must delete Checkpoints in order of creation, or oldest first
- SnapSure maintains itself even if Celerra or Servers are rebooted
- Create multiple SavVols for same PFS [Pre-4.2 code] if doing multiple Checkpoints over time
- Can extend an “active” SavVol, but not a “full” SavVol

**Note:** This is no longer true. A full SavVol can be extended manually with NAS 5.2 and higher.

--Bitmap flag on means a checkpoint blockmap is available. Bitmap flag off means last to read from PFS

--Can restore all or part of the file system to a given point-in-time using Checkpoints

**Limitations:** Checkpoints must be deleted in the order of their creation, if using multiple Checkpoints for same PFS.

## **CHECKPOINT (SNAPSURE) TERMINOLOGY:**

**Checkpoint:** Read-only logical point-in-time bitmap image of a production filesystem

**Checkpoint Window:** Time of initial Checkpoint SavVol creation to the time it becomes “invalid”, full, or deleted

**PFS Applications:** Applications that perform RW's to the PFS data blocks—normal day-to-day operations

**Checkpoint Applications:** Applications that use the ‘point-in-time’ PFS image or Snapsure File System, for RO operations

**SavVol:** Standard Celerra meta volumes which are used by Checkpoint to intercept writes to PFS. If the corresponding bitmap value is 0, then the existing PFS data is copied off to the SavVol before the file system write is allowed to proceed. Subsequent writes to the same block will see a bit value of 1 and allow the data to be written directly to the PFS (i.e., PFS data and metadata is only copied over to SavVol once, during the first change to the PFS state). This is why it is best to use Checkpoints to create a series of “SavVols” in order to capture incremental changes, etc., say on a daily basis [NAS 4.2 and higher uses single SavVol]. If PFS >than 10GB, SavVol will be created as 10GB and increments by 10GB at the 90% HWM. Use 0%HWM if you do not want SavVol to be auto extended.

## **DETERMINING SAVVOL INFO FOR PFS WITH SNAPSURE:**

**Note:** For any production file system greater than 10GB, the default SavVol size is 10GB. For any pfs under 10GB, the SavVol will equal the pfs size, unless otherwise provisioned. Each pfs will have its own dedicated SavVol.

**Note:** Default creation and extension size of the SavVol is 20GB beginning with NAS 5.6.44

**# export NAS\_REPLICATE\_DEBUG=1**

**# nas\_fs -l**

```
id    inuse type acl  volume   name      server
25      n    11  0    0     vpfs25 →While this is not the SavVol, a vpfs is created for each pfs when Checkpoints are in use
```

**# nas\_fs -i fs01-test\_ckpt1**

```
id    = 26
```

```
name   = fs01-test_ckpt1
```

**volume = vp94** →vp94 is the SavVol for pfs ‘fs01-test’—this is usually how to find the SavVol ID

**# nas\_fs -s fs01-test\_ckpt1**

```
volume: total = 2000 avail = 1880 used = 120 ( 6% ) (sizes in MB)
```

```
ckptfs: total = 1968 avail = 1968 used = 0 ( 0% ) (sizes in MB) ( blockcount = 4096000 )
```

**# nas\_volume -s vp94**

```
total = 2000 avail = 0 used = 2000 ( 100% ) (sizes in MB)
```

**Bitmap\_Blockmap:** Memory structures used by SnapSure to keep track of file system state and changes

**BITMAP:** Only the most recent checkpoint of any given file system contains the “bitmap”, which is 1 bit/8k PFS block [bit 0, 1]

**BLOCKMAP:** All checkpoints contain a “blockmap”, which is 8bytes for every 8k block in the Checkpoint

## **SAVVOL MEMORY QUOTAS:**

|     |                     |               |
|-----|---------------------|---------------|
| 507 | 204MB max mem quota | 163MB mem HWM |
| 510 | 2100MB              | 1680MB        |
| 514 | 2800MB              | 2240MB        |

**Rule of Thumb:** Checkpoints use 1MB of RAM for every 1GB PFS

## **GENERAL CHECKPOINT GUIDELINES:**

### **Rule #1:**

If the most recent checkpoint becomes damaged or deleted, all other previous checkpoints become useless

***Caution:*** Always protect the most current Checkpoint Refresh, as it contains the “bitmap” image required by all other previous Checkpoints!

- Checkpoint File Systems are created using Standard (STD) MetaVolumes & based on a Production File System (PFS)
  - Checkpoint creation quiesces & freezes the PFS while the Checkpoint bitmap volume of the PFS is being created [see NAS 5.2]
  - Checkpoints operate at the Celerra Volume Manager layer and use 8kb chunks, while actually writing blocks in 8.5kb chunks that include a 512byte header.
  - Each “Checkpoint” requires its own Meta Volume {STD}
  - Checkpoints are RO
  - Checkpoints must be mounted on same DataMover as PFS
  - Multiple checkpoints must be deleted in the order in which they were created
  - Total SavVol storage should not exceed 200GB per DataMover [old info, no longer current]
  - Memory utilization is capped at 70% during creation or extensions
  - Ideally, create the SavVol the same size as the PFS
- NOTE:** Checkpoints do not support complete file system restores

## **UNSUPPORTED VERSIONS:**

NAS 3.x and MPFS “HIGHROAD”

## **NAS 4.0 FEATURES:**

- Clients do not need to unmount during a checkpoint “refresh”
- Celerra automatically extends checkpoint by 10% of PFS size when a certain %full value is reached
- \$fs\_ckpt command replaces the "nas\_fs -Checkpoint"
- \$fs\_ckpt pfs1\_ckpt1 -refresh [updates the oldest checkpoint]

## **NAS 4.2 FEATURES:**

- Checkpoints for a single File System will use a single SavVol
- Maximum of (32) Checkpoints per file system
- Auto-extend feature at HWM if disk space is available—if not, then oldest Checkpoint is marked “invalid” & is then reused
- Auto-extend feature only works if using automatic SavVol, not manually created SavVol
- Auto-mounting of SavVol for NFS-only in 4.2.x

## **NAS 5.0 FEATURES:**

- Separate Checkpoint directories are contained within the “uxfs” for each Checkpoint created for a PFS as “.ckpt\_ckptx” entries

## **NAS 5.1M2 FEATURES:**

- Total number of checkpoints per PFS is 32
  - Not possible to modify a Checkpoint Schedule after it has been setup from Celerra Mgr or CLI
  - Some Checkpoint tasks may not be visible or manageable from Celerra Manager
  - With ATA environments that exceed 6 TB on the Celerra, SnapSure is limited to a single Checkpoint per file system
- NAS 5.2 FEATURES:**
- Subfolders now contain hidden .ckpt directories
  - Out of order deletion of checkpoints are now supported [merges blockmap entries to older checkpoint before deleting]
  - Supports 32 checkpoints per file system—submit RPQ for supporting more
  - Checkpoints will automatically be mounted after being taken
  - Can now rename a specific checkpoint
  - Checkpoint creation pauses PFS by suspending writes but not reads, while checkpoint is being created
  - Checkpoint refresh freezes PFS by suspending both reads and writes

## **NAS 5.3 FEATURES:**

- New Checkpoint Schedules Tab showing schedules, date, time, and state of each [hourly, daily, weekly, monthly intervals]
- Note:** Start On schedule can be modified, but not for Next Run, State, and Recurrence, or name of Schedule. Checkpoint Scheduling uses the APL GUI middleware layer, built on top of CS core exec layer.

## **LOCATION FOR CHECKPOINT SCHEDULES CREATED FROM WEBUI:**

### **/nbsnas/tasks/schedule**

- rw-r--r-- 1 nasadmin root 834 Nov 1 11:29 laip2.1 →Example of a Checkpoint schedule
- rw-r--r-- 1 nasadmin nasadmin 1 Nov 1 11:29 .sched\_id →Schedule ID number

## **STANDARD MODE SCHEDULING:**

Recurrence—ability to setup a recurring time, start & end dates

Hourly—one for each selected hour

Daily—up to 7 days

Weekly—up to 4 weeks

Monthly—up to 3 months

→Ability to name virtual checkpoint directories for customer—set following param:

**param cvfs virtualDirName=snapshot** (example name)

**Note:** Default is to use “ckpt” as the virtual directory name for Checkpoints on the PFS

### **VOLUME SHADOW COPY SERVICE (VSS):**

→Support for Microsoft VSS (Volume Shadow-Copy Service)

→Provides direct access to point in time copies of volumes, providing restore capability

→Support for XP & Win2k3 clients for NAS 5.3/5.4

→Windows Service is called “Volume Shadow Copy”

→VSS can be used for Full, Incremental, Differential, & Copy backups

#### **VSS uses several components:**

Requestor, Writer, Provider. The Requestor controls the backup and restore service using API calls to Windows (Legator NetWorker, Replication Manager, etc). The Writer is an Application that can support VSS, such as Exchange 2003, SQL 2005, etc., and creates datasets for the snapshot backups. The Provider (System, Software, Hardware component, such as Legato RM) actually creates the backup & presents to Host, or ‘Snapshot Shadow Copy’.

**Note:** Clients must install software to support VSS

#### **Client Software:**

→Celerra iSCSI Target Software CD [Windows Celerra VSS Provider; Linux CBMCLI pkg; AIX pkg]

→Ensure Windows system MaxRequestHoldTime parameter is changed to 600 decimal to tolerate iSCSI LUNs going offline during snapshots

→CBMCLI pkg for Linux is used to manage snapshots with cbm\_iscsi; cbm\_security; cbm\_replicate commands

#### **Two Types of Shadow Copies:**

→Plex copies (mirrored), not supported

→Differential copies, differences from original volumes are saved (supported)

#### **Celerra VSS Provider for iSCSI Version 1:**

→Supports Windows 2008 except for multiple imports; auto-recovery of transportable shadow copies; NTFS TxF transactional recovery of transportable shadow copies

→Runs as a service on Windows

→Celerra VSS provider does not support Shared Folders; Importing shadow copies to clustered servers; Plex shadow copies

→Celerra VSS does support differential Shadow Copies

→Celerra VSS allows VSS-enabled backup applications to make shadow copies of Celerra iSCSI luns

#### **MS UTILITIES:**

C:\>vssadmin list providers | list writers

C:\>vshadow -p c:\ (Using local NTFS requestor to perform backup of c: drive)

**param cifs allowSnapSureVss=0** [disables VSS support—Shadow Copy for Shared Folders (SCSF)]

**param cifs allowSnapSureVss=1** [default; enables VSS support—allows access to previous file versions via checkpoints]

### **CVFS (Checkpoint View File System):**

NAS 5.0 introduces “virtual” .ckpt entry in all directories on PFS & provides access to Snapshots from PFS

**Note:** 5.0.x contains bug in that only the first PFS created with a Checkpoint has the hidden .ckpt directory—fixed with NAS 5.0.14.0

#### **CVFS VERSION 1:** NAS 4.2

Only NFS clients can access checkpoints, and then only from the root of the PFS’s hidden .ckpt directory. CVFS1 supports DOS 6.0 file naming conventions only.

### **ENABLING/DISABLING CLIENT ACCESS TO CHECKPOINTS FOR SVFS V1:**

**param cfs showChildFsRoot=1 | 0**

**Note:** When set to 1, mounted Checkpoints will be visible to NFS clients as subdirectories of the root of the PFS—0 disables this

#### **CVFS VERSION 2:** NAS 5.1

NFS & CIFS clients can access checkpoints of the PFS via virtual links and hides the .ckpt directories when an ls -la or dir is run.

### **ENABLING/DISABLING CVFS VERSION 2 CHECKPOINT ACCESS FOR CLIENTS:**

**param cfs showHiddenCkpt=1 | 0**

**Caution:** AR57334 states that the showHiddenCkpt param applies only for CVFS v1-style checkpoints, and that setting param to 0 will disable CVFS v2 checkpoints completely

#### **PURPOSE OF CHECKPOINT:**

--As a source of data for applications that do not require live data

--Typically used for incremental backups by copying the Checkpoint to tape

--Most practical for restores of individual files or folders

**Note:** Checkpoint is not intended to be a complete Backup solution nor to restore whole file systems! Restores of whole PFS could cause SavVol to fill, which could render the Checkpoint “invalid” and unusable.

## **HOW CHECKPOINT WORKS:**

### **PFS SIDE:**

→ Writes are intercepted at the Volume level. If Blocks have not been modified since last SnapShot taken, then the blocks are first saved off to the SavVol and then the write to the PFS is allowed. If Blocks have already been modified once, then writes are executed directly to the PFS. Reads are executed against the PFS.

### **SAVVOL SIDE:**

→ Reads are intercepted at the Volume level. Writes are prohibited. If blocks have been modified since last SnapShot, then Reads are executed on SavVol. If blocks have not been modified, then reads are done against the PFS.

→ Copies are done at the Volume level and are executed as a copy on the first modify of a data block. During initial SavVol creation, a bitmap image of the production file system is created. After initial creation, changes to the Production File System [PFS] are first written to SavVol as “data blocks”, changing the bitmap values from 0 to 1--subsequent writes to the same blocks will be written to the PFS. Multiple Checkpoints can be taken to span a period of time for a production file system—mount the Checkpoint File System from any Unix Host to retrieve files. Otherwise, a Checkpoint “Restore” will restore the whole checkpoint bitmap changes.

Checkpoint file systems remain "persistent" until SavVol is filled or unless deleted. PFS can even be extended while Checkpoints are active and established.

→ By default, the SavVols can use up to a max of 20% of all available disk space on a system [tunable in the /nas/sys/nas\_param file]

## **HOW READS & WRITES OCCUR USING CHECKPOINT:**

**Writes:** After initial Checkpoint SavVol is created for the Production File System, all its Bitmap values are represented by 0's. Any new writes to the PFS are then intercepted at the Volume level--original data blocks are first moved to the SavVol, data changes are then written to the PFS, its bitmap values for the changed blocks are updated to 1's. It is important to understand that this bitmap value change from 0's to 1's are only done once for any given Checkpoint—Subsequent changes to any data blocks that already have a new value of 1 are done directly to the PFS. In otherwords, Checkpoints record only the first series of changes to data after initial creation.

**Reads:** Similar to ‘Writes’--also intercepted at the Volume level. Bitmap values of 1 are read from the SavVol, values of 0 are read from PFS.

**Performance Impact:** Some additional overhead is required to conduct the first modification to a PFS Block, as the PFS Block has to be Read and then Written to the Save Volumes, while the original modification needs to be written to the PFS Block.

## **CHECKPOINT PROCESS:**

- Creation of a Checkpoint causes a pause of the PFS for Writes only
- Conducting a Checkpoint Restore causes a pause of Reads & Writes
- Checkpoints are automatically restarted if Server is rebooted
- During Restore, a rollback Checkpoint is taken in case the restored copy was not the desired one
- Refresh operations may cause CIFS clients to have to reconnect—NFS Clients should have no trouble
- Checkpoints consume memory in order to maintain the Bitmap/Blockmap images

## **SNAPSURE V1 vs. V2:**

- V1 Checkpoints became inactive if Save Volume became full
- V2 Checkpoints inactivates the oldest Checkpoint if the SavVol becomes full
- V1 Checkpoints required deletion of oldest-to-newest in exact order
- V2 Checkpoints allows for deletion of any Checkpoint

## **SAVVOL VOLUME & SIZE:**

SavVols are created on metavolumes. Specify the SavVol by name and size in GB or % of PFS. Minimum SavVol size=64MB [NAS 5.2 min = 32MB]. Ordinarily, if PFS is less than 5GB in size, then SavVol will equal the size of the PFS. For SnapSure 4.2 and higher, only a single SavVol is used for all CheckPoints. Creation of the first CheckPoint will create a SavVol—to specify a SavVol HWM different than the default of 90%, use the following command to create SavVol before implementing CheckPoints:

**\$fs\_ckpt fs1 -Create -o %full=75**

**Note:** The default HWM for the checkpoint SavVol is 90%, which means when the SavVol is 90% full, it will be auto-extended by a default size of 10GB. HWM defaults have changed from NAS version to version—75% old versions, 90% is the HWM in later NAS versions including 5.6. The HWM can be changed by manually creating a new CheckPoint and specifying--SavVol size or % of PFS size, as well as the type of AVM volume pool, can also be specified when creating the SavVol.

### **CHANGING CHECKPOINT SAVVOL HWM FROM DEFAULT 90%:**

**# fs\_ckpt file7 -C -option %full=50**

**# fs\_ckpt file7 -l**

```
id ckpt_name creation_time inuse fullmark total_savvol_used ckpt_usage_on_savvol
57 file7_ckpt1 12/03/2009-13:14:30-EST y 50% 43% 9%
```

### # nas\_fs -info file7\_ckpt1|grep full

full(mark)= 50% → This is the Checkpoint SavVol HWM, the point at which the SavVol will auto-extend

### **SAVVOL HWM NAS 5.6:**

→ Default value is set at 90%, at which point the SavVol will auto-extend by 20GB

→ If you set the HWM to 0% (using fs\_ckpt -C –option %full=0), then the SavVol will not auto-extend. Instead, when more space is needed, the oldest checkpoint will be deleted to reclaim space, and this is repeated everytime the SavVol needs more space

### **MANUALLY EXTENDING SAVVOL:**

#### # nas\_fs -xtend <ckpt name> size=<integer>[T|G|M]

### **EXTENDING SAVVOL USING CELERRA MANAGER:**

Select the "File Systems" tab, highlight the filesystem name, rightclick and select "Checkpoints>Extend Checkpoint Storage", and on the input screen, enter the desired SavVol extension size in MB in the "Extend Size by (MB):" box, and accept all other default values.

### **NAS 5.6 SAVVOL AUTO EXTEND:**

**Note:** NAS 5.6 QOSsize represents the default auto-extension size of the SavVol

#### cmd log:

2009-12-03 14:16:32.215 db:0:27704:S: /nas/sbin/rootnas\_fs -x file8\_ckpt5 **QOSsize=20000M →20GB**

2009-12-03 14:16:35.164 db:0:27704:E: /nas/sbin/rootnas\_fs -x file8\_ckpt5 QOSsize=20000M

#### server log:

2009-12-03 14:16:52: SVFS: 4: 1: FSID:69 SavVol:127 MaxSize:0 MB %Full(hwm=30) reached (t:3140622644124770)

2009-12-03 14:16:56: ADMIN: 6: Command succeeded: volume extend 127 129

### **DETERMINING SAVVOL SIZE FOR PFS WITH NAS 5.5:**

#### # nas\_fs -size fs01-test\_ckpt1

volume: total = 2000 avail = 1880 used = 120 ( 6% ) (sizes in MB)

**ckptfs: total = 1968 avail = 1968 used = 0 ( 0% ) (sizes in MB) ( blockcount = 4096000 )**

**Note:** With production fs size of 10GB or greater, the SavVol size will be created as 10GB

### **FINDING SAVVOL INFORMATION FOR SNAPSURE CHECKPOINTS:**

**Note:** SnapSure SavVols are different from IP Replication SavVols

#### 1. Find a Checkpoint name for the Production File System:

##### # nas\_fs -i fs1

```
id      = 54
name    = fs1
acl     = 0
in_use  = True
type    = uxf
worm   = off
volume  = v135
pool    = clar_r5_performance
member_of = root_avm_fs_group_3
rw_servers= server_2
ro_servers=
rw_vdms =
ro_vdms =
auto_ext = no,virtual_provision=no
ckpts   = fs1_ckpt1
```

#### 2. Run query on a checkpoint to find Checkpoint SavVol Volume Pool name and Volume Pool File System name:

##### # nas\_fs -i fs1\_ckpt1

```
id      = 56
name    = fs1_ckpt1
acl     = 0
in_use  = True
type    = ckpt
worm   = off
volume = vp138 →Volume Pool SavVol name
pool    = clar_r5_performance
member_of = vefs55 →Volume Pool File System name
```

Alternatively, use the following Query command to locate SavVols, though if running IP Replication, will need to sort through the list as this command dumps info on IP Replication & SnapSure Checkpoint SavVols:

**# nas\_fs -query:Type=uxfs -format:"%q" -fields:Checkpoints -query:\* -format:"ID=%s, Name=%s, VolumeID=%s, VolumeName=%s\n" -fields:ID,name,VolumeID,VolumeName**

```
ID=31, Name=rep1_src_repl_restart_1, VolumeID=112, VolumeName=vp112  
ID=38, Name=rep1_src_repl_restart_2, VolumeID=112, VolumeName=vp112  
ID=43, Name=rep2_src_repl_restart_1, VolumeID=124, VolumeName=vp124  
ID=50, Name=rep2_src_repl_restart_2, VolumeID=124, VolumeName=vp124  
ID=56, Name=snapper_ckpt1, VolumeID=138, VolumeName=vp138
```

## **AUTOMATIC EXTENSION OF SAVVOLS IN NAS 5.2:**

Normally, if SavVol reaches its HWM, the SavVol will automatically extend. If an extend fails [i.e., no more storage allocated], then the remaining SavVol space above the HWM will be used. If more space is still required, the oldest checkpoint will be deleted and its space used. This pattern will repeat until all old checkpoints are consumed. If more space is required, the SavVol will become inactive and all CheckPoints will become invalid.

**Note:** Simply deleting Checkpoints in the middle of a sequence will not necessarily free up any additional space for the SavVol

## **RESOLVING SAVVOL FULL PROBLEM:**

Occasionally, due to certain problems [no more physical diskspace, code bug, etc.], the SavVol can be filled to capacity and not be able to auto-extend itself. In these situations, the various Checkpoints begin to go invalid as they fill. Following are some generic steps to recover:

1. Manually extend the SavVol once more diskspace has become available

**# nas\_volume -xtend fp1368 d1000**

**Note:** Find the savVol itself by running nas\_fs -i against any checkpoint—the SavVol will be the “volume=vp175” entry.

2. Delete any inactive Checkpoints that may have resulted, first making sure that the invalids are permanently unmounted from DM
3. Refresh the deleted Checkpoints if desired

## **RESIZING REPLICATION SAVVOL ON DESTINATION (5.6.38+) (Shrink SavVol/Resize SavVol):**

**Note:** No capability to “shrink” the SavVol. Must delete all the checkpoints for a given file system so as to delete the SavVol.

- 1) Obtain details of Replication Session on both sides
- 2) Delete destination side checkpoints
- 3) Create new checkpoint on Source file system
- 4) Refresh the Replication session on the Source side
- 5) Delete the Replication sessions on both sides (nas\_replicate –delete <session> -mode both
- 6) Create new checkpoint on Source file system
- 7) Perform incremental copy of checkpoint created in Step 3 with checkpoint created in Step 6  
\$ nas\_copy –name <copy\_name> -source –ckpt <ckpt\_step 6> -destination –fs <dest\_fs> -from\_base <ckpt\_step 3> -interconnect id=<interconnected> -overwrite\_destination -background
- 8) Confirm nas\_copy has completed, then recreate Source Replication session from info in Step 1  
\$ nas\_replicate –create <session\_name> -source –fs <source\_fs> -destination –fs <destination\_fs> -interconnect id=xxx –overwrite\_destination -background

**Note:** If SavVol needs to be deleted on Source side, then do a –reverse first, then apply the above procedure, ending with –reverse back. But, source side needs all checkpoints & internal checkpoints deleted to remove SavVol, which means Replication needs to be torn down, so this is really not practical.

## **CHECKPOINT EXTEND FAILS:**

Checkpoint extend script [extend\_ckpt] can inadvertently extend a checkpoint multiple times, leading to failure. Possible workarounds would be to increase the size of the SavVol and then tailor the “extend\_ckpt” script as follows:

```
if [$used -le $hwm ]; then  
fi
```

## **AVM STORAGE MANAGER & CHECKPOINTS NAS 5.2:**

AVM will try to match the storage profile for creating a SavVol with that of the PFS when possible. If not possible, AVM will resort to allocating storage using QOS [introduced NAS 4.2] for SavVols. Quality of Service (QoS) is a mechanism to help ensure that there is always space available in the pool for filesystem creates, extends, or checkpoints—also reflected with the flag is\_dynamic is true. QoS is only used for system-defined pools, not user-defined.

## **HIGH WATER MARKS (HWM) IN NAS 5.2:**

By default, a HWM of 90% is used with the SavVol before an extension of the SavVol occurs. Server Log will show the HWM during a CheckPoint refresh:

**2004-04-06 23:04:12: ADMIN: 4: Command succeeded: file refresh ckpt 34 260=252 hmw=90**

To conserve storage system space, you can set the HWM to 0%, meaning that the SavVol will not be automatically extended. Instead, Checkpoint will delete oldest checkpoints first to conserve space. However, the caution here is that if not monitored closely, you run the risk of running out of old checkpoints and rendering the SavVol inactive and all CheckPoints useless. Setting SavVol to 100% simply means that all space is used in the SavVol before an extension is made.

**ERROR MESSAGE WHEN CHECKPOINT CANNOT AUTO-EXTEND:**

Sep 17 14:42:59 2002 NASDB: 7; 10 /nas/bin/nas\_fs -x fs1-ckpt2

QOSsize=5000M error 3007; volume is unreachable

**CHECKPOINT OPERATION:**

**CREATING CHECKPOINTS NAS 4.2/5.1:**

**Step 1:**

Option 1. **\$fs\_ckpt fs1 -Create** [If not specified, will use first (8) STD Volumes available to create Checkpoint pool]

Option 2. **\$fs\_ckpt fs1 -C mtv1 -o %full=70**

**Note:** Using option “%full=70” will create HWM and point at which the Checkpoint will be automatically extended

Step 2: Create Mountpoint and Mount Checkpoint FileSystem to DataMover

Step 3. Export Checkpoint

**Caution:** Never permanently unmount a PFS to which a Checkpoint is still mounted!

**RENAMING A CHECKPOINT:**

1. Temp Unmount Checkpoint first: \$server\_umount server\_2 fs1\_ckpt1 /fs1\_ckpt1

2. Rename: \$server\_mount server\_2 -o cvfsname=Monday snap1\_ckpt1 /fs1\_ckpt1

**SCHEDULING CHECKPOINTS USING WEBUI:** NAS 5.1.13.0 and higher

Checkpoint refreshes can be setup using the WebUI Manager Interface or by Linux CS cron job using a script. Schedule interval using WebUI is Hourly, Daily, Weekly, Monthly, though only a single schedule can be in place for any single PFS. It is not yet possible to “change” a schedule, once created. You must remove the current schedule by changing the recurrence of the CheckPoint on the PFS to “never”, apply, and then set the new schedule. You can delete a ‘scheduled’ checkpoint if it’s in an ‘active’ status on the WebUI display, but not if it’s in a ‘pending’ status.

**NAS 5.3:**

Schedules can now be changed without deleting and recreating—change from one time to another.

**Note:** Do not schedule CheckPoint refreshes from 00-05 minutes past the hour as it may fail due to the NAS Database backup that occurs at (1) minute past each hour. Also note that Schedule times are driven by Control Station timezone.

**SCHEDULING STATES:**

**Pending:** Schedule has not yet run the first task

**Active:** Schedule has run at least one task and others are still waiting

**Paused:** Schedule has been stopped—missed tasks will not run

**Completed:** Schedule completed and no more tasks waiting

**MANUALLY DISABLING CHECKPOINTS SCHEDULED FROM WEBUI:**

Checkpoint Refreshes created from the WebUI run using “cron\_plus” and are located in the /nas/site/cron.d/nas\_users file. You can also change the ‘scheduled’ times in this file, but the WebUI will not reflect the changes.

**1. # vi /nas/site/cron.d/nas\_user** [Comment out CheckPoint refreshes]

# SnapSure Checkpoint job 1 of 1

#0 6 \*/1 \* \* nasadmin /nas/sbin/cron\_plus 75721061798813\_2 0 2147483647 -f /nas/site/fs\_wst\_exhibit\_1

# SnapSure Checkpoint job 1 of 1

#0 6 \*/1 \* \* nasadmin /nas/sbin/cron\_plus 96251064884803\_2 0 2147483647 -f /nas/site/fs\_gsmo\_eishome\_1

# SnapSure Checkpoint job 1 of 1

#0 5 \*/1 \* \* nasadmin /nas/sbin/cron\_plus 56521065591798\_2 0 2147483647 -f /nas/site/fs\_guts\_nasexp01\_1

**2. # touch /etc/crontab** [Forces changes into CS memory by reading the changed file]

**Note:** Actual refresh schedule is run by /nas/sbin/cron\_plus and the job files are located in /nas/site directory

**LOCATION OF CHECKPOINT SCHEDULES NAS 5.1 & 5.2:**

**/nas/volume/fscps**

**Note:** Control Station location for Checkpoint Schedules in NAS 5.1 and higher created from Celerra Manager.

**\$ ls -la**

-rw-r--r-- 1 nasadmin nasadmin 520 Sep 9 13:00 sched\_pfsBing

**\$ cat sched\_pfsBing**

```
<CKPT_SCHEDULE KEEP='1' PFS='pfsBing' USER='nasadmin'>
<PATTERNS>
<PATTERN JOB_ID='107931082604499_2'
UNIT='5' UNIT_STR='Month' FREQ='1'
MIN='0' HOUR='12' MDAY='1'
MONTH='*/1' WDAY='*'
TSPEC='0 12 1 */1 *'
TRANSLATION='On the 1st of every month, at noon'>
<CHECKPOINTS>
<CHECKPOINT NAME='ckpt_pfsBing_0001'
SCHEDULED_CREATE='2004/05/01-12:00:00'
SCHED_UTC='1083430800' UTC_OFFSET='18000' CREATED='False' />
</CHECKPOINTS>
</PATTERN>
</PATTERNS>
</CKPT_SCHEDULE>
```

### **CHECKPOINT SCHEDULES NAS 5.3-6.0:**

**/nbsnas/tasks/schedule**

-rw-r--r-- 1 nasadmin nasadmin 2831 Feb 19 15:50 F100\_daily.2

-rw-rw---- 1 nasadmin nasadmin 1 Feb 19 15:50 .sched\_id

srxwxrwxr-x 1 nasadmin nasadmin 0 Dec 28 12:45 .scheduler\_socket

### **SCEDULING TROUBLESHOOTING:**

**/nas/log/webui/apl\_sched.log**

Event ID 211 means Scheduled task was missed (Paused?)

Event ID 212 means Scheduled task failed

### **VERIFYING OR AUDITING CHECKPOINTS:**

**\$ nas\_fs -i -s fs1\_ckpt1**

**\$ nas\_fs -i -s fs1** [Details of PFS & its Checkpoints, how many, & how many MB used]

**\$ server\_df server\_2**

**\$ nas\_fs -l**

**Note:** Nas\_fs will show High Water Mark as “full (mark) = 70%” & % left in SavVol as “used = 50”

### **FILE SYSTEM AND CHECKPOINTS:**

**# nas\_fs -list |grep file1**

|    |   |   |   |     |             |   |
|----|---|---|---|-----|-------------|---|
| 23 | y | 1 | 0 | 90  | file1       | 1 |
| 29 | y | 7 | 0 | 100 | file1_ckpt1 | 1 |
| 36 | y | 7 | 0 | 100 | file1_ckpt2 | 1 |
| 43 | y | 7 | 0 | 100 | file1_ckpt3 | 1 |

### **NAS FS INFO FOR PFS:**

**# nas\_fs -i file1**

```
id      = 23
name    = file1
acl     = 0
in_use  = True
type    = uxf
worm   = off
volume  = v90
pool    = clar_r5_performance
member_of = root_avm_fs_group_3
rw_servers= server_2
ro_servers=
rw_vdms =
ro_vdms =
auto_ext = hwm=90%,virtual_provision=no
deduplication = On
ckpts   = file1_ckpt1,file1_ckpt2,file1_ckpt3,ckpt_daily_009,ckpt_daily_010,ckpt_daily_001,ckpt_daily_002,ckpt_daily_003
```

### **NAS FS INFO ON CHECKPOINT:**

**# nas\_fs -i file1\_ckpt1**

```

id      = 29
name    = file1_ckpt1
acl     = 0
in_use  = True
type    = ckpt
worm   = off
volume  = vp100
pool    = clar_r5_performance
member_of =
rw_servers=
ro_servers= server_2
rw_vdms =
ro_vdms =
checkpt_of= file1 Mon Nov 23 14:59:02 EST 2009
deduplication = Off
used    = 8%
full(mark)= 90%
stor_devs = APM00083201184-0011

```

**SERVER DF OUTPUT:**

```

# server_df server_2 |grep file1
file1_ckpt1  2015984  179744  1836240  9%  /file1_ckpt1
file1_ckpt3  2015984  658920  1357064  33%  /file1_ckpt3
file1_ckpt2  2015984  368808  1647176  18%  /file1_ckpt2
file1      4032984  891256  3141728  22%  /file1

```

**FILE DISPLAY COMMAND:****# .server\_config server\_2 -v "file display ckpt 23"**

```

1266613959: SVFS: 6: Checkpoint on FsID:23 (Oldest to most recent):
1266613959: SVFS: 6: CKPT Created:13 and CanCreate:97 CKPT
1266613959: SVFS: 6: Writeable Snap Created:0
1266613959: SVFS: 6: CKPT Total (Active + others):13
1266613959: SVFS: 6: Restoring PFS:FALSE FromFsid:0
1266613959: SVFS: 3: ****
1266613959: SVFS: 3: ckptID:29
1266613959: VRPL: 7: VPM::getHWM hwm=90
1266613959: VRPL: 7: VPM::getMaxVolSize size=0 (MB)
1266613959: SVFS: 6: CheckPoint is Active. MaxVolSize:0 hwm:90
1266613959: SVFS: 6: FsVol #CFSBlocks:256000, (at the time the CheckPoint was created)
1266613959: SVFS: 6: SavVol #CFSBlocks:2816000, #CFSBlocks used:213056, #CFSBlocks free:2602944
1266613959: SVFS: 6: BitMapMemUsage:0 Bytes, BlockMemUsage:24576 Bytes
1266613959: SVFS: 6: Number of Blocks in the BlockMap: 23646 ; #MemPages = 3
1266613959: SVFS: 6: Paged=Yes CKPT Type=2.5 →Checkpoint version
1266613959: SVFS: 6: ckptSizeMB:128
1266613959: SVFS: 3: ****

```

**Note:** Useful in displaying the Checkpoint order on the PFS, SavVol info, etc. Use lefthand ID number for PFS from nas\_fs -l, not Volume ID number.

**# .server\_config server\_2 -v "file display ckpt fsvol=90" (90 is the volume ID for the file system)**

```

1266614187: SVFS: 6: Checkpoint on FsVol:90 (Oldest to most recent):
1266614187: SVFS: 6: CKPT Created:13 and CanCreate:97 CKPT
1266614187: SVFS: 6: Writeable Snap Created:0
1266614187: SVFS: 6: CKPT Total (Active + others):13
1266614187: SVFS: 6: Restoring PFS:FALSE FromFsid:0
1266614187: SVFS: 3: ****
1266614187: SVFS: 3: ckptID:29
1266614187: VRPL: 7: VPM::getHWM hwm=90
1266614187: VRPL: 7: VPM::getMaxVolSize size=0 (MB)
1266614187: SVFS: 6: CheckPoint is Active. MaxVolSize:0 hwm:90
1266614187: SVFS: 6: FsVol #CFSBlocks:256000, (at the time the CheckPoint was created)

```

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!

1266614187: SVFS: 6: SavVol #CFSBlocks:2816000, #CFSBlocks used:213056, #CFSBlocks free:2602944

1266614187: SVFS: 6: BitMapMemUsage:0 Bytes, BlockMemUsage:24576 Bytes

1266614187: SVFS: 6: Number of Blocks in the BlockMap: 23646 ; #MemPages = 3

1266614187: SVFS: 6: Paged=Yes CKPT Type=2.5

1266614187: SVFS: 6: ckptSizeMB:128

1266614187: SVFS: 3: \*\*\*\*\*

### # .server\_config server\_2 -v "file stat ckpt"

DV : DeltaVol FSV : FSVol SV : SaveVol

RD : Read WS : WriteSnap WNS : WriteNoSnap

TOT : Total AT : AverageTime AVS : AverageVolumesSearched

SWNS : SmartWriteNoSnap

Global Performance Summary Statistics

| Operation | Total | AT | AVS |
|-----------|-------|----|-----|
|-----------|-------|----|-----|

|                             |      |     |    |
|-----------------------------|------|-----|----|
| Ckpt Reads From Save Volume | 6M   | 3ms | 11 |
| Ckpt Reads From PFS Volume  | 144M | 6ms | 2  |
| PFS Write Snaps             | 390M | 4ms | 1  |
| PFS Write No Snaps          | 0    | 0us | 0  |
| PFS Smart Write No Snaps    | 192M | 2ms | 6  |
| PFS Restores                | 0    | -   | -  |

Paging And Memory Statistics

| Indicator Name | Value |
|----------------|-------|
|----------------|-------|

|                               |         |
|-------------------------------|---------|
| Total Paged In                | 56367   |
| Total Pages Flushed To Disk   | 4108449 |
| Total Pages Purged            | 231005  |
| Page In Rate                  | 0       |
| Page Out Rate                 | 1       |
| Block Map Memory Quota(KB)    | 1048576 |
| Block Map Memory Consumed(KB) | 292616  |

### PRINTSTATS OUTPUT OF CHECKPOINTS FOR SERVER 3:

#### # .server\_config server\_3 -v "printstats svfs"

\*\*\*\*\*

\* Checkpoint statistics for those PFS with checkpoint FS: \*

\*\*\*\*\*

Checkpoint statistics for FS 38:

-----  
Per PFS totals:

Snapped volume:  
snappedRd 156066 2888364 18  
snappedWt 384367 235714217 613

0 SnapshotVol reads: 29 savVol reads, -29 cloneVol reads  
384367 SnappedVol writes: 384364 snapWrites, 0 nonSnapWrites

Checkpoint statistics for FS 42:

-----  
Per PFS totals:

Snapped volume:  
snappedRd 6 135 22  
snappedWt 2340 95982 41  
0 SnapshotVol reads: 0 savVol reads, 0 cloneVol reads  
2340 SnappedVol writes: 2340 snapWrites, 0 nonSnapWrites

**NOTE:** Only DeltaVol read times (thus SnapshotVol reads) include I/O time--other times are just overhead time on I/O, not I/O time.

### EXTENDING CHECKPOINT SAVVOL:

#### # nas\_fs -xtend fs1\_ckpt1 -volume mtv1\_a

**Note:** Automatic for NAS 4.x whenever HWM of SavVol is reached

NAS 4.0 extends automatically by 10% of PFS original size

NAS 4.2 extends automatically by 500MB

### RESTORING CHECKPOINTS:

Option 1: Restore files or folders manually from a remotely mounted Unix Client to the exported Checkpoint filesystem and Production File System using a Unix copy command. In the case of CIFS, use Windows Explorer to access a network Share

Option 2: Restore from commandline while PFS is mounted:

**# /nas/sbin/rootfs\_ckpt fs1\_ckpt1 -Restore**

**Note:** A new Checkpoint is automatically created in 4.2.x prior to conducting the actual “restore” operation

## **NAS 5.2 CHECKPOINT RESTORES:**

Restores can be done online to the PFS. SnapSure creates a new CheckPoint prior to the restore. Important that a HWM of 75% be indicated for the restore operation to ensure success. If the SavVol has to be automatically extended, the first Restore attempt may fail—try the operation again.

## **EXAMPLE → RESTORING FROM CHECKPOINT:**

**# /nas/sbin/rootfs\_ckpt fs1\_ckpt1 -R -o %full=75**

## **EXAMPLE OF CHECKPOINT RESTORE vs. REFRESH:**

Customer has (10) Checkpoints (ckpt1-10) of a PFS captured on a single SavVol, and wants to do a ‘Restore’ of “ckpt5”. What happens? The ‘Restore’ operation will first create a new “ckpt11” of the PFS in case the ‘Restore’ is not actually wanted, and will then conduct the ‘Restore’ without affecting any of the other checkpoints. A ‘refresh’ operation, on the other hand, will delete the oldest Checkpoint in order to create the new Checkpoint.

## **DELETING CHECKPOINTS:**

**# nas\_fs -d checkpoint\_name**

**Comment:** Normal rule is always delete the oldest Checkpoint of a PFS first. Must permanently “unmount” CheckPoints before deleting—use the –Force flag if required to unmount.

## **ALTERNATE METHOD FOR DELETING CHECKPOINT:**

**#.server\_config server\_3 -v "file delete ckpt ckptFSID PFSID SaveVol=Fsvol force"**

**#.server\_config server\_3 -v "file delete ckpt 948 102 605=569 force"**

**Note:** 948 is ckptID for ckpt fs 102. 605 is ckpt volume number and 569 is volume number for PFSID

## **CONDUCTING CHECKPOINT REFRESHES:**

--NAS 2.2 Checkpoint refresh option requires that a Checkpoint be unmounted first

--NAS 4.x Checkpoint refresh option does not require that Checkpoint be unmounted

**# fs\_ckpt fs1\_ckpt1 -refresh**

**Note:** In the case of multiple Checkpoints for a single PFS, “refreshes” must takes place on the oldest Checkpoint first. A ‘refresh’ is really a deletion and a re-creation of the checkpoint. With 5.2 out of order deletion and refresh, if the checkpoint is not the oldest, the checkpoint is compacted and merged with other checkpoints. ‘Refresh’ failures are logged in the “sys\_log”. If scheduled checkpoint refreshes fail, determine why, correct, then either manually refresh or let the next scheduled refresh occur. Do not allow ‘refreshes’ to occur while the NAS Database backup is occurring, which starts at (1) minute past each hour.

## **VERIFYING CHECKPOINTS FOR A PFS:**

**# nas\_fs -i fs1**

ckpts = fs1\_ckpt1, fs1\_ckpt2, fs1\_ckpt3

**Comment:** After a “refresh” of a checkpoint, the order will change from “left-to-right” as “oldest-to-newest”. Refreshes can also be scheduled at specific intervals using “Celerra File Server Manager”

## **VIRTUAL ENTRIES ON PFS WITH NAS 4.2:**

Production File System: FS1

Checkpoint Virtual Directory Entry for PFS: .ckpt\_fs1\_ckpt1

**Comment:** Mount the datamover’s file system from Unix Host to view or restore files from the checkpoint directory

## **CREATING CHECKPOINT CALLED “ENG CKPT1” FROM METAVOLUME “MTV1”:**

1. Create Checkpoint: **\$nas\_fs -n eng\_ckpt1 -Checkpoint eng -volume mtv1 -o %full=80**
2. Create Mountpoint, Mount & Export to same DataMover as PFS [Read-Only]
3. Auditing SavVol Space: Check Server Log HWM Entries & look at “used =65%” section of **\$nas\_fs -i eng-nb**
4. Extending a Checkpoint: **\$nas\_fs -xtend eng\_ckpt1 -volume mtv1\_a**
5. Deleting Old Checkpoints: **\$nas\_fs -d eng\_ckpt1**
6. Restoring a file from a Checkpoint FS using a Unix Host: \$cp /source/ckptfile1 /target/ckptfile1

**Note:** If not specified, Checkpoint will create a default SavVol that is 10% the size of the PFS

**Caution: For 2.2 NAS code, if you do not extend a SavVol before it becomes full, it is rendered inactive and unusable!**

## **TROUBLESHOOTING CHECKPOINTS:**

**Information Needed by Eng:** Server Log; Command Log; Crash Dump, if applicable

# cat /nas/sys/nas\_param

ckpt:10:50:20: →NAS 5.1 values

# cat /nas/sys/nas\_param

**ckpt:10:100:20:** →Default values 5.5.32.4, same since NAS 5.2

**Note:** First value of ‘10’ indicates the ‘polling rate interval’ for the Control Station in seconds; ‘100’ represents the max rate in MBps that changes will be written to the file system; ‘20’ represents the percentage of space that the system will allocate to CheckPoints & Replicator from all storage [adjust this value up as required but do not go below 20%]. If the 20% threshold is reached, errors may be logged in cmd\_log.err or nas\_log.al.err as ‘Volume is unreachable’ or ‘Volumes are not available’—if this happens, increase the 20% to a higher value.

## **NAS 5.6:**

# cat nas\_param|grep ckpt

**ckpt:10:200:20:**

### **CMD LOG.ERR (nas\_log.al.err):**

Volume is unreachable [Extend 20% storage allocation in nas\_param file]

### **SYS LOG:**

Apr 5 22:05:49 2004 SVFS:0:1 Slot 2: 1081228183: FSID:45 SavVol:260 Full(hwm)reached (t:13547361908090806)

**Note:** Look for messages related to Checkpoints, as in the SavVol HWM message here

### **CMD LOG:**

2004-04-06 23:00:05.779 db:201:18762:S: /nas/bin/fs\_ckpt ufs201\_tue\_0 –refresh

2004-04-06 23:00:08.149 db:201:18762:E: /nas/bin/fs\_ckpt ufs201\_tue\_0 –refresh

**Note:** Look for evidence that Checkpoint refresh started and completed successfully

### **SERVER LOG:**

2004-04-06 23:04:10: SVFS: 4: The filesystem:34 froze.

2004-04-06 23:04:10: SVFS: 4: currctx, SnapshotVolName:Sh26034

2004-04-06 23:04:10: SVFS: 4: oldestctx, SnapshotVolName:Sh26034 :

2004-04-06 23:04:10: SVFS: 4: pause() requested on fsid:27

2004-04-06 23:04:11: SVFS: 4: pause() done on fsid:27

2004-04-06 23:04:11: SVFS: 3: SetSnapShotInactive(): Checkpoint Delete

2004-04-06 23:04:11: SVFS: 4: D34\_0 Permanent InActive

2004-04-06 23:04:12: SVFS: 4: resume() requested on fsid:27

2004-04-06 23:04:12: CFS: 4: Resuming fs 27

2004-04-06 23:04:12: UFS: 4: Cleaners into 12548e80, cg 43dc0640, dirty 12549c40

2004-04-06 23:04:12: UFS: 3: filesystem is read-only: quotas not turned ON.

2004-04-06 23:04:12: ADMIN: 4: Command succeeded: file refresh ckpt 34 260=252 hmw=90

**Note:** Look in Server Log for details. The real PFS is FSID:27 in this example, and the Checkpoint refresh taking place is FSID:34

### **NAS EVENT FACILITY:**

NAS\_DB, Severity = Error, Event ID 211 = missed scheduled task, Event ID 212 = scheduled task fails

### **SCHEDULER LOG: [NAS 5.3]**

**/nas/log/webui/apl\_sched.log**

### **CRONTAB SCRIPTS:**

00 23 \* \* 1-5 /home/nasadmin/scripts/ckpt\_ufs201.sh > /tmp/ckpt\_ufs201.log 2>&1

20 23 \* \* 1-5 /home/nasadmin/scripts/ckpt\_ufs301.sh > /tmp/ckpt\_ufs301.log 2>&1

**Note:** Above example is not a good idea to be running Checkpoint refresh at 11:00 p.m.—reschedule for 11:15 so as not to interfere with NAS DB backups.

### **REVIEW FILE SYSTEM INFO:**

# nas\_fs -i ufs201\_tue\_0

```
id      = 34
name    = ufs201_tue_0
acl     = 0
in_use  = True
type    = ckpt
volume  = vp260
profile = symm_std
member_of = vefs28
rw_servers=
ro_servers= server_2
ckpt_of= ufs201 Tue Apr  6 23:00:07 PDT 2004
used   = 64%
full(mark)= 90%
```

**CheckPoint Information in Server log:** Logged as “SVFS” entries

### **Common Server Log Message for Checkpoint FileSystems:**

2002-06-13 19:51:26: SHADOW: 3: Error IsReadOnly from writeBlock

2002-06-13 19:51:26: SHADOW: 3: Error from write block: 26

**Note:** Normal message for Read-Only file system

## **1. Examine Checkpoint File System:**

```
$ nas_fs -i -s eng-nb [PFS=eng]
id      = 254
name    = eng-nb
acl     = 0
in_use  = True
type    = ckpt
volume  = eng-nb
rw_servers=
ro_servers= titanic
checkpoint_of= engineering Thu Mar 21 11:07:39 EST 2002
used    = 51%
full(mark)= 50%
symm_devs = 002804000222-058,002804000222-038
disks   = d84,d52
disk=d84  symm_dev=002804000222-058 addr=c0t6l1-19-0 server=titanic.....
```

## **2. Verify Status of Production File System and Checkpoints Attached:**

```
# .server_config titanic -v "file display ckpt 218" [FSID from PFS nas-fs -l]
1016817011: SVFS: 3: Checkpoint on FSID:218 (from oldest to the most recent):
1016817011: SVFS: 3: ****
1016817011: SVFS: 3: SavVol:260
1016817011: SVFS: 4: CheckPoint is Active, and Persistant. hwm:50
1016817011: SVFS: 4: FsVol #CFSBlocks:17846272, (at the time the CheckPoint was created)
1016817011: SVFS: 4: SavVol #CFSBlocks:2099561, #CFSBlocks used:1347527, #CFSBlocks free:752034
1016817011: SVFS: 4: BitMapMemUsage:2244656 Bytes, BlockMemUsage:131448 Bytes
1016817011: SVFS: 3: ****
1016817011: SVFS: 3: No more Checkpoint on FSID:218
1016817011: ADMIN: 4: Command succeeded: file display ckpt 218
3. $server config server 2 -v "printstats svfs"
```

Various statistics about checkpoints on the Server

## **4. Examine Server Logs:**

### **Checkpoint File System is Filled and Rendered INACTIVE!:**

```
2002-03-20 14:26:50: SVFS: 3: 2: FSID:23 SavVol:256 Inactive
2002-03-20 14:26:50: SVFS: 3: CKPT_INACTIVE_EVT: FSID:23 SavVol:256
2002-03-20 14:26:50: SVFS: 3: Save Vol full!! Unmounting CKpt file system 250
2002-03-20 14:26:50: CFS: 4: DNLC: erased 0 entries
2002-03-20 14:26:50: CFS: 4: Calling waitForQuiesce, refCounter 16
```

### **Checkpoint High Water Mark is Reached, Logging a Warning in the Server Log:**

```
2002-03-19 19:08:01: SVFS: 3: Recover successful for SavVol:260, FsVol:235 FsType:uxfs
2002-03-20 09:11:39: SVFS: 0: 1: FSID:218 SavVol:260 Full(hwm) reached
2002-03-20 09:11:39: SVFS: 3: CKPT_HWM_EVT: FSID:218 SavVol:260 Full(hwm) reached
```

## **5. Using VVS commands on Checkpoint volumes:**

### **# .server\_config server\_2 -v "vvs list"**

```
1201641099: VCS: 6: There are 34 version sets
1201641099: VCS: 6: Version set 118_800: type=2 refCount=4 numVersions=13 maxVersions=127
1201641099: VCS: 6: Version set 139_1362: type=2 refCount=4 numVersions=13 maxVersions=127
1201641099: VCS: 6: Version set 136_1366: type=2 refCount=4 numVersions=17 maxVersions=127
1201641099: VCS: 6: Version set 133_1370: type=2 refCount=4 numVersions=12 maxVersions=127
1201641099: VCS: 6: Version set 6250:fs6250_T18_LUN2_APM00072500829_0000: type=1 refCount=4 numVersions=1
maxVersions=2002
1201641099: VCS: 6: Version set 6251:fs6251_T18_LUN3_APM00072500829_0000: type=1 refCount=4 numVersions=1
maxVersions=2002
1201641099: VCS: 6: Version set 5086_5104: type=2 refCount=4 numVersions=1 maxVersions=127
```

### **# .server\_config server\_2 -v "vvs listversion vsid=5086\_5104"**

```
1201641217: VCS: 6: numWVersions = 0, maxWVersions = 16
1201641217: VCS: 6: vn=-1 type=1 ref=3 name=5086
```

### **# .server\_config server\_2 -v "vvs getattribute vsid=168\_171 name=RepConsistencySignature"**

```
1254868962: VCS: 7: VVS::getAttrlen : attrName=RepConsistencySignature attrLen=4
1254868962: VCS: 7: VVS::getAttribute attrName=RepConsistencySignature attrLen=4
1254868962: VCS: 6: vvs getattribute: name=RepConsistencySignature, length=4, value=[]
```

### **# .server\_config server\_2 -v "vvs setattribute vsid=764\_768 name=RepConsistencySignature value=0"**

```
1244337295: VCS: 7: VVS::setAttr attrName=RepConsistencySignature attrlen=1
1244337295: VRPL: 7: VPM::doRepTLogIO ioMode=1 ioData=2
1244337295: VRPL: 7: VPM::doRepTLogIO ioMode=1 ioData=1
```

1244337295: VRPL: 7: VPM::setVersionSetRepConsistencySignature  
1244337295: VRPL: 7: VPM::doRepConfigDataIO ioMode=1  
1244337295: VRPL: 7: VPM::doRepTLogIO ioMode=1 ioData=1  
1244337295: VCS: 6: vvs setattribute: name=RepConsistencySignature, length=1, value=0

### **CREATING CHECKPOINT:**

**\$ fs\_ckpt id=pfs1 -name ckpt1 -Create ProdVol -o %full=75** [default is 90%]

### **DELETING CHECKPOINTS:**

**\$nas\_fs -delete pfs1\_ckpt1**

### **PROBLEM DELETING CHECKPOINT:**

**Situation:** File System will not delete, shows as unmounted in data mover database files, yet shows as ‘In Use’ with nas\_fs –l.  
Running nas\_fs –d says that file system is still owned and in use by server\_2.

#vi /nas/volume/filesys

26:fs13\_ckpt1:0:y:7:::1:1061918359::0::24:

**Note:** Change ‘Y’ value to ‘N’ and remove number ‘1’ for server\_2

**#nas\_fs -delete fs13\_ckpt1**

### **ALTERNATE METHOD FOR DELETING CHECKPOINT:**

**.server\_config server\_3 -v "file delete ckpt ckptFSID PFSID SaveVol=FsVol force"**

**\$server\_config server\_3 -v "file delete ckpt 948 102 605=569 force"**

**Note:** 948 is ckptID for ckpt fs 102. 605 is ckpt volume number and 569 is volume number for PFSID

### **PREVENTING MERGE THREAD FROM RUNNING IN DART ON REBOOTS:**

**# .server\_config server\_2 -v "param fulldescription SVFS nomerge"**

SVFS.nomerge 0x020ecf18 0x00000000 0x00000000

**Note:** Setting to 1 will prevent SnapSure’s out-of-order delete process from restarting, useful if needing to delete checkpoints.

### **VIRTUAL ENTRIES ON PFS WITH NAS 5.0:**

Production File System: FS1

Checkpoint Virtual Directory Entry for PFS: .ckpt\_fs1\_ckpt1

New “hidden” Checkpoint directory: .ckpt

**Comment:** Mount the datamover’s file system from Unix Host to view or restore files from the checkpoint directory. Although permissions are inherited from the Parent directory, Checkpoint will enforce the permissions of the files themselves.

### **CHECKPOINT TIMESTAMPS:**

Checkpoints used to display in GMT time, but with NAS 5.1.19.51, timestamp represents local DM time

### **TURNING OFF CVFSv1:**

**param cfs showChildFsRoot=0**

**Note:** Newer Checkpoint style is CVFSv2—CVFSv1 entries are not hidden. A major change with CVFSv2 is that the refreshed checkpoint directory name, as found in the hidden .ckpt under the production file system root, is now updated with the time and date that DART created the refreshed checkpoint, giving the customer a better way to figure out which checkpoint is which on the CS. But note that the “ls –l” Unix or Windows timestamp of the directory itself is related to the last modification time of the root of the PFS itself, not the actual time of the refresh—so the Unix date of the directory could be days or months older than the refresh name itself.

### **Accessing Hidden Checkpoint Directory From Unix Client:**

Step 1. From Unix Host, go to File System mountpoint on DM, then \$cd .ckpt

Step 2. ls –al [Directories are listed by FS ID]

### **Accessing Hidden Checkpoint Directory from Windows Client:**

Step 1. Windows Explorer—Map drive letter “G:” to sharepoint of Production File System to which Checkpoints have been made

Step 2. Highlight Top Level of Drive Mapping from Explorer

Step 3. Go to Explorer Address Bar: Address G:\.ckpt

Step 4. You can now browse the .ckpt folder and any checkpoint folders within

### **Accessing From DOS Prompt:**

C:>\G:

G:>\cd .ckpt [dir to list]

Note: Individual Checkpoints show up as uniquely numbered folders [Jun 12 11:29 06120300.074]

## **CONFIGURING SNMP TRAPS/EMAIL NOTIFICATIONS:**

See SNMP section. Facilities and Event IDs available are as follows:

- 70      1 = HWM SavVol reached
- 2 = Checkpoint has become inactive
- 91      1 = Scheduled CheckPoint creation has failed
- 2 = Scheduled Checkpoint refresh has failed
- 137     10 = Auto-extension of SavVol failed
- 101 = Memory quota exceeded for Checkpoint creation, extension, or replication

## **CELERRA MPFS--MULTI-PATH FILE SYSTEMS old name→“HighRoad”**

### **TWO BASIC ARCHITECTURES:**

- I. EMC Celerra MPFS over Fibre Channel (NS20FC/NS40FC—CX3-40F/Gateways)
- II. EMC Celerra MPFS over iSCSI (NS40 for MPFS—CX3-40C/Celerra with CLARiiON CX3/CX4 arrays/Celerra Gateway with iSCSI MDS—CX3-20C/CX3-40C/CX4-960/CX4-240/CX4-120)

### **GENERAL MPFS INFORMATION:**

- Latest revision of MPFS 5.0 takes the best features of the SAN (High Bandwidth) and combines with the ‘data sharing’ capabilities of NAS, using either traditional Fibre Channel or both Fibre Channel & iSCSI
- Hosts send file requests to Celerra over the IP network (NFS/CIFS), and metadata passed back to client using FMP, but actual data streams are exchanged directly between Host and Storage using SAN FC connection or iSCSI-to-Switch and then to storage (DM does not do iSCSI directly)
- Clariion & Symmetrix storage allowed [Clariion since NAS 5.3]
- Linux & Windows Clients use special MPFSi software for driver and agent
- It’s estimated that 11% of the traffic consists of Control Path communications (metadata exchanges between MPFS client & DM MPFS Server), while the other 89% consists of the Data Path workload between Client and Storage via switch
- Clariion LUN trespass supported natively, does not require PowerPath
- Improved single thread IO performance
- MPFS disk protection module
- Default FMP threads=16, with max FMP threads 128 [Each thread uses 16K memory, so 128 threads=2048K]
- # server\_mpfs server\_2 –set threads=32 (persistent change)
- Default FMP port for Data Mover=4656
- MPFS Unix clients require port 6907 to listen for the MPFS Server, while Windows clients require Port 625

### **LATEST MPFS CLIENT RELEASE:** Dec 2009

v5.0.32.9 Linux

v5.0.92.6 Windows

### **5.0.20.7 LINUX MPFS RELEASE NOTES:** August 2008

- Increased support from 128 to 256 LUNs, though LINUX clients will need to use HVM to mount
- Support for Cisco MDS 9222i switch
- CSA will be integrated into the setup of NS40 for MPFS (NAS 5.6.39.5), and NS20/NS40 FC
- MPFS client support for Fleet arrays

### **HVM HIERARCHICAL VOLUME MANAGEMENT:**

- New tree-based volume structure based on FMP and minimum NAS 5.6 code, turned on by default
- Purpose is to cache more info on file to disk mappings, useful for large files with random I/O access

### **Disabling HVL Feature:**

**# mount -t mpfs -o hvl=0,rw,rsize=32768,wszie=32768,mpfs\_keep\_nfs dm2:/data1/data1 /mnt/**

→Support for Rainfinity Global Namespace Appliance (GNA)

## **EMC CELERRA MPFS: SAN & NAS INTEGRATION**

### **Theory of Operation:**

Application servers are connected to storage via the SAN, as well as to the IP network. The Servers initiate file calls (IO), but the resident MPFS agent intercepts the calls and delivers them the Celerra over IP (NFS or CIFS), with the Celerra acting as a Metadata controller, providing appropriate offset locations for the data. Subsequent Reads, Writes, & Commits are then handled directly by the FMP protocol between the MPFS client and the Storage system using block calls.

→Control Path represents the communication between Host, MPFS driver, and Celerra file systems via NFS, CIFS, or MPFS protocols for metadata and small IO operations

→Data Path represents communication between Host, MPFS driver, and Storage system via iSCSI or Fibre Channel SAN using File Mapping Protocol

1. MPFS Hosts are connected to the SAN & NAS network, and initiate file requests as either an NFS or CIFS client
2. The MPFS agent on the Hosts intercepts I/O calls and sends file access requests to Celerra via CIFS or NFS, and the Celerra retrieves the metadata information from storage
3. Metadata operations are handled by Celerra via CIFS or NFS as a “metadata” controller (CONTROL PATH), providing block allocation and locking for the Client
4. Data Reads, Writes, Commits are handled by FMP protocol directly between Client & Storage over FC or iSCSI SAN using Blocks (DATA PATH)

## **CURRENT NAMING CONVENTION:**

MPFS for iSCSI (aka MPFSi)

MPFS for Fibre-channel (aka MPFS)

MPFS for Windows v5.0 will support either FC or iSCSI MPFS environments

MPFS for Linux v4.3 will support either FC or iSCSI MPFS for RHEL and SuSE

## **LINUX & WINDOWS MPFS VERSIONS:**

→Fall 2008 Mustang MPFS release supports NS20FC (up to 90 drives and 60 MPFS FC Hosts), CSA configuration for MPFS for NS20FC and NS40FC, Fleet arrays connected to Celerra Gateway models, Rainfinity Global Namespace for Linux, Cisco MDS 9222i switch

→March 2008, MPFS v5.0.15.1 for Linux and MPFS 5.0.787.1 for Windows supports either FC or iSCSI MPFS

→March 2008 Corvette release adds support for RedHat 5, Windows XP Fibre Channel & 64-bit O/S support, HP-UX Itanium ia64 64-bit O/S support, SYR enhancements, Rainfinity Global Namespace support for Windows, small file performance improvements, Improved random I/O performance, enhanced diagnostics, fewer installation packages

→Enhanced mpfsctl command to list LUNs

→mpfsinfo tool for Windows and Linux enhanced

→mpfsinq tool enhanced for Windows and Linux

→Supports Clariion arrays using iSCSI combo cards (CX3-20C & CX3-40C)

→Support for Windows XP, 2003, XP 64-bit O/S

→Support for Linux RHEL 5.0

→128-896TB capacity support using NMFS [128TB capacity NS40 Blades; 896TB capacity for Gateways]

## **LIMITATIONS:**

→Snapsure and Replication are not supported

→Cannot mix Symm & Clariion backends

→Connectrix iSCSI-to-Fibre Channel bridges: MDS9216I, 9216A, 9506 & 9509 v1 & v2, 9513 SAN OS 2.1 (2b)/SAN OS 3.0 (2a)

→Combo card NS40 with CLARiiON CX3-20C or CX3-40C backend running Flare 22

## **CURRENT MPFS SOLUTIONS:**

→NS40 Integrated using CX3-40C iSCSI card on the array for iSCSI solution

→NS40 FC Integrated using CX3-40F array for Fibre Channel solution

→NS40G, NS80G, NSX Gateway using iSCSI solution for mid-tier [Client-to-array iSCSI, not Celerra iSCSI]

→NS40G, NS80G, NSX Gateway using iSCSI or Fibre Channel solutions for high-end [MDS iSCSI or FC switch]

## **CURRENT MPFS CLIENT O/S SUPPORT (March 2008):**

RHEL 3 U7+, RHEL 4 U4+, & RHEL 5 U1+ [all for x86, x64, ia64]

Novelle SuSE SLES 8 SP4 (x86), SLES 9 SP3+ (x86, x64, ia64), SLES 10 (x86, x64, ia64), & SLES 10 SP1 (x86, x64, ia64)

Windows 2000 SP4 (x86), 2003 R2 & SP2 (x86, x64), XP SP2 (x86, x64, iSCSI only)

Solaris 8 (SPARC), Solaris 9 (SPARC), Solaris 10 (SPARC, x86, x64)—All are for Fibre Channel support only

IBM AIX 5.2 (PowerPC), 5.3 (PowerPC)—All are for Fibre Channel support only

HP Unix 11i v1 (PA\_RISC), 11i v2 (PA\_RISC)—All are for Fibre Channel support only

## **UPDATED UNIX SUPPORT:** Dec 2009

Version 4.0.23.1 for HP-UX, AIX, Solaris; Support for 10Gbps iSCSI, increased Host support on arrays

## **HOST MPFS AGENT FAMILIES:**

MPFS for Linux v4.0; MPFS for Linux 4.1/4.2; MPFS for Linux v4.3; MPFS for Linux v5.0 (March 2008)

MPFS for Windows v4.0; MPFS for Windows v5.0 (May 2007, but updated with latest Corvette release)

MPFS for Unix v4.0 (no major update since 2004, but updated to support Solaris 10 FC-only)

## **ACTUAL MPFS VERSIONS:**

### **MPFS for WINDOWS:**

EMCmpfs.win.4.0.15.0.exe (Fibre Channel only)

EMCmpfs.win.5.0.12.100.tar.Z (FC or iSCSI)

**Note:** The 5.0 release supports Windows 2003, XP, VMWare, etc

**MPFS for LINUX (Fibre Channel & iSCSI):**

EMCmpfs.Linux.4.3.20.1.tar.Z, 128TB NMFS file systems, GA May 2007

**MPFS for UNIX (Fibre Channel only):**

EMCmpfs.hp-ux.4.0.18.3.tar.gz [AIX, HP-UX, Linux, SunOS]

**MPFS over iSCSI Version 5.0.12.0 for Windows (formerly WINROAD, Windows MPFSi, etc):**

**MPFS over Windows 5.0.12.0:** GA May 2007

- Minimum NAS version 5.4.24.8 & 5.5.27.5 for supporting W2K3 32 or 64-bit OS running MPFS v5.0.12.0 client, & 32-bit XP SP2
- Max. of 128TB storage on Data Movers under a single mountpoint using NMFS, for NAS 5.5.27.5, with maximum single file system size of 16TB [Backend hardware includes NSX, NS40, or NS80]
- Celerra can be configured for MPFS and also serve file systems for NFS & CIFS
- MPFS 5.0.12.0 supports iSCSI initiators with no more than 128 disks underneath the file system
- Note:** MPFS over iSCSI does NOT use the Celerra as an iSCSI Target. The Windows Client is the Initiator and the Connectrix/Cisco MDS SAN switch, or Clariion combo card, is the iSCSI Target and is used as the bridge between iSCSI and Fibre Channel protocols
- Recommended Windows iSCSI initiator version 2.03 or later
- Clients can connect using either MPFS over FC directly, or through the MPFS over iSCSI configuration already mentioned
- Designed to provide high performance for HPC grid computing clients
- Does not support Celerra Replicator or SnapSure on the MPFS Data Mover
- Recommended storage stripe element size is 256, with minimum 32KB
- Supported on Symm 5000, 8000, DMX
- Supported on Clariion CX300/400/500/600/700 CX3-20C CX3-40C with Flare 24 & Access Logic (?—doesn't support CX600?)
- Data LUNs can use RAID 1, 3, or 5 but Management LUNs must be 4+1 RAID 5
- Do not span file systems across Storage platforms
- MPFS does not respect RO options for Celerra mounts & exports (NFS)
- Maximum of 256 SCSI or iSCSI devices presented to Host
- MPFS over iSCSI requires iSCSI-to-Fibre Bridge Switch: Connectrix or Cisco MDS 9509/9506/9513/9216A/9216i
- Documentation will provide Quickstart Guide, Product Guide, Best Practices White Paper, Troubleshooting Guide, Tools
- MCME/MCMI, Penant, PSC, SVC tools

**CELLERRA MPFSi PIHC (MCM-E/MCM-I) TOOL FOR LINUX (Post Install HealthCheck):**

→ New tool November 2006 for verifying an MPFSi installation and configuration according to known Best Practices, that includes Celerra, Storage, Switch, & Client components—Phase I release

**Note:** With NAS 5.6.47 release and CSA support for configuring MPFS on the Data Mover and Linux client, the MCM tool support is being discontinued

→ Verifies things like Cisco MDS 2.1.2-6 firmware, zoning, NAS version, MPFSi drivers, Storage, MPFSi disks are accessible, kernel OS version, Proxy Initiators are evenly distributed on SPs and can be reached, checks AccessLogix mapping of ALUs to switch zone mappings, can run a performance daemon on Clients to obtain statistics, etc.

→ Client & Server portion of tool can run on Windows or Linux (mpfssi\_server.zip & mpfssi\_client.zip) with java1.5 JRE minimum

→ Tool intended for Field & Service engineers

→ MPFSi Monitor gui allows easy creation of a site, where the IP address & logon info for CS & Switches are entered, at which time Celerra DM's, Storage, Switch, File Systems are discovered automatically. Clients need to be added via Hostname or IP address

**PIHC Server Application:**

Used for Hosts, Celerra, Storage, & Switch information in the GUI program for the ‘site’

To run this service on a Windows client, do the following:

1. Unzip to a folder: mpfssi\_server.zip
2. JAVA\_HOME=c:\jre1.5.0\_08
3. MPFSI\_SERVER\_HOME=c:\mpfssi\_server
4. Run server.bat file

**PIHC Client Application:**

Allows reporting of config info from Linux clients, and client performance monitoring daemon

1. Unzip to a folder: mpfssi\_client.zip
2. JAVA\_HOME=c:\jre1.5.0\_08
3. MPFSI\_CLIENT\_HOME=c:\mpfssi\_gui
4. Run client.bat

**PHASE II—January 2007:**

→ Add a report function so as to be able to output the configuration and parameter settings seen in the GUI, to a report file

→ Add a scheduler function for the performance statistics, which will be useful for troubleshooting performance issues

→ Importing Host information into GUI

→ Auto configuration capability

**MCM-I/MCM-E—MPFS Configuration Manager:**

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
New name for the old PIHC tool. MCM-E is foreengineering mode (more capability), accessed same manner as Navisphere on Clariion: ctrl + shift +f12, password messner. MCM-I is for customers. Tool is used for Site Setup/Configuration, Backend & Host information, and Management of Hosts.

## **MPFS TOOLS—PENANT Pre-Sales Existing Network Analysis Tool:**

Engineering has developed a Penant (Pre-Sales Existing Network Analysis Tool) tool that will use Ethereal or Wireshark sniffer (tshark is the default) to capture network traffic, then performs a text readout that characterizes the IO in the environment. Penant.pl script will total and provide % on Reads vs. Writes for CIFS & NFS traffic, analyzes access patterns (Random vs. Sequential), analyzes amount of Data vs. Metadata transmitted, file sizes, IOPs, and overall throughput. Tool will also become useful during CS troubleshooting engagements. There is the capability to limit actual data captured by using snaplink, which truncates data packets so as to obtain more manageable captures.

## **PRE-SALES CONFIGURATION TOOL—CSC Celerra Solution Configurator:**

XLS spreadsheet tool to assist in validating and building a basic MPFS model, based on throughput needed and number of hosts in a given environment

### **MPFSi 5.0 Windows Client Setup (Win2k3 & XP):**

1. Install and connect Initiator to Target
2. Install MPFSi client software and reboot
3. Check via Explorer to see if MPFS Volume Properties are displayed & enable the MPFS shares

**Note:** Default connection uses TCP/IP, 8192 block size, make sure connection is checked Persistent; Strict Flush Policy would apply only to special applications that require each IO be committed

4. Console checks:

C:\mpfsctl version | list | stats | reset

C:\net view \\compname

C:\mpfsctl enable <compname> <sharename> | disable

### **EMC Celerra NS40 for Celerra MPFSi:**

**Note:** The NS40 used for MPFSi with the Clariion CX3-40C combo card, is a standard NS40 Integrated, pre-Napa 8 model

→May/June 2007 EMC released a standalone NS40C Integrated version of the NS40 that is configured with the iSCSI combo cards on the dedicated backend array. The actual NAS models are the NS41C-A or B or -FD, or NS42C-A or B or -FD models, which are really standard NS40 Integrateds, but targeted only for MPFS and sold with the NS40-AUXC (EMC rack) or NS40-AUXC-FD (field racked) CX3-40C array. Supports up to 120 MPFS iSCSI hosts. The Celerra itself does not serve as an iSCSI Target.

### **NS40C INSTALLATION:**

→The NS40C installation is really divided into two separate sections—the routine NS40 Integrated Celerra installation, followed by the MPFSi configuration of the Clariion storage system and MPFSi Clients

→Use the following Powerlink destination as the primary source of NS40 MPFSi information:

### **Celerra MPFS over iSCSI>Installation/Configuration:**

[http://powerlink.emc.com/km/appmanager/km/secureDesktop?\\_nfpb=true&\\_pageLabel=servicesDocLibPg&internalId=0b0140668014e854&\\_irt=true](http://powerlink.emc.com/km/appmanager/km/secureDesktop?_nfpb=true&_pageLabel=servicesDocLibPg&internalId=0b0140668014e854&_irt=true)

→The “EMC Celerra MPFS for NS40/NS40FC Integrated (Linux and Windows Clients) Quick Start Guide Oct 2007” provides information and references the Engineering documentation guides for more complete installation of the NS40 Integrated [Celerra Network Server NS40/NS40 DC NEBS Integrated Configuration Phase 1 and 2 Setup Guide and Celerra Network Server Integrated and Gateway Configuration Phase 3 Setup Guide]

→Linux or Windows clients install the MPFS client, connect to the Celerra for FMP metadata transactions using normal network IP, and have direct connections to the backend array via iSCSI for data transfers.

→The “EMC MPFS NS QuickStart Practitioner’s Guide July 25, 2007” provides more in-depth background in how to best implement the Celerra solution for MPFSi

### **MPFS PARAMETERS/RESTRICTIONS:**

- Minimum stripe size is 32k--volume not 8k aligned
- MPFS cannot support more than 4000 devices or use a disk volume with a number higher than 4000
- NAS 4.1.4.0 Requires File System stripe sizes of 32K or greater in order to mount as MPFS Volumes [32768]
- Quotas are not supported with MPFS
- Checkpoint Snapshots are not supported with MPFS
- Striping on SYMMETRIX is recommended, not on the Celerra! [Use 128k striping]
- Symm RDF supported but not Celerra SRDF
- TimeFinder supported for 3.0.16.x and above
- Use VolumeLogix for Fibre Channel installations

### **LINROAD TOPOLOGY→NEW IP SAN & IP NAS INTEGRATION:**

Similar as conventional MPFS, but builds on the availability of iSCSI technology, NFS or CIFS Clients use MPFS driver to intercept traffic and communicate to HighRoad Server, but are iSCSI-connected to Connectrix MDS9509 with iSCSI IPS module, Symmetrix or Clariion storage for high-speed data transfer, with meta-data again transported over IP Network using FMP protocol.

→Phase I using just Linux Clients will be based on RH3/SuSe 8.0

→Not a Celerra iSCSI solution

→NAS 5.4 & higher

## **CELERRA MPFS FOR iSCSI:**

→MPFSi v4.1.10.0 supports Linux RHEL 3 [Linux iSCSI initiator & Connectrix MDS 9509 or Cisco Catalyst 6509 switch with iSCSI bridge to iSCSI target]—uses only the IP-SAN switch iSCSI target, not the Celerra, CLARiiON, or Symmetrix iSCSI target

→MPFSi v4.2.12.0 supports Linux RHEL 4

→MPFSi v4.3.13.0 (Jan 2007) supports Linux RHEL 3 & RHEL 4

→Minimum one Connectrix MDS switch for FC to iSCSI bridging using SAN OS 2.1(2b) or 3.0(2a)  
Connectrix MDS 9216A, 9216i, 9506 & 9509 v1 & v2, 9513

→Both Symmetrix & Clariion backends supported, with latter using RAID 1, 3, or 5 Storage Groups

→Cannot use both MPFS & MPFSi clients on the same Host

→NAS Versions 5.4.24.5 & 5.5.22.0 support MPFSi 4.1 – 4.3 clients

→NAS Version 5.3 minimum for MPFS 4.0 clients

→iSCSI target is the SAN Switch or Clariion, not the Celerra or Storage system

**Note:** Only the NS40 Integrated is qualified with the Clariion backend (CX3-40C) for MPFS over iSCSI

→Max. of 256 luns per Host initiator for Linux RHEL 3, but 300 luns RHEL 4

## **HOW MPFS WORKS:**

→Clients connect to Storage via FC/SAN or iSCSI/FC

→Clients connect to out-of-band “metadata” controller, in this case the Celerra Server

→MPFS layer separates Metadata FS Information [Control Data] from the Data FS [Payload Data] by intercepting Client Opens, Closes, Reads, & Writes

**Note:** MPFS Client intercepts file write call and asks Data Mover to allocate blocks on disk for the file, and after extent list is sent back to the client, the client writes directly to storage disk, then informs Server when complete

→All non-data Client requests use NFS/CIFS layer for metadata information which includes file block map, attributes, directory info

→FMP—File Mapping Protocol, is used to exchange file layout information between MPFS client and Data Mover using NFS or CIFS

**Note:** The FMP Protocol sits on top of Network Access Protocols serving as a Redirector

→For SAN, metadata is transferred over IP Network, while data is transferred over SAN (MPFS over SAN)

→For IP-SAN, metadata and data is transferred over IP Network (MPFS over iSCSI) using either Cisco/Connectrix MDS Switch, or directly from the Clariion using CX3-20C/CX3-40C FC/iSCSI combo model. iSCSI requests are translated into FC Channel Data Blocks (CDBs) by the Clariion or MDS Bridge.

**Note:** A private network could be setup with a separate NIC for iSCSI traffic

**iSCSI TARGET:**                   **iSCSI INITIATOR:**

MDS Switch or

MPFS Client

Clariion array

## **Linux Client Tweaking:**

Increase max read-ahead kernel param from default 31 to 8192 by adding following line to /etc/sysctl.conf

**vm.max-readahead = 8192**

## **Linux Configuration:**

Should already have iSCSI initiator drivers installed, but if not, use iscsi-initiator-utils

Add following lines to /etc/iscsi.conf to identify port on MDS switch

DiskCommandTimeout=10

DiscoveryAddress=172.17.2.22

## **Troubleshooting:**

→Check that iSCSI virtual-target definitions include all LUNs for both SPs

→MDS Switch does not allow iSCSI port WWN to see both SPs

→Accesslogix Storage Group does not allow iSCSI port to see all LUNs from both SPs

Run iscsi-ls to see if client can see both SPs

Run mpfsinq to see if all dvols can be seen

Check /var/log/messages fil

Run mpfsdiscover to re-discover Celerra LUNs

## **MPFS CLIENT COMMANDS/CONFIG INFO:**

# **mpfsinfo -v <mpfserver>**

# **mpfsstat -d 2**

**Note:** info on I/Os issued to disk

# **mpfsinq** (active & passive paths to storage)

# **cat /proc/mpfs/params** (MPFS parameters)

# **cat /proc/mpfs/devices** (kernel devices)

# **cat /proc/mpfs/hrdp** (disk protection)

**/etc/mpfs.conf** (MPFSi configuration file, kernel parameters such as globPrefetchSize)

**/etc/sysconfig/EMCmpfs** (mpfs discover daemon & parameters—runs every 15 minutes to check for paths & trespassed luns)

#### **MPFSi Improvements:**

Extent-based flush to optimize metadata flushes to reduce Server overhead (dirty pages maintained until committed via FMP flush)  
FMP protocol enhancement to offload the volume tree traversal work to the Clients and not the Server

FMP Clients also make better use of cache to reduce RPC calls

CIFS Clients can now better determine IP Address context for multi-homed clients

#### **BEST PRACTICES INFORMATION:**

→Do not use AVM, setup User-defined storage pools

→RHEL3 supports 255 scsi devices; RHEL4 supports 300 scsi devices

→Use TCP protocol

→RAID 3 is recommended for ATA drives on Clariion

→Configure CX3 Clariion with Read cache size 256MB/SP, Write cache size 2938/SP, water marks 70% low 90% high, page size 8KB, and use 64KB stripes in (2) striped volumes across (2) pairs of (9) Luns, meta-concatenated

→For DMX-3, use 24 volumes on 24 separate disks and create 256KB striped volumes

→Create Celerra filesystems using stripe size 256kb for CLARIION and 32kb for Symmetrix

→Load balance LUNs across SPs

→Change Linux iSCSI initiator timeout in /etc/sysconfig/iscsi for reboots: ESTABLISHTIMEOUT=60

→iSCSI Best Practices Guide Connection Failure timeout values in /etc/iscsi.conf file are “ConnFailTimeout=45”

→Do not run CAVA or use multi-homed NICs on Windows system running MPFSi client

→File Systems built on either all FC or all ATA drives, not mixed [Using ATA & LCFC drives are NOT recommended]

**\$ ./ftplibboot/setup\_backend/zone -s 2 showdm** (command to collect WWN's for DMs)

Checking for Data Mover 2.

Collecting Data Mover WWNs...Done.

DM List:

DM Port:20 WWN:50:06:01:60:90:60:16:f8:50:06:01:60:10:60:16:f8

DM Port:21 WWN:50:06:01:60:90:60:16:f8:50:06:01:61:10:60:16:f8

DM Port:22 WWN:50:06:01:60:90:60:16:f8:50:06:01:62:10:60:16:f8

DM Port:23 WWN:50:06:01:60:90:60:16:f8:50:06:01:63:10:60:16:f8

**\$ ./zone -spa 10.241.168.55 -spb 10.241.168.56 showsp** (checks IP connectivity & collects WWN's SPs)

Checking network connectivity to 10.241.168.55 / 10.241.168.56 ...Done.

Collecting backend WWNs...Done.

SP List:

SP Port:A0 WWN:50:06:01:60:90:60:05:10:50:06:01:60:10:60:05:10

SP Port:A1 WWN:50:06:01:60:90:60:05:10:50:06:01:61:10:60:05:10

SP Port:A2 WWN:50:06:01:60:90:60:05:10:50:06:01:62:10:60:05:10

SP Port:A3 WWN:50:06:01:60:90:60:05:10:50:06:01:63:10:60:05:10

SP Port:B0 WWN:50:06:01:60:90:60:05:10:50:06:01:68:10:60:05:10

SP Port:B1 WWN:50:06:01:60:90:60:05:10:50:06:01:69:10:60:05:10

SP Port:B2 WWN:50:06:01:60:90:60:05:10:50:06:01:6a:10:60:05:10

SP Port:B3 WWN:50:06:01:60:90:60:05:10:50:06:01:6b:10:60:05:10

**Add following entry to /etc/auto.misc file to allow auto-mounter to mount MPFS file systems:**

mnt -fstype=mpfs /server/fs

#### **SUPPORTED MDS SWITCHES:**

MDS 9506; MDS 9509; MDS 9513; MDS 9216A; MDS 9216i (Cisco MDS SAN-OS Release 2.1(2b))

MDS 9506v2; 9509v2; 9513 and MDS SAN-OS 3.0

#### **MDS 9000 Best Practices:**

→Use proxy-initiator mode only

→Use store & forward mode only

→Use dynamic target mapping

→Use iSCSI TCP send buffer size of 16384

### **HIGHROAD DATA PROTECTION (HRDP):**

Devices that are configured for the MPFS service at boot time are protected from use by other programs using HRDP. To verify devices & protection seen by Linux host after bootup:

```
# mpfsinq  
# dd if=/dev/sdc count=1  
dd: opening '/dev/sdc': No such device
```

**Note:** Use dd to test if disk is protected by HRDP—above message return indicates device is protected

**# /usr/sbin/hrdp**

**Note:** Use above command to protect devices added after bootup

### **NOT SUPPORTED BY MPFS:**

--VirusChecker; Checkpoints; Quotas; CDMS; SRDF; Celerion FC4700-2 Storage; File Locking Policies; FSN/VLAN; I18N

--Fiber Channel Control Stations are not supported

--With NAS 5.2 and higher, Checkpoints can be used on file systems that are also in use for MPFS, though MPFS is not supported for checkpoints themselves.

### **MPFS UNIX/WINDOWS CLIENT & SW SUPPORT**

→PowerPath 4.5.x – 5.0.0 required for CLARiiON only, optional for Symmetrix

→Supports all Symm & Clariion backends, latter with RAID 1, 3 or 5 Storage Groups

→MPFS 4.0 clients do not support iSCSI for Windows or Linux

→MPFS 4.0 clients require minimum NAS Version 5.3.10.4

### **MPFS WINDOWS CLIENT Win2k/Win2k3 Fibre Channel Only:**

Version 4.0.15.0 with NAS 5.4 & 5.5 support, PowerPath 4.5.0, 4.5.1, etc.

### **MPFS CLIENTS FOR UNIX HOSTS 4.0.18.x/4.0.19.x (min. code 5.3.10.4 and above) Fibre Channel:**

NAS MPFS for Unix Clients 4.0.19.1

EMCmpfs\_aix\_4.0.19.1.iso

EMCmpfs.hp-ux.4.0.18.3.tar.Z

EMCmpfs.linux.4.0.18.3.tar.Z

EMCmpfs.sunos.4.0.18.3.tar.Z

**IRIX 6.5.9:** EMC HighRoad client 3.1 with NAS 3.0.21.3 or 3.0.22.0

**WINDOWS NT 4.0/2000:** Latest Client: EMCmpfs.win.4.0.13.1.exe

### **POWERPATH CLIENT SW:**

Version 3.0.5 NT and 3.0.6 for Windows 2000/2003

Red Hat Linux version 3.0.6 RHEL 32-bit; Improve Read performance by adding line to /etc/sysctl.conf vm.max-readahead=8192

Solaris 7, 8, 9 and HP-UX 11.11 version 4.3

### **VolumeLogix: Version 2.2.1**

**Note:** Direct-attached via SCSI or Fibre Channel. Startup scripts go into rc1.d & rc2.d. (1) mpfsd daemon runs. Add MPFS filesystems to /etc/vfstab [192.1.4.21:mpfssun - /mpfs mpfs - yes - ]. Be aware that most clients have required patch levels.  
[http://powerlink.emc.com/support/sw\\_downloads/index.jhtml](http://powerlink.emc.com/support/sw_downloads/index.jhtml) [go here to download latest MPFS clients]

### **BASIC MPFS over iSCSI CONFIGURATION STEPS:**

1. MPFS client with iSCSI HBA or sw initiator driver
2. MPFS client with single or dual NICs (NFS or CIFS for one; iSCSI for other)
3. MDS switch with FC/iSCSI configured, or Clariion with FC/iSCSI configured
4. Install MPFS client
5. Configure Data Mover and enable MPFS on file system
6. Start MPFS service on Client

### **Linux TCPDump Trace:**

Step 1. #/usr/sbin/tcpdump -s 2000 -w /tmp/dump1 host 192.1.5.44 and 192.1.5.23

Step 2. Starts capture between the two nodes; Use “ctrl + c” to stop capture

Step 3. Display Capture: #/usr/sbin/tcpdump -vvex -r /tmp/dump1 |more

### **To List MPFS File Systems:**

```
$ mount -v |grep mpfs
```

### **List Package Info:**

```
$ pkginfo |egrep -i EMCmpfs or $ pkginfo -l EMCmpfs
```

### **Look for MPFS Service:**

```
$ ps -ef |grep mpfsd
```

### **Mounting & Specifying UDP:**

```
$ mount -F mpfs -o proto=udp servername:/usr/src /usr/src
```

**MPFS Share Properties:** Defaults to “Asynchronous Flushing” for client writes. Can change to “Strict Flush Policy” if flushing after every write is important. Default “Block Size” is 8192 bytes. If SRDF used, then “Read-Only” too.

**MPFS Server Setup:**

**# server\_setup server\_2 -P mpfs -o start=n** {-o stop} [where n= number of threads to use—default=16]

**MPFS Threads:** Max.=128 FMP Port=4656 Adding, Deleting, or Setting:

**\$ server\_mpfs server\_4 -set threads=50 {-a 16; -d 16}**

**Note:** Stopping the MPFS service requires about 30 seconds for any Clients to timeout so that the service can be stopped

**Interfaces:** Process RIP—view stats using “server\_rip” command

**TAPE BACKUP SUPPORT FOR MPFS 3.0:**

Use “server\_archive” command to backup using PAX and NDMP to tape drives.

Add Following to Param File on DataMover: /nas/server/slot\_2/param

**param NDMP nbuf=<number for ndmp>** [Derived by # tape drives multiplied by 5]

**param PAX nbuf=<number for PAX>** [Derived from # tape drives multiplied by 4]

**Highroad Network Backup Support:** EMC Data Manager (EDM) 4.6 +

**SETTING UP MPFS LINUX CLIENT FOR iSCSI:**

1. Download & untar Linux MPFS client

2. ./install-mpfs [Installs, starts daemon, and discovers devices]

3. Verify rpm version & daemon:

# rpm -q EMCmpfs

# ps -ef |grep mpfssd

4. Enable MPFS license:

# mpfslicense <16\_digits>

5. Configure iSCSI:

# vi /etc/iscsi.conf and add following to end of file

PingTimeout=30

DiskCommandTimeout=30

PortalFailover=no

HeaderDigest=never

DataDigest=never

InitialR2T=no

ImmediateData=yes

DiscoveryAddress=x.x.x.x [Clariion iSCSI port configuration

# vi /etc/initiatorname.iscsi

# GenerateName=yes →Comment this line out as shown

InitiatorName=iqn.2006.com.emc.mpfsi: <linux\_name> →Add this line to file

6. Start iSCSI service & grep for iscsi process:

# service iscsi start

7. Configure disk protection

8. Verify:

# mpfsdiscover -v

9. Mount Celerra file system from Linux client & verify:

# mount -t mpfs -o tcp,mpfs\_keep\_nfs, rsize=32768, wsize=32768, noatime xxx.xxx.xxx.xxx:/export\_name /mpfs\_celerra\_mount

# mount # mpfsinq # iscsi-ls # cat /proc/mpfs/devices [Celerra devices] # cat /proc/hrdp [Disk protection] # /usr/sbin/hrdp [run to protect devices]

**SUN SOLARIS MPFS CLIENT:**

O/S: 2.6, 2.7, 2.8 #showrev -p #uname -r

Default Protocol on Sun is TCP

MPFS client software is 3.0.20\_b001 for MPFS 3.0.16.5+

**SETTING UP SOLARIS AS MPFS CLIENT:**

1. Install Software:

#/usr/sbin/pkgadd -d /cdrom/8.0 [Enter 1 at prompt, then License Key, Reboot]

2. Map Symmetrix Data Volumes to Solaris “sd.conf” file: [Values derived from BIN file]

#vi /kernel/driv/sd.conf

3. Reboot Solaris System

4. Label new Disk Drives using #format command [Labels disk only for Solaris internal use]

```
--Use #mpfslabel command to allow Solaris to access raw disk, including last (2) cylinders, if required  
#prtvtoc [Use this to verify label prior to running mpfslabel command]  
#mpfslabel /dev/rdsk/c1t0d16s2  
#prtvtoc [Use again to verify mpfslabel command]  
5. Verify MPFS Package Install & Daemon are running:  
#pkginfo -l EMCmpfs  
#ps -ef |grep mpfsd  
6. Mount Solaris to MPFS File System on DM:  
#mount -F mpfs -o rsize=32768,wsize=32768 192.10.4.28:/mpfs1 /mnt  
7. Verify Mounted MGFS FileSystem:  
#mount -v |grep mpfs  
#mpfscctl [Utility to Tune & Manage MPFS]  
#mpfscctl stats [MPFS Stats: /etc/fs/mpfs directory]  
#inq [Outputs info on MPFS Symmetrix devices]
```

**Note:** Information resides in: #/kernel/driv/sd.conf

## **LABELING MPFS VOLUMES FOR SUN SOLARIS:**

**/var/adm/messages** File Error: “Corrupt Label—Wrong magic number”

Means that the mpfs devices were not labeled correctly when setting up the SUN client and need to be labeled:

### **Step 1. LABEL VOLUMES FIRST FOR SUN HOSTS:**

- a.) #format [allows you to see mpfs devices from the Sun client]
- b.) format>label Y format>quit [Label each mpfs volume manually!!!!]

**Note:** Labeling the MPFS volumes from the Sun Client is harmless & will not destroy disk label or data

**Warning:** Failure to ensure that this is done will result in poor MPFS performance

### **Step 2. LABEL VOLUMES FOR MPFS:**

- a.) #mpfslabel /dev/rdsk/c1t0d16s2
- b.) Verify using PRTVTOC command: #prtvtoc /dev/rdsk/c1t0d16s2

## **TROUBLESHOOTING SUN MPFS CLIENTS:**

**Software Installation:** #pkginfo -l EMCmpfs

**Verify Process Running:** #ps -ef |grep mpfsd [/etc/fs/mpfs/mpfsd]

**Verify Disk Labeling:** #format #inq

**MPFS Client Version:** #mpfscctl version

**Sun Version:** #showrev -p

## **MPFS STATISTICS: # mpfscctl stats mpfscctl reset** [look for 'fallthroughs' to NFS and 'RPC errors']

Use “mpfscctl stats” to display statistics showing internal operation of the EMC HighRoad client. By default, statistics accumulate until system reboot. Use mpfscctl reset to reset the counters to 0 before executing mpfscctl stats.

**Logs:** /var/adm/messages [If var gets filled, MPFS performance will suffer] /etc/system /kernel/driv/sd.conf nfsstat

**Traces:** Use TCPDump or NetMon for Traces

## **USING MPFSCOLLECT TO GATHER ESCALATION INFO FOR MPFS:**

# mpfscollect /fs1/trace.mpfs 65536 & (binary output—run the trace during problem or problem recreation on clientside)  
# pkill mpfscollect (kill the trace process after required information is collected)

## **MPFS COMMAND SET:**

**\$ server\_setup server\_3 -P mpfs -o start=32 | stop | delete**

```
$server_mpfs ALL [FMP Threads=16 [default] FMP Open Files=468 FMP Port=4656]  
$server_mpfs server_2 -set threads=128 [Setting Threads for MPFS]  
$server_mpfs server_2 -Stats  
$server_mpfs server_2 -S file=filename [Stats for an MPFS file]  
$server_mpfs server_2 -add 16 | -delete 16 [Adding or Deleting MPFS Threads]  
$server_mpfs server_2 -default [Resets MPFS variables to defaults]  
$.server_config server_2 -v "param ip"  
$.server_config server_2 -v "param nfs"
```

**\$ server\_mpfs server\_2 -mountstatus** [Verifies mounted filesystems are MPFS compliant]

\$server\_mpfs server\_2 -add 64 | -delete 32 | -Default [to revert to default threads] | -Stats [HighRoad Statistics]

**\$ server\_mpfsstat server\_2 | ALL | -z (reset) | -file filepath | -session sessionid | -list**

**Note:** Many new options added. –sessions will show MPFS statistics by session

### \$ server\_mpfsstat server\_2 -list

server\_2 :

```
-----  
Active MPFS sessions  
(clientid/timestamp)
```

```
10.170.36.35      0 sec      0 usec
```

**Note:** -list shows active MPFS sessions on the Data Mover

### \$ server\_mpfsstat server\_2 -file fs1/test/kl44444.tmp

**Note:** Above cmd looks at statistics for specific file

|           |          |                                                                                  |
|-----------|----------|----------------------------------------------------------------------------------|
| Total     | avg msec | high msec                                                                        |
| getMap(): | 22530    | 1.08                                                                             |
|           |          | 164 → This row shows the Read I/O activity—getting FMP locak and “map” of blocks |

**Note:** Locks are shared Read

### \$ server\_mpfsstat server\_2 -session

getMap(): →Relates to Read activity; FMP locking and mapping of blocks being read; locks are shared for reads

### \$ server\_mpfsstat server\_2

allocSpace(): →These two values show Write IO traffic statistics; FMP locking and allocation of blocks for writing, locks exclusive  
commit():

### \$server\_mpfs server\_6 -mountstatus

server\_6 :

```
fs      mpfs compatible?  reason
```

```
-----
```

```
cw_fs46    mounted
```

```
fs_tcpdump yes
```

```
cw_fs30    mounted
```

```
cw_fs26    mounted
```

```
cw_fs22    mounted
```

```
root_fs_common yes
```

```
root_fs_6   yes
```

### \$ server\_mpfs server\_6 -S

server\_6 :

Server ID=server\_6

FMP Threads=16

Max Threads Used=7

FMP Open Files=3

FMP Port=4656

HeartBeat Time Interval=30

### \$ server\_nfsstat server\_2

**Note:** Output useful to see how metadata calls are going and also to see if data has fallen through to NFS—see v3read, v3write, etc.

#### Debugging Windows Client Issues:

**\$ .server\_config server\_2 -v “logsys set severity FMP=LOG\_DEBUG”**

**\$ .server\_config server\_2 -v “logsys set severity SMB=LOG\_DEBUG”**

## WINDOWS 2000 MPFS COMMAND SET:

|                              |                                                    |
|------------------------------|----------------------------------------------------|
| C:>mpfscctl /?               | [Help info on MPFS]                                |
| C:>mpfscctl enable   disable | [Establishes session & sets MPFS share properties] |
| C:>mpfscctl stats            |                                                    |
| C:>mpfscctl list             | [List of MPFS shares]                              |
| C:>mpfscctl version          |                                                    |
| C:>mpfscctl log on   off     | [Enables logging]                                  |
| C:>mpfscctl stats            | [MPFS Client Statistics]                           |

## MPFS PERFORMANCE LIMITATIONS:

Broad rule-of-thumb is that you should see a range of MPFS performance from **20-30Mbps**

**Note:** Ensure that the /var directory is not full if logging is set [re-direct logging to another location]

## ENHANCING UNIX/SOLARIS PERFORMANCE WITH MPFS:

Step 1. Mount Solaris Client to MPFS File System:

**# mount -F mpfs -o rsize=32768,wszie=32768 server\_2:/src /usr/src**

Step 2. Edit Celerra slot\_param or server param file & reboot:

**param ip setMsgSize64k=1** [configures large packet support]

### **param nfs v3xfersize=32768**

**Note:** Some switches do not support 64k size—use this param to change transfer size of data packets for NFSv3 for Reads and Writes

**# .server\_config server\_3 -v "param fulldescription ip setMsgSize64k"**

ip.setMsgSize64k 0x010ca978 0x00000001 0x00000001

### **NEW DEFAULT V3XFERSIZE NAS 5.5:**

param nfs.v3xfersize 0x00008000 0x00008000 [Hex 8000 = decimal 32768]

### **INCREASED NFSv3 PERFORMANCE:** Celerra NAS Code 2.2.37.1+

NFSv3 Transfer Sizes of 32K now supported

TCP Transmit High increased to 128kb; transmit Low now 64kb

Use server\_nfsstat and look at v3getattr, lookup, access, info statistics for MPFS-related information

Use "v3xfersize" values of 8192, 16384, 32768, 65536 for NFS Version 3 Reads & Writes

### **Changing the max number of blocks cached by NFSv3 asynchronous writes [default value is 32]:**

#### **param file asyncthreshold=32 (default)**

**Note:** Parameter related to NFSv3 asynchronous writes for filesystems and has to do with buffering for backend I/O. Default value=32. Setting value to (1) would impact system performance by flushing changes to disk more frequently, hence impacting Write Performance on ‘NFS’ Data Movers—NFS typically uses larger file transfers and a setting of 1 would slow down writes because each write action would have to be acknowledged. With the default value set, up to 32 write actions could be sent without waiting for a Celerra acknowledgement. Some sites may increase this value to (256) per Best Practice for Performance for NAS 5.5 to enhance NFSv3 performance, but this may destabilize the system and cause panics.

**Caution:** With Sun Solaris Clients (supports only 16384 transfer rates), might want to set the Celerra to also use same rate

**\$server\_config server\_x -v "param nfs v3xfersize=16384"**

### **NFSv2 PERFORMANCE:**

max rsize/wsize for NFS v2 is 8192 bytes

NFS v2 does not support 32768 byte read/write size

Confirm with nfsstat -m (if it is supported) or in a network trace

### **NT4.0/WINDOWS 2000 MPFS CLIENTS:**

--Uses UDP protocol by default

--Must use IE 5.01 with PowerPath 2.1 to prevent NT from writing disk signatures to MPFS volumes

--Must NOT write disk signatures on SYMM volumes from NT! Use PowerPath to protect SYMM volumes

[If mixed Sun/NT clients, label the Sun disks first]

--FileSystems are restricted to no more than (6) SYMM volumes for code under 3.0.15.19

--NT can access up to 120 disks on a single SCSI channel

--User limit feature not supported in NT Server Manager for Celerra MPFS

**Note: MPFS Client Software is Installed on NT Clients that will be used to Access FileSystems**

--MPFS support driver is added under “devices” and MPFS Helper is setup as an NT “service”

--Both R-1 and R-2 Symms are connected directly to the NT Servers via UW SCSI

Step 1. Install NT Client software for MPFS

Step 2. Ensure that PowerPath client for NT is also installed [prevents NT from writing signature to MPFS Volumes on Symm!!]

**Note:** Not needed for strictly NFS or CIFS environments.

Step 3. Create MPFS share from prompt or via Server Manager: c:\mpfsctl enable 193.1.21.182 mpfsshare

Step 4. Turning Logging On for Troubleshooting: c:\mpfsctl log on {off}

**MPFS DEFAULTS FOR NT:** UDP Protocol \*MPFS Enabled Block Size: 8192

**MPFS Client Setup for Unix/Solaris:** Requires direct SCSI connection to Symm & Solaris patch required if version 2.6 or lower

Step 1. Setup the “sd.conf” file so as to match the Symmetrix Device list for TID’s—ie, add devices to the table

/kernel/driv/sd.conf #vi sd.conf and add appropriate entries [see following for syntax]  
“name=”sd”class=”scsi”; target=1 lun3;

Step 2. #reboot -- -r

Step 3. Run #format command to format each new disk device and then “label” each device

Step 4. #reboot -- -r [if configured correctly, should see no errors on bootup]

Step 5. Add Unix Client Software for MPFS: Mount CD-ROM: #mount -F hsfs -o ro /dev/dsk/c0t2d0s0 /cdrom

Use Package Add Utility: #/usr/sbin/pkgadd -d /cdrom/SOLARIS/sunos-5.8

The following packages are available:

1 EMCmpfs Multiplex File System [Add License number--Installs to /etc/fs/mpfs] [\$pkginfo -l EMCmpfs]

Step 6. Mount the remote Celerra File System as MPFS:

**#mount -F mpfs -o 193.1.21.182:/mntmpfs /mnt**

**Note:** Mount NFS for troubleshooting [/var/adm/messages]. Check NFS & Disk connectivity using “format” or “inq” command  
**MPFS Client Servers set “MPFS Properties” on the mapped Celerra FileSystem Share [FS1]**

Enable MPFS    SRDF    READ Only    TCP/IP   Status: Connected

Valid Traffic: Yes   Server Name: dm32-ana0   Share Name: fs1

**MPFS Client Servers setup Remote Share Access Properties at Command Prompt:**

- a.) mpfsctl srdf add 8701E 91006 {lside rside}
- a.) mpfsctl enable dm32-ana0 fs1 SRDF RO [this is the ‘share’ command for mpfs]
- b.) mpfsctl list {stats}

### **MPFS QUOTA SUPPORT:** Introduced with NAS 5.0

Server must reserve file system blocks for Quotas via ‘allocSpace’ requests from Clients. If there are no other Shares or Write Locks on the resource, the MPFS Client obtains Write-lock from Server, conducts Writes, informs Server of completion.

**MPFS Block Policy Quotas:** Server reserves blocks for Client via allocSpace requests, issues Quota Exceeded Error to Client

**MPFS File Size Policy:** Server reserves quotas during allocSpace based on potential file size changes. Server keeps track of highest pre-allocated block offset and last block write-locked by Client.

### **MPFS AV SUPPORT:** NAS 5.0

Since FMP protocol conducts no Reads or Write operations, MPFS Client handles I/O by reading data directly from device mapping list (getmap) provided by Server. MPFS Client conducts Writes to device after obtaining pre-allocated extents information from Server (allocSpace). CheckRead and checkWrite operations are used, as well as Read/Write locking mechanisms. MPFS AV operations work mainly at CIFS SMB protocol level for MPFS Clients, substituting FMP Open and Close to call AV service.

### **NEW EMC FUSION V2.2 (Ibrix High-performance & scalable MPFSi Solution—aka Chimay):**

- Solution that will be based on Ibrix Fusion software on an Intel platform, supporting 16TB file systems and small random IOs
- Entry-level gateway appliance
- Segment Servers, running RedHat Linux Enterprise 4.0, will supply metadata information to Clients and export file systems
- Compute/Client Nodes run application that accesses file systems for NFS or CIFS Samba, using EMC Fusion Client sw
- Single Fusion Manager runs on Linux RH Ent. 4.0, & manages cluster environment
- Storage Arrays—CX, CX3, AX, Symm 8000, DMX, DMX-2, DMX-3, etc.

### **CELLERRA ANTIVIRUS AGENT (CAVA—now part of the CEPA framework):**

**GA:** NAS 2.2.39.1, CAVA 1.8.9 Agent; CAVA 3.4.0 with NAS 5.4 (however, CAVA 2.2.4 can still be used with NAS 5.4), CAVA is 3.5.8 for NAS 5.5.30.x release. CAVA 3.6.2 is being released in 5.5.32 to support new Symantec SAVSE v5.1. CAVA has undergone a name change with the release of NAS 5.6, and now falls under the CEPA framework starting with version 4.0.2.

### **CEPA RELEASES:**

NAS 5.6 CEE (aka CEPA) 4.0.2

NAS 5.6.46 CEE 4.5.1

### **CELLERRA VIRUSCHECKER:**

Virus protection of network file systems for CIFS protocol only

### **MCAFEE 8.0 VIRUSSCANING OPTIONS:**

- Delete files automatically (supported)
- Clean files automatically (supported)
- Deny access to files (not supported)

### **Server Log:**

2008-11-17 06:55:15: VC: 3: -----> Checker 10.241.169.137 answers ACCESS\_DENIED and retry is disabled

→Setting “Deny access to files” does not take any action, and will generate an Access Denied to the CIFS client or application, and allows the file to be accessed again

### **SUPPORTED AV VENDORS/ENGINES:**

→Symantec Norton AV 7.6/8.0/8.1 Corp. Edition for Windows NT/2000/Symantec NAVCE 9.0

Symantec SAVSE for NAS, v4.3.17.40 with CAVA/CEE 3.5.7.1, 3.5.8, 3.6.2, 4.0.2-4.5.1

Symantec SAVSE for NAS, v4.3.12/14 with CAVA/CEE 3.5.7.1/3.5.6, 3.5.7.1, 3.5.8, 3.6.2, 4.0.2-4.5.1

Symantec SAVSE v5.1, 5.2 CAVA 3.6.2, CEE 4.0.2-5.1—new support for ICAP (Internet Content Adaptation Protocol) protocol for SAVSE interface, new DLL and new installer changes

→NAI-McAfee NetShield 4.5 SP1 for Celerra & VirusScan 8.0i/8.5i Patch 2, 8.7i with CAVA/CEE 1.8.9, 2.2.4, 3.4.0, 3.5.6, 3.5.7.1, 3.5.8, 3.6.2, 4.0.2-4.5.1

- Trend Micro's ServerProtect for EMC NAS 5.3.1 & 5.5.8 with CAVA/CEE 1.8.9, 2.2.4, 3.4.0, 3.5.6, 3.5.7.1, 3.5.8, 3.6.2, 4.02-4.5.1
- CAI InnoculateIT 6.0 & 7.1 for Windows NT/2000 (CA eTrust r8, r8.1) with CAVA/CEE 1.8.9, 2.2.4, 3.4.0, 3.5.6, 3.5.7.1, 3.5.8, 3.6.2, 4.0.2-4.5.1
- Sophos Anti-Virus v3.82, 5.x/6.x family, v7.0.2, v7.3, v7.5, v7.6 with CAVA/CEE 1.8.9, 2.2.4, 3.4.0, 3.5.6, 3.5.7.1, 3.5.8, 3.6.2, 4.0.2-4.5.1
- Kaspersky Anti-Virus for Windows Servers Enterprise Edition v6.0, with CEE 4.5.1 and NAS 5.6.46
- NTP Software QFS for NAS version 6.0+, & Northern Parklife Northern Storage Suite for EMC v8+: NAS 5.6.36-5.6.46 CEE/CEPA versions 4.0.2-4.5.1
- Varonis DatAdvantage v4.0.6, NAS versions 5.6.45-5.6.46, CEE versions 4.5.0.4-4.5.1

## **64-bit WINDOWS O/S IS NOT CURRENTLY SUPPORTED:**

→We do not currently support 64-bit O/S for the CAVA executable (emc164335). The main symptom is that VirusChecker cannot make contact with CAVA servers that run the 64-bit O/S.

**Note:** 64-bit support introduced with NAS 5.6.45 and CEE version 4.5.0.4

## **SUPPORTED CAVA VERSIONS:**

Cava 1.8.9 supported with NAS Versions 2.2, 4.1, 4.2

Cava 2.2.4, 3.4.0, 3.5.6 supported with NAS Versions 5.x and higher

Use CAVA 3.4.0 and higher if using McAfee VirusScan 8.0i

CAVA 3.5.8, 3.6.2

CEE\_CEPA 4.0.2 (NAS 5.6)

CEE\_CEPA 4.2.1 (NAS 5.6+)

## **CAVA LIMITATIONS:**

→CAVA works only with CIFS protocol

→NFS or FTP protocols do not trigger virus checking

→Monitors entire fs or certain file extension types.

→Beginning with NAS 4.2, have the ability to invoke viruschecking during file Read

→Supports High Road file systems from NAS 5.0 and higher

→Recommended that realtime scanning of live databases not be done [do these with AV vendor sw offline]

## **CELERRA VIRUSCHECKER 3.0—CAVA NAS 5.4:**

→Enhanced notification via PopUp messages to Clients and added messages to Server Log

1) Infected file 2) DART sends PopUp notification to Client 3) DART logs time of infection & notifies of action taken by AV Engine

→Accesstimes on Celerra automatically updated whenever AV Engine conducts Virus Definition update (McAfee only right now)

**Note:** Previous CAVA versions had no user heuristics to help report back to the AV Server

## **AV SIGNATURE FILE UPDATES:**

'Scan on read if access time is less than...' can be used to require file scanning after new virus definition files have been updated on the AV client. CAVA will send the new Virus Definition signatures to DART after the AV engine notifies CAVA of the updated files, and the date will update. However, some AV Vendors do not yet update Cava properly.

### **Trend Micro:**

Starting with Trend Micro Server Protect for Celerra version 5.58 Build 1176, and Cava version 3.58, the AV application will notify CAVA when a new signature or pattern update file is available, and the signature date on the Celerra will be updated accordingly.

## **HOW DOES THE VIRUSCHECKER & CAVA SCANNING PROCESS TAKE PLACE?**

→Virus Checker Agent on DataMover monitors files and when certain conditions are met (such as when writing to a file, access time, other configuration parameters), will initiate Scan check Request to the CAVA Service that runs on the AV engine, using the ONC/RPC interface protocol, and presenting the file name path using UNC ([\\dm03\\$CHECK\filename](\\dm03$CHECK\filename)) (Path includes the default CHECK\$ share location)

→Next, the CAVA service attempts to open the file using RPC and SMB Read

→Once file is opened, the AV engine antivirus driver detects the activity, and triggers it to run a Scan. At the same time, it blocks the CAVA services "file open" request from completing until the file is properly scanned. A problem with Virus Checking is that all requests are serialized—so for large files, a single scan operation is in progress and all other requests are backlogged.

→CAVA actually passes a file's signature to the AV Server, which checks its antivirus definition files for a match. For compressed files, the whole content is passed to the AV Server for scanning.

**Note:** Default timeout for blocked file is 10 secs--thereafter file is made available.

→Once scan is complete and AV engine has taken it's action, it then releases the block on the CAVA "open file"

→CAVA performs a file request, then closes the file and sends response to DM

### **SUMMARY OF CAVA PROCESS:**

DM check request to CAVA

CAVA opens file

CAVA queries file

CAVA closes file

CAVA sends response back to DM

### **VERIFYING COMMUNICATION TO CAVA SERVICE:**

→Prior to 5.6.47, CAVA tries to connect via port 139 for CIFS, then 445 as a secondary

→After 5.6.47, CAVA tries to connect via port 445 for CIFS, then 139 as a secondary

# **/usr/sbin/rpcinfo -p 192.1.4.217** →Querying CAVA Server portmapper

```
program vers proto port
1062247 1 udp 1076
1062247 2 udp 1076 →Output abridged. CAVA RPC program is 1062247.
```

# **.server\_config server\_2 -v "pong 192.1.4.217"** →Verifying that CAVA RPC program responds from CAVA Server  
1211463226: RPC: 3: 192.1.4.217 is alive

### **EXAMPLE OF VIRUS CHECKING PROCESS:**

1. Client has file opened for writing and initiates a close on the file
2. Data Mover obtains exclusive lock and places file into virus checker queue, which is serviced by a dispatch thread
3. RPC call is made by VC to AV Server containing UNC path to file to be checked in the Queue
4. AV Server (via Checker Service) receives RPC call & generates ‘file open’ kernel request using rights of EMC V.Checking User
5. Kernel open is intercepted by VC agent and retrieves portions of locked file for inspection & followup action
- Note:** AV Server actually opens the file twice (Create AndX), once using CAVA agent, and then again using AV Client
6. Kernel call returns & AV Server Checker Service returns RPC call to Celerra.
7. Celerra receives RPC call and unlocks file

**Note:** If you do not want to use the default CIFS interface for Virus Checking, specify the CIFS interface using the CIFSServer= line in the viruschecker.conf file.

### **BEST PRACTICE VIRUSCHECKING THREADS:**

#### **1. Normal system CIFS threads allocated are 96 or 256:**

**Note:** 1GB/2GB+ on DMs. Threads used to handle CIFS requests, and virus checking requests by presenting files to CAVA

#### **2. Number of Viruschk Threads should be 25% of the total number of CIFS threads on a Data Mover:**

# **server\_setup server\_x -P viruschk -o start=64**

**Note:** Threads are used to issue VirusChecker scan requests to CAVA threads on the AV Engine. Default VirusChecking service starts with (10) threads. This value should always be less than the number of threads dedicated on CAVA servers. The thread settings are persistent.

#### **DEFAULT THREADS:**

```
# server_setup server_2 -P viruschk -o start
server_2 : done
# server_viruschk server_2
server_2 :
10 threads started.
```

#### **3. Windows AV CAVA Server Settings should be two times number of Viruschk Threads set 2@64= 128:**

#### **HKLM>Software>EMC>CAVA>Configuration>NumberOfThreads: 128**

**Note:** Default install setup value is 20 threads

#### **4. Reserved CIFS Threads used for VirusChecking:**

**Note:** Exclusively used by VirusChecking. Default=3 threads, which are used only to break deadlocks between File and VirusChecking CIFS requests

# **.server\_config server\_2 -v "param cifs maxVCThreads"**

cifs.maxVCThreads INT 0x02673858 3 3 (1,4294967295) TRUE RESTART 'Reserved threads used for Virus checking'

**Note:** If a customer had (4) CAVA servers, may want to increase this value to 4, etc.

### **DEFAULT THREAD VALUES FOR CELERRA VIRUSCHECKING:**

--Data Movers start with (10) VC threads

--AV Engines start with (20) CAVA threads

--Virus Checker reserves (3) CIFS threads for exclusive use by VC by default

## **VIRUSCHECKER AND THE DEFAULT DATA MOVER C\$ SHARE:**

Viruschecking does not scan the C\$ share of the Data Mover—this is normal and intentional behavior

## **VIRUSCHECKER TERMINOLOGY:**

**VIRUS CHECKER CLIENT:** Min. NAS 2.2.39.1+ & 'viruschecker.conf' to setup VC parameters.

1. VC Client queues files, and provides UNC path to CAVA service on AV Server based on “trigger” points
2. VC Client blocks access to CIFS files while scanning occurs
3. VC Client releases files when informed of Success or Failure of Scan by CAVA

### **Examples of Trigger Points:**

- New virus definition file on AV Engine
- File renamed on Celerra
- File copied or saved on Celerra
- File modified & closed on Celerra
- Scan on First Read is enabled [also sets up scanning based on a pre-determined time for scanning to occur]

### **VC Agent handles virus ‘infections’ based on user-defined criteria:**

- cure or repair file
- rename file
- change file extension
- quarantine file
- deleting file

VC Client is configured using ‘server\_viruschk’ command and “viruschecker.conf” file.

## **VIRUS CHECKER SERVER:**

CAVA (Celerra Antivirus Agent) Service that determines when AV Engine scans a file [CAVA Agent 1.8.9/3<sup>rd</sup> Party AV Engine] as presented by Celerra Virus Checker Client

## **CAVA (CELLERRA ANTIVIRUS AGENT):**

EMC agent running on Windows NT/2000 AV; Communicates between A/V Engine and Celerra

### **AV ENGINE:**

3<sup>rd</sup> Party Antivirus Engine running on NT/2000

### **AV USER:**

Dedicated Domain Antivirus User Account that is added to a Celerra Local Group & granted the 'Viruschecking' User Right. Use this NT Account also to install & configure NT/2000 systems running CAVA & AV Engine.

### **AV Configuration Management MMC Snap-in:**

MMC Console used to view or modify CAVA parameters located in viruschecker.conf file

### **INTERFACE PROTOCOL:**

Client/Server interface that uses ONC/RPC protocol for Scan Requests & Replies.

### **SYSTEM REQUIREMENTS:** Minimum (2) dedicated AV Servers 700Mhz/512RAM

Minimum of (2) antivirus engines, supporting up to 1000 concurrent users. 100MB Lan connection recommended. (1) RPC request per scan. Add additional AV Server for every 500 concurrent Users. For multiple AV Engines, files are scanned in a roundrobin fashion to 'load-balance' the operation.

## **KNOWN ISSUES/CONCERNS WITH AV:**

### **COMPRESSED FILES:**

- Compressed files, such as zip, require whole file be sent to AntiVirus engine for scanning, taking much longer & using more network bandwidth and resources. Also, files with unknown extensions are sent to AV Engine for scanning. HWM (High Water Mark) should be monitored to identify excessive queue buildup (Add AV Engines if consistently High)
- min. of (2) AV Engines + (1) for every 500 additional Users [Watch HWM for indicators of load—SYS\_LOG; Server Log]
- AV Engines should be dedicated systems & not used for other Windows Services, etc.
- dedicated 100MB LAN as a minimum
- no AV in NAS 3.x or HighRoad MPFS until NAS 5.0
- CIFS support only—no Unix
- NT Security only—no Share Security mode
- do not setup both AV Network scanning by AV Engines & Celerra CAVA
- ensure AV Engine SW settings mirror those on Celerra Virus Checker, such as excluded files, etc.
- always exclude Temp files and Database application files [might otherwise cause degradation of network]
- never run Virus Checker Service during data migrations!
- leave default param “maxVCThreads=3” unless directed by Engineering
- Double-zipped files may not be properly scanned [not supported by most AV Engines either]
- Outlook .PST files are not scanned by McAfee or Symantec—exclude from virus checking!

--Files from Tape Restores do not get scanned because this is not a ‘CIFS Access’ operation

**Note:** However, AR45645 identifies code issue where NDMP Restores was enabling Virus Checking on the files being restored, causing numerous problems at customer sites. Set following param to disable Scanning on NDMP Restore.

**param pax scanOnRestore=0** [NAS 5.3, 5.1.26.x, 5.2.15.x]

**# server\_param server\_2 -facility PAX -info scanOnRestore**

server\_2 :

```
name      = scanOnRestore
facility_name = PAX
default_value = 1
current_value = 1
configured_value =
user_action = reboot DataMover
change_effective = reboot DataMover
range     = (0,1)
description = Whether to scan virus on restore
```

#### **ADS (Alternate Data Streams) FILES:**

Files that use ADS involves two separate writes to disk, meaning that the file is actually scanned twice. In order to avoid scanning files with ADS, EMC recommends that mask=\*. not be used. Instead, specify the actual file extension types to scan, using mask=\*.EXE: \*.COM: \*.DOC: etc. If mask=\*. is being used, you can still exclude ADS files for particular file extension types that may contain ADS files by using the following exclude syntax: excl=\*.com\* See emc185502 for more details.

#### **FUNCTION OF THE CELERRA VIRUSCHECKER CLIENT:**

1. Queues & communicates filenames using UNC paths to CAVA Agent on AV Engine for scanning
  2. Aware of events & triggers scanning [e.g., new Virus Definition installed on AV Server; file renamed or saved, access time, etc]
- Note:** File is opened twice, first by CAVA, then by Virus Scanning Engine.

#### **PERFORMING COMPLETE FILE SYSTEM SCANNING (NAS 5.3):**

**\$server\_viruschk server\_x -fsscan fs1 -create** [starts virus scanning on file system]

**\$server\_viruschk server\_x -fsscan fs1 -delete** [stops virus scanning on fs]

**\$server\_viruschk server\_x -fsscan fs1 -list** [scanning status]

**# .server\_config server\_3 "viruschk scan start=21"**

#### **SYMANTEC SAVSE ANTIVIRUS SUPPORT (NAS 5.5):**

Symantec SAVSE (Symantec AntiVirus Scan Engine), a user-mode scanning engine with ICAP (Internet Content Adpatation Protocol) 1.0 protocol and client-side API for CAVA to use when scanning (Symantec building SAVSE support for NetApps, EMC, others, and eliminating SAVCE support)

→SAVSE listens on port 7777

→Scan requests will use SAVSE API and not CAVA kernel

→SAVSE will use event signaling on pattern updates

→Only ‘action policy=delete’ will be allowed for this version

→For CAVA 5.5 in general, the new version is 3.5.7.1

**Note:** CAVA runs as AV Service on Windows. DM runs DART AV agent as client to CAVA service, sending scan requests to CAVA per heuristic definitions. Upon startup of the CAVA service, CAVA discovers the 3<sup>rd</sup> party AV Engine in use. If Symantec SAVSE is found, CAVA will route all scan requests through the cavasav.dll library to SAVSE and calls the SAVSE API functions found in symcscan.dll. Through the use of a CAVA/SAVSE Connector, DART should be able to direct Scan requests to CAVA, and CAVA redirects to the SAVSE. In order for AV scanning of file systems to occur, the CAVA/SAVSE Connector must be used. The protocol used is Symantec Native, not RPC as in previous implementations of VirusChecking/CAVA.

#### **INSTALLING SYMANTEC SAV for NAS:**

**Note:** Symantec SAV for NAS 4.3.x uses NATIVE protocol for deleting infected files, but version NAS 5.1.x uses the ICAP protocol

1. Install the SAV for NAS software
2. Go to Start>Run>Services.msc and change the SAV for NAS service to run from the domain CAVA or antivirus User account and password
3. Open the SAV for NAS application>Configuration>select “Native protocol” for version 4.3 or select “ICAP” for version 5.1, port 1344
4. If using version 5.1, stop the SAV for NAS service, open command prompt, navigate to the install directory, run “java –jar xmlmodifier.jar –s /policies/Misc/HonorReadOnly/@value false policy.xml, then restart the SAV for NAS service

**Note:** If these steps are not done, CAVA will not accept scan requests

5. Select ‘LiveUpdate’ to download and install latest definition files

## **INSTALLING McAfee VIRUSSCAN:**

**Note:** Example based on McAfee 7.1, NAS 5.3.14.2, & Windows 2000 domains

**Caution:** Celerra Antivirus can ONLY be implemented for physical CIFS servers, not VDM CIFS servers—we use the default Data Mover rootfs to host the hidden CHECK\$ share used for viruschecking operations.

1. Create domain user for viruschecking using Active Directory Users & Computers interface: avuser
2. Install the Celerra Management MMC Snapin from the Applications & Tools CD for the NAS Version in use: CelerraCifsMgmt.exe [Version 3.00.012]

3. Install McAfee AV engine per documentation
4. Open “Virusscan On-Access Monitor”>Properties>Detection tab>Scan Files: check all

--When writing to disk  
--When reading from disk

--On network drives

What to scan:

--All files

5. Actions tab>When a threat is found:

Primary Action: Clean files automatically (if cannot clean, will append .VIR extension to file)

Secondary Action: Delete files automatically

## **THINGS TO VERIFY FOR CELERRA ANTIVIRUS SETUP:**

--For W2K Data Movers, use the Celerra CIFS Management snapin for Antivirus Management in order to assign the correct AV User the “EMC Virus Checking” privilege

**Note:** NAS 5.1 and below, and all NT 4.0 style data movers, can use the “Usrmgr.exe” Utility to set the AV privilege

--Default compname should match what viruschecker.conf file uses

--Perform AV Engine software install using AVUser account

--Install AV Engine software first, then CAVA (except for Trend Micro)

--Cava Service on the AV Servers must run under the credentials of the antivirus User account (if the service is properly changed to run under this account, a popup should display “The account .\virususer has been granted the Log On As A Service right.”

--Antivirus user account must be added to Celerra Local Administrator group

--Antivirus user account must be added to all AV Servers’ Local Administrator group

--For excluded/include files, settings on AV Servers should match that of viruschecker.conf file

## **SETTING UP AV USER ACCOUNT FOR WINDOWS 2000 CIFS SERVER:**

1. Create User account using Active Directory Users and Computers MMC

2. Create Local Group on DataMover using Active Directory Users and Computers

→Rightclick computer name>Manage>System Tools>Local Users and Groups>Rightclick and add New Group>Select Users or Groups, add the “avuser” account create in step 1.

### **Select Components to Install:**

→Unix User Management [Do not install--used only to support CIFS Active Directory Mapping tool]

→Dart Management

  --AntiVirus Management (yes)

  --Home Directory Management (Installed if using the HOMEDIR feature on Celerra)

  --Security Management (yes—this is the key program for setting up VirusChecker rights on Celerra)

3. Add domain “avuser” to Celerra CIFS Server’s local Administrators group

**Note:** Computer Management>Connect to another computer>\Cava\_server>Local Users & Groups>Groups>Administrators>Add

4. Create new Local Group on Celerra called “AVGroup” & add “avuser” to this group as well

5. Add domain “avuser” to each AV Server’s local Administrators group (Via Computer Management or ADUC)

**Note:** For troubleshooting purposes, add ‘avuser’ to ‘Domain Admins’ group if having permission issues

6. Grant following privileges or rights to the “avuser” account on the Data Mover using Celerra Management snapin:

a. Start>Programs>Administrative Tools>Celerra Management>EMC Celerra Management>Data Mover Management>

b. Rightclick and select ‘Connect to Data Mover’

c. Add the appropriate compname from the dropdown box

d. Expand ‘Data Mover Security Settings’>User Rights Assignment: Add following rights to “avuser”

→**EMC Virus Checking**

→**Backup Files & Directories**

→**Restore Files & Directories**

→**Generate Security Audits**

→**Manage Auditing and Security Log**

**Note:** Adding Virus Checking rights from command prompt

C:\Documents and Settings\Administrator>ntrights -u <domain>\cavauser -m <Data Mover IP address> +r SeVirusChecking  
Granting SeVirusChecking to europe\cavauser on 10.241.183.85... successful

7. Grant following privilege to “avuser” on each AV Server

Programs>Administrative Tools>Local Security Policy>Security Settings>Local Policies>User Rights Assignment:

**→Act as part of the operating system**

**Note:** In some cases, this right may need to be applied using the ‘Domain Security Policies’ snapin (policy dependent)

8. Install McAfee VirusScan Enterprise 7.1.0 on AV Servers

**Note:** Use the antivirus account to perform the software installation—means the account will need to be either local Administrator or Domain Admins

9. Install EMC CAVA 2.2.4 software on each AV Server using install defaults

InstallShield>Accept License>Complete>Install>Finish>answer ‘No’ to restart system

10. Using ‘Services’ program on each AV Server, configure the EMC CAVA service to run under the control of the “avuser” account:  
Services>EMC CAVA>Log On>\*This account: add “avuser” account here and enter that account’s password

**Note:** A PopUp should appear saying, “The account [avuser@dwarfs.com](mailto:avuser@dwarfs.com) has been granted the Log On As A Service right”

11. Reboot AV Servers & then verify that EMC CAVA service is running correctly

12. Configure McAfee AV Servers for Scanning Network Drives:

Open McAfee VirusScan Console>On-Access Scan properties>Highlight ‘Default Processes’>

**Detection tab:**

→When writing to disk (default setting, but make sure this is enabled)

→When reading from disk (default setting, but make sure this is enabled—AV Engine cannot scan Celerra files without this set)

→On network drives (Enable this setting—not enabled by default—VirusChecking will NOT work without)

**Actions tab:**

→Clean infected files automatically [default]

→Move infected files to a folder [default]

**Note:** With earlier versions, quarantining does not work, so change the ‘If the above action fails’ setting to →‘Deny access to file’. When the ‘deny access to file’ setting is set, and a virus is detected, the file is renamed with a .vir extension and left on the Celerra. Admin would have to check the Task>Activity Log to determine what viruses were found as no popup messages display on Client.

**MCAFEE 8.0 VIRUSSCANING OPTIONS:**

**→Delete files automatically (supported)**

**→Clean files automatically (supported)**

**→Deny access to files (not supported)**

13. Configure and Start VirusChecker service on Data Mover:

a.) Create viruschecker.conf file on Control Station and push to /.etc using server\_file server\_x –put command  
masks=\*.\*

excl=pagefile.sys:\*.tmp:\*.pst:\*.mdb:\*.ldb

addr=172.17.17.219:172.17.17.210

CIFSserver=Cava\_server

shutdown=cifs

b.) Start VirusChecker service & verify communication to AV Servers:

**\$server\_setup server\_x -P viruschk -o start**

**# server\_viruschk server\_2**

server\_2 :

10 threads started

1 Checker IP Address(es):

10.2.36.22 version=2, ONLINE at Fri Mar 4 23:23:48 2005 (GMT-00:00)

1 File Mask(s):

\*.\*

5 Excluded File(s):

pagefile.sys \*.tmp \*.pst \*.mdb \*.ldb

Share \\Cava\_server\CHECK\$

RPC request timeout=25000 milliseconds

RPC retry timeout=5000 milliseconds

High water mark=200

Low water mark=50

Scan all virus checkers every 60 seconds

When all virus checkers are offline:

Continue to work with Virus Checking and CIFS

Scan on read if access Time less than Fri Feb 25 19:42:07 2005 (GMT-00:00)

Panic handler registered for 65 chunks

14. Test effectiveness of Virus Checker setup by using Eicar Virus-test Files:

- a.) Stop VirusChecker service on data mover
- b.) Stop CAVA Service and McAfee AV Services on AV Servers
- c.) Download test files from [www.eicar.com/anti\\_virus\\_test\\_file.htm](http://www.eicar.com/anti_virus_test_file.htm)  
**eicar.com eicar.com.txt eicar\_com.zip eicarcom2.zip**
- d.) stage files on Celerra share and map drive from AV Server
- e.) Start all Services back up and access test files from Client—copy from one folder to another on the Celerra

## **TROUBLESHOOTING VIRUSCHECKER SCANNING ISSUES:**

→VirusChecker Service runs on the Data Mover & presents files to CAVA when certain rules apply [Scan on first read, writes, etc.]

→CAVA Service runs on AV Server and receives request to scan file from VirusChecker, Opens file & presents to AV Server

→McAfee AntiVirus Service receives request to scan file from CAVA & performs operation

1. Set debug logging on data mover:

**# server\_param server\_2 -facility viruschk -info Traces**

```
server_2 :  
name      = Traces  
facility_name = viruschk  
default_value = 0x00000000  
current_value = 0x00000000  
configured_value =  
user_action = none  
change_effective = immediate  
range     = (0x00000000,0xffffffff)  
description = Define traces displayed in the server log  
detailed_description
```

Define the traces displayed in the server\_log for virus checker: 0x00000001:in CFS: setCheckStatus, setCheckWriter, check Wait 0x00000002:in CIFS: createEvent, sendEvent, mustBeChecked (scan on read) 0x00000004:in Virus Checker: connectAnyServer, vc\_checkfile, stopThreads, exit, start 0x00000008:in CIFS applbnt: open, writeAsyncMsg, readMsg, close, rename 0x00000010:in Virus Checker: heartbeat the virus checker servers 0x20000000:Disable error login when a CAVA makes too many access without the EMC viruschecking privilege 0x40000000:Warnings 0xC0000000:Warnings + Errors

**# .server\_config server\_2 "param viruschk Traces=0xC0000004"** [Turns on Debug for AV for Warnings & Errors]

**Note:** After turning on viruschecker tracing, you may need to use server\_log server\_x -i in order to see trace output

**# .server\_config server\_2 "param viruschk Traces=0x00000000"** [Turns off Debug for AV]

**# server\_param server\_2 -facility viruschk -info Traces -v**

```
server_2 :  
name      = Traces  
facility_name = viruschk  
default_value = 0x00000000  
current_value = 0x00000000  
configured_value =  
user_action = none  
change_effective = immediate  
range     = (0x00000000,0xffffffff)  
description = Define traces displayed in the server log  
detailed_description
```

Define the traces displayed in the server\_log for virus checker: 0x00000001:in CFS: setCheckStatus, setCheckWriter, check Wait 0x00000002:in CIFS: createEvent, sendEvent, mustBeChecked (scan on read) 0x00000004:in Virus Checker: connectAnyServer, vc\_checkfile, stopThreads, exit, start 0x00000008:in CIFS applbnt: open, writeAsyncMsg, readMsg, close, rename 0x00000010:in Virus Checker: heartbeat the virus checker servers 0x20000000:Disable error login when a CAVA makes too many access without the EMC viruschecking privilege 0x40000000:Warnings 0xC0000000:Warnings + Errors

2. Download DebugView 4.31 from [www.sysinternals.com/Utilities/DebugView/](http://www.sysinternals.com/Utilities/DebugView/) v4.31 and configure on AV Servers:

a.) Install DebugView program

b.) Change registry settings to enable DebugView logging & Celerra's Cava Monitor display:

**HKLM>software>EMC>CAVA>Configuration:**

Verbose: 1

Debug: 1

3. Enable Cava Monitor on each AV Server via registry setting:

**HKLM>software>EMC>CAVA>Sizing:** Set 'Sizing' value to 1 for debug

**Note:** CAVAMon.exe is an executable found in Program Files>EMC>CAVA that can be used to collect Scanning stats and info

4. Startup network trace if conducting test in order to capture data

5. Open Debug View Program and Cavamon Program to be ready to capture test data
6. Begin testing with eicar files
7. Save Server Log to file
8. Verify that user account has been added to Administrators localgroup on Celerra and each AV Server
9. Verify that CAVA is running under the domain avuser account
10. Verify that avuser has been granted the EMC VirusChecker right by running following:

```
# .server_config server_2 -v 32768 "lg list" |grep -C avuser
```

```
1109977963: SMB: 4: user CORP\avuser 802a S-1-5-15-72ab4f17-4785586f-22081e61-fda1
```

```
1109977963: SMB: 4: Localgroup 'Users' (Ordinary users)
```

```
1109977963: SMB: 4: Sid = S-1-5-20-221 (gid:0x80800221)
```

```
1109977963: SMB: 4: Privileges = AV (0x0)
```

```
1109977963: SMB: 4: Members:
```

```
1109977963: SMB: 4: user CORP\avggroup 802a S-1-5-15-72ab4f17-4785586f-22081e61-fda1
```

11. Verify av user account has been granted proper rights via Celerra Management snapin

12. Verify Antivirus Software settings with customer [get screenshots if necessary]

13. Use following params to tweak number of Open Files cached & pending Virus Checking, and % Vnodes allowed:

```
# server_param server_2 -facility viruschk -info vnodeMax -v
```

server\_2 :

```
name      = vnodeMax
```

```
facility_name = viruschk
```

```
default_value = 2000
```

```
current_value = 2000
```

```
configured_value =
```

```
user_action = restart Service
```

```
change_effective = restart Service
```

```
range      = (100,4294967295)
```

```
description = Max number of vnodes that can pend on VC
```

```
detailed_description
```

Defines the max number of total vnode available in the system which can be pending on Virus Checking. An event is sent to the Control Station when the maximum is reached. The CIFS thread will be blocked until the Low Watermark is reached again.

```
# server_param server_2 -facility viruschk -info vnodeHWM -v
```

server\_2 :

```
name      = vnodeHWM
```

```
facility_name = viruschk
```

```
default_value = 90
```

```
current_value = 90
```

```
configured_value =
```

```
user_action = reboot DataMover
```

```
change_effective = reboot DataMover
```

```
range      = (0,4294967295)
```

```
description = Percentage of vnodes that can pend on VC
```

```
detailed_description
```

Defines the percentage of total vnode available in the system which can be pending on Virus Checking. An event is sent to the Control Station when the maximum is reached. The CIFS thread will be blocked until the Low Watermark is reached again.

## **SPECIFYING VIRUS CHECKER ACTIONS IN EVENT AV SERVERS ARE OFFLINE:**

**Comment:** Following settings dictate how Celerra, when configured to run Virus Checker, behaves when AV Servers are Offline

**shutdown=no**

**Note:** This is the default setup for Virus Checker and there may or may not be an entry in the viruschecker.conf file. Policy allows for continuing to check for AV Servers and does not shutdown VirusChecker or CIFS. However, with this policy set and with all AV Engines Offline, there is a potential for the system to eventually panic:

**DATA MOVER PANIC:**

The policy states that files must be checked. If something prevents us from checking a file in a timely fashion, then the file will languish until it is checked. The thread handling the file will pend until the check is complete. If the VC is offline, the check will not occur and the thread will continue to wait. Each write request pends and ties up a thread. As all threads become occupied, the new requests are queued up in the "collector" thread. The collector thread has no limit and continues to queue up requests on the assumption that things will get better. The requests occupy memory in the streams message block structure (msgb). After a time, the msgb memory will be exhausted and the system will panic when it can't get a msgb.

**SHUTDOWN=NO POLICY:**

```
# server_viruschk server_5  
-----output abridged-----
```

Scan all virus checkers every 60 seconds

**When all virus checkers are offline:**

*Continue to work with Virus Checking and CIFS*

**shutdown=cifs** [Loss of AV Servers will result in CIFS shutdown and loss of access to CIFS]

**shutdown=viruschecking** [Loss of AV Servers will result in VirusChecker Service shutting down, but not CIFS]

**HOW TO DETERMINE WHICH SHUTDOWN POLICY IS IN EFFECT:**

1. Cat viruschecker.conf file and look for “shutdown=”
2. For default “shutdown=no”, there does not need to be an entry in the viruschecker.conf file.

**\$ server\_viruschk server\_5**

**When all virus checkers are offline:**

*Continue to work with Virus Checking and CIFS*

**Note:** The above output reflects the default Shutdown Setting for Virus Checker, which is “shutdown=no”

**AV BEST PRACTICES:**

- Startup Viruschecker with 64 threads v. default 10 [start=64]
  - Startup CIFS threads at 128 v. default 32 [start=128]
  - Change default AV Server threads in registry from 20 threads to 120
  - Always ensure that CIFS threads are >than Viruschecker threads
  - Do not allow “real-time” on-line scanning of Application DataBases such as Exchange, SQL, Oracle, Access, etc. Rather, databases should be handled by Vendor-Specific database-scanning software or done offline. On-line scanning of DB files can cause db corruption and degrade Network performance (Exclude databases in viruschecker.conf “exclude” section)
- Note:** We have seen where virus checking will make Access database applications fail intermittently when performing operations over the network to a Celerra share.

**STARTUP BEHAVIOR FOR VIRUSCHECKER ON DMs:** 5.5.31.x AR98322

If using CLI or MMC to start viruschecker service, a viruschecker.enabled file will be created in /.etc. The viruscheck start line will no longer be added to the netd file. Now, whenever CIFS starts, if the viruschecker.enabled file exists, viruschecker will also start.

**DEFAULT VIRUS CHECKER PARAMETERS:**

**# server\_param server\_2 -facility viruschk -list**

|            |          |            |            |                                       |
|------------|----------|------------|------------|---------------------------------------|
| param_name | facility | default    | current    | configured                            |
| Traces     | viruschk | 0x00000000 | 0x00000000 | [Set to 0xC0000004 for debug logging] |
| vnodeLWM   | viruschk | 60         | 60         |                                       |
| chunkQuota | viruschk | 65         | 65         |                                       |
| vnodeHWM   | viruschk | 90         | 90         |                                       |
| Notify     | viruschk | 7          | 7          |                                       |
| vnodeMax   | viruschk | 2000       | 2000       |                                       |
| noRetry    | viruschk | 0x00000018 | 0x00000018 |                                       |

**# .server\_config server\_2 -v "param viruschk audit"**

viruschk\_audit INT 0x020e5e18 98 98 (0,4294967295) FALSE REBOOT 'NA'

viruschk\_Traces (0) [Set when debug mode desired]

**CHECKING VIRUS CHECKER SETUP ON DATAMOVER:**

**\$server\_viruschk server\_2** [Monitor requests in collector queue]

**\$server\_viruschk server\_2 -audit**

[Status of VC Client; # files checked; Progress of files being scanned--Use to monitor High Water Mark queue--may indicate need to add additional antivirus engines; Check "sys\_log" also]

**AV FILES LOCATED /.etc DIRECTORY:**

|            |   |      |     |                  |                    |                                    |
|------------|---|------|-----|------------------|--------------------|------------------------------------|
| -rw-r--r-- | 1 | root | bin | 0 Nov 8 18:58    | viruschecker.audit | [Contains list of unscanned files] |
| -rw-r--r-- | 1 | root | bin | 3193 Nov 7 15:52 | viruschecker.conf  |                                    |
| -rw-r--r-- | 1 | root | bin | 4 Nov 8 18:54    | viruschecker.date  |                                    |

**UPDATING VIRUS CHECK CONF FILE:**

Step 1. Edit Viruschecker.Conf File

Step 2. Push to Server: # server\_file server\_2 -put viruschecker.conf viruschecker.conf

Step 3. If Virus Checker Service is running, execute following to update “conf” in Memory:

**\$server\_viruschk server\_x -update**

**VIRUS CHECKER SERVICE COMMANDS:**

**\$ .server\_config server\_2 "viruschk scan start=21"** (Example of full viruschecking scan of file system)

**\$server\_setup server\_x -P viruschk -o start=# (# of threads to start)**

**\$.server\_config server\_x -v "param viruschk"** (displays viruschk config)

**\$.server\_config server\_x -v "viruschk stop or start"**

**\$.server\_config server\_x -v "viruschk restart" [\$.server\_setup server\_x -P viruschk -o restart]**

**\$.server\_config server\_x -v "printstats viruschk"** (same as audit)

**\$.server\_config server\_2 "param viruschk Traces=0xC0000004"** [Turns on Debug for AV]

**Note:** With 5.6, you may need to use \$ server\_log server\_x -i option to see trace output in the server log

**\$.server\_config server\_2 "param viruschk Traces=0x00000000"** [Turns off Debug for AV]

**\$.server\_config server\_2 -v "param viruschk noRetry=0xFFFFFFFF"** [Flush bad entries from queue]

**\$.server\_config server\_2 -v "param viruschk noRetry=0x00000018"** [Reset queue to normal]

**\$.server\_config server\_2 -v "ntcred user=avuser"**

**# .server\_config server\_2 -v "help viruschk"**

**Usage:**

viruschk start[=number of threads to start]

viruschk restart

viruschk stop

viruschk

viruschk audit[=filename]

<FS> is filesystem id or mount path

viruschk scan start=<FS>

viruschk scan stop=<FS>

viruschk scan status=<FS>

viruschk scan

**Description:**

'viruschk start' starts virus checker

'viruschk restart' restarts virus checker

'viruschk stop' stops virus checker

'viruschk' displays the virus checker configuration

'viruschk audit' displays the status of the virus checker

<FS> is filesystem id or mount path

'viruschk scan start=<FS>' starts scan on file system FS

'viruschk scan stop=<FS>' stops scan on file system FS

'viruschk scan status=<FS>' displays status of the scan of file system FS

'viruschk scan' display status of the scan of all file systems

**SERVER LOG DEBUG MODE FOR AV:**

2002-11-15 10:53:55: VC: 4: UNC='\\SP099NAS01\CHECK\$\adssync\adssync\eicar.com.txt'

2002-11-15 10:53:55: VC: 4: waiting to be checked

2002-11-15 10:53:55: VC: 4: [av005] UNC='adssync\adssync\eicar.com.txt'

2002-11-15 10:53:55: VC: 4: Sent to checker 131.99.75.42

2002-11-15 10:53:55: VC: 4: [av005] UNC='adssync\adssync\eicar.com.txt'

2002-11-15 10:53:55: VC: 4: Checker 131.99.75.42 answers SUCCESS

**CONDUCTING VIRUSCHECKER AUDITING ON DATAMOVER:**

**\$server\_viruschk server\_2** [SERVER STATUS]

5 threads started

**1 Checker IP Address(es):**

**addr=192.10.2.2 192.10.2.2 online** [offline would mean loss of communications to CAVA on AV Server]

**1 File Mask(s):** [masks=".\*"; Only one mask is defined; if there were multiple entries, separate with :]

**masks=".\*** (all files to be checked; or configure for specific extensions: masks="\*.ppt:\*.exe"]

**excl=\*.tmp (default=no files excluded) No File Excluded** [excl=; Lists extensions excluded from scanning, if specified]

**CIFSServer=dm03 \WDM03\CHECK\$** [CIFSServer=Netbios name of DataMover with default 'CHECK\$' viruschecker share]

**maxsize=0 (default=0 bytes) Max size of checked files=0** (maxsize=0x1000000)

[Min=0/Max=0xFFFFFFFF → 4GB file is max size allowed]

**maxThreadWaiting=0 (old default=16; new default=0) Max threads waiting for a file=16;** [Min=0, Max=0xFFFFFFFF]

**Note 1:** *Number of AV Servers \* number of threads running on each AV Server. It is now recommended not to set this field.*

**waitForTimeout=0** (Default value=0ms or off) Set timeout value to 30,000ms or higher to address certain conditions, such as migration of files--Min=0, Max=0xFFFFFFFF.

# Wait 30 seconds if the file is already opened by the viruschecker

**waitForTimeout=30000**

**Note 2:** The waitForTimeout option does not allow a check for lock conflicts at the beginning. All opens will block for a period up to the length of the timeout. If the check completes the open is processed. If the check hasn't completed and the timeout has expired, we process the open. If there is a lock conflict the sharing error is returned. If there are no lock conflicts the open will succeed (this is the instance where access is given to an unchecked file). If the timeout is set high enough it effectively removes the possibility of someone getting access to an unchecked file. However, this has the potential of making an Explorer window hang.

**Note 3:** Defines the time a thread will wait for viruschecking to complete file scan before requesting an open on the file.

**RPCRequestTimeout=25000** ms (default=25 secs) [After 5 retries, CAVA marked offline; Min=1; Max=0xF...]

**RPCRetryTimeout=5000** ms (default=5 secs) [Wait time for response to request before timeout; Must be lower than RPCRequestTimeout; Min=1; Max=0xFFFFFFFF]

**highwatermark=200** (default) [Number of files in Queue for scans; When value is exceeded, logs to /nas/log/sys\_log; Min=4 Max=0xFFFFFFFF → If consistently logging High Water Marks, consider bringing additional AV Server online]

**lowwatermark=50** (default) [Logs when this value is reached after a HWM event was triggered; Min=1]

**surveyTime=60** [default=60 seconds to scan for CAVA servers [Survey Time=interval of comm. to CAVA; Min=1; Max=0xF..]]

**Note:** DM pings AV servers every 60 secs by default

**When all virus checkers are offline:**

**shutdown=no** (default=no, but recommended setting is 'shutdown=cifs')

## VIRUSCHECKER.CONF FILE:

**# more viruschecker.conf**

```

#
# Virus Checker Configuration File
# (no blank line)
#
# masks=<list of filename extensions>
# - sets the list of filemasks that need to be checked.
#
# excl=<list of filenames or filename extensions> (optional)
# - sets the lists of filenames or filemasks that donot need
#   to be checked.
#
# addr=<list of IP addresses of virus checker servers>
# - sets the IP addresses of the VC Servers that we wish to connect to.
#
# CIFSserver=<name of the CIFS server in the dart> (optional)
# - sets the name of the CIFS server. If the parameter is not given,
#   the default CIFS server will be used.
#
# maxsize=xxxx (32 bits) (optional)
# - sets the maximum file size that will be checked.
#   Files that exceed this size will not be checked.
#   If the parameter is not given or is equal to 0,
#   it means no file size limitation.
#
# surveyTime=xxxx (optional)
# - sets the survey time frequency for scanning all the virus
#   checkers to see if they are online or offline. The unit is
#   second. The default value is 60 seconds.
#
# shutdown=<no | cifs | viruschecking> (optional)
# - sets the shutdown action when all virus checkers are offline.
#   no      : no action taken (default value)
#   cifs    : stop cifs
#   viruschecking: stop viruschecking
#
# highWaterMark=xxxx (optional)
# - set the high watermark value. When the number of requests
#   in progress is larger than "highWaterMark" value, a log
#   event will be sent to Control Station.
#   The default value is 200.
#
# lowWaterMark=xxxx (optional)—purpose is to log a descending threshold after a HWM event has been triggered)
# - set the low watermark value. When the number of requests
#   in progress is smaller than "lowWaterMark" value, a log
#   event will be sent to Control Station.

```

```

# The default value is 50.
#
# Note: The following parameters must be set only by EMC
#
# maxThreadWaiting=xxxx (Engineering now recommends that this value be set to 0)
# - set the maximum number of threads waiting for a file. If
#   the number of threads waiting for a file is already equal
#   to the maximum number, the next thread will be allowed to
#   access the file EVEN IT IS NOT VIRUS CHECKED.
#   The default value is 0. It means the function is disabled.
#
# waitTimeout=xxxx (Eng. now recommends that this value be set to 0]
# - set the time out which a thread waits for a file. If
#   the waiting time is reached, the thread will be allowed to
#   access the file EVEN IT IS NOT VIRUS CHECKED. The unit is
#   millisecond. The default value is 0. It means the function
#   is disabled.
#
# RPCRetryTimeout=xxxx (optional)
# - set the time out of RPC retry. The unit is millisecond.
#   The default value is 500.
#
# RPCRequestTimeout=xxxx (optional)
# - set the time out of RPC request. The unit is millisecond.
#   The default value is 25000.
#
# When a RPC is sent to the VC server, if the VC does answer
# after "RPCRetryTimeout", the Data Mover makes retries until
# "RPCRequestTimeout" is reached.
#
# Example:
#
# masks=*.EXE:*.COM:*.DOC:*.DOT:*.XL?:*.MD?:*.VXD:*.386:*.SYS:*.BIN
# masks=*.RTF:*.OBD:*.DLL:*.SCR:*.OBT:*.PP?:*.POT:*.OLE:*.SHS:*.MPP
# masks=*.MPT:*.XTP:*.XLB:*.CMD:*.OVL:*.DEV
# masks=*.ZIP:*.TAR:*.ARJ:*.ARC:*.Z
# addr=168.159.173.239
masks=*.*
excl=pagefile.sys:*.tmp:*.pst:*.mdb:*.ldb
addr=172.17.17.219:172.17.17.210
CIFSserver=ftwnas01
shutdown=cifs

```

**EDITING THE VIRUSCHECKER.CONF FILE:** [Defines VirusChecker parameters]**#/home/nasadmin/vi viruschecker.conf**

#viruschecker configuration file

**addr=192.10.2.2** [Up to (4)IP Addresses for A/V Servers→DM sends UNC path of file AV Server to scan]**CIFSserver=nas46** [NetBIOS name of DataMover; if not listed, the default Netbios name is used for DM]**masks=\*.\*** [Default: All files are scanned; Set to specific extensions up to 1024 characters per line]**Note:** Must include a mask or else no files will be scanned! Specify a set of extensions for file types to scan. Specify an 'exclude' list for those that you do not want to scan.**excl=pagefile.sys:\*.tmp:\*.exe:\*.com:\*.sys:\*.zip:\*\*\*\*\*** [Excluded File Extensions; \* ? metacharacters can be used]**Note:** NT limitation of up to 256 characters in a filename + extension; greater than this cannot be scanned**maxsize=n** [Sets max file size that will be scanned; Default =0 or no file size limit; Max size = 4GB]**highWaterMark=200** (default) [Number of requests in progress exceeds 200, an event is logged]**lowWaterMark=50** (default) [Triggered after HWM event—when number of requests drop to 50, event is logged]**RPCRequestTimeout=25000** (default—a suggestion change would be to set this value to 70000, or 70secs)**Note:** Celerra's default total timeout period is 25 secs before marking AV offline, which could happen during heavy loads, large files, or poor networks. Typical VC engine scan timeout is 60 secs.**RPCRetryTimeout=5000** (default)**surveyTime=60** (default) [Change time that VC Client checks to verify availability of AV Engines]**shutdown=no** [System default: Continue retrying RPC to AV servers--does not shutdown Viruschecker or CIFS]**shutdown=cifs** [If Viruschecker fails, will stop CIFS service & access to Shares]**shutdown=viruschecking** [Stops Viruschecking but Windows Clients can still access shares with CIFS]

## **EXAMPLE OF ‘SERVER\_VIRUSCHK’ OUTPUT:**

**\$ server\_viruschk server\_6**

```
server_6 :  
10 threads started  
2 Checker IP Address(es):  
169.37.139.156 online  
169.37.139.160 online  
30 File Mask(s):  
.avb *.exe *.com *.doc *.dot *.xl? *.vxd *.386 *.sys *.bin *.rtf *.obd *.dll  
.scr *.obt *.pot *.ole *.shs *.mpp *.mpt *.xtp *.xlb *.cmd *.ovl *.dev *.zip  
.tar *.arj *.arc *.z  
13 Excluded File(s):  
.ldb *.mdb *.maf *.mam *.maq *.mar *.mat *.mda *.mde *.mdn *.mdw *.mdz *.ppt  
Share \\NNYC11P20006\CHECK$  
RPC request timeout=25000 milliseconds  
RPC retry timeout=5000 milliseconds  
High water mark=200  
Low water mark=50  
Scan all virus checkers every 60 seconds  
When all virus checkers are offline:  
Continue to work with Virus Checking and CIFS  
Scan on read disable
```

## **EXAMPLE OF VIRUSCHECK AUDITING:**

**\$server\_viruschk server\_2 -a** [SERVER AUDIT OF ACTIVITY]

```
Total Requests : 78      [# requests in Queue & files processed by AV threads]  
Requests in progress : 0      [# of files in collector queue]  
NO ANSWER from the Virus Checker Servers : 0 >>>Used for Alerts  
ERROR_SETUP : 0      [Viruschecker server not setup correctly]  
min=3447 us, max=191727 us, average=31730  
FAIL : 0      [viruschecker failed to scan file]  
TIMEOUT : 0      [viruschecker timed out on file]  
0 File(s) in the collector queue  
0 File(s) processed by the AV threads      [Files processed by AV threads]
```

## **SETTING THE MAXIMUM VC THREADS FOR DATAMOVERS:** 2.2.35.4 & 2.2.39.1

Step 1. \$/nas/site/slot\_param

**param cifs maxVCThreads=5** [Breaks deadlocks if all Virus Checker threads in use or blocked]

cifs.maxVCThreads 0x02997038 0x00000003 0x00000003

Step 2. Update “viruschecker.conf” param file: \$server\_viruschk server\_x -update

**Note:** Default VCThreads=3

## **SETTING NUMBER OF THREADS FOR VC CLIENT:**

**\$server\_setup server\_x -P viruschk -o start=64**

**Note:** Change is not persistent on reboots!

## **PARAMETERS FOR VIRUSCHECKER.CONF:** Helps with performance issues when scanning

**Problem:** Check DOES NOT ANSWER, Retrying; Checker answers FAIL, Retrying File status=ok

RPCRequestTimeout=300000 [Default value is 25000ms or 25 seconds]

RPCRetryTimeout=60000 [Default is 5000ms or 5 seconds]

Low Water Mark: default=50 [Entry made in /nas/log/sys\_log]

High Water Mark: default=200 [Entry made in /nas/log/sys\_log]

RPCRetryTimeout: default=5000 milliseconds

RPCRequestTimeout: default=5 retries, or 25000 millesconds

waitForTimeout: default=0 seconds [Old value was 10]

maxThreadWaiting: default=0 [Old value was 16]

maxsize: default=0 bytes

CIFSServer: CHECK\$ share is created

surveyTime: default=60 seconds; scan interval to all CAVA servers

shutdown: default=no [continue retrying list of Servers]

**Best Practice:** Use Shutdown=cifs to minimize virus infection damage by having CIFS Service stop if VirusChecking fails

## **TROUBLESHOOTING/QUESTIONS:**

- Check NT Event Logs--Application Log
- Ping CAVA Servers to determine network connectivity
- Version of CAVA & AV Software + Virus Definition update level
- Run \$server\_netstat server\_x -i to ensure that there are no interface errors--could be duplexing issue
- "invalid access..."; ensure that Virus Checker User Account has proper rights assigned
- Ensure that the AV Server Scan Settings are matched with those on the DataMover Virus Checker Server
- Do not run Antivirus service during data migrations!
- Is User access or performance hindered by the a/v process?
- Can the issue be isolated to certain file extensions or scanning of very large files?
- Does stopping & starting CIFS or A/V Checker Service eliminate the problem?
- If CIFS appears sluggish & hung and has high CPU utilization, run "server\_profile" command
- In extreme cases, run the Server\_Panic command to create a dump file for Eng. Analysis
- Network Trace may also be required
- Use Debug Trace software program
- CIFS service must be started before Virus Checker can be started
- Known issue whereif AV Server Service dies or becomes unavailable, after the 'waitForTime' value of 10 seconds, file may become available to user once again yet did not actually get scanned.
- Datamover panics: Listing of unscanned files stored in ./etc/viruschecker.audit
- By default VC Client scans for AV Engines every 60 seconds

## **TROUBLESHOOTING VIRUS CHECKER ISSUES:**

### **I. RPC Error when starting viruschk service:**

"RPC Error from checker 192.10.3.2" [DM cannot communicate to A/V Server]  
[Verify network connectivity; Check CAVA Service on Windows Server; Check Virus Checker Service on DM;  
Check that EMCVirCk driver loaded—Event Viewer System Log; Server\_Log errors]

### **II. SMB:3: Invalid access from client USER1 to CHECK\$**

CFS:4: setCheckWriter(), FNN=0x3981bd0 (WRITE), FOF=0x6b97284, Wait=0, TO=0, Ref=3  
CFS:4: check Wait(), FNN=0x3981bd0 (CHECK), FOF=0x0, Wait=1, TO=0, Ref=2  
[Virus User account does not have proper Check Server privileges]

### **III. Virus Checker Service on DM Locks Up:**

User connected to C\$ share, modifies a file, DM informs Virus Checker of the "./" path which it cannot traverse, logging errors & possibly hanging server. Issue is that UNC path contains period within \\ entry. Server Logs bad path. Workaround--do not use hidden C\$ share to copy files to a datamover share. Code fixed at 2.2.49.1:  
SMB: 3: Checker 150.100.50.8 answers FAIL, Retrying  
SMB: 3: File status=OK  
SMB: 3: [av000] UNC=\\SERVER31\CHECK\$\.\mnt31\Share  
Properties\leicar.txt'  
SMB: 3: Checker 150.100.50.8 answers FAIL, Retrying  
SMB: 3: File status=OK  
    Checker 150.100.50.8 DOES NOT ANSWER, Retrying  
    Checker 150.100.50.8 answers FAIL, Retrying  
SMB: 3: File status=OK  
    Checker 150.100.50.8 DOES NOT ANSWER, Retrying  
SMB: 3: [av000] UNC=\\SERVER31\CHECK\$\.\mnt31\Share  
Properties\leicar.txt'

### **IV. Error when starting Virus Checker Service: # server\_setup server\_4 -Protocol viruschk -o start server\_4 :**

Not enough Vnodes to work !!  
Total Vnodes: 10240  
OfCache Vnodes: 15360  
Error: failed to complete command  
**Cause:** File may have been changed by NAS Upgrade, or was never updated  
/nas/server/slot\_4/file

**file initialize nodes=22528 dnlc=65536**

**Fix:** Change the file to the following values:

**file initialize nodes=65536 dnlc=262144**

Reboot or failover the Server. Virus Checker Service should now start.

## **V. EXCLUSION FILES:**

### **Following files are recommended for exclusion from scanning:**

.tmp; Access db files (.ldb, .mdb, etc); SQL db files (.sql, .sqr, etc); Exchange db files (.ost, .pst, etc); Oracle db files (.fmb, .frm)

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
Why? Database files in particular are continuously being written too and will degrade network performance [Scan by AV SW]

## **VI. HomeDir Files Cause Problems:**

A HomeDir config entry with the following line will cause Virus Checker problems:

emc:\*/mnt/guest/ [The extra “/” at the end of the config line is the cause of the problem!!!]

## **VII. RPCRetry & RPCRequests:**

Symptoms: Clients hang; general performance problems; VC does not answer, retrying; VC fails but logs ‘File status=OK’

Cause: AV Server does not respond to RPC requests

**Fix:** Increase parameters

RPCRequestTimeout=300000

RPCRetryTimeout=60000

## **VIII. Wait/TimeOut Error in Server Log:**

Server Log: check Wait/time-out(), FNN=0x25d98d0 (CHECK), writer=0x0, waiting=0, timeout=1Ref=6,timerPtr=0x0

Other Symptoms: Deadlock condition with VC Server; Loss of CIFS connection or CIFS Hangs

**Fix:** Ensure that there is a “waitForTimeout” value configured in viruschecker.conf; Upgrade to 2.2.57.0 +

## **IX. Issues with Office for XP:**

Users may have trouble trying to Save Excel files

Update to Office for XP SP1

## **X. AV Client Does not Have Check\$ Privilege:**

2004-05-15 13:59:48: VC: 3: Invalid access from client 129.225.82.143 to CHECK\$

2004-05-15 13:59:48: VC: 3: Client 129.225.82.143 (129.225.82.143) is not a VC server

**Note:** AV Server agent does not have correct privilege to access CHECK\$ share. Please note that this condition can actually block all SMB threads in NAS 5.2, as seen with the Server\_CIFS output. Solution is to verify AV configuration.

## **XI: Server Log Error:**

Server <ip addr> version 2 is ERROR\_AUTH → CAVA not running with EMC VChecking User account, does not have privilege to all VC Server to connect to DM. Try mapping to [\\DART\check\\$](\\DART\check$) and same error should be seen.

## **XII: Server Log Message Indicating that File has not been Scanned:**

2006-03-01 21:14:21: **VC: 3: 6:** \root\_vdm\_7\FS007\F\IFSDAT1\w5900001.www

## **OTHER WINDOWS ERRORS IN SERVER LOG:**

2007-01-28 22:33:30: VC: 3: 8: Server 10.120.17.155: ERROR\_AUTH 1115, RPC program version 3, CAVA release: 3.5.6, AV Engine: Symantec NAV CE

2007-02-03 11:21:09: VC: 3: 8: Server 10.120.17.155: ERROR\_AUTH 1208, RPC program version 3, CAVA release: 3.5.6, AV Engine: Symantec NAV CE

2007-02-03 12:58:23: VC: 3: 8: Server 10.120.17.154: ERROR\_AUTH 53, RPC program version 3, CAVA release: 3.5.6, AV Engine: Symantec NAV CE

## **WINDOWS LINK FOR WINDOWS ERRORS:**

<http://msdn2.microsoft.com/en-us/library/ms681381.aspx>

ERROR\_BAD\_NETPATH

53 The network path was not found.

ERROR\_SHUTDOWN\_IN\_PROGRESS

1115 A system shutdown is in progress.

ERROR\_EXTENDED\_ERROR

1208 An extended error has occurred.

## **HOMEDIR SCAN ISSUE:**

**emc:\*/mnt/guest/** [AntiVirus agent cannot scan files in the homedir using path similar to the line above]

## **BUG FIXES:**

\$server\_viruschk server\_x -a → created eiffel dump if output > than 16,384 characters [Fixed 2.2.46.x]

\$server\_setup server\_x -P viruschk -o stop → RPC timeout error if >than 300 files in processing Queue [Fixed 2.2.49.2]

User connected to default C\$ hidden Share and modified a file, DataMover could not handle path [Fixed 2.2.49.1]

Virus Check enabled, reboots take very long time [Fixed 2.2.49.1]

\$server\_viruschk server\_x -a → command fails or takes long time if I18N characters present [Fixed 2.2.49.1]

CIFS hangs if A/V service unavailable, eventhough configured for 'shutdown=no' [Fixed 2.2.53.4/4.0.17.1]

Viruschecker Running → user could not create Excel file [Fixed 2.2.53.5/4.0.17.1]

## **DISABLING CIFS FILENAME PATH CACHING:**

**param cifs pathCache=0**

**Note:** Setting this value to 0 forces VirusChecker to request complete paths from CIFS for files to be scanned—default is to cache this information. Some situations may occur where wrong path information is being cached on Data Mover as in following:

\$ more /nas/rootfs/slot\_2/.etc/viruschecker.audit

\mp\_users\ct\cifs\_data\rfa5630\data\maint\.\GEN\_LOG.txt

### **SETTING UP AV DEBUGGING ON AV SERVERS:**

1. Stop the EMC Checker Server Service (or CAVA Service in newer codes) on the AV Server
2. Edit Registry for CAVA Version 1.8.9:  
HKLM>Software>EMC>CheckerServer>Config: Set "verbose" and "debug" to "1"  
Edit Registry for CAVA Version 2.2.4:  
HKLM>Software>EMC>CAVA>Configuration
3. Download DebugView v4.31 [dbgview.exe] for PC's from [www.sysinternals.com](http://www.sysinternals.com) [Win32 Debug Program]
4. Open "dbgview.exe" on AV Server>Configure Log file>Click Icon to start Capture
5. Start the EMC CheckerServer Service
6. When experiencing performance problem:
  - a) - Obtain server\_log file during problem period
  - b) - Obtain output from "viruschk audit" command
  - c) - Get copy of current "viruschecker.conf" file
  - d) - Obtain output from AV Server using the 'debugview' application and saving output to logfile

**Note:** Debug View will display debugged output from kernel-mode or Win32 drivers

### **CONFIGURING EMC AV DRIVER IN WINDOWS 2000 REGISTRY:**

**HKLM>System>CurrentControlSet>Services>EMCVirCk**

**ErrorControl=1**

**Start=2**

**Type=1**

### **CONFIGURING CAVA 1.8.9 THREADS IN WINDOWS 2000 REGISTRY:**

**HKLM>Software>EMC>CheckerServer>Configuration>NumberOfThreads: *NumberOfThreads=20 (default)***

[Number of ONC-RPC threads available for VC Client; Make this setting the same as the Virus Checker Datamover setting of maxThreadWaiting=n]

### **CONFIGURING CAVA 2.2.4 THREADS IN REGISTRY:**

**HKLM>Software>EMC>CAVA>Configuration:**

**NumberOfThreads: REG\_DWORD: 0x14 [Default value=20 decimal; New rec.=64 decimal]**

**AgentType: REG\_SZ: Driver [Do not modify this value]**

### **CONFIGURING EMC AV DRIVER VIA REGISTRY:**

**HKLM>System>CurrentControlSet>Services>EMCVirCk:**

**ErrorControl=1**

**Start=2**

**Type=1**

### **SERVER LOG ERRORS:**

ERROR\_SETUP--Viruschecker Server setup incorrectly

AV\_NOT\_FOUND--Antivirus software could not be found or executed

FILE\_NOT\_FOUND--File not found or mapped – could not be checked

ACCESS\_DENIED--File access denied to the viruschecker --check AV User Account Rights

SUSPECT\_VIRUS--possible virus infection

DETECT\_VIRUS--Virus infection detected in file, type of virus not specified

MACRO\_VIRUS--Macro virus detected in file

### **EVENTS THAT TRIGGER VIRUS CHECKING:**

- DataMover Viruschecker Client determines when a file should be scanned; Sends request to scan file to CAVA via UNC path;
- CAVA opens file for Scanning; CAVA blocks User access during scan operation; Virus Scanning Engine Opens file again and scans
- Modifying & Closing an existing file
- Creating and saving a file
- Moving or Copying a file

- Restoring a file from BackUp
- Renaming a file with different extension
- Scan on Read if Access time is earlier than reference time for CIFS clients (later NAS version 4.2)

## AV FEATURES

### **SCAN ON FIRST READ:** Introduced with NAS 4.2

A time reference is stored in EMC Checker Server and CAVA uses this to determine if a file should be scanned when opened for Read. Specifically, DART uses ‘access’ time of file during an Open to determine if file should be scanned. The ‘access’ time is cross-referenced to the time stored in the config file and in the /etc/viruschecker.date file & is persistent until changed

--Typically employed after updating Virus Definition Files—will also scan compressed files such as .ZIP files

--Limited to CIFS Clients

**HIGHROAD SCAN:** To become available with NAS 5.0. Allows for VirusChecking of files written by High Road clients.

### **TURNING ON SCAN-ON-FIRST-READ (setting Reference Time):**

**\$server\_viruschk server\_x -set accesstime=0210261330.00 [YYMMDDHHMMSS]**

**Note:** Reference Time is stored in Memory and is /etc/viruschecker.date file. Scan on First Read is disabled by default, as seen when starting viruschecker with default configuration for first time:

**# server\_viruschk server\_3**

Scan on read disable

### **SETTING REFERENCE TIME TO CURRENT TIME:**

**\$server\_viruschk server\_x -set accesstime=now**

**# server\_viruschk server\_3**

Scan on read if access Time less than Wed May 18 21:16:57 2005 (GMT-00:00)

### **DISABLING REFERENCE TIME OR SCAN ON READ:**

**\$server\_viruschk server\_x -set accesstime=none or 0**

## **SIZING TOOL FOR CAVA:**

“cavamon.exe” tool is run on AV Engines to help determine load and gathers statistics; uses heuristics to determine workload and bases environment for an average 60% load per CAVA Server

### **SETTING UP SIZING TOOL:**

1. Create “cavamon.dat” file in Program Files>EMC>CAVA directory
  - add line item entry by IP or Hostname for each AV Engine Server that you wish to monitor
2. Enable Sizing Tool via Registry Entry:
  - a.) Start>Run>regedit>HKLM>software>EMC>cava>Sizing [Doubleclick and set Value to 1 to Enable]
  - b.) Doubleclick “SampleIntervalSecs” to set interval between 1-60 for CAVA to update Sizing Tool [Default = 10 secs]
3. Start Sizing/Monitoring Tool: Program Files>EMC>CAVA>cavamon.exe
  - a.) Click on “Get Stats” to begin gathering statistics
  - b.) To determine number of AV Servers recommended, click “Size” button. This Utility collects data for (10) consecutive interval periods and then displays number of recommended AV Engines

## **ANTIVIRUS MANAGEMENT FOR NAS 4.0 +:**

MMC Snap-In Module: Start>Programs>Administrative Tools>Celerra Management

**Note:** Celerra management only. AV Servers would need to be managed through the Vendor's interface

### **INSTALLING ANTIVIRUS MMC TOOL FOR CELERRA:**

Step 1. Copy executable install program from 4.0.x Celerra "/nas/cifs/CelerraVirusMgmt.exe" to WIN2K/NT 4.0 DC

Step 2. Install with defaults

## **CELERRA I18N INTERNATIONALIZATION:** *root\_fs\_common [aka .etc\_common]*

→Celerra directly supports the translation of UTF-8 (NFS & FTP Clients) & UTF-16 (CIFS Clients) without further configuration

→File & Directory names are always stored on-disk using UTF-8

→Celerra translates Files and Directory names obtained from clients into UTF-8 format for on-disk storage, and translates back into Client’s local encoding when Files or Directories are Read

→Celerra can be configured to translate various local encoding schemes

→Internationalization allows for files & directories to be created in a local language with names consistently presented to both NFS & CIFS clients

→I18N does not provide for one code page translation to another

→Non-UTF-8 NFS or FTP clients require that native encoding be identified for the Data Mover

**Byte Order Mapping:** BOM

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
Hardware platforms can present multibyte characters in different orders. Big-endian (FFFE) systems present high order byte first.  
Little-endian (FEFF) systems present low order byte first.

## **HOW CELERRA TRANSLATES & STORES UNICODE CHARACTERS:**

- I. Dart uses XLT translation table (defined xlt.cfg) to convert ASCII or Extended ASCII to Unicode (CIFS clients deliver Unicode)
- II. Dart must then convert Unicode to UTF-8 (char.-encoding scheme) to store On-Disk

**Note:** The xlt.cfg file is used for translation of NFS clients, as in the following sample

```
# cat /nas/site/locale/xlt.cfg
```

:::8859-1.txt: Any thing that didn't match above will be assumed to be latin-1

## **I18N FACTS:**

- Directories are converted when first accessed by users
- Only one Encoding type is supported per Data Mover for conversions [but up to (8) code translations can be used afterwards]
- Cannot unconvert from I18N once conversion takes place
- Supports only User-level access, not CIFS Share-level access
- Some encodings require RPQ: Latin-2; Korean/KSC; GB/GBK; Other non-UTF-8 encodings
- Limitation: Directory/Filenames limited to 255 bytes for UTF-8 formats

**CIFS I18N SUPPORT:** Celerra Supports I18N for File & Directory names created by Clients. Additionally, CIFS Clients are supported for User,Group & LocalGroup names. Share, Directory, & Domain names are only supported for ASCII. UNIX & SHARE security are not supported. Celerra Monitor is not supported.

**NFS/FTP I18N SUPPORT:** *Translates filename & directorynames--not mountpoints!*

**NAS 2.2+ & 4.0:** Supports International characters for File and Directory names.

## **I18N TERMINOLOGY:**

**INTERNATIONALIZATION:** Since Unicode can define a bit representation of every character of the world's alphabet, it allows for the ability to have consistent names & directories in various languages.

**CHARACTER SET:** The list of characters used, not fonts or Encoding

**ENCODING:** The mapping of character sets to bits used to represent data as stored on disk. Client encodings are defined by the xlt.cfg and character mapping files. All encodings are stored on Control Station in /nas/site directory. Configure DataMover encodings by DataMover; IP Subnet; IP Address; Hostname; Network protocol (NFS, FTP, All)

**UNICODE:** Unique code representation of all world language characters for computer processing. The Unicode Character Standard consists of the Basic Multilingual Plane (BMP), which consists of the first 64K characters of Unicode. The remaining Characters are contained in 17 different supplementary planes, each plane 64K in characters. UTF-32 is based on 32-bit code and references all Unicode 64k characters, and all the characters from all 17 supplementary planes.

## **UNICODE CHARACTER DATABASE:**

- Unicode version 2.0 is represented by the unidata2.txt file
- The Unicode Character Database defines default Unicode character properties and internal mappings
- File is plain ASCII text lines with fields terminated by semicolons
- Each line in the file represents the data for one encoded character in the 2.0 Unicode Standard
- So for Celerra, this file represents the Unicode definition source file

**LOCALIZATION:** Configuring machine to local language, character sets, number formatting, currency, date & time, etc.

## **UTF=UCS--UNIVERSAL CHARACTER SET-TRANSFORMATIONAL FORMAT:**

ISO 10646 defines UCS for mapping characters to most international languages onto integers between 0 and  $2^{31}$  (2,147,483,648).  
UTF-8 is a specific set of data representations that are capable of mapping UCS values into strings of ASCII hexadecimal values.  
UCS-2/UCS-4 are 16/32-bit Universal Character Sets, respectively

## **ASCII—American Standard Code for Information Interchange:**

A coding system that uses 8-bit numbers to represent letters, numbers, punctuation, & special symbols, and each character requires one byte for storage on-disk. ASCII standard applies only to codes 0-127 using lower 7 bits—8<sup>th</sup> bit is reserved for parity checking. ASCII codepage 437 and 850 are examples of ASCII codes that use high-order characters above 127.

## **UTF-8:**

On-disk disk encoding method used with Celerra, aka UCS Unicode Transformation Format. Similar to UTF-32, UTF-8 can reference the complete Unicode character space. The '8' represents 8-bit blocks used to represent each character, with a total number of blocks needed from 1-4. UTF-8 works with an algorithm that converts Unicode values to unique 1-4 byte sequences, with no embedded null characters. Latin1 is the default for Celerra, until NAS 5.4 [aka 8859-1.txt]. ASCII characters represented as 7-bit characters, and non-ASCII as 8-bit [ASCII is a subset of UTF-8, with values from 0-127]. UTF-8 supports WebNFS, FTP, TELNET, RLOGIN. ASCII values equal to or below 127 (hex 0x7F) only require 1 byte (1 8-bit block), since it uses only 7-bits to represent a character, which will be equivalent to the ASCII value itself. For characters equal to or below 2047 (hex 0x07FF) the UTF-8 representation will use two bytes. For characters from 2049-65535 (0xFFFF), UTF-8 will use 3 bytes.  
→ISO-8859-1 (Latin1) uses fixed length characters of 8 bits

## **16-bit UNICODE→UTF-16:**

UTF-16 is based on 16-bit little endian values, and is limited to 65,536 characters, [aka,UCS-2 Transformation Format], little-endian byte order, with no Byte Order mark, and no null-terminator (double byte encoding). Some of the 64k values are used to reference the 17 different supplemental Unicode planes. Used for Windows NT/2000—Windows uses Uniscribe program to reference surrogate code points for the 16-bit values reserved for the 17 supplemental Unicode planes. Celerra supports Unicode Character Database (unidata2.txt) and code page cp437.txt in the ./etc\_common/xlt directory by default--no configuration required. Celerra receives data in 16-bit format and stores on disk in UTF-8 format.

### **CODE PAGE 437:**

DOS code page used with the Windows Command prompt

Original character set of the IBM PC

Based on ASCII, but also contains International characters and Glyphs

C:\>chcp

Active code page: 437

→cp437.txt file on Celerra is the DOSLatinUS to Unicode conversion format, with DOSLatinUS in Hex, followed by Hex Unicode, then Unicode name

→UTF-16 uses character lengths of 16 bits (e.g., Windows NT)

### **EXAMPLES:**

#### **NFS/FTP UTF-8 CLIENTS:**

Celerra receives data & requires no translation to place on disk; code pages available for these clients are 8859-1; sjis; big5

#### **CIFS--WINDOWS UTF-16 CLIENTS:**

Celerra receives data in 16-bit Unicode format & translates to UTF-8 disk format & requires unidata2.txt & cp437.txt files to encode.

NFS Code Page Clients [8859-1 (Latin 1), sjis.txt, eucjp1, big5] are converted first to 16-bit Unicode from the code page stored in the /nas/site/locale directory, then stores onto disk in the UTF-8 format. When "Read Operations occur", converts Unicode to correct bytes prior to sending to Clients.

## **I18N/UTF-8 ENCODING LIMITATIONS:**

Share-Level Security not supported. Max length UTF-8 directory/file strings is 255 bytes.

Celerra Monitor/Manager not supported. Sharenames & Mountpoints not translated--remains ASCII.

### **COMMON TRANSLATION FILE SYSTEM:** Celerra specific file for translations is "xlt.cfg"

Certain character encoding configuration files located in ./etc\_common/xlt

All translation files reside in /nas/site/locale

Any changes to configuration must be made to /nas/site/locale & copied to ./etc\_common/xlt using /nas/sbin/uc\_config -update  
DataMover mounts the common file system as "root\_fs\_common", reads in the client-specific encoding translation files, and stores mappings for client encoding and UTF-8. \$/nas/rootfs/slot\_x/etc\_common/xlt is Read Only

## **COMMON TRANSLATION FILE SYSTEMS:**

**/nas/site/locale** →Contains translation files, which are then copied to DataMover “./etc\_common\xlt” using “uc\_config –update”

**./etc\_common/xlt** →Stores encoding files for interpretation by DART when I18N enabled

## **I18N DIRECTORIES: Common Translation File System Directory--→ ./etc\_common/xlt & /nas/site/locale**

**/nas/site/locale** Look for these files; cp437.txt; 8859-1.txt; sjis.txt; eucjp1.txt; big5.txt; unidata2.txt to be present

**./etc\_common/xlt** 8859-1.txt [default encoding=Latin1]; unidata2.txt & cp437.txt--supports CIFS clients [CIFS 16-bit Unicode conversions & Microsoft CIFS encoding, respectively]

**Note:** Translation files are stored in both above locations and should match between Control Station /nas/site/locale and the Data Mover in the ./etc\_common/xlt directory. Use uc\_config –update to synchronize the two locations.

**# /nas/sbin/uc\_config -i** [Use to display existing I18N configuration]

Common filesystem [root\_fs\_common] exists.

Common filesystem is presently mounted read-only on:

server\_2  
server\_3

Common filesystem is not mounted read-write anywhere.

**# /nas/sbin/uc\_config -l** [Use to list out translation files]

### **Control Station Directory /nas/site/locale:**

|            |   |          |          |        |              |             |
|------------|---|----------|----------|--------|--------------|-------------|
| -rwxrwxr-x | 1 | nasadmin | nasadmin | 10583  | Oct 18 11:35 | 8859-15.txt |
| -rwxrwxr-x | 1 | nasadmin | nasadmin | 9157   | Oct 18 11:35 | 8859-1.txt  |
| -rwxrwxr-x | 1 | nasadmin | nasadmin | 316745 | Oct 18 11:35 | big5.txt    |
| -rwxrwxr-x | 1 | nasadmin | nasadmin | 9843   | Oct 18 11:35 | cp437.txt   |
| -rwxrwxr-x | 1 | nasadmin | nasadmin | 332970 | Oct 18 11:35 | eucjp1.txt  |

```
-rwxrwxr-x 1 nasadmin nasadmin 303241 Oct 18 11:35 GB2312.txt
-rwxrwxr-x 1 nasadmin nasadmin 622893 Oct 18 11:35 jp-euc1.txt
-rwxrwxr-x 1 nasadmin nasadmin 364422 Oct 18 11:35 jp-euc.txt
-rwxrwxr-x 1 nasadmin nasadmin 293788 Oct 18 11:35 jp-pck1.txt
-rwxrwxr-x 1 nasadmin nasadmin 295686 Oct 18 11:35 jp-pck.txt
-rwxrwxr-x 1 nasadmin nasadmin 720675 Oct 18 11:35 kr-uhc.txt
-rwxrwxr-x 1 nasadmin nasadmin 173349 Oct 18 11:35 sjis.txt
-rwxrwxr-x 1 nasadmin nasadmin 444409 Oct 18 11:35 unidata2-old.txt
-rwxrwxr-x 1 nasadmin nasadmin 715838 Oct 18 11:35 unidata2.txt
-rwxrwxr-x 1 nasadmin nasadmin 1280 Sep 20 06:33 xlt.cfg
```

**Data Mover /etc common/xlt Directory:**

```
-rw-r--r-- 1 root bin 10583 Oct 18 11:34 8859-15.txt
-rw-r--r-- 1 root bin 9157 Oct 18 11:34 8859-1.txt
-rw-r--r-- 1 root bin 316745 Oct 18 11:34 big5.txt
-rw-r--r-- 1 root bin 9843 Oct 18 11:34 cp437.txt
-rw-r--r-- 1 root bin 332970 Oct 18 11:34 eucjp1.txt
-rw-r--r-- 1 root bin 303241 Oct 18 11:34 GB2312.txt
-rw-r--r-- 1 root bin 622893 Oct 18 11:35 jp-euc1.txt
-rw-r--r-- 1 root bin 364422 Oct 18 11:35 jp-euc.txt
-rw-r--r-- 1 root bin 293788 Oct 18 11:35 jp-pck1.txt
-rw-r--r-- 1 root bin 295686 Oct 18 11:35 jp-pck.txt
-rw-r--r-- 1 root bin 720675 Oct 18 11:35 kr-uhc.txt
-rw-r--r-- 1 root bin 173349 Oct 18 11:35 sjis.txt
-rw-r--r-- 1 root bin 444409 Oct 18 11:35 unidata2-old.txt
-rw-r--r-- 1 root bin 715838 Oct 18 11:35 unidata2.txt
-rw-r--r-- 1 root bin 1280 Oct 18 11:35 xlt.cfg
```

**Note:** If the two lists do not match, run uc\_config with -update option to update.

**VERIFYING DATAMOVER UTF-8 FORMAT FROM CONTROL STATION:**

1. Mount datamover root from Control Station
2. \$ls -d CA\* [Verify file names]
3. \$ls -d CA\* | od -t x1 [Obtain UTF-8 code]
4. **\$/nas/sbin/uc\_config -verify 192.168.1.100 -mover server\_2**  
**192.168.1.100 is UTF-8**

**Note:** Dart code returns UTF-8 from Control Station when Internal network is used

**VERIFYING UNICODE LOCALGROUPS DATABASE:**

Celerra uses .db.localgroups file when in Unicode mode, not localgroups.db, and contains this info:

```
# head .db.localgroups
#
# Localgroups database
#
$RELEASE:4$ →RELEASE: 4$ indicates Unicode mode
```

**CELLERRA I18N TRANSLATION FILES (CHARACTER MAPPING FILES)/nas/site/locale:**

| <u>NFS/FTP/<br/>CLIENTS</u> | <u>ENCODING/<br/>CHARACTER SET</u> | <u>TRANSLATION/<br/>FILE</u>                                                                                                                                  |
|-----------------------------|------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Yes                         | Latin-1                            | 8859-1.txt →Latin1 ASCII [Celerra default mapping file]. Conversion to UTF-8<br><b>Note:</b> Default 8-bit extension set to ASCII, aka “Latin Alphabet No. 1” |
| Yes                         | Latin-9                            | 8859-15.txt → West European languages, Euro sign, Latin-9, conversion UTF-8                                                                                   |
| Yes                         | Latin-2                            | None, RPQ →Central, East Europe, conversion UTF-8                                                                                                             |
| Yes                         | Shift-JIS or EUC-JP                | sjis.txt, eucjp1.txt, jp-euc.txt, jp-pck.txt →Japanese, conversion UTF-8                                                                                      |
| Yes                         | Korean/KSC                         | kr-uhc.txt →No UTF-8 conversion, Korean language                                                                                                              |
| Yes                         | GB or GBK                          | gb2312.txt →No UTF-8 conversion, Simplified Chinese                                                                                                           |
| Yes                         | Big 5                              | Big5.txt →Traditional Chinese with UTF-8 conversion                                                                                                           |
| Yes                         | Unicode UTF-8                      | None →No UTF-8 conversion, all languages                                                                                                                      |
| Yes                         | Other non-UTF-8                    | None →RPQ, conversion to UTF-8                                                                                                                                |

**CIFS CLIENTS**

|     |               |                                                              |
|-----|---------------|--------------------------------------------------------------|
| Yes | Unicode UCS-2 | unidata2.txt & cp437.txt →Conversion to UTF-8, all languages |
|-----|---------------|--------------------------------------------------------------|

**Note:** cp437.txt Microsoft character mapping file required for CIFS 16-bit Unicode Clients--installed by default

unidata2.txt File represents the Unicode Character Database for Unicode Version 2.0. Required by CIFS clients for 16-bit Unicode conversion & Upper-to-Lower case--installed by default

slt.cfg Celerra's translation file that defines how File & Directory names are translated into UTF-8--edit this file for the encoding format required if the default Latin1 is not to be used--for NFS/FTP Clients using Unix or CIFS

## **8859 TRANSLATIONS:**

Translation sets are given in Octal, Decimal, Hex, and Character representations, as in following “Latin-1” 8859-1:

**Oct Dec Hex Char Description**

|     |     |    |                                              |
|-----|-----|----|----------------------------------------------|
| 240 | 160 | A0 | NO-BREAK SPACE                               |
| 241 | 161 | A1 | ! INVERTED EXCLAMATION MARK                  |
| 242 | 162 | A2 | " CENT SIGN                                  |
| 243 | 163 | A3 | # POUND SIGN                                 |
| 244 | 164 | A4 | \$ CURRENCY SIGN                             |
| 245 | 165 | A5 | % YEN SIGN                                   |
| 246 | 166 | A6 | & BROKEN BAR                                 |
| 247 | 167 | A7 | ' SECTION SIGN                               |
| 250 | 168 | A8 | ( DIAERESIS                                  |
| 251 | 169 | A9 | ) COPYRIGHT SIGN                             |
| 252 | 170 | AA | * FEMININE ORDINAL INDICATOR                 |
| 253 | 171 | AB | + LEFT-POINTING DOUBLE ANGLE QUOTATION MARK  |
| 254 | 172 | AC | , NOT SIGN                                   |
| 255 | 173 | AD | - SOFT HYPHEN                                |
| 256 | 174 | AE | . REGISTERED SIGN                            |
| 257 | 175 | AF | / MACRON                                     |
| 260 | 176 | B0 | 0 DEGREE SIGN                                |
| 261 | 177 | B1 | 1 PLUS-MINUS SIGN                            |
| 262 | 178 | B2 | 2 SUPERSCRIPT TWO                            |
| 263 | 179 | B3 | 3 SUPERSCRIPT THREE                          |
| 264 | 180 | B4 | 4 ACUTE ACCENT                               |
| 265 | 181 | B5 | 5 MICRO SIGN                                 |
| 266 | 182 | B6 | 6 PILCROW SIGN                               |
| 267 | 183 | B7 | 7 MIDDLE DOT                                 |
| 270 | 184 | B8 | 8 CEDILLA                                    |
| 271 | 185 | B9 | 9 SUPERSCRIPT ONE                            |
| 272 | 186 | BA | : MASCULINE ORDINAL INDICATOR                |
| 273 | 187 | BB | ; RIGHT-POINTING DOUBLE ANGLE QUOTATION MARK |
| 274 | 188 | BC | < VULGAR FRACTION ONE QUARTER                |
| 275 | 189 | BD | = VULGAR FRACTION ONE HALF                   |
| 276 | 190 | BE | > VULGAR FRACTION THREE QUARTERS             |
| 277 | 191 | BF | ? INVERTED QUESTION MARK                     |
| 300 | 192 | C0 | @ LATIN CAPITAL LETTER A WITH GRAVE          |
| 301 | 193 | C1 | A LATIN CAPITAL LETTER A WITH ACUTE          |
| 302 | 194 | C2 | B LATIN CAPITAL LETTER A WITH CIRCUMFLEX     |
| 303 | 195 | C3 | C LATIN CAPITAL LETTER A WITH TILDE          |
| 304 | 196 | C4 | D LATIN CAPITAL LETTER A WITH DIAERESIS      |
| 305 | 197 | C5 | E LATIN CAPITAL LETTER A WITH RING ABOVE     |
| 306 | 198 | C6 | F LATIN CAPITAL LETTER AE                    |
| 307 | 199 | C7 | G LATIN CAPITAL LETTER C WITH CEDILLA        |
| 310 | 200 | C8 | H LATIN CAPITAL LETTER E WITH GRAVE          |
| 311 | 201 | C9 | I LATIN CAPITAL LETTER E WITH ACUTE          |
| 312 | 202 | CA | J LATIN CAPITAL LETTER E WITH CIRCUMFLEX     |
| 313 | 203 | CB | K LATIN CAPITAL LETTER E WITH DIAERESIS      |
| 314 | 204 | CC | L LATIN CAPITAL LETTER I WITH GRAVE          |
| 315 | 205 | CD | M LATIN CAPITAL LETTER I WITH ACUTE          |
| 316 | 206 | CE | N LATIN CAPITAL LETTER I WITH CIRCUMFLEX     |
| 317 | 207 | CF | O LATIN CAPITAL LETTER I WITH DIAERESIS      |
| 320 | 208 | D0 | P LATIN CAPITAL LETTER ETH                   |
| 321 | 209 | D1 | Q LATIN CAPITAL LETTER N WITH TILDE          |
| 322 | 210 | D2 | R LATIN CAPITAL LETTER O WITH GRAVE          |
| 323 | 211 | D3 | S LATIN CAPITAL LETTER O WITH ACUTE          |
| 324 | 212 | D4 | T LATIN CAPITAL LETTER O WITH CIRCUMFLEX     |
| 325 | 213 | D5 | U LATIN CAPITAL LETTER O WITH TILDE          |
| 326 | 214 | D6 | V LATIN CAPITAL LETTER O WITH DIAERESIS      |
| 327 | 215 | D7 | W MULTIPLICATION SIGN                        |
| 330 | 216 | D8 | X LATIN CAPITAL LETTER O WITH STROKE         |
| 331 | 217 | D9 | Y LATIN CAPITAL LETTER U WITH GRAVE          |
| 332 | 218 | DA | Z LATIN CAPITAL LETTER U WITH ACUTE          |
| 333 | 219 | DB | [ LATIN CAPITAL LETTER U WITH CIRCUMFLEX     |
| 334 | 220 | DC | \ LATIN CAPITAL LETTER U WITH DIAERESIS      |
| 335 | 221 | DD | ] LATIN CAPITAL LETTER Y WITH ACUTE          |
| 336 | 222 | DE | ^ LATIN CAPITAL LETTER THORN                 |
| 337 | 223 | DF | _- LATIN SMALL LETTER SHARP S                |
| 340 | 224 | E0 | ^- LATIN SMALL LETTER A WITH GRAVE           |

|     |     |    |   |                                      |
|-----|-----|----|---|--------------------------------------|
| 341 | 225 | E1 | a | LATIN SMALL LETTER A WITH ACUTE      |
| 342 | 226 | E2 | b | LATIN SMALL LETTER A WITH CIRCUMFLEX |
| 343 | 227 | E3 | c | LATIN SMALL LETTER A WITH TILDE      |
| 344 | 228 | E4 | d | LATIN SMALL LETTER A WITH DIAERESIS  |
| 345 | 229 | E5 | e | LATIN SMALL LETTER A WITH RING ABOVE |
| 346 | 230 | E6 | f | LATIN SMALL LETTER AE                |
| 347 | 231 | E7 | g | LATIN SMALL LETTER C WITH CEDILLA    |
| 350 | 232 | E8 | h | LATIN SMALL LETTER E WITH GRAVE      |
| 351 | 233 | E9 | i | LATIN SMALL LETTER E WITH ACUTE      |
| 352 | 234 | EA | j | LATIN SMALL LETTER E WITH CIRCUMFLEX |
| 353 | 235 | EB | k | LATIN SMALL LETTER E WITH DIAERESIS  |
| 354 | 236 | EC | l | LATIN SMALL LETTER I WITH GRAVE      |
| 355 | 237 | ED | m | LATIN SMALL LETTER I WITH ACUTE      |
| 356 | 238 | EE | n | LATIN SMALL LETTER I WITH CIRCUMFLEX |
| 357 | 239 | EF | o | LATIN SMALL LETTER I WITH DIAERESIS  |
| 360 | 240 | F0 | p | LATIN SMALL LETTER ETH               |
| 361 | 241 | F1 | q | LATIN SMALL LETTER N WITH TILDE      |
| 362 | 242 | F2 | r | LATIN SMALL LETTER O WITH GRAVE      |
| 363 | 243 | F3 | s | LATIN SMALL LETTER O WITH ACUTE      |
| 364 | 244 | F4 | t | LATIN SMALL LETTER O WITH CIRCUMFLEX |
| 365 | 245 | F5 | u | LATIN SMALL LETTER O WITH TILDE      |
| 366 | 246 | F6 | v | LATIN SMALL LETTER O WITH DIAERESIS  |
| 367 | 247 | F7 | w | DIVISION SIGN                        |
| 370 | 248 | F8 | x | LATIN SMALL LETTER O WITH STROKE     |
| 371 | 249 | F9 | y | LATIN SMALL LETTER U WITH GRAVE      |
| 372 | 250 | FA | z | LATIN SMALL LETTER U WITH ACUTE      |
| 373 | 251 | FB | { | LATIN SMALL LETTER U WITH CIRCUMFLEX |
| 374 | 252 | FC |   | LATIN SMALL LETTER U WITH DIAERESIS  |
| 375 | 253 | FD | } | LATIN SMALL LETTER Y WITH ACUTE      |
| 376 | 254 | FE | ~ | LATIN SMALL LETTER THORN             |
| 377 | 255 | FF |   | LATIN SMALL LETTER Y WITH DIAERESIS  |

## **COMPLETE 8859 ISO TRANSLATION SETS:**

The full set of ISO 8859 alphabets includes:

|             |                                                          |
|-------------|----------------------------------------------------------|
| ISO 8859-1  | west European languages (Latin-1)                        |
| ISO 8859-2  | east European languages (Latin-2)                        |
| ISO 8859-3  | southeast European and miscellaneous languages (Latin-3) |
| ISO 8859-4  | Scandinavian/Baltic languages (Latin-4)                  |
| ISO 8859-5  | Latin/Cyrillic                                           |
| ISO 8859-6  | Latin/Arabic                                             |
| ISO 8859-7  | Latin/Greek                                              |
| ISO 8859-8  | Latin/Hebrew                                             |
| ISO 8859-9  | Latin-1 modification for Turkish (Latin-5)               |
| ISO 8859-10 | Lappish/Nordic/Eskimo languages (Latin-6)                |
| ISO 8859-11 | Thai                                                     |
| ISO 8859-13 | Baltic Rim languages (Latin-7)                           |
| ISO 8859-14 | Celtic (Latin-8)                                         |
| ISO 8859-15 | west European languages (Latin-9)                        |

## **TRANSLATION/ENCODING SCHEMES:** [Modify "xlt.cfg" file to use]

|                                                   |                                                                                   |
|---------------------------------------------------|-----------------------------------------------------------------------------------|
| Using Default Translation for a Server:           | server_3:::8859-1.txt:datamover_3 defaults to latin1                              |
| Using NFS Protocol for Translation:               | :nfs:::8859-1.txt:nfs clients use latin1                                          |
| Using IP Address for Translation:                 | ::168.159.30.77::8859-1.txt:this client uses latin1                               |
| Using Subnet Address for Translation:             | ::168.159.30.0,255.255.255.0::8859-1.txt:clients on 168.159.30 network use latin1 |
| Using Hostname, DNS, or NIS Name for Translation: | ::hostone.emc.com::8859-1.txt:this host uses latin1                               |

## **UTF-8 ENCODING:** UTF=Unicode Transformation Format, an 8-bit universal character encoding standard

- UTF-8 employs algorithm to convert unicode values to a 1-4 byte sequence with no null characters
- Superset of ASCII [characters range in value 0-127 for ASCII]
- Supports WebNFS, FTP UTF-8, Telnet, Rlogin, NFS, CIFS clients without further translation
- CIFS clients use "unidata2.txt" and "cp437.txt" found in /etc\_common/xlt
- Backwards compatible with all ASCII characters as 7-bit, all non-ASCII rendered in 8-bit values
- CIFS clients obtain File & Directory info in 16-bit unicode--translates into UTF-8 format [up to 65,536 characters with unique 16-bitcode assigned to each character]
- NFS clients code page located in */nas/site/locale*

## **DEFAULT CELERRA I18N ENCODING:**

Latin-1 ASCII Encoding format for NFS/FTP [aka ISO 8859-1.txt]

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
However, using the xlt.cfg file, different & multiple encoding schemes can be supported based on DataMover, IP Subnet, IP Address, Hostname, or Network Protocol

### **EDITING XLT.CFG TRANSLATION FILE:**

Text format is comprised of (6) fields  
servername:protocol:address:reserved:translation:comment  
[server\_3:nfs:192.10.2.21::8859-1.txt:using default latin1 for translations]

### **TRANSLATIONS CONTAINED IN /NAS/SITE/LOCALE:**

***8859-1.txt, cp437.txt, 8859-15.txt, big5.txt, sjis.txt, eucjp1.txt, jp-euc.txt, jp-pck.txt***

Also found DataMover slots: #ls /nas/rootfs/slot-4/.etc\_common/xlt

Other mapping files available with RPQ

### **TRANSLATION RECOMMENDATION IF NFS ACCESS TO CIFS FILES REQUIRED:**

EMC recommends maintaining only a single character set as NFS clients can only access the first mapping entry for a specific character set.

### **DEFAULT TRANSLATIONS SUPPORTED:**

**CIFS:** File & Directory names, User & Group Names, Local Groups, but NOT Sharenames [seen as ASCII]

**NFS/FTP:** File & Directory Names, but NOT Mountpoints [seen as ASCII]

With UTF-8 turned on, translations are performed by DataMover lookups to the /.etc\_common/xlt directory

#### **How Does the Translation Work?**

NFS/FTP Client session; Celerra scans xlt.cfg file for translation match and if none found, applies default "latin-1" UTF-8

**MULTI-BYTE CHARACTERS:** NFS/FTP Clients use pure ASCII, and cannot read whole 'multi-byte' character sets but must read byte-by-byte, leading to differences in file name translations. Control Station cannot use multi-byte characters!

### **CONFIGURING I18N SUPPORT—New Install: \$/nas/sbin/uc\_config -i or -l**

**INTRO:** During Celerra software install, the default ASCII translation file and the /nas/site/locale directory are created on Control Station. Any other encoding schemes require editing of configuration files located in /nas/site/locale directory, and copying to /.etc\_common/xlt using the **\$/nas/sbin/uc\_config -update** command.

**UNIX HOSTS:** #locale [Command to display client language used]

**WINDOWS:** c:>codeagent.exe [Run utility to determine language used]

1. Verify /nas/site/locale & /.etc\_common/xlt directories are present, as well as files xlt.cfg, unidata2.txt, 8859-1.txt, cp437.txt, 8859-15.txt, big5.txt, sjis.txt, eucjp1.txt

**\$/nas/sbin/uc\_config -setup** ["Common Unicode translation subdirectory already exists"--means setup is correct]

2. Determine Client protocols and translations required for the site [e.g., non UTF-8 NFS/FTP clients require translations]

**Note:** If you cannot use the default 8859-1.txt mapping file [& require either sjis.txt; eucjp1.txt; big5.txt,etc], you will need to edit the "xlt.cfg" character-mapping file and load new translation files to the /nas/site/locale directory

3. Edit /nas/site/locale/xlt.cfg file for new translation requirements

4. Copy Translation Configuration Files to /.etc\_common/xlt using **\$/nas/sbin/uc\_config -update**

**Example:** **\$/nas/sbin/uc\_config -update big5.txt -mover server\_2** [Copies over Map file for Chinese character support]

**Caution:** If no filename is specified, all translation files including xlt.cfg are copied over

5. Verify your Configuration Files:

**#/nas/sbin/uc\_config -verify 192.10.202.12 | hostname | -mover server\_2**

[Verify client configuration by IP or Hostname against the DataMover configuration]

6. Verify that conversion will work: #/nas/sbin/uc\_config -setup ['Common Unicode translation subdirectory already exists']

7. Turning On I18N [New installs]:

**#/nas/sbin/uc\_config -on -mover server\_2**

**Caution:** Do not use this command for upgrading a Celerra from ASCII mode!

8. If an upgrade, perform steps 1-6 and then convert existing data to UTF-8

**I18N CONVERSIONS:** Conversions can be problematic. For this reason, all new Celerra Installs should have I18N enabled by default. Why? Many Customers use different types of encoding pages, or formats, such as 850 or 437, and by default, Celerra converts using a single Code Page format. Therefore, if multiple types are in use, when converting to I18N UNICODE, some files or directory names are 'changed'. If I18N were used upfront, prior to using other Code Pages and creating data on file systems, then there would be no issue.

**Important Point:** I18N Conversions do not corrupt data, but may change names of certain files & directories, which must then be manually renamed afterwards. Fortunately, there are tools available to check filesystems, based on customer code pages in use, prior

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
to conversions so as to identify any problem files. A post-conversion script can be run to correct all previously identified problem files.

**Note:** You can turn the uc\_config conversion process on again if there are files that have not been converted, for whatever reason.  
**CODE PAGES:** Support up to (8) different types of character translations--the fewer in use the better with respect to 'conversions'. However, during a conversion, only (1) local character type is allowed--after the conversion can configure for other encodings

## **I18N CONVERSION STEPS:**

1. Identify what Code Pages are being used by Customer and what the Client mix is. Run "uc\_check" script against FileSystems, using the Code pages in use to find which files/directories will be 'mis-converted'. Or, use "ls -R" and pipe output to file. Once identified, a post-conversion script can be used to repair the 'mis-converted' files.
2. Verify FileSystems under 85% capacity and number of inodes available [minimum of 10% free space required]

**Note:** Minimum of 300 inodes must be available on each filesystem to be converted.

3. Customer should conduct complete Tape BackUps of all File Systems

4. Run conversion using only a single Code Page encoding scheme—afterwards, can enable additional encodings

**CIFS Client:** c:\>uc\_check.exe -codepage cp850.txt -codepage cp437.txt \dm03\share01

**NFS Client:** ./uc\_check -codepage cp850.txt -codepage cp437.txt /mntpoint

**Note:** Preferable to run the following command from Control Station to DM as follows:

**# ls -lbR /nas/rootfs/slot\_2/fs1 >ls.out** [-b prints octal escapes for non-graphic characters]

5. Update translation files prior to running conversion:

**# /nas/sbin/uc\_config -update**

6. Kick off conversion:

**# /nas/sbin/uc\_config -convert start big5.txt -mover server\_2**

**Note:** File System access and performance may be impacted during the conversion process. Customers should try to limit access to file systems during the conversion. Takes about 1 minute per 10,000 inodes to convert a filesystem.

**Caution:** Remove the ascii filter param and do not reboot Data Movers until the convert command is started.

**Use following if enabling I18N on a prior ASCII Server & Keeping Default Code Page:**

**# /nas/sbin/uc\_config -convert start 8859-1.txt -mover server\_2**

7. Mount data mover from Control Station & start find commands against all filesystems to complete conversion of all files:

**# nohup find /mnt/fs1 -name "\*" -print > /nas/var/findlog\_fs1.txt &**

**# nohup find /mnt/fs2 -name "\*" -print > /nas/var/findlog\_fs2.txt &**

**# nohup find /mnt/fs3 -name "\*" -print > /nas/var/findlog\_fs3.txt &**

**Note:** Monitor conversion process using ps -ef or jobs command

8. After conversion completes, remove "param shadow asciifilter=1" from Data Movers and reboot.

9. Turn off the conversion process:

**# /nas/sbin/uc\_config convert -stop -m server\_2** (or would remain in progress)

10. Reboot all DataMovers

11. Run script or manually correct any mis-converted files

12. Run "uc\_check" to verify results [Running against NT Shares: c:\>uc\_check.exe -high -unusual \dm03\share01]

13. Configure additional client encoding schemes if required

## **I18N CONVERSION STEPS:**

**Note:** Conversion occurs during file/directory access by users

1. Verify /nas/site/locale & ./etc\_common/xlt directories and xlt.cfg configuration files are present

**\$/nas/sbin/uc\_config -setup** ["Common Unicode translation subdirectory already exists"--means setup is correct]

2. Determine Client protocols and translations required for the site [e.g., non UTF-8 NFS/FTP clients require translations]

**Note:** If you cannot use EMC-supplied mapping files [8859-1.txt; sjis.txt; eucjp1.txt; big5.txt], you will need to edit the "xlt.cfg" character-mapping file and provide new translation files in the /nas/site/locale directory]

3. Edit /nas/site/locale/xlt.cfg file for new translation requirements

4. Copy Translation Configuration Files to **/etc\_common/xlt** using **\$/nas/sbin/uc\_config -update**

Example1: **\$/nas/sbin/uc\_config -update big5.txt** [Updates only the configuration file referenced]

Example2: **\$/nas/sbin/uc\_config -update** [Updates all configuration files]

5. Verify your Configuration Files: **\$/nas/sbin/uc\_config -verify 192.10.202.12 | hostname | -mover server\_2**

[Verify client configuration by IP or Hostname against the DataMover configuration]

6. Upgrading the DataMover: **[Do not use \$uc\_config -on command]**

a) Start Conversion: **\$/nas/sbin/uc\_config -convert start big5.txt -mover server\_2**

b) Conversion for Whole Celerra: **\$/nas/bin/uc\_config -convert start big5.txt -mover ALL**

**Note:** Once turned on, the actual conversion occurs whenever file systems or directories are accessed. Only one encoding is supported per DataMover. Use the "find" command from Unix or NT to manually convert every file and directory immediately.

Conduct this during period of low I/O on system. FileSystems will not be available. Also, do not stop conversion prior to completion of file access by clients.

**Example:** UNIX Client: #find . / -name nosuchfilename -print  
NT Client: Start>Find>File and Folders

7. Turning Off the Conversion After Completion: #/nas/sbin/uc\_config -convert stop big5.txt -mover server\_2 [Do this only after the conversion process is done or else datamover will remain in conversion mode]

## **SAMPLE XLT.CFG FILE:**

```
# The line below identifies the meaning of each of the fields in this file
# all of these fields can have data in them except for the reserved field
# which must be left empty. (No characters, not even whitespace).
#
#<server>:<prot>:<addr>:<rsvd>:<translate>:<comment>
[server_3 or Hostname; Protocol 'All', 'NFS', or 'FTP'; ASCII Name, IP Address of Client or Subnet; name of mapping file]
#
# each of the lines below are an example of how to use some of the features
# a typical installation will use. These lines are ordered from most specific
# to least specific, since the first line that matches a client's request
# will be the one that actually applies. To adapt one of the lines below
# for a particular installation, simply remove the leading '#' and change the
# address field and translation file to be correct for the particular network.
#
# The line below shows how to identify a particular client which uses an
# encoding that is different from other clients.
#
#:168.159.6.108::8859-1.txt: This client use latin-1 (iso-8859-1)
#
# The line below is an example of how to identify a set of clients by subnet.
#
#:168.159.6.0.255.255.0::8859-1.txt:Clients on the 168.159.6.0 use latin-1
#
# the default is that all clients that don't match any of the earlier lines
# will be assumed to use latin-1
:::8859-1.txt: Any thing that didn't match above will be assumed to be latin-1
```

## **TROUBLESHOOTING I18N:** Most issues will be with the xlt.cfg map file

Use /nas/sbin/uc\_config -verify command to verify xlt.cfg settings

Common problems return "# %d" line errors for incorrect text configuration in the xlt.cfg file

Server Log entries are prefaced by either CFS, UFS, XLT subjects

## **I18N XLT FILE CORRUPTION ISSUES:**

Generally manifested by access issues to Shares that have mixed case characters. XLT files are stored in .etc\_common, and unknown actions can occasional cause corruption and loss of CIFS access to Shares. NAS 5.5 will have a maintenance release that will implement a CRC32 checksum feature that will store checksums of each XLT file in "xlt.crc". Whenever any XLT file is updated or accessed, the checksum will be compared against the xlt.crc. Upon corruption, DART will log an event, run "uc\_config -update" to copy files from /nas/site/locale, and generate a CallHome event.

**# .server\_config server\_2 -v "xlt crc <filename>"**

**Note:** Use command to verify individual XLT files

## **VERIFYING I18N CONFIGURATION ON DATA MOVER:**

**# .server\_config server\_2 -v "xlt verify 192.168.2.2"**

1040166829: LIB: 5: Querying LOCAL host file by name.

1040166829: LIB: 5: Querying NIS by name.

1040166829: LIB: 5: Querying DNS by name.

**1040166829: XLT: 3: 192.168.2.2 is UTF-8**

## **ALTERNATIVE COMMAND TO VERIFY I18N SETTINGS ON DM:**

**# /nas/sbin/uc\_config -verify 10.144.1.1 -mover server\_2**

server\_2 :

10.144.1.1, is sjis.txt

## **Determining Language/Encoding Requirements for Clients:**

1. Windows NT/2000: Run 'codeagent.exe' utility

2. Unix Clients: Run #locale

3. Windows Clients Using NFS/FTP (e.g., Hummingbird): Need encoding type from vendor

**Note:** .db.localgroups file is created when I18N is turned on and supersedes the "localgroups.db" file

## **OBSERVING UNPRINTABLE CHARACTERS FROM UNIX SYSTEMS:**

# ls -lb

### **SETTING I18N PARAMETERS FOR NDMP:**

**Note:** Set following parameter when Data Mover has I18N enabled, for all NAS versions that support I18N. This will ensure that the DM will use UTF-8 translation instead of the default “Latin 1”

**param NDMP dialect=**

## **CELERRA REPLICATION:** (Celerra Replicator/Celerra IP Replication)

→See Cognac section for Replicator V2 information

### **IP REPLICATION FEATURES:**

- Introduced with NAS 5.0, with RPQ
- Provides for Synchronous block-level replication of a file system
- NAS 5.0/5.1 does not support a Disaster Recovery mode of IP Replication
- NAS 5.1 does not support R/W on SFS file system [Upgrades from NAS 5.0 to 5.1 will require SFS be R/O]
- Provides for a READ-ONLY point-in-time async copy of a “snapshot” of the PFS to either a local or remote DataMover
- Not a Disaster Recovery Solution, but useful for a RO copy that can be used to Read or Backup to Tape
- 1-to-1 or 1-to-many IP/RDF (HighRoad); Remote Copy of FS RO & 'auto-refreshed' periodically
- Local to Remote IP Copy of a FS with periodic updating (Available via RPQ only)
- PFS must be built on slice volumes, identical in size to SFS or replication will not start [“fs1 size () is not equal to sfsfs1 size ()”]

**Comment:** Once implemented, Replication is automatic—does not require shell scripts to refresh SFS. Can copy filesystem replica copies from DM-to-DM locally, or DM-to-Remote DM.

### **BASIC IP REPLICATION FACTS TO KNOW ABOUT:**

- SavVols do not auto extend for IP Replication (for SnapSure feature, it does)
- Ports 8887/8888 must be open and unblocked on the both Target & Destination data movers
- Port 8000 must be open and available to Target & Destination Control Stations
- QOS is not implemented by default—by default, we ‘burst’ data across the link using RCP. Use QOS to slowdown or moderate the IP Replication over WANs so as not to overwork the link in high latency networks
- Always set the TO & HWM values on the Destination side at half the value as on the Source (defaults are 600/600)
- Changes are only ever written to the latest, or newest Checkpoint (never write to more than one checkpoint of same PFS)
- When there are multiple Checkpoints, they are read from Oldest to Newest, with bitmaps of each checked to see if flag is set to 1
- Except for failover & reverse, most commands are run strictly from the Source Replication side
- Source side creates checkpoints for use with fs\_copy
- Destination side creates the remote copy group structure
- Both sides create the IPFS file system
- Add reproto to boot.cfg file to enable restarting fs\_copy after DM reboots

### **USES FOR CELERRA REPLICATOR:**

- DR situations
- Backup data
- Data mining
- Software testing environments
- Migrations

### **IP REPLICATION TERMINOLOGY:**

**PFS:** Production File System

**SFS:** Secondary File System—goes through Freeze/Thaw during IP Replication sequence or when refreshes occur

**Ckpt:** check point filter

**Iosz:** 8k block size used by replication system—not changeable

**Copyiosz:** 32k size used to copy data to Save Volume—not changeable

**Delta Set:** Block modifications on PFS that are replicated to the SFS

**Log/Config Volume:** Redo Log--500MB in size. Used for recovery if PFS Server is rebooted, etc.

**Save Volume (SavVol):** Volume consisting of delta sets, which are modifications to the PFS to which blocks are saved in ‘chunks’ of 128MB. Min size of Save Volume is 3.2GB and max size is 90GB. This type of SavVol is different from SnapSure in that all changes & new data are written to SavVol from “Delta Sets” [SnapSure only writes first change to SavVol after a Checkpoint is taken]. SavVols cannot be extended on-line. Represents metavolume to which Delta sets are copied.

**Comment:** SavVol size limitation is based on the “Config File Size” limitation of 500MB

**HWM Option:** Represents the Data Size that serves as a ‘trigger’ to create a new “delta set” on the PFS side, which is then copied to the SFS side via the SavVol. IPFS automatically transfers new “delta sets” from SavVol to SFS. Represents the triggerpoint at which Replicator will update SFS with changes written to PFS.

**TimeOut Option:** Also used to trigger creation of new “delta sets” which are written to SavVol. Default=10 minutes. This value should never be set to less < 6 minutes for SFS or 3 minutes for PFS.

#### **NAS 5.2 & Higher:**

**QOS (Quality of Service):** Expressed in Kbps, originally intended as a way to prioritize among multiple Replication Sessions, which session should have greater priority, if using the same physical interface. In practice, QOS is used to limit the amount of bandwidth used by Replicator so as not to overwhelm a customer’s WAN or network link. By default, a Celerra Gbe interface will burst 128kb blocks of data at Gbe rates, in effect requiring that the network buffer up to 90% of a typical burst, resulting in packet loss and a severe drop in throughput performance due to retransmissions. When no value specified, default is to send data as fast as possible. As an example, if a WAN link is capable of 100mbps, this might be the QOS value to use for replication sessions, though even with QOS, each thread will send data at 32kb blocks.

→Each Replication Session uses 16 threads to handle data transfer

→Each thread can send 32kb blocks

**AUTORO:** Option that means when there is a hold on PFS flow, make PFS RO so as to prevent falling out of sync.

**AUOTFREEZE:** If there is an overflow on PFS, freeze the PFS so as to prevent falling out of sync.

### **REPLICATION POLICY PARAMETERS:**

**Timeout:** Generates or playsback Delta sets—default value is every 600 seconds. Value of 0 pauses replication.

**HWM:** Another threshold value of changes to PFS that generates or plays back a Delta set—default = 600MB (0 MB pauses)

**Note:** Replication Policies are created on PFS and SFS. First policy processed creates the Delta set on primary or copies to secondary (i.e., Delta Sets are created when either the HWM or Timeout policy is met).

### **DATA FLOW CONTROL TRIGGERS:**

→If SavVol becomes full, replication service stops copying to primary SavVol but continues to cache PFS changes until caught up

→Svc will hold up Delta sets if current Delta Set not yet written to SFS (i.e., Playback rate too slow on SFS side)

→If Delta Set is not transferred from Primary to Secondary SavVol, errors logged on primary Control Station: “Network down”

### **REPLICATION FUNCTIONS WITH NAS 5.4:**

→Destination dto= and dhwm= can now be set from the Source side

→Contains version 2.5 Checkpoints, which can have up to 64 Checkpoints

→Bitmaps remain in memory, but blockmaps are paged out to disk

→Upgrades from NAS 4.2 & Higher converts to new 2.5 checkpoint format

→Data Mover is capped at 1GB for checkpoint use

### **REPLICATION MEMORY STRUCTURES:**

→Replication uses both a “bitmap” and “block list” to track changes

→Replication uses “bitmap” in memory on PFS, from which delset changes are created, and then moved across to Target side. For new checkpoints, the “bitmap” values are all 0, then changes are written to the checkpoint (original fs block) and the bitmap value goes to 1. The “blocklist” tracks block changes via between original block that is modified and the block number in checkpoint. The “blocklist” is rebuilt during reboots from the info contained in all the checkpoints

### **BASIC IP REPLICATION TROUBLESHOOTING:**

# **nas\_cel -l** to list out Control Station IP’s, if remote IP replication is in use

# **fs\_copy -l** to observe if any copy sessions are in progress, run command on both sides to validate

# **fs\_copy -i id=76** [id=76 taken from fs\_copy -list and depends on which side you run the cmd]

# **fs\_replicate -l** to observe current state of replication sessions, run command on both sides to validate

# **fs\_replicate -i id=264 -v 10** to observe state of playback service for delta sets, etc. (10 indicates max lines to display)

**Note:** Command outputs in (3) Sections, with times in GMT.

Top Section = Source Side, regardless of whether cmd issued on Source or Target--lists HWM/TO values, as well as SavVol Status

Middle Section = Target Side, similar in info as the Top Section

Last Section = Info on network status and transfer rates for the replication session

### **IPREPv1 TROUBLESHOOTING COMMANDS:**

“**volcast create | delete |display**”

“**iprep\_svc start | stop | display | transfer | policy |stat | detail**”

“**replica |start | stop | recover | suspend | resume**”

**“playback start | stop | recover |suspend | resume”**

**“ipfscopy start | stop | display”**

## **COMMAND SYNTAX NAS 5.5:**

**# .server\_config server\_2 -v "replicate"**

```
1148667835: DPSVC: 4: replicate { start
    srcTargetName=<srcIqn> srcLun=<srcLun>
    [srcWBranchName=<srcWBranchName>]
    [dstTargetName=<dstIqn>] [dstLun=<dstLun>]
    [dstWBranchName=<dstWBranchName>]
    alias=<repAliasName>
    dstDicAuthAlias=<dstDicAuthenticationAliasName>
    [{verName=<fromVersionName> |
        verNum=<fromVerNumStr> }]
    [srcDataIp=<srcIp>] dstDataIp=<dstIp>
    [srcCtrlIp=<srcCtrlIp>] [dstCtrlIp=<dstCtrlIp>]
    [transportType=<rcp>]
    [crc=<True | False>] [qos=<kbytes/sec>]
    [applabel=<labelStr>] mode=<sync | async>
| abort
    {repName=<repName> | alias=<repAliasName> }
    [local=<True | False>]
    [internalStop=<True | False>] mode=<sync | async>
| mark
    {repName=<repName> | alias=<repAliasName> }
    {verName=<versionName> | verNum=<verNumStr>}
    mode=<sync | async>
| unmark
    {repName=<repName> | alias=<repAliasName> }
    {vername=<versionName> | verNum=<verNumStr>}
    mode=<sync | async>
| modify
    {repName=<repName> | alias=<repAliasName> }
    {[newAlias=<newRepAliasName>]
    [|srcDataIp=<srcIp>] [dstDataIp=<dstIp>]
    [srcCtrlIp=<srcCtrlIp>] [dstCtrlIp=<dstCtrlIp>]
    [crc=<True | False>] [qos=<kbites/sec>] }
    mode=<sync | async>
| validateip
    [srcCtrlIp=<srcIp>] dstCtrlIp=<dstIp>
    dstDicAuthAlias=<dicAuthenticationAliasName>
    mode=<sync | async>
| failover
    {repName=<repName> | alias=<repAliasName> }
    [srcCtrlIp=<srcIp>] [failoverType=<now>]
    mode=<sync | async>
| info
    {repName=<repName> [infoId=<infoId>] |
        alias=<repAliasName> [infoId=<infoId>] |
        <> }
    mode=<sync | async>
| deleteinfo
    {repName=<repName> infoId=<infoId> |
        alias=<repAliasName> infoId=<infoId> }
    mode=<sync | async>
| find
    {repName=<repName> | alias=<repAliasName> }
    mode=<sync | async>
| list
```

```

mode=<sync | async>
| display
  {repName=<repName> | alias=<repAliasName> ] |
    <> }
    displayType<info | stat>
  mode=<sync | async>
| copyreverse
  {repName=<repName> | alias=<repAliasName> }
  verName=<versionName>
  mode=<sync | async> }

# .server_config server_2 -v "vcs listversionset" | "vcs listversion vsid=<vsidStr>"
1148667962: VCS: 4: There are 0 version sets

```

## **QUERYING REMOTE CELERRA TO VERIFY HTTP COMMUNICATIONS:**

**# nas\_cel -l**

| id | name  | owner | mount_dev | channel       | net_path           | CMU |
|----|-------|-------|-----------|---------------|--------------------|-----|
| 0  | laip2 | 0     |           | 10.241.168.63 | APM000306008720000 |     |
| 2  | laip1 | 0     |           | 10.241.168.51 | APM000238010400000 |     |

**# nas\_cel -exec id=2 "nas\_disk -l"**

| id | inuse | sizeMB | storageID-devID     | type  | name       | servers                      |
|----|-------|--------|---------------------|-------|------------|------------------------------|
| 1  | y     | 11263  | APM00023801040-0000 | CLSTD | root_disk  | 1,2                          |
| 2  | y     | 11263  | APM00023801040-0001 | CLSTD | root_ldisk | 1,2                          |
| 3  | y     | 2047   | APM00023801040-0002 | CLSTD | d3         | 1,2                          |
| 4  | y     | 2047   | APM00023801040-0003 | CLSTD | d4         | 1,2-----output abridged----- |

## **FS REPLICATE SWITCHES AND USAGE:**

**# fs\_replicate -list | -info | -start | -abort | -refresh | -modify | -failover -o now -o sync | -resync | -reverse | -restart | -suspend**

**# fs\_replicate -list**

→lists out active replication sessions, from Source to Dest, & vice versa; inactive state indicates replication or playback service failure

**# fs\_replicate -info id=1 | -v 10**

→Use to obtain specific info on the health & state of current replication sessions & playback

**# fs\_replicate -start**

→Used to start replication for the first time, or to start again after a previously aborted replication session, requires that valid checkpoints are available. Destination fs must be rawfs & mounted RO. Options include to=, hwm=, qos=, and others to control transport and playback characteristics.

**# fs\_replicate -abort fs\_source,fs\_dest:cel=laip2**

→Run on Source side to stop replication & playback service.

**Note:** fs\_replicate -abort can be run separately on both Source and Destination side to abort replication, but the preferred method is to run the abort for both sides simultaneously

**#fs\_replicate -abort root\_fs\_vdm\_vdm\_source, root\_fs\_vdm\_vdm\_dest:cel=laip2**

**# fs\_replicate -refresh**

→Use to playback delta sets with options affecting to=, hwm=, and playuntildelta=; useful for kickstarting delta set copies from Source to Destination. Also used to generate new datasets on source, to be copied to destination—change HWM and Timeout values as needed. Use following option to playback only a specific number of datasets before stopping:

**\$fs\_replicate -refresh dest -option playuntildelta=91**

**#/nas/sbin/rootfs\_replicate -refresh root\_fs\_vdm\_server\_2\_vdm01 -o to=0,hwm=0**

**# fs\_replicate -modify**

→Use to modify to= (TimeOut values), hwm= (HighWaterMark), & qos= (bandwidth used in kbs/sec). If changing Source HWM & Timeout values without specifying Destination flow control values, the destination will inherit the Source values. Other -modify options are autorole=yes, autofreeze=yes.

**# fs\_replicate -modify root\_fs\_vdm\_vdm\_dest -o to=300,hwm=300**

**# fs\_replicate -modify fs\_source -o to=300,hwm=300**

### **# fs\_replicate -failover -o now**

→Used to perform Disaster Recovery failover when Source side is not available—executed from Destination side as asynchronous operation, meaning that there will be data loss. –failover -o now option conducts immediate failover without copying over source delta sets or playing back existing delta sets on target-target side goes RW immediately. At the completion of the failover, Replication is aborted and not running. If no failover options are specified, is called a ‘default’ failover, and incurs data loss, but does replay existing datasets on Target side before going RW. Use the –resync option to restart replication and copy incremental changes back to the original source in preparation for failing back (-reverse).

**Note:** Disaster failover options are –o now or default without any switches—run from Destination side.

### **# fs\_replicate -resync**

→Used to restart replication after disaster failover. –resync will first attempt to synchronize any incremental changes between Source & Destination replication sessions—original Source is running Playback Service and original Destination side is running the Replication Service. Without autocopy switch, default is to attempt to incrementally resync vdm rootfs and production file systems. If resync incremental copy fails, may need to use autofullcopy=yes option, which recopies all data back to file systems. Can also use options such as autofullcopy=yes, to=, hwm=, qos=, autoro=, autofreeze=. Use –reverse to change the flow of replication sessions. **Note:** The resync option does not change the direction of the current replication sessions—use the –reverse command to reverse the direction so that replication flows from original Source to Target, or vice versa. The –resync feature is broken in NAS 5.5. Basically, if a –resync is used, a full copy is performed because the the dataset query was moved from CS to DART, and DART does not record the correct dataset number on the destination checkpoint. This issue is fixed in NAS 5.5.24.0 and higher—see AR79069.

### **# fs\_replicate -failover -o sync**

-o sync option will synchronize Source & Destination file systems before failing over to prevent data loss—it does so by putting Source side file systems in RO state, copies all datasets from Source to Target side, then replays all datasets on Target side before Target goes RW—this is also known as a planned “administrative” failover that is synchronous and has no data loss. Executed from Source side & only if both Source & Destination sides are up and running.

### **# fs\_replicate -reverse**

```
# /nas/sbin/rootfs_replicate -reverse root_fs_vdm_vdm_dest:cel=laip2:if=rep_dest  
root_fs_vdm_vdm_source:if=rep_source
```

**Note:** Make sure you plug in the file system that is being failed over to first, then the source!

→Use this switch to perform a graceful failover from Source to Destination or from Destination to Source. This command is intended to failover replication without aborting any ongoing sessions. Also use options to=, hwm=, qos=, autoro=, autofreeze=. The –reverse essentially performs a resync of the file system (both Source & Destination fs becomes RO while changes are being transmitted) and then reverses the direction of the replication session. You would use the –reverse to perform an “administrative” or “planned” failover so as not to have any data loss, or to fallback to the original Source side to recover after a disaster failover and –resync was performed. Execute from the side that is currently running the replication service and not the playback service.

### **# fs\_replicate -restart**

→Use for planned downtime. Intended to restart replication after a –suspend was issued, or if replication is out-of-sync—as long as differential checkpoints exist, replication should be able to be restarted. Uses fs\_copy to send a differential checkpoint copy over to destination and then restart replication service, can also specify options for SavVol.

### **# fs\_replicate -suspend**

→Suspends the replication service for planned maintenance periods, creates source-side checkpoint, which is used to restart when the –restart command is run. Stops replication & playback services, but leaves fs in restartable state. –suspend must be followed by a –restart, which incrementally copies the checkpoint of source to destination.

## **SETTING UP IP REPLICATION WITH CIFS VDMs (NAS 5.5)**

### **DEBUG LOGGING:**

```
# export NAS_REPLICATE_DEBUG=1  
# export NAS_XML_DEBUG=1
```

**Note:** Dumps XML communications between DART & CS to screen and to log

### **1. VERIFY IP CONNECT BETWEEN LOCAL & REMOTE CONTROL STATIONS [LAIP1/LAIP2]:**

```
# ping 10.241.168.63
```

PING 10.241.168.63 (10.241.168.63) from 10.241.168.51 : 56(84) bytes of data.

64 bytes from 10.241.168.63: icmp\_seq=0 ttl=64 time=215 usec

**# nas\_cel -list**

```
id    name      owner mount_dev channel net_path      CMU
0    laip1      0          10.241.168.51 APM000238010400000
```

**2. SETUP REPLICATION RELATIONSHIP BETWEEN CONTROL STATIONS:**

Note: Run from Source side first, then target

**# /nas/sbin/nas\_rdf -init laip2 10.241.168.63** (Initializing Remote CS)

Contact laip2 ... is alive

Please create a new login account to manage RDF site laip2

New login: rep\_admin

New password:

Retype new password:

Changing password for user rep\_admin

passwd: all authentication tokens updated successfully

Please enter the passphrase for RDF site laip2:

Passphrase:

rep\_admin

Retype passphrase:

rep\_admin

operation in progress (not interruptible)...

id = 2

name = laip2

owner = 0

device =

channel =

net\_path = 10.241.168.63

celerra\_id = APM000306008720000

passphrase = rep\_admin

Note: Run from Source side to create “rep\_admin” with “rep\_admin” password--create passphrase as the account name  
sys log entry:

Feb 7 12:23:59 2006 OTHERS:7:1 nas\_rdf -nolog -init laip2 10.241.168.63

**# /nas/sbin/nas\_rdf -init laip1 10.241.168.51** (Initializing Local CS)

Contact laip1 ... is alive

Please create a new login account to manage RDF site laip1

New login: rep\_admin

New password:

Retype new password:

Changing password for user rep\_admin

passwd: all authentication tokens updated successfully

done

Please enter the passphrase for RDF site laip1:

Passphrase:

rep\_admin

Retype passphrase:

rep\_admin

operation in progress (not interruptible)...

id = 2

name = laip1

owner = 0

device =

channel =

net\_path = 10.241.168.51

celerra\_id = APM000238010400000

passphrase = rep\_admin

**3. VERIFY LOCAL & REMOTE CS REGISTRATIONS:**

**# nas\_cel -l**

```
id    name      owner mount_dev channel net_path      CMU
0    laip1      0          10.241.168.51 APM000238010400000
2    laip2      0          10.241.168.63 APM000306008720000
```

**# nas\_cel -l**

```
id name owner mount_dev channel net_path CMU
0 laip2 0 10.241.168.63 APM000306008720000
2 laip1 0 10.241.168.51 APM000238010400000
```

**Note:** The passphrase is stored on each Control Station in a file named after the Celerra serial number, in the following directories:  
/nas/http/replication/inbound/APM000306008720000  
/nas/http/replication/outbound/APM000306008720000

#### **4. SETUP DATA MOVER INTERFACES ON SOURCE & DESTINATION SYSTEMS:**

**Note:** For CIFS VDM replication, both Source & Destination CIFS interfaces must have the same name—IP addresses can be different

##### **SOURCE:**

```
reps protocol=IP device=cge0
inet=192.1.6.201 netmask=255.255.255.0 broadcast=192.1.6.255
UP, ethernet, mtu=1500, vlan=0, macaddr=8:0:1b:42:15:9d
```

##### **DESTINATION:**

```
reps protocol=IP device=cge0
inet=192.1.6.202 netmask=255.255.255.0 broadcast=192.1.6.255
UP, ethernet, mtu=1500, vlan=0, macaddr=8:0:1b:42:4e:3e
```

#### **5. SETUP DNS ON SOURCE & DESTINATION DATA MOVERS:**

**\$ server\_dns server\_2 <domain> <ip\_address>** (repeat on Destination mover)

#### **6. SETUP NTP TIME SERVICES ON SOURCE & DESTINATION SYSTEMS:**

**Note:** CS times must be within 10 minutes!

**# server\_date server\_2 timesvc start ntp 192.1.4.217**

server\_2 : done

**# server\_date server\_2 timesvc**

server\_2 :

Timeservice State

time: Tue Feb 7 12:58:05 EST 2006

type: ntp

sync delay: off

interval: 60

hosts: 192.1.4.217,

**Note:** Repeat NTP setup on Destination Data Mover

#### **7. SETUP LOCAL & REMOTE CONTROL STATIONS AS NTP CLIENT TO TIME SERVICE:**

**# /sbin/chkconfig ntpd --list**

ntpd 0:off 1:off 2:off 3:off 4:off 5:off 6:off

a.) Start NTP daemon on Source & Dest CS

**# /sbin/chkconfig --level 345 ntpd on**

**# /sbin/chkconfig ntpd --list**

ntpd 0:off 1:off 2:off 3:on 4:on 5:on 6:off

b.) Edit /etc/ntp.conf file--add the network NTP Server & comment out the other two lines as shown below:

#server 127.127.1.0 # local clock

#fudge 127.127.1.0 stratum 10 [Commented out so as not to point to itself as the primary NTP Service]

**server 192.1.4.217 minpoll 8 ## NTP Server**

c.) Add IP Address of NTP Server (s) to /etc/ntp/step-tickers file:

192.1.4.217

d.) Restart NTPD Daemon on Control Station:

**# /sbin/service ntpd restart**

```
Shutting down ntpd: [ OK ]
Synchronizing with time server: [ OK ]
```

Starting ntpd:

```
# ps -ef |grep ntp
ntp 15364 1 0 14:00 ? 00:00:00 ntpd -U ntp
root 16098 9015 0 14:01 pts/1 00:00:00 grep ntp
```

**# /sbin/service ntpd status**

ntpd (pid 28905) is running...

e.) Manually set Date & Time on Control Station:

**1. #date -s "4/09/03 16:15:15"**

**2. #/usr/sbin/setclock**

3. #date → Wed Apr 9 16:15:30 MDT 2003

Note: Repeat steps a-e for Destination Control Station, trying to get CS times as close to the Data Mover as possible

## **8. CONFIGURE INTERNAL USERMAPPER>PRIMARY ON DESTINATION DM & SECONDARY SERVICE ON SOURCE:**

Note: Setup Primary Usermapper service on Destination mover and Secondary service on Source mover

### **DESTINATION:**

```
# server_usermapper server_2 -enable
```

```
# server_usermapper server_2
```

server\_2 : Usrmapper service: Enabled

Service Class: Primary

### **SOURCE:**

```
# server_usermapper server_2 -enable primary=192.1.6.203
```

```
# server_usermapper server_2
```

server\_2 : Usrmapper service: Enabled

Service Class: Secondary

Primary = 192.1.6.203 (c) →"c" means that Secondary has contacted the Primary service

## **9. PREPARE SOURCE & DESTINATION VDM CIFS ROOT FILE SYSTEMS:**

Note: This step prepares the CIFS VDM Containers and loads the service at the appropriate state. CIFS Servers will be created at a later step on the Source side, and can only be created if the VDM is in a 'loaded' state.

### **a.) create Source & Dest VDM Containers:**

#### **SOURCE SIDE:**

```
# nas_server -name vdm_source -type vdm -create server_2 (source side)
```

id = 1

name = vdm\_source

acl = 0

type = vdm

server = server\_2

rootfs = root\_fs\_vdm\_vdm\_source →Actual name of rootfs

I18N mode = UNICODE

mountedefs =

member\_of =

status :

defined = enabled

actual = loaded, ready

Interfaces to services mapping:

```
[root@laip1 ntp]# nas_server -a -i -vdm
```

id = 1

name = vdm\_source

acl = 0

type = vdm

server = server\_2

rootfs = root\_fs\_vdm\_source

I18N mode = UNICODE

mountedefs =

member\_of =

status :

**defined = enabled**

**actual = loaded, ready** →Normal state of VDM container on Source side

Interfaces to services mapping:

```
[root@laip1 ntp]# server_mount server_2
```

server\_2 :

root\_fs\_2 on / ufs,perm,rw

root\_fs\_common on /.etc\_common ufs,perm,ro

root\_fs\_vdm\_source on /root\_vdm\_1/.etc ufs,perm,rw

```
# /nas/sbin/rootnas_fs -size root_fs_vdm_vdm_source
```

total = 111 avail = 111 used = 0 ( 0% ) (sizes in MB) ( blockcount = 262144 )

volume: total = 128 (sizes in MB) ( blockcount = 262144 )

#### **DESTINATION SIDE:**

**# nas\_server -name vdm\_dest -type vdm -create server\_2 -setstate**

**mounted pool=clar\_r5\_performance -option fstype=rawfs** → Destination VDM rootfs created Temp Unloaded state  
id = 1

name = vdm\_dest  
acl = 0  
type = vdm  
server = server\_2  
rootfs = root\_fs\_vdm\_dest  
I18N mode = UNICODE  
mounteddfs =  
member\_of =  
status :  
**defined = enabled**  
**actual = temporarily unloaded**

Interfaces to services mapping:

**[root@laip2 ntp]# nas\_server -a -i -vdm**

id = 1  
name = vdm\_dest  
acl = 0  
type = vdm  
server = server\_2  
rootfs = root\_fs\_vdm\_dest  
I18N mode = UNICODE  
mounteddfs =  
member\_of =  
status :  
defined = enabled  
**actual = temporarily unloaded**

Interfaces to services mapping:

**# server\_mount vdm\_dest**

vdm\_dest :

**Note:** The vdm root is NOT mounted on the VDM server, and shows up as <unmounted> and temporarily unloaded. Once the Source/Destination VDMs are replicated, and the destination rootfs is converted to uxf, the destination VDM becomes mounted.  
**[root@laip2 ntp]# server\_mount server\_2**

server\_2 :

root\_fs\_2 on / uxf,perm,rw  
root\_fs\_common on /.etc\_common uxf,perm,ro  
root\_fs\_vdm\_dest on /root\_vdm\_1/.etc rawfs,perm,ro,<unmounted>

**# /nas/sbin/rootnas\_fs -size root\_fs\_vdm\_vdm\_dest**

total = 111 avail = 111 used = 0 ( 0% ) (sizes in MB) ( blockcount = 262144 )

volume: total = 128 (sizes in MB) ( blockcount = 262144 )

## **10. VERIFY SOURCE & DESTINATION VDMs:**

a. Source Side:

**# nas\_server -a -i -vdm**

id = 1  
name = vdm\_source  
acl = 0  
type = vdm  
server = server\_2  
rootfs = root\_fs\_vdm\_source  
I18N mode = UNICODE  
mounteddfs =  
member\_of =  
status :  
**defined = enabled**  
**actual = loaded, ready**

Interfaces to services mapping:

**#server\_cifs server\_2**

CIFS service of VDM **vdm\_source (state=loaded)**

Home Directory Shares DISABLED

**# nas\_server -l -vdm**

```
id    acl server mountedfs   rootfs name
1     0   1           164   vdm_source
```

**# pwd**

/nas/server/slot\_2

**# cat vdm**

vdm add id=1 fs=164 name=vdm\_source

**# server\_df server\_2**

server\_2 :

| Filesystem             | kbytes | used | avail  | capacity | Mounted on       |
|------------------------|--------|------|--------|----------|------------------|
| root_fs_vdm_vdm_source | 114592 | 624  | 113968 | 1%       | /root_vdm_1/.etc |

**/nasmdc/quota/slot\_2**

drwxr-xr-x 3 root bin 80 Feb 7 13:37 root\_vdm\_1

**/nasmdc/quota/slot\_2/root\_vdm\_1/.etc****# ls -la**

```
drwxr-xr-x 2 root bin 1024 Feb 7 13:37 audit
-rw-r--r-- 1 root bin 42 Feb 7 13:37 .db.localgroups
dr-xr-xr-x 2 root bin 80 Feb 7 13:37 .etc
drwxr-xr-x 2 root bin 80 Feb 7 13:37 .filefilter
drwxr-xr-x 2 root root 8192 Feb 7 13:32 lost+found
drwxr-xr-x 2 root bin 80 Feb 7 13:37 secmap
drwxr-xr-x 5 root bin 1024 Feb 7 13:37 shares
```

**New Rootfs Created from VDM creation:****# nas\_fs -i id=164**

```
id      = 164
name   = root_fs_vdm_vdm_source
acl     = 0
in_use  = True
type    = uxf
worm   = off
volume  = v113309385
pool    = clar_r5_performance
member_of = root_avm_fs_group_3
rw_servers= server_2
```

**b. Target Side:****# nas\_server -a -i -vdm**

```
id      = 1
name   = vdm_dest
acl     = 0
type    = vdm
server  = server_2
rootfs  = root_fs_vdm_vdm_dest
I18N mode = UNICODE
mountedfs =
member_of =
status   :
```

**defined = enabled**

**actual = temporarily unloaded** → This is the normal state until vdm rootfs is diff. copied over, i.e., rootfs becomes uxf vs. rawfs  
 Interfaces to services mapping:

**#server\_cifs server\_2**

CIFS service of VDM vdm\_dest (state=loaded)

Home Directory Shares DISABLED

**# nas\_server -l -vdm**

```
id    acl server mountedfs   rootfs name
1     0   1           68    vdm_dest
```

**# pwd**

```
/nas/server/slot_2
# cat vdm
vdm add id=1 fs=68 name=vdm_dest
# server_df server_2
server_2 :
Filesystem      kbytes      used      avail capacity Mounted on
root_fs_vdm_vdm_dest
    114592       624     113968   1% /root_vdm_1/.etc
/nasmed/quota/slot_2
drwxr-xr-x  3 root  bin      80 Feb  7 13:37 root_vdm_1
/nasmed/quota/slot_2/root_vdm_1/.etc
# ls -la
drwxr-xr-x  2 root  bin      1024 Feb  7 13:37 audit
-rw-r--r--  1 root  bin      42 Feb  7 13:37 .db.localgroups
dr-xr-xr-x  2 root  bin      80 Feb  7 13:37 .etc
drwxr-xr-x  2 root  bin      80 Feb  7 13:37 .filefilter
drwxr-xr-x  2 root  root     8192 Feb  7 13:37 lost+found
drwxr-xr-x  2 root  bin      80 Feb  7 13:37 secmap
drwxr-xr-x  5 root  bin      1024 Feb  7 13:37 shares
New VDM Rootfs:
# nas_fs -i id=68
id      = 68
name   = root_fs_vdm_vdm_dest
acl     = 0
in_use = True
type   = udfs
worm   = off
volume = v174
pool   = clar_r5_performance
member_of = root_avm_fs_group_3
rw_servers= server_2
```

## **11. CREATE SOURCE & DESTINATION DATA FILE SYSTEMS FOR REPLICATION:**

**Note:** Source & Target replication file systems must be identical in size. Use the “samesize” option from CLI to create correct file system sizes on destination system and to specify “rawfs” since Celerra Manager does not create ‘rawfs’. For fs\_copy to work, the destination file systems need to be rawfs.

### a.) Create Source File Systems using CLI:

**# nas\_fs –name fs\_source –create size=50G pool=clar\_r5\_performance –option slice=yes**

#### cmd\_log shows fs\_source creation:

```
2006-02-07 14:11:50.721 db:0:28844:S: nas_fs -name fs_source -type udfs -create
size=50000M pool=id=3 storage=SINGLE worm=off -auto_extend yes -hwm 90% -max_size 75000M -vp no -option
slice=y,mover=server_2
2006-02-07 14:12:00.351 db:0:28844:E: nas_fs -name fs_source -type udfs -create
size=50000M pool=id=3 storage=SINGLE worm=off -auto_extend yes -hwm 90% -max_size 75000M -vp no -option
slice=y,mover=server_2
2006-02-07 14:12:02.654 server_2:0:28964:S: server_mountpoint server_2 -create /fs_source
2006-02-07 14:12:03.096 server_2:0:28964:E: server_mountpoint server_2 -create /fs_source
2006-02-07 14:12:05.081 server_2:0:29003:S: server_mount server_2 -option rw fs_source /fs_source
2006-02-07 14:12:06.995 server_2:0:29003:E: server_mount server_2 -option rw fs_source /fs_source
```

#### **#nas\_fs -l**

```
165      y  1  0  113309388 fs_source      1
```

#### **# nas\_fs -s fs\_source**

```
total = 49236 avail = 49235 used = 0 ( 0% ) (sizes in MB) ( blockcount = 102400000 )
volume: total = 50000 (sizes in MB) ( blockcount = 102400000 )
```

#### **# server\_df server\_2**

```
server_2 :
```

```
Filesystem      kbytes      used      avail capacity Mounted on
fs_source      50417984      576     50417408   0% /fs_source
```

**Note:** nas\_log.al.rll → Log shows file system deletions and changes conducted from CLI but NOT from WebUI

### b.) Mount Source file systems on VDM Server:

**[root@laip1 ntp]# server\_mount vdm\_source fs1 /fs1**

```
vdm_source : done  
[root@laip1 ntp]# server_mount vdm_source fs2 /fs2  
vdm_source : done
```

c.) Create destination file systems using CLI & samesize option:

```
# nas_fs -name fs_dest -type rawfs -create samesize=fs_source:cel=laip1 pool=clar_r5_performance -option slice=yes
```

**Note:** If using Celerra Manager to create the fs, the file system must be unmounted and then converted to rawfs. Normal state of file systems are rawfs and NOT mounted to the Destination VDM Server.

```
[root@laip2 log]# /nas/sbin/rootnas_fs -T rawfs root_fs_vdm_vdm_dest -F (was able to convert rootfs of VDM back to rawfs)
```

```
# nas_fs -T rawfs fs_dest -F
```

-----output abridged-----

## **12. VERIFY SIZES OF SOURCE & DEST FILE SYSTEMS:**

**SOURCE:**

```
# nas_fs -s id=165  
total = 49236 avail = 49235 used = 0 ( 0% ) (sizes in MB) ( blockcount = 102400000 )  
volume: total = 50000 (sizes in MB) ( blockcount = 102400000 )
```

**DEST:**

```
# nas_fs -s id=69  
total = 49236 avail = 49235 used = 0 ( 0% ) (sizes in MB) ( blockcount = 102400000 )  
volume: total = 50000 (sizes in MB) ( blockcount = 102400000 )
```

**Note:** At this point, all file systems on target side are created, unmounted, in rawfs state, with VDM service temporarily unloaded

## **13.) CREATE SOURCE CIFS VDM, JOIN TO DOMAIN, START CIFS, & CREATE SHARES:**

a.) # server\_cifs vdm\_source -add compname=source,domain=2k3.pvt.dns,interface=rep\_source

b.) # server\_cifs vdm\_source -Join compname=source,domain=2k3.pvt.dns,admin=tmatta

```
# server_cifs vdm_source
```

```
vdm_source :  
CIFS service of VDM vdm_source (state=loaded)  
Home Directory Shares DISABLED  
Usermapper auto broadcast enabled  
Usermapper[0] = [127.0.0.1] state:active (auto discovered)  
DOMAIN 2K3 FQDN=2k3.pvt.dns SITE=Default-First-Site-Name RC=3  
SID=S-1-5-15-bfc56af5-d6cf8701-5f67b1a3-fffffff  
>DC=GEORGE(192.1.4.217) ref=3 time=0 ms (Closest Site)  
CIFS Server SOURCE[2K3] RC=2  
Full computer name=source.2k3.pvt.dns realm=2K3.PVT.DNS  
Comment='EMC-SNAS:T5.5.18.2'  
if=rep_source l=192.1.6.204 b=192.1.6.255 mac=8:0:1b:42:15:9d  
FQDN=source.2k3.pvt.dns (Updated to DNS)
```

c.) Start CIFS Service:

```
# server_setup server_2 -P cifs -o start
```

d.) Create CIFS Shares for Source VDM Server:

```
# server_export vdm_source -P cifs -n fs_source /fs_source
```

```
# server_export vdm_source -P cifs -n fs_scli /fs_scli
```

```
# server_export vdm_source
```

```
vdm_source :  
share "fs_source" "/fs_source" maxusr=4294967295 umask=22  
share "fs_scli" "/fs_scli" maxusr=4294967295 umask=22
```

## **14. REPLICATE VDM ROOTFS FROM SOURCE TO DESTINATION:**

**SOURCE:**

a.) Create first baseline checkpoint of VDM root:

```
# /nas/sbin/rootfs_ckpt root_fs_vdm_vdm_source -Create
```

```
operation in progress (not interruptible)...id      = 164
```

```
name     = root_fs_vdm_vdm_source
```

```
acl      = 0
```

```
in_use   = True
```

```
type     = uxf
```

```
worm    = off
```

```
volume   = v113309385
```

```
pool     = clar_r5_performance
```

```

member_of = root_avm_fs_group_3
rw_servers= server_2
ro_servers=
rw_vdms =
ro_vdms =
auto_ext = no,virtual_provision=no
ckpts  = root_fs_vdm_vdm_source_ckpt1
b.) fs_copy the baseline rootfs checkpoint to destination:
# fs_copy -start root_fs_vdm_vdm_source_ckpt1:if=rep root_fs_vdm_vdm_dest:cel=laip2:if=rep -option convert=no

Note: Interface names should be the same
operation in progress (not interruptible)...id      = 168
name    = root_fs_vdm_vdm_source_ckpt1
acl     = 0
in_use   = True
type     = ckpt
worm    = off
volume   = vp113309395
pool     = clar_r5_performance
member_of = vpf167
rw_servers=
ro_servers= server_2
rw_vdms =
ro_vdms =
checkpt_of= root_fs_vdm_vdm_source Tue Feb 7 15:24:28 EST 2006
ipfs    = root_fs_vdm_vdm_source_ckpt1_ipfs1
used    = 6%
full(mark)= 90%
stor_devs = APM00023801040-0013,APM00023801040-0022,APM00023801040-0025,APM00023801040-002A
disks   = d22,d17,d31,d21
disk=d22 stor_dev=APM00023801040-0013 addr=c16t1l3      server=server_2
disk=d22 stor_dev=APM00023801040-0013 addr=c0t1l3      server=server_2
disk=d17 stor_dev=APM00023801040-0022 addr=c0t2l2      server=server_2
disk=d17 stor_dev=APM00023801040-0022 addr=c16t2l2      server=server_2
disk=d31 stor_dev=APM00023801040-0025 addr=c16t2l5      server=server_2
disk=d31 stor_dev=APM00023801040-0025 addr=c0t2l5      server=server_2
disk=d21 stor_dev=APM00023801040-002A addr=c0t2l10      server=server_2
disk=d21 stor_dev=APM00023801040-002A addr=c16t2l10      server=server_2
id     = 68
name    = root_fs_vdm_vdm_dest
acl     = 0
in_use   = True
type     = rawfs
worm    = off
volume   = v174
pool     = clar_r5_performance
member_of = root_avm_fs_group_3:laip2
rw_servers=
ro_servers= server_2
rw_vdms =
ro_vdms =
backup_of = root_fs_vdm_vdm_source Tue Feb 7 15:24:28 EST 2006
auto_ext = no,virtual_provision=no
ipfs    = root_fs_vdm_vdm_dest_ipfs1:laip2
stor_devs = APM00030600872-0006,APM00030600872-000B,APM00030600872-000E,APM00030600872-0015
disks   = d22,d12,d18,d8
disk=d22 stor_dev=APM00030600872-0006 addr=c16t1l15      server=server_2
disk=d22 stor_dev=APM00030600872-0006 addr=c0t1l15      server=server_2
disk=d12 stor_dev=APM00030600872-000B addr=c0t1l10      server=server_2
disk=d12 stor_dev=APM00030600872-000B addr=c16t1l10      server=server_2

```

```

disk=d18 stor_dev=APM00030600872-000E addr=c16t1l7    server=server_2
disk=d18 stor_dev=APM00030600872-000E addr=c0t1l7    server=server_2
disk=d8  stor_dev=APM00030600872-0015 addr=c0t1l2     server=server_2
disk=d8  stor_dev=APM00030600872-0015 addr=c16t1l2     server=server_2
IP Copy remaining (%) 100..Done.
done

```

**Note:** Only took a few minutes to accomplish for rootfs

c.) Start Replication for Source VDM rootfs (NOT the CKPT):

```
# fs_replicate -start root_fs_vdm_vdm_source:if=rep root_
fs_vdm_vdm_dest:cel=laip2:if=rep
```

→If run without any to or hwm indicated, sets TO & HWM to 600 by default

-----output abridged-----

### **Verify Rootfs Files:**

#### SOURCE:

```
# fs_replicate -l
```

Local Source Filesystems

| Id  | Source     | FlowCtrl | State  | Destination         | FlowCtrl | State    | Network |
|-----|------------|----------|--------|---------------------|----------|----------|---------|
| 173 | root_fs_vd | inactive | active | root_fs_vdm_vdm_des | inactive | inactive | alive   |

Local Destination Filesystems

| Id | Source | FlowCtrl | State | Dest. | FlowCtrl | State | Network |
|----|--------|----------|-------|-------|----------|-------|---------|
|----|--------|----------|-------|-------|----------|-------|---------|

#### SOURCE NETD FILE:

cifs start

```

rcproute add group=173_APM000238010400000 nexthop=192.1.4.143 fromhop=192.1.6.204
iprepsvc start 173_APM000238010400000 vol=113309402
migThreadPool action=startPool

```

#### SOURCE FS REPLICATE -i:

```
# fs_replicate -i id=173
```

```

id          = 164
name        = root_fs_vdm_vdm_source
fs_state    = active
type        = replication
replicator_state = active
source_policy   = NoPolicy
high_water_mark = 600
time_out      = 600
current_delta_set = 1
current_number_of_blocks = 2
flow_control   = inactive
total_savevol_space = 1048576 KBytes
savevol_space_available = 917504 KBytes (Before Flow Control)
id          = 68
name        = root_fs_vdm_vdm_dest:laip2
type        = playback
playback_state = inactive → Playback is inactive until the incremental checkpoint is sent & rootfs is converted to uxs
high_water_mark = N/A
time_out      = N/A
current_delta_set = N/A
flow_control   = inactive
total_savevol_space = 1048576 KBytes
savevol_space_available = 917504 KBytes (Before Flow Control)
outstanding delta sets: <None>
communication_state = alive
current_transfer_rate = ~ 896 Kbits/second
avg_transfer_rate   = ~ 896 Kbits/second
source_ip          = 192.1.6.204
source_port         = 59174
destination_ip     = 192.1.4.143
destination_port    = 8888
QOS_bandwidth      = 0 kbytes/sec

```

Note: All times are in GMT. Block size is 8 KBytes.

### **SOURCE FILESYS FILE TO THIS POINT:**

```
164:root_fs_vdm_vdm_source::0:3::y:1:113309385:1:::0:167,68@2,172:23::0:0:  
165:fs_source::0:3::y:1:113309388:v1:::0::23::90:75000:  
166:fs_scli::0:1::y:1:113309391:v1:::0::23::0:0:  
167:vpfs167::0:0::n:11:113309395::1139343866:164:2::23:168:  
168:root_fs_vdm_vdm_source_ckpt1::0:4::y:7::1:1139343868::0::167::0:0:  
171:vpfs171::0:0::n:11:113309402::::0::23:172:  
172:root_fs_vdm_vdm_source_rvfs::0:1::n:9:113309403::1139345597:164:4:173:171::0:0:  
173:root_fs_vdm_vdm_source_rvfs_ipfs1::0:1::y:5:0::1:1139345623:172:9::174::0:0:  
174:173_APMM000238010400000::0:1::n:100:0::::0::173,74@2:
```

### **DESTINATION:**

#### **# fs\_replicate -l**

Local Source Filesystems

| Id | Source | FlowCtrl | State | Destination | FlowCtrl | State | Network |
|----|--------|----------|-------|-------------|----------|-------|---------|
|----|--------|----------|-------|-------------|----------|-------|---------|

Local Destination Filesystems

| Id | Source | FlowCtrl | State | Dest. | FlowCtrl | State | Network |
|----|--------|----------|-------|-------|----------|-------|---------|
|----|--------|----------|-------|-------|----------|-------|---------|

```
74 root_fs_vdm_vdm_sou inactive active root_fs_vd inactive inactive alive
```

### **DEST NETD FILE:**

cifs start

volmcast create 73 173\_APMM000238010400000 185

### **DEST FILESYS FILE TO THIS POINT:**

```
68:root_fs_vdm_vdm_dest::0:7::y:5:174::1:1139343868:164@2:11:73:24::0:0:  
69:fs_dest::0:5::n:5:177::::0::24::90:75000:  
70:fs_dcli::0:3::n:5:180::::0::24::0:0:  
72:vpfs72::0:0::n:11:185::::0::24:73:  
73:root_fs_vdm_vdm_dest_pvfs::0:1::n:10:186::1139345614:68:5:74:72::0:0:  
74:root_fs_vdm_vdm_dest_pvfs_ipfs1::0:1::y:5:0::1:1139345637:73:7::174@2::0:0:
```

d.) Create 2<sup>nd</sup> Checkpoint on Source VDM rootfs, to be used for Incremental copy:

#### **# /nas/sbin/rootfs\_ckpt root\_fs\_vdm\_vdm\_source -Create**

operation in progress (not interruptible)...

```
id      = 175  
name    = root_fs_vdm_vdm_source_ckpt2  
acl     = 0  
in_use  = True  
type    = ckpt  
worm   = off  
volume  = vp113309395  
pool    = clar_r5_performance  
member_of = vpfs167
```

e.) Copy Incremental changes from 2<sup>nd</sup> Checkpoint compared to 1<sup>st</sup> Checkpoint to establish Replication Session:

**Note:** This time, copying changes between ckpt1 and ckpt2 to the destination rootfs VDM, and converting the destination vdm to udfs

#### **# fs\_copy -start root\_fs\_vdm\_vdm\_source\_ckpt2:if=rep root**

#### **\_fs\_vdm\_vdm\_dest:cel=laip2:if=rep -fromfs root\_fs\_vdm\_vdm\_source\_ckpt1**

operation in progress (not interruptible)...

#### **# fs\_copy -l**

Local Source Filesystems

| Id | Source | Destination | Status | %Remaining | CommState |
|----|--------|-------------|--------|------------|-----------|
|----|--------|-------------|--------|------------|-----------|

Local Destination Filesystems

| Id | Source | Destination | Status | %Remaining | CommState |
|----|--------|-------------|--------|------------|-----------|
|----|--------|-------------|--------|------------|-----------|

```
75 root_fs_vdm_vdm root_fs_vdm_vdm N/A N/A down →Normal during diff. fs_copy operation
```

**Note:** Output on DEST side when Incremental FS-COPY is underway on SOURCE

#### **# fs\_replicate -l**

Local Source Filesystems

| Id | Source | FlowCtrl | State | Destination | FlowCtrl | State | Network |
|----|--------|----------|-------|-------------|----------|-------|---------|
|----|--------|----------|-------|-------------|----------|-------|---------|

Local Destination Filesystems

| Id | Source | FlowCtrl | State | Dest. | FlowCtrl | State | Network |
|----|--------|----------|-------|-------|----------|-------|---------|
|----|--------|----------|-------|-------|----------|-------|---------|

```
74 root_fs_vdm_vdm_sou inactive active root_fs_vd inactive active alive
```

**Note:** fs\_replicate -list shows session alive and well

**15. VERIFY THAT REPLICATION SESSION IS RUNNING:****SOURCE:**

```
# fs_replicate -l
Local Source Filesystems
Id Source FlowCtrl State Destination FlowCtrl State Network
173 root_fs_vd inactive active root_fs_vdm_vdm_des inactive active alive
Local Destination Filesystems
Id Source FlowCtrl State Dest. FlowCtrl State Network
# fs_replicate -i id=173
id = 164
name = root_fs_vdm_vdm_source
fs_state = active
type = replication
replicator_state = active
source_policy = NoPolicy
high_water_mark = 600
time_out = 600
current_delta_set = 4
current_number_of_blocks = 2
flow_control = inactive
total_savevol_space = 1048576 KBytes
savevol_space_available = 917504 KBytes (Before Flow Control)
id = 68
name = root_fs_vdm_vdm_dest:laip2
type = playback
playback_state = active
high_water_mark = 600
time_out = 600
current_delta_set = 0
flow_control = inactive
total_savevol_space = 1048576 KBytes
savevol_space_available = 524288 KBytes (Before Flow Control)
```

outstanding delta sets:

Delta Source\_create\_time Blocks

```
----- -----
3 2006-02-07 16:33:23 2
2 2006-02-07 16:23:23 2
1 2006-02-07 16:13:23 2
0 2006-02-07 16:03:23 2
communication_state = alive
current_transfer_rate = ~ 896 Kbits/second
avg_transfer_rate = ~ 896 Kbits/second
source_ip = 192.1.6.204
source_port = 59174
destination_ip = 192.1.4.143
destination_port = 8888
QOS_bandwidth = 0 kbytes/sec
```

Note: All times are in GMT. Block size is 8 KBytes.

**DESTINATION SIDE:**

```
# fs_replicate -l
Local Source Filesystems
Id Source FlowCtrl State Destination FlowCtrl State Network
Local Destination Filesystems
Id Source FlowCtrl State Dest. FlowCtrl State Network
74 root_fs_vdm_vdm_sou inactive active root_fs_vd inactive active alive
# fs_replicate -i id=74
id = 164
name = root_fs_vdm_vdm_source:laip1
fs_state = active
type = replication
```

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!

```

replicator_state      = active
source_policy         = NoPolicy
high_water_mark      = 600
time_out              = 600
current_delta_set    = 4
current_number_of_blocks = 2
flow_control          = inactive
total_savevol_space  = 1048576 KBytes
savevol_space_available = 917504 KBytes (Before Flow Control)
id                   = 68
name                 = root_fs_vdm_vdm_dest
type                 = playback
playback_state       = active
high_water_mark      = 600
time_out              = 600
current_delta_set    = 0
flow_control          = inactive
total_savevol_space  = 1048576 KBytes
savevol_space_available = 524288 KBytes (Before Flow Control)
outstanding delta sets:
Delta Source_create_time  Blocks
----- -----
3   2006-02-07 16:33:23  2
2   2006-02-07 16:23:23  2
1   2006-02-07 16:13:23  2
0   2006-02-07 16:03:23  2
communication_state   = alive
current_transfer_rate = ~ 896 Kbits/second
avg_transfer_rate     = ~ 896 Kbits/second
source_ip             = 192.1.6.204
source_port            = 59174
destination_ip        = 192.1.4.143
destination_port       = 8888
QOS_bandwidth         = 0 kbytes/sec
$ nas_server -a -i -vdm
id                   = 3
name                 = vdm_dest
acl                  = 0
type                 = vdm
server               = server_2
rootfs              = root_fs_vdm_vdm_dest
I18N mode = UNICODE
mountedfs = fs_dest,fs_dcli
member_of =
status   :
defined = enabled
actual = mounted
Interfaces to services mapping:
interface=rep_source :cifs

```

**Note:** The VDM Replication session between Source & Target rootfs is established

# **server\_mount server\_2**

root\_fs\_vdm\_vdm\_dest on /root\_vdm\_1/etc udfs,perm,ro → VDM rootfs is now mounted on Server\_2

## **16. REPLICATING DATA FILE SYSTEMS FROM SOURCE TO DESTINATION:**

### **SOURCE FILE SYSTEMS:**

#**server\_mount server\_2**

fs\_source on /root\_vdm\_1/fs\_source udfs,perm,rw

fs\_scli on /root\_vdm\_1/fs\_scli udfs,perm,rw

### **MOUNT DESTINATION DATA FILE SYSTEMS AS RAWFS RO ON VDM SERVER:**

a.) Mount rawfs data file systems on Destination side to VDM as Read Only:

[root@laip2 root\_vdm\_1]# **server\_mountpoint vdm\_dest -create /fs1**

vdm\_dest : done

**[root@laip2 root\_vdm\_1]# server\_mountpoint vdm\_dest -create /fs2**

vdm\_dest : done

**[root@laip2 root\_vdm\_1]# server\_mount vdm\_dest -option ro fs1\_d /fs1**

vdm\_dest : done

**[root@laip2 root\_vdm\_1]# server\_mount vdm\_dest -option ro fs2 /fs2**

vdm\_dest : done

**CREATE BASELINE CHECKPOINTS ON SOURCE FILE SYSTEMS:****b.) →Create baseline Checkpoints for each Source file system:****\$ fs\_ckpt fs\_source -Create**

operation in progress (not interruptible)...id = 165  
 name = fs\_source  
 acl = 0  
 in\_use = True  
 type = uxfs  
 worm = off  
 volume = v113309388  
 pool = clar\_r5\_performance  
 member\_of = root\_avm\_fs\_group\_3  
 rw\_servers= server\_2  
 ro\_servers=  
 rw\_vdms = vdm\_source  
 ro\_vdms =  
 auto\_ext = hwm=90%,max\_size=75000M,virtual\_provision=no  
 ckpts = fs\_source\_ckpt1

**\$ fs\_ckpt fs\_scli -Create**

operation in progress (not interruptible)...id = 166  
 name = fs\_scli  
 acl = 0  
 in\_use = True  
 type = uxfs  
 worm = off  
 volume = v113309391  
 pool = clar\_r5\_performance  
 member\_of = root\_avm\_fs\_group\_3  
 rw\_servers= server\_2  
 ro\_servers=  
 rw\_vdms = vdm\_source  
 ro\_vdms =  
 auto\_ext = no,virtual\_provision=no  
 ckpts = fs\_scli\_ckpt1

**c.) →Copy the baseline Checkpoint (first) of each PFS over to Destination:****\$ fs\_copy -start fs\_source\_ckpt1:if=rep fs\_dest:cel=laip2:if=rep -option convert=no,monitor=off**

operation in progress (not interruptible)...id = 179

name = fs\_source\_ckpt1  
 acl = 0  
 in\_use = True  
 type = ckpt  
 worm = off  
 volume = vp113309407  
 pool = clar\_r5\_performance  
 member\_of =  
 rw\_servers=  
 ro\_servers= server\_2  
 rw\_vdms =  
 ro\_vdms = vdm\_source  
 ckpt\_of= fs\_source Wed Feb 8 13:13:26 EST 2006 -----output abridged-----  
 member\_of = root\_avm\_fs\_group\_3:laip2

**\$ fs\_copy -start fs\_scli\_ckpt1:if=rep fs\_dcli:cel=laip2:if=rep -option convert=no,monitor=off**

operation in progress (not interruptible)...id = 181  
name = fs\_scli\_ckpt1  
acl = 0  
in\_use = True  
type = ckpt  
worm = off  
volume = vp113309411  
pool = clar\_r5\_performance -----output abridged-----  
member\_of = root\_avm\_fs\_group\_3:laip2

**d.) Verify status of FSCopy:**

\$ **fs\_copy -l**

Local Source Filesystems

| Id         | Source                 | Destination          | Status         | %Remaining | CommState           |
|------------|------------------------|----------------------|----------------|------------|---------------------|
| <b>182</b> | <b>fs_source_ckpt1</b> | <b>fs_dest:laip2</b> | <b>started</b> | <b>100</b> | <b>transferring</b> |
| <b>184</b> | <b>fs_scli_ckpt1</b>   | <b>fs_dcli:laip2</b> | <b>started</b> | <b>100</b> | <b>transferring</b> |

**Note:** This output shows fs\_copy status during first fs\_copy of the first checkpoint, prior to starting the fs\_replication session.

\$ **fs\_copy -i id=182**

session\_id = 182  
source = fs\_source\_ckpt1  
destination = fs\_dest:laip2  
copy\_group = 182\_APM000238010400000  
status = started  
reason = N/A  
percent\_remaining = 100  
total\_blocks = 102400000  
copied\_blocks = 0  
communication\_state = transferring  
source\_ip\_address = 192.1.6.204  
target\_ip\_address = 192.1.4.143

\$ **fs\_copy -i id=184**

-----output abridged-----

**e.) Check Logs for FS\_COPY Progress:**

**OTHER LOGS TO CHECK FOR VERIFYING FS COPY STATUS:**

**nas\_log.al.tran** (source CS)

**nas\_log.al.remote** (on source CS)

**SOURCE SIDE NETD WITH ALL (3) FS NOW:**

reproute add group=173\_APM000238010400000 nexthop=192.1.4.143 fromhop=192.1.6.204  
reproute add group=182\_APM000238010400000 nexthop=192.1.4.143 fromhop=192.1.6.204  
reproute add group=184\_APM000238010400000 nexthop=192.1.4.143 fromhop=192.1.6.204  
ipfscopy start fsid=179 rcpg=182\_APM000238010400000 token=1139422631 address=0 noconvert=1  
ipfscopy start fsid=181 rcpg=184\_APM000238010400000 token=1139422813 address=0 noconvert=1  
iprepvc start 173\_APM000238010400000 vol=113309402

**CALCULATING FS COPY PROGRESS FOR DATA FS:**

[root@laip1 nasadmin]# .server\_config server\_2 -v "ipfscopy display fsid=181 rcpg=184\_APM000238010400000 token=1139422813"

1139424347: VMCAST: 5: ipfscopy fsid=181 rcpg=184\_APM000238010400000 token=1139422813 address=0  
1139424347: VMCAST: 5: DisplayIPFSCopy(): Token:1139422813 Group:184\_APM000238010400000 FSID:181  
1139424347: VMCAST: 4: ipfscopy:display() Token:1139422813 Group:184\_APM000238010400000 FSID:181 fromFSID:0  
1139424347: VMCAST: 4: VolMCast::display() Name:181\_184\_APM000238010400000 VolName:Sh113309411181  
NodeName:184\_APM000238010400000 refcount:1  
1139424347: VMCAST: 4: VolMCastSender::display() NextBlock:0 **TotalBlock:20971520** remote\_communication:Transferring  
**NextBeingProcessed:16784768**

1139424347: RCPD: 4: rcp:: Group:184\_APM000238010400000

1139424347: RCPD: 4: rcp:: Target\_num:0 Target\_ip:192.1.4.143 Local\_ip:192.1.6.204 Local\_port:59303

1139424347: VMCAST: 4: DisplayIPFSCopy(): State:0 ReasonCode:0

1139424347: ADMIN: 4: Command succeeded: ipfscopy display fsid=181 rcpg=184\_APM000238010400000 token=1139422813

**Note:** NextBeingProcessed divided by TotalBlock = 80% transfer done. Even with NAS 5.5, the traditional way to measure progress of the fs\_copy just doesn't work. Fs\_copy -l showed 100% remaining and 0% copied, but "ipfscopy" shows real progress.

**SERVER MOUNT WITH DEBUG SET SHOWS VARIOUS SERVICES IN ACTION FOR REPLICATION:**

# **export NAS\_REPLICATE\_DEBUG=1**

**# export NAS\_XML\_DEBUG=1**

**Note:** Dumps XML communications between DART & CS to screen and to log

**# server\_mount server\_2**

server\_2 :

```
root_fs_vdm_vdm_source on /root_vdm_1/etc ufs,perm,rw
fs_source on /root_vdm_1/fs_source ufs,perm,rw
fs_scli on /root_vdm_1/fs_scli ufs,perm,rw
root_fs_vdm_vdm_source_ckpt1 on /root_fs_vdm_vdm_source_ckpt1 ckpt,perm,ro
root_fs_vdm_vdm_source_rvfs_ipfs1 on /root_fs_vdm_vdm_source_rvfs_ipfs1 ipfs,temp,ro,<unmounted>
root_fs_vdm_vdm_source_ckpt2 on /root_fs_vdm_vdm_source_ckpt2 ckpt,perm,ro
fs_source_ckpt1 on /root_vdm_1/fs_source_ckpt1 ckpt,perm,ro
fs_scli_ckpt1 on /root_vdm_1/fs_scli_ckpt1 ckpt,perm,ro
fs_source_ckpt1_ipfs1 on /fs_source_ckpt1_ipfs1 ipfs,temp,ro,<unmounted>
fs_scli_ckpt1_ipfs1 on /fs_scli_ckpt1_ipfs1 ipfs,temp,ro,<unmounted>
```

**SOURCE:****# nas\_fs -l**

```
164 y 1 0 113309385 root_fs_vdm_vdm_sou 1
165 y 1 0 113309388 fs_source v1
166 y 1 0 113309391 fs_scli v1
167 n 11 0 0 vpf167
168 y 7 0 113309395 root_fs_vdm_vdm_sou 1
171 n 11 0 0 vpf171
172 n 9 0 113309403 root_fs_vdm_vdm_sou
173 y 5 0 0 root_fs_vdm_vdm_sou 1
174 n 100 0 0 173_APM000238010400
175 y 7 0 113309395 root_fs_vdm_vdm_sou 1
178 n 11 0 0 vpf178
179 y 7 0 113309407 fs_source_ckpt1 v1
180 n 11 0 0 vpf180
181 y 7 0 113309411 fs_scli_ckpt1 v1
182 y 5 0 0 fs_source_ckpt1_ipf 1
183 n 100 0 0 182_APM000238010400
```

**DEST SIDE:**

[nasadmin@laip2 nasadmin]\$ **fs\_copy -l**

Local Source Filesystems

| Id | Source | Destination | Status | %Remaining | CommState |
|----|--------|-------------|--------|------------|-----------|
|----|--------|-------------|--------|------------|-----------|

Local Destination Filesystems

| Id | Source | Destination | Status | %Remaining | CommState |
|----|--------|-------------|--------|------------|-----------|
|----|--------|-------------|--------|------------|-----------|

```
76 fs_source_ckpt1 fs_dest N/A N/A alive
77 fs_scli_ckpt1:l fs_dcli N/A N/A alive
```

[nasadmin@laip2 nasadmin]\$ **fs\_copy -i id=76**

```
session_id = 76
source = fs_source_ckpt1:laip1
destination = fs_dest
copy_group = 182_APM000238010400000:laip1
status = N/A
reason = N/A
percent_remaining = N/A
total_blocks = N/A
copied_blocks = N/A
communication_state = alive
source_ip_address = 192.1.6.204
target_ip_address = 192.1.4.143
```

[nasadmin@laip2 nasadmin]\$ **fs\_copy -i id=77**

```
session_id = 77
source = fs_scli_ckpt1:laip1
destination = fs_dcli
copy_group = 184_APM000238010400000:laip1
status = N/A
reason = N/A
```

```
percent_remaining = N/A
total_blocks      = N/A
copied_blocks     = N/A
communication_state = alive
source_ip_address = 192.1.6.204
target_ip_address = 192.1.4.143
```

**VOLMCAST ENTRIES CREATED IN NETD FILE FOR ALL (3) FILE SYSTEMS:**

```
volmcast create 73 173_APM000238010400000 185
```

```
volmcast create 69 182_APM000238010400000 177
```

```
volmcast create 70 184_APM000238010400000 180
```

```
[root@laip2 nasadmin]# export NAS_REPLICATE_DEBUG=1
```

```
[root@laip2 nasadmin]# server_mount server_2
```

```
server_2 :
```

```
root_fs_vdm_vdm_dest_pvfs_ipfs1 on /root_fs_vdm_vdm_dest_pvfs_ipfs1 ipfs,temp,ro,<unmounted>
root_fs_vdm_vdm_dest on /root_vdm_1/etc uxfss,perm,ro
fs_dest on /fs_dest rawfs,perm,ro,<unmounted>
fs_dcli on /fs_dcli rawfs,perm,ro,<unmounted>
fs_dest_ipfs1 on /fs_dest_ipfs1 ipfs,temp,ro,<unmounted>
fs_dcli_ipfs1 on /fs_dcli_ipfs1 ipfs,temp,ro,<unmounted>
```

f.) →Start Replication Sessions for Production File Systems after 1<sup>st</sup> fs\_copy completes:

```
[root@laip1 nasadmin]# fs_replicate -start fs_source:if=rep fs_dest:cel=laip2:if=rep
```

```
operation in progress (not interruptible)...id = 165
```

```
name = fs_source
```

```
acl = 0
```

```
in_use = True
```

```
type = uxfss
```

```
worm = off
```

```
volume = v113309388
```

```
pool = clar_r5_performance
```

```
member_of = root_avm_fs_group_3
```

```
rw_servers= server_2
```

```
ro_servers=
```

```
rw_vdms = vdm_source
```

```
ro_vdms =
```

```
auto_ext = hwm=90%,max_size=75000M,virtual_provision=no
```

```
ckpts = fs_source_ckpt1
```

```
replicas = fs_source_rvfs
```

```
ip_copies = fs_dest:laip2
```

```
stor_devs = APM00023801040-0027,APM00023801040-0020,APM00023801040-0019,APM00023801040-001E
```

```
disks = d32,d16,d25,d15
```

```
disk=d32 stor_dev=APM00023801040-0027 addr=c16t217 server=server_2
```

```
disk=d32 stor_dev=APM00023801040-0027 addr=c0t217 server=server_2
```

```
disk=d16 stor_dev=APM00023801040-0020 addr=c0t210 server=server_2
```

```
disk=d16 stor_dev=APM00023801040-0020 addr=c16t210 server=server_2
```

```
disk=d25 stor_dev=APM00023801040-0019 addr=c16t119 server=server_2
```

```
disk=d25 stor_dev=APM00023801040-0019 addr=c0t119 server=server_2
```

```
disk=d15 stor_dev=APM00023801040-001E addr=c0t1114 server=server_2
```

```
disk=d15 stor_dev=APM00023801040-001E addr=c16t1114 server=server_2
```

```
id = 69
```

```
name = fs_dest
```

```
acl = 0
```

```
in_use = True
```

```
type = rawfs
```

```
worm = off
```

```
volume = v177
```

```
pool = clar_r5_performance
```

```
member_of = root_avm_fs_group_3:laip2
```

```
rw_servers=
```

```
ro_servers= server_2
```

```
rw_vdms =
```

```

ro_vdms =
backup_of = fs_source Wed Feb 8 13:13:26 EST 2006
auto_ext = hwm=90%,max_size=75000M,virtual_provision=no
replicas = fs_dest_pvfs
stor_devs = APM00030600872-0008,APM00030600872-000D,APM00030600872-0012,APM00030600872-0017
disks = d21,d11,d17,d7
disk=d21 stor_dev=APM00030600872-0008 addr=c16t1l13 server=server_2
disk=d21 stor_dev=APM00030600872-0008 addr=c0t1l13 server=server_2
disk=d11 stor_dev=APM00030600872-000D addr=c0t1l8 server=server_2
disk=d11 stor_dev=APM00030600872-000D addr=c16t1l8 server=server_2
disk=d17 stor_dev=APM00030600872-0012 addr=c16t1l5 server=server_2
disk=d17 stor_dev=APM00030600872-0012 addr=c0t1l5 server=server_2
disk=d7 stor_dev=APM00030600872-0017 addr=c0t1l0 server=server_2
disk=d7 stor_dev=APM00030600872-0017 addr=c16t1l0 server=server_2
done

```

[root@laip1 nasadmin]# **fs\_replicate -start fs\_scli:if=rep fs\_dcli:cel=laip2:if=rep**

-----output abridged-----

### **VERIFY PRODUCTION FS REPLICATION STATUS ON SOURCE:**

#### **SOURCE SIDE:**

[root@laip1 log]# **fs\_replicate -l**

Local Source Filesystems

| Id | Source | FlowCtrl | State | Destination | FlowCtrl | State | Network |
|----|--------|----------|-------|-------------|----------|-------|---------|
|----|--------|----------|-------|-------------|----------|-------|---------|

**254 fs\_scli inactive active fs\_dcli:laip2 inactive inactive alive** → State of Replication session after 1<sup>st</sup> checkpoint has been copied over to destination, and right after the Replication Session has been started

#### **SOURCE SIDE:**

[root@laip1 nasadmin]# **fs\_replicate -l**

Local Source Filesystems

| Id | Source | FlowCtrl | State | Destination | FlowCtrl | State | Network |
|----|--------|----------|-------|-------------|----------|-------|---------|
|----|--------|----------|-------|-------------|----------|-------|---------|

173 root\_fs\_vd inactive active root\_fs\_vdm\_vdm\_des inactive active alive

188 fs\_source inactive active fs\_dest:laip2 inactive inactive alive

192 fs\_scli inactive active fs\_dcli:laip2 inactive inactive alive

Local Destination Filesystems

| Id | Source | FlowCtrl | State | Dest. | FlowCtrl | State | Network |
|----|--------|----------|-------|-------|----------|-------|---------|
|----|--------|----------|-------|-------|----------|-------|---------|

#### **DEST SIDE:**

[root@laip2 nasadmin]# **fs\_replicate -l**

Local Source Filesystems

| Id | Source | FlowCtrl | State | Destination | FlowCtrl | State | Network |
|----|--------|----------|-------|-------------|----------|-------|---------|
|----|--------|----------|-------|-------------|----------|-------|---------|

Local Destination Filesystems

| Id | Source | FlowCtrl | State | Dest. | FlowCtrl | State | Network |
|----|--------|----------|-------|-------|----------|-------|---------|
|----|--------|----------|-------|-------|----------|-------|---------|

74 root\_fs\_vdm\_vdm\_sou inactive active root\_fs\_vd inactive active alive

80 fs\_source:laip1 inactive active fs\_dest inactive inactive alive

83 fs\_scli:laip1 inactive active fs\_dcli inactive inactive alive

**Note:** Again, both PFS will show as Network “inactive” until the 2<sup>nd</sup> (incremental) checkpoint is copied over

g.) → Create Incremental 2<sup>nd</sup> Checkpoint for each PFS being setup for IP Replication:

[root@laip1 nasadmin]# **fs\_ckpt fs\_source -Create**

operation in progress (not interruptible)...id = 165

name = fs\_source

acl = 0

in\_use = True

type = uxf

worm = off

volume = v113309388

pool = clar\_r5\_performance

member\_of = root\_avm\_fs\_group\_3

rw\_servers= server\_2

ro\_servers=

rw\_vdms = vdm\_source

ro\_vdms =

auto\_ext = hwm=90%,max\_size=75000M,virtual\_provision=no

ckpts = fs\_source\_ckpt1,fs\_source\_ckpt2

```

replicas = fs_source_rvfs
ip_copies = fs_dest:laip2
stor_devs = APM00023801040-0027,APM00023801040-0020,APM00023801040-0019,APM00023801040-001E
disks = d32,d16,d25,d15
disk=d32 stor_dev=APM00023801040-0027 addr=c16t2l7 server=server_2
disk=d32 stor_dev=APM00023801040-0027 addr=c0t2l7 server=server_2
disk=d16 stor_dev=APM00023801040-0020 addr=c0t2l0 server=server_2
disk=d16 stor_dev=APM00023801040-0020 addr=c16t2l0 server=server_2
disk=d25 stor_dev=APM00023801040-0019 addr=c16t1l9 server=server_2
disk=d25 stor_dev=APM00023801040-0019 addr=c0t1l9 server=server_2
disk=d15 stor_dev=APM00023801040-001E addr=c0t1l14 server=server_2
disk=d15 stor_dev=APM00023801040-001E addr=c16t1l14 server=server_2
id = 194
name = fs_source_ckpt2
acl = 0
in_use = True
type = ckpt
worm = off
volume = vp113309407
pool = clar_r5_performance
member_of = vpfsl78
rw_servers=
ro_servers= server_2
rw_vdms =
ro_vdms = vdm_source
checkpt_of= fs_source Wed Feb 8 16:18:22 EST 2006
used = 2%
full(mark)= 90%
delta_number= 1
stor_devs = APM00023801040-0013,APM00023801040-0022,APM00023801040-0025,APM00023801040-002A
disks = d22,d17,d31,d21
disk=d22 stor_dev=APM00023801040-0013 addr=c16t1l3 server=server_2
disk=d22 stor_dev=APM00023801040-0013 addr=c0t1l3 server=server_2
disk=d17 stor_dev=APM00023801040-0022 addr=c0t2l2 server=server_2
disk=d17 stor_dev=APM00023801040-0022 addr=c16t2l2 server=server_2
disk=d31 stor_dev=APM00023801040-0025 addr=c16t2l5 server=server_2
disk=d31 stor_dev=APM00023801040-0025 addr=c0t2l5 server=server_2
disk=d21 stor_dev=APM00023801040-002A addr=c0t2l10 server=server_2
disk=d21 stor_dev=APM00023801040-002A addr=c16t2l10 server=server_2

```

[root@laip1 nasadmin]# fs\_ckpt fs\_scli -Create

-----output abridged-----

h.) →Perform incremental fs\_copy to Destination for each PFS, letting fs's convert to ufs on Destination side:

[root@laip1 nasadmin]# fs\_copy -start fs\_source\_ckpt2:if=rep fs\_dest:cel=laip2:if=rep -fromfs fs\_source\_ckpt1 -option monitor=off

-----output abridged-----

[root@laip1 log]# fs\_copy -l

Local Source Filesystems

| Id  | Source          | Destination   | Status     | %Remaining | CommState |
|-----|-----------------|---------------|------------|------------|-----------|
| 250 | fs_source_ckpt2 | fs_dest:laip2 | converting | 0          | N/A       |

[root@laip1 nasadmin]# fs\_copy -start fs\_scli\_ckpt2:if=rep fs\_dcli:cel=laip2:if=rep -fromfs fs\_scli\_ckpt1

-----output abridged-----

IP Copy remaining (%) 100..Done.

Done

Note: At this point, CIFS VDMs and Production File Systems are replicating between Source & Destination Celerras

## **DESTINATION SIDE DURING DIFFERENTIAL 2<sup>nd</sup> CHECKPOINT FS COPY OPERATION:**

[root@laip2 log]# fs\_copy -l

Local Source Filesystems

```

Id Source Destination Status %Remaining CommState
Local Destination Filesystems
Id Source Destination Status %Remaining CommState
113 fs_scli_ckpt2:1 fs_dcli N/A N/A down →What 2nd fs_copy looks like on dest. side while the differential copy is in progress. Good to document.
[root@laip2 log]# fs_replicate -l
Local Source Filesystems
Id Source FlowCtrl State Destination FlowCtrl State Network
Local Destination Filesystems
Id Source FlowCtrl State Dest. FlowCtrl State Network
102 root_fs_vdm_vdm_sou inactive active root_fs_vd inactive active alive
107 fs_source:laip1 inactive active fs_dest inactive active alive
112 fs_scli:laip1 inactive active fs_dcli inactive active alive →During 2nd fs_copy and conversion operation, the Replication Session status actually changes before the differential copy is completed, meaning that the differential 2nd copy actually converts file system to uxfs first before the copy completes.

```

Comment: At this point, replication is up and running with VDM rootfs and both PFS file systems replicating normally

## IP REPLICATION SESSIONS AND CONFIGURATION FILES FOR IP REP WITH CIFS VDMs:

### SOURCE:

```

[root@laip1 nasadmin]# fs_replicate -l
Local Source Filesystems
Id Source FlowCtrl State Destination FlowCtrl State Network
173 root_fs_vd inactive active root_fs_vdm_vdm_des inactive active alive
188 fs_source inactive active fs_dest:laip2 inactive active alive
192 fs_scli inactive active fs_dcli:laip2 inactive active alive
Local Destination Filesystems
Id Source FlowCtrl State Dest. FlowCtrl State Network
[root@laip1 nasadmin]# fs_replicate -i id=188
id          = 165
name        = fs_source
fs_state    = active
type        = replication
replicator_state = active
source_policy = NoPolicy
high_water_mark = 600
time_out    = 600
current_delta_set = 3
current_number_of_blocks = 22400
flow_control = inactive
total_savevol_space = 5120000 KBytes
savevol_space_available = 5111808 KBytes (Before Flow Control)
id          = 69
name        = fs_dest:laip2
type        = playback
playback_state = active
high_water_mark = 600
time_out    = 600
current_delta_set = 2
flow_control = inactive
total_savevol_space = 5120000 KBytes
savevol_space_available = 5111808 KBytes (Before Flow Control)
outstanding delta sets:
Delta Source_create_time Blocks
-----
2 2006-02-08 16:37:49 1
communication_state = alive
current_transfer_rate = ~ 832 Kbits/second
avg_transfer_rate   = ~ 832 Kbits/second
source_ip           = 192.1.6.204
source_port          = 59322

```

destination\_ip = 192.1.4.143

destination\_port = 8888

QOS\_bandwidth = 0 kbytes/sec

Note: All times are in GMT. Block size is 8 KBytes.

[root@laip1 nasadmin]# fs\_replicate -i id=192

id = 166

name = fs\_scli

fs\_state = active

type = replication

replicator\_state = active

source\_policy = NoPolicy

high\_water\_mark = 600

time\_out = 600

current\_delta\_set = 3

current\_number\_of\_blocks = 1

flow\_control = inactive

total\_savevol\_space = 1048576 KBytes

savevol\_space\_available = 917504 KBytes (Before Flow Control)

id = 70

name = fs\_dcli:laip2

type = playback

playback\_state = active

high\_water\_mark = 600

time\_out = 600

current\_delta\_set = 2

flow\_control = inactive

total\_savevol\_space = 1048576 KBytes

savevol\_space\_available = 917504 KBytes (Before Flow Control)

outstanding delta sets:

Delta Source\_create\_time Blocks

-----

2 2006-02-08 16:40:10 22398

communication\_state = alive

current\_transfer\_rate = ~ 37440 Kbytes/second

avg\_transfer\_rate = ~ 35008 Kbytes/second

source\_ip = 192.1.6.204

source\_port = 59323

destination\_ip = 192.1.4.143

destination\_port = 8888

QOS\_bandwidth = 0 kbytes/sec

Note: All times are in GMT. Block size is 8 KBytes.

#### SOURCE NETD:

rcproute add group=173\_APM000238010400000 nexthop=192.1.4.143 fromhop=192.1.6.204

rcproute add group=188\_APM000238010400000 nexthop=192.1.4.143 fromhop=192.1.6.204

rcproute add group=192\_APM000238010400000 nexthop=192.1.4.143 fromhop=192.1.6.204

iprepsvc start 173\_APM000238010400000 vol=113309402

iprepsvc start 188\_APM000238010400000 vol=113309418

iprepsvc start 192\_APM000238010400000 vol=113309426

#### SOURCE FILESYS ENTRIES:

164:root\_fs\_vdm\_vdm\_source::0:3::y:1:113309385:1:::0:167,68@2,172:23::0:0:

165:fs\_source::0:6::y:1:113309388:v1:::0:178,69@2,187:23::90:75000:

166:fs\_scli::0:4::y:1:113309391:v1:::0:180,70@2,191:23::0:0:

167:vpfs167::0:1::n:11:113309395::1139343866:164:2::23:168,175:

168:root\_fs\_vdm\_vdm\_source\_ckpt1::0:4::y:7::1:1139343868::0::167::0:0:

171:vpfs171::0:0::n:11:113309402:::0:23:172:

172:root\_fs\_vdm\_vdm\_source\_rvfs::0:1::n:9:113309403::1139345597:164:4:173:171::0:0:

173:root\_fs\_vdm\_vdm\_source\_rvfs\_ipfs1::0:1::y:5:0::1:1139345623:172:9::174::0:0:

174:173\_APM000238010400000::0:1::n:100:0:::0::173,74@2:

175:root\_fs\_vdm\_vdm\_source\_ckpt2::0:4::y:7::1:1139347436,3::0::167::0:0:

178:vpfs178::0:1::n:11:113309407::1139422402:165:2::23:179,194:

```

179:fs_source_ckpt1::0:4::y:7:::v1:1139422406::0::178::0:0:
180:vpfs180::0:1::n:11:113309411:::1139422450:166:2::23:181,195:
181:fs_scli_ckpt1::0:4::y:7:::v1:1139422455::0::180::0:0:
186:vpfs186::0:0::n:11:113309418::::0::23:187:
187:fs_source_rvfs::0:1::n:9:113309419:::1139432863:165:4:188:186::0:0:
188:fs_source_rvfs_ipfs1::0:1::y:5:0::1:1139432888:187:9::189::0:0:
189:188_APM000238010400000::0:1::n:100:0::::0::188,80@2:
190:vpfs190::0:0::n:11:113309426::::0::23:191:
191:fs_scli_rvfs::0:1::n:9:113309427:::1139432994:166:4:192:190::0:0:
192:fs_scli_rvfs_ipfs1::0:1::y:5:0::1:1139433030:191:9::193::0:0:
193:192_APM000238010400000::0:1::n:100:0::::0::192,83@2:
194:fs_source_ckpt2::0:4::y:7:::v1:1139433502,1::0::178::0:0:
195:fs_scli_ckpt2::0:4::y:7:::v1:1139433522::0::180::0:0:

```

**SOURCE MOUNTS:**

```

[root@laip1 log]# server_mount vdm_source
vdm_source :
fs_source on /fs_source udfs,perm,rw
fs_scli on /fs_scli udfs,perm,rw
fs_source_ckpt1 on /fs_source_ckpt1 ckpt,perm,ro
fs_scli_ckpt1 on /fs_scli_ckpt1 ckpt,perm,ro
fs_source_ckpt2 on /fs_source_ckpt2 ckpt,perm,ro
fs_scli_ckpt2 on /fs_scli_ckpt2 ckpt,perm,ro
[root@laip1 log]# server_mount server_2
server_2 :
root_fs_vdm_vdm_source on /root_vdm_3/etc udfs,perm,rw
fs_source on /root_vdm_3/fs_source udfs,perm,rw
fs_scli on /root_vdm_3/fs_scli udfs,perm,rw
root_fs_vdm_vdm_source_ckpt1 on /root_fs_vdm_vdm_source_ckpt1 ckpt,perm,ro
root_fs_vdm_vdm_source_rvfs_ipfs1 on /root_fs_vdm_vdm_source_rvfs_ipfs1 ipfs,temp,
p,ro,<unmounted>
root_fs_vdm_vdm_source_ckpt2 on /root_fs_vdm_vdm_source_ckpt2 ckpt,perm,ro
fs_source_ckpt1 on /root_vdm_3/fs_source_ckpt1 ckpt,perm,ro
fs_scli_ckpt1 on /root_vdm_3/fs_scli_ckpt1 ckpt,perm,ro
fs_source_rvfs_ipfs1 on /fs_source_rvfs_ipfs1 ipfs,temp,ro,<unmounted>
fs_source_ckpt2 on /root_vdm_3/fs_source_ckpt2 ckpt,perm,ro
fs_scli_rvfs_ipfs1 on /fs_scli_rvfs_ipfs1 ipfs,temp,ro,<unmounted>
fs_scli_ckpt2 on /root_vdm_3/fs_scli_ckpt2 ckpt,perm,ro

```

**DEST:**

```
[root@laip2 nasadmin]# fs_replicate -l
```

Local Source Filesystems

| Id | Source | FlowCtrl | State | Destination | FlowCtrl | State | Network |
|----|--------|----------|-------|-------------|----------|-------|---------|
|----|--------|----------|-------|-------------|----------|-------|---------|

Local Destination Filesystems

| Id | Source | FlowCtrl | State | Dest. | FlowCtrl | State | Network |
|----|--------|----------|-------|-------|----------|-------|---------|
|----|--------|----------|-------|-------|----------|-------|---------|

```
74 root_fs_vdm_vdm_sou inactive active root_fs_vd inactive active alive
```

```
80 fs_source:laip1 inactive active fs_dest inactive active alive
```

```
83 fs_scli:laip1 inactive active fs_dcli inactive active alive
```

```
[root@laip2 nasadmin]# fs_replicate -i id=80
```

```

id = 165
name = fs_source:laip1
fs_state = active
type = replication
replicator_state = active
source_policy = NoPolicy
high_water_mark = 600
time_out = 600
current_delta_set = 3
current_number_of_blocks = 22400
flow_control = inactive
total_savevol_space = 5120000 KBytes
savevol_space_available = 5111808 KBytes (Before Flow Control)

```

```
id          = 69
name        = fs_dest
type        = playback
playback_state = active
high_water_mark = 600
time_out    = 600
current_delta_set = 2
flow_control = inactive
total_savevol_space = 5120000 KBytes
savevol_space_available = 5111808 KBytes (Before Flow Control)
outstanding delta sets:
```

Delta Source\_create\_time Blocks

```
----- -----
2 2006-02-08 16:37:49 1
communication_state = alive
current_transfer_rate = ~ 832 Kbits/second
avg_transfer_rate   = ~ 832 Kbits/second
source_ip           = 192.1.6.204
source_port         = 59322
destination_ip      = 192.1.4.143
destination_port    = 8888
QOS_bandwidth       = 0 kbytes/sec
```

Note: All times are in GMT. Block size is 8 KBytes.

[root@laip2 nasadmin]# fs\_replicate -i id=83

```
id          = 166
name        = fs_scli:laip1
fs_state    = active
type        = replication
replicator_state = active
source_policy = NoPolicy
high_water_mark = 600
time_out    = 600
current_delta_set = 3
current_number_of_blocks = 1
flow_control = inactive
total_savevol_space = 1048576 KBytes
savevol_space_available = 917504 KBytes (Before Flow Control)
id          = 70
name        = fs_dcli
type        = playback
playback_state = active
high_water_mark = 600
time_out    = 600
current_delta_set = 2
flow_control = inactive
total_savevol_space = 1048576 KBytes
savevol_space_available = 917504 KBytes (Before Flow Control)
```

outstanding delta sets:

Delta Source\_create\_time Blocks

```
----- -----
2 2006-02-08 16:40:10 22398
communication_state = alive
current_transfer_rate = ~ 37440 Kbits/second
avg_transfer_rate   = ~ 35008 Kbits/second
source_ip           = 192.1.6.204
source_port         = 59323
destination_ip      = 192.1.4.143
destination_port    = 8888
QOS_bandwidth       = 0 kbytes/sec
```

Note: All times are in GMT. Block size is 8 KBytes.

**DEST NETD:**

```
volmcast create 73 173_APM000238010400000 185
volmcast create 79 188_APM000238010400000 191
volmcast create 82 192_APM000238010400000 197
```

**DEST FILESYS:**

```
68:root_fs_vdm_vdm_dest::0:10::y:1:174::1:1139347436:164@2:11:73:24::0:0:
69:fs_dest::0:12::y:1:177::1:1139433502:165@2:11:79:24::90:75000:
70:fs_dcli::0:10::y:1:180::1:1139433522:166@2:11:82:24::0:0:
72:vpfs72::0:0::n:11:185::::0::24:73:
73:root_fs_vdm_vdm_dest_pvfs::0:1::n:10:186::1139345614:68:5:74:72::0:0:
74:root_fs_vdm_vdm_dest_pvfs_ipfs1::0:1::y:5:0::1:1139345637:73:7::174@2::0:0:
78:vpfs78::0:0::n:11:191::::0::24:79:
79:fs_dest_pvfs::0:1::n:10:192::1139432878:69:5:80:78::0:0:
80:fs_dest_pvfs_ipfs1::0:1::y:5:0::1:1139432900:79:7::189@2::0:0:
81:vpfs81::0:0::n:11:197::::0::24:82:
82:fs_dcli_pvfs::0:1::n:10:198::1139433020:70:5:83:81::0:0:
83:fs_dcli_pvfs_ipfs1::0:1::y:5:0::1:1139433041:82:7::193@2::0:0:
```

**DESTINATION MOUNTS:**

```
[root@laip2 log]# server_mount vdm_dest
vdm_dest :
fs_dest on /fs_dest udfs,perm,ro
fs_dcli on /fs_dcli udfs,perm,ro
[root@laip2 log]# server_mount server_2
server_2 :
root_fs_vdm_vdm_dest_pvfs_ipfs1 on /root_fs_vdm_vdm_dest_pvfs_ipfs1 ipfs,temp,ro
,<unmounted>
root_fs_vdm_vdm_dest on /root_vdm_2/.etc udfs,perm,ro
fs_dest_pvfs_ipfs1 on /fs_dest_pvfs_ipfs1 ipfs,temp,ro,<unmounted>
fs_dest on /root_vdm_2/fs_dest udfs,perm,ro
fs_dcli_pvfs_ipfs1 on /fs_dcli_pvfs_ipfs1 ipfs,temp,ro,<unmounted>
fs_dcli on /root_vdm_2/fs_dcli udfs,perm,ro
```

**PERFORMING ADMINISTRATIVE IP REPLICATION FAILOVER WITH CIFS VDM SERVERS:**

**Note:** An administrative failover is a “planned” failover using the –reverse switch that incurs no data loss because the event is synchronous, unlike the -failover options. The following example assumes that normal IP Replication is flowing one-way from both VDM rootfs and data file systems, SOURCE to TARGET Celerra. The following procedure can be used to failover, then failback. Always run the –reverse from the active side running the replication service.

**SOURCE SIDE:****1. Unload Source VDM CIFS Server from Loaded to Mounted State (Stops the CIFS Service):**

```
$ nas_server -vdm source_vdm -setstate mounted
```

```
status :
defined = enabled
actual = mounted
```

**Note:** If VDM fails to change to ‘mounted’, try downing the interface first, then run –setstate mounted, then bring interface back up  
**TARGET SIDE:**

**2. Ensure that Target CIFS Server interface is up:**

```
# server_ifconfig server_2 -a
# server_ifconfig server_2 reps up
```

**SOURCE SIDE:****3. Failover rootfs VDM from active side running Replication (Source) to the Target side:**

```
[root@laip1 /]# /nas/sbin/rootfs_replicate -reverse root_fs_vdm_dest:cel=laip2:if=reps
root_fs_vdm_source:if=reps
```

**Note:** Both Source and Destination interfaces were online and needed when doing the –reverse method

**TARGET SIDE:****4. Load VDM CIFS Server on Target Side:**

```
$ nas_server -vdm dest_vdm -setstate loaded
```

```
status :
defined = enabled
```

actual = loaded, active

**Note:** Verify that DNS has been updated, and update manually if required (See server\_cifs output to see if dns was auto. updated).

**5. Manually start the CIFS service if required**—it may be that 5.4 requires manual start after the –setstate loaded command.

#### SOURCE SIDE:

#### **6. Failover Source file systems to Target side using Reverse Method Only:**

[root@laip1 /]# **fs\_replicate -reverse fs1\_d:cel=laip2:if=reps fs1:if=reps**

**Note:** A default –failover without options, or with the “now” option, would be used for disaster recovery failover, would incur data loss, and would be executed from the Target side. Remember to specify the Destination file system and Celerra in the first part of the command, and the Source file system at the end.

#### **7. Update Timeout & HWM Values on Source file systems:**

\$ **fs\_replicate -modify root\_fs\_vdm\_source -o to=300,hw=3**

\$ **fs\_replicate -modify fs1\_source -o to=300,hw=15**

#### **8. Update Timeout & HWM Values on Target file systems:**

\$ **fs\_replicate -modify root\_fs\_vdm\_target -o to=600,hw=6**

\$ **fs\_replicate -modify fs1\_target -o to=600,hw=30**

**Note:** The following method did not work and may be obsolete, even though still referenced in the man pages. After running the –failover –o sync, the replication session did not failover and went inactive N/A on the Source. Had to unmount the restart checkpoint on Destination side for the fs2 file system, then permanently unmount fs2, remount as ro, then convert to rawfs. On Source side, created new checkpoint, restarted replication for fs2, created another checkpoint, then copied the incremental over and this restored the Replication session.

[root@laip2 log]# **fs\_replicate -failover fs2:cel=laip1 fs2 -o sync**

#### **HOW TO DETERMINE IF REPLICATION SESSION IS WORKING:**

\$ **fs\_replicate -i id=77 -v 10**

```
id          = 25
name        = profiles_fs
fs_state    = active
type        = replication
replicator_state = active
source_policy = ReadOnly
high_water_mark = 0
time_out    = 0
current_delta_set = 46178
current_number_of_blocks = 0
flow_control = inactive
total_savevol_space = 10736640 KBytes
```

**savevol\_space\_available = 0 KBytes (Before Flow Control)**

communication\_state = alive

**current\_transfer\_rate = ~ 0 Kbits/second →Indicates that replication is broken and will need to be restarted**

**avg\_transfer\_rate = ~ 0 Kbits/second**

```
source_ip    = 164.56.49.82
source_port  = 65384
destination_ip = 129.124.103.82
destination_port = N/A
QOS_bandwidth = 3072 kbytes/sec
```

#### **NORMAL REPLICATION SESSION:**

\$ **fs\_replicate -l**

108 profiles\_f inactive active profiles\_fs:plcl1ea inactive active alive

\$ **fs\_replicate -i id=108**

```
id          = 25
name        = profiles_fs
fs_state    = active
type        = replication
replicator_state = active
source_policy = ReadOnly
high_water_mark = 30
```

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!

```

time_out      = 600
current_delta_set      = 212
current_number_of_blocks = 828
flow_control      = inactive
total_savevol_space    = 10736640 KBytes
savevol_space_available = 10616832 KBytes (Before Flow Control)
id              = 32
name            = profiles_fs:plc11ea0
type             = playback
playback_state    = active
high_water_mark   = 15
time_out         = 300
current_delta_set = 212
flow_control      = inactive
total_savevol_space = 10736640 KBytes
savevol_space_available = 10616832 KBytes (Before Flow Control)
outstanding delta sets: <None>
communication_state = alive
current_transfer_rate = ~ 277299 Kbits/second
avg_transfer_rate = ~ 205375 Kbits/second
source_ip        = 164.56.49.82
source_port       = 49154
destination_ip    = 129.124.103.82
destination_port   = 8888
QOS_bandwidth     = 3072 kbytes/sec

```

## **CELERRA IP REPLICATION TROUBLESHOOTING:**

### **# export NAS\_REPLICATE\_DEBUG=1**

- Remote IP replication transfers commands via HTTP
- Check nas\_log.al.remote to see commands issued from source CS to destination CS
- Check nas\_log.al.tran to see commands from CS to DART, and replies
- Use nas\_fs -l and nas\_fs -i on file systems setup for replication
- Check cmd\_log, cmd\_log.err, server\_log, & sys\_log to debug fs\_copy issues, on both Source & Destination sides
- Replication must be started while destination fs is rawfs
- Both Source and Destination file systems must be exactly same size and number of blocks
- Changes on source must not be greater than the ability of delta sets to be transmitted and committed to destination
- SavVol must be of adequate size to store changes prior to being written to destination
- Celerra supports 16 replication sessions per DM
- NAS 5.1/5.2 uses Control Station command to convert destination fs from rawfs to udfs
- NAS 5.3+ uses DART code to convert destination fs from rawfs to udfs
- I18N or ASCII must be set the same for both Source & Target sides for VDMs
- Do not rename VDMs after failover or suspend operations
- Source & Target Control Stations must have time sync within 10 minutes or HTTP communications will fail (nas\_cel -l)

**Note:** Use date command to correct time/date skew

→ Verify that passphrase is valid on Source & Target

### **# nas\_cel -info id=1** (run on Source)

passphrase=iprep

### **# nas\_cel -exec id=1 “nas\_cel -info id=1”** (run on Target)

Passphrase=iprep

→ Local CS failover on Source may interrupt replication traffic (NA status)

### **\$ fs\_replicate -l**

Local Source Filesystems

| Id   | Source     | FlowCtrl | State  | Destination | FlowCtrl | State      | Network |
|------|------------|----------|--------|-------------|----------|------------|---------|
| 1307 | uslist20_f | inactive | active | N/A         | N/A      | <b>N/A</b> | alive   |

## **THINGS TO DO AND CHECK TO IMPROVE IP REPLICATION/FSCOPY PERFORMANCE:**

I. Adjust TCP Window size for each Celerra replication session—default code value is 0, which is actually a windowsize of 128kb. The highest value allowed is 2MB, or 2097152.

**Note:** With NAS 5.6, the max RCP tcpwindow size allowed is 1MB, or 1048576.

**param RCP tcpwindow=400000 (Source)**  
**param RCP tcpwindow=300000 (Destination)**

**Note 1:** Network latency is one of the key factors behind poor IP Replication performance. The TCP Window defines the amount of data that can be outstanding on the network, as put on the wire by the sender, before stopping and waiting for an acknowledgement from the Receiving side. Used by Replication service and FS\_Copy. You would increase the RCP tcpwindow size for high latency networks with low retransmits. TCPWindow size is also considered the HWM setting for TCP.

**Note 2:** The RCP tcpwindow should be set differently on both Source and Destination data movers. On Source side, this parameter is used as the HWM for replication TCP connections, while on the Destination side, it is used as the TCP window size for the Replication tcp connections.

**GENERAL FORMULA TO DETERMINE RCP TCPWINDOW:**

TCP Window Size / Latency = Max Throughput

**GENERAL FORMULA TO DETERMINE WINDOW SIZE:**

**Example:** Need throughput of 10MB/sec on the network, with a Round Trip delay of 100ms

Slowest link speed \* Network Latency = TCP Window Size

**Solution:** Multiply 100ms (.1sec) \* 10MB (desired throughput) = 1MB tcpwindow size (1048576)

10MB/s \* .1 second (100ms) = 1MB Window Size

**param RCP tcpwindowlowat**

**Note 1:** This param represents the Low Water Mark for the TCPWindow as it relates to Replication & FS\_Copy Sessions. During a TCP transmission session, if the HWM is hit, then further TCP transmissions are halted until the TCP write buffer data is acknowledged by the Receiving side, down to the point where the LWM value is reached—only then will new transmissions occur. Default param value is 0, which represents 64kb. Max value is 1MB. Ordinarily, you would increase this param value in conjunction with an increase in the RCP TCPWindow size (aka, the HWM value) so as to reduce the ‘wait’ time required to acknowledge all the data held in the TCP buffer down to the LWM before resuming the sending of more data.

**Note 2:** For Cognac, the LWM will change to remain in proportion to the HWM (TCP Window) value, should this value be changed

**II. Set QOS on Each Replication Session to limit burstiness of DART**

**Note:** Originally, QOS value was wrong, was in kilobytes per second. Now fixed and represents kilobits per second.

**III. For Packet Loss networks, consider setting Slow Retransmission timer to 400ms vs. normal 1.5sec**

**#server\_netstat server\_2 -s**

**Note:** Look for TCP packets retransmitted—used if Client Reads are being impacted. If the rate of retrans is > than .01%, could be a network issue to look at. For Write issues, retransmission rates would need to be calculated on the Client. For NFS using UDP, check RPC for retransmissions.

**param tcp fastRTO=1**

**Note:** Reducing the time for the Retransmission timer reduces overall wait times when dropped packets are occurring

**IV. Reduce Network Bursting of Data Movers:**

**Note:** Default is to ACK for every 2 packets received. Setting to 8 helps to increase throughput between ACKs.

**param tcp maxburst=8**

**V. TCP Send Window:**

Param is not enabled by default. When set, will limit Data Mover TCP transmission window size to specified value, regardless of what Client advertises for tcp window. When set to 0, DM will match the max transmit windows to the max receive window of the client.

**param tcp sndcwnd=0**

**VERIFYING IPREPSENDER SERVICE AND RCPD SERVICE:**

**\$server\_config server\_2 -v "iprepsvc display 149\_APM000338015350000 vol=858"**

**\$server\_config server\_2 -v "rcproute display group=149\_APM000338015350000"**

**\$server\_config server\_2 -v "iprepsvc start 149\_APM000338015350000 vol=858"** (restarting IPRepSender svc)

**# .server\_config server\_2 -v "rcproute display"**

1165504352: RCPD: 4: rcp:: Group:95\_APM000238010400000

1165504352: RCPD: 4: rcp:: Group:91\_APM000238010400000

**IPREPSENDER SERVICE LOG MESSAGES:**

→sys\_log will note that IPRepSender is Shutdown

→Server Log will note VRPL: 3: Chunk Header contains Bad Magic:RBLK MAGIC ad:0 vol:858

→nas\_event –list –f VRPL will show IP Rep Svc Network events

**EXAMPLE OF REVERSING THE ROOTVDM FROM DEST TO SOURCE CELERRA:**

**[root@laip1 /]# /nas/sbin/rootfs\_replicate -reverse root\_fs\_vdm\_dest:cel=laip2;if=reps  
root\_fs\_vdm\_source;if=reps**

operation in progress (not interruptible)...

```

id      = 31
name    = root_fs_vdm_source
acl     = 0
in_use  = True
type    = uxfss
worm    = off
volume  = v105
pool    =
rw_servers=
ro_servers= server_2
rw_vdms =
ro_vdms =
ckpts   = root_fs_vdm_source_ckpt1,root_fs_vdm_source_ckpt2
ip_copies = root_fs_vdm_dest:laip2
stor_devs = APM00023801040-0010
disks   = d7
disk=d7 stor_dev=APM00023801040-0010 addr=c0t1l0      server=server_2
disk=d7 stor_dev=APM00023801040-0010 addr=c16t1l0      server=server_2
id      = 23
name    = root_fs_vdm_dest
acl     = 0
in_use  = True
type    = uxfss
worm    = off
volume  = v90
pool    = clar_r5_performance
member_of = root_avm_fs_group_3
rw_servers=
ro_servers= server_2
rw_vdms =
ro_vdms =
backup_of = root_fs_vdm_source:laip1 Tue Oct 17 10:41:24 EDT 2006
stor_devs = APM00030600872-0016
disks   = d8
disk=d8 stor_dev=APM00030600872-0016 addr=c16t2l9      server=server_2
disk=d8 stor_dev=APM00030600872-0016 addr=c0t2l9      server=server_2
operation in progress (not interruptible)...
done

```

## **EXAMPLE OF ACTIVE REPLICATION SESSIONS ON SOURCE & TARGET SIDES:**

### **SOURCE NY REPLICATION SESSIONS:**

**\$ nas\_cel -l**

| Id | name      | owner | mount_dev | channel        | net_path           | CMU |
|----|-----------|-------|-----------|----------------|--------------------|-----|
| 0  | usilnas4a | 0     |           | 141.202.1.81   | APM000433036020000 |     |
| 1  | uslinas4a | 500   |           | 130.200.10.174 | APM000433035980000 |     |

**\$ fs\_replicate -l**

Local Source Filesystems

| Id   | Source     | FlowCtrl | State  | Destination         | FlowCtrl | State  | Network                                                |
|------|------------|----------|--------|---------------------|----------|--------|--------------------------------------------------------|
| 1316 | usilst20_f | inactive | active | usilst20_ftpu_ftp:u | inactive | active | transferring →Replications sessions from NY to Chicago |
| 1308 | usilst20_c | inactive | active | usilst20_ca_cadocs: | inactive | active | alive                                                  |
| 1340 | usilst20_c | inactive | active | usilst20_ca_invoice | inactive | active | alive                                                  |

**1320 usilst20\_n inactive active usilst20\_niku\_files inactive active alive**

Local Destination Filesystems

| Id   | Source              | FlowCtrl | State  | Dest.      | FlowCtrl | State  | Network                                       |
|------|---------------------|----------|--------|------------|----------|--------|-----------------------------------------------|
| 1305 | uslist20_ftp_ftp:us | inactive | active | uslist20_f | inactive | active | alive →Replication session from Chicago to NY |

### **SOURCE NETD FILE:**

```
cifs add compname=USILCIFS9A domain=CA.COM netbios=usilcifs9a interface=usilcifs9a comment="Islandia NAS - USILCIFS9A"
```

```
rcproute add group=1308_APM000433036020000 nexthop=130.200.63.30 fromhop=141.202.63.30
```

```
rcproute add group=1316_APM000433036020000 nexthop=130.200.63.30 fromhop=141.202.63.30
```

**rcproute add group=1320\_APM000433036020000 nexthop=130.200.63.30 fromhop=141.202.63.30**

```
rcproute add group=1340_APM000433036020000 nexthop=130.200.63.30 fromhop=141.202.63.30
```

```
iprepsvc start 1308_APM000433036020000 vol=2484
```

```
iprepsvc start 1316_APM000433036020000 vol=2496
```

**iprepsvc start 1320\_APM000433036020000 vol=2502**

```
iprepsvc start 1340_APM000433036020000 vol=2532
```

```
volmcast create 1304_565_APM000433035980000 2478
```

## **SOURCE FILE SYSTEM REPLICATION DETAILS BY ID #:**

**\$ fs\_replicate -i id=1320 -v 5** (Lists last five deltaset)

```
id          = 509
name        = usilst20_niku_filestore
fs_state    = active
type        = replication
replicator_state = active
source_policy = NoPolicy
high_water_mark = 300
time_out     = 600
current_delta_set = 75
current_number_of_blocks = 1
flow_control = inactive
total_savevol_space = 1048576 KBytes
savevol_space_available = 917504 KBytes (Before Flow Control)
id          = 499
name        = usilst20_niku_filestore:uslinas4a
type        = playback
playback_state = active
high_water_mark = 600
time_out     = 600
current_delta_set = 74
flow_control = inactive
total_savevol_space = 1048576 KBytes
savevol_space_available = 917504 KBytes (Before Flow Control)
```

outstanding delta sets:

| Delta | Source_create_time | Blocks |
|-------|--------------------|--------|
|-------|--------------------|--------|

```
----- -----
74 09/26 12:37:00 1
communication_state = alive
current_transfer_rate = ~ 13312 Kbits/second
avg_transfer_rate = ~ 13516.8 Kbits/second
source_ip          = 141.202.63.30
source_port         = 64245
destination_ip     = 130.200.63.30
destination_port   = 8888
QOS_bandwidth      = 0 kbytes/sec
```

| Source | Destination |
|--------|-------------|
|--------|-------------|

| Delta | Create Time | Dur | Blocks | Playback Time | Dur | Blocks | DSinGroup |
|-------|-------------|-----|--------|---------------|-----|--------|-----------|
|-------|-------------|-----|--------|---------------|-----|--------|-----------|

```
-----|-----|-----|-----|-----|-----|-----|-----|
74 09/26 12:37:00 0 1
73 09/26 12:26:59 0 1 09/26 12:32:56 0 1 1
72 09/26 12:21:41 0 1 09/26 12:22:55 0 1 1
71 09/26 12:11:41 0 1 09/26 12:12:55 0 1 1
70 09/26 12:01:41 0 1 09/26 12:02:55 0 1 1
```

## **SOURCE FILE SYSTEM ENTRIES FILESYSTEM FILE:**

**# cat /nas/volume/filesys |grep -i niku**

```
1320:usilst20_niku_filestore_rvfs_ipfs1:0:y:5:0::1:1127693508:1319:9::1321:
```

```
1319:usilst20_niku_filestore_rvfs:0:n:9:2503:::1127693499:509:4:1320:1318:
```

```
509:usilst20_niku_filestore:0:y:1:1266:1:::0:1319,499@1:23:
```

# **nas\_fs -l |grep -i niku** →This is original source file system that is being replicated to Chicago

509 y 1 0 1266 usilst20\_niku\_files 1

**\$ nas\_fs -s id=509**

total = 8066 avail = 7666 used = 400 ( 4% ) (sizes in MB) ( blockcount = 16777216 )

volume: total = 8192 (sizes in MB) ( blockcount = 16777216 )

### **TARGET CHICAGO REPLICATION SESSIONS:**

**\$ nas\_cel -l**

|    |           |       |           |                |                    |     |
|----|-----------|-------|-----------|----------------|--------------------|-----|
| id | name      | owner | mount_dev | channel        | net_path           | CMU |
| 0  | uslinas4a | 0     |           | 130.200.10.174 | APM000433035980000 |     |
| 1  | usilnas4a | 500   |           | 141.202.1.81   | APM000433036020000 |     |

**\$ fs\_replicate -l**

Local Source Filesystems

|     |            |          |        |                     |          |        |         |
|-----|------------|----------|--------|---------------------|----------|--------|---------|
| Id  | Source     | FlowCtrl | State  | Destination         | FlowCtrl | State  | Network |
| 565 | uslist20_f | inactive | active | uslist20_ftp_ftp:us | inactive | active | alive   |

→Replication from Chicago to NY

Local Destination Filesystems

|     |                      |          |        |            |          |        |              |
|-----|----------------------|----------|--------|------------|----------|--------|--------------|
| Id  | Source               | FlowCtrl | State  | Dest.      | FlowCtrl | State  | Network      |
| 575 | usilst20_ftpu_ftpu:u | inactive | active | usilst20_f | inactive | active | transferring |

→Replications from NY to Chicago

599 root\_fs\_vdm\_server\_inactive active root\_fs\_vd inactive active alive

593 usilst20\_ca\_invoice inactive active usilst20\_c inactive active alive

569 usilst20\_ca\_cadocs: inactive active usilst20\_c inactive active alive

578 usilst20\_niku\_files inactive active usilst20\_n inactive active alive

### **TARGET SIDE NETD FILE:**

```
cifs add compname=USLICIFS9A domain=CA.COM netbios=uslicifs9a interface=uslicifs9a comment="Lisle NAS - USLICIFS9A"
rcproute add group=565_APM000433035980000 nexthop=141.202.63.30 fromhop=130.200.63.30
iprepsvc start 565_APM000433035980000 vol=1475
volmcast create 568 1308_APM000433036020000 1481
volmcast create 574 1316_APM000433036020000 1493
volmcast create 577 1320_APM000433036020000 1499
volmcast create 592 1340_APM000433036020000 1529
```

### **TARGET SIDE REPLICATION DETAILS BY FILE SYSTEM #:**

**\$ fs\_replicate -i id=578 -v 5**

|                          |                                       |
|--------------------------|---------------------------------------|
| id                       | = 509                                 |
| name                     | = usilst20_niku_filestore:usilnas4a   |
| fs_state                 | = active                              |
| type                     | = replication                         |
| replicator_state         | = active                              |
| source_policy            | = NoPolicy                            |
| high_water_mark          | = 300                                 |
| time_out                 | = 600                                 |
| current_delta_set        | = 77                                  |
| current_number_of_blocks | = 1                                   |
| flow_control             | = inactive                            |
| total_savevol_space      | = 1048576 KBytes                      |
| savevol_space_available  | = 917504 KBytes (Before Flow Control) |

|                         |                                       |
|-------------------------|---------------------------------------|
| id                      | = 499                                 |
| name                    | = usilst20_niku_filestore             |
| type                    | = playback                            |
| playback_state          | = active                              |
| high_water_mark         | = 600                                 |
| time_out                | = 600                                 |
| current_delta_set       | = 76                                  |
| flow_control            | = inactive                            |
| total_savevol_space     | = 1048576 KBytes                      |
| savevol_space_available | = 917504 KBytes (Before Flow Control) |

outstanding delta sets:

Delta Source\_create\_time Blocks

---

|                     |                |   |
|---------------------|----------------|---|
| 76                  | 09/26 12:57:00 | 2 |
| communication_state | = alive        |   |

```

current_transfer_rate = ~ 14336 Kbits/second
avg_transfer_rate    = ~ 13524.8 Kbits/second
source_ip            = 141.202.63.30
source_port          = 64245
destination_ip       = 130.200.63.30
destination_port     = 8888
QOS_bandwidth        = 0 kbytes/sec
|   Source           |   Destination
Delta|Create Time   Dur  Blocks|Playback Time Dur  Blocks DSinGroup
-----+-----+-----+-----+-----+-----+-----+-----+
75  09/26 12:47:00 0  1  09/26 12:52:56 0  1  1
74  09/26 12:37:00 0  1  09/26 12:42:56 0  1  1
73  09/26 12:26:59 0  1  09/26 12:32:56 0  1  1
72  09/26 12:21:41 0  1  09/26 12:22:55 0  1  1
71  09/26 12:11:41 0  1  09/26 12:12:55 0  1  1

```

### **TARGET SIDE FILE SYSTEM DETAILS:**

**\$ cat /nas/volume/filesys|grep -i niku**

```

577:usilst20_niku_filestore_pvfs:0:n:10:1500:::1127693501:499:5:578:576:
578:usilst20_niku_filestore_pvfs_ipfs1:0:y:5:0::1:1127693509:577:7::1321@1:
499:usilst20_niku_filestore:0:y:1:1337::1:1127693522:509@1:11:577:23:

```

**\$ nas\_fs -l |grep -i niku**

```

499  y  1  0  1337  usilst20_niku_files 1

```

**\$ nas\_fs -s id=499**

```

total = 8066 avail = 7666 used = 400 ( 4% ) (sizes in MB) ( blockcount = 16777216 )
volume: total = 8192 (sizes in MB) ( blockcount = 16777216 )

```

### **SOURCE SIDE:**

→Flow Control ACTIVE indicates that source side cannot transfer delta sets to target side (i.e., networking issues, etc.)

**Note:** Important point here is that source SavVol cannot accept more delta sets while in this state

→If State is INACTIVE, this means that the replication session is broken for that fs

### **DESTINATION SIDE:**

→An INACTIVE state could indicate a problem, or just that an fs\_copy is in progress and the target fs is still rawfs and not converted  
→Transferring means that the IP Rep Sender is sending data across the link

### **IP REPLICATION PROCESSES TO KNOW ABOUT:**

IP Rep Sender Service →Moves delta sets from Source to Target sides

IP Replication Service →Triggered by HWM/TO values

### **FS COPY COMMAND:**

**\$ fs\_copy -I**

Local Source Filesystems

| Id   | Source          | Destination     | Status  | %Remaining | CommState    |
|------|-----------------|-----------------|---------|------------|--------------|
| 1061 | usilst20_ftpu_f | usilst20_ftpu_f | started | 95         | transferring |

**Note:** Source pfs is being successfully copied to destination file system, which in this case, is named the same. During an fs\_copy, data is organized into tables of contiguous blocks of data (anywhere from 8MB up to 25GB). After each table is transferred, fs\_copy pauses to wait for remote Control Station to acknowledge that data has been received and written to the destination SavVol, before starting to transfer the next table.

**Caution:** Do not extend PFS while fs\_copy is running, NAS 5.3 and below

### **REMOVING INVALID FILE SYSTEM REFERENCES FROM SOURCE & DESTINATION:**

#### **Destination Backup Pointer:**

**# /nas/bin/nas\_cmd @fs\_ip -C <dst\_fs\_name>**

#### **Source Backup Pointer:**

**# /nas/bin/nas\_cmd @fs\_ip -U <src\_fs\_name> -f <dst\_fs\_name>:CMU=<CMU\_id>**

**Note:** CMU\_id is obtained from nas\_cel –list—last field in output

### **TO VALIDATE FS COPY % REMAINING, DO THE FOLLOWING:**

- Locate current fs\_copy session running in the source side netd file:

# cat netd

**ipfs copy start fsid=1060 rcpg=1061\_APM000433036020000 token=1124007449 address=0 qos=40960**

2. Run following command to obtain VolMCastSender information:

```
$ .server_config server_2 -v "ipfsycop display fsid=1060 rcpg=1061_APM000433036020000 token=1124007449"
```

```
1124031588: VMCAST: 5: ipfsycop fsid=1060 rcpg=1061_APM000433036020000 token=1124007449 address=0
1124031588: VMCAST: 5: DisplayIPFSCopy(): Token:1124007449 Group:1061_APM000433036020000 FSID:1060
1124031588: VMCAST: 4: ipfsycop:display() Token:1124007449 Group:1061_APM000433036020000 FSID:1060 fromFSID:0
1124031588: VMCAST: 4: VolMCast::display() Name:1060_1061_APM000433036020000 VolName:Sh20441060
NodeName:1061_APM000433036020000 refcount:1
1124031588: VMCAST: 4: VolMCastSender::display() NextBlock:102882528 TotalBlock:3758096384
```

remote\_communication:Transferring **NextBeingProcessed:103864896**

```
1124031588: RCPD: 4: rcp:: Group:1061_APM000433036020000
```

```
1124031588: RCPD: 4: rcp:: Target_num:0 Target_ip:130.200.63.30 Local_ip:141.202.27.118 Local_port:56566
```

```
1124031588: VMCAST: 4: DisplayIPFSCopy(): State:0 ReasonCode:0
```

3. Divide the NextBeingProcessed block by TotalBlock to obtain a percentage of what has actually been copied over:

**103864896 / 3758096384 = .027 or 2.7% completed**

### **EXTENT TABLES:**

FSCopy organizes data into ranges of data to be copied, called “extents”. The extent tables can be as large as 25GB in size, or up to 1024 entries—each entry in the table represents at least one 8kb file system block. The “NextBlock” value is the extent starting point for the current fscopy session, and the “NextBeingProcessed” block represents the actual block in progress. Extents are how data mover keeps track of fscopy status, say if a reboot occurs—the fscopy operation would pickup at last known NextBlock value and begin recopying from that point. One of the problems with trying to use fs\_copy -i to determine fs\_copy progress is that it only updates for each “extent” copied, and could therefore be 25GB between updates. Once a given Extent Table is copied and acknowledged on the destination side, the next Extent can begin to be copied.

**\$ fs\_replicate -l**

Local Destination Filesystems

| Id  | Source              | FlowCtrl | State  | Dest.      | FlowCtrl | State  | Network |
|-----|---------------------|----------|--------|------------|----------|--------|---------|
| 427 | usilst20_ftpu_ftp:u | inactive | active | usilst20_f | inactive | active | alive   |

**Note:** Above shows state of replication as up and running normally from source side

**# fs\_replicate -i id=25**

```
id          = 25
name        = usilst20_ftpu_ftp
fs_state    = active
type        = replication
replicator_state = active
source_policy = NoPolicy
high_water_mark = 600
time_out     = 600
current_delta_set = 66
current_number_of_blocks = 3063
flow_control = inactive
total_savevol_space = 187904000 KBytes
savevol_space_available = 187826176 KBytes (Before Flow Control)
id          = 289
name        = usilst20_ftpu_ftp:uslinas4a
type        = playback
playback_state = active
high_water_mark = 600
time_out     = 600
current_delta_set = 66
flow_control = inactive
total_savevol_space = 187904000 KBytes
savevol_space_available = 187826176 KBytes (Before Flow Control)
outstanding delta sets: <None>
communication_state = alive
current_transfer_rate = ~ 382188 Kbits/second
avg_transfer_rate    = ~ 298668 Kbits/second
source_ip            = 141.202.63.30
source_port          = 56525
destination_ip       = 130.200.63.30
```

destination\_port = 8888

QOS\_bandwidth = 40960 kbits/sec

**Note:** Use fs\_replicate -i on both sides to see more information of state of current replication session per fsid, especially to check playback activity, etc.

## **WHAT OCCURS AFTER FS COPY COMPLETES:**

- Copy completes, nas\_event collector finds entry from sys\_log to post
- the ipfs1 file system services are removed from both Source & Destination sides
- the destination fcpg file entry is removed from filesystem file
- the rcproute info is removed from both boot.cfg files

## **CLEANING UP FAILED FS COPY SESSION:**

1. Set debug mode on both Source & Target Sides:

**#export NAS\_REPLICATE\_DEBUG=1**

2. Use server\_mount to locate ipfs-mounted file systems and server\_umount to unmount from Source first, then Target

3. Examine /nas/volume/filesys file and use following commands to delete ipfs file systems and fs\_groups:

**#fs\_group -d id=34** (deletes fcpg group associated with the fs\_copy operation that failed)

**#nas\_fs -d id=33** (deletes IP replication service for checkpoint wgc3\_ckpt1\_ipfs1)

**#nas\_fs -d id=35** (deletes IP replication service for wgc3\_sfs\_ipfs1)

**Note:** If deletion fails, may need to manually edit /nas/volume/filesys file to change respective entries for above examples to “n” in the line for not-in-use, and then remove the reference count located at the 2<sup>nd</sup> to last item in the row: :29:7

4. Also check netd & boot.cfg to see that rcp entries have been removed

5. Make sure that IP Replication Debug Logging is turned off prior to commencing fs\_copy operations:

**# .server\_config server\_2 -v "logsys get severity VMCAST"**

1143253431: LIB: 4: Server log severity for facility VMCAST is 4

**Note:** If VMCAST logging is set to 4, then the logging level is normal and debug logging is not enabled

**# .server\_config server\_2 -v "logsys get severity RCPD"**

1143254192: LIB: 4: Server log severity for facility RCPD is 4

**\$ .server\_config server\_2 "logsys set severity VMCAST=LOG\_PRINTF"** (sets debug logging back to level 4)

**\$ .server\_config server\_2 "logsys set severity RCPD=LOG\_PRINTF"**

**Note:** Run commands on both Source and Destination sides

## **TURNING UP IP REPLICATION DEBUG LOGGING:**

**\$ .server\_config server\_x -v "logsys set severity VRPL=LOG\_DBG3"**

**\$ .server\_config server\_x -v "logsys set severity VMCAST=LOG\_DEBUG"**

**\$ .server\_config server\_x -v "logsys set severity RCPD=LOG\_DEBUG"**

**\$ .server\_config server\_x -v "logsys set severity VMCAST=LOG\_PRINTF"** (Set this prior to running fs\_copy)

**\$ .server\_config server\_x -v "logsys set severity VMCAST=LOG\_ERR"**

**Note:** The above command turns off IP Replication Debug logging. If VMCAST debug is invoked with 5.2 code, can actually prevent fs\_copy operation from completing correctly and may need to restart fs\_copy.

## **BEST METHOD FOR CHECKING LOGGING LEVELS:**

**# .server\_config server\_2 -v "logsys get severity RCPD"** (Logging level 4 is normal, no debug)

1143254192: LIB: 4: Server log severity for facility RCPD is 4

## **DEBUGGING FS COPY OPERATIONS:**

**\$ .server\_config server\_x -v "logsys set severity RCPD=LOG\_DEBUG"**

**\$ .server\_config server\_x -v "logsys set severity RCPD=LOG\_PRINTF"**

## **HOW TO DETERMINE IF IP REPLICATION (FS COPY) IS RUNNING OR HUNG:**

→Check sys\_log to see if fs\_copy completion message is posted (once posted, triggers fs\_copy abort & cleanup)

→verify that DM is listening to port 8888

→Run following commands at intervals and then calculate to see if progress is being made.

**Note:** “ipfscopy” is run only on the SOURCE side and the session is found in the netd file

**\$ .server\_config server\_2 -v "ipfscopy display fsid=96 rcpg=97\_APM000306008700000**

**token=1058286129"**

1058291610: VMCAST: 5: ipfscopy fsid=96 rcpg=97\_APM000306008700000 token=1058286129 address=0

1058291610: VMCAST: 5: DisplayIPFSCopy(): Token:1058286129 Group:97\_APM000306008700000 FSID:96

1058291610: VMCAST: 4: ipfscopy:display() Token:1058286129 Group:97\_APM000306008700000 FSID:96

1058291610: VMCAST: 3: VolMCast::display() Name:96\_97\_APM000306008700000 VolName:Sh13496

NodeName:97\_APM000306008700000 refcount:1

1058291610: VMCAST: 4: VolMCastSender::**display()** **NextBlock:5376** TotalBlock:209715200

state:VMCAST\_SENDER\_BROADCAST\_IN\_PROGRESS1058291610: VMCAST: 4: DisplayIPFSCopy(): State:01058291610:

ADMIN: 4: Command succeeded: ipfscopy display fsid=96 rcpg=97\_APM000306008700000 token=1058286129

**Note:** Find the appropriate “rcpg=“ line in the NETD file and run command a few times to see if “NextBlock:5376” has changed—if it has, then IP Replication is in progress. If not, then it might be hung.

### **PREVENTING RESTART OF FS COPY:**

Comment out ipfscopy line in the netd file to prevent fs\_copy sessions from being restarted (might do this in cases of rolling panics)

### **HOW TO BREAKOUT OF A HUNG FS COPY SESSION:**

# .server\_config server\_x "ipfscopy stop fsid=x rcpg=x token=x address=x" (see netd file)

# .server\_config server\_x "ipfscopy start fsid=x rcpg=x token=x address=x noconvert=1" (netd)

### **CHANGING HTTP TIMEOUT VALUE BETWEEN SOURCE & TARGET CONTROL STATIONS:**

**Note:** fs\_copy or fs\_replicate operations can timeout, causing them to hang. For a temporary workaround, increase HTTP timeout value for the Source Control Station from 300 seconds to 600 seconds in the /nas/http/conf/httpd.conf file--issue has been addressed with NAS 5.3.22.0 and 5.4.19.0 by increasing the default HTTP communication timeout value to 600 seconds (10 minutes):

1. #vi /nas/http/conf/httpd.conf and find the line associated with "Timeout 300"--should be on or around line 106:

104 # Timeout: The number of seconds before receives and sends time out.

105 #

106 Timeout 300 -->Change this value to 600

2. Identify and Kill the master Apache httpd process:

# ps fax |egrep httpd

```
18504 pts/3 S 0:00      \_ egrep httpd
17186 ? S 0:00 /nas/sbin/httpd -D HAVE_PERL -D HAVE_SSL -f /nas/http
17187 ? S 0:00 \_ /nas/sbin/httpd -D HAVE_PERL -D HAVE_SSL -f /nas/
17188 ? S 0:00 \_ /nas/sbin/httpd -D HAVE_PERL -D HAVE_SSL -f /nas/
```

# kill 17186

**Note:** Do not use kill -9 to stop the Apache httpd daemon

# ps fax |egrep httpd

```
18824 pts/3 S 0:00      \_ egrep httpd
```

3. Restart Apache httpd process using following command:

# /nas/sbin/httpd -D HAVE\_PERL -D HAVE\_SSL -f /nas/http/conf/httpd.conf

4. Verify Apache process:

# ps fax |egrep httpd

```
19571 pts/3 S 0:00      \_ egrep httpd
19519 ? S 0:00 /nas/sbin/httpd -D HAVE_PERL -D HAVE_SSL -f /nas/http
19521 ? S 0:00 \_ /nas/sbin/httpd -D HAVE_PERL -D HAVE_SSL -f /nas/
```

### **REPLICATOR FLOW CONTROL:**

Replication Flow Control will automatically kick in under the following situations;

1.) Delta Sets on destination side are about to be overwritten by newer delta sets. Flow Control kicks in. Situation can occur if the Playback rate on the Remote side is too slow

2.) If Delta Sets cannot be transferred from Primary SavVol to Secondary side SavVol, flow control kicks in. This can happen if there is a loss of network connectivity between Source and Destination sides.

3.) SavVol becomes full on Primary side, flow control kicks in. Could happen if Primary Write rate is too fast.

### **LOCAL REPLICATION vs. REMOTE REPLICATION:**

→Local Replication is done using the same data mover on the local Celerra (Loopback Replication) or different data movers on the local Celerra—No remote Celerras are involved. Replication synchronization is started between PFS & SFS via the fs\_copy command. Subsequent PFS block changes after the initial copy are saved by the Replication Service to delta sets which are then copied to the shared SavVol used by PFS & SFS. Replication Playback Service plays back completed delta sets to the SFS.

→Remote Replication is setup between Local and Remote Celerras. Replication synchronization is started between PFS & SFS via the fs\_copy command. Subsequent PFS block changes after the initial copy are saved by the Replication Service to delta sets which are then copied to the primary SavVol. The Remote Replication Service transfers completed delta sets to the remote SavVol, where the Replication Playback Service plays back each delta set to the SFS.

### **FILE SYSTEM EXTENSIONS WHILE REPLICATION IS RUNNING:**

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
The default behavior is that when the Source file system is extended, replication will be suspended, the destination file system will extend first, followed by the source fs, with the replication session restarting

### **EXTEND SUCCEEDS ON DESTINATION BUT FAILS ON SOURCE:**

If the Destination side extended, but Source failed, try the following src\_only option to extend:

**# nas\_fs -xtend Groupware\_A size=102400M pool=clar\_r5\_performance -option src\_only,slice=y**

### **DESTINATION FS WILL NOT EXTEND DUE TO POOL SPACE:**

Try manually extending sfs using –Force flag and different storage pool, then extend pfs using source –only option

### **FILE SYSTEM EXTENSION WHEN USING SLICED VOLUMES vs. NON-SLICE VOLUMES:**

If the original file system was created without using the Slice=y option, then automatic file system extension using AVM will fail with the following error, and must be extended manually from the CLI. AVM uses the slice=yes option by default, but since the original file system was created without slice=y, the manual CLI command must be used instead:

CFS:3:101 fs auto extension failed: /nas/sbin/rootnas\_fs -x hadata3-1 size=10000M -o slice=n Error 3136: Should be sliced as hadata3-1 is a member of a replication session, fsname=hadata3-1 (Slot3:1159398824:)

**Resolution:** Manually extend the file system and specify the slice=y

**# nas\_fs -x hadata3-1 size=10000M -o slice=y**

### **REPLICATION SUSPEND MODE:**

**Note:** NAS 5.3 introduces ability to suspend a replication session and then use the –restart operation to restart replication later, and is meant as a way to temporarily stop replication and leave it in a state where it can be restarted. Also, you cannot suspend replication when upgrading from 5.3 to 5.4 and restart replication without converting remote file systems to rawfs and renaming the appropriate remote service to root\_suspend\_ckpt\_25\_1 (where 25=remote filesys id & 1=remote cel id)

**1. Set refresh policy on Source and then Destination file system** to prevent an out-of-synchronization issue from occurring while replication is in a suspended state:

**# fs\_replicate -refresh pfs -o to=0,hwm=0**

**# fs\_replicate -refresh sfs -o to=0,hwm=0**

**2. Suspend replication on file system:**

**# fs\_replicate -suspend <srcfs> <dstfs>[:cel=<cel\_name>]**

### **SUSPEND OPERATION EXAMPLE:**

**# fs\_replicate -suspend fs\_scli fs\_dcli:cel=laip2**

operation in progress (not interruptible)...

```
id          = 166
name        = fs_scli
fs_state    = active
type        = replication
replicator_state = active
source_policy = NoPolicy
```

Generating new checkpoint...

operation in progress (not interruptible)...id = 166

```
name      = fs_scli
acl       = 0
in_use   = True
type     = uxfss
```

```
id      = 283
name   = root_suspend_ckpt_98_2
```

#### **cmd log:**

```
2006-04-04 07:45:10.243 db:0:17902:S: /nas/sbin/rootfs_replicate -refresh fs_source -o to=0,hwm=0
2006-04-04 07:45:10.431 db:0:17902:E: /nas/sbin/rootfs_replicate -refresh fs_source -o to=0,hwm=0
2006-04-04 07:45:31.261 db:0:18234:S: /nas/sbin/rootfs_ckpt fs_source -name root_suspend_ckpt_97_2 -Create -option
automount=yes
2006-04-04 07:45:34.403 db:0:18234:E: /nas/sbin/rootfs_ckpt fs_source -name root_suspend_ckpt_97_2 -Create -option
automount=yes
2006-04-04 07:45:42.654 db:0:18474:S: /nas/sbin/rootfs_replicate -refresh fs_source -o to=0,hwm=0
2006-04-04 07:45:42.985 db:0:18474:E: /nas/sbin/rootfs_replicate -refresh fs_source -o to=0,hwm=0
2006-04-04 07:46:24.489 db:0:18941:S: /nas/sbin/rootfs_replicate -abort fs_source
2006-04-04 07:46:36.347 db:0:18941:E: /nas/sbin/rootfs_replicate -abort fs_source
2006-04-04 07:46:44.569 db:0:19185:S: rootfs_ip -U fs_source -f fs_dest:CMU=APM000306008720000
```

2006-04-04 07:46:45.829 db:0:19185:E: rootfs\_ip -U fs\_source -f fs\_dest:CMU=APM000306008720000

**Server Log:**

```
2006-04-04 07:45:52: SVFS: 4: pause() requested on fsid:165
2006-04-04 07:45:52: SVFS: 4: D282113309388_0: createBlockMap PBM root=0 keys=0 h=0 nc=0
2006-04-04 07:45:52: VRPL: 4: 113309480: Allocating chunk:2 Add:525312 Chunks:78
2006-04-04 07:45:52: SVFS: 4: <SNAPSURE DELTA_NUMBER="490"/>
2006-04-04 07:45:52: SVFS: 4: resume() requested on fsid:165
2006-04-04 07:45:52: CFS: 4: Resuming fs 165
2006-04-04 07:45:52: ADMIN: 4: SnapSure command build ID 282 SavVol 113309480 FsVol 113309388 HWM 90 succeeded
2006-04-04 07:46:03: CFS: 4: Resuming fs 165
2006-04-04 07:46:03: VRPL: 4: Replication::Valid v:ClSn113309388 Delta:490 ad:5768192 g:490 nc:1
2006-04-04 07:46:03: ADMIN: 4: Command succeeded: replica start srcvol=113309388 savevol=113309491 cfgvol=113309492
to=0 hwm=0
2006-04-04 07:46:47: VRPL: 4: 5: IPRepSender stop succeeded on gpe:275_APM000238010400000 vol:113309491.
2006-04-04 07:46:47: ADMIN: 4: Command succeeded: iprepvc stop 275_APM000238010400000 vol=113309491
2006-04-04 07:46:47: ADMIN: 4: Command succeeded: rcproute remove group=275_APM000238010400000
2006-04-04 07:46:55: VRPL: 4: StopStart in progress.
2006-04-04 07:46:55: VRPL: 4: StopFinish in progress.
2006-04-04 07:46:55: VRPL: 3: FS will be paused.
2006-04-04 07:46:55: VRPL: 3: FS is paused.
2006-04-04 07:46:55: VRPL: 4: RedoLog::~RedoLog() unregisterSelf from panicHandler for:113309388
2006-04-04 07:46:55: VRPL: 3: RedoLog::~RedoLog() unregisterSelf from panicHandler for:113309388 failed
2006-04-04 07:46:55: VRPL: 3: Prepare to delete LogFilterVol:LF113309388 with no io in progress!!!
2006-04-04 07:46:55: VRPL: 3: ~LogFilter Destructor for LF113309388
2006-04-04 07:46:55: VRPL: 3: FS is paused.
2006-04-04 07:46:55: CFS: 4: Resuming fs 165
```

**Note:** Suspend creates a checkpoint of the source, copies the delta changes over using refresh to the destination, then aborts replication session

**3. Restart replication:**

**# fs\_replicate -restart pfs:if=pfsRep sfs:cel=secondaryCS:if=sfsRep –option hwm=600,dhwm=300,to=600,dto=300**

**Note:** The –restart switch is the only viable method for resuming replication from a –suspend operation. The restart will create a new checkpoint of the Source file system and then do a differential fs\_copy to the destination file system between the latest checkpoint and the suspend checkpoint that was taken at the time that the file system was suspended. Make sure that Source file system policies are reset to a level greater than the destination so as to prevent an out-of-sync replication, which could require a complete breakdown and setup of replication between the Source and Destination file systems.

**USING THE SUSPEND COMMAND FOR REPLICATION**

- Moving either PFS or SFS from one mountpoint or data mover to another
- Changing IP addresses on data mover interface on Source or Target system
- Changing size of replication SavVol
- Upgrading sites where IP Replication is established, by suspending Production File System replication

**TECH SUPPORT GUIDANCE ON HANDLING REPLICATION SITES FOR NAS UPGRADES:**

1. Take manual checkpoint of each replication file system on the source prior to Upgrade
2. Perform NAS upgrade
3. If replication does not survive upgrade, perform abort and manual restart, which the manual checkpoint will allow you to do

**OVERVIEW OF FS REPLICATION OPERATIONS:**

**#fs\_replicate -l  
#fs\_replicate -i fs01\_1  
#fs\_replicate -i fs01\_1 -v 15**

**FUNDAMENTAL IP REPLICATION STEPS FOR FILE SYSTEMS:**

1. Create rawfs on Destination side of equal size to that of Source side, and mount RO
2. Conduct initial Checkpoint of production fs on Source side using fs\_ckpt fs01 –Create
3. Startup initial fs\_copy operation and wait until it completes--fs\_copy –start convert=no monitor=off

**Note:** Observe status of fs\_copy using -list

4. Startup Replication process using fs\_replicate -start

5. Create 2<sup>nd</sup> checkpoint of production file system using `fs_ckpt fs01 -Create`
  6. Copy over the incremental checkpoint changes using `fs_copy -start -o monitor=off` and allow the file system to be converted from rawfs to uxfss
  7. FS replication state should now be established and operational between Source and Destination

#### **VERIFYING REPLICATION BETWEEN SOURCE AND DESTINATION:**

```
# fs_replicate -l | -i
```

## Local Source Filesystems

Id Source FlowCtrl State Destination FlowCtrl State Network  
 35 fs16fsfs inactive active fs16:RDF1\_52 inactive active alive → Example shows normal IP Replication established  
 39 root fs vd inactive active root fs vdm 1:RDE1 inactive active alive

**Note:** FlowCtrl inactive is normal and is only used if special switches for flow control are invoked. The key with this output is if both sides of the IPP application service are showing a State of 'Active' and Network of 'Alive'.

```
# fs replicate -l -dst
```

```
# fs replicate -l -src
```

## WHAT IF REPLICATION SESSIONS BECOME INACTIVE?

→ Refer to primus solutions emc116166, emc180572, & emc173039, for troubleshooting and restarting Replication V1 sessions that have gone ‘inactive’.

#### **VERIFYING PLAYBACK SERVICE & DELTA SETS:**

The Playback Service reads delta set changes to the destination SavVol and updates destination file system with changes. Use the following command to verify the existence of outstanding delta sets and playback:

```
# fs replicate -i fs01 data -v 15
```

```

# ls -Replicate -fso1_data -v 13
id          = 41
name        = fs01_data
fs_state    = active
type        = replication
replicator_state = active
high_water_mark = 600
time_out    = 600
current_delta_set = 29912
current_number_of_blocks = 1800
flow_control = inactive
total_savevol_space = 20480000 KBytes
savevol_space_available = 20447232 KBytes (Before Flow Control)
id          = 28
name        = fs01_rep:Rochester-EMC-cs0
type        = playback
playback_state = active
high_water_mark = 300
time_out    = 300
current_delta_set = 29912
flow_control = inactive
total_savevol_space = 20480000 KBytes
savevol_space_available = 20447232 KBytes (Before Flow Control)
outstanding delta sets: <None>
communication_state = alive
current_transfer_rate = ~ 21479 Kbits/second
avg_transfer_rate = ~ 15503.8 Kbits/second
source_ip      = 168.173.44.32
source_port     = 1436
destination_ip = 10.100.44.77
destination_port = 8888
|   Source      |   Destination
Delta|Create|Time|Dur|Blocks|Playback|Time|Dur|Blocks|DSinGroup
----|-----|----|---|-----|-----|---|---|-----|-----
29911 09/24 18:34:33 1 848 09/24 18:37:55 0 848 1
29910 09/24 18:24:33 2 2595 09/24 18:27:55 0 2595 1
29909 09/24 18:14:33 2 3940 09/24 18:22:54 1 3940 1
29908 09/24 18:04:33 1 932 09/24 18:07:54 0 932 1
29907 09/24 17:54:33 1 2228 09/24 17:57:53 0 2228 1
29906 09/24 17:44:33 1 3474 09/24 17:52:52 1 3474 1
29905 09/24 17:34:33 3 3598 09/24 17:42:52 0 3598 1
29904 09/24 17:24:33 1 1328 09/24 17:27:51 0 1328 1
29903 09/24 17:14:33 1 1251 09/24 17:17:51 0 1251 1
29902 09/24 17:04:33 2 1703 09/24 17:07:50 1 1703 1
29901 09/24 16:54:33 2 2193 09/24 16:57:50 0 2193 1
29900 09/24 16:44:33 1 1435

```

29899 09/24 16:34:33 1 1421 09/24 16:52:49 1 2676 2  
29898 09/24 16:24:33 7 17299 09/24 16:47:47 2 17299 1  
29897 09/24 16:14:33 3 3398 09/24 16:22:45 0 3398 1

## **CHANGING DATAMOVER INTERFACE WHEN IP REPLICATION IS RUNNING & ACTIVE:**

1. Refresh PFS to=0, hwm=0 [cuts dataset X]
2. Checkpoint the PFS (ckpt1) [references dataset X+1]
3. Refresh the PFS to=0, hwm=0 [cuts dataset X+1]
4. Refresh the SFS to=10, hwm=0
5. Wait until PFS dataset from Step 3 is replayed on Secondary
6. Abort replication on PFS, then run abort on SFS
7. Convert SFS to rawfs
8. Change the NIC interface to the new interface on each side & then restart Replication
9. Create another checkpoint of PFS (ckpt2)
10. Do differential fs\_copy of ckpt1 to ckpt2

## **VERIFYING COMMS & PASSPHRASE BETWEEN SOURCE & DESTINATION SIDES:**

### **1. FROM SOURCE:**

**# nas\_cel -l**

```
id      name      owner mount_dev channel   net_path      CMU
0      Mankato-EMC-cs00          168.173.44.31 APM000334005350000 [Destination Celerra]
1      Rochester-EMC-cs0500     10.100.44.76  APM000334005360000 [Source Celerra]
```

**# nas\_cel -info id=0**

```
id      = 0
name    = Mankato-EMC-cs0
owner   = 0
device   =
channel  =
net_path = 168.173.44.31
celerra_id = APM000334005350000
```

**# nas\_cel -info id=1**

```
id      = 1
name    = Rochester-EMC-cs0
owner   = 500
device   =
channel  =
net_path = 10.100.44.76
celerra_id = APM000334005360000
passphrase = rdfadmin
```

### **2. FROM DESTINATION:**

**\$ nas\_cel -info id=1**

```
id      = 1
name    = Mankato-EMC-cs0
owner   = 500
device   =
channel  =
net_path = 168.173.44.31
celerra_id = APM000334005350000
passphrase = rdfadmin
```

## **IP REPLICATION FAILBACK FROM SFS TO PFS:**

1. Bring the production IP address on the SFS interface down and bring the PFS interface up (bring “UP” the alternate IP address interface on SFS)

2. Use autofullcopy to resync SFS VDM rootfs back to PFS VDM rootfs—baseline copy

**# fs\_replicate -resync root\_fs\_vdm\_pfs\_vdm1:cel=sourceCS root\_fs\_vdm\_sfs\_vdm1 -o autofullcopy=yes**

**Note:** The –resync operation goes through the following sequences & finishes by cleaning up checkpoints:

Starting baseline copy...

Starting replication...

Generating new checkpoint

Starting diff copy

3. Conduct Failback from SFS Side using new “-reverse” switch

**# fs\_replicate -reverse root\_fs\_vdm\_pfs\_vdm1:cel=sourceCS root\_fs\_vdm\_sfs\_vdm1**

4. Change VDM rootfs state on PFS side to “loaded” and SFS side to “mounted”

5. Conduct Copy of Data File System from SFS to PFS Side using –resync switch

**# fs\_replicate -resync pfs:cel=sourceCS sfs**

**Note:** Change production pfs back to “rawfs” type if required using –T switch and change mount to “ro” using mount command, though pfs remains <unmounted>

6. Conduct Failback using –reverse switch

**# fs\_replicate -reverse pfs:cel=sourceCS sfs**

7. Startup CIFS service on PFS side and note that Replication & Playback policies have reverted to default of 600secs & 600MB

8. Use new –modify switch to change Secondary PFS policies back to 300secs/300MB

**# fs\_replicate -modify sfs -o to=3,hwm=3**

9. Use –modify switch to change Source PFS policies back to 600secs/600MB (if required)

**# fs\_replicate -modify pfs -o to=6,hwm=6**

10. Re-export PFS Side “pfs” share by unexporting all Shares on PFS and reexporting “pfs” share

## **SUSPENDING REPLICATION & FAILBACK SERVICE ON PFS SOURCE SIDE:**

**# fs\_replicate -suspend pfs sfs:cel=destCS**

**Note:** Operation creates a suspend checkpoint

**# fs\_replicate -i pfs**

**Note:** Returns error: Error 2242: pfs : replication/playback is not set up

## **VERIFYING REPLICATION:**

**# fs\_replicate -list**

**# fs\_copy -i id=45**

**# server\_df pfs\_vdm1**

**# server\_df sfs\_vdm1**

## **VERIFYING ROOTFS REPLICATION INFO:**

**# fs\_replicate -info root\_fs\_vdm\_pfs\_vdm1**

## **NAS 5.3 FEATURE ALLOWS CONVERSION OF UXFS TO RAWFS:**

**# /nas/sbin/rootnas\_fs -T rawfs root\_fs\_vdm\_sfs\_vdm1 -Force**

**Note:** SFS rootfs needs to be rawfs to support replication

## **NAS 5.2 & DR1 REPLICATION:**

--Introduces differential fs\_copy and incremental failback

**Note:** SFS can now be read/write after failover and then incrementally re-synchronized prior to restoring to PFS side.

--Now supports extension of SFS if PFS is extended

--CIFS asynchronous disaster recovery is now supported

--Supports loopback replication [i.e., source and destination FS on same DM]

--Supports remote replication from one Celerra to another

--Introduces VDM support for failing over CIFS Server from Source to Target Celerra

**Note:** Correct terminology for IP Replication is “Data Recovery”, not “Disaster Recovery”

## **DIFFERENTIAL FS COPY WITH NAS 5.2:** Differences between first and second checkpoints are copied over to SFS

convert=no [Option to not convert destination fs to uxfis upon completion]

qos=<kbytes> [limit transfer speed to kbs/sec]

monitor=off [return to command prompt after executing, runs command in background]

**-fromfs <fsname>** [copy delta between old and newer checkpoint] →-fromfs command always indicates a differential copy

**1. #fs\_ckpt pfs -C**

**2. #fs\_copy -start pfs.ckpt1 sfs -o convert=no**

**3. #fs\_replicate -start pfs sfs savsize=X**

**4. #fs\_ckpt pfs -C**

## 5. #fs\_copy –start pfs.ckpt2 sfs –fromfs pfs.ckpt1

### **CELLERRA IP REPLICATION SAVVOL:**

- Minimum size 1GB and max 500GB
- Remote replication uses a SavVol on Source and on Destination sides for each Replication Session
- Local replication uses only a single SavVol per Replication Session on the local Celerra
- File system changes are stored in DM memory until they can be written to the SavVol
- File system changes are stored in the SavVol until a delta set can be accumulated, which is then transmitted via IP to destination
- Changes are stored in destination SavVol until they can be written to the SFS

### **FINDING SAVVOL VOLUME POOL NAME, VOLUME POOL FILESYSTEM NAME, AND SIZE:**

1. # export NAS\_REPLICATE\_DEBUG=1
2. Find Replication sessions & query to find SavVol sizes:

```
# fs_replicate -list
```

Local Source Filesystems

| Id | Source | FlowCtrl | State  | Destination    | FlowCtrl | State  | Network |
|----|--------|----------|--------|----------------|----------|--------|---------|
| 98 | fs02   | inactive | active | dst_fs02:nyip2 | inactive | active | alive   |
| 91 | fs01   | inactive | active | dst_fs01:nyip2 | inactive | active | alive   |

```
# fs_replicate -i id=98
```

total\_savevol\_space = 25600000 KBytes **-->25GB**

savevol\_space\_available = 25559040 KBytes (Before Flow Control)

```
# fs_replicate -i id=91
```

total\_savevol\_space = 25600000 KBytes

savevol\_space\_available = 25559040 KBytes (Before Flow Control)

3. Locate name of Replicas for each replicated production file system:

```
# nas_fs -i fs02
```

ckpts = fs02\_ckpt1,fs02\_ckpt2

replicas = fs02\_rvfs

```
# nas_fs -i fs01
```

ckpts = fs01\_repl\_restart\_1,fs01\_repl\_restart\_2

replicas = fs01\_rvfs

4. Use replicas to find the Volume name and Volume Pool File System name for the Save Volume:

```
# nas_fs -i fs02_rvfs
```

sav\_volume= vp189

member\_of = vpf96

```
# nas_fs -i fs01_rvfs
```

sav\_volume= vp140

member\_of = vpf89

5. Use following query to verify SavVol size:

```
# nas_fs -s vpf96
```

total = 25000 (sizes in MB) ( blockcount = 51200000 )

```
# nas_fs -s vpf89
```

total = 25000 (sizes in MB) ( blockcount = 51200000 )

6. Query Volume Pool name to verify vpf96 name:

```
# nas_volume -i vp189
```

id = 189

name = vp189

acl = 0

in\_use = True

type = pool

chunk\_size = 128

volume\_set = v188

disks = d21

cInt\_filesys= vpf96

```
# nas_volume -i vp140
```

id = 140

name = vp140

acl = 0

in\_use = True

type = pool

```
chunk_size = 128
volume_set = v139
disks     = d11,d18
clnt_filesys= vpfs89
```

### **USING XML QUERIES TO OUTPUT SAVVOL NAMES & SIZES:**

```
# nas_fs -query:Type=uxfs:ReplicationState=source,destination -format:"%q" -fields:ReplicationSessions -query:* -format:"RepID=%s, SavVolID=%s, SavVolName=%s, SavVolSize(MB)=%s, SavVolSizeUsed(MB)=%s\n" -
fields:Id,SrcSavVolID,SrcSavVolName,SrcSavVolSizeMB,SrcSavVolUsedMB
RepID=98, SavVolID=189, SavVolName=vp189, SavVolSize(MB)=25000, SavVolSizeUsed(MB)=40
RepID=91, SavVolID=140, SavVolName=vp140, SavVolSize(MB)=25000, SavVolSizeUsed(MB)=40
RepID=105, SavVolID=151, SavVolName=vp151, SavVolSize(MB)=25000, SavVolSizeUsed(MB)=40
# .server_config server_2 -v "volpool display vol=117"
# .server_config server_2 -v "volpool display vol=117 detail"
```

### **PREVENTING SAVVOL FROM GOING INACTIVE DURING MAINTENANCE ON SOURCE:**

- Set Data Mover param 'freeze' value to 1 on Source side:

```
param VRPL freeze=1
```

- Set Source Side policy to: to=0,hwm=256 (allows DeltaSets to remain full of data)
- Create ckpt on source side (pfs\_ckpt1)
- Check for "delta\_number" in output of "nas\_fs -i pfs\_ckpt1"
- On target side wait for delta\_number to be played back to SFS
- Target side can then be taken down

### **PROBLEM WITH LARGE DELTA SETS:**

#### CaseStudy:

Customer sets high HWM, and high system activity results in very large Deltatasets accumulating for the PFS—435GB. The problem with large deltatasets is that the system may not be able to copy the changes to the SavVol fast enough. The system has (10) VRPL.ncopythreads allocated for this activity by default. In this example, the large deltatasets are causing IO's to become blocked during the copy operation, and if this lasts (5) minutes or more, the Data Mover will panic. The problem is exacerbated in that the Panic Handler cannot recover, and the Server gets into a rolling panic situation and the IP Replication session becomes Inactive.

#### Resolution:

See ARs 70711/98226/105863 and Primus emc172332. The code solution is to increase the number of threads that can be allocated for VRPL.ncopythreads, with a max of 32 allowed. Other steps that could be taken would be to decrease the HWM values so as to create smaller deltatasets. At a minimum, increase the VRPL.ncopythreads to (32) for any sites that exhibit the following panic header:

**>>PANIC: I/O not progressing LastVol touched LF313 Kind 6 (ptr=0x97e86604)**

DART stack trace:

```
0xdb283e90: 0x13d85c waitForReboot+0x90
0xdb283eb0: 0x13da59 fault_dump+0x67
0xdb283ed0: 0x13d959 PANIC+0x29
0xdb283ee0: 0x717513 _ZN9RawIO_IPR13checkProgressEv+0x10d
0xdb283fd0: 0x7173fc _ZN17IrpProgressThread5startEv+0x6
0xdb283fe0: 0x140ff0 Sthread_startThread_internal+0xe
```

#### Checking for Number of ncopythreads:

```
$ .server_config server_2 -v "replica display srcvol=<srcvol_ID>
```

#### Monitoring Deltaset copy progress:

```
$ .server_config server_2 -v "replica stat srcvol=<srcvol_ID>
```

### **CONFIGURING SAVVOL SIZE:**

#### /nas/sys/nas\_param

```
#replication:<percent_of_fs_vol>
local_rep_save_vol: 10: chunk_size=32768 [Chunk size is fixed at 128MB]
remote_rep_save_vol: 10: chunk_size=32768 [10=10% or default size of SavVol based on 10% of PFS Size—this is adjustable]
Comment: PFS & SFS “SavVol” sizes must be the same—changing the nas_param default for SavVol is not honored when establishing Replication Sessions using Celerra Mgr—will revert to the standard 10% of PFS size when creating SavVol
```

#### NAS 5.6 EXAMPLE:

```
# cat /nas/sys/nas_param
```

replication:10:

### **ERROR WHEN TRYING TO START REPLICATION:**

```
# fs_replicate -start fs1 fs2:cel=cel2
```

Error 3024: no free disks are available

**Note:** Problem can occur if there isn't enough space (10% of fs size is needed) to create the SavVol in the respective system pool. Use the “savsize=20480” option to specify a SavVol size that can be created—or, add disk space to appropriate system pool.

### **NETWORK ISSUES & SAVVOL:**

Default replication policy settings allow delta sets to be transferred to SavVol until the SavVol runs out of space, at which point replication is aborted, but the primary file system continues to function properly. There is an optional policy to disallow further writes to the PFS if the SavVol fills.

### **SAVVOL FULL ISSUE:**

**Error 4019:** failed to complete command—#fs\_replicate –refresh [Command fails with this message]

Problem would occur if SFS is not copying changes over fast enough to prevent SavVol from filling on PFS side. By default, PFS SavVol is only 10% the size of the PFS. If this issue occurs, abort Replication and reconfigure to create a larger SavVol, possibly changing HWM and TO values on each side as well.

**Rule of Thumb:** HWM & TO values should be greater on PFS than SFS side.

### **DETERMINING AMOUNT SAVVOL SPACE AVAILABLE:**

# **fs\_replicate -l** [obtain id number of any file system]

# **fs\_replicate -i id=xxx**

total\_savevol\_space = 1048576 KBytes

savevol\_space\_available = 917504 KBytes (Before Flow Control)

### **SAVVOL POLICIES THAT CAN BE SET FOR IP REPLICATION:**

\$ .server\_config server\_2 -v "param VRPL freeze"

VRPL.freeze INT 0x018f7d58 0 0 (0,1) TRUE NONE 'Enables freeze of primary when savVol full'

**Note:** Setting value to 1 would halt all IO to source PFS in order to preserve synchronization state, and can also be invoked with fs\_replicate –modify switch using “autofreeze=yes”

\$ .server\_config server\_2 -v "param VRPL readonly"

VRPL.readonly INT 0x018f7d98 0 0 (0,1) TRUE NONE 'Determines action when changes can't be tracked'

**Note:** Setting value to 1 changes source PFS to Read-Only to preserve state, also invoked via –modify switch, “autorw=yes”

### **THREE POSSIBLE OUTCOMES WHEN SAVVOL BECOMES FULL:**

→PFS will freeze, preventing any further Reads or Writes, if param VRPL freeze is set to 1

→PFS will become RO if param VRPL readonly is set to 1

→Once SavVol is full, and changes can no longer be written to DM memory, replication sessions that depend on the SavVol will become inactive

### **(3) DISTINCT PHASES OF IP REPLICATION WITH NAS 5.2:**

**FAILOVER→RESYNC to REVERSE IP REPLICATION DIRECTION→FAILBACK**

#### **NAS 5.2 IP REPLICATION FAILOVER:**

# **fs\_replicate –failover PFS:cel=NS600S SFS –o sync**

**Note:** Conducts failover from PFS to SFS and stops replication service [-o sync = primary is up; -o now = primary down].

#### **RESYNCING SFS SIDE WITH NAS 5.2—DIFFERENTIAL COPY:**

# **fs\_replicate –Resync PFS:cel=NS600S SFS**

**Note :** Incrementally copies and synchronizes the SFS, which is RW, to the PFS when in a “failover” state

#### **NAS 5.2 IP REPLICATION FAILBACK:**

# **fs\_replicate –failback PFS:cel=NS600S SFS**

**Note:** When conducting failback, nas\_cel is used to talk to destination Apache Server, a CGI script is initiated, then CGI Eiffel code executes, which assembles the fs\_replicate –start command from the remote to destination side to reverse the replication process.

#### **NAS 5.2 VERIFY REPLICATION:**

# **fs\_replicate –info PFS | SFS**

**Note:** By default, IP Replication takes place at 10 minute mark or after 600MB have been written to PFS, whichever occurs first

#### **NAS 5.2 CONDUCTING IP REPLICATION REFRESH:**

1. # **fs\_replicate –refresh SFS –o to=300, hwm=300**

**Note:** Default playback service values are 600

2. # **fs\_replicate –refresh PFS –o to=600, hwm=600**

#### **NAS 5.2 IP REPLICATION ENHANCEMENTS:**

→Auto extension of SFS when PFS is extended

→Special filesystems now hidden from view: vefs, ipfs, rufs, pufs

→Aborting an fs\_copy should now work

## **NAS 5.2 IP REPLICATION LIMITATIONS:**

- HighRoad works on PFS only, not SFS
- SRDF/TimeFinder/FS cannot be used in conjunction with IP Replication on same filesystems [except SRDF Sync DR]
- Do not use BCV's for either PFS or SFS used in IP Replication
- CDMS is not supported and MGFS file systems cannot be replicated

## **DEFAULT FS REPLICATE FAILOVER PROCESS:**

1. All deltas played back on secondary
2. Stop playback on SFS & remount SFS as “rw”
3. Refresh PFS to=0, hwm=0 & remount PFS “ro”

## **FS REPLICATE FAILOVER NOW:**

1. Stop playback on SFS & remount SFS “rw”
2. Refresh PFS to=0, hwm=0 & remount PFS “ro”

**Note:** Failover command must be run on SFS Control Station

## **FS REPLICATE FAILBACK:**

1. Refresh PFS to=0, hwm=256 & refresh SFS to=10, hwm=256
2. Wait for dataset to be identical for PFS/SFS
3. Refresh PFS to=128, hwm=128
4. Again wait for dataset to be identical for PFS/SFS
5. Remount SFS “ro”
6. Refresh PFS to=0, hwm=0
7. Wait for dataset to be identical for PFS/SFS
8. Stop playback on SFS and replicator on PFS
9. Start replication from SFS to PFS
10. Remount PFS “rw”

**Note:** Failback must be run from side that is “rw” & effectively changes the side that does replication

## **CELLERRA REPLICATOR COMMANDS:**

|                        |                                                                                                 |
|------------------------|-------------------------------------------------------------------------------------------------|
| #fs_replicate          | [Used to Start/Abort/Refresh IP Replication]                                                    |
| #fs_copy               | [Used to copy data from PFS Checkpoint to SFS during initial setup]                             |
| #nas_cel/nas_rdf -init | [Used to initialize Control Stations]                                                           |
| #nas_fs -i pfs/sfs     | [Used to verify Delta Sets on PFS or SFS [Delta Set increments each time TO or HWM takes place] |

## **Other Commands:**

|                                       |                                                                              |
|---------------------------------------|------------------------------------------------------------------------------|
| #fs_replicate -failover -o sync   now | [sync=sync pfs to sfs after failover; now=disaster on primary, failover now] |
| #fs_replicate -info                   | [Current state of replicator]                                                |
| #fs_replicate -failback               | [failback to PFS]                                                            |
| #fs_replicate -reverse                | [new syntax with NAS 5.3 that replaces –failback]                            |

## **LOCAL vs REMOTE IP REPLICATION:**

→ Local Replication produces RO copy of PFS for export by same or other Data Movers (aka Loopback Replication), but shares SavVol between Source & Destination Servers

→ Remote Replication does same thing for Data Movers at remote locations but uses separate SavVol for Source & Destination

## **REMOTE REPLICATION PROCESSES:**

- Initial replication manually synchronizes both PFS and SFS after starting Celerra Replicator for first time
- Addresses of subsequent block changes to PFS are sent to Log & and applied to SFS SavVol after sync is completed
- Replication Svc creates Delta set to source primary SavVol (record of all changes to PFS for each Delta Set trigger)
- Replication Svc transfers Delta set to target secondary SavVol (one Delta Set per FS is processed at a time)
- Replication Playback Svc reads Delta set and updates SFS

## **CREATING SFS THAT IS IDENTICAL IN SIZE TO PFS:**

**Note:** As stated in earlier notes, IP Replication will not work if the SFS is not exactly the same size as the PFS. The following two procedures only work with NAS Codes 5.1.14.0 or higher

## **CONVERT NUMBER OF BLOCKS IN EXISTING META VOLUME ON FS TO MBYTES:**

1. Find meta volume(s) that filesystem is built upon:

**#nas\_fs -i fs01 →mtv1**

2. Find Volume ID # of Meta:

### #nas\_volume -i mtv1 →id =237

3. Obtain number of blocks in Meta:

#nas\_volume -query:id=237 -format:'%s\n' -fields:blocks →141418496

4. Multiply number Blocks by 512 and then divide by 1048576:  $141418496 * 512 = 72406269952 / 1048576 = 69052$  MB

5. Build SFS:

#nas\_fs -n sfs -t rawfs -c size=69052MB profile=clar\_r5\_performance -o slice=y

### CREATING SFS FROM SLICES:

1. Create slice of desired size: \$nas\_slice -n slc1 -c str1 1 0 -src mtv1:cel=remcs0

2. Create meta from stripe: \$nas\_volume -n mtv1 -c slc1

3. Create SFS FileSystem: \$nas\_fs -n sfs -t rawfs -c mtv1

### VERIFYING THAT PFS & SFS HAVE SAME FS & BLOCK SIZES:

1. #nas\_volume -query:Name==v99 -fields:Blocks (File System size)

40960

2. #nas\_volume -query:Name==v99 -fields:BlockSize (Block Size)

512

### IP REPLICATION PROCESSES:

**rvfs:** replication volume file system (Service that creates & copies ‘delta sets’ to SavVol & manages ‘Redo Log’ buffer for SavVol)

**pvfs:** playback volume file system (Service that copies from Save Volume to Secondary File System—SFS)

**vvfs:** volume pool file system (Service used for replication by allocating blocks for the volumes used in replication)

**ipfs:** ip file system (Service that handles data for IP transfer; IPFS Service checks for new Delta Sets in SavVol every second)

### REPLICATION SESSION EXAMPLE:

|                   |                              |
|-------------------|------------------------------|
| pfs_rvfs1         | [Source Replication Service] |
| pfs_rvfs1_ipfs1   | [Source IP Transfer service] |
| sfs_pv vfs1       | [Remote Playback Service]    |
| sfs_pv vfs1_ipfs1 | [Remote IP Transfer service] |

### ABORTING AN FS COPY SESSION:

Run from Source side  
# fs\_copy -abort usilst20\_ftpu\_ftp\_ckpt3

operation in progress (not interruptible)...id = 1060

name = usilst20\_ftpu\_ftp\_ckpt3

### ABORTING FS REPLICATION SESSION:

# fs\_replicate -abort fs\_source,fs\_dest:cel=laip2

**Note:** This is preferred abort method, clean up both sides simultaneously when possible

# fs\_replicate -abort new1

operation in progress (not interruptible)...id = 152

name = new1

acl = 0

in\_use = True

type = uxfs

worm = off

volume = v113309367

pool = clar\_r5\_performance

member\_of = root\_avm\_fs\_group\_3

rw\_servers= laip1-2

ro\_servers=

rw\_vdms =

ro\_vdms =

auto\_ext = no,virtual\_provision=no

ckpts = new1\_ckpt1,new1\_ckpt2

ip\_copies = repdest:laip2

**Note:** Run on Source and Destination Control Stations to cleanup entries in netd file and boot.cfg for replication service and sessions

### CLEANING UP FAILED –abort REPLICATION SESSIONS MANUALLY:

#### PRIMARY SIDE:

Need to know repGroupName, SavVol\_id, SrcVol\_id, & CfgVol\_id in order to abort a file system replication session

**SECONDARY SIDE:**

Need to know VolMCastName, MultiCastNode, SrcVolSec\_id, SavVolSec\_id, CfgVolSec\_id in order to abort the Secondary side.

**Note:** Derive source information using nas\_fs -i srcfs\_rvfs\_ipfs1 & nas\_fs -i rvfs output or from the contents of the /nas/volume/filesys, netd, and ufs file of server. Derive Secondary information using nas\_fs -i destfs\_pvfs and nas\_fs -i destfs\_pvfs\_ipfs1. Depending on whether abort was partially successful, not all entries may be present in filesys file. May also need to reboot Data Mover if commands do not execute.

**EXAMPLE FOR SOURCE SIDE CLEANUP:**

I. Use following DART commands to cleanup

|                                                                           |                                             |
|---------------------------------------------------------------------------|---------------------------------------------|
| \$server_config server_2 -v "iprepvc stop <rcpGroupName> vol=<SavVol_id>" | [stops Replication Service on Primary side] |
| \$server_config server_2 -v "rcproute remove group=<rcpGroupName>"        | [removes RCP route entry]                   |
| \$server_config server_2 -v "replica stop srcvol=<SrcVol_id>"             | [Stops replication on Primary]              |
| \$server_config server_2 -v "volpool stop vol=<CfgVol_id>"                | [Stops VolPoolManager on config volume]     |
| \$server_config server_2 -v "volpool stop vol=<SavVol_id>"                | [Stops VolPoolManager on SavVol]            |

II. Remove any required entries from netd file for iprepvc and rcproute

III. Remove volpool & replica entries from ufs file if required [used for replication recovery]

IV. Cleanup CS database

a.) # export NAS\_REPLICATE\_DEBUG=1

b.) # server\_umount server\_2 -p /srcfs\_rvfs\_ipfs1

c.) Verify that /nas/volume/filesys entry for srcfs\_rvfs\_ipfs1 shows “n” for inuse flag and has blank ref count—edit if required  
201:srcfs\_rvfs\_ipfs1:0:n:5:0::1143306484:200:9::202: [inuse & ref count fields, respectively]

d.) # nas\_fs -d srcfs\_rvfs\_ipfs1 [Verify that this entry has been removed from filesys file]

e.) # fs\_replicate –abort srcfs [Verify that session is aborted using fs\_replicate –list]

**EXAMPLE FOR DESTINATION SIDE CLEANUP:**

I. Use DART commands:

|                                                                                             |                                      |
|---------------------------------------------------------------------------------------------|--------------------------------------|
| \$server_config server_2 -v "volmcast delete <VolMCastName> <MultiCastNode> <SavVolSec_id>" | [deletes Volmcast object]            |
| \$server_config server_2 -v "playback stop srcvolsec=<SrcVolSec_id>"                        | [Stops playback service]             |
| \$server_config server_2 -v "volpool stop vol=<SavVolSec_id>"                               | [Stops VolPoolManager on SavVol]     |
| \$server_config server_2 -v "volpool stop vol=<CfgVolSec_id>"                               | [Stops VolPoolManager on Config Vol] |

II. Remove netd entries for Volmcast object

III. Remove volpool entries for SavVol & CfgVol, and playback entry from ufs file

IV. Cleanup CS database

a.) # export NAS\_REPLICATE\_DEBUG=1

b.) # server\_umount server\_2 -p /destfs\_pvfs\_ipfs1

c.) Make sure that filesys entry shows ‘n’ for inuse, has no reference count, and does not have remote reference  
125:destfs\_pvfs\_ipfs1:0:n:5:0::1143306476:124:7::: [inuse & ref count fields, respectively]

d.) # nas\_fs -d destfs\_pvfs\_ipfs1

e.) # fs\_replicate –abort destfs [Verify using fs\_replicate –list]

**# NAS\_REPLICATE\_DEBUG=1;export NAS\_REPLICATE\_DEBUG****VIEWING PFS IP REPLICATION PROCESSES:**

**\$nas\_fs -i pfs**

|                     |                                                                                                  |
|---------------------|--------------------------------------------------------------------------------------------------|
| vpfss20             | [Represents Save Volume on PFS side—now seen when doing nas_fs -i on PFS file system checkpoint] |
| pfs_rvfs1           | [Replication Service]                                                                            |
| pfs_rvfs1_ipfs1     | [IP Transfer service]                                                                            |
| 22_0028060001810047 | [FS Group for the IP Transfer]                                                                   |

**VIEWING SFS IP REPLICATION PROCESSES:**

**\$nas\_fs -i sfs**

|                 |                                      |
|-----------------|--------------------------------------|
| vpfss21         | [Represents Save Volume on SFS side] |
| sfs_pvfs1       | [Playback Service]                   |
| sfs_pvfs1_ipfs1 | [IP Transfer service]                |

**TROUBLESHOOTING IP REPLICATION—RESTRICTIONS/LIMITATIONS:**

--Server Log, sys\_log, & /nas/http/error\_log

--Date on both Local and Remote Control Stations must be the same

--‘Passphrase’ passwords must also be the same on Local & Remote Control Stations

--Must use Celerra WebGUI to change Control Station Hostnames & Timezones for IP Replication to work

--PFS & SFS must be identical in size & use the same “slice” size

--‘Read Only server not found for pfs’ might mean that Checkpoint Volume was not correctly specified

--‘Replication Inactive’ in Server Log requires stopping replication service, adjusting Save Volume and Timeout/HWM values, and

restarting the Replication process—could occur if changed Data Size is greater than SavVol size  
 --Volume Pool is required for IP Replication  
 --Remote SFS must be created as “rawfs” to work [Converts to “uxfs” after FS\_Copy completes]  
 --IP Replication works for SYMMETRIX, FC4700, and CX600  
 --Backup SFS to tape if HWM=0 & TimeOut (to)=0 during replication  
 --CIFS IP Replication requires setup similar to a ‘migration’ [ie., lgdup, sharedup, etc]  
 --Can specify different DataMover to transfer Replication data to reduce loads on busy Production Servers  
 --Published Transfer rate on 510 DM is 10MBps [Actual rates may vary→Fast Ethernet=8-10MBps; Gigabit=25-30MBps]  
**Example:** At 1MBps, will take (10) hours to copy 36GB data  
 --IP Replication will become Inactive if the Delta Sets become bigger than the Save Volume  
 --Replication will fail if PFS & SFS are not identical, with “pfs size () is not equal to sfs size ()” error  
 --DataMover reboots or failovers should not stop ‘Replication’ process, though panics may result in need to start over  
 --PFS cannot be extended while running Replication Service

### **TROUBLESHOOTING IP REPLICATION ISSUES:**

```
# export NAS_DB_DEBUG=1
# export NAS_REPLICATE_DEBUG=1 [/nas/log/nas_log.al.tran | nas_log.al.trace | nas_log.al.remote]
# env |grep NAS_DB
NAS_DB=/nas
NAS_DB_DEBUG=1
#export NAS_REPLICATE_DEBUG=1 [turns on Debug; use 0 to turn off]
#echo $NAS_REPLICATE_DEBUG (Use to verify that variable has been exported)
```

**1**

**Note:** After setting debug mode, use nas\_fs -l to see if there are any \_pvfs or \_pvfs\_ipfs1 sessions, which would indicate that replication is running. Introduced with NAS 5.2

\$ .server\_config server\_2 “logsys set severity CMD-LOG\_DEBUG” →Cognac replication debug setting

# nas\_event -list -f VRPL

|    |                                                   |
|----|---------------------------------------------------|
| id | description                                       |
| 0  | Replication ok                                    |
| 1  | Replication on Source Filesystem Inactive         |
| 2  | Resync asked by (previous) Destination Filesystem |
| 3  | Source Filesystem Switch Delta on HWM             |
| 4  | Destination Filesystem in Error                   |
| 5  | IP Rep Svc failed - Transport                     |
| 6  | IP Rep Svc NetWork or Receiver Down               |
| 7  | IP Rep Svc Network or Receiver Up                 |
| 8  | Rep Svc Source Filesystem Flow on Hold            |
| 9  | Rep Svc Source Filesystem Flow Resumed            |
| 10 | Source Filesystem Frozen                          |
| 11 | Source Filesystem Thawed                          |
| 12 | Last Delta Replayed on Destination Filesystem     |
| 13 | Redo Buffer close to overflow event               |
| 14 | Source mounted RO                                 |
| 15 | Source mounted RW                                 |
| 16 | Replication in error                              |

# .server\_config server\_3 -v "param VRPL"

| Name                | Location   | Current    | Default    |
|---------------------|------------|------------|------------|
| VRPL.freeze         | 0x02115d18 | 0x00000000 | 0x00000000 |
| VRPL.nchunkreserved | 0x01fc6480 | 0x00000002 | 0x00000002 |
| VRPL.ncopythreads   | 0x01fc65d8 | 0x0000000a | 0x0000000a |
| VRPL.readonly       | 0x02115d58 | 0x00000000 | 0x00000000 |

### **EXAMPLE OF HIDDEN VOLUMES:**

**1.** # nas\_fs -i ckpt\_ebf1\_0041

```
id      = 392
name   = ckpt_ebf1_0041F
acl     = 0
in_use  = True
type    = ckpt
volume  = vp147
```

```
pool      = clar_r5_economy
member_of = [volume is hidden by default]
2. #NAS_REPLICATE_DEBUG=1;export NAS_REPLICATE_DEBUG=1
3. # nas_fs -i ckpt_efb1_0041
id      = 392
name    = ckpt_efb1_0041
acl     = 0
in_use  = True
type    = ckpt
volume  = vp147
pool    = clar_r5_economy
member_of = vfps104 [volume no longer hidden]
```

### **VOLUME POOL DATA:**

```
#.server_config server_2 -v "volpool display vol=285 detail"
```

```
2004-11-19 09:12:13: VRPL: 4: SuperBlockHeader ChunkSize=262144, NChunkInVol=7,
2004-11-19 09:12:13: VRPL: 4: NextChunkToAllocate=263168, NextVolGenCount=1,
2004-11-19 09:12:13: VRPL: 4: Allocate Mode=WRAP AROUND LastUsedChunk=263168.
2004-11-19 09:12:13: VRPL: 4: Header Bad:1024.
2004-11-19 09:12:13: VRPL: 4: Header Bad:1311744.
2004-11-19 09:12:13: VRPL: 4: Header Bad:1573888.
2004-11-19 09:12:13: ADMIN: 4: Command succeeded: volpool display vol=285 detail
```

**Note:** 285 = vol id of SavVol—command will also output details of Checkpoints, i.e., any volume in a VolPool

### **DISPLAYING SAVVOL INFORMATION FOR IP REPLICATION:**

```
# nas_fs -query:Type=uxfs:ReplicationState=source,destination -format:"%q" -fields:ReplicationSessions -query:* -format:"RepID=%s, SavVolID=%s, SavVolName=%s, SavVolSize(MB)=%s, SavVolSizeUsed(MB)=%s\n" -fields:Id,SrcSavVolID,SrcSavVolName,SrcSavVolSizeMB,SrcSavVolUsedMB
```

```
RepID=35, SavVolID=117, SavVolName=vp117, SavVolSize(MB)=1024, SavVolSizeUsed(MB)=128
RepID=47, SavVolID=129, SavVolName=vp129, SavVolSize(MB)=1024, SavVolSizeUsed(MB)=128
RepID=35, SavVolID=117, SavVolName=vp117, SavVolSize(MB)=1024, SavVolSizeUsed(MB)=128
RepID=47, SavVolID=129, SavVolName=vp129, SavVolSize(MB)=1024, SavVolSizeUsed(MB)=128
```

```
# cat /nas/volume/filesys |grep -i vfps
```

```
30:vpfs30::0:21::n:11:112:::1189024275:23:2::24:31,38:
35:vpfs35::0:1::n:11:117:::0::24:36,37: →(Blue = SavVol volume ID, Red = Checkpoints)
42:vpfs42::0:21::n:11:124:::1189024538:25:2::24:43,50:
47:vpfs47::0:1::n:11:129:::0::24:48,49: →(Blue = SavVol volume ID, Red = Checkpoints)
```

```
# .server_config server_2 -v "volpool display vol=117"
```

```
# .server_config server_2 -v "volpool display vol=117 detail"
```

```
1189098276: STORAGE: 5: opening volume 117, ref 4, ioObj b3defb84
1189098276: VRPL: 5: VolPoolManager::reference() now refCount:4.
1189098276: VRPL: 5: VolPoolManager::dereference() now refCount:3.
1189098276: STORAGE: 5: closing volume 117, ref 5, ioObj b3defb84
1189098276: VRPL: 4: SuperBlockHeader ChunkSize=262144, NChunkInVol=7,
1189098276: VRPL: 4: NextChunkToAllocate=1049600, NextVolGenCount=123,
1189098276: VRPL: 4: Allocate Mode=WRAP AROUND LastUsedChunk=1573888.
1189098276: VRPL: 4: IPRepSenderMailBox:
1189098276: VRPL: 4: PlaybackMailBox:
1189098276: VRPL: 4: ind:0 Ad:787456 VolGen:122 ChunkSize:262144
1189098276: VRPL: 4: Replication Chunk header on address 1024.
1189098276: VRPL: 4: Owner=, ValidFlag=1, DeltaSetStatus=1, DeltaSetGenCount=119,
1189098276: VRPL: 4: VolumeGenCount=119, ChunkNum=0, NChunkInDeltaSet=1,
1189098276: VRPL: 4: NextDeltaSet=263168, NextChunk=263168, FirstChunkInDeltaSet=1024,
1189098276: VRPL: 4: NBlockInChunk=1, NBlockOnFsVolume=20480000, NBlocksModified=1,
1189098276: VRPL: 4: DataSize=8192, ChunkSize=262144, TOCFlushedFlag=1,
1189098276: VRPL: 4: CheckSumActive=1, ChunkHeaderCheckSum=1656619566, TOCCheckSum=16,
1189098276: VRPL: 4: DeltaSetSuspend=0 -----output abridged-----
```

### **DISPLAYING IP REPLICATION ROUTING INFO BETWEEN LOCAL AND TARGET SYSTEMS AND PORT USED:**

**# .server\_config server\_2 -v "rcproute display group=34\_APM000421031830000"**

1106757717: RCPD: 4: rcp:: Group:34\_APM000421031830000

1106757717: RCPD: 4: rcp:: Target\_num:0 Target\_ip:10.241.169.58 Local\_ip:10.241.169.54 Local\_port:57694

### **USING RCPROUTE TO RENEW IP ROUTE:**

**# .server\_config server\_2 "rcproute add group=128\_CK2000411006400000 nexthop=172.24.16.119 fromhop=172.24.16.111"**

### **DELETING AN RCPROUTE FROM SOURCE SIDE:**

**\$ .server\_config server\_2 "rcproute remove group=76\_APM000417003390000"**

**Note:** Use to remove data mover memory reference when the Group no longer exists on the Destination side

### **STOPPING INVALID IPREP SENDER SERVICES ON SOURCE SIDE:**

**\$ .server\_config server\_2 "iprepsvc stop 76\_APM000417003390000"**

**Note:** Might use this to delete from Dart memory if the actual database reference has been deleted but IPRepsvc remains

### **DESTINATION PLAYBACK DISPLAY COMMAND:**

**\$ .server\_config server\_x -v "playback display srcvolsec=143"**

**\$ .server\_config server\_x -v "playback stat srcvolsec=143"**

### **REVIEWING DELTA SET INFORMATION BETWEEN TARGET & SOURCE SIDES:**

#### **SOURCE SIDE:**

\$ grep fs04 /nas/volume/filesys

25:fs04:0:y:1:**536**:1:::0:76,24@1,134::: →File System in question is “fs04” with volume ID 536

**\$ .server\_config server\_2 -v "replica stat srcvol=<vol\_id>"**

**\$ .server\_config server\_2 -v "replica stat srcvol=369"**

1142384207: VRPL: 5: Entering replication.

1142384207: VRPL: 4: RepControl::Stat() Begin

1142384207: VRPL: 4: LogFilterVol::Stat() Begin

1142384207: VRPL: 4: 03/14/2006 04:44:04 CT:0 DS:16956 VolGC:39354 DSAd:15205376

NBMod:6.

**\$ .server\_config server\_2 -v "replica display srcvol=536"**

1132351729: VRPL: 5: Entering replication.

1132351729: VRPL: 4: RepControl::Audit() Begin

1132351729: VRPL: 4: RepControl:: Flow\_Control:Inactive PFS\_State:Active

1132351729: VRPL: 4: LogFilterVol::Audit() Begin

1132351729: VRPL: 4: LogFilterVol:: srcvol=536 savevol=672 iosz=8192 copyiosz=32768 hwm=600 to=600 checksum=TRUE

1132351729: VRPL: 4: LogFilterVol:: state:2 CurDeltaSet:3524

1132351729: VRPL: 4: LogFilterVol:: nInTOC:16853, hwmbModified:76800, LastBlockCopied:4215537712

1132351729: VRPL: 4: LogFilterVol:: Available\_Space\_Before\_Flow\_Control:335151104 (KBytes)

1132351729: VRPL: 4: LogFilterVol:: nBlockRemainingForCopy:0

1132351729: VRPL: 4: LogFilterVol:: HighCapacityMode:1 SystemCapacityMode:1

1132351729: VRPL: 4: LogFilterVol:: TOC audit

1132351729: VRPL: 4: NewBlockMap::Audit() Begin

1132351729: VRPL: 4: NewBlockMap:: MemUsage:315536 bytes

1132351729: VRPL: 4: Number of FS Blocks in BlockMap = 16853

1132351729: VRPL: 4: NewBlockMap::Audit() End

1132351729: VRPL: 4: LogFilterVol:: OldTOC audit

1132351729: VRPL: 4: RedoLog::Audit() Begin

1132351729: VRPL: 4: RedoLog:: CfgName:673, RedoKey:0x437e4fa1:0xc6e20

1132351729: VRPL: 4: RedoLog:: ChunkFirstBlock:1024, CurBlock:9620726743040

1132351729: VRPL: 4: RedoLog::Audit() End

1132351729: VRPL: 4: LogFilterVol::Audit() End

1132351729: VRPL: 4: RepControl:: PFS policy:NoPolicy

1132351729: VRPL: 4: RepControl:: hwmPolicy=600 toPolicy=600 pfsPolicy:NoPolicy

#### **DESTINATION SIDE:**

\$ grep fs04 /nas/volume/filesys

24:fs04:0:y:1:189::1:1131393468:25@1:11:63,106:25:: →File system “fs04” is volume ID 189

**\$ .server\_config server\_x -v "playback stat srcvolsec=<vol\_id>"**

**\$ .server\_config server\_2 -v "playback display srcvolsec=189"**

```
1132350613: VRPL: 5: Enter Playback srcvol 189 savevol cfg threads 10
1132350613: VRPL: 4: PlaybackFilter::Audit() Begin
1132350613: VRPL: 4: PlaybackFilter:: srcvols=189 savevol=240 cfgvol=241 threads=10 hwm=200 to=300 checksum=TRUE
1132350613: VRPL: 4: PlaybackFilter:: state:0 CurDeltaSet:3522
1132350613: VRPL: 4: PlaybackFilter:: CurVolGenCount:39219, CurrentChunkHeaderAd:226493440
1132350613: VRPL: 4: PlaybackFilter:: playuntildelta:NotDone
1132350613: VRPL: 4: PlaybackFilter:: SuspendAsked:False, SuspendDone:False, SuspendNow:False
1132350613: VRPL: 4: PlaybackFilter:: NBlocksRemainingForCopy:0
1132350613: VRPL: 4: PlaybackFilter:: hwmPolicy=200 toPolicy=300
1132350613: VRPL: 4: PlaybackFilter:: HighCapacityMode:1 SystemCapacityMode:1
```

### **RE-TRANSMITTING DELTASET CHUNKS FROM R1 TO R2 SIDE MANUALLY:**

1. **`$.server_config server_x "iprepsvc transfer <rcp_group> vol=<savvol_id>`**

**`add=<addr_of_chunk_to_transfer>`**

**Note:** rcp\_group name can be found in netd file

2. Stop playback on Secondary Side and start Playback Recovery Service:

**`$.server_config server_x -v "playback stop srcvols=xxx"`**

**`$.server_config server_x -v "playback recover srcvols=xxx"`**

**`# .server_config server_2 -v "volpool display vol=113 detail"`**

### **IPREPSVC DISPLAY NAS 5.3:**

**`# .server_config server_2 -v "iprepsvc display 107_CK2000411006400000 vol=243"`** [Found in Source netd file]

```
1100373872: VRPL: 4: IPRepSender::Audit() Begin
1100373872: VMCAST: 4: VolMCast::display() Name:243_107_CK2000411006400000 VolName:243
NodeName:107_CK2000411006400000 refcount:1
1100373872: VMCAST: 4: VolMCastSender::display() NextBlock:20186112 TotalBlock:52428800 remote_communication:Alive
NextBeingProcessed:20186176
1100373872: RCPD: 4: rcp:: Group:107_CK2000411006400000
1100373872: RCPD: 4: rcp:: Target_num:0 Target_ip:172.24.16.119 Local_ip:172.24.16.111 Local_port:61972
1100373872: VRPL: 4: IPRepSender:: CurrentIPMailNum:6645 vol:243 ad:0.
1100373872: VRPL: 4: IPRepSender:: QOS:0 kbytes/sec
1100373872: VRPL: 4: IPRepSender:: nextVolGenCount:6645 lastVolGenCount:6644
1100373872: VRPL: 4: IPRepSender::Audit() End
1100373872: ADMIN: 4: Command succeeded: iprepsvc display 107_CK2000411006400000 vol=243
```

### **DISPLAYING IPREPSVC STATUS:**

**`# .server_config server_2 -v "iprepsvc display 34_APM000421031830000 vol=134"`**

```
1106757497: VRPL: 4: IPRepSender::Audit() Begin
1106757497: VMCAST: 4: VolMCast::display() Name:134_34_APM000421031830000 VolName:134
NodeName:34_APM000421031830000 refcount:1
1106757497: VMCAST: 4: VolMCastSender::display() NextBlock:0 TotalBlock:2097152
remote_communication:Alive NextBeingProcessed:0
1106757497: RCPD: 4: rcp:: Group:34_APM000421031830000
1106757497: RCPD: 4: rcp:: Target_num:0 Target_ip:10.241.169.58 Local_ip:10.241.169.54 Local_port:57694
1106757497: VRPL: 4: IPRepSender:: CurrentIPMailNum:5747 vol:134 ad:1024.
1106757497: VRPL: 4: IPRepSender:: QOS:0 kbytes/sec
1106757497: VRPL: 4: IPRepSender:: nextVolGenCount:1856687 lastVolGenCount:1856686
1106757497: VRPL: 4: IPRepSender::Audit() End
```

**`# .server_config server_2 -v "iprepsvc stat 34_APM000421031830000 vol=134"`**

1106757592: VRPL: 4: IPRepSender:: Stat() Begin.

1106757592: VRPL: 4: IPRepSender:: Stat() End.

**Note:** Displays the current State of the IP Replication Service—iprepsvc info is derived from Source data mover netd file

### **USING IPREPSVC TO RESTART IPREPSENDER SERVICE FOR REPLICATION:**

**`# .server_config server_2 "iprepsvc start 128_CK2000411006400000 vol=273"`**

### **USING REPLICA DISPLAY:**

```
$ .server_config server_2 -v "replica display srcvol=113"  
# .server_config server_2 -v "replica stat srcvol=113"  
# .server_config server_2 -v "volpool display vol=113 detail"
```

## **CLEANING UP VOLMCAST SERVICE ON DESTINATION AFTER UNSUCCESSFUL ABORT:**

```
#.server_config server_2 -v "volmcast delete 27 403_APM000426033100000 129"
```

**Note:** Or, comment out these lines in netd file and reboot server. Volmcast is the fs\_copy receiving service that runs on destination.

```
#.server_config server_2 -v "volmcast display 27 403_APM000426033100000 129"
```

```
#/nas/bin/nas_cmd@rootfs_ip -Remove fs_aspen_e_ipfs1
```

**Note:** Another method for deleting the Volmcast service on destination side?

## **BANDWIDTH PERFORMANCE SETTINGS:**

```
$ .server_config server_x -v "param IPREP qos=0"
```

**Note:** Default setting is 0 for unlimited bandwidth, but in practice needs to be set if multiple IP Replication sessions are being used.

## **INCREASING REDO LOG SIZE:**

### **Server Log Messages:**

RedoBuf should be extended or Config Vol to slow for vol:LF150

**Cause:** Frequency of writes to the PFS is more than the redolog can handle. The redolog keeps track of the changes to the PFS in case there is a Server reboot. To increase the redolog change the param VRPL nredobuf to 8 (default is 4)—requires reboot.

### **param VRPL nredobuf=8**

**Note:** This parameter is hidden in 5.4 and is only visible when you enter the actual param name—output below shows default and actual setting is 4. This parameter is only applicable on the SOURCE replication side.

```
$ .server_config server_2 -v "param VRPL nredobuf"
```

```
VRPL.nredobuf INT 0x01f0f4c0 44 (0,4294967295) FALSE REBOOT 'NA'
```

## **TAPE BACKUPS:**

→BackUp to tape from SFS using hwm=0,to=0 options while conducting BackUps

## **USING DRIVE OR TAPE BACKUPS TO SETUP DESTINATION REPLICATION:**

→Basic mechanics are to backup the initial Checkpoint of the production file system, and then restore to the destination file system (mounted as rawfs) from disk or from tape backup, aka tape-silvering (using NDMP)

→NDMP environmental variable must be VLC=y when backing up the initial checkpoint

→Restore the backup using the following format as the file system name (Volume ID of rawfs file system):

```
/.celerra_vol1_<#>
```

→Perform the Incremental copy using the following:

```
$ fs_copy -start Eng_User_ckpt2:if=ace0 Eng_User:cel=DR_Site:if=s2_fsn2 -fromfs Eng_User_ckpt1 -Force -option monitor=off
```

→Startup the replication session using the following:

```
$ fs_replicate -start deptvie:if=VIECIFSSV0002 deptvie:cel=BRICS1:if=BRICIFSSV0001 -tape_copy
```

## **USING THE TAPE TRANSPORT METHOD FOR REPLICATION V1 (Tape Silvering):**

### **Prerequisites:**

→Must have a valid NDMP infrastructure on Source & Destination Data Mover

→Restore operation requires rawfs file system of same size as Source, mounted on Destination Server

1. Take checkpoint of production file system

```
$ fs_ckpt fs1 -Create
```

2. Set the NDMP environmental variable that applies to the backup software in use

VLC=y

3. Backup checkpoint to tape

4. Transport tape to destination

5. Create destination file system as rawfs and same size as Source

6. Determine the rawfs file system volume ID using nas\_fs -list

7. Perform NDMP restore of checkpoint using following syntax:

```
/.celerra_vol_<fs_volume_ID>
```

8. Start Replication between Source and Destination

```
$ fs_replicate -start src_fs1 dst_fs1:cel=cs110
```

9. Create 2<sup>nd</sup> Checkpoint of Source file system

```
$ fs_ckpt fs1 -Create  
10. Perform incremental copy using fs_copy and allow Destination system to convert from rawfs to ufs  
$ fs_copy -start src_fs1_ckpt2 dst_fs1:cel=cs110 -fromfs  
src_fs1_ckpt1 -Force -option monitor=off  
11. Verify using fs_replicate -list
```

**Need following param set on destination side:**

```
PAX.allowVLCRestoreToUFS 0x01152350 0x00000001 0x00000000
```

```
# server_param server_2 -facility PAX -info allowVLCRestoreToUFS -verbose
```

```
server_2 :
```

```
name      = allowVLCRestoreToUFS  
facility_name = PAX  
default_value = 0  
current_value = 0  
configured_value =  
user_action = reboot DataMover  
change_effective = reboot DataMover  
range     = (0,1)  
description = Whether to allow VLC backup to be restored to UFS file system  
detailed_description
```

Specifies whether to allow VLC backup to be restored to UFS file system. Set it to 1 to allow VLC backup to be restored to UFS file system

## **DETERMINING FILE SYSTEM MAGIC NUMBER:**

```
# .server_config server_2 -v "file verify ufs 96"
```

```
1227108012: STORAGE: 7: opening volume 96, ref 2, ioObj e7fef704  
LVol:referenceInternal(96) refCnt=2, nVol=1  
LVol:referenceInternal(95) refCnt=3, nVol=1  
LVol:referenceInternal(93) refCnt=4, nVol=1  
LVol:referenceInternal(86) refCnt=5, nVol=0  
1227108012: STORAGE: 7: closing volume 96, ref 3, ioObj e7fef704  
1227108012: UFS: 6: filesystem_magic_number:[0x17-0x4919e393-0x0-0xd33b]  
1227108012: UFS: 6: filesystem creation time Tue Nov 11 19:57:07 2008
```

## **REPLICATION TRAFFIC:**

--Uses TCP only, with default ports set as 8888

### **NOT SUPPORTED FOR IP REPLICATION:**

--MPFS (5.3 supports running MPFS on Source side, but not on Target side)

--SRDF

--No Timefinder/FS or SnapSure Restore to PFS/SFS

--One-to-many IP Replication is NOT supported

--Non-built-in LocalGroups are not supported if using LGDUP from (1) Server to another—User permissions will be affected. Only built-in or Domain-wide accounts can be moved over while retaining User or Group permissions on PFS to RO copy of SFS

## **CELLERRA GPO (GROUP POLICY OBJECTS): NAS 5.0, 5.1 & higher**

Support limited to Kerberos Maximum Clock Skew information. GPO Daemon to run [gpod]. On startup of CIFS, Server will read GPO from cache file [/etc/gpo.cache] and then retrieve latest GPO settings from AD using machine account--refreshed every 90 minutes. GPOs are updated on CIFS Startup, every 90 minutes, & manually using \$server\_security command.

### **Commands:**

```
$server_security server_2 -i -policy gpo [Querying GPO settings for Server]
```

```
$server_security server_2 -update -policy gpo [Updating settings]
```

```
$ .server_config server_2 -v "gpo update server=mmc"
```

```
$server_security server_2 -i -p gpo server=compname [Settings for a specific compname]
```

```
$ server_security server_6 -i -p gpo
```

```
server_6 :
```

Server: peter1

Domain: t2dom3.local

Kerberos Max Clock Skew (minutes): 5

Group Policy Refresh interval (minutes): 90

GPO Last Update Time: Wed Feb 26 22:51:38 EST 2003

### **COMMENT ABOUT GPO QUERY MECHANISMS ON DATA MOVER:**

1. When using server\_security –update –policy gpo CLI commands, DART uses a MAC subsystem to get updates
2. When using MMC Snapins, an XML interface is used to get the updates
3. GPO's have become more commonly used to enforce certain domain requirements, especially since release of W2K3

### **ACCESSING WINDOWS GPO EDITOR:**

Start>Run>gpedit.msc>Group Policy Object Editor>Computer Configuration>Administrative Templates>DNS Client

### **GPO POLICY ISSUES:**

NAS code has experienced memory leaks with GPO option and sometimes requires turning off GPO with “param cifs gpo=0”, usually related to large size of user rights credentials. Disabling GPO reverts DM to default Kerberos Clock Skew of (5) minutes vs. whatever the normal Domain Policy value is set to.

Specifically, there is a bug in NAS 5.1 in which we did not correctly check for Expired Kerberos Tickets & flush from cache, particularly if we were making an LDAP query to the GPO and an update was already in progress--CIFS process would hang. Function being called is the “krb5\_cc\_retrieve\_cred”. Event can happen in domains where max lifetime of service ticket is shorter than max lifetime of a user ticket. Resolution is to disable the GPO parameters by setting to 0 for gpo=0 & gpocache=0.

### **IMPACT OF DISABLING GPO SETTINGS ON DM:**

The impact of disabling GPO is that the Data Mover will use a default of 5 minutes for the Kerberos maximum Clock Skew setting, instead of the domain assigned value. Since 5 minutes is the default (and used by most domains), this should not be an issue. DART cannot currently make use of any other GPO value other than Kerberos clock synchronization.

#### **GPO Parameters:**

**param cifs gpo=1** [Enabled by default; gpo=0 for off]

**param cifs gpocache=1** [Enabled by default; gpocache=0 to turn off]

**param cifs gpo.messages=1** [Messaging is not enabled by default--messages=0 for off; Set to 1 to start logging]

**Note:** Windows 2000 Group Policies are applied to Objects at the Site, Domain, and OU Levels, and can be thought of as policies that are applied to Computers and Users during system startup or User LogOn

### **DELETING & REFRESHING DM GPO CACHE ONLINE:**

1. Disable GPO Cache

\$server\_config server\_2 -v “param cifs gpocache=0”

2. Delete GPO Cache

# mv /nas/quota/slot\_2/.etc/gpo.cache gpo.old

3. Restart GPO Cache

\$server\_config server\_2 -v “param cifs gpocache=1”

4. Update GPO Cache

\$ server\_security server\_2 –update –policy gpo

### **QUERYING EFFECTIVE GPO SETTINGS:**

**\$ .server\_config server\_2 -v "gpocache mview\_dm2 dump"**

Enable

Defined Privileges: SeBackup, SeRestore, NetworkLogon,

Everyone Privileges:

Authenticated users Privileges:

UNKNOWN S-1-5-20-22b.BUILTIN:S-1-5-20-22b

\* NetworkLogon,

1153831039: LIB: 5: nsswitch: Parsing file

1153831039: KERNEL: 5: >>919247 getting first hold on e915305c, nActive = 104084

1153831040: LIB: 5: nsswitch: Parsing line #

1153831040: LIB: 5: nsswitch: Parsing line # /etc/nsswitch.conf

1153831040: LIB: 5: nsswitch: Parsing line passwd: ldap files nis

-----abridged-----

1153831040: SMB: 7: Unix user 'tmatta' found with 'SECMAP:tmatta.2k3' UID=32769

1153831040: SMB: 6: FindUserId:Access\_Password 'tmatta',1=8001 T=0 (OK)

USER 0x8001 tmatta.2K3:S-1-5-15-bfc56af5-d6cf8701-5f67b1a3-46c

\* SeBackup, SeRestore,

## **CREATING GPO POLICY TO RENAME ADMINISTRATOR & GUEST ACCOUNTS:**

1. Admin Tools>ADUC>Properties>Group Policy>New><enter name for policy>: enter, close
2. Admin Tools>ADUC>Properties>Group Policy>Edit>expand Computer Configuration>Windows Settings>Security Settings>Local Policies>Security Options>doubleclick “Accounts: Rename administrator account”, check the “Define this policy” box, rename the account, click o.k.
3. Repeat above step to rename Guest account

## **EXPIRED KERBEROS TICKET PROBLEM: 5.1.9.4**

If DM uses expired Kerberos ticket, would panic—workaround is to set the following:

```
param cifs gpo=0
param cifs gpocache=0
param shadow stream=0
```

**Note:** May need to set above params in cases where DM continuously panics

## **ALTERNATE DATA STREAMS:**

- support introduced for up to 1000 streams per file
- no support for Streams on directories
- c:>more <testfile:stream[If alternate data streams exist, will see in output]
- Use the “lads.exe” tool to see Streams [www.heysoft.de/nt/ntfs-ads.htm]

## **Server Log Error:**

CFS:3: attempt to set a zero-length file stream name

Clients receive Semaphore timeout message

Error resulted after migration of files using EMCopy when a null stream is being sent to Celerra—locks all CIFS threads

## **Resolution is to Disable Streams Feature:**

```
param shadow stream=0
```

## **SNTP SUPPORT:**

- Implemented on DART as a client
- defacto ‘time’ standard for networks that must keep time synchronized

## **TIMESVC DEBUG LOGGING:**

```
"logsys set severity TIMESRVC=LOG_NOTICE"
"logsys set severity TIMESRVC=LOG_INFO"
"logsys set severity TIMESRVC=LOG_PRINTF"
```

“logsys set severity TIMESRVC=LOG\_PRINTF”

```
# .server_config server_2 -v "logsys get severity TIMESRVC"
```

(log level 4 is normal, non-debug mode)

1143254411: LIB: 4: Server log severity for facility TIMESRVC is 4

## **Testing SNTP Services:**

--tuna.us.dg.com/pub/dart/sntp: nptestsrv.pl—perlscript NTP test server utility

## **CIFS & NFS FTP:**

- Access to Celerra via GUI and CLI, respectively
- ftpd daemon runs by default on DM
- Time Out is 900 secs.
- For CIFS FTP, requires use of HomeDir Service running and configured on DMs

## **DART MAPPING RESOLUTION ORDER FOR USER/GROUP AUTHENTICATION:**

### **MAPPING RESOLUTION ORDER NAS 5.1 & BELOW:**

1. Global Data Mover Sid Cache (DART resolves Group SIDs and maps to GIDs)
2. Local Password & Group files (DART resolves by User/Group names and maps to UIDs/GIDs, respectively)
3. NIS Client Service (DART resolves by User/Group names and maps to UIDs/GIDs, respectively)
4. Active Directory Mapping Utility (DART resolves by User/Group SIDs and maps to UIDs/GIDs, respectively)
5. UserMapper Service (External-only) (DART resolves by User/Group names to UIDs/GIDs, respectively)

### **MAPPING RESOLUTION ORDER NAS 5.2 & ABOVE:**

1. SecMap Persistent Cache [/etc/secmap] (DART resolves by Group/User SIDs and maps to GIDs/UIDs, respectively)
2. Global Data Mover Sid Cache (DART resolves Group SIDs and maps to GIDs)
3. Local Password & Group files (DART resolves by User/Group names and maps to UIDs/GIDs, respectively)
4. NIS Client Service (DART resolves by User/Group names and maps to UIDs/GIDs, respectively)

5. Active Directory Mapping Utility (DART resolves by User/Group SIDs and maps to UIDs/GIDs, respectively)
6. UserMapper Service (External or Internal) (DART resolves by User/Group SIDs and maps to UIDs/GIDs, respectively for Internal UserMapper, and by Name-to-UID/GID prior to NAS 5.2)

**Note:** Do not use uppercase characters in passwd files or NIS database—use lowercase ASCII for Celerra

## **CELERRA DATA MOVER MANAGEMENT TOOLS:** NAS 5.0 and higher

CNM—Celerra Native Manager [New interface to manage & perform limited operations on multiple Celerra Cabinets]

Celerra Web Manager→GUI interface to manage & configure Celerra File Servers on a single Celerra

Antivirus Management Snap-in to manage virus checking parameters, stop & start service

### **Prerequisites to using Celerra Home Directory Management Snap-in:**

1. Must have CIFS up and running
2. Must export the Top Level Share for the Home Directories
3. Set permissions on directories after creating
4. Snap-in can be used to associate a User with a Home Directory

Audit Policy Snap-in—Apply audit logging for DataMover Security Events

User Rights Assignment Snap-in—Manage Users or Groups with logon rights or privileges on DataMover

--JRE 1.4.0 Celerra Monitor

--Usermapper 3.0.5 for Windows Systems

## **CELERRA APPLICATIONS & TOOLS CD:** (Apps & ToolsCD)

### **CONTENTS & VERSIONS:**

**Note :** First introduced with NAS 5.0

#### **NAS 5.4.14.3:**

USERMAPPER: Version 3.1.1 [Actual versions are 3.1.4 for Linux & Solaris, with version 3.1.1 for Windows NT/2000]

CELERRA CIFS MANAGEMENT MMC SNAPINS: Version 4.00.003, Windows 2000/XP, NAS 5.4.0.5135 or higher

JRE 1.4.2 FOR CELERRA MANAGER/MONITOR: Version 2.3, NAS 5.1.12.0 +, supports Windows NT/2000/XP,Linux,Solaris

## **CELERRA CIFS MANAGEMENT MMC CONSOLE SNAPINS:**

### **Select Components:**

--Dart Management

- \* Antivirus
- \* HomeDir
- \* Data Mover Security Management (Audit Policy & User Rights Assignment)

--UNIX User Management

**Note:** Unix User management requires following param: **param cifs useADMap=1**

Celerra Installations that do not plan to use the ADMap Tool should ensure that this value is set to 0. User access issues could result if the AD Schema has not been extended and this parameter is set.

## **WINDOWS SNAPINS AVAILABLE AFTER INSTALLATION OF CIFS MANAGEMENT TOOLS:**

→[Programs>Celerra Tools>About Celerra CIFS Management](#)>Version information about packages installed

### **CELERRA UNIX USER MANAGEMENT:**

→[Programs>Celerra Tools>Celerra UNIX Attributes Migration>Celerra UNIX Attributes Source Location](#) : Migrate

Users/Groups from : \* Nis server : [Enter Info in Boxes : server name/address Domain name]

o Unix password and group files : [passwd file Browse group file Browse]

>[Celerra UNIX Attributes Users/Groups Migration](#) : Users Tab Groups Tab [select Domain by checkbox to add accounts to]

→[Programs>Administrative Tools>Celerra Management>EMC Celerra Management>Celerra UNIX User](#)

[Management>Select local or remote Domain in which to store UNIX attributes>Users or Groups](#) [add users/groups manually or by browsing to file—also delete users/groups from this GUI]

### **PURPOSE:**

With installation of Celerra UNIX User Management program, extensions are added into Active Directory “schema” that Customers can now use to centralize and manage Celerra UNIX Users or Groups—designed for “mixed mode” environments. Use these tools to assign, remove, or modify UNIX attributes for Users, Groups on local or remote Domains. Use Local Domain if only one Domain present, trusts not allowed, and there is no need to centralize UNIX User information further. Use Remote Domain if you have multiple Domains, bi-directional trusts, and want to centralize UNIX User administration.

**Note:** After importing Celerra UNIX users and groups into AD, customer no longer needs any other UID/GID mapping service—no Usermapper, no NIS database service, no local passwd & group files on data mover!

## **MANAGING CELERRA USERS OR GROUPS USING AD:**

→Users/Groups can be migrated or imported into AD from a NIS Server’s database, from source passwd and group files, and also individually from the extensions added into Active Directory for Users and Computers>T2dom3.local>Users>boyo>Properties “Celerra UNIX User” tab>Delete, Modify, or Add UID & domain GID mapping for User manually, or by browsing to NIS server or passwd file. Users/Groups can be added or deleted individually from Celerra Management>EMC Celerra Management>Celerra UNIX User Management>Users or Groups  
→Users/Groups can be deleted from the Celerra UNIX Management via Active Directory Users and Computers>T2dom3.local>EMC Celerra>User Mapper>Groups or Users>highlight users or groups and delete

## **CELERRA MANAGEMENT SNAP-IN TOOLS:**

### **Unix User Management:**

**Note:** **param cifs useADMap=1** [Requires param to work with AD]

### **Dart Management:**

- AntiVirus Management
- Home Directory Management
- Security Management

## **USER RIGHTS ASSIGNMENT CONSOLE:**

Programs>Administrative Tools>Celerra Management

→Celerra UNIX User Management

→DART Management: Antivirus; HomeDir; DART Security Settings (Audit Policy; User Rights Assignments)

### **Purpose:**

--Allows for management of Users & Groups’ rights on Celerra

**Note:** These rights cannot be managed via the Local Policy Setting GUI as is done on native Windows 2000 Servers

## **UNIX USER MANAGEMENT CONSOLE:**

### **Purpose:**

--Networks using only Windows Users will map SID’s to UNIX UID/GID’s for file system access

--Networks using mixed-mode Windows & Unix access will map to UNIX UID/GID’s

**Comment:** When Windows 2000 Users connect to Celerra, the DART queries the AD Server to see if User is assigned a UID/GID. It also queries Active Directory to determine what Windows Groups are mapped to UNIX GID’s

### **Tools: /nas/cifs/CelerraUnixUserMgmt.exe**

CIFS Migration Tool—Migrate UID/GID’s of Windows Users & Groups to Active Directory from NIS or Passwd Files  
Tool is run and Users & Groups are listed in tree format—an Administrator can then select which mappings get added to AD  
UNIX User Tool—Adds extension to Active Directory for Users so that UID/GID’s can be specified [local domain only]  
DART UNIX User Management—Manage Users & Groups from Local and Trusted NT Domains

## **I. INSTALLING CELERRA “UNIX USER MGMT” MMC:** Install on Windows 2000 Server

1. Installshield ‘Welcome to the installshield for CelerraUnixUserMgmt V1.01.002
2. License agreement Yes
3. Start Copying Files (Default Installation): C:\Program Files\EMC\CelerraUnixUserMgmt V1.01.002
4. Installshield wizard complete

**Note:** Creates MMC snap-in to Active Directory by adding a new tab called “Celerra UNIX User” under the properties for any ‘User’.

## **II. USING UNIX USER MIGRATION TOOL:**

1. Set following parameter for DataMover to enable Migration tool: NAS 4.2.5.2/NAS 5.0.9.2

**param cifs useADMap=1**

2. Start>Programs>Celerra CIFS Management>Celerra Unix User Migration

a.) Celerra CIFS Migration Tool

b.) Migrate Users/Groups from:

\* NIS domain: \_\_\_\_\_

o Unix passwd & group file

    Passwd File: [browse to select file]

    Group File: [browse to select file]

c.) The local domain’s Active Directory does not contain an attribute database for Celerra UNIX User management. Do you wish to initialize the local database? Yes No

d.) Select “local” on remote domain to store user accounts

    Unix User Domain

    \* Store UNIX user attributes in local domain. Share local attribute database with domains

    \* Store UNIX user attributes in a remote domain: [dropdown bar to select]

e.) Celerra CIFS Users and Groups [New Containers created under OU>EMC Celerra>User Mapper]

|                                                                                 |                  |
|---------------------------------------------------------------------------------|------------------|
| <b>Users</b>                                                                    | <b>Group</b>     |
| _x_NT_BIW_MASTER                                                                | _x_NT_BIW_MASTER |
| f.) ADUC>EMC Celerra>UserMapper>Groups/Users [List of Groups or Users migrated] |                  |
| g.) Also creates entry for CIFS output:                                         |                  |

**Active Directory usermapper's domain: “not yet located”**

**Active directory usermapper's domain: "not found"**

**Note:** Running the ‘Migration Tool’ creates a new ‘OU’ in Active Directory under ‘EMC Celerra’ called ‘User Mapper’, with (2) new Containers called “GROUPS” & “USERS”. Within these new containers are the Users and Groups that were migrated into AD, from either the NIS DB or other Passwd & Group files. Properties of Users or Groups in these ‘containers’ contain the following information: Domain Name; SID; UID or GID; and description—‘emc-com-usrMapperUser’.

**Purpose:** Allows for integrated mapping of UID/GIDs to DM on a per Domain basis in AD

**Prerequisites:** Run tool on AD Server at ‘Root/Parent Domain’ as User with Domain Admins and Schema Admins membership. If using NIS, NIS domain must be in same broadcast domain as AD Server—uses Ypbnd service. If using tool across multiple Domains or Trusted Forests, the Unix User Migration Tool needs to be installed in each Domain.

#### **EXAMPLE OF ACTIVE DIRECTORY TOOL IN PLACE:**

CIFS Server (Default) REX[T2DOM3]

Full computer name=rex.t2dom3.local realm=T2DOM3.LOCAL

**Active directory usermapper's domain: t2dom3.local**

Comment='EMC-SNAS:T5.1.9.4'

if=ana0 l=192.10.3.28 b=192.10.3.255 mac=0:6:2b:3:6a:65

FQDN=rex.t2dom3.local (Updated to DNS)

### **III. ADMINISTERING UNIX USERS:**

1. Start>Programs>Administrative Tools>Celerra Management

Celerra Management Tool>Celerra Management>Celerra UNIX User Management>\*NT\_BIW\_MASTER>Users/Groups>Properties

#### **UNIX User Attributes:**

**Username:** jcarr

**UID:** 33724

**GID:** 30000

**Note:** NAS 5.4 documentation says that you can only ‘manage’ these properties locally in the domain where they are located

### **NAS FEATURES & GA CODE:**

**Note:** Most information on NAS versions prior to 5.1 is archived

### **NAS 5.0: NOTABLE NEW FEATURES:**

#### **META FILESYSTEM SUPPORT:**

Introduces minimal support for "meta" FileSystems up to 32 TB—that is, taking multiple smaller filesystems & exported as (1) Meta to Clients [Purpose is to support easier FSCK maintenance & extends, which could be done on the individual smaller FileSystems]; FS will be comprised of 'cells' with a primary cell containing directory structure; User sees only (1) file system

#### **VirusChecking MPFS:**

Virus Checking support only on MPFS clients, not the DM, for HighRoad configurations

#### **CELLERRA IP REPLICATION:** Introduced 5.0 with RPQ

#### **DART NFS CLUSTERS:**

--Provides for Read/Write Sharing of FS to multiple Hosts

--Provides for Load Balancing and Bandwidth on demand

--Control Station Manages the “cluster”

#### **iSCSI SUPPORT:**

Block Services over IP as a replacement for direct-connect SCSI (no sharing or encryption), but is complex and requires more ‘Standards’ be developed; will facilitate HighRoad deployments; Gigabit-Ethernet Network will share data (SCSI over IP) and Metadata (MPFS) traffic, eliminating costly FC cards for each Server; Subject to ECC Management Control;

Support for Exchange Server using CIFS or iSCSI v. Direct-Attached Storage

Support for Snapsure and MPFS

#### **PERSISTENT BLOCK RESERVATIONS (PBR):** NAS 5.0.11.2

→The Persistent Block Reservation counter stores “reserved” block values in the Superblock structure

→During copy operations, the Data Mover will reserve the blocks required to complete the write, and also allows the system to recover after a Server reboot, since the PBR’s remain in the Superblock, until the blocks transition to “used”

→A file system can have Used, Free, or Reserved blocks

→A PBR is a per file system counter that reserves space for dense files, such that any write from offset 0 to EOF (End of File) will succeed even if there is no space left on the disk (i.e., uses its reserved space)

→The PBR feature was introduced to allow CIFS Applications to reserve space for files so that out-of-space errors do not occur [Unix systems use sparse-files by default while CIFS does not—this feature will allow CIFS to use sparse file semantics by reserving blocks of space for files to be written to]

### **PBR-SPARSE FILE SUPPORT TURNED OFF BY DEFAULT:** To Enable...

**param cifs createDenseFiles=1**

**param ufs createDenseFiles=1**

**Note:** Migration of sparse files supported by rsync

### **NAS 5.1 SUPPORT:** EOSL Aug 2005

--IPDART NS600 now supported

--CNM, WEB UI/Monitor, Secure NFS by RPQ only

### **DETERMINING BACKEND STORAGE SYSTEMS FOR NAS 5.1:**

**\$nas\_symm -l →nas\_symm -i serial#**

**\$nas\_storage -l**

### **NAS 5.1 FEATURES:**

#### **CONTROL STATION HTTPS & X509 SSL CERTIFICATES:**

HTTPS uses HTTP over SSL for secure communication over the network. For Celerra Control Stations, the expectation is that private internal customer networks are in use, where the use of public X509 certificates (Verisign, etc) are not required, and that self-signed certificates are adequate.

→CS uses HTTP over SSL to protect traffic using self-signed X509 certificates

→CS uses Apache, OpenSSL, and mod\_ssl to handle HTTPS traffic

**/nas/http/conf**

**/nas/http/conf/httpd.conf** (Apache config file) | **ssl.crt** (certificate) | **ssl.key** (RSA private key)

**/nas/sbin/nas\_config -ssl** [Use to create new certificate]

#### **CREATING NEW SSL CERTIFICATE ON CS:**

1. **# /nas/sbin/nas\_config -ssl**

Installing a new SSL certificate requires restarting the Apache web server.

Do you want to proceed? [y/n]: y

New SSL certificate has been generated and installed successfully.

**Note:** Script calls following perl script to actually change certificate and restart Apache Server

**/nas/http/nas\_ezadm/bin/gen\_ssl\_cert.pl**

2. Reboot CS—only sure way of getting Apache properly restarted

**Note:** New installs or Control Station replacements may require regeneration of Security Certificate

### **CELLERRA AVM (Automatic Volume Management):** Introduced NAS 5.1, ATA Support 5.1.18.3

AVM is a volume creation & management tool for ease of use when creating file systems and to provide performance best practices

**Note:** AVM is invoked when using Celerra Manager, or from the CLI if specified

**# nas\_fs -n fs04 -c meta04 size=<integer>[G|M] profile=profile**

#### **AVM & STORAGE TEMPLATES:**

--storage templates are not part of AVM, but work with them

--storage templates consist of definition files for creating LUNs from specific RAID types, and scripts which create RAID groups and bind LUNs

--storage templates are not used with SAN Gateway Arrays, only with Integrated Arrays

--storage templates for fibre channel disks create 2 LUNs per RG, one owned by SPA and the other SPB

--storage templates for ATA drives create 2 LUNs per RG, both owned by either SPA or SPB, supporting only RAID-5

### **CELLERRA STORAGE POOLS:**

Containers that hold storage for use by file systems, checkpoints, & other Celerra objects, for both system-defined & user-defined pools

#### **SYSTEM-DEFINED STORAGE POOLS:**

**symm\_std**

Highest performance and availability using Symmetrix STD disk volumes in RAID 1 configurations

**symm\_ata**

Highest performance at low cost using Symm ATA disks in RAID 1 configurations

**symm\_std\_rdf\_src**

Highest performance and availability using storage mirrored for SRDF sites or locally using TimeFinder/FS

**symm\_std\_rdf\_tgt**

Highest performance and availability using storage mirrored for SRDF sites or locally using TimeFinder/FS on R2STD disk volumes

**symm\_ata\_rdf\_src**

Highest performance and availability at low cost using mirrored storage to SRDF sites or locally using TimeFinder/FS on R1ATA

**symm\_ata\_rdf\_src**

Highest performance and availability at low cost using mirrored storage to SRDF sites or locally using TimeFinder/FS on R2ATA

**symm\_ssd**

High performance uses in 3+1 or 7+1 Raid5 configurations [DMX4 5773/5.6.37.x], not supported with TimeFinder/FS or SRDF

**clar\_r1**

High performance & low cost volumes using Clariion CLSTD from RAID 1 mirrored-pair disk groups

**clar\_r6**

High availability at low cost using CLARiiON CLSTD disks from 4+2, 6+2, or 12+2 RAID 6 disk groups

**clar\_r5\_performance**

Medium performance at low cost using CLSTD disk volumes from 4+1 RAID 5 disk groups

**clar\_r5\_economy**

Medium performance lowest cost using CLSTD disk volumes from 8+1 RAID 5 disk groups

**clarata\_archive**

Archival performance lowest cost using CLATA drives in RAID 5 configuration

**clarata\_r3**

Archival performance lowest close using LCFC, SATA II, or CLATA disk drives in RAID 3 configurations

**clarata\_r6**

High availability at low cost using CLATA disk volumes from 4+2, 6+2, or 12+2 RAID 6 disk groups

**cm\_r1**

High performance & availability at low cost. Uses CLARiiON CMSTD disk volumes from R5 4+1 for MView/S

**cm\_r5\_performance**

Medium performance & availability low cost. Uses CLARiiON CMSTD disk volumes from R5 4+1 MView/S

**cm\_r5\_economy**

Medium performance & availability lowest cost. Uses CLARiiON CMSTD disk volumes from R5 8+1 MView/S.

**cm\_r6**

High availability low cost using CMSTD disk volumes from 4+2, 6+2, or 12+2 RAID 6 disk groups for MView

**cmata\_r6**

High availability low cost using CMATA disk volumes from 4+2, 6+2, or 12+2 RAID 6 disk groups for MView

**cmata\_archive**

Archival retrieval only, using CLARiiON ATA drives in R5 for MView/S

**cmata\_r3**

Archival retrieval only at lowest cost, using CLARiiON ATA drives in R3 for MView/S

**clar\_ssd\_r5**

High performance using RAID 5 4+1 or 8+1 with (2) luns

**cmssd\_r5**

High performance MView volumes using RAID 5 4+1 or 8+1

**User-Defined Storage Pools:**

Can create User-defined pools from Celerra Manager, but are based on single disk type

**DISK TYPES:**

CLSTD—standard CLARiiON disk volumes

CLATA—CLARiiON ATA volumes

STD—standard Symmetrix disk volumes

R1STD—Symmetrix FC disk volumes used for mirroring storage as in SRDF, etc.

R2STD—“

ATA—standard symmetrix disk volumes built on SATA drives

R1ATA—Symmetrix SATA disk volumes for mirroring storage as in SRDF, etc.

R2ATA—“

**TWO TYPES AVM PROFILES:**

Volume Profiles—defines how disk volumes are aggregated and placed into a pool, applies only to system-defined pools

Storage Profiles—description of disk volumes, used by a volume profile, matches disks to a certain expected disk type, system-defined pools

**AVM POOLS:**

- Containers for Volumes with similar characteristics [can consist of disks, stripes, slices, metavolumes]
- Association between pools and profiles
- A volume can only be a member of a single pool
- Pool members must be of the same device type

**System-defined Storage Pools vs. User-defined Storage Pools**

|                                     |                                                         |
|-------------------------------------|---------------------------------------------------------|
| Volume Profiles predefined          | No Volume Profile                                       |
| Dynamic, but can be disabled        | Not dynamic, user must create and populate storage pool |
| Storage added/removed automatically | Storage added or removed manually                       |

**Comment:** NAS 5.2.9.6 calls the “profile” a “pool”: pool=<pool> and the pool options are the same as with NAS 5.1. System-defined pools use default 8k stripe size

**ABOUT AVM:**

- Automatic creation and deletion of volumes during file system operations—rule set for defining a “pool of storage”
- Pools are made from profile disk volumes
- Requires creation of rule set called a “profile” that defines a “pool of storage” in order for WEB UI to be able to administer FS
- AVM is invoked for either Symmetrix or Clariion BackEnds by Web UI, by CheckPoints, and by CLI when creating FS using “size” and “profile” parameters
- Clariion uses disk drives in pairs for “profiles”—one for each SP
- Disk Volume profiles, RAID type, and spindle counts must match

**Note:** AVM automatically claims volumes—in order to keep a reserve of volumes, might need to create stripes/metis via CLI

**AVM BACKEND ENHANCEMENTS NAS 5.4:**

- More flexibility in creating Raid Groups and configurations
  - Setup\_clariion script for user-defined templates, used only for Integrated Systems
  - Mixed shelves of different RAID configurations [such as RAID1 for Logs and RAID5 for tablespace for Oracle solutions]
  - Script will prompt for template to use when adding new enclosure
  - New RAID 3 configuration for Backup-to-Disk (B2D) ATA solutions [clarata\_r3]
  - AVM to prevent creation of file system that spans multiple arrays
- Note:** GUI will warn, CLI does not
- Extend System pool by size
  - Extend User pool by volumes
  - No mix of RAID 3 & RAID 5 on same shelf

**CREATING FILE SYSTEM FROM WEBUI USING AVM PROFILE:**

\$ nas\_fs -n fs09 -c size=32 profile=clar\_r5\_economy -o slice=y, mover=server\_2 [32=32GB]

**Note:** Profiles are a set of rules to aggregate storage

# nas\_fs -name cdms -type uxfs -create size=50000M pool=clar\_r5\_performance storage=SINGLE  
worm=off -option slice=y,mover=server\_2 (NAS 5.4)

**CREATING FILE SYSTEM FROM CLI USING AVM PROFILE:**

\$ nas\_fs -n fs10 -c size=10G profile=symm\_std -o slice=y, mover=server\_2

**Note:** Can also create Checkpoints using the AVM rules set. Do not need to specify “slice=y”—this will create a file system by striping (4) d volumes into a meta. When using “slice=y”, AVM creates stripes first, then metas, then slices, metas again, then creates the File System.

# nas\_fs -name migrate -create size=1000M pool=clar\_r5\_performance -o slice=y

```
id      = 63
name    = migrate
acl     = 0
in_use  = False
type    = uxfs
worm   = off
volume  = v203
pool    = clar_r5_performance
member_of = root_avm_fs_group_3
rw_servers=
ro_servers=
rw_vdms =
ro_vdms =
stor_devs = APM00040303779-0011
disks   = d8
```

## # nas\_fs -s migrate

total = 1000 (sizes in MB) ( blockcount = 2048000 )

## WEB UI LIMITATIONS:

- cannot extend filesystems
- cannot setup 32-bit GIDs
- cannot setup Quotas
- cannot export with “anon=0” option
- when setting up default network route, must use 0.0.0.0 in “Destination” field!
- File Systems created from CLI cannot be managed with this interface [Requires AVM profile ]

## TROUBLESHOOTING HTTPD FOR WEB UI:

1. Grep for process: **#ps -awlx |grep httpd** [Shows complete path of service]
  2. Kill processes: #pkill -9 12543 [Should kill all httpd processes]
  3. Check to see if CS is listening on Port 8000: #netstat -alp
- | PID/Program Name          | State |
|---------------------------|-------|
| tcp 0 0 *:8000 *:* LISTEN |       |
4. Restarting HTTPD daemon: **#/nas/sbin/httpd -D HAVE\_PERL -D HAVE\_SSL -f /nas/http/conf/httpd.conf** [NAS 5.1.9.4]

## CELERRA SECURE NFS IN NAS 5.1:

**Overview:** Secure NFS provides Kerberos-based User and Data authorization & is designed to provide strong authentication using secret-key crypto. The service allows a Client process to prove its identity without sending User information across the network. It further assures the identity of sender and recipient. Clients must run Sun Solaris 5.8 or Hummingbird Version 7 to support Kerberos. Celerra uses RPCSEC\_GSS & Kerberos protocols and runs only with NFS over TCP.

## NFS CLIENT SUPPORT FOR SECURE NFS:

- Supported by SUN Solaris 8, 9, 10 using SEAM v1.01 [SUN Enterprise Authentication Mechanism]
- Supported by Hummingbird Maestro v7.2

## OTHER INFO:

- RPQ Only for NAS 5.1
- SecMap caching can be used with SecureNFS, beginning with NAS 5.2
- NFS User and data authentication via RPC calls using Kerberos 5 RPCSEC\_GSS framework and NFSv3 over TCP/IP
- Security services are AUTH\_SYS (UID/GID) & RPCSEC\_GSS (API)—sys and krb5 for SecureNFS
- Single Kerberos REALM only for Secure NFS
- KDC Server runs (2) Services and contains the encrypted database of principals and keys
- AUTH\_SYS or AUTH\_UNIX—authentication via UID/GID [Server assumes this info passed to it is correct]
- AUTH\_DH—DES key exchange
- AUTH\_KERB4—Kerberos version 4 DES key exchange
- RPCSEC\_GSS—Generic Security Services API—encryption of user credentials and data

## Keys:

- User Key is 1-way hashing of User password
- Service key from random or known password
- “Principal” is any User or Service that is using the KDC Service [username@t2dom2.com; nfs/server,t2dom2.com@t2dom2.com]

## How Does Secure NFS Session Work?

1. Client contacts KDC for session ticket
2. KDC replies with two messages encrypting first with User’s Key & second with NFS Server’s Key
3. Client decrypts and stores Session Key and replies to Server with new message verifying identity
4. Server stores Session Key and validates Client’s identity

## Security Options Available:

- sys [Default implementation--no authentication or encryption performed by Server—Client supplies credentials]
- krb5 [Server requires authentication through Kerberos—EMC’s implementation of SecureNFS]

## Unsupported:

- krb5i—RPC Data Checksum header ensures data integrity
- krb5p—Encrypting data before sending to Client to ensure data integrity
- Windows 2000 Kerberos Server NOT supported for SecureNFS

## CONFIGURING SOLARIS 2.8 FOR KDC SERVICES:

- Install SEAM v1.01 using custom install option
  - Install KRB5 encryption package [Solaris uses KRB4 by default]
  - Configure DNS & NTP
1. Setup the /etc/krb5/krb5.conf file with Principal Hosts & REALM name
    - /usr/krb5/sbin/kdb5\_util create -r HOSTS.PVT.DNS -s [libdefaults]

```
default_realm=HOSTS.PVT.DNS
[realms]
HOSTS.PVT.DNS = {
    kdc=sun1.hosts.pvt.dns
    kdc=sun1.hosts.pvt.dns
    admin_server=sun1.hosts.pvt.dns
    default_domain=hosts.pvt.dns  }
[domain_realm]
.hosts.pvt.dns=HOSTS.PVT.DNS
hosts.pvt.dns=HOSTS.PVT.DNS
```

### **Windows 2000 Note:**

SecureNFS requires setting up of Sun KDC Realm before setting up Windows 2000 CFS Server

2. Edit /etc/krb5/kadm5.acl file

**#\*/admin@HOSTS.PVT.DNS\*** [comment out so that all Admins do not have KDC privileges—first \* indicates any Admin]  
**szg30/admin@HOSTS.PVT.DNS\*** [Last \* indicates All Permissions]

3. Add Passwd for Kerberos Administrator to Keytab file: /usr/krb5/sbin/kadmin.local

4. Add Principals using “ktadd system1” or “ktrem system2”

(Use ‘kadmin’ utility to create keytab entries for all NFS servers)

#kadmin -k krb5.keytab nfs/dm3.pvt.dns

5. Start KDC daemon: /etc/init.d/kdc start

/etc/init.d/kdc.master start

6. Create gsscred\_db database for User Credentials: Maps UID/GID to Users

#gsscred -m kerberos\_v5 -a [creates db using all Users from /etc/passwd file]

#gsscred -m kerberos\_v5 -u 71 -r [Remove UIDs of System Users, etc from db]

#gsscred -m kerberos\_v5 -n root/yucca.pvt.nds -u 0 -a [Create gsscred for root of each host]

#gsscred -l

7. Reviewing NFS or ROOT Principals;

#kadmin: listprincs nfs\* llistprincs root\*

### **Key Directories & Files:**

/etc/krb5            /var/krb5

kpropd.acl [Key propagate ACLs for KDCs]

kdc.conf file: kdc\_ports = 88, 750

#kutil: read\_kt /etc/krb5/kadm5.keytab [Read Keytab files]

#kutil: read\_kt /etc/krb5/krb5.keytab [Principals with Keytab entries]

### **TroubleShooting SecureNFS:**

→Increase DM logging, Verify NTP service, Verify DNS service, Verify principals & keytab entries, Verify gsscred\_db, take snoop traces, Verify that Realm name matches KDC name

#kdestroy [destroys logged in User’s KDC ticket]

#kinit [regenerates new KDC ticket for user]

#klist [Ticket cache]

\$server\_config server\_3 -v “gsscred list” |grep quota3 or \$server\_config server\_3 -v “gsscred -l”

\$server\_config server\_3 -v “rpcgssctl statserv prog=100003 vers=3 all”

\$server\_config server\_3 -v “rpcgssctl showclnt prog=100003 vers=3 handle=2”

### **USING GSSCRED:**

**Note:** Creates mapping between security principal name and local unix User ID

**# gsscred -m kerberos\_v5 -a** [Creates credentials for all users located in /etc/passwd or NIS DB]

# gsscred -m kerberos\_v5 -n root/dm3.realm.com -u 0 -a [Adding credentials for root user of Client]

# gsscred -m kerberos\_v5 -n username -u 1400 -r [Removing credentials for a User]

### **CONFIGURING SECURENFS ON CELERRA:**

1. Data Mover name must match DNS Name

2. Setup NTP Service on DM

3. Setup NIS Service on DM [or use local Passwd/Group files]

4. Setup DNS Service on DM

5. Add Kerberos Realm to DM:

\$server\_kerberos server\_2 -add

realm=t2dom2.com,kdc=sun.t2dom2.com,kadmin=sun.t2dom2.com, domain=t2dom2.com, defaultrealm

6. Add Service Principal Names to Control Station:

```
$server_kerberos server_2 -kadmin -p root/admin@sun.t2dom2.com
7. Generate Keytab File on Sun Server: #ktadd nfs/server_3.t2dom2.com
8. Generate Kerberos Principal to UID Map: gsscred -a -m [Adds entries to NIS table]
$fncreate -t org -o org/
$vi /etc/gss/gsscred.conf [Add line: xfn_nis]
$gsscred -m kerberos-v5 -a
$gsscred -n krbusr1 -u 5000 -c "kerberos user" -m kerberos_v5 -a
9. Upload Keytab and Gsscred Map: $server_kerberos server_3 -p krb5.keytab krb5.keytab lgsscred_db gsscred_db
10. Export File System for SecureNFS: $server_export server_3 -o sec=krb5:root=clients,sec=sys:ro /fs211. Remote Mount to Solaris: #mount -f nfs -o sec=krb5 server_3:/fs2 /mnt
12. Access Filesystem as valid user $klist
Note: Minimum files on DM: ./etc/gsscred_db ./etc/krb5.keytab
```

### **KRB5 SECURITY OPTIONS FOR SERVER EXPORT:**

ro, rw=, ro=, root=, access=

### **TROUBLESHOOTING SECURE NFS ISSUES:**

- Increase Server Logging
- Verify NTP & DNS setup
- Verify principals, gsscred\_db, and keytab entries
- Verify realm names match KDC

#### **DNS Domain Configured Incorrectly:**

RPC: 3: RpcGssServer::acquireCred:gss\_import\_name ([name=nfs@dm2,type=13089288](#)) failed  
RPC: 3: GSS-API major error: An invalid name was supplied  
RPC: 3: GSS-API minor error: Configuration file does not specify default realm

#### **Wrong Principal in Request:**

RPC: 3: RpcGssClient::accept:gss\_accept\_sec\_context failed  
RPC: 3: GSS-API major error: Miscellaneous failure  
RPC: 3: GSS-API minor: Wrong principal in request

#### **Uname Problem:**

User authenticates to NFS Server but new files do not contain UID/GID

Solution: Update gsscred database or NIS/Passwd files

SECURITY: 3: Access\_GsscredDatabase::interpret: No uname found for uid 5000

#### **Permission Denied:**

Expired Kerberos credentials; NTP Time is off; DNS Entries not setup properly; Kerberos not setup correctly

#### **Kadmind Daemon Not Running:**

#/etc/init.d/kdc.master start

#### **SUN Tools:**

\$kutil

\$klist; kinit; kdestroy

\$/usr/krb5/sbin/kadmin

### **NAS 5.1 LIMITATIONS:**

- SRDF not supported for NS600 BackEnd
- TimeFinder/FS or R2 Backup not supported for NS600
- HighRoad enhancements for CAVA & Snapsure not supported for NS600

#### **LOG CHANGES & CALLHOMES:**

```
/nas/log/nas_rdf.log [New Log]
/nas/log/ERRSFILE.LOG & ERRSFILE.LOG.old [Call Home Files for NS600]
#cat /nas/log/CH*
/nas/var/log/log_config_unknown_021201040012.gz
```

### **NAS 5.1M BEST PRACTICES & PERFORMANCE TUNING FOR CLARIION BACKENDS:**

- When possible do not extend File Systems on same spindles
- Set Clariion stripe size to 8kb, Set Symmetrix NFS stripe size to 32kb and CIFS stripe size to 8kb, HighRoad to 256kb
- Do not stripe file systems across two LUNs from same RAID group
- Stripe multiple DAE2 file systems across LUNs of same RAID type and configuration
- Never build file systems directly on Symm metavolumes or stripes, always use Celerra Stripe volumes
- HighRoad and NFS workloads, use 16 volumes in each stripe set
- Ensure two LUNs are bound for each RAID group

--RAID 1 has better performance than RAID 5, though optimal RAID 5 performance is using 4+1 R5 LUNs

--Increase NFS threads in heavy workload environments from 192 to 512

#vi /nas/server/slot\_x/netd

### **nfs start openfiles=44032 nfsd=512**

--Write Intensive Environments, disable Level 2 Oblocks

### **param cifs capabilities=0xA379**

--Configure Read Cache after code upgrades using:

#/nas/sbin/setup\_backend & #nas\_raid cache configure [32-256MB]

### **2GB Memory per SP:**

Recommended Write Cache allocation—1465MB

Recommended Read Cache allocation—32MB

### **4GB Memory per SP:**

Recommended Write Cache allocation—3072MB

Recommended Read Cache allocation—146MB

--Bind LUNs with following parameters to maximize cache

-rc 1 [enables read cache on LUNs]

-wc 1 [enables write cache on LUNs] -wc 0 [disables write cache]

-aa [adjusts auto assignment value, used in failover; -aa 0 for NS600 & -aa 1 for Celerra/CX600]

--Tune CIFS threads if required

--Enable preallocation for CIFS to preallocate space before writing to new files [**param cifs prealloc=6**]

--Tune asynchronous I/O on Clarion

**param cfs closeDelay=30**

**param cfs deleteDelay=1**

**param cfs maxPendingClose=127000**

**param file secureReadSide=1**

--Solaris Client Mounting for NFS Operations using NFS v3 with TCP

**\$mount -o vers=3,proto=tcp,rsize=32768,wszie=32768 192.10.0.111:/fs01 /mnt**

--Solaris 2.7/2.8 Tuning [Add to /etc/system and reboot]

**set nfs:nfs3\_nra=8**

**set nfs:nfs3\_max\_threads=32**

--Adjust write flush parameters on Solaris

**\$echo "flush\_freq/Z 0" |adb -kw [Turns off flushing for better write performance]**

**\$echo "flush\_freq/Z 800000" |adb -kw [Optimal setting for most environments]**

**\$echo "flush\_freq/Z 1000000" |adb -kw [Sets flush frequency at 16M]**

--Linux Client Mounting for NFS v3 using TCP

**\$mount -o nfsvers=3,tcp,rsize=32768,wszie=32768 192.10.0.111:/fs01 /mnt**

--Linux Client Tuning for Read-Ahead threads

**\$echo 8> /proc/sys/vm/min-readahead**

**\$echo 256> /proc/sys/vm/max-readahead**

## **WINDOWS 2003 SUPPORT LIMITATIONS—see Primus emc77391:**

**SMB Signing:** Problem with Windows 2003 & SMB Signing involved with use of Kerberos tickets, frequently seen as failures to ‘Join’ the domain. Only NAS 5.1.18.8, with mods referenced here, can interoperate with Win2K3 DC’s for Windows 2003 domains.

### **NAS 5.2 provides native support for SMB Signing:**

→Disable GPO Settings [Security Settings>Security Options]

→Digitally sign communications (always) [disable]

→Digitally sign communications (if client agrees) [disable]

**Group Policy Objects:** Disable in Windows 2003 domains

→Add following param to data movers: **param cifs gpo=0**

**Dynamic DNS:** Authenticated Updates are not supported in Windows 2003

→Disable Dynamic DNS or Enable Non-Authenticated DNS Updates

**param dns updateMode=0** [If the choice is to disable Dynamic DNS]

DNS Zone>Properties>General Tab>Change “Allow Dynamic Updates?” to “Yes” for non-authenticated updates

### **DNS With Top Level Zones:**

Top Level zones are used if the Windows Domain is a single word domain, such as in “EMC.COM”. In Domains of this type, Celerra must have GPO enabled to allow for updates to the Top Level Domain Zones:

GPO Editor>Administrative Templates>Network>DNS Client>Update Top Level Domain Zones

## **NS600 PEER POWER CONTROL: NAS 5.1.10.0 +**

**Support:** NAS 5.1.4.0; Hidden CS utilities added 5.1.6.0; Complete CS reset commands added 5.1.10.0

### **Purpose:**

Cycle, shutdown, or restore power to panicked or hung DM via a peer datamover when serial or network connections won't work.

Allows for reset and powerdown of DM's connected to NS600 from CS

Allows Control Station to issue reset commands & obtain getreason for DMs over Serial path if network path not working

**Note:** Traditional Celerra Cabinets have more hardware controls built-in, so this will solve the problems for NS600. NS600 has only a single Network & Serial connection to DM

### **Prerequisites for Using Peer Power Control:**

NS600 DM boards that support this feature—Calandor2 boards [Peer Power Control is using additional hardware & SW to perform]

Peer DM must be at reason code 4 or 5

Requires use of either Serial [DART Console] or Network [XML] connection to DM

Underlying connection is made using Telnet, therefore make sure that DNS is correctly setup in /etc/resolv.conf

**Note:** HTTP XML MAC interface is replacing the traditional RPC Client-Server protocol

## **SERIAL LINE COMMUNICATION BETWEEN CS & DM:**

**\$ /nas/tools/.server\_tty -c 2 “peer\_powerctrl action=reset”**

**Note:** .server\_tty are an older set of commands, no longer supported with 5.5

**\$ /nbsnas/tools/.server\_peer\_powerctrl**

**Note:** Current method of running Peer Power Control Commands

**XML Command Interfaces:** Conducted over HTTP—don't know how to invoke this interface with NAS 5.5

checkPowerctrl [Verifies whether peer power control hardware is present]

isPeerDMInserted [DM inserted]

rebootPeer [Cycles power to peer data mover]

shutdownPeer [Shutdowns power to DM]

restorePeer [Restores power to DM]

### **Typical Peer Power Management Configuration Commands:**

**\$server\_cpu server\_x -reboot | -halt now**

**\$t2reset pwroff | pwrone | hldrst -s 2 | rlsrst**

recover\_slot

install scripts

**Note:** CS processes peer power control commands over XML network interface or serial DART Console. A reset command first performs a network shutdown over the mac\_shutdown interface, then issues CS code to shutdown hardware

## **USING PEER POWER CONTROL COMMANDS:**

**#/nbsnas/tools/.server\_peer\_powerctrl**

.server\_peer\_powerctrl <-chkhw | -chkins | -reboot | -shutdown | -restore> <slot>

**Note:** Commands are not useable on NSX systems

**/nas/tools/.server\_peer\_powerctrl -chkhw 2** [verifies that peer power control hardware is present in slot\_2]

**/nas/tools/.server\_peer\_powerctrl -reboot 2** [Reboots server\_3]

**/nas/tools/.server\_peer\_powerctrl -chkins** [checks to see if peer power control hardware is present]

**Note:** No output indicates that all is well

**# /nbsnas/tools/.server\_peer\_powerctrl -chkhw 4**

host 192.168.1.4 is not responding! Is it up??

.server\_peer\_powerctrl: Data Mover server\_4 is not responding

**Note:** There is no physical server in slot\_4 in the above example

**/nas/tools/.server\_peer\_powerctrl -chkins 2 | 3** [checks to see if peer DM is inserted]

**/nas/tools/.server\_peer\_powerctrl shutdown** [powers off DM]

**/nas/tools/.server\_peer\_powerctrl -restore** [restores DM power]

**Note:** Peer power control commands use HTTP/XML to perform tasks and telnet to create communications link. Be aware that if DNS configurations are incomplete or wrong, telnet, and therefore peer power control commands may fail—see example:

**# /nas/tools/.server\_peer\_powerctrl -chkins 3**

host server\_3 is not responding! Is it up??

.server\_peer\_powerctrl: Data Mover server\_3 is not responding

## **USING T2TTY COMMANDS—Serial or RPC communications between CS & DM:**

→The t2tty facility can be used to communicate to the Data Mover, from the Control Station, using either Serial communication, or Internal Network interface using RPC (RPC would be only facility available for NSX, NS40, or NS80 systems—no Serial connections)

→t2tty facility can be used to test Serial connections on system, and the IPMI connection between CS0 & CS1

### # /nasmcd/sbin/t2tty -n

2 →Indicates two active serial connections (i.e., Data Mover 2 and Data Mover 3)

Note: Used to count the active Serial connections for ttyS4 – ttyS9

### # /nasmcd/sbin/t2tty -t

ttyS4 on

ttyS5 on

ttyS6 off

ttyS7 off

Note: Used to check state of Serial connection for Serial ports

### # /nas/sbin/t2tty -s 3

Slot-3 (ttyS5) on

### # /nasmcd/sbin/t2tty -m 192.168.3.101

IPMI Ping to 192.168.3.101 Succeeded

Note: Use to check state of IPMI connection between CS0 & CS1

→t2tty facility can also be used to initiate a Flash upgrade of the Data Mover BIOS/POST

Flash Data Mover in slot (2-5) with image file(s) (BIOS and/or POST) -F forces overwrite, disregarding version numbers

### # t2tty -f 2 image\_file {image\_file}

→Use t2tty to force PXE boot of Data Mover

### # /nasmcd/sbin/t2tty -p 2

## EXAMPLES:

### RPC METHOD:

#### # /nasmcd/sbin/t2tty -C 2 "ifconfig" →Uses RPC to send & receive output from DART to Control Station

Note: For NSX Systems, the RPC interface allows upper or lowercase: -c “ifconfig” or -C “ifconfig”

### SERIAL PORT METHOD:

#### # /nasmcd/sbin/t2tty -c 2 "logsys add output console" →Uses Serial communication between DART & Control Station

#### # /nasmcd/sbin/t2tty -c 2 "ifconfig"

#### # /nasmcd/sbin/t2tty -c 2 "logsys delete output console"

Note: Make sure to change output to console or the “ifconfig” command will not produce expected output on screen. Make sure to delete the output to console after you are done using t2tty. May want to consider just using -C for the RPC method.

### /nas/sbin/t2tty or /nasmcd/sbin/t2tty

Celerra\IP Serial Connection Tester....

Usage:

Check serial connection signal on ttyS4 thru ttyS9

t2tty -t

Check serial connection signal to a given slot (2-5)

t2tty -s 2

Force Data Mover in slot (2-5) to PXE boot

t2tty -p 2

Send dart command "cmd" to data mover in slot (2-5)

t2tty -c # "cmd"

for example: t2tty -c 2 "ifconfig"

### # /nasmcd/sbin/t2tty -c 3 "camshowconfig"

camshowconfig

CAM Devices on scsi-32:

TID 00: 0:d0+ 1:d1+ 2:d2+ 3:d3+ 4:d4+ 5:d5+

TID 01: 0:d6+ 1:d7-

### # /nasmcd/sbin/t2tty -c 3 "fcip bind show"

fcip bind show

\*\*\* Persistent Binding Table \*\*\*

Chain 0000: WWN 5006048000000000 HBA 0 N\_PORT Bind Pending

Chain 0016: WWN 5006048000000000 HBA 1 N\_PORT Bind Pending

Chain 0032: WWN 5006016610602235 HBA 2 SP-a6 Bound

Chain 0048: WWN 5006016e10602235 HBA 3 SP-b6 Bound

\*\*\* Dynamic Binding Table \*\*\*

Chain 0000: WWN 0000000000000000 HBA 0 ID 0 Inx 00:81 Pid 0000 S\_ID 000000 Non

Chain 0016: WWN 0000000000000000 HBA 1 ID 1 Inx 01:81 Pid 0016 S\_ID 000000 Non

Chain 0032: WWN 5006016610602235 HBA 2 ID 2 Inx 02:00 Pid 0032 S\_ID 0000ef Sys

## **OUTPUTTING INFORMATION TO CONSOLE USING T2TTY COMMANDS:**

**Note:** Changing Server output to local console allows t2tty commands to output to screen.

**# /nasmcd/sbin/t2tty -c 2 "logsys add output console"**

**CONSOLE>**

**# /nasmcd/sbin/t2tty -c 2 "logsys delete output console"**

**Note:** Run this command after running t2tty commands to delete output to console

**# /nas/sbin/t2tty -c 2 "ifconfig" "ifconfig 192.168.1.2 down" "ifconfig 192.168.1.2 up"**

reason\_code=05

ifconfig

Devices:

fxp0 dmtu=1500, dmac=8:0:1b:43:b:63

loop dmtu=65536, dmac=0:0:0:0:0

Interfaces:(3)

e130 on fxp0 l=192.168.1.2 n=255.255.255.0 b=192.168.1.255 DNIF UP

    mtu=1500, dmtu=1500, vlid=0, mac=8:0:1b:43:b:63 dmac=8:0:1b:43:b:63

e131 on fxp0 l=192.168.2.2 n=255.255.255.0 b=192.168.2.255 DNIF UP

    mtu=1500, dmtu=1500, vlid=0, mac=8:0:1b:43:b:63 dmac=8:0:1b:43:b:63

loop on loop l=127.0.0.1 n=255.0.0.0 b=127.255.255.255 UP

    mtu=32768, dmtu=65536, vlid=0, mac=0:0:0:0:0:0 dmac=0:0:0:0:0:0

## **TURNING OFF CONSOLE OUTPUT:**

**# /nas/sbin/t2tty -c 2 "logsys set output disk=root\_log\_2"**

**# /nasmcd/sbin/t2tty -c 2 "logsys delete output console"**

## **CELERRA NAS 5.2:**

NAS 5.2.9.6 GA February 9, 2004; Supports the CNS-14, CFS-SE systems, Celerra NS family and Celerra NS-G (direct and switched connections) family products. Current revision 5.2.22.1. EOL announced as February 28, 2006 for this code family.

### **Caution:**

Be aware that NAS 5.2.15.3 conducts automatic conversion to MPD format on all file systems and changes directory last modified times. Also, this revision supports FLARE 14.

### **Requirements:**

--Symm microcode 5267.42/52.29 Symm 4.0/4.8; 5568.62.23 for Symm 5.0/5.5; 5669.48.25/5670.26.28/5670.66.62 for Symm 6.0

--Flare 8.50/8.51 for FC4700; 2.05.1.60.5.xxx or 2.06.400.5.xxx for CX400; 2.06.500.5.xxx for CX500; 2.05.0.60.5.xxx for

NS600/600S & 2.05.1.60.5.xxx for NS600G/700G & 2.06.600.5.xxx; Flare 2.06.700.5.xxx for CX700

### **Limitations:**

--Supports only Flare 12 & 13 for Clariion backends!

### **Features:**

--Internal Usrmapper

--Introduces Multiprotocol Directories (MPD—DIR3) that will replace the old Shadow file directory format

--Windows 2003 Domain support, SMB Signing, Dynamic DNS Updates, GPO Policies, Kerberos over TCP

--NDMP Version 4 support

--SYR support

--ADS support for directories

--GPO, NFS Security, UNIX ACL tool, CDMS, CIFS Replication, SnapSure, virtual DM for CIFS, enhancements

--Introduces NS630 with 3Ghz Pentium 4 and 533 Mhz Front Side Bus, 4GB RAM

--Supports Clariion with Flare 12 and 13 [Flare 11 or lower is not allowed]

### **NAS 5.2 SUPPORT:**

CFS-14; CFS-SE; NS600/G ; NS700/G and higher requires 5.2

## **NAS 5.2 BOOT OPTIONS FOR INSTALLS, UPGRADES**

--Ability to conduct ‘unattended’ installs and upgrades via the use of a ‘kickstart’ or ‘serialkickstart’ command and the “ksnas.cfg” file

### **BOOT OPTIONS:**

Install

Serialinstall

Kickstart [Run d:\setup -f ksnas.cfg for Golden Eagle from Service Laptop CD-ROM]

Serialkickstart [Use for NS600 unattended installs]

**Note:** CFS-14 Eagle, SE, NS600, NS600G, NS700, or NS700G will require editing of ‘ksnas.cfg’ file on the boot floppy, whereas the CNS-14 Golden Eagle requires editing of ‘ksnas.cfg’ file from the install CD-ROM onto a c:\Temp directory on the service laptop.

D:\NetworkInstall>setup -f ksnas.cfg

## **CONTROL VOLUME CHANGES:**

From NAS 5.2.7.0 and higher, all new installs will require a size change from 4G to 11G on the first two Control Volumes [Captive Clariion backend control luns will remain same until next release of 5.2 and then this will be auto calculated. Clariions on SAN will require manual configuration]. Purpose of this change was to allow for more room for the Internal Usrmapper database.

## **ASYNC DATA RECOVERY FOR CELERRA REPLICATOR:**

Allows use of replication between file systems for disaster recovery by allowing failover/failback with incremental resyncs and baseline checkpoint on SFS. CIFS is supported with Virtual Data Mover module to transfer Share ACLs, Usrmapper database, GPO, AV, local groups, and password file information. DR means that Primary site goes down, Secondary site is activated and SFS can be made Read/Write. Then, when Primary is back online a resync operation occurs without making full copies of filesystems necessary and replication can be re-established.

## **INTERNAL USRMAPPER:**

Usermapper automatically starts and runs on Server\_2 for new installations. No manual configuration required and other data movers will discover the Usrmapper service via broadcast over the internal IP network. External Usrmapper configurations can also be migrated, or imported, to an Internal Usrmapper configuration, and vice versa.

## **SNAPSURE ENHANCEMENTS:**

- Subfolders now contain hidden .ckpt directories
- Out of order deletion of checkpoints are now supported, persistent after DM reboots
- Supports 64 checkpoints per file system
- Checkpoints will automatically mount after being taken
- Can now rename a specific checkpoint

## **CELERRA MULTIPROTOCOL DIRECTORY (Celerra MPD—aka DIR3):**

**Note:** Introduced NAS 5.2.x. NAS 5.2.15.3 automatically converts all file systems to MPD format during Upgrade and last modified date on directories are updated to the date the conversion took place. NAS 5.5 will no longer allow old COMPAT Shadow structure. With DIR3, the Unix name, DOS m8.3 name, and Windows M256 long name, are all stored in the directory block for files. Hash algorithms are used to aid in searches & listings. Hash is built using file name, inode number, and file's position within the directory. With COMPAT, the traditional SHADOW file implementation is used, and only the Unix file names are stored in the directory block.

## **PREVENTING AUTOMATIC MPD CONVERSION:**

List all file systems to exclude from conversion in [\*\*# nas\\_fs -l | awk 'NR>1{print \\$6}' > /nas/site/fs\\_exclude\*\*](#), listing each file system name on a separate line

**Note 1:** To create a list of file systems for the EXCLUDE file, run:

**# nas\_fs -l | awk 'NR>1{print \$6}' > /nas/site/fs\_exclude**

**Note 2:** AR71786 describes a problem with NAS 5.4 Upgrades in which the fs\_exclude file is ignored because it is now looking for fs\_exclude file with the following format:

### **NEW FS EXCLUDE FORMAT 5.4:**

fs1:fs2:fs3 (now requires colon separation between file system names, all on a single line in the file!)

→ This behavior will be corrected in a future 5.4 release so as to accommodate the original fs\_exclude format, which was to have each file system listed separately on separate lines, without colons—see primus emc94312

### **ORIGINAL FS EXCLUDE FILE FORMAT:**

fs1

fs2

fs3

## **MANUAL MPD CONVERSION SYNTAX HAS CHANGED FOR NAS 5.4:**

**\$ nas\_fs -translate Shane\_mpd -mpd start** (original syntax nas\_fs –translate fs1 start)

→ Support for both Shadow File and new MPD FS [Enhanced FS Directory “DIR3”]

→ Ability to change from Shadow [COMPAT] to MPD [DIR3] Online

→ New installs and upgrades will create/convert File Systems to MPD, but at sites with known Shadow problems, this process can be stopped on designated file systems

→ Both NFS and CIFS directories will contain names within file system—current algorithm to translate between NFS and & CIFS names was dependent on order in which files were created, and Shadow rebuilds do not guarantee that the order will be saved.

**Note:** Main difference is that the CIFS FS metadata was stored outside of the file system with the “SHADOW” file [CIFS attributes, CIFS long & short names, UNIX name] whereas the MPD stores the metadata within the File System in a new three-name directory format composed of NFS Name, CIFS Long Name (M256), and CIFS Short Name (8.3)

#### **PATH CACHING PARAMETER:**

DIR3 file systems do not contain shadow files, yet still rely on legacy shadow routines such as name lookups and path checking to perform work. Enable or disable path caching using following:

**param cifs PathCache=0**

#### **MPD CONVERSION PROBLEMS:**

MPD conversion routine has had a bug since day one such that if the inlineValidate complains about a “.” Inode as not the same as a containerInumber, then the xlate process cannot fix because it does not have permission to read containers to validate this issue. So, MPD conversion fails with Mangled directory entry panic. See AR50292.

##### **1. MPD Conversion Fails to Complete on File Systems:**

###### **Solution:**

→FSCK/ACLCHK affected file systems, mount to Standby after converting to a Primary Server, complete MPD conversion, mount back to production Server.

**Comment:** This would not be the first course of action to take for failed MPD conversions—this would be a last resort step in conjunction with an Engineering recommendation!

##### **2. MPD Fails Because of Shadow File Corruption:**

--#nas\_fs -i xxx -o mpd shows Translation\_error\_code = 28

--Server Log shows ‘MPDTranslate of /test: failure on dir inode 679125, status = InternalError

**Solution:** see emc103638

##### **3. MPD Conversion Fails and Server Log Indicates Quota Issues:**

--Failover & fallback Server, have customer reconfigure affected User Quotas, stop and restart conversion, should succeed

--similarly for Tree Quotas issue, one case required customer extension of Hard Quota on Tree Quotas to complete conversion

##### **4. MPD Conversion Fails with I18N characters and Shadow File Translation Issues:** see AR48340

2004-11-13 05:52:03: SHADOW: 3: update: Unable to translate name: 5

2004-11-13 05:52:03: SHADOW: 3: aA39497a938582a918485A3KA3A7A3--

2004-11-13 05:52:05: UFS: 3: 6: MPDTranslate of /home: failure on dir inode 276314, status = InvalidArgument

**# nas\_fs -i vol1 -o mpdtail**

Multi-Protocol Directory Information

```
Default_directory_type = DIR3
Needs_translation      = True
Translation_state       = Not Requested
Percent_nodes_scanned  = 99%
Has_translation_error  = True
Translation_error_code  = 5
Translation_error_message = %InvalidArgument%
```

- a. Get the list of files or directories that failed during the Conversion from the Server Log
- b. Locate the path for each file/directory by running the following for the file name or the inode:

**# find -iname mail997715f00050099.mai**

./Backups/dan/tmp/mail997715f00050099.mai

**Note:** Command run from Control Station mounted to Server over External interface:

**# find /emcfind -inum 125228 -print >/inode1 &**

**# cat inode1**

/emcfind/home/Faculty/KHS/kgonzalez/ZSpeech/Introduction speech/INTRODUCTION SPEECH GRADE SHEET.doc

c. Conduct a Unix rename of the file or in the case of a directory, cd to the directory above the problem directory, move the offending directory to a temporary directory name.

d. Restart MPD Conversion:

**#nas\_fs -translate fs1 start**

**Note:** Pre-NAS 5.4 syntax to manually convert File Systems to MPD

**# nas\_fs -translate fs1 -mpd start**

**Note:** New syntax for manually converting File Systems to MPD with NAS 5.4

e. After the MPD Conversion completes, rename the directories and/or files back to their original names

##### **5. After MPD Conversion CIFS Clients cannot see all Directories:**

Problem where the root directory of a file system was hashed before the xlt tables were available. See AR50408. Fix would be to unmount and remount file system to rebuild hash table correctly. Also, running cifs update, shadow fix, etc., can help resolve Shadow problems that prevent successful translations.

**PROCEDURE FOR CORRECTING MOST FAILED MPD CONVERSIONS/TRANSLATIONS:**

**Note:** Use following procedure if MPD translation fails with InternalError, InvalidArgument, or NameTooLong errors

**1. Use Server Log to find the problem inode(s):**

```
2005-03-02 18:36:57: ADMIN: 4: Command succeeded: file MPDTranslate ufs filesys on 18
2005-03-02 18:36:58: UFS: 3:10: MPDTranslate of /fs1: shadow getAllNames of daily plans.doc failed in upgrade of ino 125246,
status = InternalError
2005-03-02 18:36:58: UFS: 3: 6: MPDTranslate of /fs1: failure on dir inode 125246, status = InternalError
2005-03-02 18:37:02: UFS: 3:10: MPDTranslate of /fs1: shadow getAllNames of ~WRL0516.tmp failed in upgrade of ino 738231,
status = InternalError
2005-03-02 18:37:02: UFS: 3: 6: MPDTranslate of /fs1: failure on dir inode 738231, status = InternalError
2005-03-02 18:42:50: UFS: 3: MPDTranslate of /fs1: completed translation of 0 of 101905 dir inodes without 100 percent success.
See log for errors
```

**Note:** In most cases, the problem Inode number is actually the directory for where the filename that is printed in the Server Log is located. See AR47080 for more details. Sometimes, during DIR3 MPD translation, the XLATE process queries each file system directories' SHADOW file structure, and in this case, when the SHADOW file is broken, returns an InternalError, preventing the translation from completing.

**2. Conduct XOR calculation of Inode number from Server Log:**

**Note:** From the Inode number provided in the Server Log, conduct XOR calculation to convert the Inode number to the 'correct value' when conducting a file system search from the rootfs of the data mover. For example, when conducting find from /nas/rootfs/slot\_2/fs01, or when mounting the root of the Server from the Control Station (#mount server\_2: /find), you would need to search by the XOR'ed value. If you had the permission to mount, from the Control Station, as server\_2:/fs1 /find, then the find operation would use the same inode value as listed in the Server Log, but in most cases, you will not have the permission to do so:

a.) Find fsid value of file system mounted to /fs1 using #nas\_fs -l

**Note:** Left column of nas\_fs -l output represents fsid value, in this case fsid=26

b.) Inode number from Server Log: 738231

c.) Enter value in calc as decimal, click on XOR function, enter fsid value of "26", and the click on enter: **738221 is the XOR value**

**Note:** XOR the FSID and INODE number from Server Log using AWK in script to get true Inode number

**\$ echo “44449 45” | awk '{print(xor(\$1,\$2)) }'**

44428

**3. Mount root file system of data mover from Control Station:**

**# mount server\_2:/ /find**

**4. Conduct Inode search based on XOR value, 738221, from "/" on Control Station, and print to file:**

**# find /find/fs1 -inum 738221 -print >xor\_inode &**

**Note:** The find operation can take many hours, depending on size and layout of file system.

**5. Check the output file once the find job has completed:**

```
# cat xor_inode
# /find/fs1/.ckpt_v0l_tue_ckpt/pliu/CCTemplate/0224-Wholesale-LNC_Success.doc
/find/fs1/.ckpt_v0l_mon_ckpt/pliu/CCTemplate/0231-Wholesale-OCRA_FailureNack.doc
# /find/fs1/pliu/CCTemplate/IC
```

**Note:** Ignore checkpoint paths and concentrate on the path that prints for the production file system.

**6. Verify that the inode found translates into the directory called "IC", by going to the directory above "IC" and running an ls -ail:**

```
#pwd
/find/fs1/pliu/CCTemplate
#ls -ail
738221 drwxr-xr-x 5 80172 wheel 3072 May 4 21:06 IC
```

**7. Conduct Shadow fix on the problem directory:**

**#.server\_config server\_2 "shadow fix \fs1\pliu\CCTemplate\IC"**

1115402638: ADMIN: 4: Command succeeded: shadow fix \fs1\pliu\CCTemplate\IC

**Note:** Shadow Fix zeroes out Shadow file completely, eliminating the conflict, subsequently allowing the MPD translation to run. Be aware that Shadow Fix will not succeed where there is no residual Shadow File, a key indicator.

**# .server\_config server\_2 -v "shadow fix '\PMC\_FS01\_E\$\Data\CO390\Users\JWKleparsk\My Docs'"**

1112312141: CFS: 3: shadow fix failed: NotFound

**Note:** Above output indicates that it could not find a Shadow File to fix!

**8. Restart MPD translation & Observe Successful Conversion:**

**# nas\_fs -translate fs1 start**

**# nas\_fs -translate fs1 -mpd start**

**Note:** NAS versions 5.3.14.0 & 5.4.6.0 contain syntax change for conducting manual MPD conversions. AR71461 submitted to doc.

### # nas\_fs -i fs1 -o mpd |tail

Multi-Protocol Directory Information

```
Default_directory_type = DIR3
Needs_translation = False
Translation_state = Completed
Has_translation_error = False
```

### **ALTERNATIVE FIND METHOD TO BYPASS NEED FOR XOR'ing INODE NUMBER:**

**Note:** If you can successfully export and mount the file system in question directly from the Control Station, then you do not need to run the XOR method outlined above, as the Server Log inode number will be the actual inode number used in the find

#### **1. Export file system with anon=0 rights:**

```
# server_export server_2 -o anon=0 /cifs1
```

server\_2 : done

```
# server_export server_2
```

server\_2 :

```
export "/cifs1" anon=0
```

#### **2. Make temporary mountpoint on Control Station and mount DM file system:**

```
# mkdir /mnt/cifs1
```

```
# mount server_2:/cifs1 /mnt/cifs1
```

```
# mount
```

server\_2:/cifs1 on /mnt/cifs1 type nfs (rw,addr=192.168.1.2)

#### **3. Run find using same inode number as seen in Server Log:**

```
# find /mnt/cifs1 -inum 163205 -print >cifs1_inode &
```

### **COMMON ERRORS AND CONDITIONS SEEN:**

#### **Server Log InternalError:**

2005-03-02 18:36:57: ADMIN: 4: Command succeeded: file MPDTranslate uxf5 filesys on 18

2005-03-02 18:36:58: UFS: 3:10: MPDTranslate of /fs1: shadow getAllNames of daily plans.doc failed in upgrade of ino 125246, status = InternalError

2005-03-02 18:36:58: UFS: 3: 6: MPDTranslate of /fs1: failure on dir inode 125246, status = InternalError

2005-03-02 18:37:02: UFS: 3:10: MPDTranslate of /fs1: shadow getAllNames of ~WRL0516.tmp failed in upgrade of ino 738231, status = InternalError

2005-03-02 18:37:02: UFS: 3: 6: MPDTranslate of /fs1: failure on dir inode 192485, status = InternalError

2005-03-02 18:42:50: UFS: 3: MPDTranslate of /fs1: completed translation of 0 of 101905 dir inodes without 100 percent success.

See log for errors

#### **Server Log InvalidArgument error:**

2005-05-04 21:06:17: UFS: 3:10: MPDTranslate of /fs1: shadow getAllNames of Image Caf addon\_welcome.doc failed in upgrade of ino 738231, status = InvalidArgument

2005-05-04 21:06:17: UFS: 3: 6: MPDTranslate of /fs1: failure on dir inode 738231, status = InvalidArgument

2005-05-04 21:20:53: UFS: 3: MPDTranslate of /fs1: completed translation of 0 of 306542 dir inodes without 100 percent success.

See log for errors

#### **NAS FS OUTPUT:**

errors during a running translation

```
# nas_fs -i fs1 -o mpd |tail
```

```
Needs_translation = True
Translation_state = Running
Percent_nodes_scanned = 2%
Has_translation_error = True
Translation_error_code = 28
Translation_error_message = %InternalError%
Skipped_dir_inodes = 1
Current_inode = 706176
Entries_in_current_inode = 0 of 0
```

InternalError error:

```
# nas_fs -i fs1 -o mpd
```

Multi-Protocol Directory Information

```
Default_directory_type = DIR3
Needs_translation = True
```

```
Translation_state      = Failed
Percent_nodes_scanned = 99%
Has_translation_error = True
Translation_error_code = 28
Translation_error_message = %InternalError%
```

InvalidArgumentException:

```
# nas_fs -i fs1 -o mpd |tail
```

Multi-Protocol Directory Information

```
Default_directory_type = DIR3
Needs_translation     = True
Translation_state      = Failed
Percent_nodes_scanned = 99%
Has_translation_error = True
Translation_error_code = 5
Translation_error_message = %InvalidArgumentException%
```

```
$ nas_fs -i ufs01 -o mpd
```

Multi-Protocol Directory Information

```
Default_directory_type = DIR3
Needs_translation     = True
Translation_state      = Failed
Percent_nodes_scanned = 99%
Has_translation_error = True
Translation_error_code = 6
Translation_error_message = %NameTooLong%
```

### **Server log server 3 -a -s**

2005-08-13 12:51:00: UFS: 3:10: MPDTranslate of /appsmp: direnter of rt.jarOrig failed for dir into 119253, status = NameTooLong

2005-08-13 12:51:00: UFS: 3: 6: MPDTranslate of /appsmp: failure on dir inode 1 19253, status = NameTooLong

### **SERVER LOG ERRORS FOR ADS PROBLEM:**

UFS: 3: 7: MPDTranslate of /filesystem: Dir corruption in inode 21: Corrupted inode number 20 found for ".", status = InternalError

UFS: 3: 6: MPDTranslate of /filesystem: failure on dir inode 21, status = InternalError

UFS: 3: MPDTranslate of /filesystem: completed translation of 0 of 7 dir inodes without 100 percent success. See log for errors

**Note:** Alternate Data Streams translation issues resolved in AR49041, NAS 5.2.17.1, 5.3.12.0, 5.4.2.0

### **COMMENTS ABOUT XOR:**

When accessing DM file system over internal IP network or from by mounting its root, you will need to XOR the inode value printed in the Server Log to find the actual problem directory (s) where the Shadow File is corrupt. Both examples below are mounting rootfs of Data Mover over Internal IP network:

```
#cd /nas/rootfs/quota/slot_x/fs1
```

```
#mount server_2:/ /find
```

### **IMPORTANT COMMENTS ABOUT SHADOW REaddir & SHADOW FIX:**

```
# .server_config server_2 -v "shadow readdir"
```

```
\PMC_FS01_E$\Data\H0060\Users\MAWeed\AppData\Microsoft\Office\Recent"
```

1113219816: CFS: 5: start scavenging

1113219816: CFS: 5: scavenged 101 forks to free 208 maps in 12553511 ticks

**Note:** Shadow Readdir fails because of a problem with the Shadow File in this directory. Shadow Readdir command succeeds even for DIR3 style directories, this is normal.

```
# .server_config server_2 -v "shadow fix '\PMC_FS01_E$\Data\C0390\Users\JWKleparsk\My Docs'"
```

1112312141: CFS: 3: shadow fix failed: NotFound

**Note:** Shadow Fix command will log above results if the directory has already been converted to DIR3 & does not contain a residual Shadow File. You know you have the correct directory found when the command does succeed, as in following output:

```
# .server_config server_2 -v "shadow fix"
```

```
\PMC_FS01_E$\Data\H0060\Users\MAWeed\AppData\Microsoft\Office\Recent"
```

1113219869: ADMIN: 4: Command succeeded: shadow fix

\PMC\_FS01\_E\$\Data\H0060\Users\MAWeed\AppData\Microsoft\Office\Recent

**Note:** In some cases, the command will not run because of spaces or other peculiarities in the path—when all else fails, try to determine DOS 8.3 name for the next directory in the path and substitute that name:

```
# .server_config server_7 -v "shadow readdir '\data\userdata\SFISHE3\Sandy's files\Action Plans'"
```

1131570657: CFS: 3: shadow fix: getAlternateName failed: NotFound

```
# .server_config server_7 -v "shadow readdir \data\userdata\SFISHE3\SANDY'~1\ACTION~1"
```

### **MPD BENEFITS:**

-- MPD is expected to improve file system performance & maintain more consistent metadata

--No Shadow file rebuilds

### **MPD ERRATA:**

With NAS 5.3, Upgrades will automatically convert file systems to DIR3 format. Use the “-no\_shadow” switch when translating pure NFS file systems so as not to incur space and inode usage with the building of a Shadow directory. MPD directory format contains UNIX Name, CIFS Long & DOS 8.3 Short Names. NAS 5.5 will not allow file systems to remain as COMPAT.

### **VERIFYING DIR3 FILESYSTEMS WITH NAS 5.5:**

```
# /nas/sbin/rootnas_fs -i root_fs_common
```

Multi-Protocol Directory Information: not available

**Note:** NAS 5.5 does not properly return the DIR3 value using nas\_fs -i or the XML query. Output below from NAS 5.4:

```
# nas_fs -query:"name=root_fs_common" -format:"%s %s\n" -fields:Name,DirectoryType
root_fs_common DIR3
```

### **DETERMINING FS FORMAT—‘COMPAT’ SHADOW vs. ‘DIR3’ MPD FORMAT:**

1. Turn on NAS FS Option for MPD:

```
#export NAS_DB_DEBUG=1
```

2. #nas\_fs -info fs1 –option mpd

#### Multi-Protocol Directory Information

|                       |                 |                                                                                       |
|-----------------------|-----------------|---------------------------------------------------------------------------------------|
| <b>directory_type</b> | <b>= DIR3</b>   | [DIR3 means that FS is using MPD format]                                              |
| needs_translation     | = False         | [Means all directories are of the same type; ‘True’ means some of each DIR3 & COMPAT] |
| translation_state     | = not requested | [Means a translation has never been requested on this FS]                             |
| has_translation_error | = False         |                                                                                       |

#### **EXAMPLES:**

**# nas\_fs -i vol1 -o mpd|tail**

Multi-Protocol Directory Information

|                           |                     |
|---------------------------|---------------------|
| Default_directory_type    | = DIR3              |
| Needs_translation         | = True              |
| Translation_state         | = Not Requested     |
| Percent_nodes_scanned     | = 99%               |
| Has_translation_error     | = True              |
| Translation_error_code    | = 5                 |
| Translation_error_message | = %InvalidArgument% |

**\$ nas\_fs -i fs1 -o mpd |tail**

Multi-Protocol Directory Information

|                        |             |
|------------------------|-------------|
| Default_directory_type | = DIR3      |
| Needs_translation      | = False     |
| Translation_state      | = Completed |
| Has_translation_error  | = False     |

### **CONVERTING “COMPAT” SHADOW FORMAT TO MPD “DIR3” FORMAT:**

1. Verify Current File System Directory Format & Obtain Inode and File System Space Info:

**\$ nas\_fs -i fs94 -o mpd**

|        |         |
|--------|---------|
| id     | = 21    |
| name   | = fs94  |
| acl    | = 0     |
| in_use | = True  |
| type   | = uxfss |
| volume | = mtv04 |
| pool   | =       |

```
rw_servers= server_2
ro_servers=
rw_vdms =
ro_vdms =
symm_devs = 002806000123-0010
disks = d12
disk=d12 symm_dev=002806000123-0010 addr=c0t119-16-1 server=server_2
```

#### **Multi-Protocol Directory Information**

```
Default_directory_type = COMPAT
Needs_translation = False
Translation_state = Completed
Has_translation_error = False
```

**\$ server\_df server\_2 -i |grep fs94**

```
fs94 10325744 9546816 778928 92% /mp04
```

**Note:** Do not proceed with conversion unless you have at least 10% Inodes and Space available

2. Change Directory Format Type with Following Command:

**#nas\_fs -modify fs94 directory\_type=DIR3 | directory\_type=COMPAT**

3. Start File System Format Conversion:

**#nas\_fs -translate fs94 start | -no\_shadow** [Will not create a Shadow directory when translating from COMPAT]

4. Monitor Conversion Process:

**\$ nas\_fs -i fs94 -o mpd**

#### **Multi-Protocol Directory Information**

```
Default_directory_type = DIR3
Needs_translation = True
Translation_state = Running
Percent_nodes_scanned = 4%
Has_translation_error = False
Skipped_dir_inodes = 0
Current_inode = 50720
Entries_in_current_dir = 0
```

5. Verify in Server Log:

**\$ server\_log server\_2 |grep -i mpd**

```
2004-05-07 14:11:29: ADMIN: 4: Command succeeded: file MPDTranslate uxfss filesys on 21
```

```
2004-05-07 14:11:48: UFS: 3: 6: MPDTranslate of /fs94: Successfully completed translation of 0 of 2349 dir inodes
```

**Caution:** You can actually convert back to COMPAT format but it is not supported at this time and should not be done

### **CONVERTING DIR3 FILE SYSTEM TO COMPAT WITH NAS 5.5:**

**#.server\_config server\_4 -v "file dirType uxfss filesys DIR\_COMPAT 266"**

### **TROUBLESHOOTING DIR3 CONVERSIONS:**

**Note:** By default, a translation will “pause” at 99% capacity of the file system or inode value. To resume translation, do following:

1. Extend File System or delete data to reach 90% [Translation is supposed to resume if the 90% threshold is reached]

2. If translation does not resume, permanently unmount and then remount File System

**Note:** Possible translation states: running; not requested; never; pending; queued; paused; completed; failed

3. Do not run MPD conversions with Quotas turned on—disable quotas first, then do the conversion

4. Directory Translation Error with multiple stream attributes—code is unaware of streams & fails with following:

```
2004-08-18 02:40:28: UFS: 3: 7: MPDTranslate of /volumes-mnt: Dir corruption in inode 64828183: Corrupted inode number
```

```
64828181 found for ".", status = InternalError
```

```
2004-08-18 02:40:28: UFS: 3: 6: MPDTranslate of /volumes-mnt: failure on dir inode 64828183, status = InternalError
```

**Note:** Primus emc94552, AR49041, to be corrected in NAS 5.2.17.x, 5.3.12.x and higher

### **MPD PARAMETERS:**

**\$ .server\_config server\_2 -v "param ufs" |grep -i xlate**

```
ufs.xlateMinThreads 0x01170074 0x0000000a 0x0000000a [Default = 10 threads minimum]
```

```
ufs.xlateMaxThreads 0x01170070 0x0000000a 0x0000000a [Default = 10 threads]
```

```
ufs.xlateSpaceLowCeiling 0x01170088 0x000003e8 0x000003e8
```

[Default = 1000 = 10% free space before resuming]

```
ufs.xlateSpaceHighCeiling 0x01170084 0x00000064 0x00000064
```

[Default Value=100 dec. = 99% Full]

## **GPO ENHANCEMENTS:**

Max lifetime for kerberos user tickets

Domain audit account logon events, account management, directory service access, user rights, etc.

## **UNIX ACL TOOL:**

Allows for NFS User to view & modify CIFS ACLs for files [emcgetsd and emcsetsd] from a Unix Host [HPUX, Linux, Solaris] or the Control Station. Obtain ‘emcgetsd’ and ‘emcsetsd’ from Applications & Tools CD>CifsTools>unixacltools/linux. Must set param cifs acl.extacl=32 in order for the “emcsetsd” tool to work.

## **CELERRA VDM:**

### **NAS 5.2 VIRTUAL DATA MOVER (VDM):**

#### **Purpose:**

--Separates CIFS Services and File Systems into virtual containers to allow for administration of separate groups of CIFS Servers and

to enable replication of CIFS environments between physical Data Movers [NAS 5.3]

--VDM will be logically connected between network interfaces and the file system as a separate CIFS instance that can be transferred from DM to DM.

--VDM will contain the audit log, local groups, and shares database within the .etc directory of each file system. A VDM server can only export a directory on its own file system or nested file system. Default root filesystem size of VDM will be 128MB, same as current root file systems. VDM relationship to rootfs will be saved on CS.

--Internal Usermapper required for VDM.

--Can have a group of VDM's per file system.

--VDM container can have more than one CIFS Services

## **LISTING VDMs ON CELERRA:**

**# nas\_server -a -i -vdm**

```
id      = 3
name    = vdm_houfsrv0
acl     = 0
type    = vdm
server  = houons7041dm2
rootfs  = root_fs_vdm_houfsrv0
I18N mode = UNICODE
mountedfs = houfsrv0,ckpt_daily_houfsrv0_001
member_of =
status   :
defined = enabled
actual = loaded, active
```

Interfaces to services mapping:

```
interface=10_30_161_228 :cifs
interface=10_30_161_61 :cifs
```

**# nas\_server -l -vdm**

```
id      acl server mountedfs rootfs name
3       0      1       165,166,238,240 225  vdm_source
```

## **ISSUE WHEN VDM NAMES ARE ALL UPPERCASE:**

**# nas\_server -i -a -vdm**

```
id      = 1
name    = CLKPTDM04
acl     = 0
type    = vdm
server  = server_4
rootfs  = root_fs_vdm_CLKPTDM04
```

**# server\_export CLKPTDM04**

**Note:** In this particular example, had to use all Uppercase characters when retrieving VDM information from CLI. AR74879

## **CONTENTS OF VDM DATABASE:**

--local groups database  
--shares database for file systems dedicated to the VDM  
--CIFS server configuration [compnames, interface names, etc]  
--Home directory information  
--Auditing and Event log information  
--Kerberos information for Servers in VDM  
--Secmap cache—secure cache

**Unsupported VDM Features:**

No individual NTP configuration—must use single NTP for data mover globally  
FTP to VDM not supported  
Failover between VDMs not supported  
CDMS with CIFS supported, but not CDMS with NFS  
NFS exports are not supported  
Server\_uptime, cpu, standby commands are not supported on individual VDMs

**SETTING UP A VDM CIFS SERVER CONTAINER ON THE DATA MOVER:**

**1. CREATE VDM CONTAINER:**

# **nas\_server -n vdm\_tm -type vdm -create server\_2**

```
id      = 1
name    = vdm_tm
acl     = 0
type    = vdm
server  = server_2
rootfs  = root_fs_vdm_1
I18N mode = UNICODE
mountedfs =
member_of =
status   :
defined = enabled
actual = loaded, ready
```

Interfaces to services mapping:

**Note:** For VDM creation, only two “states” are allowed, setstate “loaded” or “mounted”. When the –setstate switch is not specified, the default behavior is to create the VDM in a loaded state. The “-fs” switch is used when the rootfs for the VDM has already been manually created for use with a VDM.

# **nas\_server -a -i -vdm**

```
id      = 1
name    = vdm_dest
acl     = 0
type    = vdm
server  = server_2
rootfs  = root_fs_vdm_vdm_dest
I18N mode = UNICODE
mountedfs =
member_of =
status   :
defined = enabled
actual = temporarily unloaded → This is normal condition of target VDM Server running rootfs as rawfs—changes to “actual=mounted” when replication is running and rootfs changes to uxfss
```

**2. VERIFY NEW VDM:**

# **nas\_server -vdm -l**

```
id    acl server mountedfs  rootfs name
1     0    1          25    vdm_tm
```

# **nas\_fs -l**

```
25    y  1  0  155  root_fs_vdm_1  1
#/nas/sbin/rootnas_fs -i root_fs_vdm_1
```

```
id      = 25
name    = root_fs_vdm_1
acl     = 0
in_use  = True
type    = uxfss
```

volume = v15

### # /nas/sbin/rootnas\_volume -i v155

id = 155

name = v155

acl = 0

in\_use = True

type = meta

volume\_set = s87

disks = d9

clnt\_filesys= root\_fs\_vdm\_1

### # /nas/sbin/rootnas\_slice -i s87

id = 87

name = s87

acl = 0

in\_use = True

slice\_of = v135

**offset(MB)= 128** →Default size VDM rootfs also 128MB, as is data mover rootfs

size (MB)= 128

volume\_name = s87

### # server\_cifs server\_2 [output abridged]

---

CIFS service of VDM vdm\_tm (state=loaded)

Home Directory Shares DISABLED

## **3. CREATE MOUNTPOINT AND MOUNT FILE SYSTEM TO VDM:**

### # server\_mountpoint vdm\_tm -c /fs1

vdm\_tm : done

### # server\_mount vdm\_tm fs1 /fs1

vdm\_tm : done

### # server\_mount vdm\_tm

vdm\_tm :

fs1 on /fs1 ufs,perm,rw

## **4. CREATING CIFS SERVERS WITHIN A VDM CONTAINER:**

### # server\_cifs vdm\_tm -add compname=virtuous, domain=mouse.com, interface=47

vdm\_tm : done

**Note:** Default CIFS server is not compatible with VDMs. Each CIFS server of a VDM must have its own interface defined.

#### **SERVER LOG:**

2004-06-07 15:56:21: SMB: 4:[vdm\_tm] CIFS Server VIRTUOUS[] created (0)

2004-06-07 15:56:21: SMB: 4:[vdm\_tm] Full computer name virtuous.mouse.com, Realm MOUSE.COM

2004-06-07 15:56:22: ADMIN: 4:[vdm\_tm] Command succeeded: :1 cifs add compname=VIRTUOUS domain=MOUSE.COM interface=47

## **5. JOINING CIFS SERVERS TO DOMAIN WITH VDM CONTAINER:**

### # server\_cifs vdm\_tm -Join compname=virtuous, domain=mouse.com, admin=administrator

vdm\_tm : Enter Password:\*\*\*\*\*

done

#### **SERVER LOG:**

2004-06-07 15:59:13: KERBEROS: 4:[vdm\_tm] Upgading file krb5.conf from version 0 to verion 2

2004-06-07 15:59:13: SMB: 4:[vdm\_tm] DomainJoin::getAccountGuid: Account GID:12c20259-d604-4719-951c-adfa6d944b69

2004-06-07 15:59:13: ADMIN: 4:[vdm\_tm] Command succeeded: :1 domjoin compname=virtuous domain=mouse.com

admin=administrator password=23233D193D37252D init

2004-06-07 15:59:13: SMB: 4:[vdm\_tm] >DC=MICKEY(10.241.169.16) R=5 T=0 ms S=0,1/-1

2004-06-07 15:59:13: USRMAP: 4:[vdm\_tm] Usermapper[10.241.169.43] now available

## **6. CREATING CIFS SHARE ON CIFS SERVER WITHIN VDM CONTAINER:**

### # server\_export vdm\_tm -P cifs -n Virtuous /fs1

vmd\_tm : done

### # server\_export vdm\_tm

vmd\_tm :

share "Virtuous" "/fs1" maxusr=4294967295 umask=22

**Note:** Cannot stop/start CIFS Service on a VDM, only on the physical Server\_2

## **7. VERIFYING VDM CIFS CONFIGURATION:**

### # nas\_server -info -vdm vdm\_tm

```
id      = 1
name    = vdm_tm
acl     = 0
type    = vdm
server   = server_2
rootfs  = root_fs_vdm_1
I18N mode = UNICODE
mountedfs = fs1
member_of =
status   :
defined = enabled
actual = loaded, active
Interfaces to services mapping:
interface=47 :cifs
# server_cifs vdm_tm
vmd_tm :
96 Cifs threads started
Security mode = NT
Max protocol = NT1
I18N mode = UNICODE
CIFS service of VDM vdm_tm (state=loaded)
-----output abridged-----
CIFS Server VIRTUOUS[MOUSE] RC=2
Full computer name=virtuous.mouse.com realm=MOUSE.COM
Comment='EMC-SNAS:T5.2.14.0'
if=47 l=10.241.169.47 b=10.241.169.255 mac=8:0:1b:42:38:24
FQDN=virtuous.mouse.com (Updated to DNS)
```

## **INSPECTING CIFS CONFIGURATION WITH VDM CIFS SERVERS:**

**Note :** Must inspect several different files and outputs in order to determine the layout of CIFS vs. VDM-CIFS Servers

### **\$server\_cifs server\_2**

```
DOMAIN DOM_1 RC=4
SID=S-1-5-79057000-1e31512e-2e75ae2-ffffffff
DC=LOGON-SERVER2(141.90.95.26) ref=1 time=0 ms
>DC=LOGON3(141.90.95.105) ref=4 time=0 ms
CIFS Server TRENDMICRO[DOM_1] RC=2 →Default CIFS Server for Server_2
Comment='EMC-SNAS:T5.3.10.4'
if=net4 l=141.90.95.115 b=141.90.95.255 mac=8:0:1b:42:64:fb
```

### **CIFS service of VDM vdm1 (state=loaded) →First VDM Container for Server\_2**

```
Home Directory Shares DISABLED
Default WINS servers = 141.90.94.106
DOMAIN DOM_1 RC=8
SID=S-1-5-79057000-1e31512e-2e75ae2-ffffffff
>DC=LOGON-SERVER2(141.90.95.26) ref=3 time=140 ms
>DC=LOGON3(141.90.95.105) ref=30 time=0 ms
```

### **CIFS Server CLUSTER1[DOM\_1] RC=14 →CIFS Server for VDM1**

```
Comment='EMC-SNAS:T5.3.10.4'
if=net2 l=141.90.95.113 b=141.90.95.255 mac=8:0:1b:42:64:fb
```

### **CIFS Server CLUSTER2[DOM\_1] RC=10 →CIFS Server for VDM1**

```
Comment='EMC-SNAS:T5.3.10.4'
if=net0 l=141.90.95.111 b=141.90.95.255 mac=8:0:1b:42:64:f9
```

## **NETD FILE CONTAINS CIFS SERVER INFO:**

### **\$ cat /nas/server/slot\_2/netd**

```
cifs add usermapper=127.0.0.1
cifs add wins=141.90.94.106
cifs add netbios=TRENDMICRO domain=DOM_1 interface=net4
```

## **VDM FILE CONTAINS INFO FOR VDM's:**

### **\$ cat /nas/server/slot\_2/vdm**

### **SEE RESPECTIVE VDM.CFG FILE FOR INFO ON CIFS SERVICES FOR VDM's:**

**\$cat /nasmcd/quota/slot\_2/root\_vdm\_1/etc/vdm.cfg**

```
cifs add netbios=CLUSTER1 domain=DOM_1 interface=net2  
cifs add netbios=CLUSTER2 domain=DOM_1 interface=net0
```

**CIFS VDM CONFIGURATION FILES:** /nas/server/vdm/<vdm\_name>

**# server\_cifs vdm\_1 -add compname=vdm\_1, domain=w2k.pvt.dns, interface=vdm -comment vdm server**

```
vdm_1 : done
```

**# server\_cifs vdm\_1 -Join compname=vdm\_1, domain=w2k.pvt.dns, admin=administrator**

```
vdm_1 : Enter Password:*****
```

```
done
```

**# server\_cifs vdm\_1**

```
-----output abridged-----
```

CIFS Server VDM\_1[W2K] RC=2

Full computer name=vdm\_1.w2k.pvt.dns realm=W2K.PVT.DNS

Comment='hello vdm'

if=vdm l=192.1.4.223 b=192.1.4.255 mac=0:60:16:c:51:46

FQDN=vdm\_1.w2k.pvt.dns (Updated to DNS)

Password change interval: 0 minutes

Last password change: Tue Jun 10 12:40:15 2008 GMT

Password versions: 2

**# cat /nas/server/vdm/vdm\_1/export.shares**

```
# 1213101225
```

share "vdm\_share" "/vd़m1" umask=022 maxusr=4294967295

share "vdm\_comments" "/vd़m1" umask=022 maxusr=4294967295 comment="Sample comments for CIFS VDM Server"

**# cat /nas/server/vdm/vdm\_1/vdm.cfg**

```
cifs add compname=VDM_1 domain=W2K.PVT.DNS interface=vdm comment="hello vdm"
```

### **CREATING VDM ON SPECIFICALLY CONFIGURED FILE SYSTEM:**

**1. \$ nas\_slice -n svdm2 -c d7 512** [Default size of VDM root is 128MB when configured by system]

**2. \$ nas\_fs -n fsvdm2 -c mvdm2**

**3. \$ nas\_server -name vdm2 -type vdm -create server\_2 -setstate loaded -fs fsvdm2**

**Note:** Must create a VDM root file system with this method for Clariion backends

### **VDM STATES:**

**\$ nas\_server -i -v vdm2**

```
id      = 6  
name    = vdm1  
acl     = 0  
type    = vdm  
server  = server_2  
rootfs  = root_fs_vdm_6  
I18N mode = UNICODE  
mountedfs =  
member_of =  
status   :  
  defined = enabled  
  actual = loaded, ready
```

Interfaces to services mapping:

### **MOVING VDM FROM ONE DM TO ANOTHER:**

**\$ nas\_server -vdm vdm02 -move server\_4**

### **CREATING VDM IN LOADED STATE:**

**1.) \$ nas\_server -name vdm01 -type vdm -create server\_2 -setstate loaded -fs vdm\_root\_fs1**

**Note:** Loaded state means a normally running CIFS server with configuration mounted Read-Write with Cifs Service running. In this example, the VDM root filesystem is created on the file system specified.

2.) Create mountpoint and mount file system

```
$server_mountpoint vdm01 -create /vdmfs01           $server_mount vdm01 vdmfs01 /vdmfs01
```

3.) Create CIFS service and Join to Domain:

```
$ server_cifs vdm01 -add compname=vdm01, domain=vdm01.network.com, interface=ace0  
$ server_cifs vdm01 -Join compname=vdm01, domain=vdm01.network.com, admin=administrator
```

4.) Create Shares:

```
$ server_export vdm01 -P cifs -n vdm01 /vdmfs01
```

#### **VDM IN UNLOADED STATE(temp unmount of vdm root & permanent unmount of vdm root):**

```
1.) $ nas_server -vdm vdm01 -unload
```

```
2.) $ server_umount vdm01 -all
```

**Note:** Root filesystem is unmounted and CIFS Service is not running

#### **CHANGING VDM STATE FROM LOADED TO MOUNTED:**

```
# nas_server -vdm vdm_tm -setstate mounted
```

```
id      = 1  
name    = vdm_tm  
acl     = 0  
type    = vdm  
server  = server_2  
rootfs  = root_fs_vdm_1  
I18N mode = UNICODE  
mountedfs = fs1  
member_of =  
status   :  
    defined = enabled  
    actual  = mounted
```

Interfaces to services mapping:

#### **LOADING VDM SERVER ON SERVER\_2:**

```
# nas_server -vdm vdm_1 -setstate loaded server_2
```

```
id      = 1  
name    = vdm_tm  
acl     = 0  
type    = vdm  
server  = server_2  
rootfs  = root_fs_vdm_1  
I18N mode = UNICODE  
mountedfs = fs1  
member_of =  
status   :  
    defined = enabled  
    actual  = loaded, active
```

Interfaces to services mapping:

#### **VDM IN MOUNTED STATE:**

```
$nas_server -name vdm01 -type vdm -create server_2 -setstate mounted -fs vdm_root_fs1
```

**Note:** Use this mode when CIFS services do not need to be running, such as for Celerra Replication on Secondary side.

Configuration file system is mounted read-only. Default config file system size is 128MB, but can be created larger. CIFS Services are not running.

#### **DELETING VDM SERVER:**

```
$ nas_server -vdm vdm01 -setstate loaded | tempunloaded | permunloaded server_2
```

```
# server_umount vdm1 -p -all
```

```
# nas_server -vdm vdm1 -setstate permunloaded server_2
```

```
id      = 6  
name    = vdm1  
acl     = 0  
type    = vdm  
server  =  
rootfs  = root_fs_vdm_6  
I18N mode = ASCII  
mountedfs =  
member_of =  
status   :
```

defined = enabled  
actual = permanently unloaded

Interfaces to services mapping:

**Note:** This operation will delete the VDM Server and associated vdm\_rootfs

### **DELETING & CHECKING STATUS OF VDM:**

# nas\_server -info -vdm vdm1

# nas\_server -list -vdm

```
id    acl  server  mountedfs   rootfs  name
1     0    1        37       vdm1
2     0    1        43       vdm2
```

# nas\_server -delete vdm1

### **VERIFYING SIZE OF VDM ROOTFS:**

# /nas/sbin/rootnas\_fs -s root\_fs\_vdm\_6

total = 128 (sizes in MB)

### **DELETING VDM ROOTFS:**

# /nas/sbin/rootnas\_fs -d root\_fs\_vdm\_6

```
id      = 35
name    = root_fs_vdm_6
acl     = 0
in_use  = False
type    = uxf
volume  = v137
rw_servers=
ro_servers=
rw_vdms =
ro_vdms =
symm_devs = APM00040303779-0012
disks   = d9gg
```

### **RENAMING & MOVING VDM:**

\$nas\_server -rename vdm01 vdm001

\$nas\_server -vdm vdm01 -move server\_3

### **NDMP BACKUP SUPPORT:**

NDMP version 4 support introduced

Multi-threaded PAX (Portable Archive Xchange) enhancements made to increase BackUp/Restore performance

**Note:** Goal is to meet NetApps 104GB/hour backup with (2) tape drives

### **WINDOWS 2003 SUPPORT:**

Support for SMB Signing for both Kerberos and NTLMSSP

Kerberos over TCP

Joining Windows 2003 Domains as member servers [Windows Server 2003, Windows Server 2003 family interims, Windows Server 2003 .NET]

### **REDUCED PERMISSION JOIN:**

Prior to 5.2, only a person who was a member of the Domain Admins group in the domain to be Joined could Join a DM At 5.2.14.100 and higher, anyone with the appropriately delegated rights to Join a Server can Join the DM to the Domain

#server\_cifs server\_2 -Join compname=dm1, domain=domain.com, admin=delegateduser@domain.com

### **PRE-CREATED COMPUTER ACCOUNT JOINS:**

Authority to Join a ‘pre-created’ computer account can now be allowed to a User or Group from different domains in the same forest

#server\_cifs server\_2 -Join compname=dm1, domain=domain.com, admin=admin -option reuse

### **JOINING WHEN DNS AND WINDOWS DOMAIN NAMES ARE DIFFERENT (Disjointed Namespace)**

**Note:** Original 4.0 code used a “dns” switch to allow for a ‘dis-jointed’ DNS to AD Domain—later revisions removed this switch NAS 5.2 re-introduces the option as follows to allow for successful Joins when DNS and AD Kerberos Realm names are different

\$server\_cifs server\_2 -add compname=dm1, domain=domain.com, interface=cge0, dns=bigcompany.com

\$server\_cifs server\_2 -Join compname=dm1, domain=domain.com, admin=administrator

**Note:** Delegated users must still belong to an Admin Group to use this feature

## **EXAMPLE SYNTAX FOR DISJOINED NAMESPACE JOIN:**

**\$ server\_cifs server\_2 -Join**

**compname=schnas01.rmd.navistar.com, domain=ad.navistar.com, admin=yytxg**

## **CREATING CIFS SERVICE WITH DNS DOMAIN AND KERBEROS ONLY AUTH:**

**\$ server\_cifs server\_2 -add compname=BYCFS1DM2A03839,**

**domain=IUSER.IROOT.ADIDOM.COM, interface=DM2\_fsn0\_if1, dns=nat.bt.com, authentication=Kerbeos**

### **NAS 5.2.14 JOIN OPTIONS:**

**param cifs djAddAdminToLg=0** [Set this value to 1 to enable]

**Note:** Use this parameter if you want the user performing the Join to be added to default local Administrator's group on DM. You would want to use this switch if the User is not a Domain Admins account [i.e., is a reduced permission user] and needs to be able to "manage" the compname after the Join.

**param cifs djUseKpasswd=1** [Code default is 1 for using Kerberos kpasswd protocol to set/change server passwd; Setting this value to 0 will enable Joins to succeed for Reduce Permission Users by using the SAMR authentication for setting machine password]

**param cifs djEnforceDhn=1** [Set this value to 0 to prevent DNS updates to Windows domain]

### **EXPANDED NOTE ON JOIN FAILURES WITH REDUCED PERMISSION ACCOUNTS:**

Normal Join process with W2K/W2K3 & CIFS servers is to set the machine account password with the DC at the end of the JOIN using the KPASSWD utility. Any operation using the "server\_cifs -J" process uses the same JOIN function, and will require KPASSWD to be executed at the end of the JOIN. In this situation, the KPASSWD operation is being attempted over UDP which has a datagram size limit of 2K. The security information for the user exceeds this 2K limit due to the number of groups of which the user is a member, causing the user authentication to fail. DART does not currently support the use of TCP for the KPASSWD protocol and therefore cannot recover from an oversized KPASSWD UDP packet. To correct this, modify the djUseKpasswd param to enable SAMR over TCP instead of KPASSWD.

#### **Example of Error Seen for Failed Join:**

Error Code / Message: krb5\_set\_password: krb5\_rd\_setpw\_rep returned

## **DISJOINED NAMESPACE/REDUCED PERMISSIONS JOIN/MANAGE PROCEDURE:**

Step 1. Ensure data movers are running NAS 5.2.14.100 or higher

Step 2. Ensure data movers have the following parameters set and have been rebooted:

**param cifs djAddAdminToLg=1**

**Note:** Adds User account used in Join to local Administrators group on Data Mover for delegated management for non-Domain Admin users--required for delegated management of Compname.

**param cifs djUseKpasswd=0**

**Note:** By setting value to 0, use of Kpasswd protocol to set CIFS server password is disabled, and SAMR authentication (MS\_RPC) is used instead, which allows a User with Reduced Permissions, to Join the compname and set the machine's password.

**param cifs djEnforceDhn=0**

**Note:** When set to 0 prevents DNS updates to Windows domain during Join process

Step 3. Add the DNS and NTP Services to the Data Mover

**\$server\_dns server\_2 -p tcp foobar.world.com 10.241.169.16**

**\$server\_date server\_2 timesvc start ntp 10.241.169.16**

Step 4. Add the NIS or Usermapper client service to the Data Mover, but not both (NIS & Usermapper combined is not allowed)

**\$server\_cifs server\_2 -add usermapper=10.241.169.43**

**\$server\_nis server\_2 hosts.pvt.dns 10.241.168.21**

Step 5. Created two Windows Users called "Reduced" & "Reduced2" and added to a new Domain Local Group called "Reduced Manage"

Step 6. Created Celerra OU's called GCM & Servers

Step 7. Rightclicked on each new OU from within ADUC Interface and selected "Delegate Control" wizard to grant manage rights to the Domain Local Group "Reduced Manage", and repeated for each of the two Users "Reduced" and "Reduced2" as well.

---

Delegate Control>Next>Add>add Users and Groups>Tasks to Delegate, selected 'Custom...'>Selected 'Delegate control of This folder, existing objects in this folder, and creation of new objects in this folder'>Permissions, selected all three checkboxes and granted 'Full Control' to complete delegated control.

Step 8. Manually created CIFS Compnames "Mini" and "Mini2" in OU GCM>Servers and granted the following right to the "Reduced Manage" group via the wizard:

"The following user or group can join this computer to a domain. User or group: Reduced Manage"

Step 9. Manually created Forward and Reverse Lookup Zone entries for Host and PTR records respectively, for ‘Mini’ & ‘Mini2’

Step 10. Created CIFS Service for each compname:

**\$server\_cifs server\_2 –add compname=mini, domain=mouse.com,interface=cge0,dns=foobar.world.com**

**Note:** Use “dns=” syntax only when the FQDN of the Windows domain is different from the DNS domain

**\$server\_cifs server\_2 –add compname=mini2, domain=mouse.com,interface=cge0,dns=foobar.world.com**

Step 11. Started CIFS Service on Data Mover

Step 12. Joined each Compname to the Domain (Joins do not map a User’s Group SIDs, only their User SID):

**\$server\_cifs server\_2 –Join**

**compname=mini.foobar.world.com, domain=mouse.com, admin=reduced, ou=”ou=Servers:ou=GCM”**

**\$server\_cifs server\_2 –Join**

**compname=mini2.foobar.world.com, domain=mouse.com, admin=reduced2, ou=”ou=Servers:ou=GCM”**

**Note:** Join command adds the User to the Administrators localgroup for the Compname—only a single User can be added during the Join. Manually add additional Users to Administrators localgroup file from a privileged account such as Domain Admins, or logged in as the User already added to the group. Above syntax for ‘compname=’ is required to denote the different DNS domain, while the ‘domain=’ entry denotes the Windows Fully Qualified domain name. Also note that CIFS output shows “Update of “A” record failed during update...”—this is normal message since the DNS entries were added manually to the FWD and Reverse Lookup Zones.

Step 13. Verify access and management rights to new Compnames

- a.) Log into AD Domain as user ‘Reduced’, open Computer Management>Connect to another Computer>\\\mini
- b.) Add additional users as required to local Administrators group and save

**Caution:** Users and Groups found in AD can be successfully added to the local Administrators group, but keep in mind that unless the Users and respective Groups are mapped via Usermapper or NIS, they will not have the ability to ‘manage’ the compname. In otherwords, Users and Groups can be added to the data mover’s localgroups file without being mapped with a UID/GID. This is why a delegated user might be able to Join a compname to the domain successfully but might not be able to ‘manage’ the compname after the Join—the User’s Group memberships are not mapped until attempting to ‘manage’ the compname for the first time. In otherwords, a Join never maps a User’s Group memberships.

### **TEARING CIFS DOWN BETWEEN TESTS:**

1. Stop CIFS service
2. Delete CIFS configuration: \$server\_setup server\_2 –P cifs –o delete
3. Delete computer accounts from AD (if repeating tests immediately, then use new Compnames for testing)
4. Delete DNS entries from Forward and Reverse Lookup Zones for compnames
5. If Usermapper running, stop, disable, delete records if required, remove pointer to Usermapper from CIFS, if required
6. Delete SecMap cache directory: #rm –Rf secmap
7. Delete all krb5\* files from /.etc directory on Data Mover
8. Delete all .db.localgroups, .db.localgroups.bak, localgroups.db files from /.etc directory
9. Go to /nas/server/slot\_2 directory and delete all krb5\* and .db.localgroups files
10. Reboot data mover and verify that no entries exist in the /.etc/secmap or /.etc/usrmapper directories or in any of the above-mentioned files in /nas/server/slot\_2
11. Build new configuration for next series of tests

### **SMB SIGNING CONFIGURATIONS:**

1. **param cifs smbsigning=1** 10 disables [with parameter set, enables DM Server & Client signing and overrides GPOs]

**Note:** Enabled by default with NAS 5.2. Disable this parameter if joining a Windows 2000 domain that does not have SMB Signing enabled.

2. GPO Settings apply to all machines in domain, controlling both Server & Client signing

Default Domain Security Settings>Security Settings>Local Policies>Security Options

--Microsoft network client: Digitally sign communications (always) →Disabled by default

--Microsoft network client: Digitally sign communications (if server agrees) →Enabled by default

--Microsoft network server: Digitally sign communications (always) (Server requires signing)→Disabled by default

--Microsoft network server: Digitally sign communications (if client agrees) →Disabled by default

3. In lieu of GPO settings, Windows systems use Server/Client-side signing with Registry entries

### **SERVER-SIDE SIGNING:**

HKLM>System>CurrentControlSet>Services>lanmanserver>parameters:

enablesecuritysignature 1 [Enables]

requiresecuritysignature 1 [Requires that SMB signing be used]

### **CLIENT-SIDE SIGNING:**

HKLM>System>CurrentControlSet>Services>lanmanworkstation>parameters:

enablesecuritysignature 1 [Enables]

requiresecuritysignature 1 [Requires that SMB signing be used]

## **TERMINOLOGY:**

Disabled—Client or Server does not support SMB signing

Enabled—Client or Server supports SMB signing but is not required for all transactions

Required—Client or Server requires SMB signing for all transactions

## **NAS 5.3 BORDEAUX:**

--NAS 5.3.10.4 GA Aug 4, 2004, supporting CNS-14, CFS-SE, Celerra NS, and Celerra NSG systems

--Support for NAS High and midrange CNS, CFS, NS600/G, NS700/G, using NAS 5.2 as base code

--Current release 5.3.24.200, compatible with Flare 16 & 19

**Restrictions:** NS500, NS500G, & NS704G requires use of NAS 5.3 as minimum NAS code, while 506 CM/CS Models and below cannot be used with 5.3

## **Features:**

Snapsure enhancements, VSS Integration; Replicator enhancements, Replication Silvering, etc.; Home Dir enhancement; Quotas, ACL Tools; On Demand AV; CSV Reporting, etc.; New TFTP protocol to be included; Variable length subnet masks; NSB-4 Support, Exterminator, CNS TOE support [TCP Offload Engines on NIC cards]; ILM Data Protection; iSCSI Exchange; LDAP iPlanet; iSCSI Target; Celerra Manager to have enhanced statistical displays, as well as capability to export the stats

## **MICROSOFT SHADOW COPY FOR SHARED FOLDERS (SCSF):**

→Support for Windows 2003/XP Clients to list, view, copy, restore files from SnapSure Checkpoints

→Requires installation of SCSF software on XP Client, not Win2k3

→Celerra support enabled by default

→No support for Windows 2000 Clients

→Not the same thing as MS VSS

## **ISSUES:**

XP/Win2k client stations hang when trying to access CVFS checkpoints on 5.4.19.5, when traversing NMFS structure—AR74087

## **iSCI SUPPORT:**

--iSCSI is a transport protocol for sending SCSI packets over TCP/IP networks

--iSCSI requires iSCI Host Initiator to encapsulate SCSI commands, data, status and send to iSCSI Target (storage device)

--Two topologies are Native iSCSI (Celerra is a native iSCSI solution—uses no Fibre Channel) and Bridging iSCSI (uses Ethernet & Fibre Channel—Servers attach to network using Ethernet and Storage devices connect via Fibre Channel)

--iSCSI targets can be configured on DM using either CLI or Celerra Manager

## **Celerra supports following iSCSI Components:**

--stopping & starting of iSCSI service on DM

--Creation of iSCSI targets to include portal groups & LUN masking

--Creation & management of iSCSI LUNs—each LUN emulates a SCSI disk device

--Support for IP Naming Service (iSNS) Client on Data Mover, a service that includes Naming, Discover, Login

--Support for CHAP authentication by iSCSI Initiators when challenged by iSCSI targets before access is granted to LUNs

**Note:** `param iscsi RequireChap=0` (Chap is off by default=0)

## **Support for iSCSI Host-based Solutions:**

(Solutions that use Celerra iSCSI targets that add iSCSI snap capabilities for backing up and restoring data stored on iSCSI LUNs)

**iSCSI SnapSure Manager for Exchange 2000** (stores Exchange system/log files & storage groups on separate iSCSI LUNs)

**Note:** Functionality includes ability to restore or delete Storage Groups/log/system files, promote storage groups for mailbox restores, demote storage groups, schedule group snaps

**SnapSure Manager for iSCSI**—point-in-time iSCSI snaps of iSCSI LUNs using Windows GUI

**Note:** Functionality includes creation, deletion, restoration of iSCSI luns, promotion or demotion of iSCSI snaps, displaying snaps of iSCSI luns

## **FILE SYSTEM LINKING:**

--Ability to access several file systems from Single share using UNIX absolute symbolic links created by 'Root'

**param shadow followabsolutpath=1** [Set this param to enable file system linking and reboot DM]

## **SETTING UP FILE SYSTEM LINKING TO SOLARIS:**

1. Create & Mount file systems

2. Export each fs using "root=192.168.25.140"

3. Set param and reboot Server

4. On Solaris, mkdir /fs1 and /fs2

5. From Solaris, #mount 192.168.25.122:/fs1 /fs1 and /fs2

6. From Solaris, #mkdir /fs1/subdir1 /fs2/subdir2 and copy data to /fs2/subdir2

7. Create Symbolic link from /fs1/subdir1 to /fs2/subdir2

**#cd /fs1/subdir1;ln -s /fs2/subdir2 subdir2**

8. Verify link in /fs1/subdir1:

**#ls -l [subdir2 → /fs2/subdir2]**

10. Test access to Link in Windows using Explorer [\dm2\fs1\ subdir2\kernel]

#### **ERROR CONDITIONS:**

- If target does not exist, or param is not set, directories are seen as files
- If target is a link and we enter a loop, the link is seen as a file
- If symbolic link has not been created by ‘root’, link might be seen as a file

#### **TFTP PROTOCOL SUPPORT:**

Text or binary transfers using get or put commands via UDP Port 69, up to 33MB total for any single file transfer

TFTP uses UGO ‘Other’ permissions for Read or Write

Secure TFTP drops Client into specific toplevel directory without cd capability

#### **CONFIGURING CELERRA TFTP:**

**\$server\_tftp server\_2 -service -start | -stop**

**\$server\_tftp server\_2 -set -path /fs1 -readaccess all -writeaccess -all** [Define access to directories]

**\$server\_tftp server\_2 -service -status | -stats**

**\$ server\_tftp server\_2 -service -status**

server\_2 :

Tftp Running

**\$ server\_tftp server\_2 -service -stats**

server\_2 :

Attempted Transfers:5363

Successful Transfers:1890

createdthrds:5363

deletedthrds:5363

timedoutthrds:20

TotalBinds:5363

TotalUnbinds:5185

BindFailures:416

InvalidAttempts:3258

AttemptedReadTransfers:5363

SuccessfulReadTransfers:1890

AttemptedWriteTransfers:0

SuccessfulWriteTransfers:0

**\$server\_tftp server\_2 -info | -clear**

Export filesystem for remote access

Create temporary /tftp directory on Unix host

**Note:** Only one path can be defined by default—use **param tftp pathno=10** to increase number of paths from 1 to 10. TFTP supports up to 16 concurrent connections.

#### **Transferring From Unix Client:**

#tftp –i 192.168.25.122 [Use –i for binary files]

tftp> bin

tftp> get filename

tftp>quit

#### **Transferring From Windows Client:**

C:>tftp 192.168.25.122 get passwd

### **CELERRA FILEMOVER/CELERRA DHSM (Distributed Hierarchical Storage Management):**

NAS 5.3 API version 1.0

NAS 5.4 API version 1.1

NAS 5.5 API version 1.2 with HTTP v.1.1 XML digest authentication, and BAR (Bulk Attribute Retrieval)

NAS 5.6 turns on Digest authentication by default

#### **SUPPORTED POLICY/MIGRATION ENGINES:**

##### **ARKIVIO auto-stor v4.3:**

→NAS 5.4.18 and above; ONTAP 7.0.6 or higher; Windows NT, 2000, 2003, 2008; Solaris 2.6, 7, 8, 9, 10; HP-UX 19, 11, 11i; IBM AIX 4.3.3, 5.x; Linux; Apple; VMWare

→Secondary File Servers running most versions of Centera, Celerra, NetApps, etc.

##### **ATEMPO Digital Archive 2.3:**

→NAS 5.6.43 or higher; see support matrix

##### **CENTERA FileArchiver v3.5:**

Centera is a data storage platform that uses a data access mechanism called CAS (Content Addressable Storage), which creates unique handles for each object stored—no directories, no pathnames, no URLs

→Centera Gen3/Gen4 or Dell 2850 as Policy Engine runs on Linux as CUA (Centera Universal Access)

→Supports NFS or CIFS

→Secondary storage Centera-only

→Does not support digest authentication for DHSM API

#### **CENTERA v4.0 CFS/CUA SUPPORT:**

Minimum NAS versions are 5.4.24.8 & 5.5.27.5.

**Note:** v4.0 Support information is not posted on the Celerra Matrix, only in the Centera PSB.

#### **DISKXTENDER FOR NAS v3.1:**

→Policy engine running on Windows 2003, Solaris 9 & 10, Redhat Enterprise 4.0

→NAS 5.5 or 5.6 required

#### **SYMANTEC ENTERPRISE VAULT v7.0 FSA:**

→NAS 5.5 or 5.6, refer to Symantec.com for information

#### **SMARTMOVE v.2.3 Enigma Data Systems (SmartMover 2.3.2):**

→NAS 5.2 + Windows 2003; Solaris 8, 9, 10; Linux RedHat 3.x or 4.x; SUSE Linux 8 or 9; Supports FLR

#### **RAINFINITY FILE MANAGEMENT APPLIANCE (FMA) VERSION 7.2 (& 7.3):**

→IP-based File level archiving solution

→Offerings are FMA, FMA-HA (High Availability), and GFV (Global File Virtualization)

→Data Movers communicate using HTTP and resolving FQDNs via DNS. Celerra can directly recall data from storage using CIFS or NFS protocols. If on Centera storage, the EMC Centera SDK is called.

#### **DHSM FILEMOVER INFORMATION:**

→NAS 5.3.12.0 supports primary & secondary DHSM file systems on same Data Mover using external interface

→Recommended NAS 5.3.21.2 for 5.3 and 5.4.18.3 for 5.4

→DHSM is a feature that uses an external Policy Engine (Windows Enigma Server or FMA) to migrate infrequently used files from primary storage (Celerra) to less-expensive hierarchical secondary storage, and then set the file ‘offline’, meaning that it becomes a stub file, with the absolute path to the Secondary Storage defined in the stub inode

**Note:** Enigma Server reads files from Celerra & transfers to secondary storage, leaving “stub” file on Celerra. Stub files contain attributes & metadata. Clients access “stub” files and Celerra will read file from secondary server.

--DHSM migrates files based on type, size & access times, migrates only regular files, and only data, not metadata

--Migrated files are converted to stub files, held on the Primary storage in order to recall or read from secondary storage when requested from NFS, CIFS, or FTP clients

--DHSM is part of ILM (Information Life-Cycle Management) for Migration of data

--another DHSM product is Arkivio Auto-Stor

#### **CELLERRA FILEMOVER/DHSM RULES:**

→Must use Celerra as primary storage in order to have a Celerra as part of a secondary storage solution

→FileMover can support a two hop configuration, but Policy Engines may not yet support

→FileMover requires I18N Data Mover

→Port 5080 must be opened on network between DM and Policy Engine

→Default HTTP threads = 20, with 15 threads dedicated to FileMover API (up to 15 simultaneous FileMover API requests from Policy Engine)

→HTTP protocol is used to communicate between Policy Engine and Celerra FileMover API, for both NFS or CIFS connections, when converting file to stub file after the migration

→\$ server\_http\_server\_2 –modify dhsm –threads 30 (max number threads is 30)

→Every file that gets migrated requires the Policy Engine to initiate the migration

→But, files that are recalled do not involve the use of the Policy Engine, just the rules that have been setup on the storage systems

**Note:** The exception is if secondary storage is on Centera, which would require a Policy Engine to recall the file, since Celerra has no concept of the clip ID which Centera uses to store files. Celerra would provide the inode stub info to the Policy Engine, which would recall the file and deliver to Celerra.

→ There can be multiple FileMover connections for a primary file system, but each connection must go to a Secondary storage target

→If using offline folders, set –read\_policy\_override option to prevent files from being migrated back to primary storage

#### **DHSM ATTRIBUTES:**

--File migration most commonly based on size and access times

--File migration and management using Open standards to other Celerra Servers or File Servers

--Only regular files are migrated, not links or directories

#### **Stub Files:**

Files become stubs after migration—stub file resides locally on DM & contains metadata & attributes. Stub files store modification times (mtime) of files, name, permissions, timestamps, size in 8k, offline path, file size, user data string, Policy engine ID string, and Read recall policy. Orphan files are those files on Secondary storage that do not have Stub file data on Celerra.

- Uses CDMS building blocks for files offline
- HTTP uses 20 threads by default for DHSM API requests
- Uses distributed data but accessible from single point on Celerra [transparent data recall]
- File attributes and metadata are stored on Primary Store as a “stub” file, which uses 1 inode and consumes an 8k block (mtime)
- Policy-Based defined by Administrators
- Celerra VDMs can be used as Primary and Secondary Stores [use UNC paths]
- AntiVirus options can skip Offline file scanning or use Scan on First Read using pass-through mode
- HighRoad Clients use NFS or CIFS protocol for Offline files
- Quotas uses additional inode for offline and uses pass-through reads when reaching Quota Limits
- SnapSure defaults to pass-through reads
- Do not allow quotas or file system to become full as Server will revert back to ‘Passthrough Read’

**Note:** For migrated data, select migration policy to ‘Recall Data’ if the desire is to bring files ‘online’ or read data in ‘pass-through’ mode to leave offline

### **STUB FILE ATTRIBUTES:**

- Stub files reside on the Primary storage file system
- Contains offline path to Secondary storage for recall or pass-through
- Contains mtime of file and original file size, though stub file itself is only an 8k block
- Contains file name and permissions
- Contains Recall Policy and policy engine ID
- Contains User data string
- Windows Clients see offline file with clock graphic icon, and file sizes bracketed in DOS directory output [ ]
- CIFS Redirector timeout increases for Stub files from 45 secs to 1000 seconds

**Note:** Stub and Secondary file MTIME & File Size must be the same, or else the file will become orphaned

### **FILEMOVER API FUNCTION CALLS:**

- FMOpenSession: open a session with a FileMover server
- FMCloseSession: close a session
- FMGetApiAttrs: retrieve version and attributes
- FMGetFileAttributes: retrieve online and offline attributes
- FMSetOfflineAttributes: get secondary connection information
- FMGetErrorString: get string related to FileMover API error code

### **DHSM API CALLS:**

- DHSM\_GET\_API\_ATTRS →Query version and capabilities of API  
./get\_api\_attrs -user -password 10.241.169.13 /
- DHSM\_GET\_ATTRS →Get attributes of a file  
./get\_attributes 10.241.169.13 /dm2fs1/dir1/pax.tar
- DHSM\_SET\_ATTRS →Convert online file to stub file after copying to secondary storage  
.set\_attributes -user dhsm\_user -p dhsm -v 1004503452 10.241.169.13 /dm2fs1/dir1/pax.tar nfs://192.168.10.30/nfsv3\_fs/pax.tar
- DHSM\_GET\_CONNECTION\_LIST →List of secondary connections (new with 5.4)  
.get\_connection\_list -u dhsm -p dhsmdhsm 10.241.16.13 cifs://cifs7dm2.w2003.celerra.com
- DHSM\_GET\_BULK\_ATTRIBUTES →Attributes of files/directories in a tree (NAS 5.5)—filter by Size, Timestamps, Type, State
- DHSM\_QUERY\_BULK\_RETRIEVAL →status of running BAR (Bulk Attribute Retrieval) job, up to 4 per DM
- DHSM\_ABORT\_BULK\_RETRIEVAL →abort running BAR job

**\$ .server\_config server\_2 -v “mgfs action=query fsid=100”**

**\$ .server\_config server\_2 -v “dhsm action=query\_fsoptions fsid=100”**

**\$ .server\_config server\_2 -v “dhsm action=modify\_conn”**

**\$ .server\_config server\_2 -v “dhsm action=query\_conn fsid=18”**

|                                                                                                                                                                                | Total KBytes | Used KBytes | Total inodes | Used inodes |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|-------------|--------------|-------------|
| 1146223007: MGFS: 4: Dart cid: 379624240                                                                                                                                       | 304644192    | 46310974    | 2695734      |             |
| 1146223018: MGFS: 4: -----                                                                                                                                                     |              |             |              |             |
| 1146223018: MGFS: 4: remote 0: 212099072                                                                                                                                       | 156467200    | 0           | 0            |             |
| 1146223018: MGFS: 4: type=CIFS state=2 rpolicy=15 wpolicy=1 cifs=\gb0008cs0086.uk.hibm.hsbc\ARKIVIO_CIFS\localServer=ukhibmdata04.uk.hibm.hsbc account=ukhbeu\NASArch passwd=* |              |             |              |             |

**\$ .server\_config server\_2 -v “cifsclient audit”** →Auditing Data Mover Client connections for DHSM

1134574620: SMB: 4:

Server:CLUSFS001NASTST.us.cly (10.55.9.92) Client:CLUSFS001NASTST.US.CLY (10.55.9.92)

1134574620: SMB: 4: NativeOS: EMC-SNAS:T5.3.20.1

1134574620: SMB: 4: NativeLanman: NT1

1134574620: SMB: 4: Status: Connected

1134574620: SMB: 4: Auth mode: kerberos

```
1134574620: SMB: 4: Capabilities: 8000e3fd
1134574620: SMB: 4: Encrypted password: Yes
1134574620: SMB: 4: Protocol: NT LM 0.12
1134574620: SMB: 4: Refcount: 3
1134574620: SMB: 4: List of sessions:
1134574620: SMB: 4: * Name: us.cly\hsadmin
1134574620: SMB: 4: uid: 3f
1134574620: SMB: 4: Vc: 0
1134574620: SMB: 4: RefCount: 3
1134574620: SMB: 4: List of connections::
1134574620: SMB: 4: * Name:zummy_archive
1134574620: SMB: 4: tid:: 3f
1134574620: SMB: 4: RefCount: 2
1134574620: SMB: 4: Service: (NULL)
1134574620: SMB: 4: Mounted by: us.cly\hsadmin
```

## **DHSM ARCHITECTURE:**

Primary Server Store = Celerra

NAS Clients access data via Celerra

Secondary Server Store = data migrated to a File Server

Policy Engine runs on CIFS or NFS Client, has access to both Primary and Secondary Server Stores

DHSM API's run XML over HTTP Port 5080 [DHSM\_SET\_OFFLINE\_ATTRS & DHSM\_GET\_ATTRS]

## **POLICY ENGINE:**

Searches Primary Store for files that match defined policies (rules) and transfers files to Secondary Storage and marks files offline [absolute path, attributes, and metadata info is stored in offline inode (Opaque Inode) on Primary Store--called stub files]

**Note:** CIFS default for accessing offline files is 1000 secs.

## **DHSM FILE READ RECALL POLICIES:**

**Passthrough:** Client accesses stubfile on Primary storage, Celerra recalls data from Secondary storage as a Read operation and passes to Client. Stub file is preserved and data is not recalled to Primary Storage. Read-Only file systems, such as checkpoints, use Passthrough Reads automatically.

**Full Recall:** Client access file on Primary storage, Celerra recalls entire file and writes back to Primary Storage, then passes to Client, and stub file is converted back to a normal file.

**Partial Recall:** Recalls only data needed by client, but moves data to Primary Storage. If whole file is Read, stub file reverts to a normal file.

**Note:** DHSM Policies are specified when the CIFS or NFS Connection is established. File System Read policies take precedence over Connection or Stub file policy settings, while Connection Read policies take precedence only over Stub file policies.

## **File Retrieval Process:**

- 1) Client request made to Celerra
- 2) Celerra reads stub file attributes
- 3) Celerra retrieves info on file from Secondary storage and compares mtime and file data size to verify the correct file
- 4) File is read back using Full, Partial, or Passthrough

**Note:** Default behavior is to recall all files to primary storage if the DHSM connection is deleted

## **DHSM WRITE MODIFY POLICY:**

**Modify Recall:** Complete file recalled from Secondary to Primary, Client modifies file on Primary.

→CIFS, NFS, FTP clients that try writing to an offline, or secondary storage file (modify), must first have the entire contents of the file recalled to local storage before the modification can be made

## **DHSM MIGRATION POLICIES:**

These policies are handled by the Policy Engine

## **ANTIVIRUS APPLICATIONS & DHSM BEHAVIOR:**

→Be aware that some Virus applications may update last access times on files, which may defeat purpose of DHSM policies in that files that are always updated will not migrate properly (Symantec with option ‘Skip offline files’ will skip stub files; McAfee has an option to scan migrated files, but does not work & will not scan such files)

→CAVA will skip offline stub files and not scan, therefore not change last access times

→McAfee automatic scan will preserve last access times—a manual scan, however, will update last access times, but not for stub files

→Symantec versions cause file recalls because of read behavior & requires special settings to address

    Disable Realtime Protection or Auto-Protect; clear the Opened for backup option; select Skip offline files option, etc.

→Write operations proceed normally

→With “Scan-on-Read” mode set, data is read in passthrough mode from AV Server—this setting is recommended or else the AV application may change the last access times on files

→Using “-fsscan” option, Stub files are not scanned by default. However, if ‘-fsscan –create offline’ option is used, CAVA will allow scanning of Stub Files from the AV Server in ‘readthrough’ mode.

**Note:** Stub File attributes are stored as an SMB Offline file attribute, which is set on all migrated files by default. If full file system scans are anticipated, EMC recommendation would be to set –read\_policy\_override to passthrough option for fs\_dhsm options.

## **BACKUP APPLICATIONS & DHSM BEHAVIOR:**

→CIFS performs backups by default using –backup passthrough as offline options, which backs up stub files as 0-byte files with extended acls, and does not recall the whole file from Secondary storage

→Be aware that some Backup applications can change the last access time, which may defeat the purpose of DHSM policies in that the file access times will always be updated [Veritas NetBackup & Legato NetWorker do not change last access times for stub files]

→General Unix backups would recall all files

### **CONFIGURING BACKUP OPTION ‘PASSTHROUGH’:**

**# fs\_dhsm -modify <fsid> -backup passthrough**

## **SPECIAL CONSIDERATIONS FOR NFS APPLICATIONS & DHSM:**

CIFS works with DHSM using DOS attributes called OFFLINE, or FILE\_ATTRIBUTE\_OFFLINE, for zero-byte stub files. NFS is not DOS aware and the following settings are recommended in order to ensure that stub files are not recalled inadvertently:

1. Change –read\_policy\_override option to passthrough, at filesystem or connection level, and stub files will not be recalled
2. Alternatively, create a Checkpoint and backup files from a Read-Only file system (reading of stub files defaults to passthrough)
3. Use NDMP backups, which handles stub files properly

→NDMP can perform offline backups of attributes or data using Variables [EMC\_OFFLINE\_DATA=Y or EMC\_OFFLINE\_DATA=N]

## **OTHER APPLICATIONS TO CONSIDER WHEN USING DHSM POLICIES:**

→By their nature, Web applications will recall files read—avoid use of web bots that scan websites—will cause large file recalls!

→Microsoft Utilities, CACLS & ATTRIB will not recall files—reads metadata; Robocopy requires /B switch to prevent file recall

→Unix find and touch Utilities do not recall files, but grep, head, tail will cause files to be recalled

→IP Replication would recall data if stub files are copied

## **VIEWING OFFLINE ATTRIBUTES FROM CIFS CLIENT:**

**# fs\_dhsm -modify cifs\_common**

cifs\_common:

state = enabled  
offline attr = on →EMC recommends keeping this enabled, otherwise Windows explorer can inadvertently recall files  
popup timeout = 0  
backup = passthrough  
read policy override = none  
log file = on  
max log size = 10MB

**# fs\_dhsm -modify cifs\_common -offline\_attr off** (turns off capability of viewing offline attributes for files)

cifs\_common:

state = enabled  
offline attr = off

## **FS DHSM COMMAND & TOOLS:**

**Note:** fs\_dhsm is used to create, delete, or modify NFS, CIFS, or HTTP connections to remote hosts. It also displays and configures DHSM properties on Celerra file systems.

**# fs\_dhsm** (DHSM management, status)

**# server\_http** (API security access)

**# /nas/sbin/server\_user** (user id to access policy engine)

### **\$/nas/tools/dhsm/get\_attributes**

**Note:** get\_attributes retrieves information on file attributes using the Data Mover DHSM API

### **\$ /nas/tools/dhsm/set\_attributes**

**Note:** set\_attributes can set a file offline using Data Mover DHSM API or change Retention State

**\$ ./set\_attributes -u rsadmin -p Stud3nt123 -w "0" server\_2 '/podE/LICENCE.TXT'**

## **DHSM COMMANDS:**

**# fs\_dhsm -list**

|    |            |
|----|------------|
| id | name       |
| 29 | fs01       |
| 30 | fs03       |
| 83 | fs01_snap1 |

**# fs\_dhsm -i fs01**

```
fs01:
state      = enabled
offline attr = on
popup timeout = 0
backup      = passthrough
read policy override = none
log file    = on
max log size = 10MB
```

**\$ fs\_dhsm -connection <file\_systemname> -info**

```
podE:
state      = enabled
offline attr = on
popup timeout = 0
backup      = passthrough
read policy override = none
log file    = on
max log size = 10MB
cid        = 0
type       = CIFS
secondary   = \\win08-e.rain.emc\share\
state      = enabled
read policy override = none
write policy = full
local_server = STORAGE.RAIN.EMC
admin       = rain.emc\rsadmin
wins        =
cid        = 1
type       = CIFS
secondary   = \\win03-e.rain.emc\share\
state      = enabled
read policy override = none
write policy = full
local_server = STORAGE.RAIN.EMC
admin       = rain.emc\rsadmin
wins        =
```

**# get\_attributes -u dhsm1 -p password 10.55.9.91 /root\_vdm\_1/userhome10/fs/adams2/presnotes.doc**

```
digest 10.55.9.91 dhsm1 password /tmp/input21797
issue_request: server IP 10.55.9.91
issue_request: datagram socket open
issue_request: bind successful
issue_request: connect successful
issue_request: local port = 55315, local addr = 10.55.9.91
issue_request: ha1=bb1af13d7105efe3493082425a2649fb
Sending 148 bytes ***
POST /dhsm HTTP/1.0
```

**USING get\_attributes TO VIEW FILE ATTRIBUTES:****\$ /nas/tools/dhsm/get\_attributes**

```
get_attributes [ -u user_name ] [ -p password ] [ -h handle ] [-w] [-a] [-S CA_cert_file] [-V http_version] primary_server primary_file
  a: Print ACLs for the file
  h: <handle> get the attributes by handle instead of <primary_file>
  w: Print WORM state of the file
  S: <CA_cert_file> Use SSL. Do server certificate verification with the Certificate authority certificates in <CA_cert_file>
  V: <http_version> use the specified version, HTTP/1.0 or HTTP/1.1. Default is HTTP/1.0
```

**\$ /nas/tools/dhsm/get\_attributes -u rsadmin -p Stud3nt123 -w server\_2 '/podE/LICENCE.TXT'**

**Note:** Following output shows an archived file with Stub Retention set

```
digest server_2 rsadmin Stud3nt123 /tmp/input7379
open_connection: server IP 128.221.252.2
open_connection: streaming socket open
```



**# server\_http server\_2 -info**

```
server_2 : done
DHSM FACILITY CONFIGURATION
Service name      : EMC File Mover service
Comment          : Service facility for getting DHSM attributes
Active           : False
Port             : 5080
Threads          : 16
Max requests     : 300
Timeout          : 60 seconds
ACCESS CONTROL
Allowed IPs      : any
Authentication   : none ,Realm : DHSM_Authorization
Allowed user     : everybody
```

## SSL CONFIGURATION

```
Mode            : OFF
Persona         : default
Protocol        : default
Cipher          : default
```

**Special Usage Commands for DHSM:****# export NAS\_DB\_DEBUG=1****# fs\_dhsm****special usage:**

```
-info [-format {default|parsable}]
|-modify [-state legacy] [-verify {on|off}] -Force
```

**VIEWING CIFS SHARE INFORMATION FOR DHSM CONNECTIONS:****# fs\_dhsm -c share -l**

```
id    name      cid
49    share      2
```

**# fs\_dhsm -c share -i 2**

```
share:
state      = enabled
offline attr = on
popup timeout = 0
backup      = passthrough
read policy override = none
log file    = on
max log size = 10MB
cid        = 2
type       = CIFS
secondary   = \\pplzna3a.osbprod.osfc.net\sharearch$\\
state      = enabled
read policy override = none
write policy = full
local_server = pplzna3a.osbprod.osfc.net
admin      = osbprod.osfc.net\eccadmin
wins       = v
```

**DEBUGGING DHSM CIFS CONNECTIONS:****\$ .server\_config server\_x -v “param NTsec logonTraces=99”****FINDING, DELETING, & RECREATING DHSM CONNECTION:****# .server\_config server\_x -v “queryCid action=onDiskCidQuery fsid=65”**

1197775248: MGFS: 4: ----- cid information -----

1197775248: UFS: 6: inc ino blk cache count: nInoAllocs 9: inoBlk a3338904

1197775248: MGFS: 4: cid=1 fsid=65 type=CIFS state=2 rpolicy=15 wpolicy=1 cifs=\\nasarch\_oeste.san.corp\VFS0030\_ARCH\localServer=vnasfs0012.gabp.bsch account=gabp.bsch\t000020 passwd=\*

**# .server\_config server\_x -v “disconnect fsid=65 cid=1”**

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
**# fs\_dhsm -c id=65 -type cifs -admin “gabp.bsch\\\admin” –secondary \\\\nasarch\_oeste.san.corp\\VFS0030\_ARCH –local\_server “vnasfs0012.gabp.bsch” –password exchange1**

## **OFFLINE MTIME IN STUBFILE DOES NOT MATCH REMOTEMTIME:**

The offline Mtime no longer matches the remoteMtime for many files and customer has users that cannot access certain files.  
Turn off Mtime verification via following:

**# fs\_dhsm -modify <<fs\_name>> -verify off**

## **CONFIGURING FILEMOVER/DHSM ON CELERRA:**

1. Create Primary and Secondary file systems of same size on different Servers & Export for CIFS

**# nas\_fs -n fs1 -c size=512M pool=symm\_std -o slice=y**

2. Create MD5 User Accounts on Server if Digest Access Authentication is going to be used for FileMover:

**# /nas/sbin/server\_user server\_2 -add -md5 -passwd dhsm\_user**

**Note:** Use unique UID/GID and min. password length 7 characters—password must be stored locally on DM for access

3. Configure access to FileMover API on Data Mover for Policy Engine IP Address:

**# server\_http server\_2 -append dhsm -hosts 192.168.25.45**

4. Configure digest access authentication to HTTP Server:

**# server\_http server\_2 -append dhsm -users valid -hosts 192.168.25.45,192.168.25.122,192.168.25.123**

**Note:** valid option allows all Users in passwd file to digest authenticate. If no Users are specified, digest authentication is turned off.  
Specify users via comma separated list for specific digest authentication.

5. Enable the Primary Storage file system for DHSM:

**# fs\_dhsm -modify fs1 -state enabled**

6. If not already done, setup CIFS on Primary and Secondary Servers

7. Create (CIFS/NFS) DHSM connection from the Primary File System to the Secondary File System:

### **CIFS CONNECTION:**

**# fs\_dhsm -connection fs1 -create -type cifs -admin ‘nas2000.lab\Administrator’ –secondary \\sec.celerra5.emc.com\\sec’ –local\_server dm2 –password nasadmin**

**Note:** If Primary & Secondary storage systems are in different Windows domains, then the –admin switch must be used. Do not use DFS share to create CIFS connection.

### **NFS CONNECTIONS:**

**# fs\_dhsm -connection fs1 -create -type nfsv3 –secondary 10.241.169.13:/nfsv3\_fs -proto UDP -useRootCred true | false**

**Note:** The default credentials used when accessing Secondary Storage via Stubfiles using NFS, are to use the credentials presented by the owner of the file when retrieving the files. When set to “true”, the credentials used are as “root” user.

8. Modify & Verify DHSM Connections:

**# fs\_dhsm -connection fs1 -modify 0 -read\_policy\_override partial**

**Note:** Implications here are that the –modify needs to be run after creating the connection if you need to specify Read & Write Recall policies

9. Configure HTTP server on DM to accept Celerra FileMover API connections from Host:

**\$ server\_http server\_2 -append dhsm -users <user\_name> -hosts <ip\_addr\_policy\_engine>**

**# fs\_dhsm -connection fs01 -info**

fs01:

```
state          = enabled
offline attr   = on
popup timeout  = 0
backup         = passthrough
read policy override = none
log file       = on
max log size   = 10MB
cid            = 0
type           = CIFS
secondary      = \\fei0018a030.hq.ferg.com\\DMC\
state          = enabled
read policy override = none
write policy    = full
local_server    = FEI0018S009.DS.FERG.COM
```

```
admin      = ds.ferg.com\backup
wins      =
cid       = 1
type      = CIFS
secondary = \\\fei0018s012.hq.ferg.com\ArchiveHome$\_
state     = enabled
read policy override = none
write policy = full
local_server = fei0018s009.ds.ferg.com
admin     = ds.ferg.com\backup
wins     =
```

**# fs\_dhsm -connection -info sec id=0**

**# server\_http server\_3 -info dhsm**

server\_3 :

DHSM:

users: <all valid users>  
allowed IPs: 172.16.105.180

threads: 20

9. DHSM Logging:

**# fs\_dhsm -modify fs1 -log on | off**

**CREATING CONNECTION FOR NFS:**

**# fs\_dhsm -connection fs01 -create -type nfsv3 -secondary nfs\_server:/dest\_path**

**MODIFYING CONNECTIONS:**

**# fs\_dhsm -connection fs1 -modify 0 -read\_policy\_override partial**

**DELETE CONNECTION:**

**# fs\_dhsm -connection fs1 -delete -recall\_policy check | no | yes** (yes will recall files before deleting connection)

**BACKING UP FILES:**

**# fs\_dhsm -modify fs1 -backup offline | passthrough** [backs up stub files, CIFS only]

**# fs\_dhsm -modify fs1 -read\_policy\_override full | partial | passthrough | none** [backs up files using NFS]

**MODIFYING HTTP THREADS:**

**# server\_http server\_2 -modify dhsm -threads 40**

**DISABLING DHSM:**

**# fs\_dhsm -modify fs1 -state enabled | disabled**

**MODIFYING DHSM WITH DISKXTENDER FOR VDMs:**

**Source Celerra:**

CIFS Server = vdma, Share = vdm\_source, IP = 10.32.22.34, File System = vdm\_fs\_1 on vdm\_1, FQDN path

[\\vdma.nas2000.lab\\vdm\\_source](\\vdma.nas2000.lab\\vdm_source)

**Target Windows 2000 Server:**

Nasdc2000, domain = nas2000.lab, IP = 10.32.22.240, admin=administrator, target directory = c:\vdmdest, target share = <\\vdmdest>,  
FQDN path to target share = <\\nasdc2000.nas2000.lab\\vdmdest>

**Modifying Celerra:**

**# server\_http server\_4 -modify dhsm -users valid -hosts 10.32.22.240**

**# /nas/sbin/server\_user server\_4 -a -md5 -passwd dhsmuser**

**# fs\_dhsm -modify vdm\_1\_fs1 -state enabled**

**# fs\_dhsm -connection vdm\_1\_fs1 -create -type cifs -admin 'nas2000.lab\administrator' -secondary '\\\nasdc2000.nas.lab\\vdmdest' -local\_server vdma**

**Setting up DiskXtender:**

--Setup all DX Data Manager services to run using Log On As NAS2000\Administrator account

--Using Data Manager for Celerra>Source dirs: setup Source path ([\\vdma.nas2000.lab\\vdm\\_source](\\vdma.nas2000.lab\\vdm_source)) & Host Name  
(vdma.nas2000.lab)

--Using Data Manager for Celerra>Destination dirs: setup Destination path (<\\nasdc2000.nas2000.lab\\vdmdest>), Path to primary storage (/vdmdest/), & Host Name (nasdc2000.nas2000.lab)

--Rules>setup Rule criteria as "-name \*.txt"

--Migration methods>setup Readback method (full), Migration Mode (DHSM), & Username (dhsmuser)

--Jobs>setup Rule (vdmtext), Migration method (vdmmigrate), Destination (vdmtarget), Time Limit (minutes) (0), Source (vdmsource)

**SMARTMOVE ENIGMA DATA SYSTEMS DHSM POLICY HOST:**

1. Open Enigma Management Console and add ‘DHSM Hosts’ by compname and port 5080, selecting ‘Enable secure HTTP, and testing with ‘Test Connection’ button
2. Configure>Destinations>Add: add remote Secondary Server by name and define secondary file system path using UNC
3. On Primary Store ‘Share’, create folders and migrate data for testing to these folders
4. Add Container to Enigma: Configure>Containers>Add: Container name, UNC path to container [\dm2\prim\container1], select ‘Migrate Celerra’ radial button
5. Select ‘Update>Projects>Update All’ and verify that files are added
6. Create Filters: Migration>Filters [Use Exclude if selecting file types such as \*.pdf for a filter or Include for all files]
7. Add Rule to Filter: Migration>Rules>Add
8. Add a Job: Migration>Jobs>Add
9. Execute the Job and check Secondary Store Share for migrated files [View Primary Store for Offline Inodes for migrated files]

### **FILEMOVER UTILITIES (developed by IST group):**

--fmexport (bulk file exporter, copies data to Secondary & converts regular files to offline stubs; scriptable, works with Policy Engine)  
--fmimport (bulk file importer, copies stub files from one location to Celerra—NFS or CIFS; scans existing directory tree to get file attributes, creates 0-length file, transfers permissions, then replaces 0-length file with stub file pointing to Secondary source)  
Issues: links, namespace of files; can use xfrperms to transfer Windows permissions  
--fmxfer (Unix copy that is stub file aware). Retrieves attributes of files from Source, creates zero-length file on Primary and transfers permissions from Source, then converts file to Stub while pointing to same Secondary storage. Preserves hard & soft links.  
--fmlist (Unix ls that displays stub files and offline attributes)

### **PROBLEMS RECALLING FILES FROM SECONDARY STORAGE:**

→There is an issue where Rainfinity FMA preserves file permissions when copying to Secondary Storage. This means that if files are Read-Only, Celerra will not be able to recall them unless the FileMover connection account has backup privileges [e.g., member of local Backup Operators group, or member of Domain Admins], and also has the “backupintent” flag set.

→AR135811 introduced hard-coded BackupIntent flag, and NAS 5.6.46 introduced a param that can be changed

**# .server\_config server\_2 -v "param mgfs useBackupIntentMode"**

mgfs.useBackupIntentMode INT 0x03534f5c 1 1 (0,1) FALSE NONE 'The parameter decides whether DHSM connection should use backupIntentMode or not.'

### **TROUBLESHOOTING COMMANDS:**

fs\_dhsm →DHSM management on Celerra

server\_http →Security management on Celerra

server\_user →DHSM User creation on Celerra

get\_attributes →DHSM API query call to get state of files and return attributes

set\_attributes →DHSM API call to convert online files to Stub or update offline attributes of stub files (# ./set\_attributes -u dhsm\_user -p dhsm\_user -v 1063404534 (mtime of file) 192.168.4.53 /path/file.tar nfs://10.241.168.78/path/file.tar)

get\_api\_attrs →DHSM API queries to check API access and functionality (# ./get\_api\_attrs -u dhsm -p dhsmdhsm 10.64.200.53 /PATH...PATH="/")

get\_connection\_list →DHSM API query for list of connections configured on Data Mover (# ./get\_connection\_list -u dhsm -p dhsmdhsm 10.64.200.172 cifs://celdm2.w2k.celerra.com/fs1)

### **TROUBLESHOOTING FILEMOVER/DHSM:**

Check Server Log and DHSM logs (pfs/.etc), increase logging levels, verify Policy Engine, Data Mover configuration and connections, use get\_attributes commands, etc.

#### **1. Set following debug logging levels:**

**\$ .server\_config server\_2 "logsys set severity DHSM=LOG\_DBG3"**

**\$ .server\_config server\_2 "logsys set severity MGFS=LOG\_DBG3"**

**\$ .server\_config server\_2 "logsys set severity HTTPD=LOG\_DBG3"**

**Note:** Use the following syntax to turn off debug logging when completed with log collection and troubleshooting

**\$ .server\_config server\_2 "logsys set severity DHSM=LOG\_PRINTF"**

**# .server\_config server\_2 -v "printstats dhsm"**

SetOfflineAttrs : 0

GetAttrs : 0

GetAttrsFail : 0

SetOfflineAttrsFail : 0

GetApiAttrs : 0

GetApiAttrsFail : 0

```
GetConnectionList      : 0
GetConnectionListFail : 0
GetBulkAttrs          : 0
GetBulkAttrsFail      : 0
QueryBulkRetreival   : 0
QueryBulkRetreivalFail: 0
AbortBulkRetreival   : 0
AbortBulkRetreivalFail: 0
```

## **2. Check DHSM Log on the Primary Storage File System in the hidden .etc directory:**

```
.etc]# ls -la
-r--r--r-- 1 root bin 7144638 Sep 22 09:06 dhsm.log
```

**Note:** (5) generations of the 10MB log are kept. Increase log size:

**\$ fs\_dhsm -modify fs01 -max\_log\_size 20 [20MB]**

→SET\_OFFLINE indicates that migrated file was replaced by stub

→REMOVE indicates that stub file was deleted

→GET\_ONLINE indicates that stub file on primary was recalled from secondary

**# head dhsm.log**

```
91760*: SET_OFFLINE: 1.1d.1d198e.4296ec2f, pe_id: NONE, offline_path:
```

```
\\\fei0018a030.hq.ferg.com\DMC\fei0018s009.hq.ferg.com\home\HQ\epuckett\My Documents\Vendor Queries\Queries for Thom Riley\4294967325-1907086-1117187119 ... Sun Sep 4 13:14:05 2005
```

## **3. Check Primus & AR's before escalating issues**

## **4. Check Policy engine, Server and DHSM file system logs**

## **5. Verify NFS or CIFS exports**

## **6. Verify connections from DM to Secondary using fs\_dhsm -c <fs01> -i**

## **7. Use get\_api attrs script to check connectivity to API**

## **8. Use get\_attributes script to check files**

## **9. User server http to verify various parts of DHSM configuration that relate to HTTP**

**Note:** With NAS 5.6, the default Authentication for DHSM is set to “Digest”. You may need to modify the security access settings as the following example shows:

**# server\_http server\_2 -info**

```
server_2 : done
DHSM FACILITY CONFIGURATION
Service name      : EMC File Mover service
Comment          : Service facility for getting DHSM attributes
Active           : False
Port             : 5080
Threads          : 16
Max requests     : 300
Timeout          : 60 seconds
ACCESS CONTROL
Allowed IPs      : any
Authentication   : digest ,Realm : DHSM_Authorization
Allowed user     : nobody
```

**# server\_http server\_2 -modify dhsm -authentication none**

```
server_2 :
DHSM FACILITY CONFIGURATION
Service name      : EMC File Mover service
Comment          : Service facility for getting DHSM attributes
Active           : False
Port             : 5080
Threads          : 16
Max requests     : 300
Timeout          : 60 seconds
ACCESS CONTROL
Allowed IPs      : any
Authentication   : none ,Realm : DHSM_Authorization
Allowed user     : everybody
```

## **MS EXCEL FILES & DHSM BEHAVIOR:**

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
Excel files have a tendency to be recalled when a stub file is accessed by a Windows client. Workaround to correct this behavior--  
modify Excel files as ReadOnly.

## **VERIFIER MISMATCH OFFLINE FAILURE:**

### **Server Log:**

DHSM: 6: API failed: VERIFIER\_MISMATCH

→See AR78669, but basically, DiskXtender will fail to complete setting files offline if DART returns more than 40  
VERIFIER\_MISMATCH errors, which can occur when DM updates file attributes after DX has cached previous values. DART  
returns the mismatch error because the file attributes are different than DX. Setting mappingErrorAction=1 may allow offline  
operation to succeed in the face of many unknown SIDs.

## **NETGROUP ACCESS WITH NFS EXPORTS:**

**Purpose:** Netgroups are groups of Hosts and/or Users used to restrict access to shared NFS exports

--Hosts and Subnet groups have preference over Netgroups

--When two or more netgroups are defined, the first matching entry is used for granting access

--Negation can be used to specify hosts to be excluded from access

**Note:** Negated entries should be at beginning of list and do not negate other netgroup lists

**#server\_export server\_2 -o rw=-netgroup1:netgroup2 /fs1**

**#server\_export server\_2 -o rw=-host1:netgroup1 /fs1**

--If both RW and RO are used in Export, result depends on following params:

**param nfs secureExportMode=0** [Provides RW access in case of conflict]

**param nfs secureExportMode=1** [Provides RO access in case of conflict]

### **EXAMPLES:**

rw=host1 ro=netgroup1 [If host1 is a member of netgroup1, then host1 will have RW access due to host preference]

rw=netgroup1 ro=subnet1 [host1 is a member of both groups, host1 will have only RO access due to subnet preference]

rw=-netgroup1:netgroup2 [host1 is member of both groups, then host1 will not have RW access]

rw=netgroup2:-netgroup1 [host1 is member of both groups, then host1 will have RW access]

## **FILE EXTENSION FILTERING (Celerra File Filtering):**

**Intro:** Can now filter what file extension types are allowed to be saved to a network Share on Data Mover

### **FILE FILTERING CIFS PARAM:**

**# server\_param server\_2 -facility cifs -info enableFileFiltering**

server\_2 :

```
name      = enableFileFiltering
facility_name = cifs
default_value = 0x00000007
current_value = 0x00000007
configured_value =
user_action = none
change_effective = immediate
range     = (0x00000000,0xffffffff)
description = Bit vector to enable various forms of file filtering
```

**param cifs enableFileFiltering=7**

**Note:** File Filtering, with PopUps and Auditing, is enabled by default. 0=disabled; 1=filtering enabled; 3=filtering + Popups;

5=Filtering & Auditing; 7=Filtering, Auditing, & Popups

### **File Extension Filtering Requirements/Limitations:**

→works only with CIFS protocol (NFS Users cannot be controlled with this feature)

→requires security NT policy

→File Filtering uses combination of file extension and ACLs to work

→File Filtering done at DM or Share level, not subfolder level

→filters are created in the /nasmcd/quota/slot\_2/.filefilter directory on the DM

# pwd

**/nasmcd/quota/slot\_2**

[root@nyip1 slot\_2]# ls -la

**drwxr-xr-x 2 root bin 80 Sep 25 14:10 .filefilter**

### **CREATING FILTER FILES IN .FILEFILTERING FOLDER:**

--You must create the 'filter' file from a Windows client

--As Administrator, connect to [\\datamover\c\\$](\\datamover\c$) share and go to the .filefilter folder and create the filter file inside this folder

**Note:** The actual “filter” file is a zero byte file with the chosen extension identified as the name of the file itself.

--As Administrator, set ACL permissions on the actual “filter” file itself to further control behavior

#### **FILTER FILE EXAMPLES:**

**Note:** The name after the first @ sign indicates the Sharename—if a 2<sup>nd</sup> @ sign is used, the name that follows is the Compname  
**/slot\_2/.filefilter/jpg@isaac | exe@isaac | allfiles@isaac lhtm@isaac@isaac**

**htm@isaac@compname →htm=extension to filter on | @isaac=sharename | @compname=compname**

1. Create a filter file to prevent any .jpg files from being saved on a share called “Isaac” on Server\_2 by using notepad to create an empty file with a name “jpg@isaac”. Do not allow notepad to save the file with the .txt extension. The actual file is on the extension called “jpg” for the share called “isaac”. With this setting, and file filtering params in effect, the only persons allowed to copy files with the ‘.jpg’ extension to this Share will be the actual Owner of the filter file, in this case an Administrator. To further refine access, you can then define ACL’s on the filter file to allow or disallow Windows Users or Groups from being able to copy .jpg files to the share. For example, if you grant a user called “NAS1” a specific ACL to allow RW privileges to “jpg@isaac”, that User will be able to copy jpg files to this Share, and so on.

2. Create a filter file to prevent copying of any files to a Share called ‘isaac’: allfiles@isaac. Only the Owner of the filter file will be able to copy files to this Share. You could then set permissions to allow specific Users or Groups to be able to copy files to the Share. Add specific file filters with ACE entries to allow for specific file types to be copied to a Share for specified Users/Groups, etc.

3. To allow all files to be copied to a share called ‘isaac’ on a Server called ‘compname’:

allfiles@isaac@compname

4. To allow only a specific extension to be copied:

exe@isaac@compname

**Note:** Multiple filter files should be tested to ensure the desired effect. For example, creating a filter ‘allfiles@isaac’ to deny all files, then creating an allow filter for ‘exe@isaac’ will not result in exe files to be copied

5. To allow only xls files to be copied to a share, create following filter files:

allfiles@isaac

xls@isaac [Set ACL on this file to all Everyone Full Control—will allow copy or creation of xls files to this share]

noext@isaac [Set ACL on this file to Everyone Full Control—file needed because of temp files used by excel process]

#### **RESERVING A SHARE FOR SPECIFIC FILE TYPES:**

a.) Create the Filter Filename called allfiles@sharename@compname (restricts all file types from Share)

b.) Prevent files without extensions from being added to share: noext@sharename@compname

c.) Specify the actual file type for each extension that will be allowed on the Share: extension\_name@sharename@compname and then make sure to set ACLs on the filter files

#### **SPECIAL FILE FILTER NAMES REQUIRED TO SUPPORT MS WORD, EXCEL, PPT, etc:**

allfiles@sharename →No ACL assigned

noext@sharename →No ACL assigned

doc@sharename →Assign ACE for User/Group access

tmp@sharename “ “

asd@sharename “ “

wmf@sharename “ “

rtf@sharename “ “

xls@sharename →Assign ACE for User/Group access for Excel files

ppt@sharename →Assign ACE for User/Group access for Powerpoint—also requires the tmp entry as shown above for Word

#### **FILE FILTERING BEHAVIOR TESTING:**

→allfiles@isaac@compname allows any user to copy any file type to Isaac Share on compname ‘Compname’

→allfiles@isaac denies any user except owner, or someone added to the ACL on the filter file with RW perms, the ability to copy files to Isaac Share

**Note:** This file filter is restrictive--other filter files added to allow extensions to be copied will be denied [allfiles@issac + exe@isaac or exe@isaac@compname]

→allfiles@isaac denies all Users, except for a user called “NAS1”, who has been granted Full Control permissions on the filter file called “allfiles@isaac” in the .filefilter directory

→jpg@isaac allows User to copy all files to Share except for jpg files

→jpg@isaac@compname allows users to copy all files, including jpg’s, to the Isaac Share on ‘Compname’

#### **FILE FILTERING PARAMETERS & OBSERVED BEHAVIOR:**

0: filtering disabled—No Filtering, No PopUps, No Event Logging

1: file filtering enabled—Filtering enabled, No PopUps, and No Event Logging on Client system

3: file filtering + popups enabled—Filtering enabled, PopUps enabled, and Event Logging on Client system occurs

5: file filtering + auditing enabled, no popups—Filtering enabled, No PopUps, and No Event Logging on Client system

7: file filtering + popups + auditing enabled [Default Celerra setting]—Filtering enabled, PopUps enabled, and Event Logging on Client occurs

**Note:** Messenger Service is required in order for Clients to be able to receive a PopUp message—runs by default on Win2k, does not on Win2k3. Must reboot Server if changing file filter param policy.

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
**FILE FILTERING CLIENT ERROR MESSAGE** [Not a PopUp message—message results when Filter policy is in force]:

Error Copying File or Folder

Cannot copy JUNKMAIL: Access is denied

Make sure the disk is not full or write-protected and that the file is not currently in use.

**FILE FILTERING POPUP MESSAGE FROM MESSENGER SERVICE:**

Message from Iprep2 to Sneezy on 3/29/2005 3:12:30PM

File Extension not allowed

Share Isaac

File \JUNKMAIL.LKO

Operation create

## **USING CIFSMSG.TXT FILE TO CREATE CUSTOM POPUP NOTIFICATION MESSAGES:**

→cifsmsg.txt file can be used to customize PopUp messages for File Filtering, CAVA, Quotas, DHSM [CIFS only]

→In order to use PopUp messages, the Windows client must run the Messenger service

→The cifsmsg.txt file is meant only to customize the actual popup message generated by the feature

→Last line for the Warning or Error should have a period

### **CUSTOMIZING POPUP MESSAGES FOR QUOTAS:**

1.) Create and/or edit a file called “cifsmsg.txt” to create the custom popup.

#

# Warnings for status NoSpace, QuotaExceeded and

# GroupQuotaExceeded

#

#

\$warning.NoSpace=

You're about to run out of space on your home share

Please initiate cleanup of unwanted and old files now to

prevent future disruption

.

\$warning.QuotaExceeded=\$warning.NoSpace

\$warning.GroupQuotaExceeded=\$warning.NoSpace

#

#

# Errors for status NoSpace, QuotaExceeded and GroupQuotaExceeded

#

#

\$error.NoSpace=

Your home share has reached the allotted quota limit

Please initiate cleanup of unwanted and old files now

Contact the ANR help desk @ ext 5800 for any needed assistance

.

\$error.QuotaExceeded=\$error.NoSpace

\$error.GroupQuotaExceeded=\$error.NoSpace

2.) Push file out to /.etc of DM using server\_file -put

3.) Stop and Start CIFS service

## **TOE TCP OFFLOAD ENGINES ON NIC CARDS:**

NAS 5.3 introduces TOE (TCP Offload Engines) with Alacritech Accelerator 10/100/1000 Base-T or 1000Base-SX NICs, introduced with 514 DM's and designed to offload NFS, CIFS, and FTP connections

**Note:** Drawback of TOE is that network captures from DART do not see TOE traffic—must use other sniffing equipment. Disable TOE with following parameter. Additionally, TOE devices can panic, causing a SLIC dump to logfile called “/slicdump” ~16MB.

**param TOE StdNicMode=1**

**Note:** Current recommendation is to disable TOE offload engine for TOE cards by setting value to 1—see AR44627

**\$ .server\_config server\_2 -v "param TOE"**

| Name | Location | Current | Default |
|------|----------|---------|---------|
|------|----------|---------|---------|

|                |            |            |            |
|----------------|------------|------------|------------|
| TOE.StdNicMode | 0x011f9654 | 0x00000000 | 0x00000000 |
|----------------|------------|------------|------------|

|                    |            |            |            |
|--------------------|------------|------------|------------|
| TOE.framechk.hiwat | 0x011f97d4 | 0x00020000 | 0x00020000 |
|--------------------|------------|------------|------------|

|                    |            |            |            |
|--------------------|------------|------------|------------|
| TOE.framechk.lowat | 0x011f97d8 | 0x00008000 | 0x00008000 |
|--------------------|------------|------------|------------|

## **ESCALATING TOE ISSUES :**

**1. Server Log will indicate when a TOE dump has been produced :**

2005-07-05 09:28:25: DRIVERS: 3: slicNicDevice: TOE (ace0) panic detected. Saving TOE core image to '/slicdump'  
2005-07-05 09:28:25: DRIVERS: 3: 6: slicNicDevice: TOE (ace0) panic detected. Saving TOE core image to '/slicdump'  
2005-07-05 09:28:25: DRIVERS: 4: slicDumpManager: Dump requested for ace0  
2005-07-05 09:28:25: DRIVERS: 4: slicDump: HALT Receive Processor  
2005-07-05 09:28:26: DRIVERS: 4: slicNicDevice: output(): CARD\_DOWN! Packet dropped!  
2005-07-05 09:28:26: DRIVERS: 4: last message repeated 20 times  
2005-07-05 09:28:26: DRIVERS: 4: slicDump: halt receive sequencer failed  
2005-07-05 09:28:26: DRIVERS: 4: slicDump: Write CoreHeader len[168] offset[0]  
2005-07-05 09:28:26: DRIVERS: 4: slicNicDevice: Dump complete

2. When escalating to EE, run the following script—collects information from sys\_log, server\_log, server\_sysconfig, server\_netstat, etc., and retrieves TOE dumps, and bundles tarred gzip file in /tmp/Toe\_Materials\_050718\_1507

**#/nas/tools/toe\_support\_materials server\_2**

/tmp/Toe\_Materials\_050717\_2057.tar.gz

/tmp/slicdump.gz

3. Debugging TOE Firmware Issues :

**param slic diagFW=1**

4. Disabling TOE Card :

**param TOE StdNicMode=1**

**COMMANDS USED TO DERIVE TOE INFORMATION :**

**\$ .server\_config server\_2 -v "slic ace0 showlist"** →Shows connections on TOE card

**\$ .server\_config server\_2 -v "slic ace0 showatkt"**

**\$ .server\_config server\_2 -v "slic ace0 showadapter"**

**\$ .server\_config server\_2 -v "slic ace0 showcard"**

**\$ .server\_config server\_2 -v "slic ace0 showoffload"**

**\$ .server\_config server\_2 -v "slic ace0 showints"**

**BROADCOM BCM GIGE DEVICES:**

Currently, Broadcom NICs are used in NS600/NS700 & have (6) copper ports

New ‘Ajaguar’ card to have (6) copper, (2) optical 1GB fibre, and (1) copper management port for NS700 uses 5704 chipset

**Note:** When a port is set to (1) GB, duplexing must always be set to full. Txflow & Rxflow are only used at 1Gbps. Linkneg is on by default and must be disabled on switches that do not support linkneg. Support for MAC flow control added.

**NETWORKING STATISTICAL LOAD BALANCING USING TCP OR IP ADDRESSES:**

--TCP and IP Statistical load balancing [TCP uses source and dest. IP address and tcp port number; IP uses source & dest. IP address]

--Configure load balancing via param file or server\_sysconfig command

**param trunk perTrunkLB=0** [Sets default stat. load balancing based on IP addr., vs. MAC addresses in previous versions]

**\$server\_sysconfig server\_2 -v -n trk1 -create trk -option “device=cge0,cge1 lb=tcp”** [Load balancing TCP]

**\$server\_sysconfig server\_2 -v -info trk1**

**\$server\_sysconfig server\_2 -v -n link1 -create trk -option “device=cge0,cge1 protocol=lacp lb=tcp”** [Load balancing LACP]

**\$server\_sysconfig server\_2 -v -info link1**

**VLSM—VARIABLE LENGTH SUBNET MASKS:**

--support for optional subnet mask lengths [DART now supports VLSM in the Routing Table]

--new Broadcom NIC driver support for flow control

**#server\_sysconfig server\_x rxflowctl=enable | disable      txflowctl=enable | disable**

**TCP CONNECTION HASHING:**

New default maxStreams=65536 [16k to 64k increase]

NAS 5.3 caches the ‘last connection’ in cache

**CELLERRA MONITOR 5.3:**

HomePage Statistics tab

CIFS Shares, Data Movers, NFS Exports, Network, File Systems, Volumes, Control Station, Network Storage systems

Capability of exporting “lists” & “graphs” into .csv files by rightclicking>Export Data [comma-separated]

Graphical displays of past and current status of DM’s, FS’s, Storage Systems, Networks

Notifications of DM resource issues can be done using email or SNMP traps [High Memory & CPU Usage] via Polling

Notifications of FS issues are done using email/SNMP traps [File System Usage & File System Projection]

## **NEW NAS SERVER -QUERY INTERFACE WITH CELERRA MANAGER:**

**\$ /nas/bin/nas\_server -query:name=server\_2 -fields:MemoryUsage -format:"%s\n"**

62

**Note:** Use above format to make queries for quotas, file system info, etc.

## **USING SERVER CIFS TO MIGRATE SIDs FROM ONE DOMAIN TO ANOTHER:**

**\$server\_cifs server\_x -Migrate fs1 -acl Disney:nb=walt:if=ana0 mouse:nb=Minnie:if=fsn0**  
**\$server\_cifs server\_x -Migrate walt -localgroup disney:nb=walt:if=ana0 mouse:nb=Minnie:if=fsn0**  
**\$server\_cifs server\_x -Replace fs1 -acl Disney:nb=walt:if=ana0**  
**\$server\_cifs server\_x -Replace walt -localgroup Disney:nb=walt:if=fsn0**

**Note:** -Migrate updates SIDs from a Source domain to SIDs of a target domain by matching User and Group account names. – Replace updates all SIDs on a file system with target SIDs and localgroups information.

## **SNAPSURE ENHANCEMENTS:**

--more than one schedule per PFS is now supported, and schedules can be modified

**Note:** 5.1 and 5.2 will convert old Schedules to new format automatically on upgrades, 4.2 will require manual conversion

--New ‘PAE’ feature to provide for memory blockmaps if RAM greater than 4GB [Physical Address Extension], to be handled by the VSM [Virtual Space Manager]

**Note:** This requires special PAE Data Movers and will support up to 64 Checkpoints per PFS

--New ‘VSS’ [Virtual Shadow-Copy Service] support—based on Microsoft’s VSS feature found in NetWin products

## **FS REPLICATE/IP REPPLICATION ENHANCEMENTS:**

→Replication of CIFS environments from (1) DM to another [VDM root file system and PFS are replicated at same time to SFS, mounted RW by User, becomes a replicated VDM]

**Note:** Rules are that mounts must be identical on PFS and SFS; Source & Target VDMs must be in same Windows Domain; Interfaces must also be identical & use same IP Address

→Ability to change DART TCP Window Size to support replication and fs\_copy

→nas\_fs supports “cloning” through use of ‘samesize=’ syntax to create PFS and SFS of same size

→Full copy on a failed resync

→Initial Copy Support for Disk or Tape transport to startup Remote IP Replication Sites

→ -suspend and -restart commands to allow for changes in SavVol, unmount filesystems and move, change IPs on interfaces, etc.

**New commands:** fs\_replicate -suspend | -restart | -reverse (replaces the ‘failback’ command) | -modify | -refresh | -resync

### **fs\_replicate -suspend:**

The –suspend command is run on the PFS side and creates a checkpoint newer than the SFS, causing Replication and Playback to shut down and stops replication processes—leaves SFS as rawfs. Command allows for changes to SavVol size, change of mountpoints, or change of IP Addresses.

### **fs\_replicate -restart:**

Also run on PFS side. Checks to see if a Suspend has occurred and if so, will use the suspend checkpoint to initiate an incremental restart of replication. Data from point of suspend on PFS side will be copied to remote side. Checks to see if PFS and SFS are in a restartable condition. Verifies that SFS is same size as PFS and that SFS is rawfs.

### **fs\_replicate -modify:**

Does not create a deltabset or attempt playback. Makes changes to HWM, Timeouts, autorotate or autofreeze immediate.

### **fs\_replicate -refresh:**

Takes new deltabset and sets new timeout, HWM policies

### **fs\_replicate -resync:**

Allows for full copy of fs to be sent back to primary side if incremental resync failed.

## **INITIAL COPY SUPPORT (aka, Disk & Tape Silvering):**

### **TAPE LEVEL COPY:**

→Special NDMP backup and restore using Volume Level Copy (VLC) using NDMP Variable VLC=y

→Volume level backup only of RO filesystems such as Checkpoints

→Restore from tape is made to rawfs file systems only

→File system limited to 2TB in size for this variable

### **DISK LEVEL COPY:**

→Transporting CX300 array to site

**Note:** Concept here is that a bulk copy of a PFS to the SFS can be done via Tape or Disk images, then synced up using IP Replication. For example, a 1TB file system could take 50 days to replicate over a T-1 line.

## **NDMP ENHANCEMENTS:**

--Increased Restore and Backup performance using multi-threading for PAX

--Better directory traversal and large file handling

### **Remote NDMP Support with Veritas:**

Heterogeneous 3-way backup over network. Data Mover is not connected to Tape Server and acts as Data Server.

### **THREE NEW THREAD GROUPS:**

NASA—Thread responsible for writing file header information, reading file data, and writing to buffer. Sends metadata to backup software and activates multiple threads, puts buffer in stat buffer pool. Allows multiple NASA threads for large file support. Pre-NAS 5.3 allowed only single NASA thread per file.

NASS—Threads responsible for traversing file system and providing metadata for each file and directory--returns metadata to stat buffer pool for each file and directory. Directory Traversal bottlenecks could occur with many small CIFS files or highly fragmented file system.

NASW—Threads gather data from buffer pool, write to tape or send to remote NDMP system.

### **New PAX Tuning Commands & GUI Interface:**

**#server\_param server\_x -facility <name> -list <paramname> | -modify <paramname> <value>**

[paxStatBuff; paxWriteBuff; paxWriteToTape; paxWriteToArch; paxnPrefetch; paxnThread; paxnFTSThreads; paxReadBuff]

**#server\_pax server\_x -stats | -reset | -verbose**

--New PAX param added to configure PAX block size: PAX.readWriteBLockSizeInKB (64 kb default—256 maximum value)

**Note:** Issue with EDM in which this value needs to be set specifically to 128

### **PARAMETER TUNING INTERFACE:**

**\$server\_param server\_x -facility PAX -info paxWriteBuff**

**Note:** Outputs current, configured, and maximum values for a param

**\$server\_param server\_x -facility PAX -modify nThread -value 8** [Sets PAX threads to 8]

**\$server\_param ALL -facility PAX -modify nThread -v 8**

**\$ server\_param server\_2 -f PAX -l**

server\_2 :

| name         | facility | default | current | configured |
|--------------|----------|---------|---------|------------|
| paxWriteBuff | PAX      | 64      | 64      |            |
| paxStatBuff  | PAX      | 128     | 128     |            |
| nFTSThreads  | PAX      | 8       | 8       |            |
| nThread      | PAX      | 64      | 64      |            |
| nPrefetch    | PAX      | 8       | 8       |            |
| nRestore     | PAX      | 16      | 16      |            |

**Note:** New default values for PAX in 5.3

### **AVM ENHANCEMENTS:**

#### **AVM Profiles:**

→Storage Pools—containers used for storage management [User-defined and System-defined]

→Profiles—set of rules to aggregate storage

→Striping only occurs in the building of a pool, never in the contents of a pool

#### **User-Defined Pools:**

--Create manually with specialized volume striping, etc, aggregate into pools

--Volume Profiles do not apply to ‘user-defined’ pools

--Cannot be dynamically extended or shrunk

--Storage must be explicitly added or removed

#### **System-Defined Pools:**

--Uses Volume Profile and Storage Profile

--Aggregates disk volumes into Pools automatically

--Volume profile that allows disks to be added and released automatically

**Note:** Storage profiles match disks to certain profile and compares to template to see if they match. Volume profiles define how new disk volumes are added to a system-defined storage pool.

#### **Volume Profile:**

--system-defined pools only

--defines how disks are aggregated and placed into pools

--only (1) volume profile per system-defined pool

#### **Storage Profile:**

--system-defined pools only

--description of disk volumes

--only Symm STD disks, not BCV's

#### **STORAGE POOL:**

--A container used to aggregate disk volumes into pools

--can consist of disks, stripes, slices, metavolumes

--A volume can only be a member of one pool

--Volumes must be of the same type as other members

### **MANAGING AVM STORAGE POOLS:**

--CLI

--Celerra Manager GUI

### **AVM GUIDELINES:**

--File Systems can be extended from only one storage pool

--Use multiple disk volumes from one Backend and not multiple backends at the same time

--Create user-defined pools if there are different Storage systems [System-defined might span different storage Backends]

### **STORAGE POOLS:**

**#nas\_pool -create pool\_1 | -delete pool\_1**

**#nas\_pool -list**

**#nas\_pool -size pool\_1 #nas\_pool -i pool\_1**

**#nas\_pool -modify id=11 -default\_slice\_flag y**

**Note:** CLI behavior is to have the default\_slice\_flag on by default—set to “n” if not desired—used to control whether volumes can be sliced or not. For example, if member volumes have already been built on slices, you might want to use “n” to disable.

**#nas\_pool -modify symm\_std -is\_dynamic y**

**Note:** -is\_dynamic indicates whether pool is allowed to automatically Add or Remove member volumes. Applies only to system-defined pools. Setting the value to no will prevent QoS mechanism from grabbing additional storage luns for use. An Error 3024: No free disks are available would be indicated.

**# nas\_pool -modify symm\_std -is\_greedy y**

**Note:** System-defined pools where volumes will be created from new space before using space from existing member volumes. A pool that is not greedy will first use up all space from existing pool member volumes before using other volumes.

**# nas\_pool -append pool\_1 -volumes d40 | -remove** [to remove a volume from a pool]

**Note:** Adds a volume to the pool indicated

### **USING AVM TO CREATE SYSTEM-DEFINED POOL FOR FILE SYSTEM:**

**# nas\_fs -name fs\_system -create size=100G pool=symm\_std**

### **USING AVM TO CREATE USER-DEFINED POOL:**

**# nas\_pool -create -name user\_pool -volumes d40,d41,d42,d43**

### **USING AVM TO CREATE FS FROM USER-DEFINED POOL:**

**# nas\_fs -name fs\_system -create size=100G pool=user\_pool -o slice=y**

**Note:** Must create the user-defined pool first!

### **EXTENDING FILE SYSTEM USING AVM:**

**nas\_fs -xtend fs1 size=50G**

**Note:** NAS 5.3 is NOT backend aware, meaning that AVM may or may not choose to extend a file system across different Storage systems! Recommendation would be to create user-defined storage pools that are built on specific systems. To be doubly certain, create a temp meta that uses all free disks on the backends you do not want to extend to! Then conduct the extension and delete the temp meta.

### **CREATING NEW FILE SYSTEM FROM EXISTING POOL:**

**# nas\_pool -list**

```
id    inuse  acl   name
1     n      421   symm_std
2     n      421   clar_r1
3     n      421   clar_r5_performance
4     y      421   clar_r5_economy  [Pool exists]
8     n      421   symm_std_rdf_src
10    n      421   clarata_archive
```

**# nas\_pool -i clar\_r5\_economy**

```
id          = 4
name        = clar_r5_economy
description = CLARiiON RAID5 8plus1
acl         = 421
in_use     = True
clients    = IPrep_src,vpfs24,pfs,vpfs30,vpfs32,boeing
members    = v110,v114,v121,v124,v128,v131,v139
default_slice_flag = True
```

```
is_user_defined = False [Pool is a system-defined pool]
disk_type = CLSTD
server_visibility = server_3,server_2
volume_profile = clar_r5_economy_vp
is_dynamic = True
is_greedy = True
# nas_pool -size clar_r5_economy
id = 4
name = clar_r5_economy
used_mb = 303736
avail_mb = 462512 [Pool shows room available to create file systems]
total_mb = 766248
potential_mb = 218928
```

### **# nas\_pool -size id=3 -slice y**

**Note:** Use this to see pool size information based on Slice information

```
# /nas/sbin/rootnas_fs -n new_group -create size=2G pool=clar_r5_economy -o slice=y
```

```
id = 47
name = new_group
acl = 0
in_use = False
type = uxf
volume = v152
pool = clar_r5_economy
member_of = root_avm_fs_group_4
rw_servers=
ro_servers=
rw_vdms =
ro_vdms =
stor_devs = APM00042103183-0018
disks = d11
```

### **NOT ALLOWED TO EXTEND FILE SYSTEM FROM POOL:**

S: nas\_fs -xtend fs1 size=29000M pool=symm\_Std

E: nas\_fs -xtend fs1 size=29000M pool=symm\_std: : volume(s) are not available

**Note:** Though message may be legitimate, in many cases, it may only mean that an entire disk is not available to perform the extension. Check the output of nas\_pool to determine if there really is any space left and then use the slice=yes option:

### **\$ nas\_pool -s symm\_std -slice y**

```
id = 1
name = symm_std
used_mb = 315040
avail_mb = 30200 →Total space left in pool
total_mb = 345240
potential_mb = 0
```

```
# nas_fs -xtend fs1 size=29000M pool=symm_std -o slice=y
```

### **SRDF FILE SYSTEM:**

If a file system is to be mirrored for SRDF, must use symm\_std\_rdf\_src storage pool. AVM will allocate space from other volumes for the remotely mirrored file system.

### **HIGHROAD ENHANCEMENTS:**

--Highroad now supports all Clariion Backend configurations (prior to this only Symmetrix)

--As before, Highroad functions with large file transfers only—metadata and small files go over production network using normal NFS or CIFS protocols

--New AVM now supports use of 256k stripes created from CLI by using Celerra Manager to combine stripes into metavolumes for filesystem creation or storage pools (prior GUI only supported stripe sizes of 8k & 32k)

### **Required Components for Clariion Support:**

--Access Logix for Array, HighRoad and PowerPath software for clients

### **CONFIGURING MPFS WINDOWS CLIENT:**

1. Connect Client to Storage, configure AccessLogix and StorageGroups (create StorageGroups, rightclick, add LUNs)
2. From Windows Client, open Computer Management>Disk Management>Rescan Disks, but DO NOT write disk signature!

3. Install PowerPath and enable Failover Policy on array (1)
4. Install HighRoad for Windows Client
5. Reboot Windows Client, open DOS prompt and run cmd: c:\mpfsinq
6. Create file system from Clariion Luns that are presented, after building meta(s) manually

**# nas\_volume –name hrstripe –create –Stripe 262144 d7,d8,d13,d14** (Use stripe in meta or Volume Pool)

7. Create Share or Export
8. Enable MPFS on Data Mover using server\_setup command

**# server-setup server\_2 –P mpfsf –o start=16**

9. Enable MPFS on Share via Properties tab (Enable MPFS) or Export for NFS using mount option for MPFS

**# mount –F mpfs 192.168.25.143:/hrfs2 /mpfs**

**# mpfslabel /dev/rdsk/emcpower1c** [run for Sun Solaris Clients]

**# server\_mpfs server\_2 –mountstatus**

**# server\_mpfsstat server\_2**

**# server\_mpfs server\_2 -Stats**

#### **HIGHROAD LIMITATIONS:**

- NFS v2 not supported
- Unicode for Share names not supported
- CIFS mandatory locks are not supported
- Per Seat licenses for Windows restricts access to shares on only a single CIFS server per data mover

### **CELERRA NAS 5.4 CHAMPAGNE:**

Introduction of Celerra NSX platform at this code version. NAS 5.4.14.3 GA 11 April 2005, latest version 5.4.31.2. Supports CNS-14, CFS-SE, NS, NSG platforms. Last maintenance version 5.4.31.2 for Flare 26 support, except for Raid 6, which is not supported. Also, 5.4.31.2 does not support DMX-4. Upgrade path is from NAS 5.1 & higher only. Sites should be running Flare 16 or Flare 19, but the latter requires a minimum 5.4.18.3 NAS. NAS 5.4.26 introduces support for Flare 24. Only 5.4 and higher will support DMX3 Symm7 hardware. Supports Symm microcode 5567.52.29, 5566.46.31, 5267.42.29, 5266.43.30, & 5568.56.22, DMX code 5669.47.25 & 5670.27.29.

### **NAS 5.4 IMPROVED GATEWAY INSTALLATIONS:**

- Manual installation required public IP addresses, PXE booting DM's to obtain WWN's, creating Storage Groups, Raid Groups, Binding LUNs, Registering Initiators/Hosts, Mapping LUNs to Storage Group, Configuring Zoning, etc.
- Automatic installation requires IP addresses, switch username and password, but when script runs, Data Movers are PXE booted to acquire WWNs, Storage Groups, Raid Groups, LUNs WWN Registration, LUN Mappings, and Switch zoning [DM auto-senses switch speeds while SP's must be set manually] are done automatically via script and user replies

**Note:** Must use WWN Zoning, have all (4) paths cabled for both DMs, and available storage for Control Luns (Raid5 4+1) from 5 disks. For each DM, BE-0 port will be connected to SPA & SPB, and BE-1 will be connected to a port on SPA & SPB. Switches are zoned per E-Lab rules, DM's WWNs cannot be part of any pre-existing Zones. AccessLogix must be installed and enabled on Clariion before install can succeed:

“Rechecking for existing System RG/Lun: Invalid system raid group detected. Recovery must be performed. Type ‘C’ to continue”

**Note:** Open NaviSphere, connect to SP, rightclick “Storage Systems”>properties>Storage Access \_\_Access Control Enabled

### **INTEGRATED (aka captive) NS INSTALLS:**

→ Installation script automatically configures disks 0-5, creates Control LUNs, creates two data LUNs (16 & 17 decimal HLU), and runs nas\_diskmark to discover and add the new devices to the Celerra database

**Note:** The use of the User Defined template does not work from Celerra Manager, but does from the CLI, as shown below

→ The use of setup\_clariion to configure additional storage, does so on a shelf-by-shelf basis

→ The setup\_clariion command cannot be used on FC-enabled models, only Navi tools

### **ADDING SPECIFIC DAE SHELF TO CELERRA SYSTEM:**

**# /nas/sbin/setup\_backend/setup\_clariion2 -S setup -e 0\_1 SL7E1081700022**

**Note:** If one of the attached DAE's is missing all their disks, the setup\_clariion2 command may fail and require the use of a specific DAE Enclosure identity to add the shelf. See emc200483.

#### **Example of Adding Disks to existing configuration:**

**# /nas/sbin/setup\_clariion -init**

The system displays:

Found CLARIION(s) APM00044604xxx

Setup CLARIION APM00044604xxx storage device...

System 192.168.1.200 is up

System 192.168.1.201 is up

Clariion Array: APM00044604xxx Model: CX700 Memory: 3967

**Enclosure(s) 0\_0,1\_0,0\_1,1\_1 are installed in the system.**

-----output abridged-----

**Continue to configure APM00044604xxx as User\_Defined [yes or no]?:** yes

The system displays:

The following 3 template(s) available:

1. ATA\_RAID5\_HS\_6+1\_6+1

2. ATA\_RAID3\_HS\_4+1\_8+1

3. None

### **Enclosure 1\_1**

Please select a template in the range of 1-3 or 'q' to quit: 1

Enclosure info:

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14

1\_1: 320 320 320 320 320 320 320 320 320 320 320 320 320 320 320

**ATA \*HS \*14 \*14 \*14 \*14 \*14 \*14 \*15 \*15 \*15 \*15 \*15 \*15 R5**

"\*" indicates a diskgroup/spare which will be configured

Template: User\_Defined

Summary:

**2 disk group(s) are created. 14,15**

**1 spare(s) are created. 202**

-----output abridged-----

Do you want to continue and configure as shown [yes or no]?: yes

Enclosure 0\_0.

Enclosure 1\_0.

Enclosure 0\_1.

### **Enclosure 1\_1.**

**Created spare 202**

**Created disk group 14, luns 31,33**

**Created disk group 15, luns 30,32**

Binding complete.

All luns are created successfully!

## **Use WebUI or CLI to add additional LUNs:**

**#/nas/sbin/setup\_clariion -init**

## **NAS 5.4 FEATURES:**

iSCSI NBS support for MS Exchange, etc.

File Level Snaps

Asynchronous Disaster Recovery for Celerra Replicator—Allows Secondary Side to sync with Primary after restore

Multiprotocol Directory Structure (Unified Unix/Windows Directory—no more Shadow file?)

Virtual DM [CIFS]

Linux Upgrades; Modularized NAS Code; Parameter Tuning Tool

Unlimited DNS domains can be specified (lifts the 3 DNS domain limitation)

## **NFS Cluster:**

DMs export same FS Read-Write multiple times, with (1) Server controlling the Metadata & Data, all others control data only  
Bulk Archiving (HSM-lite)

## **NBS Block Services:**

Block access over IP; Celerra Local Disk (CLD); Support for Snapshots; Exchange 2000 Email Support

NS600G Phase 2—Gateway Server for both NAS & SAN environments; uses CX600 Back-End

Level II Oblock Support

NDMP v4 Support

Internal Usrmapper running on DM

Connect Home Support—similar to Symmetrix model

Out of Order Checkpoints can be deleted

## **DFS ( System) ROOT (standalone only):**

**Note:** There are two types of DFS Servers: Domain Root DFS, stores DFS hierarchy in AD, & StandAlone Root DFS, stores DFS hierarchy locally on the Server. DFS provides mapping and uniform naming convention for collections of servers, shares, and files within a single directory tree—idea is to group shared folders from different physical servers into the same logical DFS namespace. A

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
DFS Root is the DFS service containing DFS links pointing to Shared folders, or DFS targets, on the network. Celerra provides DFS Targets when hosting a DFS Standalone Root, using DFS links to point to the Shared folders (aka targets).

- DFS Root can be stored on local Server as Standalone DFS Root or within Active Directory as Domain DFS Root.
- Celerra only supports Stand-alone DFS root as either single DFS Root with W2K or multiple DFS Roots with W2K3 Servers
- DFS Root on a global share [XP or Win2k3 only] can be viewed from any CIFS server on the Data Mover
- use Windows tools dfsutil.exe (use to create DFS root) and dfscmd.exe (use to manage DFS tree)

**Note** NAS does not currently support dfs query for siteinfo with Win2k version of dfsutil--Win2k3 version of dfsutil does not have this issue—AR70360:

**Failure:**

**C:>dfsutil /sitename:cel23cge0**

Above syntax works for Win2k3 version of dfsutil and Celerra, but not Win2k version--client connects to the svrsvc pipe and issues a "NetrDfsManager ReportSiteInfo" command. We return a DCE RPC Fault of 0x1c010002, or "Range Error".

**Success:**

**C:>dfsutil /sitename:cdms-source**

Microsoft(R) Windows(TM) Dfs Utility Version 4.0

Copyright (C) Microsoft Corporation 1991-2001. All Rights Reserved.

Site for cdms-source is Default-First-Site-Name

- Disable DFS on DART via registry: HKLM>SW>EMC>DFS>Enable=0
- Normal DFS Management should be done from MMC Console
- Celerra DFS does not support FRS (File Replication Service)
- Do not create DFS root when access-checking policy is UNIX or SECURE—DFS links cannot be created with UNIX rights
- dfscmd.exe tool can be used to administer DFS root content, creating & deleting links, but not to delete the DFS tree

**Note:** Default DFS cache value 1800 secs for TTL (time that Client will cache list from DFS Root Server)

**Debug:**

**\$server\_config server\_x -v “dfs audit” | “dfs check” | “dfs repair”** (repairs share attribute information)

**CREATING DFS ROOT ON CELERRA USING MMC or CLI TOOL:**

1. Open MMC DFS gui (Welcome to the New Root Wizard) and select next
2. Enter Host Server server name information as fully qualified name of CIFS server
3. Enter Root Type as \*Stand-alone root
4. Enter unique DFS Root Name and any comments
5. If Root Share is not already created, enter the path for the new DFS Root Share: C:\dfs\_root

**CLI EXAMPLE:**

**C:>dfsutil /AddStdRoot /Server:DM2\_cge0 /Share:dfs\_root**

**Note:** Both dfscmd.exe and dfsutil.exe tools are found in Win2k/Win2k3 Support Tools.

**CELERRA WIDE LINKS (Widelink) NAS 5.4:**

**What are Wide Links?**

Combined with the functionality provided by DFS, Celerra CIFS Servers can allow Clients to resolve UNIX absolute paths via symbolic links stored on the Celerra using a DFS Root service that maps the UNIX mountpoint path to the Server:\share\path. Allows CIFS to follow Symbolic Links between file systems on different Servers using DFS redirector (DM sends redirect to Clients)

- Wide Links is enabled by default on the Celerra when CIFS is running
- To disable Wide Links, enter 0 for following Data Mover Registry entry using Regedit.exe from Windows Client:  
HKLM>Software>EMC>DFS>Enable: 0
- A feature that works in conjunction with a DFS share. When a link contains absolute path, we return a DFS Referral to CIFS client with path of the target. Purpose of Widelink is to allow CIFS Clients to resolve absolute symbolic links using DFS <\\name\\share> paths, similar to way Unix clients can resolve absolute links normally via mountpoints. DFS only redirects directories, so DFS link must consist of a directory pathname.

→Widelink Server and Share specified in Registry: Set WideLink Share in Registry, Stop & Restart CIFS Service  
HKLM>SW>EMC>WideLink>Share <\\server\\sharename>

→Feature works with File System Linking when param shadow followabsolutepath=1 is set

**NAS 5.5.26.0 and higher extends param to allow CIFS Users to follow non-root owned symbolic links:**

**# server\_param server\_3 -facility shadow -info followabsolutepath -v**

```
server_3 :  
name      = followabsolutepath  
facility_name = shadow  
default_value = 0  
current_value = 0  
configured_value =  
user_action = none  
change_effective = immediate
```

range = (0,3)

description = Whether to follow the absolute path in symlinks

**param shadow followabsolutpath=0**

→Neither bit 0 or 1 is set and CIFS cannot follow Unix symbolic link paths

**param shadow followabsolutpath=1**

→=1 means that bit 0 is set & symbolic links are enabled--root-owned absolute path symlinks can be followed by CIFS Users

**param shadow followabsolutpath=3**

→=3 means that bit 0 & 1 are set, symbolic links are enabled, and CIFS Users can follow any absolute path symlink (See AR84916)

**CONFIGURING WIDE LINKS:**

1. Create Shares to be used for Wide Link

2. Configure Data Mover for DFS Standalone Root using MMC DFS Utility

DFS>Action>Show Root>enter compname of DM>Select & rightclick DFS root to be used>New Link>Enter Link name & Path to target

3. Create a second New Link for the Remote Server, specifying Link Name and Path to target

4. Make sure that Wide Link shares are listed in registry:

**HKLM>Software>EMC>WideLink>Share: \\server\sharename**

5. Stop and restart CIFS Service

**Requirements/Limitations:**

→Must have DFS Standalone Root shares configured on the Data Mover

→Symbolic links with absolute paths are seen as directories from Explorer

→Use WideLink only on directories

→Pathname in link must be the same between CIFS and UNIX names

→Supports CIFS VDM's

**CELERRA LOCAL USER SUPPORT**

**“Local Users Supported” or “Standalone” modes for CIFS (Beginning with NAS 5.4):**

**Purpose:** Allows for the creation of Local User accounts on a CIFS Data Mover, either on a CIFS Server joined to a Windows Domain (Local Users Supported mode), or as a completely Standalone CIFS server operating within a Workgroup environment (i.e., no Domain authentication).

**Benefits/Restrictions of Local User Support on CIFS Servers:**

→Administrators could access the CIFS Server if Domain authentication is not available, as long as you log into the local CIFS server using its Administrator account.

→No need for Domain structure when using Local Users created on DM, similar to Workgroup

→Encrypted passwd NTLM authentication

→Use ACLs with Local Groups and Local Users (up to 128 Users can be created)

→Can also use Domain structure and mix Local Users, just as in native Windows Servers

→Designed to replace Share and Unix Security modes

→Can only be used for CIFS access, not Unix access

→Local User UIDs are automatically assigned from a pre-determined range

→Unicode must be turned on for this feature to work

→Data Mover authentication must be NT

→Local Users Support does not support Kerberos authentication, so attempts to do so from hard-coded Clients, will fail—Celerra will use NTLM V1/V2

→Default Local User accounts are Administrator and Guest and are automatically generated when setting up Local Users Support (Up to 128 other local Users can be created—Guest is disabled by default)

→Once enabled, Local Users Support cannot be disabled (it would need to be completely broken down)

**PopUp Message if trying to create the 129<sup>th</sup> Local User:**

*Local Users and Groups*

*X The following error occurred while attempting to create the user bottste on computer 10.217.214.9:*

*Not enough server storage is available to process this command.*

**Two Modes of Local User CIFS Server Support**

1. Local Users Supported mode whereby existing domain CIFS Server is configured with Local User support/accounts

**CIFS Output for this mode:**

CIFS Server CIFS\_SERVER1[W2K] RC=4 (**local users supported**)

2. Local Users Supported mode (Standalone mode) completely outside any Domain authentication

**CIFS Output for this mode:**

CIFS Server(**standalone**) SERVE\_ALONE[EMC] RC=2

## **SETTING UP STANDALONE CIFS SERVER USING CLI:**

### **1. Create Standalone CIFS Server:**

```
# server_cifs server_2 -add standalone=serve_alone,workgroup=EMC,interface=192-1-10-22,local_users
```

server\_2 : Enter Password:\*\*\*\*\*

Enter Password Again:\*\*\*\*\*

Done

### **2. Verifying Standalone server:**

```
# server_log server_2 -s |tail
```

2009-01-29 15:38:32: SMB: 6: CIFS Server SERVE\_ALONE[EMC] created (0)

2009-01-29 15:38:32: LGDB: 6: {Administrator} password must change (C=0x1c982518c9f9400 LS=0x0

MA=0x8000000000000000)

2009-01-29 15:38:32: ADMIN: 6: Command succeeded: cifs add standalone=SERVE\_ALONE workgroup=EMC

localAdminPassword=\*\*\*\*\* interface=192-1-10-22

2009-01-29 15:39:49: LGDB: 6: Database consolidated

```
# server_cifs server_2
```

CIFS Server(standalone) SERVE\_ALONE[EMC] RC=2

### **3. Use Windows client to “Change Password” from above temporary password to a permanent password**

#### **a. Log into Windows client**

#### **b. Use ctrl + alt + delete sequence and select “Change Password”**

#### **c. Enter IP address of standalone CIFS server in “Log on to” box**

#### **d. Enter temporary password in “Old Password” box and then enter a new password**

### **4. Managing Local User accounts on the Standalone CIFS Server:**

**a. Enter the CIFS name in the local Windows system’s Host file:** c:\WINNT>System32>Drivers>etc>hosts file (Add to lmhosts file if Browsing is required on a network or if netbios Name Resolution is required and WINS is not established on the subnet)

**b. Issue the Net Use command to provide the security context for the Windows logon session:**

```
C:>net use \\serve_alone /user:serve_alone\administrator <password>
```

The command completed successfully.

**c. Open Computer Management>Connect to another computer: Name [Must enter CIFS server name, not IP address]**

**Note:** You will not be prompted for username and password. Console will open and local groups database will be manageable on the Server. Please note that the security credentials are only valid for the existing logon session. The net use step would have to be reissued to any system from which access is desired, and each time the Windows system is logged in to.

## **SETTING UP STANDALONE CIFS SERVER USING CELERRA MGR:**

Celerra Manager>Data Movers>server\_2>CIFS>CIFS Servers>New

→Select Standalone option, specify NetBIOS name and Workgroup, Set Local Admin Password, assign interface

→Start CIFS Service on separate screen or via CLI

```
# server_cifs server_2
```

CIFS Server(standalone) LONER[LONER] RC=2

Comment='EMC-SNAS:T5.6.37.6'

if=cge0 l=192.1.4.222 b=192.1.4.255 mac=0:60:16:c:51:3e

Password change interval: 0 minutes

```
/nasmed/quota/slot_2/.etc
```

-rw-r--r-- 1 root bin 1270 May 8 10:02 .db.5.localgroups

```
# cat .db.5.localgroups
```

# Localgroups database

#

\$RELEASE:5\$

@LONER:1:1000:AdMcAdMcAdMcIdAdAdAdAdAdAdAdAdAdAdAdAdAdAdAdAdAdMcAdMcAdMcBdMcAdMcIdAdAdAdAdAdAdA  
dAdAdAdAdAdAdAdAdAdAdAd:S-1-5-15-32434d45-2201-3bc5d9de-ffffffffff

#

# Localgroups of server LONER

#

LONER:\*Administrator:0x1f4:0x1007f:\

Built-in account for administering the computer/domain:\

%Administrator|11610#544|=KdMhAhIbDgFlDoNgFpMdHcCmDfOINmPbDjKdLlHdCdBnFgCiOlCfHoMcHfPoAfHpKgFaKd|

LONER:\*Guest:0x1f5:0x8:\

Built-in account for guest access to the computer/domain:\

%Guest|117|0#546|=KdCiIdDjNkKpCnNmGkBoEkEbDpLoPnIfHpKdAkBoAfHmKfHbAaOiKkDnElFdFlEbEaOoKd|

LONER:Administrators:0x220:0x1007f:\

Members can fully administer the machine:\



User name: Administrator

Log on to: <Compname or IP address of CIFS server>

Old Password: <enter passwd used in Step 1>

New Password: <new\_passwd>

Confirm New Password: <new\_passwd>

[PopUp window] “Your password has been changed.”

**Note:** In some cases, only the Compname will work in the “Log on to” box, not IP Address.

#### **4. Results of passwd change are not logged in Server Log but are seen in the Local Users group db:**

%Administrator|1|16|0|#544|=KdMnOgPhFpMgMfAkPpAfOfIiLcMjNiJpInKdOhGoOkHbApNaMjDoFdMeCmOoDgPaPeHIKd|21c  
9824baccc54801

#### **5. Managing Local User accounts for existing domain CIFS Server where Local Users support has been enabled:**

##### **Using Computer Management Snapin to access from a computer within the Domain:**

a. Open the Computer Management MMC snapin on any computer within the same Domain as the CIFS server>rightclick Computer Management>Connect to another computer>Name: <enter CIFS server name or IP address> and manage local users and groups

**Note:** As long as the CIFS name or IP is resolvable with DNS (generally would be since its also a member of the Domain), there is no need to add the CIFS server name and IP address to the local \winnt\system32\drivers\etc\hosts file.

##### **Using Computer Management Snapin to access from a computer outside the Domain:**

- a. Add the CIFS server name and IP address to the Windows local \winnt\system32\drivers\etc\hosts file
- b. Use the following command to establish the security context from the Windows management station:

**C:\>net use \\cifs\_server1 /user:cifs\_server1\administrator <password>**

The command completed successfully.

- c. Open Computer Management>Connect to another computer>Name: <enter Name or IP address>

**Note:** No username and password prompt will appear. Local groups can be accessed and modified.

#### **DELETING STAND-ALONE CIFS SERVER:**

**# server\_cifs server\_x –delete standalone=dm2\_local**

#### **DELETING CIFS SERVER ENABLED FOR LOCAL USERS SUPPORT:**

**# server\_cifs server\_x –delete compname=cifs\_server1 –remove\_localgroup**

**Note:** Use the –remove\_localgroup option only if you want to delete the CIFS instance from the .db.5.localgroups database.

#### **RESETTING THE LOCAL USERS/STANDALONE CIFS SERVER PASSWORD:**

**# .server\_config server\_2 -v "lg admin passwd=<temp\_passwd> vs=<standalone\_CIFS>"**

**Note:** This resets the temporary password to a known password, useful for staring over when logging in from a Windows client to change the temporary to a permanent password.

#### **LOCAL USERS/STANDALONE CIFS SERVER USES NEW LOCALGROUPS DB FILE:**

-rw-r--r-- 1 root bin 1361 Jan 29 2009 /nasmcdb/quota/slot\_2/.etc/.db.5.localgroups

#### **MOVING STANDALONE NETBIOS SERVER TO VDM CONTAINER:**

##### **1. Setup standalone netbios CIFS server:**

**# server\_cifs server\_2 -add standalone=squib,workgroup=emc,interface=squib,local\_users**

server\_2 : Enter Password:\*\*\*\*\*

Enter Password Again:\*\*\*\*\*

done

##### **2. Move standalone server to VDM container:**

**# server\_cifs server\_2 -move netbios=squib vdm1**

server\_2 : done

##### **3. Verify:**

**# server\_cifs vdm1**

CIFS Server(standalone) SQUIB[EMC] RC=2

Comment='EMC-SNAS:T5.6.43.4'

if=squib l=192.1.10.24 b=192.1.10.255 mac=0:60:16:1f:ad:ca

Password change interval: 0 minutes

#### **CELLERRA NT CREDENTIALS FOR UNIX ACCESS:**

When specifying the “ntcredential” option in the file system mount, NT Credentials will be built when NFS users access file system objects. The “ntcredential” is built from matching UID & GIDs to SIDs, merging Unix and Windows groups together. Main difference from pure Windows credential is that this credential does not contain local groups information.

## **PURPOSE OF NT CREDENTIALS FEATURE:**

- More consistent permissions on file system objects, regardless of access protocol
- Cache to store credentials, reducing access-checking times
- Unix access rights managed by ACL
- Eliminates 16-group limitation for Unix secondary groups

## **SETTING UP NT CREDENTIALS:**

**\$ server\_mount server\_2 -o accesspolicy=NT,ntcredential fs1 /fs1**

fs01 on /fs1 uxt,perm,rw,accesspolicy=NT,ntcredential

## **STEPS REQUIRED TO BUILD AN NT CREDENTIAL FOR NEW USER:**

1. Applies only to NFS access
2. DART checks UID of incoming User against existing NT Credential Cache
3. If UID is not in the cache, tries mapping UID to Windows SID
4. If UID-to-SID match fails, will try UID-to-Name match from local passwd file or NIS. If name is found, retrieves SID from DC.
5. If name found, but SID cannot be matched up, then a failed mapping entry gets put into cache, and file access reverts back to traditional Unix credentials
6. If name to sid match is found, retrieves group list for User from DC
7. Builds new Windows NT credential for NFS user and places credential into cache
8. Performs access checking

## **NT CREDENTIAL CACHE:**

→default ttl 20 minutes **param nfs NTcred.TTL=20**

**# server\_param server\_2 -facility nfs -i NTcred.TTL**

server\_2 :

```
name          = NTcred.TTL
facility_name = nfs
default_value = 20
current_value = 20
configured_value =
user_action   = none
change_effective = immediate
range         = (0,4294967295)
description   = NT credential TTL (in minutes) for NFS connections before renewing it
```

→default size 1009 entries: **param nfs NTcred.size=1009**

**# server\_param server\_2 -facility nfs -i NTcred.size**

server\_2 :

```
name          = NTcred.size
facility_name = nfs
default_value = 1009
current_value = 1009
configured_value =
user_action   = reboot DataMover
change_effective = reboot DataMover
range         = (0,4294967295)
description   = Global NT credential cache size for NFS connections
```

**\$ .server\_config server\_2 -v "ntcredcache list"**

List of the nt credential cache entries

-----  
uid 201 gid 65534 Mapping failed \*\* 1169 sec before expiration

List of the nt credential cache entries

-----  
uid 32768 gid 32768 User SID S-1-5-15-bfc56af5-d6cf8701-5f67b1a3-46c \*\* Expired\*\*

uid 32880 gid 32768 User SID S-1-5-15-bfc56af5-d6cf8701-5f67b1a3-4c3 \*\* 1130 sec before expiration

**Note:** Accessed fs from CS as ‘nasadmin’—gid 201 cannot be properly resolved—in this case, negative mappings are cached

**\$ .server\_config server\_2 -v "ntcredcache dump=201"**

==== UNIX NT credential entry dump

uid 201 credp 0xd5c81a04 (stamp=14930 sec): Valid for 1009 sec

uid 201 gid 65534 Mapping failed

**# .server\_config server\_2 -v "ntcredcache dump=98769"**

==== UNIX NT credential entry dump

uid 98769 credp 0xe0c11e04 (stamp=1895219 sec):\*\* Expired \*\*

1205946918: SMB: 6: USER 'W2K\ (mapped=)

Auth=UNIX CredCapa=0x0

SID = S-1-5-15-242a3a09-6bc5c62-3f32a78a-5e2f

1205946918: SMB: 6: PRIMARY = S-1-5-15-242a3a09-6bc5c62-3f32a78a-201

1205946918: SMB: 6: Priv=0x0,0x0 DefOpt=0x0 Adm=0, Backup/Adm=0 (Bkp=0) NTCapa=0x0

1205946918: SMB: 6: NT2UNIX: 4 groups

1205946918: SMB: 6: gid=0x30d49 S-1-5-15-242a3a09-6bc5c62-3f32a78a-200

1205946918: SMB: 6: gid=0x8004 S-1-5-15-242a3a09-6bc5c62-3f32a78a-5e2b

1205946918: SMB: 6: gid=0x30d50 S-1-5-15-242a3a09-6bc5c62-3f32a78a-201

1205946918: SMB: 6: gid=0xa S-1-5-12-2-a

1205946918: SMB: 6:

1205946918: SECURITY: 6: cred at 0xe0c11e04: uid=0x181d1, gid=0xa, inuid=98769

1205946918: SECURITY: 6: 3 other gids:

0x30d49 0x8004 0x30d50

### # .server\_config server\_2 -v "ntcredcache status"

NT credential cache status

Enable:1

DefaultSize:1009 →Default cache size, can be increased to a very large number of records

Hits:79

Miss:5

Count:1 Collisions:0 →Count:1 shows total mappings cached, which is one in this case

Total of collisions:0 Total of hash OK:2 Compare:81

### # .server\_config server\_2 -v "ntcredcache fulldump"

### \$ .server\_config server\_2 -v "ntcredcache reset"

Note: Use this command to flush entries from the ntcredentials cache

## MANUALLY BUILDING AN NTCREDENTIAL FOR A KNOWN UID:

### \$ .server\_config server\_2 -v "unixntcred uid=32773"

1145395999: SMB: 4: UnixId=0x8005 srvDomainSID=S-1-5-15-bfc56af5-d6cf8701-5f67b1a3-46c Srv='LAIP1\_DM2A' \\george.2k3.pvt.dns

1145395999: SMB: 3: getTrustedDCAddr:FQDN theTrustedDCName=george.2k3.pvt.dns nb=GEORGE dns=2k3.pvt.dns fqdn=george.2k3.pvt.dns

1145395999: SMB: 3: hepDNS:getTrustedDCAddr found IP=192.1.4.217 for DC=GEORGE

1145395999: SMB: 3: getTrustedDCAddr:use new DC=0xd90401c0 ip=192.1.4.217 name=GEORGE dom=2K3 dc=0xd4ea7404

1145395999: SMB: 3: setDCInformation:Use DC=GEORGE domain=2K3 Srv=LAIP1\_DM2A

1145395999: SMB: 4: >DC=GEORGE(192.1.4.217) R=10 T=1 ms S=0,1/-1

1145395999: KERBEROS: 7: send\_as\_request

1145395999: KERBEROS: 7: \_krb5\_use\_dns: use dns = yes

1145395999: KERBEROS: 7: krb5\_locate\_srv\_dns: for service \_kerberos at realm 2K3.PVT.DNS, # KDC: 1

1145395999: KERBEROS: 7: krb5\_locate\_srv\_dns: name george.2k3.pvt.dns

-----abridged-----

1145395999: KERBEROS: 5: Warning: send\_as\_request: Realm 2K3.PVT.DNS - KDC 0.0.0

.0 returned error: Additional pre-authentication required (25)

1145395999: KERBEROS: 7: send\_as\_request

1145395999: KERBEROS: 7: \_krb5\_use\_dns: use dns = yes

1145395999: KERBEROS: 7: krb5\_locate\_srv\_dns: for service \_kerberos at realm 2K3

.PVT.DNS, # KDC: 1

1145395999: KERBEROS: 7: krb5\_locate\_srv\_dns: name george.2k3.pvt.dns

-----abridged-----

1145395999: SMB: 5: checkDCBlob:TicketFlags doesn't match 17/3

1145395999: SMB: 7: RID=0201 GID=32775 A:7 U:2 ='Domain Users'

1145395999: SMB: 7: User 2K3\ primary GID set to 32775 (Unix=0xffff)

1145395999: SMB: 7: RID=0200 GID=32774 A:7 U:2 ='Domain Admins'

1145395999: SMB: 7: Primary=201 Nb=2 isValid=1 isAdmin=0

1145395999: VC: 5: abortCheckWait(smb\_share=0x0)

1145395999: SMB: 4: USER '2K3\'

Auth=UNIX CredCapa=0x0

SID = S-1-5-15-bfc56af5-d6cf8701-5f67b1a3-46c

1145395999: SMB: 4: PRIMARY = S-1-5-15-bfc56af5-d6cf8701-5f67b1a3-201

1145395999: SMB: 4: Priv=0x0,0x0 DefOpt=0x0 Adm=0, Backup/Adm=0 (Bkp=0) NTCapa=0x0  
1145395999: SMB: 4: NT2UNIX: 2 groups  
1145395999: SMB: 4: gid=0x8007 S-1-5-15-bfc56af5-d6cf8701-5f67b1a3-201  
1145395999: SMB: 4: gid=0x8006 S-1-5-15-bfc56af5-d6cf8701-5f67b1a3-200  
1145395999: SMB: 4:  
1145395999: SECURITY: 4: cred at 0xd4e9a004: uid=0x8005, gid=0xffffe, inuid=32773  
1145395999: SECURITY: 4: 2 other gids:  
0x8007 0x8006

**param cifs acl.extendExtraGid=0** →disables mapping Secondary UNIX groups to NT users

**param cifs acl.extendExtraGid=1** →Enables NT User to add Unix secondary groups to their credential, from NIS or ./etc

#### STORED ACL INFORMATION:

SIDs, GIDs, UIDs

#### **BENEFITS:**

- Consistent permissions on file system object
- Cache that stores NT Credentials, performance
- Management of UNIX access rights from Windows ACLs
- Lifts 16-group UNIX limitation for Users

#### TRADITIONAL CIFS BEHAVIOR FOR USERS:

- SIDs are only thing checked from ACLs

#### TRADITIONAL NFS BEHAVIOR FOR USERS:

- UID & GIDs are only thing checked for Unix credentials

#### NEW NT CREDENTIAL FOR UNIX PURPOSE:

→NT Credentials built for Unix User based on UID/GID-to-SID match of User/Groups in order to allow NFS access based on ACL checking (and to lift the 16-group UNIX limitation). NT Credential replaces the Unix credential for NFS access checking.

**Note:** NT & SECURE access policies means that ACLs are checked using UID/GID mappings as stored in the ACL

→With NT for Unix Access, the credentials are built from the UID and SID equivalent, plus the SIDs of all groups that the User is a member of, then added to cache with ttl of 20 minutes to reduce access-checking times & increase performance

→Purpose to manage UNIX access rights using Windows ACL

→NT Credential is built where possible, negative lookups cached, failures will result in Unix credential (old style)

**param nfs NTcred.winDomain=emc.com**

**Note:** Use of this param is required if there are multiple CIFS Servers on a Data Mover and defines the default Windows domain to be used in the environment. Param allows DM to use name from UID/GID to retrieve SIDs from DC. Failure to build an NT credential will revert to normal UNIX credentials.

**param nfs NTcred.size=1009**

**Note:** This cache contains a default value of up to 1009 NT Credentials, and/or negative UID/GID entries that could not be mapped to SIDs. This value is configurable. Cache entries TTL is 20 minutes by default. Oldest, ‘unused’ entries expire first.

**param nfs NTcred.TTL=20**

**Note:** Default value is 20 minutes for caching of NT Credentials.

**param cifs acl.extendExtraGid=0**

**Note:** When disabled (default), the built NT credential for the User is based only on Windows groups, not Unix.

**param cifs acl.extendExtraGid=1**

**Note:** When enabled, NT Credentials are built using both Active Directory Groups and Unix Secondary Groups, using either NIS, local passwd/group files, or LDAP.

#### HOW SID LOOKUPS ARE CONDUCTED:

1. The Unix UID is received & matched to SID using Secmap Cache directly, or using Passwd/Group files, or NIS by using “user.domain” or “user” name query to DC’s

2. Once User SID is determined, DM queries DC’s for list of Group SIDs, if not found in SecMap cache

#### **Misc:**

Accesspolicy=MIXED | MIXED\_COMPAT

param nfs NTcred.winDomain=name\_of\_domain

param nfs NTcred.size=1009 (default entry)

param nfs NTcred.TTL=20 (default = 20 minutes)

“ntcredcache status” →Statistics on cache access

| reset →Flushes cache | list →Lists UID/SIDs in cache and negative mappings | fulldump →UID/SIDs & GID/SIDs

| dump=uid →List credentials for particular UID | remove=uid →remove from cache

#### CELLERRA MIXED & MIXED COMPAT ACCESSPOLICIES (new with NAS 5.4):

- Original Access Policy models had Unix Mode bits and CIFS ACLs managed and checked independent of each other, with no synchronization between the two
- New Access Policy called “Mixed” in which UNIX ownership mode bits are translated into ACE’s “Owner, Group, & Everyone”,
- MIXED Policy synchronizes UNIX & Windows permission as closely as possible using algorithm to translate UNIX rights (mode bits) into ACLs and Windows ACLs into UNIX rights.

### **MIXED ACCESSPOLICY:**

- Works with either Compat or DIR3 file systems
- UNIX ownership mode bits are translated into (3) ACEs: Owner, Group, Everyone (MIXED\_COMPAT does not translate Group)
- ACL and UNIX mode bits are synchronized to create a single ‘merged’ list of attributes
- Mixed Access policy information is stored within the File System, not on the Control Station, but controlled by CS not DART?
- a Chmod will overwrite the equivalent ACL setting
- a SetACL will try to create equivalent UNIX Mode bits
- When determining access rights, DART will always use ACLs, for both NFS & CIFS clients, but only for MIXED
- Last protocol that modifies a file system object’s security is used to generate the new ACL, and that ACL will be used when granting or denying access to both NFS or CIFS clients
- Mixed Access Policy designed to be used with NFSv4
- Slightly modified “Mixed\_Compat” access policy is to be used when displacing NetApp systems. Mixed\_Compat does not translate UNIX groups into Group ACEs, and instead creates an Everyone ACE for the UNIX group information
- File Systems must go through a translation process when changing to “Mixed” or “Mixed\_Compat”
- Once translated to Mixed accesspolicy, cannot properly revert to the original accesspolicy

### **MIXED TRANSLATION OF UNIX MODE BITS INTO ACLS:**

- (3) ACL entries are created based on the UNIX mode bits of Owner, Group, Other: File Owner, Group, Everyone
- Delete/Change permissions & Take Ownership are set for the Owner but not for Everyone or other Groups

### **MIXED TRANSLATION OF ACLS INTO UNIX MODE BITS:**

- Windows ACL Deny option is not translated, but the Allow ACL option is (same for both MIXED & MIXED\_COMPAT)
- UNIX Owner mode bits are built from Owner entry, the ACE of the file/directory Owner, and the Everyone Group ACE
- UNIX Group mode bits are built from Group entry, the ACE of the primary group Owner, and the Everyone Group ACE
- UNIX Other mode bits are built from Other ACE, all ACE’s different from User/Group, and the Owner/Group ACE

### **MIXED COMPAT ACCESSPOLICY:**

- When determining access rights, DART will use the ACL when checking NFS or CIFS access, but only when there is an EXPLICIT ACL, meaning that the file or directory perms were last set or changed from a CIFS Client
- When a file or directory permission was last modified by an NFS client, access checking for both NFS and CIFS is done against the UNIX mode bit permissions—an ACL gets created based on the Mode bits, but is not enforced
- Access checking is therefore independent of protocol
- The last protocol used to set permissions or ownership on a FS object is used to recreate the ACL for that object, but only when the ACL is created or modified from CIFS does it become EXPLICIT, meaning that access checking is done against the ACL and not UNIX mode bits
- Celerra synchronizes and merges Unix and Windows permissions by translating UNIX rights to ACLs, and WINDOWS rights to Unix mode bits

#### **EXAMPLE:**

If a CIFS User1 is given RWX permissions to a file, an ACL is created, and UNIX mode bits would read –RWX for Other, but in reality, only User1 would have RWX from NFS, and nobody else in the Other category would

→For MIXED or MIXED\_COMPAT, the default group should be set for Windows users, as this group will become the primary UNIX group when file system objects are created from CIFS

- NetApps does not map group SIDs to UNIX GIDs
- UNIX Mode bits are only translated into ACLs as Owner and for the Everyone group
- Since UNIX groups are not translated into a Group ACE, the Everyone group ACE is generated for the ACL
- Delete/Change permission and Take Ownership always set for File Owner and never for Everyone

### **MIXED COMPAT TRANSLATION OF ACLS INTO UNIX MODE BITS:**

- Windows ACL Deny option is not translated, but the Allow ACL option is (same for both MIXED & MIXED\_COMPAT)
- Builds None, Owner, and granted ACEs into Group and Other UNIX mode bits (different than MIXED policy)

### **MIXED COMPAT TRANSLATION OF UNIX MODE BITS INTO ACLS:**

- Only two entries are created in the ACL: Owner & Everyone (MIXED policy builds (3) entries in ACL)
- An Everyone Group ACE is created from Group mode bits, since other groups are not translated with this policy
- The Other mode bits are ignored and not used when building the ACL
- Delete/Change permissions and Take Ownership are set for the File Owner but not for the Everyone group

## **CHANGING EXISTING FILE SYSTEM ACCESSPOLICY TO MIXED:**

1. Set Accesspolicy:

**#server\_mount server\_x -o accesspolicy=MIXED | MIXED\_COMPAT fs1 /fs1**

2. Conduct Translation:

**#nas\_fs -translate fs01 -access\_policy start -to MIXED -from NATIVE | NT | UNIX | SECURE**

3. Verify Translation:

**# nas\_fs -translate fs01 -access\_policy status**

1145387101: CFS: 7: conversion in progress : 829/830 inodes done 99 % done

status=In progress

percent\_inode\_scanned=99

**# nas\_fs -translate fs01 -access\_policy status**

1145387319: UFS: 6: inc ino blk cache count: nInoAllocs 2: inoBlk d5a73584

status=Completed

## **SERVER LOG ENTRY:**

2006-04-18 15:04:49: ADMIN: 4: Command succeeded: acl database=/fs01 convertAccessPolicy start old\_policy=NATIVE

2006-04-18 15:05:01: SVFS: 4: D304131\_0: allocDataBlock: transition index state

2006-04-18 15:06:17: CFS: 4: conversion done status 0 : 829/830 inodes done 99% done

## **MIXED or MIXED COMPAT ACLDUMP INFORMATION:**

**# server\_mount server\_2**

fs01 on /1 uxf,perm,rw,accesspolicy=MIXED\_COMPAT | MIXED

**# ls -la lhead**

drwxrwxrwx 5 32773 32773 1024 Apr 17 10:27 copy

**# chmod 755 copy**

**# .server\_config server\_2 -v "acl dump=/fs01/copy"**

1145388087: SMB: 3: Dump with no ThreadCtx

Dump of rights of /fs01/copy

===== UNIX =====

USER 0x8005 GROUP 0x8005 mode=rwxr-xr-x

===== NT =====

aclId=0xd

controlSummary = 0x8004

**Owner=+USER 0x8005 S-1-5-12-1-8005 →+USER shows directory permissions created or modified by NFS Client**

**Owner=USER 0x8005 S-1-5-12-1-8005 →USER shows directory permissions created or modified by CIFS client**

Group=GROUP GID=0x8005 S-1-5-12-2-8005

**Note:** MIXED & MIXED\_COMPAT are special accesspolicies in which the last protocol that modifies permissions on files or folders, is what is used to generate the resultant ACL. But, a big difference with MIXED is that ACLs are always used for access-checking, whether NFS or CIFS client. MIXED\_COMPAT access-checking is dictated by the last protocol that modified the permissions.

## **CELERRA KERBEROS CREDENTIAL MANAGEMENT:**

### **Dump Cache:**

**#server\_kerberos server\_2 -ccache**

server\_2 :

Dumping credential cache

Names:

Client: WBCROOT\$@STJOSEPHSWB.COM

Service: WBC-DC.STJOSEPHSWB.COM

Target: HOST/WBC-DC.STJOSEPHSWB.COM@STJOSEPHSWB.COM

Times:

Auth: 04/07/2006 04:38:12 GMT

Start: 04/07/2006 04:38:12 GMT

End: 04/07/2006 14:37:53 GMT

Flags: PRE\_AUTH,OK\_AS\_DELEGATE

Encryption Types:

Key: rc4-hmac-md5

Ticket: rc4-hmac-md5

-----abridged-----

### **Flush Cache:**

**#server\_kerberos server\_x –ccache –flush**

krb5.account → File contains Computer account information

krb5.keytab → File contains keys for encryption/decryption of Kerberos Service Tickets

### **CELERRA PARAM TUNING:**

Objective is to be able to view and tune Server parameters from either CLI or GUI

**#server\_param server\_2 –facility cifs –info acl.extendExtraGid -verbose**

**#server\_param server\_2 –facility NFS –modify acl.checkacl –value 0**

### **CELERRA NDMP B2D (Backup-to-Disk):**

→ Built on dedicated “cartridge” ufs file systems using VTLU [Virtual Tape Library Unit]

→ NDMP2D can be employed as two-way or three-way solution using Data Management Application (DMA) to control backup and recovery of an NDMP host, which is accomplished in NAS 5.4 through use of VTLU for Data Mover

→ Both Legato Networker and Veritas Netbackup are ISV backup software solutions

### **TWO-WAY NDMP B2D SOLUTION:**

DMA NDMP Backup Server uses NDMP to tell Data Mover to backup local VTLU file system—indexing information is sent to DMA Host over network. DMA Host never touches actual data. Optionally, data mover could have physical tape drives attached and backup to tape.

### **THREE-WAY NDMP B2D SOLUTION:**

DMA Host instructs source DM to begin backup, which sends data over network to DM containing the VTLU. Indexing information sent over network to DMA host. Optionally, DM with VTLU could have physical tape drives attached for writing backups to tape.

### **B2D (BackUp-to-Disk):**

Growing trend to use server-based arrays using cheap large capacity ATA drives for backups

Backups generally faster and more reliable

### **CONFIGURING VTLU FOR B2D ON CELERRA USING CELERRA MANAGER:**

1. VTLU properties, set Chain desired for Servers [Default chain is 0]
2. Create File System for use by Tape Storage
3. Create VTLU on Server using VTLU>new>Server\_2>accept defaults and record target & lun info for Robot Device
4. Create virtual Tape Media Storage on VTLU file system: VTLU>properties>create Tape Media Storage

**Note:** Select VTLU file system, create Media Tapes, specify size, specify ‘vault’ for location if you intend to move from Celerra Mgr to Import/Export, use Barcode Prefix (2GBtapes). Backup SW alone moves tapes from or to the ‘slot’ location.

5. VTLU>Properties>Tapes: Record Target & Luns for each Media Tape created

6. Move Media from Vault to Import/Export Slots using “Insert” (so that backup software can utilize)

**Note:** Move back to Vault using “Eject” button.

### **CONFIGURING VTLU FOR B2D ON CELERRA USING CLI:**

**1. Create file system for use as VTLU Storage**

**2. Set Chain for Server if desired & Verify:**

**# server\_vtlu server\_2 –service –set –chain 2**

**# server\_vtlu server\_2 –service -info**

**3. Create VTLU with default settings & Verify:**

**# server\_vtlu server\_2 –tlu -new**

**# server\_vtlu server\_2 –list**

server\_2 :

|    |        |          |          |               |             |
|----|--------|----------|----------|---------------|-------------|
| id | vendor | product  | revision | serial_number | device_name |
| 2  | ADIC   | Scalar24 | 1.0      | gkzp9492k5    | c9000t010   |

**4. Create Tape/Media & Verify:**

**# server\_vtlu server\_2 –storage –new vtlu\_fs3 –tlu 3 –tapesize 2G –tapes 4 –barcodeprefix 2GBTapes – destination vault**

**\$ server\_vtlu server\_2 -storage -list 2**

server\_2 :

|        |            |                |
|--------|------------|----------------|
| tlu_id | filesystem | barcode_prefix |
| 2      | NDMP_VTLU  | 0022           |

**5. List out Tape Drives and record Target & Luns for Backup SW Configuration:**

**# server\_vtlu server\_3 -tape -list 3**

server\_3 :

|          |            |              |          |             |
|----------|------------|--------------|----------|-------------|
| barcode  | filesystem | capacity(GB) | location | source_slot |
| 00220000 | NDMP_VTLU  | 400          | drive:1  | 0           |

```
00220002 NDMP_VTLU 396 slot:2 2
00220003 NDMP_VTLU 396 slot:3 3
00220004 NDMP_VTLU 396 slot:4 4
```

### # server\_vtlu server\_2 –drive –list 3

```
server_3 :
drive_id device_name serial_number status tape_barcode
0 c9000t011 h9SJIC000 empty
1 c9000t012 xFgSmIC000 in_use 00220000
```

#### EXTENDING TAPE STORAGE:

1. Extend File System first:

```
# nas_fs –xtend vtlu_fs3 size=4G pool=symm_std –o slice=y
```

2. Extend Tape Storage:

```
# server_vtlu server_2 –storage –extend vtlu_fs3 –tlu 3 tapesize 1G –tapes 4
```

#### CONFIGURING CELERRA FOR B2D BACKUPS:

1. Verify that NDMP & PAX params are set correctly

```
# server_param server_2 –facility NDMP –info –all
```

```
# server_param server_2 –facility PAX –info –all
```

**Note:** Modify param if required using –modify paxWriteBuff –value 128 syntax

2. Create NDMP account on DM

```
# server_user server_2 –add –md5 –password ndmp
```

#### VTLU PRINCIPLES:

(4) Parts of a VTLU Tape Library: Vault, Import/Export Slots (0-12); Regular Slots (0-500); (4) Tape Drives

--Tapes move from the Vault through the Import/Export Slots to the Regular Slots. The best situation is to load up to (12) tapes in the Regular Slots, as this is the only location that the Backup software can manage them—Backup sw cannot touch tapes in the Vault.

--File Systems are created for each VTLlibrary using the VTLU GUI

--An “svtl” database is created on the root of each data mover, consisting of a tables directory that holds VTLU configuration information in the following directories in the form of links: connections, drives, elements, inquiry, libraries, servers

→connections directory holds information on robots & drives

→drives directory holds compression information, which is off by default for VTLU configs

**Note:** SW compression only gains 1:1.75 compression ratio—can only be set using .server\_config, not GUI

→elements directory holds element info on each Library

→inquiry directory holds EMC and svtlRobot information

→libraries directory shows configuration: 1 4 32 8 0 1603737314040820 (robots, drives, slots, Import/Exp slots, vault slots, and timestamp)

```
$ .server_config server_2 –v "svtl"
```

```
$ server_param server_2 -f svtl -l
```

server\_2 :

| name            | facility | default | current | configured |
|-----------------|----------|---------|---------|------------|
| discardTapeData | svtl     | 0       | 0       |            |
| dbLocation      | svtl     | '/svtl' | '/svtl' |            |

#### Checking Robot Configuration:

```
$ .server_config server_2 -v "svtl tlusInfo tluld=1"
```

```
$ .server_config server_2 -v "svtl probeAll devType drive details=yes" (query each device)
```

```
$ .server_config server_2 -v "svtl probe device svtl0001 details=yes"
```

#### Checking VTLU Tape Configuration:

```
$ .server_config server_2 -v "svtl showDrvParams tluld=1"
```

```
$ .server_config server_2 -v "svtl setDrvParams tluld=1 drive=0 compression=zlib maxRate=10"
```

```
$ .server_config server_2 -v "svtl configChains startChain=50"
```

#### Checking for Free Space on Tapes:

```
$ .server_config server_2 -v "svtl storageInfo path=/vtlu_3"
```

**Note:** Shows all tapes for the VTLU file system—this info cannot be seen from GUI—Backup sw can see free space if tapes are in “slots”

#### VTLU BEST PRACTICES:

→Always configure as many “slots” as there are tapes and load all tapes to slots, because Backup SW can see tapes only in the “slots”

→After configuration, all operations should normally be done from the Backup sw side

**NAS 5.4 CELERRA CWORM (Write Once Read Many):** Compliance & Enterprise Levels

**FILE LEVEL RETENTION CAPABILITY:** New marketing name for this feature

File Systems that can be committed to a CWORM state with retention period—default Celerra Retention Period is indefinite

**Note:** Files show default infinity with “December 31, 1969” timestamp. Current release does not require CWORM License. CWORM file system must be created on a new ufs file system, and is applied on a per-file basis only

CWORM can be used on NFS or CIFS implementations

CWORM cannot be converted back from committed state, or otherwise reverted, but can be deleted by an administrator

**CWORM FILE SYSTEM STATES:**

WORM\_CLEAN: CWORM file state is CLEAN (files can be modified, etc)—initial state of newly created fs before files committed

WORM\_OK: (cannot be modified or renamed--committed) Files are RO, cannot be modified, moved, extended, renamed.

WORM\_EXPIRED: (Commit interval expires (Retention), files still cannot be modified/renamed--can be deleted by Owner/Admin)

**DETERMINING CWORM STATE OF FILES:**

**\$ .server\_config server\_2 -v "file query worm\_state /fs1/cworm/file.txt"**

1112722242: CFS: 4: cworm\_state of file /fs1/cworm/file.txt = WORM\_OK

**CREATING CWORM FILE SYSTEM & COMMITTING FILES:**

1. Create Cwom File System Type:

**# nas\_license -create cworm=<key\_code>**

**# nas\_fs -n worm\_fs -create v132 worm=enterprise**

**# nas\_fs -n worm\_fs -create size=500M pool=<pool> worm=enterprise -o slice=y**

2. Set Retention period using Unix Utime Utility (if specific expir. Time desired other than default ‘infinite’):

**# touch -at 201011200510 filename**

3. Commit file to CWORM by changing to R-Only or R-X-Only in Unix, or setting file attributes to RO in Windows:

**# chmod 444 filename** [If file is an executable, set to 555]

**# chmod 555 filename**

**Note:** If you forget to specify the retention period, the default will be ‘indefinite’, however, user can set new retention period

4. Verify using following:

**# nas\_fs -i cworm\_fs1**

**# ls -ul**

**Verifying Whether File System CWORM or not:**

**# grep file5 /nas/volume/filesys**

28:file5::0:1::y:1,c:97:1::::0::24::0:0:-1,-1: →Blue highlight fields indicate CWORM properties

**# nas\_fs -i file5**

id = 28

name = file5

acl = 0

in\_use = True

type = ufs

worm = compliance with no protected files

worm\_clock= Fri Nov 13 10:55:26 EST 2009

worm Max Retention Date= No protected files created

**# .server\_config server\_2 -v "file query 28"** (28 = fsid from nas\_fs -l)

1258127987: CFS: 6: isWorm = 1

**CWORM COMMIT:** Clearing Write bits!

→You cannot delete or modify (write), or overwrite a committed file. Unix will say permission denied, Windows will say “Cannot create the....file” and will prompt you to save file as a new file. If you try to change attributes on Windows file, you will get ‘Access is denied’ message. Please note that a copied CWORM file can be written to, and therefore is not in a committed state.

**Note:** CWORM enforces DOS RO bit mode for Windows (param cifs ReadOnly.Comp) and R-W Mode for UNIX

**CWORM & WINDOWS:**

1. Use Windows setTime Utility to set Retention period via CIFS

2. Commit file to CWORM by changing to RO using Windows property tab

**CWORM INTEROPERABILITY/SUPPORT:**

SnapSure and Replicator can be restored to a cworm fs after confirmation message answered

CDMS Migrations can be done to CWORM file system

Backup restores are not allowed to overwrite existing files

HighRoad supported

CAVA (scan files prior to committing)

NFS Clusters/Nested Mounts (but root of nested mounts cannot be CWORM file system)

**NOT SUPPORTED CWORM:**

DHSM, NFSv4

**CELERRA NMFS (NESTED MOUNT FILE SYSTEM):**

- Intro:** New file system type 102 that contains the nested root file system (RO) and its component file systems underneath, with purpose to provide ability to manage a collection of file systems under a single mountpoint.
- Creation of nested mountpoints only at the data mover root. Underneath the mountpoints are the component, or real file systems
  - Feature allows Celerra to present a single virtual filesystem with multiple diff. fs underneath (nested)
  - Clients will see component file systems as an Export for NFS or Share for CIFS
  - Server\_df will show total size of NMFS (combines file system sizes that are nested)
  - File System Export permissions are inherited from NMFS root level, but if permissions are exported on individual file systems, they override nested mountpoint export permissions when there is a conflict (Enforced at local file system level)
  - Cannot create mountpoints on the NMFS root (RO Root)
  - CIFS will be RO unless component file systems are properly Shared
  - No HardLinks or renaming across nested file systems
  - NFSv3 or higher only

**NMFS OPERATION:**

- Space is managed by each component file system, but will display at NMFS
- Quotas managed only at component file system
- Cannot have hard links, renaming, or copying of files between component file systems
- Component file systems must be ‘Shared’ for CIFS RW—otherwise, component file systems will inherit RO of NMFS root
- Server\_df should show total aggregated file system space, and individual file system components

**NESTING UXFS FILE SYSTEMS UNDER ANOTHER UXFS FILE SYSTEM:**

- Evidently, our earlier NAS versions (prior to 5.5.27) allowed the nesting of one filesystem underneath another, but this is not encouraged or really supported
- NMFS is not the same thing as nesting one file system under another. Rather, it is the nesting of multiple ufs file systems under a common /nmfs file system object and mountpoint, with each ufs file system having its own mountpoint defined under the toplevel NMFS mountpoint

/nmfs  
/nmfs/fs1  
/nmfs/fs2, etc.

**Hidden Option to Nest One File System under Another:**

**# server\_mount vdm2 -nested vdm\_fs9 /nmfs\_vdm2/test/vdm\_fs8/vdm\_fs9**

vdm2 : done

**FAST MOUNT ISSUE WITH NMFS & NESTED FILE SYSTEMS:** NAS 5.6.47 & emc233009

vdm\_fs9 on /root\_vdm\_2/nmfs\_vdm2/test/vdm\_fs8/vdm\_fs9 ufs,perm,rw,<unmounted>

**server\_log:**

2010-02-05 14:31:25: NMFS: 3: getNodeByName: thr 0x22ff08b0 findNode vdm\_fs8 NULL, status 5 - InvalidArgument  
2010-02-05 14:31:25: CFS: 3: file mount ufs rw /root\_vdm\_2/nmfs\_vdm2/test/vdm\_fs8/vdm\_fs9 123=42 rw: failed. Error NotFound

2010-02-05 14:31:25: UFS: 4: 48: The file system 42 mount failed during failover because the file system was corrupt.

**Sys log:**

Feb 5 14:31:25 2010:DART:UFS:WARNING:48:Slot 2:::1265398285:The file system 42 mount failed during failover because the file system was corrupt.

**Workaround—Disable the maxParallelMount parameter and reboot:**

# vi /nas/server/slot\_2/param  
param cfs maxParallelMount=0

→ Or, if already at 5.6.47, simply reissue the mount command using the –nested option, then add the param, and reboot

**SETTING UP CELERRA NMFS FOR NFS:**

1. Create the ‘User’ file systems using Celerra Manager, then permanently unmount the file systems
2. Create the NMFS object:

**# nas\_fs -name nmfs -type nmfs -create**

```
id      = 36
name    = nmfs
acl     = 0
in_use  = False
type    = nmfs
worm   = off
volume  = 0
```

```
pool      =
rw_servers=
ro_servers=
rw_vdms   =
ro_vdms   =
auto_ext = no,virtual_provision=no
deduplication = unavailable
stor_devs =
disks     =
```

**Note:** No actual volumes are used when creating the NMFS object

3. Create the NMFS object mountpoint:

**# server\_mountpoint server\_2 -create /nmfs**

server\_2 : done

4. Mount the NMFS object to the NMFS mountpoint as “Read-Only”:

**# server\_mount server\_2 -o ro nmfs /nmfs**

server\_2 : done

5. Export the NMFS object for NFS:

**# server\_export server\_2 -P nfs /nmfs**

server\_2 : done

6. Create mountpoints for the component file systems that will be nested under the NMFS ‘file system’:

**# server\_mountpoint server\_2 -create /nmfs/nmfs1**

server\_2 : done

**# server\_mountpoint server\_2 -create /nmfs/nmfs2**

server\_2 : done

7. Mount the component file systems to the NMFS

**# server\_mount server\_2 nmfs1 /nmfs/nmfs1**

server\_2 : done

**# server\_mount server\_2 nmfs2 /nmfs/nmfs2**

server\_2 : done

8. Verify:

**# server\_mount server\_2**

-----abridged-----

```
nmfs on /nmfs nmfs,perm,ro
nmfs1 on /nmfs/nmfs1 udfs,perm,rw
nmfs2 on /nmfs/nmfs2 udfs,perm,rw
```

## **SETTING UP NMFS FOR CIFS VDM:**

1. Create CIFS VDM container:

**# nas\_server -name vdm1 -type vdm -create server\_2**

2. Create NMFS object:

**# nas\_fs -name nmfs\_vdm -type nmfs -create**

3. Create NMFS object mountpoint on the VDM:

**# server\_mountpoint vdm1 -create /nmfs\_vdm**

4. Mount the NMFS object to the NMFS mountpoint RO on the VDM:

**# server\_mount vdm1 -o ro nmfs\_vdm /nmfs\_vdm**

5. Export the NMFS object for CIFS on the VDM:

**# server\_export vdm1 -P cifs -name nmfs\_vdm /nmfs\_vdm**

6. Create mountpoints for the component file systems on the VDM:

**# server\_mountpoint vdm1 -create /nmfs\_vdm/vdm\_fs1**

**# server\_mountpoint vdm1 -create /nmfs\_vdm/vdm\_fs2**

7. Mount the component file systems on the VDM:

**# server\_mount vdm1 vdm\_fs1 /nmfs\_vdm/vdm\_fs1**

**# server\_mount vdm1 vdm\_fs2 /nmfs\_vdm/vdm\_fs2**

8. Verify:

**# server\_mount vdm1**

vdm1 :

```
nmfs_vdm on /nmfs_vdm nmfs,perm,ro
```

vdm\_fs1 on /nmfs\_vdm/vdm\_fs1 ufs,perm,rw

vdm\_fs2 on /nmfs\_vdm/vdm\_fs2 ufs,perm,rw

## **DETERMINING NMFS OBJECTS:**

**\$ nas\_fs -l**

```
141  y  102 0  0      nas_nmfs2      1 →NMFS filesystems are type "102"  
144  y  102 0  0      home2        1
```

**\$ nas\_fs -i nas\_nmfs3**

```
id      = 146  
name    = nas_nmfs3  
acl     = 0  
in_use  = True  
type    = nmfs
```

**\$ .server\_config server\_2 -v "file mountdisplay"**

Current Mounted File Systems are:

uxfs rw /nmfs/nmfs2 113=35 rw →Shows that this ufs file system is mounted to the NMFS object 'nmfs'

uxfs rw /nmfs/nmfs1 111=34 rw

nmfs ro /nmfs 0=36 ro

## **DATA MOVER CAPACITY & FSCK SUPPORT:**

→NAS 5.4 will support up to 16TB file size on NS700 & Hammerhead DM's [but limited by SCSI to 2TB metavolumes]

→507 DM models or lower, are not supported

→Uses Intelligent Caching Scheme

→iSCSI LUNs will increase from 1TB max to 2TB max size

→FSCK will run integrated ACLCHK at the same time

→Up to two concurrent nas\_fsck's can be run per DM

# nas\_fsck -aclchkonly [Option for running aclcheck only—cannot be run with fs exported]

→Server may initiate an autofsck if corruption is detected—system will panic and run autofsck

**# nas\_fsck -fsckonly** [Option for running fsck only]

**param ufs skipFsck=1** [Set value to 1 to disable auto-fsck for corrupted file systems on reboots—corrupted file system will remain unmounted while all other file systems are mounted and needs manual fsck and remounting]

**param ufs corrupt=0** [disabling auto-fsck]

## **CELLERA SNAPSURE ENHANCEMENTS NAS 5.4:**

→Support for up to 64 checkpoints per fs, checkpoint Version V2.5

→Memory limitations of checkpoints have been lifted, now only restricted by SavVol size. Prior to V2.5, BlockMaps resided in DM memory. Now BlockMaps are paged to and from the SavVol.

→NAS Upgrades will convert Checkpoints to V2.5

→Checkpoint & Replicator are limited to 1GB memory—Deltasets for Replicator are limited to 8GB in size

→High Capacity Mode for >2TB file systems once systems running NAS 5.4 and all Checkpoints converted to V2.5

→V1 Checkpoints [i.e., NAS 4.2] will not convert—upgrade will fail

**#server\_sysstat ALL -blockmap** [page in rate page out rate →average pages into & out of SavVol for last 3 minutes]

**\$ server\_sysstat server\_2 -blockmap** (enhanced memory for checkpoints--25% physical RAM reserved for blockmap)

server\_2 :

```
total paged in      = 0  
total paged out     = 17416  
page in rate        = 0  
page out rate       = 0  
block map memory quota = 1048576(KB)  
block map memory consumed = 16272(KB)
```

**file systems greater than 2TB = disallowed**

**Note:** Above 'disallowed' message indicates that all checkpoints and replication delta sets have not yet been converted to high capacity version 2.5 format after a NAS 5.4 upgrade, meaning that the Control Station is still in low capacity mode and file systems cannot be extended past 2TB in size

## **VERIFYING DART HIGH CAPACITY MODE:**

**\$ .server\_config server\_5 -v "param SVFS useBtree"**

SVFS.useBtree INT 0x018f66b8 1 1 (0,4294967295) FALSE REBOOT 'NA'

**\$ .server\_config server\_5 -v "param blockmap conversion"**

blockmap.conversion INT 0x018f69d8 1 1 (0,4294967295) FALSE REBOOT 'NA'

## **CELLERRA MANAGER/MONITOR ENHANCEMENTS:**

- Navigation Tree enhancements with blinking status icon for new events; Tooltip text with number of alerts, hardware issues; Acknowledge Status of events to turn off blinking
- Status Monitor is a standalone version of navigation tree that can be launched—no expansion nodes, use as Monitoring window
- Alert Enhancements for CS, VC, FS Full & Usage predictions, Control LUNs [launch Navisphere]
- Default Monitor polling is 5 minutes, default polling for File System space and inode usage is 10 minutes

## **NAS 5.4 IP REPLICATION ENHANCEMENTS:**

- Introduction of GUI to setup and manage replication [CLI management uses fs\_replicate, fs\_copy, fs\_ckpt]
- Source/Destination timeout values at minimum of 600 secs—Timeout value is interval that Replication Svc uses between delta set creation and/or playbacks
- Source/Destination HWM represents Delta Set size in MB in which Replication Svc will be triggered to replay a delta set on SavVol. Default value is 600
- Source & Destination requires use of “rdfadmin” passphrase
- SavVol size for replication is 1-500GB, but VDM replication defaults to 1GB
- Must first establish connections between replication sites using shared passphrase
- GUI can setup file systems and VDMs
- Replication can be setup in two basic modes:

Immediate Method → Baseline source copied to destination, replication then started, and differential copy completed

Physical Transport → Requires creation of checkpoint backed up to media which is shipped to destination site and then loaded to the destination file system—when loaded, replication will then create the differential copy and start replication [data file systems only]

→ Destination VDMs & Filesystems are created under Destinations tab

→ Failover always initiated from Destination side

### **TERMS:**

#### **Delta Set:**

Source file system block updates [or changes] that are used to update destination file system. Minimum size of delta set is 128MB and max size is 8GB

#### **SavVol:**

Volume used to store copied data blocks from source fs locally and a remote SavVol for storing changes until Playback commits changes to remote file system

#### **VDM:**

Ability to administer and replicate a CIFS server from one location to another

#### **Replication Policies on Source:**

- Continue fs access even if replication will become out-of-sync [for whatever reason]
- Upon exceeding SavVol size and buffer changes, stop writes to buffer but allow reads [AutoRO]
- Upon exceeding SavVol size and buffer changes, stop all reads and writes to buffer [AutoFreeze]

**Note:** After changing timeout, hwm, or policy, changes will show as ‘pending’ until an update occurs

#### **IP REPLICATION COMMANDS/OPTIONS:**

Refresh → creates new delta set on demand, same as what replication svc will do when reaching timeout or HWM, starts playback of delta set on destination side

Suspend → Stops delta set playback and the replication service, but leaves in restartable state, starts playback of delta set on dest.

Restart → Restarts a suspended replication or resyncs a failed over or out-of-sync replication—source & destination will be resynced

**Note:** Run on Source side, but failed-over replication restart is done on Destination side.

Reverse → Reverses direction of Replication, run on Destination side [can also reset local timeout, hwm, bandwidth, & source policy]

Failover → fails over replication from Source to Destination only, but always issued from Destination side

**Note:** Restarting replication may require that Source & Destination timeouts, HWM, Interfaces, Bandwidth, SavVol name & size, and Source Policy be redefined. In otherwords, a Suspend of Failover does not maintain replication attributes. With a Reverse, local timeout and HWM and Source Policy can be specified.

#### **OUT OF SYNC REPLICATION RESTART [RESTARTABLE CHECKPOINTS]:**

CS script runs at 25 minutes past the hour to guarantee a restart point for out-of-sync replication restart

1. Manually create src\_fs\_repl\_restart\_1 and src\_fs\_repl\_restart\_2 checkpoints on each Production File System
2. Cron job takes over and refreshes the oldest checkpoint every hour

#### **Runs as Cron job /nas/sbin/refresh\_ckpt:**

```
# cat /etc/cron.d/nas_sysgrep refresh  
25 * * * * root /nas/sbin/refresh_ckpt >/dev/null 2>&1
```

**Note:** refresh\_ckpt script checks to see if replication source file systems have two restart checkpoints. If two already exist by name [\_repl\_restart\_1 and \_repl\_restart\_2], then refreshes the oldest checkpoint

Check sys\_log to verify that script refreshes are succeeding

#### **MOUNT ATTRIBUTES—CELLERRA MANAGER:**

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
Path; Data Mover; File System; Read Only; Multi-Protocol Access Policy; Virus Checking status; CIFS Oplocks in use; NT Credential; Direct Writes Enabled (writes to disk without caching); Prefetch Enabled; Multi-Protocol Locking Policy; CIFS Sync Writes Enabled (Yes means synchronous writes); CIFS Notify Enabled;  
Requires use of IE or Netscape in order to login

### **USER/GROUP QUOTAS:**

- Can manage User and Group Quotas from Celerra Manager
- Separate User/Group quotas can be applied per Tree quota [in addition to User/Group quotas at FS level]
- Tree quotas can have User/Group defaults; Soft quota grace periods; Deny Disk Space option; Flags to control errors/events
- Note:** If File System and Tree User quotas are enabled, the most restrictive Quota will be applied at FS level. For example, if a User has a User Tree quota of 100MB, but a file system User Quota of 50MB, then total allowable quota for the User will be 50MB  
\$nas\_quotas –on –fs fs01 –path /tree2 [Turns on User Quota for Tree Quotas]
- Still can only set User & Group quotas on a file system with TreeQuotas set, but empty!
- Also, if QuotaPolicy is set to Blocks, you cannot change to FileSize

### **CELERRA VSS FOR iSCSI (Volume Shadow Copy):**

Runs as a Windows 2003 service, providing interface between Volume Shadow Copy Service (VSS) and Celerra iSCSI snapshot iSCSI target & LUN snapshot. Provides ability of VSS aware applications to create shadow copies of iSCSI Luns using Celerra iSCSI snapshot technology

- Point-in-time RO copy of single or multiple volumes
- VSS are RO by default and limited to 512 copies per volume and 1000 snapshots per LUN
- Requires use of Windows 2003 as Volume Shadow Copy Service with Celerra VSS Provider Service
- Purpose is for Backups and Rapid Recoveries
- Celerra supports two VSS Usage Models: Backups & Data Transport
- Benefits are consistent copies of application data without impact, and can backup open files
- One strategy would be to create the VSC copy, backup to tape, then delete VSC
- Another strategy would be to transport VSC to other servers for testing, etc.
- Celerra iSCSI target provides data protection, snapshot file versioning, for iSCSI storage devices

### **VSS COMPONENTS:**

VSS Volume Shadow Copy Service → Windows 2003 [EMCVSSPack.exe]—coordinates actions of Requestors, Providers, Writers  
VSS Requestors → Backup Applications, such as Legato Networker 7.1 or Veritas BackupExec 9.0 that invoke a Shadow Copy

VSS Writers → Applications such as Exchange or SQL that store persistent information, and participate in Shadow Copy sync

VSS Providers → Hardware & Software Providers initiate Shadow Copies & own the Shadow Copy data

Celerra VSS Hardware Provider → NBS Server & iSCSI Targets, Storage Backend does work of creating Shadow Copy using LUNs

**Note:** CBMAPI.DLL (Celerra Block Management API) used to create, promote, delete iSCSI LUNs exported by DART, between iSCSI initiator and NBS server and Win2k3 Server

Software VSS Provider → Owns Shadow Copy data & initiates shadow copies of volumes--Intercepts & processes I/O requests in software layer between filesystem & volume manager

Celerra Provider & MS iSCSI Software Initiator & SnapAPI DLL

### **HOW SHADOW COPIES WORK FOR CELERRA:**

- Requester makes request and Celerra Provider determines if iSCSI LUNs are supported
- VSS Service tells Provider to prepare snapshot sets for LUNs
- VSS Service asks Provider to commit snapshot sets & CBMAPI works with backend to create snapshots
- VSS Provider provides snapshot device info and VSS Service saves information to XML file
- VSS Provider software for EMC is “EMC\_ISCSI\_UTILS\_1\_6\_2”

### **VSS LOGS & TROUBLESHOOTING:**

Event Logs

DebugView utility

C:\...\VssPrvOut.log

Debug info with CelerraVssProvider /debug | nodebug

Restart Volume Shadow Copy Service and Celerra VSS Provider Service on Windows 2003 system

### **VSS BACKUP LIMITATIONS:**

Veritas NetBackup & Backup Exec iSCSI LUNs become RO after Shadow copy → Resolved with VSP 1.04

Legato NetWorker cannot select individual SQL db's or Exchange info Stores for backups [whole only]

### **VSS REQUESTORS:**

EMC Legato NetWorker 7.1

EMC Snapview Integration MS Exchange

### **VSS WRITERS:**

Exchange 2003 & SQL Server 2000

### **VSS PROVIDERS:**

EMC Clariion & Symmetrix

## **NAS 5.4 NDMP2D ‘Backup-to-Disk’ ENHANCEMENTS:**

Purpose of NDMP is to provide a protocol for controlling Backup & Recovery of data, as well as low-level control of tape devices and SCSI media changers.

#### **TLU (TAPE LIBRARY UNIT):**

→ TLU has four elements: Robot, Drive, Slot, & Import/Export Elements

#### **VTLU COMPONENTS:**

**VTLU:** A software-based Virtual Tape Library Unit that resides on a DM filesystem & emulates the physical TLU [aka, SVTL]

Purpose of VTLU is to allow for compatibility between NDSMP and Disk-based storage [as opposed to tape-based storage]

**Barcode Prefix:** User-defined string to identify virtual tapes

**Slots:** Virtual tapes in the VTLU (move virtual tapes from virtual slots to virtual drives)

**Vault:** Holds virtual tapes, not visible from Backup software

**Import/Export Slots:** Slots that hold virtual tapes, using CLI or Celerra Mgr, move to & from Vault, or to & from tape slots

#### **VTLU Limitations:**

Limited to 65536 elements and tapes per TLU

#server\_vtlu server\_x -tlu -new -robot -slots [Configuring new VTLU]

#server\_vtlu server\_x -storage -new fs01 -tlu tlu\_id [Add storage to existing VTLU]

#server\_vtlu server\_x -storage -import fs01 -tlu tlu\_id -destination [import configured file system to a VTLU]

#server\_vtlu server\_x -storage -export fs01 -tlu tlu\_id [Export storage from a VTLU]

#server\_vtlu server\_x -tape -insert | -eject -tlu tlu\_id [Insert or Eject a tape]

#server\_vtlu server\_x -drive -unmount -tlu tlu\_id -id drive\_id [Unmounting a drive]

#server\_vtlu server\_x -service -info | -tlu -list | -tlu -info tlu\_id

#server\_vtlu server\_x -storage -list tlu\_id

#server\_vtlu server\_x -tape -list tlu\_id | -tape -info [Querying VTLU tapes]

#server\_vtlu server\_x -drive -list tlu\_id [list VTLU drives] -drive -info drive\_id -tlu tlu\_id [information on a drive]

#### **Recommendations:**

→ Use dedicated file systems for VTLU storage

→ Create virtual tapes to use up all available space

→ Use RAID-3 LUNs for ATA drives

→ Do not use 2 LUNs from same RAID group when making virtual tape storage

→ Do not use LUNs of different RAID types when making virtual tape storage [cartridge file system]

→ Best restore performance is with 64k striping

## **SUN JAVA SYSTEM DIRECTORY SERVER:** (aka; Sun One Directory Server, iPlanet, Java System Directory Server)

#### **INTRO:**

Purpose of iPlanet is to allow for LDAP-based entity lookup based on Directory schema outlined in RFC 2307(DIT—Directory Information Tree), and implemented with the Sun ONE Directory Server. iPlanet LDAP Directory Services would be a good way to integrate the Naming Services for environments that have both Windows and Unix systems. LDAP provides common language for Clients & Servers to communicate. LDAP applications can authenticate, search, add, delete, and modify directory entries. LDAP searches are in accordance with RFC 2254.

**Note:** Traditional entity queries use NIS, NIS+, DNS, local passwd/group files for Naming Services

#### **SUN ONE DIRECTORY SERVER COMPONENTS:**

→ Sun One Directory Server 5.1 was originally called iPlanet

→ Sun One Directory Server 5.2 replaces iPlanet

→ Backend LDBM database; LDAP Server protocol; authentication & access control; administrative interfaces

→ dse.ldif file describes LDAP configuration, entries, & attributes in LDAP Data Interchange Format (LDIF)

#### **SUPPORTED SERVER PLATFORMS:**

→ UNIX systems, Windows NT/2000

#### **CELERRA CLIENT PLATFORM:**

→ DART must be configured to use an iPlanet domain name and at least (1) iPlanet Server Host for the iPlanet domain

→ Celerra makes LDAP queries to Sun ONE Directory Server via LDAPv3 per RFC 2251 using ‘gethostbyname’ query. SunOne Is meant to complement or replace the existing Naming Services: NIS, DNS, Local Files, AD

→ Celerra will connect to Sun ONE using ‘client profile’ to obtain LDAP info, schema, list of preferred Servers

→ Requires use of configuration for Domain, Server Location, Client Profile, and NSSWITCH file

→ Ability to use LDAP v3 Server Controls (latter resides on Sun ONE Server):

**Server Side Sorting Request:** Ability to ask LDAP server to sort the reply

**Virtual List View Request:** Ability to return only a set number of entries from sorted list

→ Celerra can control order and use of various Naming Services via the nsswitch.conf file (Current entity queries are passwd, group, hosts, netgroup, usermapper), and now add LDAP as a Naming Service

→ Celerra, as an LDAP client, does NOT listen to TCP/UDP ports, but opens connections using TCP to the Sun ONE Server

→ Specific LDAP queries will be for Username [getpwnam()]; UID [getpwuid()]; Groupname [getgrnam()]; GID [getgrgid()]; Hostname or IP [gethostbyname() or gethostbyaddr()]; Netgroup name [using NIS YP Match calls]

→Configure DM as iPlanet Client using \$ server\_ldap server\_x –set –domain domain.com –servers xxxx.xxxx.xxxx.xxxx, which also starts the LDAP Service on the DM

→Manage DM as iPlanet Client using \$ server\_ldap

## **USE OF NSSWITCH.CONF FILE FOR DATA MOVERS AND USER/GROUP MAPPINGS:**

**\$ cat nsswitch.conf**

passwd: files ldap nis

group: files ldap nis

### **iPLANET STRUCTURE:**

→Objects stored as directory entry

→Each entry has multiple attributes & are stored in hierachical tree form

## **CELERRA iPLANET (SUN ONE DIRECTORY SERVICE) GENERAL OVERVIEW:**

→Version 5.2 LDAP Directory Service

→Sun ONE supports use of SASL, GSS-API, TLS, simple password, and anonymous authentication methods (Celerra supports only the Simple ‘Proxy’ Authentication & Anonymous Authentication methods)

→Supports CIFS & NFS/FTP

→Celerra uses iPlanet as NIS replacement [passwd, group, hosts, netgroup databases] using Simple Password & Anonymous authentication

→If NIS is in use, LDAP domain will be used as Sun ONE Directory Domain

→DART limited to single LDAP domain & single NIS domain per physical Server [NIS objects passwd, group, hosts, netgroup]

→Recommended use of DNS for host records

→NSSwitch.conf file must be configured and copied to /.etc directory on DM or iPlanet LDAP queries will not occur

→Sun ONE Directory Server DOES NOT register with DNS

## **USE OF CLIENT CONFIGURATION PROFILE:**

Client profiles reside on the iPlanet server and can be specified using the –profile option to include additional configuration attributes

**\$ server\_ldap server\_2 –set –domain pvt.emc.com –servers 172.16.21.10 -profile celerra\_profile**

Client Profiles contain schema mapping, authentication info, list of servers, preferred server

### **CLIENT PROFILE ATTRIBUTES:**

→Client attributes are preferred and alternate servers

→Search path

→Profile TTL

→Object class and attribute mapping

→Authentication method

→Order of Servers is Preferred, Alternate, and Configuration

## **CELERRA PROFILE ISSUE WITH BASE DNs:**

→Celerra establishes a default base DN based on the default domain to which it is configured

→Celerra does not know how to set a specific profile for a base DN or to search for available base DNs [AR76154 77525]

## **Appendix A Configuring Celerra Naming Services:**

preferredServerList—List of Server Addresses & Port numbers in the order the DM should contact

defaultServerList—if preferredServerList fails or is not configured, establish connection with same Server as used in ldap config.

defaultSearchBase—defines the base for directory searches

defaultSearchScope—default is one for one-level search, and sub for whole sub-tree searches

authenticationMethod—LDAP bind method to use when contacting DSA (Directory Server Agent)—Celerra supports Simple only

credentialLevel—what type of credentials DUA (Directory User Agent—DM) should use—Celerra supports Anonymous & Proxy

profileTTL—defines time interval before DUA should reload and reconfigure itself—set to zero does not reload

-----many other attributes that Celerra does not yet support-----

## **LDAP iPLANET AUTHENTICATION SUPPORT:**

→Data Mover uses Anonymous Logon by default if a Bind Distinguished Name was not specified during server\_ldap configuration

→Data Mover uses Simple Proxy Authentication if –binddn option is used during server\_ldap configuration. Use of –binddn allows user to specify the DN of the identity used to bind the service, which otherwise defaults to the Domain Manager. Use of –p will prompt User for the directory bind password during initialization.

**\$ server\_ldap server\_2 –set –p –domain pvt.emc.com –servers 172.16.21.10 -binddn**

**“uid=admin,ou=Administrators,ou=TopologyManagement,o=NetscapeRoot”**

## **SPECIFYING NIS DOMAIN:**

If the NIS domain name is different from the iPLANET domain name, use the following syntax to specify the NIS domain name

**\$ server\_ldap server\_2 –set –domain pvt.emc.com –servers 172.16.21.10 -nisdomain nsg**

## **CONFIGURING iPLANET FOR CELERRA:**

1. Recommended that DNS be configured for Host Name resolution on Control Station

2. Configure nsswitch.conf file for Celerra with ldap switch and put on /.etc directory

4. Export file system for NFS [specify netgroups if appropriate]

5. Configure Data Mover for LDAP:

**# server\_ldap server\_2 -set -p -domain hosts.pvt.dns -servers 192.1.4.209,192.1.4.201 -profile celerra1**

**-binddn "uid=admin,ou=Administrators,ou=TopologyManagement,o=NetscapeRoot"**

**#server\_ldap server\_2 -set -domain firstlogic.net -servers 172.16.5.148,172.16.5.149 -binddn**

**"cn=proxyagent,ou=profile,dc=firstlogic,dc=net"**

Password “nasadmin”

6. Verify LDAP setup:

**# server\_ldap server\_2 -info -v**

server\_2 :

LDAP domain: hosts.pvt.dns

State: Configured - Connected

NIS domain: hosts.pvt.dns

Proxy (Bind) DN: uid=admin,ou=Administrators,ou=TopologyManagement,o=NetscapeRoot

Profile Name: celerra1 -----output abridged-----

→Need minimum of domain name and configuration server address

**#server\_ldap server\_2 -set -domain nfs.lab -servers 10.241.169.150 -profile Celerra\_profile**

**#server\_ldap server\_2 -info -v | -clear (stops & deletes config) | -service -start | -stop | -status**

**#server\_ldap server\_2 -lookup -user user1 | -uid | -group | -gid | -hostbyname host1 | -netgroup netgroup1**

**# server\_export server\_2 -P nfs -n ldap\_fs1 -o root=clients /ldap\_fs1** [Example of exporting to Netgroup]

→Use /etc/nsswitch.conf file to configure precedence of naming services

passwd: files ldap

group: files ldap

hosts: dns ldap files

netgroup: ldap

## **MANAGING DM AS iPLANET CLIENT:**

**\$ server\_ldap server\_2 -service -status | -stop | -start | -clear** (permanently stops LDAP iPlanet service)

**\$ server\_ldap server\_2 --set** (use to make configuration changes on existing service)

## **MIGRATING FROM NIS DOMAIN TO SUN ONE:**

1. Migrate passwd, group, hosts, netgroup tables from NIS to Sun ONE Directory Server

2. Delete NIS Client service on Data Mover

3. Configure Sun ONE Directory service for Celerra

4. Test

## **TROUBLESHOOTING SUN ONE DIRECTORY SERVICES:**

**\$ server\_ldap server\_2**

Error 2100: usage: server\_ldap { <movername> | ALL }

**-set [-p] -domain <FQDN>**

**[ -servers <ip\_addr>[:<port>][,<ip\_addr>[:<port>]...]]**

**[ -profile <profile\_name> ]**

**[ -nisdomain <NIS\_domain> ]**

**[ -binddn <bind\_DN> ]**

**| -clear**

**| -info [-verbose]**

**| -service { -start | -stop | -status }**

**| -lookup { -user <username>**

**| -group <groupname>**

**| -uid <uid>**

**| -gid <gid>**

**| -hostbyname <hostname>**

**| -netgroup <groupname> }**

where 'profile' is Sun ONE Directory Server Client Profile

'binddn' is DN of the identity used to bind to the service

**\$ server\_ldap server\_2 -lookup -user ndmp**

server\_2 :

user: ndmp, uid: 999, gid: 101

**\$ server\_ldap server\_2 -lookup -group facilities**

server\_2 :

Group name: facilities, gid: 2010

**\$ server\_ldap server\_2 -lookup -uid 3019**

server\_2 :

user: stevez, uid: 3019, gid: 1010

**\$ server\_ldap server\_2 -lookup -gid 1010**

server\_2 :

Group name: domain users, gid: 1010

**\$ cat /nasmcd/quota/slot\_2/.etc/passwd**

ndmp:34M8hLCSxDGLY:999:101:ndmp backup user:/userdata/users/ndmp:

**\$ server\_ldap server\_2 -service -status**

server\_2 :

LDAP service active

**Note:** Not 100% sure, but secmap cache does not appear to store any entries as originating from “ldap”, just “nis”, even though customer was not using NIS client or NIS in the options for the nsswitch.conf file.

**\$ server\_ldap server\_2 -info**

server\_2 :

LDAP domain: firstlogic.net

State: Configured - Connected

NIS domain: firstlogic.net

Proxy (Bind) DN: cn=proxyagent,ou=profile,dc=firstlogic,dc=net

Profile Name: default

Profile TTL: 43200 seconds

Next Profile update in 27974 seconds

Connected to LDAP server address: 172.16.5.148 - port 389

**\$ server\_ldap server\_2 -info -v**

server\_2 :

LDAP domain: firstlogic.net

State: Configured - Connected

NIS domain: firstlogic.net

Proxy (Bind) DN: cn=proxyagent,ou=profile,dc=firstlogic,dc=net

Profile Name: default

Profile TTL: 43200 seconds

Next Profile update in 27918 seconds

Profile modification timestamp: 20051005162649Z

Connected to LDAP server address: 172.16.5.148 - port 389

LDAP configuration servers:

Server address: 172.16.5.148 - port: 389

Server address: 172.16.5.149 - port: 389

LDAP default servers:

Server address: 172.16.5.148 - port: 389

Domain naming contexts:

cn=changelog

dc=firstlogic,dc=net

o=NetscapeRoot

Domain supported authentication mechanisms:

EXTERNAL

GSSAPI

DIGEST-MD5

Directory Base DN: dc=firstlogic,dc=net

Domain default search Scope: single-level

'passwd' DN:

ou=people,dc=firstlogic,dc=net - search scope single-level

'group' DN:

ou=group,dc=firstlogic,dc=net - search scope single-level

'hosts' DN:

ou=hosts,dc=firstlogic,dc=net - search scope single-level

'netgroup' DN:

ou=netgroup,dc=firstlogic,dc=net - search scope single-level

LdapDomainSunOne::searchDomainRoot: Root domain subentry DN: ou=Groups, dc=firstlogic,dc=net

LdapDomainSunOne::searchDomainRoot: Root domain subentry DN: ou=People, dc=firstlogic,dc=net

LdapDomainSunOne::searchDomainRoot: Root domain subentry DN: ou=Special Users,dc=firstlogic,dc=net

LdapDomainSunOne::searchDomainRoot: Root domain subentry DN: ou=group,dc=firstlogic,dc=net-----output abridged-----

## **HOW MAPPINGS APPEAR IN SECMAP DATABASE FOR iPLANET:**

sid S-1-5-15-bfc56af5-d6cf8701-5f67b1a3-4c0 mapping

**Mapped from ldap** on Tue May 16 09:16:37 2006

user=b275(fde8)

Name=2K3\W2kU45685

## **RPM INSTALL & UPGRADE FOR STORAGE API:**

New RPM install and upgrade program to address need for new storage API's between NAS release cycles

EMC Solutions Enabler = SymAPI, aka Widesky

NaviCli/Agent = Clariion

Primary and Standby Control Station packages

Compatible with NAS 5.4 and higher releases only

Disable NAS Services prior to running

## **REPLACING NAS STORAGE API ON 5.5 SYSTEM:**

1. Copy the desired nasStorage API rpm version to the /tmp directory from the NAS CD-ROM:

**# cp /mnt/cdrom/EMC/nas/emcnas-5.5.31-1.i386.rpm /tmp**

**Note:** Or find the correct nasStorageAPI rpm file from the /nas/sys/apis directory

2. Extract the NAS rpm package:

**# cd /tmp & extract the NAS rpm package**

**# rpm2cpio emcnas-5.5.31-1.i386.rpm | cpio --extract --make-directories \*nasStorageAPI\*.i686.rpm &**

3. Unzip & untar the nasStorageAPI:

**# cd /tmp/nas/sys/apis**

**# tar -zxfv nasStorageAPI-6.0-3.tar.gz**

**# cd /tmp/nas/sys/rpms;ls**

nasStorageAPI-6.0-3.i686.rpm

3. Erase the old API:

**# rpm --erase nasStorageAPI**

4. Install the new nasStorageAPI:

**# rpm --install nasStorageAPI-6.0-3.i686.rpm**

5. Run PAHC to ensure that the storage api check passes

**# /nas/bin/nas\_checkup**

Checking if NBS devices are accessible..... Pass

**Note:** The nasStorageAPI provides Navicli/Naviagent and Solutions Enabler functionality when the Control Station interfaces with Symmetrix or Clariion backend storage systems. Alternatively, use the /var/sadm/pkg/nasStorageAPI\_mgr to erase, install, or upgrade the StorageAPI rather than using rpm.

## **INSTALLING, UPGRADING, ERASING STORAGE API NAS 5.6 SYSTEMS:**

**# /var/sadm/pkg/nasStorageAPI\_mgr/nasStorageAPI\_mgr**

**Note:** BASH script for Upgrading, Installing, or Erasing the nasStorageAPI on the system

## **UPGRADING STORAGE API WITH NAS RUNNING NORMALLY—Single CS:**

1. Copy the appropriate nasStorageAPI-7.1-8.tar.gz update to /tmp and make /nsa subdirectory

**# cd /tmp;mkdir nsa;cd nsa**

2. Untar the StorageAPI file in the nsa sub-directory

**# tar zxvf ..../nasStorageAPI-7.1-8.tar.gz**

**# ls -la**

-rwxr-xr-x 1 root root 8859 Jan 27 14:12 nasStorageAPI\_mgr

drwxr-xr-x 2 root root 4096 Jan 27 14:12 rpms

drwxr-xr-x 2 root root 4096 Jan 27 14:12 scriptlets

drwxr-xr-x 2 root root 4096 Jan 27 14:12 symcli

3. Run the storage api manager script

**# ./nasStorageAPI\_mgr -u**

Checking Backend Storage System(s) Version(s)

Backend Storage System(s) Version(s) OK

Stopping NAS Services...

rmdir: `/nas/etc/Navisphere/backup': Directory not empty

Did not delete /nas/etc/Navisphere/backup due to the presence of generated files.

If these files are no longer necessary, please delete manually.

Starting EMC Solutions Enabler install. This may take several minutes.

Removed EMC Solutions Enabler semaphores 0

EMC Solutions Enabler install complete

Re-starting NAS Services.....

### **LISTING OUT CONTENTS OF RPM PACKAGE:**

# rpm -ql nasStorageAPI

/etc/Navisphere

/nas/etc/Navisphere

/nas/etc/Navisphere/Navimon.cfg

/nas/etc/Navisphere/SupportedFlareRevisions

/nas/etc/Navisphere/agent.config

/nas/opt/Navisphere/bin

/nas/opt/Navisphere/bin/classic\_navicli

/nas/opt/Navisphere/bin/nas\_navi

/nas/opt/Navisphere/bin/naviagent

/nas/opt/Navisphere/bin/navicli

/nas/opt/Navisphere/bin/naviseccli

/nas/sbin/classic\_navicli

/nas/sbin/get\_backend\_status

/nas/sbin/nas\_navi

/nas/sbin/naviagent

/nas/sbin/navicli

/nas/sbin/naviseccli

/nas/sys/Navimon.cfg

/nas/sys/agent.config

/nas/sys/clariion\_be\_callhome.tpl

/nas/sys/nas\_be\_callhome.tpl

/nas/sys/requirements

/nas/sys/symm

/nas/sys/symm/options

/nas/sys/symm/symapi\_licenses.dat

/nas/tools/CelerraServiceFlag-01.01.5.001-xpfree.ena

/nasmcd/sbin/classic\_navicli

/nasmcd/sbin/nas\_navi

/nasmcd/sbin/navicli

/nasmcd/sbin/naviseccli

/opt/Navisphere

### **LOCATION OF NAS STORAGE API ON CS:**

**#** /nas/sys/apis

# ls

nasStorageAPI-6.0-3.tar.gz

# cat /nas/sys/apis/nasStorageAPI.env | head -8

#Primary NAS Storage API RPM Environment Variables:

NAS\_STORAGE\_API\_NAME=nasStorageAPI

NAS\_STORAGE\_API\_VERSION=7.0

NAS\_STORAGE\_API\_RELEASE=12

NAS\_STORAGE\_API\_TARGET=i686

NAS\_STORAGE\_API\_RPM=nasStorageAPI-7.0-12.i686.rpm

NAS\_STORAGE\_API\_TAR=nasStorageAPI-7.0-12.tar.gz

NAS\_STORAGE\_API\_REQUIREMENTS="#nasStorageAPI-7.0-12

# rpm -qa nasStorageAPI

nasStorageAPI-6.0-3

### **UPGRADING NAS STORAGE API AFTER UPGRADE FAILURE:**

**Note:** If NAS 5.4 upgrade fails with following message, download the latest nasStorageAPI and install Backend Storage Requirements Check Failed:

INSTRUCTIONS: Upgrade NAS Storage API

1. Download nasStorageAPI from TS2 website (An API required for primary CS and also one for Standby CS)
2. Upgrade RPM: # rpm --upgrade nasStorageAPI-0.0-6.i386.rpm
3. Query to verify: **# rpm --query --info nasStorageAPI**

#### **StorageAPI Install & Upgrade:**

#rpm --install nasStorageAPI-0.0-3.i386.rpm [Primary CS]

#rpm --install nasStorageAPI-sby-0.0-3.i386.rpm [Standby CS]

#rpm --upgrade nasStorageAPI-0.4.i386.rpm [Primary CS]

#rpm --upgrade nasStorageAPI-sby-0.4.i386.rpm [Standby CS]

**StorageAPI Erase:** Run only in emergency if needed to uninstall

**# rpm --erase nasStorageAPI**

#### **StorageAPI Query:**

**# rpm --query -info nasStorageAPI**

```
Name      : nasStorageAPI          Relocations: (not relocateable)
Version   : 6.0                  Vendor: EMC
Release   : 3                   Build Date: Mon 06 Aug 2007 06:03:03 PM EDT
Install date: Tue 21 Aug 2007 12:01:08 PM EDT  Build Host: NasBuild5
Group     : System Environment/Base  Source RPM: nasStorageAPI-6.0-3.src.
rpm
Size      : 83359399            License: EMC
Packager  : EMC
URL       : http://www.celerra.isus.emc.com/top_level/top_docs.htm
Summary   : API used to interface with Celerra backend storage systems.
Description: API used to interface with Celerra backend storage systems. The API provides NaviCli/Agent and EMC Solutions
Enabler functionality.
```

**# rpm --query -provides nasStorageAPI** [From NAS 5.5.31.1 system]

```
NaviCli_Agent = 6.26.0.2.14-1
EMC_Solutions_Enabler = V640
FC4700_Software_Support >= 8.51
FC4700_Software_Support < 8.53
CXx00_Software_Support >= 2.19
CXx00_Software_Support <= 2.26
CX3-x0_Software_Support >= 2.24
CX3-x0_Software_Support <= 2.26
Symm5267_Software_Support > 17
Symm5567_Software_Support > 31
Symm5568_Software_Support > 23
Symm5669_Software_Support > 36
Symm5670_Software_Support > 23
Symm5671_Software_Support > 31
Symm5771_Software_Support > 68
Symm5772_Software_Support > 55
nasStorageAPI = 6.0-3
```

#### **SOLUTIONS ENABLER:**

→A host-based API used to manage & monitor Symm, CLARiiON, and Celerra storage devices

→SE 7.1 is expected to be IPv6 capable, Q4 2009

→ECC and SYMAPI are host-based applications built on the SE

→SYMAPI used by Celerra to monitor Symmetrix and Clariion storage

#### **CAVA ENHANCEMENTS:**

→Automatic recognition of AV software updates

→Popup notifications of infected files and action taken by AV engine

#### **CELERRA NAS 5.5 BURGUNDY:**

5.5.19-4 GA March 17, 2006; Current release 5.5.42

#### **NAS 5.5 BURGUNDY RELEASE:**

#### **UNSUPPORTED ITEMS:**

→SCSI connections to backends no longer supported, though is still provided to Tapes

- 507 DM or below
- FDDI/ATM connections not supported
- FC4700 no longer supported
- Symm 5 FC

### **NS350 ENTRY LEVEL LOW-END NAS PLATFORM:**

- Ease of Use/Installation initiative designed with Placemat and Configuration Wizard—No fresh install required!
- NAS 5.5.19 and Flare 19 minimum using CX300 array
- Introduced with Single or dual-Data Mover configurations, but Single Control Station only
- Sold as Integrated model only, but upgradeable to Gateway NS500G
- Old model NS350-AUX-FD has been discontinued—no more “fresh install” platforms
- Factory-install models are all that are available now

### **NS350FC:**

You may encounter the term “NS350 FC”, which is a misnomer. What is meant is that an NS350 Integrated can be manually upgraded to become a Fibre Channel enabled NS500G. There is no additional software charge for this option. This is the “origin” of the so-called “All-in-one” models that are really being introduced in a public way starting with the NS20. The big marketing spiel is that one platform will support FC Hosts, iSCSI Hosts, and NAS Hosts. The “Celerra NS350...Series Hardware Upgrade Procedures Guide” contains the complete process for upgrading an NS350 to an NS500G—the use of the EMCNAS Software CD and upgrade script is required: [/mnt/cdrom/EMC/nas/setup\\_hw\\_upgr\\_aux2fc](/mnt/cdrom/EMC/nas/setup_hw_upgr_aux2fc).

### **Data Mover Specs (same hardware as NS500 DM):**

Processor = Intel Pentium 4  
Processor speed (MHz) = 1600  
Total main memory (MB) = 4031  
Mother board = Exterminator XP  
Bus speed (MHz) = 533  
Bios Version = 3.62  
Post Version = Rev. 02.25

### **RSA/ENVISION SOLUTION:**

→The special RSA/enVision solution used to ship on the NS350 platform, and in 2008 changed over to the NS20 Integrated platform. It will soon be shipping on the NS-120 Integrated. The current issue is that the NS20 system is being implemented without Proxy ARP (No CSA applied) or public IP address assignments for the SPs, which is contrary to the productline itself. Product will move over to the NX4 platform at some point(?).

→RSA enVision is a security information and event management solution (SIEM)

### **RSA enVision LS Series NS20 Solutions:**

RSA NAS 3500 →NS20 with 15 disks 3.5TB capacity; Dual Blades; CIFS-based  
RSA NAS 7000 →NS20 with 30 disks 7TB capacity; Dual Blades; CIFS-based

### **LINUX KERNEL UPDATE:**

→Kernel 2.4.9-34 will be replaced by 2.4.20-28.7 (Still RedHat 7.2) as a drop-in update, i.e., no recompiling of kernel is required (main purpose of update are to address security updates, bug fixes, and provide more hardware support, such as USB ports, and also to be at same kernel version as Blackbird project). Incidentally, will add hyper-threading capability, though code is still uniprocessor.

#### **\$ uname -r**

2.4.20-28.5504.EMC

→KLMs (Kernel Loadable Modules) will be supported at runtime, and supports standard insmod, lsmod, modprobe, and rmmod commands (other KLMs are ctap (serial), nd (network drivers), and SNARE audit (System iNtrusion Analysis & Reporting Environ.)

→Control Station considered an appliance with an embedded Linux kernel

→USB port support on Control Station—do NOT reboot CS while USB port is in use

SimpleTech STI-USB2FD/1GB USB 2.0 Flash Drive

### **USING LINUX USB PORTS WITH CELERRA CONTROL STATIONS:**

1. Plug flash drive into Control Station USB port
2. Run /sbin/fdisk -l to identify the dynamic sda number that was assigned.

**Note:** For NS Models, the sda device will most likely be the only such local SCSI device seen, as in the following output from an NS700 system:

```
# /sbin/fdisk -l
```

Disk /dev/sda: 32 heads, 63 sectors, 993 cylinders

Units = cylinders of 2016 \* 512 bytes

| Device    | Boot | Start | End | Blocks  | Id | System |
|-----------|------|-------|-----|---------|----|--------|
| /dev/sda1 |      | 1     | 992 | 999813+ | 6  | FAT16  |

3. Create mountpoint for usb device and mount:

```
# mkdir /usb
```

```
# mount /dev/sda1 /usb
# df -T
Filesystem Type 1k-blocks Used Available Use% Mounted on
/dev/sda1 vfat 999552 14160 985392 2% /usb
# mount
usbdevfs on /proc/bus/usb type usbdevfs (rw)
# cd /usb
# ls -la
drwxr-xr-x 2 root root 16384 Dec 1 2004 CruzerLock 2
drwxr-xr-x 2 root root 16384 Mar 22 13:12 my files
4. Copy files, etc.
5. Umount usb device from Control Station and Verify!
# umount /usb
# df -T
Filesystem Type 1k-blocks Used Available Use% Mounted on
/dev/hda3 ext3 2063536 889600 1069112 46% /
/dev/hda1 ext3 31079 2751 26724 10% /boot
none tmpfs 257448 0 257448 0% /dev/shm
/dev/nde1 ext3 1818352 787776 938204 46% /nbsnas
/dev/nda1 msdos 136368 36212 100156 27% /nas/dos
/dev/ndf1 ext3 1818352 121500 1604480 8% /nas/var
/dev/hda5 ext3 2063504 673660 1285024 35% /nas
```

**WARNING:** Never reboot the Control Station while the USB drive is engaged and mounted as it could corrupt the USB drive and cause the Control Station not to boot properly. Also, before removing the USB Drive, make sure you manually unmount from the Control Station.

**Note:** The only “supported” thumbdrive at this time is the SimpleTech STI-USB 2FD/1GB USB 2.0 Flash Drive, with EMC 053-001-348 USB part number. Other drives may work, such as the ‘SanDisk Cruzer mini’ 1GB USB 2.0 Flash Drive with an NS700 Control Station, but are not supported.

#### **CONTROL STATION USB PORTS:**

CFS, CNS, & NS600 Control Stations have version USB 1.1 ports that have transfer rates of 1.5MB/sec  
NS500/700 & NSX Control Stations (Falcon and Chivas CS, respectively) use USB 2.0 ports with transfer rates of 40MB/sec. Falcon CS has (3) USB ports on back of system to far left. Chivas CS has (2) USB port in back of system to the middle of the unit, and (2) ports in the front of the system to the far right. NS600 Control Station has (2) ports in the center rear of the unit, but buried behind the Allied Telesyn Ethernet switch.

#### **NTFS DRIVER SUPPORT FOR 5.5:**

→NTFS RO file system driver for use in special cases for iSCSI recovery

1. Loading NTFS driver

**#/sbin/modprobe ntfs**

**#mount -t ntfs -o ro /dev/xxx /mnt/zzz**

#### **NAS STORAGE API:**

# pwd

**/nas/sys/apis**

# ls -la

```
-rwxrwxr-x 1 nasadmin nasadmin 46887850 Nov 9 15:33 nasStorageAPI-0.0-18.i686.rpm
-rwxrwxr-x 1 nasadmin nasadmin 728 Nov 9 15:33 nasStorageAPI.env
-rwxrwxr-x 1 nasadmin nasadmin 21 Nov 15 16:13 nasStorageAPI.ok
```

**Note:** API interfaces with backend storage and enables NaviCli Agent and EMC Solutions Enabler functionality

#### **NAS PRE UPGRADE HEALTH CHECK TOOL (PUHC):**

**# /nas/tools/check\_nas\_upgrade -pre | -up | -debug | -version**

**Note:** Errors will result in the upgrade NOT being allowed to proceed until corrected. Pre-Upgrade check tool can be used as stand-alone healthcheck tool.

**# /nas/bin/nas\_storage -check -all**

```
/nas/log/nas_log.al
2006-01-26 12:46:44.260 db:0:7000:S: /nas/bin/nas_storage -check -all
2006-01-26 12:47:26.404 db:0:7000:E: /nas/bin/nas_storage -check -all
```

**# /nas/tools/check\_nas\_upgrade -up** →Triggered by NAS upgrade script

**# /nas/tools/check\_nas\_upgrade -pre** →More checks than -up, run as part of CCA process

Check Version: 5.5.19.1

Check Command: /nas/tools/check\_nas\_upgrade

Check Log : /nas/log/check\_nas\_upgrade.Mar-17-09:41:54.log

-----Checks-----

Control Station: Checking if standby is down..... Pass  
Control Station: Checking if NIS is stopped..... Pass  
Control Station: Checking that no cron jobs are scheduled..... Fail  
Control Station: Checking there is no exclusion file for MPD translation... Pass  
Data Movers : Checking boot files..... Pass  
Data Movers : Checking network connectivity..... Pass  
Data Movers : Checking status..... Pass  
Data Movers : Checking MAC address..... Pass  
Data Movers : Checking if using standard DART image..... Pass  
Data Movers : Checking if primary is active..... Pass  
Data Movers : Checking if root filesystem has enough free space..... Pass  
Data Movers : Checking if hardware is supported..... Pass  
Storage System : Checking if access logix is enabled..... N/A  
Storage System : Checking that no hot spares are rebuilding..... Pass  
Storage System : Checking that no hot spares are in use..... Pass  
Storage System : Checking if microcode is supported..... Pass  
Control Station: Checking if NBS configuration exists..... Pass  
Control Station: Checking if NBS clients are started..... Pass  
Control Station: Checking if NBS devices are accessible..... Pass  
Control Station: Checking if NBS service is started..... Pass  
Control Station: Checking if enough free space exists ns..... Pass  
Storage System : Checking if FLARE is supported..... Pass  
Storage System : Checking if FLARE is committed..... Pass  
Control Station: Checking integrity of NASDB..... Pass  
Control Station: Checking if Symapi data is present..... Pass  
Control Station: Checking if Symapi is synced with Storage System..... Pass  
Storage System : Checking that no disks or storage processors are faulted.. Pass  
Storage System : Checking disk high availability access..... Pass  
Storage System : Checking no disks or storage processors are failed over... Pass  
Storage System : Checking disk emulation type..... Pass  
Storage System : Checking disks and storage processors read cache enabled.. Pass  
Storage System : Checking disks and storage processors write cache enabled. Pass  
Control Station: Checking if NAS Storage API is installed correctly..... Pass

---

/nas/log: -rw-r--r-- 1 root root 3636 Mar 17 09:43 check\_nas\_upgrade.Mar-17-09:41:54.log

Overall status code is 0

## **TO RUN PUHC HEALTHCHECK:**

**[/nas/tools/nas\\_checkup](#)**

### **CONTROL STATION CHANGES:**

- nasStorageAPI upgrade capability on Control Station (/nas/sys/apis)
- USB 2.0 thumb drive support, NTFS file system driver support
- Intended as a recovery tool for iSCSI
- Plug in thumb drive, mount the drive

**# mkdir -p /mnt/usb**

**# mount /dev/sdx1 /mnt/usb**

**# umount /mnt/usb**

**Note:** Always unmount USB drive before rebooting or unplugging the USB drive

→ Load NTFS driver to kernel separately for RO purposes only

#/sbin/modprobe ntfs

#mount -t ntfs -o ro /dev/xxx /mnt/zzz

### **GSS-API ERROR DECODING:**

- Creation of a GSS-API interface to decode internal error codes related to GSS-API, and its components, such as Kerberos
- Provides for a NULL-terminated decoded error string when returning error codes
- Subsystems CIFS, DNS, LDAP, RPCGSS will be modified to use gss\_error\_decode for Kerberos error decoding
- convertGssError and logGssError are modified

## **LDAP SIGNING SUPPORT:**

- LDAP Integrity Checking requires that messages are signed using Kerberos authentication
- LDAP Client & Server authentication is conducted via Kerberos and SASL (Simple Authentication and Security Layer) protocol, which describes how data is carried, and allows Security Layer to be negotiated (Integrity, Privacy, or None)
- Note:** The Integrity & Encryption Provider is Kerberos, which uses the GSS-API. Kerberos will check signatures and decrypt messages, as required. An algorithm is used for Signing and Encryption, from Kerberos, using krb5.conf file algorithms, such as DES, RC4, etc.
- LDAP Privacy Protection requires that messages are signed and encrypted using Kerberos
- LDAP Signing, or Integrity Checking, is requested by all Win2k/2k3 Servers, and DART 5.5 when using LDAP sessions, but not enforced on Clients
- Dart will use LDAP Integrity checking if DC uses & clients request—param value is 2
- Dart can only perform message encryption if param value of 4 is set
- Win2k/2k3 does not perform Message encryption by default
- LDAP used for Joins, Unjoins, Server Passwd Changes, GPO Updates, ADMapping LDAP Queries, IPsec policies, adding Services to a CIFS Server, iPlanet LDAP Services and queries

### **Win2k registry to enforce LDAP message signing:**

HKLM>System>CurrentControlSet>Services>NTDS>parameters>LdapServerIntegrity: [0 for no signing; 2 for signing]

→ Win2k3 contains Domain Policy & Domain Controller Policy

**None:** Data Signing not required to bind to Server, but if Client requests, Server will support signing

**Require signing:** LDAP signing must be used between Client & Server, unless Transport Layer Security/SSL is used (TLS/SSL)

**Not defined:** setting not enabled or disabled

## **DART LDAP PARAMETERS:**

CIFS Stop & Restart will put param change in place

param ldap SecurityLayer=0 → Signing & Encryption disabled—if DC or Servers require LDAP Security, LDAP bind will be rejected with LdapErr: DSID-OC090169

param ldap SecurityLayer=1 → Negotiate whatever Security layer the Server/DC is proposing

param ldap SecurityLayer=2 → Default Celerra value, Server enforces Signing with Clients

**param ldap SecurityLayer=4** → Enforces LDAP Encryption (Privacy)

**# server\_param server\_2 -facility ldap -info SecurityLayer -v**

server\_2 :

```
name      = SecurityLayer
facility_name = ldap
default_value = 2
current_value = 2
configured_value =
user_action = reboot DataMover
change_effective = reboot DataMover
range     = (0,4)
description = 0=No security layer, 1=Server proposes, 2=Integrity (default), 4=Privacy
```

## **AUTOMATED SAN INSTALL PART II:**

→ Automated zoning & Storage Group setup will be done for hardware upgrades, supporting all NS Celerras & NSX Hammerhead, Gateway direct-connect or fabric-connect (Does not support CNS or CFS cabinets)

→ If NS Celerras boot from a Symm, script can be used to add a Clariion backend

→ Direct-connect will support only 2Gbs/sec speeds for CX500/600/700, while fabric-connect will support CX300/500/600/700

→ Up to (4) fabric switches supported, using models Brocade, Cisco, or McData

→ Does healthcheck first of Array, as it will not succeed if there are faults (requires Write Cache enabled, Access Logix, etc):

**# /nas/sbin/check\_requirements**

→ Single data mover can be added:

**Note:** Script will not work if Zones or WWN Initiator Records or StorageGroup already exists!

## **AUTOMATED SAN INSTALL FUNCTIONALITY:**

### **ADDING DATA MOVER TO CABINET:**

**Note:** Script is invoked during setup\_slot to automatically add fibre DM fibre ports to existing Storage Group

**# /nas/sbin/setup\_slot -init -gateway\_auto\_config <slot#>**

Starting PXE Service....done

1<sup>st</sup> Reboot ....0000000333333314

Stopping PXE Service...done

2<sup>nd</sup> Reboot....000000033333344

Asks for Switch info and login prompt

Configures Zones and Storage Group records

3<sup>rd</sup> Reboot....discoveries disks

4<sup>th</sup> Reboot--Stops PXE Service

#### **ADDING ADDITIONAL CLARIION ARRAY & LUNS TO CELERRA:**

##### **# /nas/sbin/add\_clariion -init**

→New array must have AccessLogix enabled, write cache, two ports per SP for Celerra, and Root user

→Script prompts addition of SPA & SPB IP addresses for array, fibre switch information, creates new Storage Group and adds all Data Movers to the new SG.

##### **# /nas/sbin/setup\_clariion -init <serial\_no>**

**Note:** Run this after the add\_clariion to add new Storage Luns to Celerra database

#### **TROUBLESHOOTING LOGS:**

/nas/log/nas\_raid.log

/nas/log/zone.log

#### **SYMANTEC SAVSE ANTIVIRUS SUPPORT NAS 5.5—see Celerra Antivirus section:**

#### **FSDB CHECK TOOL:**

Tool to allow for inspection of on-disk structures for problem troubleshooting. For example, if a checkpoint fs is corrupt, the tool can show if the corruption exists on the PFS or the SavVol. Tool will obtain volume layer info from DM—which presumes that this info is correct. Tool will be a fs\_db executable on the CS using RPC to communicate to DART, and require linking to work (similar to mac\_db tool).

##### **Using the fs\_db Tool:**

→Inspect superblock:

##### **# fs\_db server\_2 readsb 120 16 (metavolume id=120, fsid=16)**

→Inspect cylinder group headers:

##### **# fs\_db server\_2 readcg 120 1**

→Inspect inode for given file:

##### **# fs\_db server\_2 readi 120 14 (metavolume id=120, inode=14)**

→Inspect metadata or data block:

##### **# fs\_db server\_2 readfsb 120 4479 /tmp/block.out (reads block 4479 & outputs to file)**

#### **EXAMPLE—VERIFYING STATE OF SUPERBLOCKS ON A FILESYSTEM:**

\$ server\_mount server\_2 -v <fs\_name>

\_fs\_db server\_2 readsb <volid> psb (PSB is located on Sector 16 of the file system)

\_fs\_db server\_2 readsb volid asb (ASB is located on Sector 32 of the file system)

\_fs\_db server\_2 readsb volid 8194

\_fs\_db server\_2 readsb volid 16386

#### **USING FS\_DB TOOL:**

1. While in the /nas/tools directory, create following link:

##### **# ln -s fs\_db \_fs\_db**

lrwxrwxrwx 1 root root 5 Feb 13 12:08 \_fs\_db -> fs\_db

2. Finding volume ID for file system “file4” (or use nas\_fs –list to see volume ID associated with the file system)

##### **# grep file4 /nas/volume/filesys**

26:file4::0:4::y:1:**93:1**::0::24::0:: (vol ID and Server\_ID, respectively)

3. Run fs\_db tool to obtain superblock information:

##### **# /nas/tools/\_fs\_db server\_2 readsb 93 psb →93 is the Volume ID of the file system “file4”**

About to read superblock from volume 93 and sector 16

ncyl=250

magic=0xdabadface →This magic value indicates file system is not a FLR Compliant file system [0xdcbaabc = FLR Compliant]

version=2

ncg=16

size=128000

bsize=8192

nbfree=108108

nifree=12589

4. Remove link when done with activity

#### **CAM MEMORY REDUCTION:**

→Reducing memory used by CAM subsystem

→Most memory used for caching inquiry data from target-lun addresses on backend (up to 256 connections)

→scsi\_dev object will hold INQUIRY data, and backend can be queried for updates

→I/O queue implemented at volume layer (starting with 5.4)

## **CELERRA MANAGER INDICATIONS MANAGER ENHANCEMENT:**

**Note:** Whole purpose is to increase WebUI and API performance via new caching mechanism

→Indication Mgr subsystem will provide clients (JServer, Celerra Mgr, API) with more objects to track persistent state changes

→JServer receives indications via core schema (Port 8887, use telnet to verify port)

→API & Celerra Mgr will receive indications via APL schema (Port 8886, use telnet to local CS to verify port)

### **Logs:**

/nas/log/apl\_log.al.indication (Indication Mgr service on CS)

/nas/log/webui/apl\_ind.log (apl\_ind\_mgr service on CS)

### **Troubleshooting:**

→Telnet to ports 8886 or 8887 to verify ports

→Set NAS\_DB\_DEBUG=1 and restart apl\_indication\_mgr and core\_indication\_mgr

## **CHECKPOINTS/SNAPSURE:**

→max number checkpoints increased to 96/file system, 64/fs for 510 DM—fs cyl group caching redesigned

→max number file systems per DM/Celerra 2048, except NSX, which has 2048 fs/DM and 4096/NSX

→max number replication sessions 64 for NSX and 32 for all other cabinets, from 16 previously

### **Cylinder Group Cache Statistics:**

**\$server\_config server\_2 -v "printstats cgcache"**

## **CELERRA CCMD--COMMON MESSAGE DEFINITION:**

**\$ nas\_message -info <msg\_id>**

**# nas\_message -info 13421904928**

id = 13421904928

severity = ERROR

component = CS\_CORE

facility = cabinetdr

baseid = 1056

brief description = Operation not permitted. Command must be run from remote administration account.

full description = This error message may be reported when the /nas/sbin/nas\_mview command to manage the remote environment is run from the local NAS administration account.

recommended action = Rerun the command from remote administration account.

→Message IDs can range from 1 – 68719476735 (1-65534)

## **CCMD PURPOSE:**

Purpose to provide common framework for storing and presenting messages. Post-Burgundy, messages will be stored as raw message data. Currently, all error messages are embedded in DART as strings, which are called by logmsg to create logEvents, and from Control Station sys\_log as strings. Goal is to convert sys\_log and many DART logmsg formats to CCMD format. Burgundy is creating a CCMD framework and CCDM message compiler, though enhancing and converting message definitions will be ongoing in further code releases. CCMD messages are 64-bit.

### **Four Major Components:**

**WebUI:** Browser-based interface using TCP and XML

**API:** Gets messages from CS via SAL and transmits to WebUI (JServer is part of this). Interface between WebUI presentation layer and core CS code.

**CS CORE:** Messages from DART; transmits some messages from DART to API via SAL.

**DART:** CS requests are sent via MAC/RPC or MAC/XML interfaces. DART also uses MMC to send messages to hosts.

**Note:** Feature limited to those messages translated into this framework

## **CONCEPTUAL CCMD FRAMEWORK:**

DART Cmd → HTTP → XML Dispatcher → MAC Handler → XML\_CMD\_Handler → Cmd (logmsg debug) → logmsg (CCMD) → Server Log

DART Facility → logmsg will output any debug messages to Server Log, just as it does now—current design is to not format debug messages into CCMD

DART Facility → logevents will be generated in CCMD format, which go into a DART memory buffer

The Control Station Facility will use MAC/XML to collect messages from memory buffer and puts into the ‘NAS Event Collector’ → then reports events to ‘NAS\_Eventlog’ → which then get recorded to sys\_log. Sys\_log will be viewable by a CCMD tool to read the new format (currently uses .server\_config MAC/RPC interface to retrieve DART messages)

DART Facility → if Server Log messages become CCMD cataloged, then DART will generate logmsg and record to Server Log

### **NAS EVENTLOG DAEMON:**

→Daemon receives messages from DART via “nas\_eventcollector”

→Daemon receives messages from CS via ReportEvent and PostEvent

## **WHEN ARE CCMD MESSAGES LOGGED AND WHEN ARE THEY NOT LOGGED?**

There are CCMD messages generated by User or System activity that would display to console only [Query-based non-state change actions], which would not be logged anywhere on the Celerra, and then there are Transaction-based actions [Transaction-based actions are “state” change actions that take locks to make system changes] that would log CCMD Error messages in sys\_log or server\_log when they occur. It’s not clear whether Warning messages would ever be logged.

**EXAMPLES OF MESSAGES NOT LOGGED in SYS LOG (even with nas logviewer -v switch):**

# **nas\_diskmark -m -a**

Discovering storage (may take several minutes)

done

**Warning 1771681587:** tid/lun= 1/4 type= disk sz= 0 val= -99 info= DGC RAID 5 03241400140014NI: invalid LUN size

**Note:** Even though a nas\_diskmark would be a transaction-based command, in this case, the diskmarking portion of the code still executed, but the invalid LUN itself would not be diskmarked, which is considered a Warning and not an event that would be logged.

# **nas\_diskmark -m -a**

Discovering storage (may take several minutes)

**Warning: 17716815787:** server\_2 c0t0l7 skipping reserved LUN id,

APM00073801838 stor\_dev=0x0007 (7 decimal), maps to a reserved host LUN id 0x7.

**VIEWING CCMD RECORDS IN SERVER LOG:** NAS 5.6

**\$server\_log server\_2 -v >slog2.ccmd**

Then used nas\_logviewer to view the contents of the file

**\$server\_log server\_2 -i**

**Note:** A hidden option that outputs server log entries in Unix epoch time

**VIEWING CCMD FORMATED SYS LOG ENTRIES:** NAS 5.6

**# /nas/bin/nas\_logviewer -v /nas/log/sys\_log >system.log |more** (CCMD events from sys\_log to file or screen)

facility = Callhome

baseid = 206

type = EVENT

argument name = arg0

argument value = /opt/connectemc/poll/RSC\_FCNHH050500031\_021307\_221346460.xml

argument type = 8

brief description = Processing CallHome event file @ /opt/connectemc/poll/RSC\_FCNHH050500031\_021307\_221346460.xml.

full description = For details, inspect log file at /opt/connectemc/logs/ConnectEMC .

recommended action =

**# /nas/bin/nas\_logviewer -t /nas/log/sys\_log >sys1.log** (called “terse” format, pulls entries into timestamp readable file)

Feb 14 13:08:20 2007:96108609742:Processing CallHome event file @

/opt/connectemc/poll/RSC\_FCNHH050500031\_021407\_092908311.xml.

**# /nas/bin/nas\_logviewer -f /nas/log/sys\_log >sys2.log**

Feb 14 13:14:30 2007:CS\_PLATFORM:Callhome:INFO:206::::Processing CallHome event file @

/opt/connectemc/poll/RSC\_FCNHH050500031\_021407\_001706374.xml.

**# /nas/bin/nas\_logviewer sys\_log**

**CCMD CATALOG LOCATION ON CS:**

**/nas/etc/catalog/en\_US**

**Use following to redirect Catalog location:**

**NAS\_CATALOG\_DEBUG=/tmp**

**MMC EXTENSIONS:**

→Win2k3 R2 adds Action Pane in MMC version 2.1

**Debug Logging for MMC Issues:**

Set registry: HKLM>Software>Emc>Celerra>MMC>Configuration>Debug: Set value 0x4

Configure DebugView on system and run in capture mode while reproducing problem

**CIFS EMC TOOLS:**

→New tool fstoolbox.exe to list, remove, move, take ownership of files and directories for specific User and manage Quotas

**NAS 5.5 COMPUTER ACCOUNT PASSWORD CHANGE & HISTORY FEATURE:**

**How the Computer Account Password Feature works:**

DART generates a computer account password during the initial Join process, and is renewed as scheduled (Windows default every 30 days), or manually using –Join resetserverpasswd command. The computer’s password is transferred to the KDC and stored in AD & on the DM, using either Kerberos Set Password protocol [KPASSWD] or MSRPC SAMR Set Password protocol. NAS 5.5 introduces ability to keep password history of last (2) passwords, which will be stored under version numbers that match Kerberos key version information. The DM produces encryption/decryption Kerberos keys based on its computer account password—these keys are regenerated during Join process, during password renewal, or when restarting CIFS service. Clients will be able to connect to DM

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
using Kerberos service tickets generated from a KDC using either the current or previous password version. When decrypting Client Kerberos Service Tickets, the DM will use its computer account password, newest, then oldest, to decrypt the ticket. Whenever a Kerberos key is generated, DM cache will be searched for principal name and key version—if that fails, then all passwords will be searched to generate new keys for Kerberos principal.

### **NAS 5.5 PASSWD CHANGE FEATURE IS DISABLED BY DEFAULT:**

**# server\_param server\_2 -facility cifs -info srwpwd.updtMinutes**

```
server_2 :  
name      = srwpwd.updtMinutes  
facility_name = cifs  
default_value = 0  
current_value = 0  
configured_value =  
user_action = reboot DataMover  
change_effective = reboot DataMover  
range      = (0,4294967295)  
description = Time interval between password changes
```

**NAS 5.5 PASSWD HISTORY ENABLED BY DEFAULT** (requires updtMinutes param set to become functional):

**# server\_param server\_2 -facility cifs -info srwpwd.maxHistory**

```
server_2 :  
name      = srwpwd.maxHistory  
facility_name = cifs  
default_value = 2  
current_value = 2  
configured_value =  
user_action = none  
change_effective = immediate  
range      = (1,10)  
description = Number of password kept when password changes
```

#### **Pre-NAS 5.5 Method:**

→ Password stored in krb5.keytab file, which were looked up as needed by DM Kerberos client

→ Enforce legacy keytab file behavior and disable password history:

**param cifs srwpwd maxHistory=0**

**Note:** Default setting is 2, to keep newest and previous passwd in krb5.account file. Change value to store up to 10 passwords.

#### **KEY FEATURES:**

#### **PASSWORD HISTORY & KERBEROS KEY CACHE**

→ Data Mover will keep two passwords in a new Kerberos key cache, oldest will be destroyed after each password renewal—passwords are stored in krb5.account file for NAS 5.5

→ NAS 5.5 will support Windows 2003 GPO Policies:

Domain Member: Maximum machine account password age

Domain Member: Disable machine account password changes

→ Pre NAS 5.5, passwords were stored in the krb5.keytab file and read every time a Kerberos key needed to be generated

→ Key cache will be populated during Join, whenever CIFS is started, or when password is changed, and from keytab file if authenticating NFS Users

→ Kerberos Keys are generated on the fly and cached in memory, deleted when Unjoined, when CIFS service is stopped, etc.

→ Domain policy default is to renew password every 30 days

→ cifs srwpwd.updtMinutes, default changes every 30 days, though passwd change is disabled by default

→ server\_cifs server\_2 –Join –o resetserverpasswd (use to change machine password)

→ For UNIX systems, computers do not have accounts, and keys generated by KDC will be stored in DM's keytab file using 'kadmin'

#### **NEW CIFS OUTPUT FOR PASSWORD HISTORY FEATURE:**

CIFS Server LAIP1\_DM2A[2K3] RC=2

Full computer name=laip1\_dm2a.2k3.pvt.dns realm=2K3.PVT.DNS

Comment='EMC-SNAS:T5.5.18.2'

if=laip1-2a l=192.1.6.202 b=192.1.6.255 mac=8:0:1b:42:15:9e

FQDN=laip1\_dm2a.2k3.pvt.dns (Updated to DNS)

Password change interval: 0 minutes

Last password change: Mon Feb 6 17:35:46 2006 GMT

Password versions: 3, 2

**Note:** During the Join, a temporary machine account password based on the machinename is used, but never actually saved, hence the minimum password version starts at 2, not 1.

### **SETTING UP PASSWD CHANGE FREQUENCY & PASSWORD HISTORY:**

**Note:** Passwd change feature is disabled by default with NAS 5.5. To enable Passwd Change History, run following command to specify value, in minutes. Command also updates server param file & requires Data Mover reboot.

**\$ server\_param server\_2 -facility cifs -modify srpwd.updtMinutes -value 120**

server\_2 : done

Warning 17716815750: server\_2 : You must reboot server\_2 for srpwd.updtMinutes changes to take effect.

\$ cat /nas/server/slot\_2/param

**param cifs srpwd.updtMinutes=120**

**# server\_param server\_2 -facility cifs -modify srpwd.maxHistory -value 5**

server\_2 : done

**# server\_param server\_2 -facility cifs -info srpwd.maxHistory**

server\_2 :

name = srpwd.maxHistory

facility\_name = cifs

default\_value = 2

current\_value = 5

configured\_value = 5

user\_action = none

change\_effective = immediate

range = (1,10)

description = Number of password kept when password changes

### **WINDOWS 2003 ACCESS-BASED ENUMERATION SUPPORT (ABE):**

Windows 2003 SP1 provides a new properties tab called “Access-based Enumeration” which, when enabled (“Enable access-based enumeration on this shared folder” or all Shared folders on a computer), allows administrators to set ACLs on folders and files so that the right to view and access the subfolders or files under the Share can be restricted—minimum of Read access required. Win2k3 provides a GUI, CLI, or API to utilize this feature—ABEUI.msi. Celerra provides a Windows executable called “emcabe.exe” version 1.03 which can be downloaded from the Applications & Tools CD—its function is similar to Windows “abecmd” tool.

**C:\emcabe tool>emcabe.exe**

EMCABE 01.03

emcabe [/E [/D [/G] [/T <servername>] [/A] [/S <sharename>] [/H]

/E Enables ABE on specified shares

/D Disables ABE on specified shares

/G Get ABE status on specified shares

/T <servername> Target server name

/S <sharename> Specifies the name of the shared folder on which to run the command

/A run the command on all shares of the target server

/? display this usage message

### **ENABLING ABE ON CELERRA FILE SHARE:**

**C:\emcabe tool>emcabe /E /T \\mview\_dm3 /S nyip2**

EMCABE 01.03

ABE is now enabled on share nyip2

### **STATUS OF ABE SHARES ON COMPNAME:**

**C:\emcabe tool>emcabe /G /T \\mview\_dm3 /A**

### **CELLERRA MIRRORVIEW/SYNCHRONOUS DR(MView/S):**

**Note:** Introduced with NAPA\_3 NAS 5.5.23.2, RPQ only. With the release of NAS 5.6.47, MView is now considered a “GA” product, no RPQ required.

#### **What is MirrorView/S DR for Celerra?**

A Celerra implementation that uses the new CLARiiON MirrorView functionality (Flare 16/19) to maintain synchronous copies of LUNs mirrored from a Source Celerra/CLARiiON system to a Target Celerra/CLARiiON system, and used as a Disaster Recovery (DR) solution, though the solution can also be used as a way to perform planned or unplanned maintenance by failing over to the DR location, or for testing, etc. The Source LUN is called the Primary Image and the Destination LUN is the Secondary Image. Together, the Primary and Secondary form a mirrored relationship. MirrorView can be setup in an Active/Passive or Active/Active configuration, providing full-copy synchronous remote mirroring of LUNs between arrays. Delta set technology tracks changes between transfer cycles.

#### **MIRRORVIEW SPECIFICS:**

→CLARiiON backend DR solution for Celerra, similar to SRDF, though MView consists of binaries built into NAS, vs. shell scripts (possible framework for SRDF in future)

- Synchronous mirroring of Control & Data Luns between Primary Images (mirrors) on Source array & Secondary Images on Target array
- Disaster Recovery Campus solution up to 60km distance
- MirrorView is Clariion licensed software used to copy data between Local & Remote LUNs using delta set technology to maintain synchronous mirrors
- Active/Active and Active/Passive configurations allowed (former just means each Celerra has an Active MView configuration with its own Primary LUNs mirrored to a CG with corresponding Secondary LUNs onTarget/Passive side)
- Only Gateway NS, NS40, & NSX Celerra platforms in any combination using only CX/CX3 Clariion arrays [no integrated Celerras for MirrorView]
- Multi Storage Systems allowed (if qualified CLARiiONs), but all LUNs must reside on (1) Backend—other backends used only if there is local storage & other DM's available

→Write Intent Log (WIL) LUNs are used to facilitate sync between arrays & to prevent need for full sync if a link goes down or SP crashes—tracks in-flight writes and stores ondisk. WIL LUNs are not referenced during normal operations—WILs are Private LUNs, not in Storage Group, not visible to Celerra—(1) set of WIL LUNs for Source & Dest Arrays

- (1) 128MB WIL private Lun required for each SP dedicated for MirrorView software on both Source & Target CLARiiON arrays
- WIL LUNs are Bitmaps that indicate which portions of the primary image are different from the secondary image, and assists in resyncing after link interruptions or SP crash and recovery [Recovery is either automatic or manual]

**Note:** Changes in the bitmaps are transferred during the resync vs. the entire LUN

→Fracture Log is a bitmap of 64kb chunks that is only put into use if the link between Primary & Secondary goes down for over 10 seconds, which is the heartbeat interval—changes are automatically tracked in this bitmap once a ‘fracture’ occurs—as a Best Practice, when configuring the Backend for MirrorView, provision the MView link to automatically recover when the link is restored, as opposed to a manual resync using nas\_mview –resume. Fracture Log stores differences between images in SP memory.

→Synchronous means that a Host writes to a local storage LUN, and the data is then copied via MView fibre link to Secondary storage before local storage acknowledges write back to Host—all this needs to happen before next write can occur

## **MIRRORVIEW INFORMATION:**

- New synchronous mirroring solution used for DR to remote Celerra in Active/Passive or Active/Active configurations
- Similar to traditional Celerra SRDF DR solution using Symmetrix backend
- This feature will be limited to NS & NSX “gateway” configurations only with CX arrays
- One difference from Symm SRDF is that remote devices (Secondary LUN images) are not visible to Data Movers on Destination side until after failover has occurred (i.e, Secondary LUNS are offline)
- nas\_mview commands are built into DART to Initialize, Activate, & Restore MirrorView DR (nas\_rdf is composed of scripts)
- DR solution up to 60km
- NAS 5.5.27 will support Secure NaviCLI and requires that nas\_mview –init be run to update the security file since the format has changed with this version
- Synchronous mirroring on storage array to replicate control and user data
- Storage and DM failover/failback

→Only a single Consistency Group (devicegroup) per Active side is allowed for Celerra solution with dedicated FC link between Primary & Secondary arrays

**Note:** Consistency Groups are required since there is a possibility that an SP could fail, and Flare 19 supports Consistency Groups to allow LUNs to fracture if heartbeat communications are lost between Primary & Secondary image (heartbeat every 10 seconds)

## **HOW DATA WRITES ARE HANDLED WITH MIRRORVIEW:**

1. Data is written to the Source array by a Host
2. MirrorView copies data to Remote array cache
3. Remote array performs CRC check on data in cache and sends acknowledgement to source array
4. Write-acknowledgement is returned to Host from Source array before next write is allowed

## **MIRRORVIEW REQUIREMENTS:**

- Required NAS SW=NAPA\_3 NAS 5.5.23.2 release, or higher, RPQ until NAS 5.6.47
- Required SW=MirrorView, Navisphere Mgr, & Access Logix on Arrays
- Minimum EMC Solutions Enabler V6.0.1 [/nas/symcli/bin/symcli], but NAS 5.5.23.2 comes with SYMAPI Version V6.2.1.0
- FLARE 19 Patch 034 CLARiiON array support with MirrorView license on each array
- Celerra MirrorView/S solution is somewhat different than native CLARiiON MirrorView [e.g., Celerra only supports single CG, fewer LUNs for mirroring, and Synchronous mode, etc.]
- A Single CLARiiON Consistency Group (CG) contains all the Primary & Secondary LUN images for the Celerra MView configuration

**Note:** True for Active/Passive, but there could be two CG's if using Active/Active setup

- From Celerra perspective, CG is called a ‘devicegroup’ [synonymous]
- LUNs on source must be in single Storage Group on Source array
- LUNs on destination must be in single Storage Group on Target array
- A single CLARiiON security domain is required between arrays

- Source & Target LUNs must be of same RG Type, Same size, owned by same SP, use same Disk Types, same AVM type, for each respective side's Mirrored LUNs
- Best Practice recommends single LUN per Raid Group
- Depending on array type, max of (13) Data LUNs & (3) Control LUNs allowed in the Consistency Group
- New installs preferred, but upgrading a platform for MView is allowed
- For dual-CS Celerras, CS1 must be powered off during Activate and Restore operations, and never run nas\_mview or nas\_devicegroup commands on CS1
- IP communications are used CS-to-CS and CS-to-SPs
- Control Station clock times must be within 10 minutes of each other
- MirrorView links between SPs must be on the highest numbered SP port [Either port 1 on CX500 or port 3 on CX700]
- Each SP must have a separate zone between the source & target SP ports for SPA, and same for SPB for the “mirror” ports [SP A mirror source → SP A mirror destination]
- Management of MirrorView is done via CLI only, no WebUI support
- IP comm ports 80/8000 for HTTP, port 6839 for NaviCLI, and port 443 for Naviseclli
- Each SP is zoned to remote SPA and SPB ports, and vice versa—array-to-array communications are via dedicated Fibre Channel links, usually configured on highest SP port
- Local file systems are allowed only if there are non-MirrorView Data Movers available [i.e., Celerra does not have to be pure MView; could have NS704 with (2) local DMs and (2) DMs doing MView, for example, but keep in mind that the configuration between local backend Storage and MirrorView Backend Storage must be kept separate]
- Replication, SnapSure, MPFS, autofs extend features are not Supported with MView
- Source & Destination Data Movers must have same network interfaces--there is a special procedure to mask extra NIC devices on one side vs. the other side (e.g., NSX has more interfaces than NS series)
- Rule-of-thumb; run nas\_mvview –init after any Backend or MView configuration changes are made after initial setup (do not make config changes while in failed over state—will present problems)
- Support for NS/NS, NSX/NSX, NS/NSX, with Clariion CX/CX3 backends only (no mixed backends such as NS + Symmetrix,etc)
- Only SAN-attached NS platforms, no integrated NS platform support
- Active/Passive and Active/Active configurations
- Data Movers must boot from the CLARiiON which is also used for the MirrorView devicegroup (i.e., a MView Celerra must boot from an eligible Clariion array), no Symmetrix storage allowed
- Consistency groups can only have LUNs from single Storage Group
- Requires MirrorView license on each array & single MirrorView/Sync Consistency Group for Active/Passive, two CG for Active/Active configurations
- Only the Global User account is available for MirrorView/S commands (local accounts are not allowed)
- Only single consistency group allowed for Active/Passive, consisting of all data volumes and (3) Control Luns [0, 1, 4, root\_disk, root\_ldisk, and /nas, respectively]
- Max. (13) User Luns for CX600/700 & (5) User Luns for CX400/500
- Preferred setup is for new installations, but upgrading to MirrorView/S can be accomplished if LUN requirements allow (cannot have mixed local and MirrorView storage on the same Celerra)
- MirrorView/Sync -activate or -restore commands can only be issued when CS1 is powered off, for sites with dual CS
- Requires NAS 5.5, Flare 19 (Saturn) for introduction of Consistency Groups, Solutions Enabler v6.0.1 or higher, MirrorView license & SW, Navisphere Manager, Access Logix on each array
- Max distance 60km between sites—Campus solution
- Requires IP communication between Control Stations and between Control Stations & Clariion SPs
- Clariion uses FC port 3 on CX600/CX700 on each SP for MirrorView & port 1 for CX400/500 (highest SP port on the array)
- Clariions also must communicate between sites using IP
- Standby Data Movers must have same or superset of Primary Data Mover's network configuration (use hidden\_interfaces param to mask any interfaces that should not be seen by each side)
- Required ports for IP communications are Port 80 & 8000 for HTTP, Port 6389 for NaviCLI, Port 443 for Navicli.jar (Naviseclli)

## **MIRRORVIEW RESTRICTIONS:**

- If first failover or restore fails to complete all operations, rerun –activate or –restore again (this is critical information to know!)
- Auto File System extension not support
- Celerra Replication, Celerra Replicator for iSCSI, & MPFS are not supported with MirrorView
- SnapSure is supported only if the SavVol is built on MirrorView LUNs, but Schedules will not work in failed over state
- Checkpoints can be used, but SavVol should be built on MirrorView luns
- Synchronous MirrorView may impact write operations, but should not impact Reads
- Only single CG per Active setup, and only max. of 13 Data LUNs supported (as opposed to CLARiiON itself, which does more)
- Keep CS1 out of the picture when doing MView operations!
- MAN pages are missing for nas\_mvview & nas\_devicegroup in NAS 5.5.23.2
- Requires MirrorView/S license on Clariion backend, not on Celerra side

→Celerras can run with dual control stations, but must be on CS0 when invoking Failover or Failback, and is recommended that CS1 be powered off before running nas\_mview commands

→Basic iSCSI LUNs are not supported until NAS 5.5.24.x release

### **NAS 5.5.24.x iSCSI LUN SUPPORT WITH MVIEW:**

→Destination LUN copy cannot be accessed by iSCSI host until after activating and failing over to the destination side

→Same or different iSCSI host can be used to access the LUN after failover, but during failover, lun is not available

→If performing a failover with iSCSI LUNs, stop all Windows applications on the iSCSI Client that are using the Celerra; Clear the Session information to logoff the iSCSI initiator session under target properties; Failover; Run fsck or chkdsk before accessing the iSCSI LUN after failover

### **COMMUNICATION CHANNELS USED:**

→Array-to-Array uses dedicated Fibre Channel link

→Celerra CS-to-CS, and Celerra-to-SP communications will be IP over LAN or WAN links to either local or remote Array

### **DISK TYPES FOR MIRRORVIEW:**

CMSTD—Clarion mirrored standard device (nas\_mview –init will change from CLSTD to CMSTD)

CMATA—Clarion mirrored ATA device

**Note:** Setup via Navisphere Manager interface or navicli

### **MIRRORVIEW AVM STORAGE POOLS:**

cm\_r5\_economy (8+1 mirrored RAID5)

cm\_r5\_performance (4+1 mirrored RAID5)

cm\_r1 (mirrored RAID1)

cmata\_archive (6+1 mirrored RAID5)

cmata\_r3 (4+1 or 8+1 mirrored RAID3)

**Note:** Different pool types allowed provided LUNs are configured with same AVM pool type between sites

### **RAID Types:**

RAID 1, 4+1 RAID5, 8+1 RAID5 for FC, 6+1 RAID5, RAID3 4+1or 8+1 for ATA, SATA II, LCFC

→Max. fibre channel storage depends on disk size, from 1.7 – 13.9 TB on CX700 4+1 RAID5, 3.4 – 26.0TB on CX600 8+1 RAID5, to .6 – 5.4 TB on CX500 4+1, to 1.3 – 10TB on CX500 Raid5 8+1.

→ATA storage 6+1 RAID5 6.3-8.3TB for CX500 & 16.5-21.6TB for CX700

### **DEVICEGROUP vs. CONSISTENCY GROUP:**

→When referring to MirrorView devices from the Celerra database perspective, we refer to devices in a “devicegroup”

→When referring to MirrorView devices from Clarion perspective, we are talking about the same devices in a “Consistency Group”

### **CELERRA MIRRORVIEW CLI COMMANDS (no Celerra Manager support):**

#### **# nas\_cel**

**# /nas/bin/nas\_devicegroup** (Run as nasadmin/root or DRAdmin/root, depends if system is failed over or not; Never run on CS1)

**Note:** Will not see mirrored devices unless logged in as DR account when failed over

**# /nas/sbin/nas\_mview** (Run –init nasadmin/root but –active/-restore as DRadmin/root; Never run on CS1)

**Note:** Run –info as nasadmin/root unless failed over, then use DR account/root to run -info

\$ java -jar /nas/opt/Navisphere/bin/navicli.jar -AddUserSecurity -password nasadmin -scope 0

\$ /nas/sbin/navicli -h 10.241.168.52 mirrorview (Use for info only, not configuration; must be nasadmin, not root to run)

# export NAS\_DB\_DEBUG=1 (use for troubleshooting & special command usages)

# export CLARAPI\_DEBUG=1 (use for troubleshooting MirrorView CG operations)

---

#### **# nas\_cel –create**

→Used to establish communication paths & passphrase trust between Control Stations—run on both Local and Remote systems

**# nas\_devicegroup -list** (list devicegroups) | **-info -all** (summary info) | **-acl** (NAS\_DB access control for device group) | **-suspend** (suspend mirroring, Secondary consistent but no updating) | **-resume** (resumes mirroring state and synchronizes Secondary with Primary)

**Note:** nas\_devicegroup used to manage MirrorView devices (equivalent to the Clarion Consistency Group) from Celerra db

→Run as nasadmin/root or DRaccount/root, depending on whether system is failed over or not

#### **# nas\_devicegroup -list**

| ID | name            | owner | storage ID     | acl | type  |
|----|-----------------|-------|----------------|-----|-------|
| 2  | 872_MViewCGroup | 0     | APM00030600872 | 0   | MVIEW |
| 3  | 172_AP_MView    | 500   | APM00030600872 | 0   | MVIEW |

**# nas\_devicegroup -info id=3 -sync no**

```

name      = 172_AP_MView
description = Active Passive MirrorView Consistency Group
uid       = 50:6:1:60:90:60:7:BB:0:0:0:0:0:0:0
state     = Consistent
role      = Primary
condition = Active
recovery policy = Automatic
number of mirrors = 9
mode      = SYNC
owner     = 500

```

### **mirrored disks = root\_disk,root\_ldisk,d5,d7,d8,d9,d10,d11,d12,**

```

local clarid = APM00030600872
remote clarid = APM00023700172
mirror direction = local -> remote

```

# **nas\_devicegroup** with special usage options after setting #export NAS\_DB\_DEBUG=1

special usage:

```

| -list [-backend]
| -info {<name>}lid=<id>|-all} [-backend] [-sync [yes|no]]
| -suspend {<name>}lid=<id>} [-backend]
| -resume {<name>}lid=<id>} [-backend]
| -add <name>
| -delete {<name>}lid=<id>}
| -failover {<name>}lid=<id>} [-backend] [-Force]
| -failback {<name>}lid=<id>} [-backend]

```

### **# nas\_devicegroup -list -backend**

| ID | name            | owner | storage ID     | acl | type  |
|----|-----------------|-------|----------------|-----|-------|
| 1  | 172_AP_MView    | 0     | APM00030600872 | 0   | MVIEW |
| 2  | 872_MViewCGroup | 0     | APM00030600872 | 0   | MVIEW |
| 3  | 872_MViewCGroup | 0     | APM00023700172 | 0   | MVIEW |
| 4  | 172_AP_MView    | 0     | APM00023700172 | 0   | MVIEW |

# **/nas/sbin/nas\_mview -init <cel\_name> | -info | -activate | -restore** (built into CLI—converts CLSTD to CMSTD during init)

special usage: | mcdrestore

**Note:** -init prepares source and destination Celerras for MirrorView DR & configures Data Mover standby relationships. Rerun the -init command anytime the configuration of the Celerra has been changed. Run command as root on the SOURCE Control Station.

→Run -init as nasadmin/root, run -activate/-restore as DRaccount/root, and -info depends on whether failed over or not

### **[root@nyip2 mvadmin]# /nas/sbin/nas\_mview -info**

\*\*\*\*\* Device Group Configuration \*\*\*\*\*

```

name      = 172_AP_MView
description = Active Passive MirrorView Consistency Group
uid       = 50:6:1:60:90:60:7:BB:0:0:0:0:0:0:0
state     = Consistent
role      = Primary
condition = Active
recovery policy = Automatic
number of mirrors = 9
mode      = SYNC
owner     = 500
mirrored disks = root_disk,root_ldisk,d5,d7,d8,d9,d10,d11,d12,
local clarid = APM00030600872 →This example shows failed over from Source to Target side
remote clarid = APM00023700172
mirror direction = local -> remote

```

\*\*\*\*\* Servers configured with RDFstandby \*\*\*\*\*

```

id      = 1
name    = server_2
acl     = 0
type    = nas
slot    = 2
member_of =

```

```
standby =
RDFstandby= slot=2
status :
defined = enabled
actual = online, active
***** Servers configured as standby *****
id      = 2
name    = server_3
acl     = 2000, owner=mvadmin, ID=501
type    = standby
slot    = 3
member_of =
standbyfor=
status   :
defined = enabled
actual = online, ready
```

**# export NAS\_DB\_DEBUG=1**

### **MIRRORVIEW CONTROL STATION COMMANDS:**

**# export NAS\_DB\_DEBUG=1** (troubleshooting & special command usages)

**# export CLARAPI\_DEBUG=1** (troubleshooting MirrorView CG operations)

**\$ java -jar /nas/opt/Navisphere/bin/navicli.jar -AddUserSecurity -password nasadmin -scope 0**

**\$ /nas/sbin/navicli -h 10.241.168.52 mirrorview** (Info only & not to configure; must be nasadmin—root account will not work)

**# /nas/sbin/nas\_mview –activate:** Used to failover from primary lun images to secondary lun images on remote CLARiiON, shuts down local data movers & fails over NASDB to RDF Data Movers on remote Celerra, always run from the Passive Celerra [all LUNs in devicegroup (Consistency Group) failover together]

**# /nas/sbin/nas\_mview –restore:** Used to fallback to Source CLARiiON & Active Celerra--run from Passive Celerra (on occasion, may need to complete restore by running /nasmcd/sbin/nas\_mview –restore on source side if previous restore was incomplete)

**# ls -la /nas/sbin |grep nas\_mview** (Link to nas\_cmd, issued only from Destination side for –activate or –restore operations)

lrwxrwxrwx 1 nasadmin nasadmin 14 Sep 5 14:32 nas\_mview -> ../bin/nas\_cmd

**# ls -la /nasmcd/sbin |grep nas\_mview** (Special executable for /nasmcd/sbin restore on Source)

-rwxr-x--- 1 root nasadmin 9492 Sep 5 14:09 nas\_mview

**Rule of Thumb:** If first –restore fails, run a 2nd time from Destination side. However, as last resort, or as directed by Error message, run the –restore from Source side using special executable /nasmcd/sbin/nas\_mview –restore

### **RUNNING NAVICLI MIRRORVIEW COMMANDS FROM CONTROL STATION:**

**Note:** Setup the security context first, then run navicli commands as nasadmin

**\$ java -jar /nas/opt/Navisphere/bin/navicli.jar -AddUserSecurity -password nasadmin -scope 0**

**\$ /nas/sbin/navicli -h 10.241.168.52 mirrorview -listlog**

Storage Processor: SP A

Lun Number: 150

Storage Processor: SP B

Lun Number: 151

### **SOME USER ACCOUNTS THAT ARE REQUIRED FOR MVIEW OPERATIONS:** (know when to use)

→Global CLARiiON domain account (needed for MView init; single Domain concept—see Backend Setup Guide—setup using Navisphere)

→nasadmin & root account passwords are still used at proper times

→DRadmin (special MView admin account) account is required—can be given any name [e.g., mvadmin, dradmin, etc.]

nasadmin & root accounts used on Source MView Celerra for normal management of Celerra

**Note:** DR admin account does not exist on Source unless you have an Active/Active MirrorView configuration

→nasadmin & su root used on Target Celerra for nas\_mview -init

→DRadmin account is used on Target Celerra to perform –activate, failover or –restore failback operation, & to administer/manage the Celerra when in a failed over state (dradmin account is set for NASDB at /nas/rdf/500, the location of NASDB after failover)

# env

USER=mvadmin

PATH=/bin:/usr/bin:/usr/local/bin:/usr/bin/X11:/usr/X11R6/bin:/home/mvadmin/bin:

**Note:** Do not use nasadmin account to reboot or perform any changes while in failed over state (may present problems during restore)

## **OVERVIEW OF MVIEW INITIALIZATION:**

- Prompts for global Clariion user account info
- Runs diskmark on Source & Destination Celerras
- Validates MirrorView/S storage configuration
- Prompts for Source & Destination Data Mover configuraiton
- Prompts for DR Admin account
- Configures Source & Destination systems

**Note:** See MirrorView Troubleshooting section for information on navicli and naviseccli MirrorView commands

## **Basic MirrorView Configuration Overview—Parts I, II, III**

- I. MirrorView Backend Setup: Configuring Arrays for MView (aka Engineering Backend Setup Guide)
- II. MirrorView Celerra Setup: Initializing & configuring Celerra for either Active/Passive, or Active/Active (aka MirrorView Tech Module)
- III. Conducting MirrorView Failover, Restore, Suspend Operations

## **PART I. MIRRORVIEW BACKEND SETUP:**

### **SUMMARY OF STEPS:**

1. Create CLARiiON Global domain account on either array
2. Establish MirrorView comms links between arrays
3. Create & allocate WIL LUNs on both arrays
4. Prepare Source LUNs (aka Primary LUN images)
5. Prepare Target LUNs from Source side (aka Secondary LUN images)
6. Create ‘Remote Mirrors’ on Primary Source LUNs
7. Add Secondary Mirror Images to respective Source Remote Mirror Images
8. Wait for LUNs to synchronize before continuing...
9. Add Source LUNs & Target LUNs to respective Storage Groups
10. Create CG on Source (Active side) & add Source Remote Mirrors to CG

### **I. CONFIGURE ARRAYS FOR MIRRORVIEW DR—ACTIVE/PASSIVE:**

#### **1. Create Clariion Global Administrative Domain:**

**Note:** Both Clariions must be in same Global Administrative Domain (Use nasadmin user and nasadmin as default password; can select Systems via IP Address or scan a Subnet for IPs; assign one storage system as domain master, & other storage system to domain—can perform these steps from either Storage system, doesn’t matter)

**Navisphere>File>Setup Domain>Configure Domain>IP Address of System>Selected Systems>o.k.**

**Navisphere>File>Setup Domain>Select Master>enter IP address of Clariion system to change Domain Master** (One array is assigned the Domain Master role)

#### **From Naviseccli:**

```
# /nas/opt/Navisphere/bin/naviseccli -h mview_source -user nasadmin -password nasadmin -scope 0 domain -setmaster  
10.241.168.57  
# /nas/opt/Navisphere/bin/naviseccli -h mview_source -user nasadmin -password nasadmin -scope 0 domain -add  
10.241.168.52
```

#### **2. Establish MirrorView communication links between arrays:**

- a) (2) fibre channel zones required, SPA-to-SPA, SPB-to-SPB point-to-point or fabric on highest numbered SP port
- b) Then, the logical communication path between storage systems via Navisphere must be created

**Navisphere>rightclick array>MirrorView>Manage Mirror Connections>Unconnected/Unknown Systems>select>Enable**

**Note:** Highest numbered port is Port 3 on CX600/700 and Port 1 on CX400/500. This procedure can be conducted on either Source or Target system and it automatically creates the logical link path between Source & Target SP’s in each direction.

**# /nas/opt/Navisphere/bin/naviseccli -h 10.241.168.57 (Source--SPA) -user nasadmin -password nasadmin -scope 0 mirror -enablepath 10.241.168.52 (Target—SPA)**

**Note:** Command activates all the necessary paths between all SPs: SPA local to SPA remote; SPA remote to SPA local; SPB local to SPB remote; SPB remote to SPB local

#### **3. Create & allocate 128MB WIL (Write Intent Log) LUNs (aka Private LUNs) on Source & Target arrays—(1) each for SPA & SPB on each side:**

**a.) Create two RAID GROUPs for Source & Destination Arrays:**

**Navisphere>Rightclick array>Create Raid Group>select (5) disks>Initiate create RAID Group operation? yes**

**b.) Create two WIL LUNs and bind each WIL into a separate RAID GROUP for Source & Destination Arrays:**

**Raid Groups>RAID Group 30 [Unbound] >Bind LUN** >Ensure RAID Type is set to RAID 5, assign LUN ID number (150), keep Number of LUNs to bind as 1, select the desired SPA, enter 128 for LUN Size and select \*MB for megabytes, apply

**Raid Groups>RAID Group 31 [Unbound] >Bind LUN** >Assign LUN ID (151), Select desired SPB, enter LUN Size 128MB, apply

**Note:** Select each new RAID Group & bind one WIL Private LUN into each respective RAID Group—one for SPA and one for SPB.

**Binding WIL Luns via CLI:**

\$ navicli -h spa bind r5 100 -rg 2 -rc 1 -wc 1 -sp a -sq mb -cap 128

\$ navicli -h spb bind r5 101 -rg 3 -rc 1 -wc 1 -sp b -sq mb -cap 128

**c.) Allocate WIL LUNs for Source and Destination Arrays:**

Source Navisphere>rightclick array>MirrorView>Allocate Write Intent Log>select Available LUNs 100 & 101 and move to Selected LUNs>o.k.

Destination Navisphere>rightclick array>MirrorView>Allocate Write Intent Log>select Available LUNs 150 & 151 and move to Selected LUNs>o.k.

**Allocating WIL Luns via CLI:**

\$ /nas/opt/Navisphere/bin/naviseccli -h spa\_source -user nasadmin -password nasadmin -scope 0 mirror -sync -allocatelog -spa 100 -spb 101

\$ /nas/opt/Navisphere/bin/naviseccli -h spa\_destination -user nasadmin -password nasadmin -scope 0 mirror -sync -allocatelog -spa 150 -spb 151

**Note:** Move luns 100 & 101 to “Selected LUNS” and click o.k., Confirm Allocate Intent Log—Yes. WIL records which areas of primary image differ from secondary image and assists in resynchronization after failure. WIL LUNs are not part of the mirrored configuration and should not be created from the same RAID group as the Control LUNs. Create WIL LUNs in separate RAID Groups (SPA WIL LUN & SPB WIL LUN)—note that WIL LUNs are NOT added to the Storage Group & are therefore not directly visible to Celerra—used only in the background as “Private LUNs” for MirrorView functionality. Repeat for Destination array.

**d.) Repeat a-c steps on DESTINATION Array**

**4. Prepare Primary LUN Images on Source (Control Luns HLU 0, 1, 4, & up to (13) user LUN images):**

**Note:** Create Data LUNs if required, and add to Celerra Storage Group on Source. Must bind at least one Data lun using HLU 16 or higher.

**Example:** Created new RAID Group and 5 disks. Rightclicked new RAID Group and selected BIND LUN, created ALU lun numbers 90-95, choose “Default Owner Auto” for LUN Owner assignments.

**a.) Create RAID GROUP for Data Luns:**

Navisphere>Create Raid Groups and Bind User LUNs

**b.) Bind LUNs, select LUNS, select Default Owner Auto**

Navisphere>rightclick on new RAID GROUP>Bind LUN>Set number of LUNs, Default Owner Auto, LUN Size, o.k.

**c.) Add User LUNs to Storage Group:**

Navisphere>Storage Groups>Celerra\_NYIP1>rightclick and “Select Luns”, making sure to assign correct HLU numbers

**Note:** Assign correct HLU numbers to Data LUNs (> than HLU 16)—at this point, you could still remove from Storage Group and re-add with correct HLU assignments. All Control & Data LUNs should be in the correct Storage Group on the Source at this time. This procedure assumes that Control Luns have already been added to the Storage Group as part of the initial NAS installation.

**Binding and Assigning Data LUNs to Storagegroup via CLI:**

\$ navicli -h spa bind r5 16 -rg 4 -rc 1 -wc 1 -sp a -sq gb -cap 10 (capacity 10Gb)

\$ navicli -h spa storagegroup --addhlu -gname Celerra\_Mirrorview -hlu 16 -alu 16

**Note:** Created single Data LUN 16—repeat above steps to create additional Data Luns

**5. Create Secondary Image LUNs using Source Navisphere for Control & Data Luns:**

**Note:** As with Steps 1-4, this is done from Navisphere on the SOURCE Clarion system, but the Secondary Image LUNs themselves will be created on the Destination array. This ensures that all Secondary images will match with same size, the same SP, and be of the same RAID type as the Source.

**a.) Create Secondary Image Luns for Control Luns:**

Navisphere>expand SPA>rightclick LUN 0 (choose correct LUN number in display)>MirrorView>Create Secondary Image

LUN [Secondary Storage system is displayed] >Select an ALU LUN number “Secondary Image LUN” box >Select existing RAID Group or New RAID Group, ensuring that the RAID Group type is identical to Source. Repeat process for Source Control Luns 1 & 4. Could also create LUN directly from the Raid Group section of Navisphere. Please keep in mind that the LUN numbers seen in Navisphere are ALU numbers—therefore, LUN 0 could be a different number entirely from that displayed in Navisphere.

**Note:** Later, when adding the Secondary Control Luns and Data Luns to the Remote system’s Storage Group, the HLU numbers will become Secondary Image LUN number 6, 7, 9, respectively—data HLU numbers must be >than 16)

**Using CLI to create Secondary Image LUN 0:**

\$ navicli -h spa bind r5 17 -rg 3 -rc 1 -wc 1 -sp a -sq gb -cap 11

**Note:** Creates ALU 17 same size as source control lun 0

**b. Create Secondary Image Luns for Data Luns:**

Navisphere>expand SPA>rightclick data LUN 91, 93, 95 >MirrorView>Create Secondary Image LUN>Secondary Image

LUN: 91 | 93 | 95 >Existing or New RAID Group

**Navisphere>expand SPB>rightclick data LUN 90, 92, 94 >MirrorView>Create Secondary Image LUN>Secondary Image LUN: 90 | 92 | 94 >Existing or New RAID Group**

**Note:** The option to select ‘Use Write Intent Log’ seems to be only available when creating the Remote Mirrors and not when creating the Secondary LUN Images—but, the system by default will select “Use Write Intent Log” so it doesn’t seem to matter.

**Using CLI to create Secondary Image for Data Luns:**

**\$ navicli -h spa bind r5 20 -rg 3 -rc 1 -wc 1 -sp a -sq gb -cap 10**

**Note:** Select Raid Group ID and number of disks, then apply. After RAID Group is created, go back and add Data Lun 91 as Secondary Image LUN, select correct Raid Group that was just created, make sure RAID 5 is selected in Select RAID Type, click o.k. to bind LUN 91. Repeat process for each Data Lun to be added to the Raid Group.

## **6. Create Remote Mirrors for each Primary Image for Control (0, 1, 4) & Data LUNs (Source side):**

**Navisphere>SPA>rightclick each primary image LUN>MirrorView>Create Remote Mirror>Mirror Type:**

**Synchronous>Enter a Name: Lun 0 Primary>Select ‘Use Write Intent Log’>Select LUN 0>Apply**

**Note:** Remote Mirrors hold changes made to primary data blocks for each lun, stored either in SP memory or WIL. Perhaps an easier way to ensure that correct LUNs are being used, go to **RAID Group>LUN X>MirrorView>Create Remote Mirror>**check the **Use Write Intent Log box, assign Name ‘Lun 0 Primary’**, apply. Repeat for all LUNs to be mirrored

**Creating Remote Mirror of Source LUN using CLI:**

**\$ /nas/opt/Navisphere/bin/naviseclli -h spa -user nasadmin -password nasadmin -scope 0 mirror -sync -create -name spa\_m0 -lun 0 -description "mirror for src LUN 0" -usewriteintentlog**

## **7. Add Secondary Image LUNs to their respective Source Remote Mirrors:**

**Navisphere>Array>Remote Mirrors>rightclick LUN 0>Add Secondary Image>Select appropriate matching LUN 6>**Set recovery policy to Automatic & Initial Sync Required—LUN will go FAULTED, then to TRANSITIONING state, as LUNs synchronize. We are adding the remote lun that matches up with lun 0, and so on, in this step.

**Note:** Each Source Image LUN, aka Remote Mirror, will need to have an equivalent Secondary image LUN matched up with it. Repeat steps and match Source Data Lun number to corresponding Secondary Image Lun number. This is the step where the Source & Destination LUNs become ‘paired’. Each LUN pair must go through a synchronization process before the destination luns can be added to the Remote Storage Group. For this step, the chosen ALU numbers for each of the Secondary Image Control & Data Luns does not matter—might be useful to label the Luns in Navisphere, however, to keep track of ALU, System, HLU numbers.

**Adding Destination LUN Image to Mirror from CLI:**

**\$ /nas/opt/Navisphere/bin/naviseclli -h spa -user nasadmin -password nasadmin -scope 0 mirror -sync -addimage -name spa\_m0 -arrayhost 10.6.1.112 -lun 17**

**Note:** The option to select ‘Use Write Intent Log’ seems to be only available when creating the Remote Mirrors and not when creating the Secondary LUN Images—but, the system by default will select “Use Write Intent Log” so it doesn’t seem to matter.

## **8. Wait for Synchronization of Secondary Images to Remote Mirrors to complete before proceeding:**

**Expand Remote Mirrors>select LUN & properties>Secondary Image tab>monitor State and %Synchronized**

**Checking progress from Celerra CLI:**

```
# nas_devicegroup -info id=3 -sync no
name      = 172_AP_MView
state     = Synchronizing
```

When completed, Data Luns show up as State: Synchronized with %Synchronized: 100

Control Luns show up as State: Consistent with %Synchronized: Not Applicable or 100

## **9. Add Remote Array LUNs to Remote Storage Group & assign correct HLU numbers:**

After synchronization has completed, add Control Luns and Data Luns from Remote array to the Secondary Array Storage Group on the Remote Secondary Side using Navisphere. Make sure to assign the correct HLU numbers at this time—Control LUN HLU assignments are 6, 7, 9. Data HLU Lun numbers must be >than decimal 16.

**Navisphere>Storage Groups>Celerra\_nyip2>highlight & Select LUNs>Add LUNs to Storage Group>**Assign correct HLU numbers for Control LUNs 6, 7, & 9 before applying—all data HLU numbers must be >than decimal 16

**→Source Control Luns HLU 0, 1, 4 must be matched up with Secondary Image Luns HLU 6, 7, 9, respectively [DOS, UFSLOG, NBSNAS, respectively]**

**Adding Control Luns to Destination Storage Group from CLI:**

**\$ navicli -h element-spa storagegroup -addhlu -gname Celerra\_element -hlu 6 -alu 17**

**Note:** Assigns destination LUN 0, hlu 6, to destination Storagegroup as alu 17

## **10. Create MirrorView/S Consistency Group on Source Array & Add Remote Mirrors to Group:**

**Navisphere>Array>Consistency Group>Create Group>Mirror Type:Synchronous >Name: SVT MirrorView>Select all the Available Remote Mirrors>Recovery Policy automatic**

**Note:** Limited to (16) total Remote Mirror luns for NS600/700 and (8) total Remote Mirror luns for NS400/500. Can also create the Consistency Group from **Navisphere>Array>rightclick>MirrorView>Create Group**

**Create Consistency Group on Source & Add Remote Mirrors using CLI:**

**\$ /nas/opt/Navisphere/bin/naviseclli -h spa -user nasadmin -password nasadmin -scope 0 mirror -sync -creategroup -name cg\_source -description "consistency group from src to dst"**

\$ /nas/opt/Navisphere/bin/navisecl -h spa -user nasadmin -password nasadmin -scope 0 mirror -sync -addtogroup -name cg\_source -mirrorno cg\_m17

**Note:** Repeat above command to add Remote Mirrors for all System & Data LUNs to the Consistency Group

## **CONFIGURING MIRRORVIEW DR ON CLARIION SYSTEMS—ACTIVE/ACTIVE:**

### **ACTIVE/ACTIVE MIRRORVIEW SETUP:**

#### **ACTIVE/ACTIVE BACKEND CONFIGURATION:**

→There will be (1) Consistency Group (Celerra devicegroup) for each active side—make sure the CG names are unique

**Note:** See emc138372 if a Consistency Group has been renamed in Navisphere after MirrorView has been configured and initialized

→Do steps 1-11 above for Backend configuration to one Celerra/Clariion combination, then repeat steps 4-11 for the other Celerra/Clariion combination—each side will have a Consistency Group—an Active/Active setup really means that you have an Active/Passive relationship on each Celerra

**Note:** A recommendation would be to carefully name all Luns in Navisphere and map out the ALU/HLU Lun assignments for each MirrorView configuration and its respective Storage system

## **CONFIGURATION OF CONTROL LUNS BETWEEN SOURCE & DESTINATION:**

LUN 0 11GB Primary Image LUN maps to 11GB Secondary Image LUN 6

LUN 1 11GB Primary Image LUN maps to 11GB Secondary Image LUN 7

LUN 4 2GB Primary Image LUN maps to 2GB Secondary Image LUN 9

## **PART II. MIRRORVIEW CELERRA SETUP:**

### **OVERVIEW**

1. Register CS's with each other using nas\_cel
2. Initialize config on Target Celerra (Passive side) using nas\_mview -init (nasadmin/root account) and configure RDF Server relationships when prompted

### **INITIALIZING MIRRORVIEW/S CONFIG BETWEEN SOURCE & DESTINATION CELERRAS FOR ACTIVE/PASSIVE MVIEW:**

#### **Preparatory Information and Best Practices:**

→As Best Practice, run primary Internal Usermapper on Source server\_2—db and complete functionality will follow when conducting failover to Destination side

→Know IP Addresses of both Control Stations, management accounts and passwords, passphrase account and passwords, and create a new DR Admin account for MirrorView

→Must know Clarrion management authentication, Consistency Group, Mirrors (Primary & Secondary Images), and Storage Groups (Source and Destination)

→Celerra times must be within 10 minutes of each other

→Use nas\_cel -create to identify each Celerra to the other

→Passphrase used, and must be same on both Sides (6-15 characters)

→Configure remote communication

→Rerun -init on Destination Side after any config or DM relationship changes

→Run **nas\_mview -init** from destination control station (Prompts for Global User account, runs diskmarks on both source & destination, validates MirrorView/S configuration, prompts for DM configuration, prompts for DR admin account, configures source & destination systems)

**Note:** -init prepares source and destination Celerras for Mirror/View DR, but is run on the Destination side

#### **1. Register Control Stations for Remote Communications:**

##### **a.) Register Source CS to Remote CS:**

# **nas\_cel -name nyip2 -create 10.241.168.48 -passphrase mvadmin**

operation in progress (not interruptible)...

```
id      = 1
name    = nyip2
owner   = 0
device  =
channel =
net_path = 10.241.168.48
celerra_id = APM00030600872000F
passphrase = mvadmin
```

##### **b.) Register Destination CS to Source CS:**

# **nas\_cel -name nyip1 -create 10.241.168.50 -passphrase mvadmin**

operation in progress (not interruptible)...

```
id      = 1
name    = nyip1
owner   = 0
device  =
channel =
net_path = 10.241.168.50
celerra_id = APM000237001720000
passphrase = mvadmin
```

**Note:** Leaving the –name option off gives default name of “Dom1”. To rename, using –modify command:

**#nas\_cel -modify dom1 -passphrase mvadmin -name dom2**

**c. Verify results of Registration process:**

**# nas\_cel -l**

```
id  name      owner mount_dev channel net_path      CMU
0   nyip1     0          10.241.168.50 APM000237001720000
1   nyip2     0          10.241.168.48 APM00030600872000F
```

**Note:** Run on each Celerra to verify that the one is registered with the other

**# nas\_cel -info id=1**

```
id      = 1
name    = nyip2
owner   = 0
device  =
channel =
net_path = 10.241.168.48
celerra_id = APM00030600872000F
passphrase = mvadmin
```

**Note:** Both Source & Destination Control Stations should have the following passphrase files and contents

**/nas/httpreplication/inbound**

**/nas/httpreplication/outbound**

```
# cat APM000237001720000
```

```
mvadmin
```

```
/nas/httpreplication/inbound
```

```
/nas/httpreplication/outbound
```

```
# cat APM00030600872000F
```

```
mvadmin
```

**2. Initialize the MirrorView Configuration—Run from Destination Celerra:**

**# /nas/sbin/nas\_mview -init nyip1**

Celerra with MirrorView/Synchronous Disaster Recovery

Initializing nyip1 --> nyip2

Contacting nyip1 for remote storage info

Local storage system: APM00030600872

Remote storage system: APM00023700172

Enter the Global CLARiiON account information →Global Administrator account created in Setup of MirrorView Clariion security

**Username: nasadmin**

Password: \*\*\*\*\* Retype your response to validate

Password: \*\*\*\*\*

Discovering storage on nyip1 (may take several minutes)

Setting security information for APM00030600872

Discovering storage APM00023700172 (may take several minutes)

Discovering storage (may take several minutes)

Contacting nyip1 for remote storage info

Gathering server information...

Contacting nyip1 for server capabilities...

Analyzing server information...

Source servers available to be configured for remote DR

- 
- 1. server\_2:nyip1
  - 2. server\_3:nyip1 [ local standby ]
  - v. Verify standby server configuration
  - q. Quit initialization process

c. Continue initialization

**Select a nyip1 server: 1**

Destination servers available to act as remote standby

- 
- 1. server\_2:nyip2
  - server\_3:nyip2 [ local standby ]
  - b. Back

**Select a nyip2 server: 1**

Source servers available to be configured for remote DR

- 
- 1. server\_2:nyip1 [ remote standby is server\_2:nyip2 ]
  - 2. server\_3:nyip1 [ local standby ]
  - v. Verify standby server configuration
  - q. Quit initialization process
  - c. Continue initialization

**Select a nyip1 server: 2**

Destination servers available to act as remote standby

- 
- server\_2:nyip2 [ is remote standby for server\_2:nyip1 ]
  - 2. server\_3:nyip2 [ local standby for remote standbys ]
  - b. Back

**Select a nyip2 server: 2**

Source servers available to be configured for remote DR

- 
- 1. server\_2:nyip1 [ remote standby is server\_2:nyip2 ]
  - 2. server\_3:nyip1 [ remote standby is server\_3:nyip2 ]
  - v. Verify standby server configuration
  - q. Quit initialization process
  - c. Continue initialization

**Select a nyip1 server: c**

Standby configuration validated OK

Enter user information for managing remote site nyip1

**Username: mvadmin**

Password: \*\*\*\*\*\* Retype your response to validate

Password: \*\*\*\*\*\*

**Initializing Active-->Passive (nyip1-->nyip2)**

Do you wish to continue? [yes or no] yes

Updating MirrorView configuration cache

Setting up server\_3 on nyip1

Setting up server\_2 on nyip1

Rebooting server\_2 on nyip2 as standby ... done

Creating user account mvadmin

Setting acl for server\_3 on nyip2

Setting acl for server\_2 on nyip2

Updating the Celerra domain information

Creating device group SVT MirrorView DR on nyip2

Creating device group SVT MirrorView DR on nyip1

done

**Note 1:** When prompted, configure the appropriate Servers on Source & Target sides, then type “c” to continue initialization, and setup the special User Account and Password to be used for MirrorView Operations. Notice that the ACLs for servers are set to 2000, a device group is created on each Celerra based on the Consistency Group name used earlier, and the Servers rebooted. Use the –v to verify the Standby Server configuration after the Servers have been selected.

**Note 2:** Keep in mind that Data Mover relationships and hardware compatibility will be enforced during initialization sequence.

Check MirrorView Tech Module for more information.

**Using Dissimilar Platforms with Different Network Interface Hardware—Source NS702G vs. Destination NSX:**

NS702G has (6) copper cge ports, NSX has only (5) copper cge ports. To logically hide the cge6 port on NS702G, enter following line on Source side /nas/site/nas\_param file:

**hidden\_interfaces:cge6:**

**3. Verify MirrorView Initialization on Destination (Passive Side):**

**Note:** nas\_devicegroup and nas\_mview commands may not output anything on the Destination side, unless the destination is running as the DR site. The –activate is initiated from the Destination side, and after failover, the following commands would be used to verify the state of MirrorView on Celerra.

### # cat /etc/passwd |grep mvadmin

```
mvadmin:x:50001:201::/home/mvadmin:/bin/bash
```

**Note:** This account is only created on the Destination side of an Active/Passive MirrorView configuration

### # nas\_devicegroup -list

```
ID name          owner storage ID    acl type
1  SVT MirrorView DR  50001  APM00030600872 0  MVVIEW
```

### # /nas/sbin/nas\_mview -info

### # nas\_devicegroup -info id=1

Sync with CLARiiON backend ..... done

```
name      = SVT MirrorView DR
description =
uid       = 50:6:1:60:90:60:7:BB:0:0:0:0:0:0:0:0
state     = Consistent
role      = Secondary
condition = Active
recovery policy = Automatic
number of mirrors = 9
mode      = SYNC
owner     = 50001
mirrored disks =
local clarid = APM00030600872
remote clarid = APM00023700172
mirror direction = local <- remote
```

### # nas\_cel -l

```
id  name      owner mount_dev channel net_path      CMU
0   nyip2     0        10.241.168.48 APM00030600872000F
1   nyip1     50001  /dev/ndj1  /dev/ndg  10.241.168.50 APM000237001720000
```

### # nas\_server -l

```
id  type acl slot groupID state name
1   4   2000 2      0   server_2
2   4   2000 3      0   server_3
```

### # nas\_server -a -i

```
id      = 1
name    = server_2
acl     = 2000, owner=mvadmin, ID=50001
type    = standby
slot    = 2
member_of =
standbyfor=
status   :
defined = enabled
actual = online, ready
id      = 2
name    = server_3
acl     = 2000, owner=mvadmin, ID=50001
type    = standby
slot    = 3
member_of =
standbyfor= server_2
status   :
defined = enabled
actual = online, ready
```

## **4. Verify MirrorView Initialization on Source (Active Side):**

### # nas\_devicegroup -list

```
ID name          owner storage ID    acl type
```

1 SVT MirrorView DR 0 APM00023700172 0 MVIEW

## # nas\_devicegroup -info id=1

Sync with CLARiiON backend ..... done

name = SVT MirrorView DR  
 description =  
 uid = 50:6:1:60:80:60:5:35:0:0:0:0:0:0:0:0  
 state = Consistent  
 role = Primary  
 condition = Active  
 recovery policy = Automatic  
 number of mirrors = 9  
 mode = SYNC  
 owner = 0  
 mirrored disks = root\_disk,root\_ldisk,d5,d7,d8,d9,d10,d11,d12,  
 local clarid = APM00023700172  
 remote clarid = APM00030600872  
 mirror direction = local -> remote

## # /nas/sbin/nas\_mview -info

\*\*\*\*\* Device Group Configuration \*\*\*\*\*

name = SVT MirrorView DR  
 description =  
 uid = 50:6:1:60:80:60:5:35:0:0:0:0:0:0:0:0  
 state = Consistent  
 role = Primary  
 condition = Active  
 recovery policy = Automatic  
 number of mirrors = 9  
 mode = SYNC  
 owner = 0  
 mirrored disks = root\_disk,root\_ldisk,d5,d7,d8,d9,d10,d11,d12,  
 local clarid = APM00023700172  
 remote clarid = APM00030600872  
 mirror direction = local -> remote

\*\*\*\*\* Servers configured with RDFstandby \*\*\*\*\*

id = 1  
 name = server\_2  
 acl = 1000, owner=nasadmin, ID=201  
 type = nas  
 slot = 2  
 member\_of =  
 standby = server\_3, policy=auto  
 RDFstandby= slot=2

status :  
 defined = enabled  
 actual = online, active

id = 2  
 name = server\_3  
 acl = 1000, owner=nasadmin, ID=201  
 type = standby  
 slot = 3

member\_of =  
 standbyfor= server\_2  
 RDFstandby= slot=3  
 status :  
 defined = enabled  
 actual = online, ready

\*\*\*\*\* Servers configured as standby \*\*\*\*\*

No servers configured as standby

## # nas\_cel -l

id name owner mount\_dev channel net\_path CMU

```
0 nyip1 0 10.241.168.50 APM000237001720000
1 nyip2 0 10.241.168.48 APM00030600872000F
```

**# nas\_server -l**

```
id type acl slot groupID state name
1 1 1000 2 0 server_2
2 4 1000 3 0 server_3
```

**# nas\_server -a -i**

```
id = 1
name = server_2
acl = 1000, owner=nasadmin, ID=201
type = nas
slot = 2
member_of =
standby = server_3, policy=auto
RDFstandby= slot=2 →MirrorView Standby
status :
defined = enabled
actual = online, ready
id = 2
name = server_3
acl = 1000, owner=nasadmin, ID=201
type = standby
slot = 3
member_of =
standbyfor= server_2
RDFstandby= slot=3 →MirrorView Standby
status :
defined = enabled
actual = online, ready
```

**# nas\_pool -l**

```
id inuse acl name
3 n 0 clar_r5_performance
13 n 0 cm_r5_performance →New MirrorView performance pool
```

**# nas\_disk -l**

```
id inuse sizeMB storageID-devID type name servers
1 y 11263 APM00023700172-0000 CMSTD root_disk 1,2 →New CMSTD type when MirrorView init. completed
2 y 11263 APM00023700172-0001 CMSTD root_ldisk 1,2
3 y 2047 APM00023700172-0002 CLSTD d3 1,2
4 y 2047 APM00023700172-0003 CLSTD d4 1,2
5 y 2047 APM00023700172-0004 CMSTD d5 1,2
6 y 2047 APM00023700172-0005 CLSTD d6 1,2
7 n 10239 APM00023700172-005B CMSTD d7 1,2 -----output abridged-----
```

**Note:** Data Movers should be set to 100FullDuplex while CS interface should be set to Auto. By default, after the initialization, a new cm\_r5\_performance Storage Pool is created to combine all the data luns used in MirrorView.

**# cat /nas/server/server\_setup**

```
2:0:1:1:3,1000,0,:  
1:2:1:3:2,1000,0,:  
Note: Shows local Standby Server relationship on Source & MView RDF Standby relationship on Destination side
```

## **INITIALIZING MIRRORVIEW/S CONFIG BETWEEN SOURCE & DESTINATION CELERRAS FOR ACTIVE/ACTIVE MVVIEW:**

**ACTIVE/ACTIVE SETUP PRECAUTION:** mvadmin UID numbers need to be different on each Celerra. The DR Admin account UID needs to be different on each Celerra CS that has been assigned the Active MView role—use following procedure to check and/or make sure account UIDs are different. Or, just follow the Active/Active steps in succeeding slides—let UID assignment default to UID 500 for first Active side setup, then create dummy account on other Control Station using same UID so that the –init process will choose the next higher UID value for the 2nd Active side (e.g., UID 501).

**Note:** If running NIS on Control Station, the default UID will be 50001, otherwise, the default UID assigned by the –init script is 500. The purpose for the separate UIDs ensures that the correct DR Admin account administers each side correctly when performing –active or –restore commands.

1. Check UID for DR Admin account on first Control Station setup as Active side
2. Make sure remote Active side uses the same UID account as a “dummy user account”—this will ensure that the ‘other’ side does not use the same UID value when creating the DR Account, which is the whole point of this procedure.

# /usr/sbin/useradd -u 500 drdummy (this will ensure that –init script will use next UID, 501, on remote side)

# tail /etc/passwd

drdummy:x:500:500::/home/drdummy:/bin/bash

3. Run –init on the remote Active side and the next UID to be assigned to the DR Admin account would be 501

### **ACTIVE/ACTIVE INITIALIZATION STEPS:**

- 1. Check to see if UID 500, 501, or 5001 have been used on either Linux Control Station:**

# tail /etc/passwd

nasadmin:x:201:201::/home/nasadmin:/bin/bash

emailhome:x:300:99::/:/bin/sh →Last UID assigned was 300

- 2. Register each Control Station with the other using the same Passphrase:**

[root@nyip1 rdf]# nas\_cel -name nyip2 -create 10.241.168.48 -passphrase mvadmin

operation in progress (not interruptible)...

```
id      = 2
name    = nyip2
owner   = 0
device  =
channel =
net_path = 10.241.168.48
celerra_id = APM00030600872001D
passphrase = mvadmin
```

[root@nyip2 outbound]# nas\_cel -n nyip1 -create 10.241.168.50 -passphrase mvadmin

operation in progress (not interruptible)...

```
id      = 1
name    = nyip1
owner   = 0
device  =
channel =
net_path = 10.241.168.50
celerra_id = APM000237001720000
passphrase = mvadmin
```

- 3. Verify Passphrase and Control Station-to-Control Station Registration:**

[root@nyip1 site]# nas\_cel -list

| id | name  | owner | mount_dev | channel | net_path      | CMU                |
|----|-------|-------|-----------|---------|---------------|--------------------|
| 0  | nyip1 | 0     |           |         | 10.241.168.50 | APM000237001720000 |
| 2  | nyip2 | 0     |           |         | 10.241.168.48 | APM00030600872001D |

[root@nyip2 site]# nas\_cel -list

| id | name  | owner | mount_dev | channel | net_path      | CMU                |
|----|-------|-------|-----------|---------|---------------|--------------------|
| 0  | nyip2 | 0     |           |         | 10.241.168.48 | APM00030600872001D |
| 1  | nyip1 | 0     |           |         | 10.241.168.50 | APM000237001720000 |

# cat /nas/httpreplication/inbound/APM00030600872001D (From Source system)

mvadmin

# cat /nas/httpreplication/outbound/APM000237001720000 (From Remote system)

mvadmin

**Note:** When the CS registers the remote Celerra, a file containing the passphrase is named after the Remote Clariion serial number and stored in both the /nas/httpreplication/inbound and outbound directories. As a troubleshooting note, if a prior DR setup had been in place at one time, there might be multiple passphrase files.

/nas/httpreplication/inbound | outbound

# cat /nas/site/cshosts (From Source system)

2:nyip2:0:10.241.168.48:::APM00030600872001D:

0:nyip1:0:10.241.168.50:::APM000237001720000:

# cat /nas/site/cshosts (From Remote System)

1:nyip1:0:10.241.168.50:::APM000237001720000:

0:nyip2.localdomain:0:10.241.168.48:::APM00030600872001D:

- 4. Run MirrorView Initialization on First Side for Active/Active Setup:**

**Note:** Example shows -init from Remote Celerra “NYIP2” to establish the MirrorView relationship with Local Celerra “NYIP1” as the Primary, with Server\_2 local and Server\_2 remote as the RDFStandby server

[root@nyip2 outbound]# /nas/sbin/nas\_mview -init nyip1

Celerra with MirrorView/Synchronous Disaster Recovery

**Initializing nyip1 --> nyip2**

Contacting nyip1 for remote storage info

Local storage system: APM00030600872

Remote storage system: APM00023700172

Discovering storage on nyip1 (may take several minutes)

Setting security information for APM00030600872

Discovering storage APM00023700172 (may take several minutes)

Discovering storage (may take several minutes)

Contacting nyip1 for remote storage info

Gathering server information...

Contacting nyip1 for server capabilities...

Analyzing server information...

Source servers available to be configured for remote DR

---

1. server\_2:nyip1

2. server\_3:nyip1

v. Verify standby server configuration

q. Quit initialization process

c. Continue initialization

Select a nyip1 server: 1

Destination servers available to act as remote standby

---

1. server\_2:nyip2

2. server\_3:nyip2

b. Back

Select a nyip2 server: 1

Source servers available to be configured for remote DR

---

1. server\_2:nyip1 [ remote standby is server\_2:nyip2 ]

2. server\_3:nyip1

v. Verify standby server configuration

q. Quit initialization process

c. Continue initialization

Select a nyip1 server: c

Standby configuration validated OK

Enter user information for managing remote site nyip1

Username: mvadmin

Password: \*\*\*\*\* Retype your response to validate

Password: \*\*\*\*\*

Active/Active configuration

**Initializing (nyip1-->nyip2)**

Do you wish to continue? [yes or no] yes

Updating MirrorView configuration cache

Setting up server\_2 on nyip1

Rebooting server\_2 on nyip2 as standby ... done

Creating user account mvadmin

Setting acl for server\_2 on nyip2

Updating the Celerra domain information

Creating device group AP\_MView on nyip2

Creating device group AP\_MView on nyip1

done

## **5. Verifying MirrorView Configuration on Remote Side:**

[root@nyip2 outbound]# tail /etc/passwd

mvadmin:x:500:201::/home/mvadmin:/bin/bash

**Note:** All MirrorView -init, -activate, and -restore commands are run on the Remote Celerra, and the DR Account “mvadmin” is created on the Remote Control Station

[root@nyip2 outbound]# /nas/bin/nas\_devicegroup -list

```
ID name          owner storage ID  acl type
1 AP_MView      500  APM00030600872 0  MVIEW
```

[root@nyip2 outbound]# /nas/bin/nas\_devicegroup -info id=1

Sync with CLARIION backend ..... done

```
name        = AP_MView
description = Active Passive MirrorView Consistency Group
uid        = 50:6:1:60:90:60:7:BB:0:0:0:0:0:0:0
state       = Consistent
role        = Secondary
condition   = Active
recovery policy = Automatic
number of mirrors = 9
mode        = SYNC
owner       = 500
mirrored disks =
local clarid = APM00030600872
remote clarid = APM00023700172
mirror direction = local <- remote
```

[root@nyip2 outbound]# /nas/sbin/nas\_mview -info

\*\*\*\*\* Device Group Configuration \*\*\*\*\*

No device group configured

\*\*\*\*\* Servers configured with RDFstandby \*\*\*\*\*

No servers configured with RDFstandby

\*\*\*\*\* Servers configured as standby \*\*\*\*\*

```
id      = 1
name    = server_2
acl     = 2000, owner=mvadmin, ID=500
type    = standby
slot    = 2
member_of =
standbyfor=
status   :
defined = enabled
actual  = online, ready
```

[root@nyip2 outbound]# nas\_server -l

| id | type | acl         | slot | groupID | state    | name |
|----|------|-------------|------|---------|----------|------|
| 1  | 4    | <b>2000</b> | 2    | 0       | server_2 |      |
| 2  | 1    | 1000        | 3    | 0       | server_3 |      |

Note: Remote Server\_2 has been assigned an ACL of 2000, indicating that it is an RDFStandby for the other side

[root@nyip2 outbound]# nas\_server -a -i

```
id      = 1
name    = server_2
acl     = 2000, owner=mvadmin, ID=500
type    = standby
slot    = 2
member_of =
standbyfor=
status   :
defined = enabled
actual  = online, ready
id      = 2
name    = server_3
acl     = 1000, owner=nasadmin, ID=201
type    = nas
slot    = 3
member_of =
standby  =
status   :
```

defined = enabled

actual = online, ready

**[root@nyip2 outbound]# nas\_pool -list**

| id | inuse | acl | name                |  |  |
|----|-------|-----|---------------------|--|--|
| 3  | n     | 0   | clar_r5_performance |  |  |
| 13 | n     | 0   | cm_r5_performance   |  |  |

**[root@nyip2 outbound]# nas\_disk -list**

| id | inuse | sizeMB | storageID-devID     | type  | name       | servers |
|----|-------|--------|---------------------|-------|------------|---------|
| 1  | y     | 11263  | APM00030600872-001D | CMSTD | root_disk  | 1,2     |
| 2  | y     | 11263  | APM00030600872-001E | CMSTD | root_ldisk | 1,2     |
| 3  | y     | 2047   | APM00030600872-001F | CLSTD | d3         | 1,2     |
| 4  | y     | 2047   | APM00030600872-0021 | CLSTD | d4         | 1,2     |
| 5  | y     | 2047   | APM00030600872-0022 | CMSTD | d5         | 1,2     |
| 6  | y     | 2047   | APM00030600872-0023 | CLSTD | d6         | 1,2     |
| 7  | n     | 5119   | APM00030600872-0011 | CMSTD | d7         | 1,2     |
| 8  | n     | 5119   | APM00030600872-0013 | CMSTD | d8         | 1,2     |
| 9  | n     | 5119   | APM00030600872-0012 | CMSTD | d9         | 1,2     |
| 10 | n     | 5119   | APM00030600872-001C | CMSTD | d10        | 1,2     |

**[root@nyip2 server]# cat /nas/server/servers**

1:server\_2:2000:4:2::0:2:

2:server\_3:1000:1:3::0:3:

**[root@nyip2 server]# cat /nas/server/server\_setup**

**[root@nyip2 rdf]# pwd;ls -la**

/nbsnas/rdf

```
drwxr-xr-x 2 root root 4096 Jul 13 15:30 500 →this is the directory where /dev/ndj1 is mounted after failing over to Remote
-rw-r--r-- 1 root root 0 Jul 11 14:02 vfstab
```

**[nasadmin@nyip2 nasadmin]\$ cat /nas/site/cshosts**

1:nyip1:500:10.241.168.50:/dev/ndj1:/dev/ndg:APM000237001720000:

0:nyip2.localdomain:0:10.241.168.48:::APM00030600872001D:

## **6. Verify MirrorView Configuration on Local Side:**

**[root@nyip1 site]# nas\_devicegroup -list**

| ID | name     | owner | storage ID     | acl | type  |
|----|----------|-------|----------------|-----|-------|
| 2  | AP_MView | 0     | APM00023700172 | 0   | MVIEW |

**[root@nyip1 site]# nas\_devicegroup -info id=2**

Sync with CLARiiON backend ..... done

name = AP\_MView  
description = Active Passive MirrorView Consistency Group  
uid = 50:6:1:60:80:60:5:35:0:0:0:0:0:0:  
state = Consistent  
role = Primary  
condition = Active  
recovery policy = Automatic  
number of mirrors = 9  
mode = SYNC  
owner = 0  
mirrored disks = root\_disk,root\_ldisk,d5,d7,d8,d9,d10,d11,d12,  
local clarid = APM00023700172  
remote clarid = APM00030600872  
mirror direction = local -> remote

**[root@nyip1 site]# /nas/sbin/nas\_mview -info**

\*\*\*\*\* Device Group Configuration \*\*\*\*\*

name = AP\_MView  
description = Active Passive MirrorView Consistency Group  
uid = 50:6:1:60:80:60:5:35:0:0:0:0:0:0:  
state = Consistent  
role = Primary  
condition = Active  
recovery policy = Automatic

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!

```
number of mirrors = 9
mode      = SYNC
owner     = 0
mirrored disks = root_disk,root_ldisk,d5,d7,d8,d9,d10,d11,d12,
local clarid = APM00023700172
remote clarid = APM00030600872
mirror direction = local -> remote
***** Servers configured with RDFstandby *****
id      = 1
name    = server_2
acl     = 1000, owner=nasadmin, ID=201
type    = nas
slot    = 2
member_of =
standby =
RDFstandby= slot=2
status   :
defined = enabled
actual = online, ready
***** Servers configured as standby *****
No servers configured as standby
```

[root@nyip1 site]# nas\_server -l

```
id  type acl slot groupID state name
1   1   1000 2       0   server_2
2   1   1000 3       0   server_3
```

[root@nyip1 site]# nas\_server -a -i

```
id      = 1
name    = server_2
acl     = 1000, owner=nasadmin, ID=201
type    = nas
slot    = 2
member_of =
standby =
RDFstandby= slot=2
status   :
defined = enabled
actual = online, ready
id      = 2
name    = server_3
acl     = 1000, owner=nasadmin, ID=201
type    = nas
slot    = 3
member_of =
standby =
status   :
defined = enabled
actual = online, ready
```

[root@nyip1 site]# cat /nas/server/servers

```
1:server_2:1000:1:2::0:2:
2:server_3:1000:1:3::0:3:
```

[root@nyip1 site]# cat /nas/server/server\_setup

```
1:0:1:1:2,0,0,:;
```

**Primary\_id: Backup\_id: Component: Policy: Slot#, ACL, State:**

**Note:** This file is all important and defines the Local/Remote local standby or RDF standby relationships between Servers

[root@nyip1 site]# nas\_pool -list

```
id  inuse acl  name
3   n    0    clar_r5_performance
13  y    0    cm_r5_performance
```

[root@nyip1 site]# nas\_disk -l

```

id inuse sizeMB storageID-devID type name servers
1 y 11263 APM00023700172-0000 CMSTD root_disk 1,2
2 y 11263 APM00023700172-0001 CMSTD root_ldisk 1,2
3 y 2047 APM00023700172-0002 CLSTD d3 1,2
4 y 2047 APM00023700172-0003 CLSTD d4 1,2
5 y 2047 APM00023700172-0004 CMSTD d5 1,2
6 y 2047 APM00023700172-0005 CLSTD d6 1,2
7 y 10239 APM00023700172-005B CMSTD d7 1,2
8 y 10239 APM00023700172-005D CMSTD d8 1,2
9 y 10239 APM00023700172-005F CMSTD d9 1,2
10 y 10239 APM00023700172-005A CMSTD d10 1,2
11 y 10239 APM00023700172-005C CMSTD d11 1,2
12 y 10239 APM00023700172-005E CMSTD d12 1,2

```

[root@nyip1 rdf]# pwd;ls -la

/nbsnas/rdf

```
-rwxrwxr-x 1 nasadmin nasadmin 0 Apr 5 12:59 vfstab
```

[root@nyip1 site]# cat /nas/site/cshosts

```
2:nyip2:0:10.241.168.48:::APM00030600872001D:0:
```

```
0:nyip1:0:10.241.168.50:::APM000237001720000:
```

#### **7. Create Dummy User Account on Local Control Station using Following procedure:**

**Note:** For an Active/Active MirrorView/S configuration, each DR Admin account should have a unique UID. This ensures that the correct configuration is displayed when logged into either Control Station as the DR Admin account.

##### **a. Remote CS UID for MirrorView is 500**

```
[nasadmin@nyip2 slot_2]$ tail -1 /etc/passwd
mvadmin:x:500:201::/home/mvadmin:/bin/bash
```

##### **b. Create dummy account on Source side**

[root@nyip1 slot\_2]# /usr/sbin/useradd dradmin\_ghost -u 500

[root@nyip1 slot\_2]# tail -1 /etc/passwd

```
dradmin_ghost:x:500:500::/home/dradmin_ghost:/bin/bash
```

**Note:** After –init of the other side of the Active Active MView setup, this is the UID that was assigned:

```
dradmin_ghost:x:500:500::/home/dradmin_ghost:/bin/bash → Dummy account, same UID as used on first Active MView setup
```

```
mvadmin:x:501:201::/home/mvadmin:/bin/bash → MVAdmin account on other Active MView Setup
```

#### **8. Run MirrorView Initialization on Remote Side for other half of the Active/Active Setup:**

[root@nyip1 /]# /nas/sbin/nas\_mview -init nyip2

Celerra with MirrorView/Synchronous Disaster Recovery

Initializing nyip2 --> nyip1

Contacting nyip2 for remote storage info

Local storage system: APM00023700172

Remote storage system: APM00030600872

Discovering storage on nyip2 (may take several minutes)

Setting security information for APM00023700172

Discovering storage APM00030600872 (may take several minutes)

Discovering storage (may take several minutes)

Contacting nyip2 for remote storage info

Gathering server information...

Contacting nyip2 for server capabilities...

Analyzing server information...

Source servers available to be configured for remote DR

server\_2:nyip2 [ is remote standby for server\_2:nyip1 ]

2. server\_3:nyip2

v. Verify standby server configuration

q. Quit initialization process

c. Continue initialization

Select a nyip2 server: 2

Destination servers available to act as remote standby

server\_2:nyip1 [ remote standby is server\_2:nyip2 ]

2. server\_3:nyip1

b. Back

Select a nyip1 server: 2

Source servers available to be configured for remote DR

-----  
server\_2:nyip2 [ is remote standby for server\_2:nyip1 ]  
2. server\_3:nyip2 [ remote standby is server\_3:nyip1 ]

v. Verify standby server configuration

q. Quit initialization process

c. Continue initialization

Select a nyip2 server: c

Standby configuration validated OK

Enter user information for managing remote site nyip2

Username: mvadmin

Password: \*\*\*\*\*\* Retype your response to validate

Password: \*\*\*\*\*\*

#### Active/Active configuration

#### Initializing (nyip2-->nyip1)

Do you wish to continue? [yes or no] yes

Updating MirrorView configuration cache

Setting up server\_3 on nyip2

Rebooting server\_3 on nyip1 as standby ... done

Creating user account mvadmin

Setting acl for server\_3 on nyip1

Updating the Celerra domain information

#### Creating device group 872\_MViewCGroup on nyip1

Creating device group 872\_MViewCGroup on nyip2

done

### **MIRRORVIEW CONFIGURATION OUTPUT FOR ACTIVE ACTIVE SETUP:**

[root@nyip1 log]# cat /nas/server/server\_setup

1:0:1:1:2,0,0,:

[root@nyip1 log]# nas\_server -l

| id | type | acl  | slot | groupID | state    | name |
|----|------|------|------|---------|----------|------|
| 1  | 1    | 1000 | 2    | 0       | server_2 |      |
| 2  | 4    | 2000 | 3    | 0       | server_3 |      |

[root@nyip1 log]# nas\_cel -l

| id | name  | owner | mount_dev | channel       | net_path           | CMU                |
|----|-------|-------|-----------|---------------|--------------------|--------------------|
| 0  | nyip1 | 0     |           | 10.241.168.50 | APM000237001720000 |                    |
| 2  | nyip2 | 501   | /dev/ndj1 | /dev/ndg      | 10.241.168.48      | APM00030600872001D |

[root@nyip1 log]# nas\_devicegroup -list

| ID | name            | owner | storage ID     | acl | type  |
|----|-----------------|-------|----------------|-----|-------|
| 2  | AP_MView        | 0     | APM00023700172 | 0   | MVIEW |
| 3  | 872_MViewCGroup | 501   | APM00023700172 | 0   | MVIEW |

[root@nyip1 log]# /nas/sbin/nas\_mview -info id=2

\*\*\*\*\* Device Group Configuration \*\*\*\*\*

|                   |                                                 |
|-------------------|-------------------------------------------------|
| name              | = AP_MView                                      |
| description       | = Active Passive MirrorView Consistency Group   |
| uid               | = 50:6:1:60:80:60:5:35:0:0:0:0:0:0:0:0          |
| state             | = Consistent                                    |
| role              | = Primary                                       |
| condition         | = Active                                        |
| recovery policy   | = Automatic                                     |
| number of mirrors | = 9                                             |
| mode              | = SYNC                                          |
| owner             | = 0                                             |
| mirrored disks    | = root_disk,root_ldisk,d5,d7,d8,d9,d10,d11,d12, |
| local clarid      | = APM00023700172                                |
| remote clarid     | = APM00030600872                                |
| mirror direction  | = local -> remote                               |

\*\*\*\*\* Servers configured with RDFstandby \*\*\*\*\*

```

id      = 1
name    = server_2
acl     = 1000, owner=nasadmin, ID=201
type    = nas
slot    = 2
member_of =
standby =
RDFstandby= slot=2
status   :
defined = enabled
actual  = online, ready
***** Servers configured as standby *****
id      = 2
name    = server_3
acl     = 2000, owner=mvadmin, ID=501
type    = standby
slot    = 3
member_of =
standbyfor=
status   :
defined = enabled
actual  = online, ready

```

[root@nyip2 nasadmin]# cat /nas/server/server\_setup  
2:0:1:1:3,0,0,:;

[root@nyip2 nasadmin]# nas\_server -l

| id | type | acl  | slot | groupID | state    | name |
|----|------|------|------|---------|----------|------|
| 1  | 4    | 2000 | 2    | 0       | server_2 |      |
| 2  | 1    | 1000 | 3    | 0       | server_3 |      |

[root@nyip2 nasadmin]# nas\_cel -l

| id | name  | owner | mount_dev | channel       | net_path           | CMU                |
|----|-------|-------|-----------|---------------|--------------------|--------------------|
| 0  | nyip2 | 0     |           | 10.241.168.48 | APM00030600872001D |                    |
| 1  | nyip1 | 500   | /dev/ndj1 | /dev/ndg      | 10.241.168.50      | APM000237001720000 |

[root@nyip2 nasadmin]# nas\_devicegroup -list

| ID | name            | owner | storage        | ID | acl   | type |
|----|-----------------|-------|----------------|----|-------|------|
| 1  | AP_MView        | 500   | APM00030600872 | 0  | MVIEW |      |
| 2  | 872_MViewCGroup | 0     | APM00030600872 | 0  | MVIEW |      |

[root@nyip2 nasadmin]# /nas/sbin/nas\_mview -info id=2

\*\*\*\*\* Device Group Configuration \*\*\*\*\*

|                   |                                         |
|-------------------|-----------------------------------------|
| name              | = 872_MViewCGroup                       |
| description       | =                                       |
| uid               | = 50:6:1:60:90:60:7:BB:1:0:0:0:0:0:0:0  |
| state             | = Consistent                            |
| role              | = Primary                               |
| condition         | = Active                                |
| recovery policy   | = Automatic                             |
| number of mirrors | = 7                                     |
| mode              | = SYNC                                  |
| owner             | = 0                                     |
| mirrored disks    | = root_disk,root_ldisk,d5,d7,d8,d9,d10, |
| local clarid      | = APM00030600872                        |
| remote clarid     | = APM00023700172                        |
| mirror direction  | = local -> remote                       |

\*\*\*\*\* Servers configured with RDFstandby \*\*\*\*\*

```

id      = 2
name    = server_3
acl     = 1000, owner=nasadmin, ID=201
type    = nas
slot    = 3
member_of =
standby =

```

```
RDFstandby= slot=3
status :
defined = enabled
actual = online, ready
***** Servers configured as standby *****
id      = 1
name    = server_2
acl     = 2000, owner=mvadmin, ID=500
type    = standby
slot    = 2
member_of =
standbyfor=
status :
defined = enabled
actual = online, ready
```

### **III. CONDUCTING MIRRORVIEW FAILOVER, RESTORE, SUSPEND OPERATIONS:**

#### **ACTIVATING MIRRORVIEW DR FAILOVER ACTIVE/PASSIVE:**

##### **OVERVIEW**

1. Run –init from Target side after any backend changes
2. Run –activate or –restore operations as DR account/root
3. Run nas\_devicegroup –suspend or –resume when needing to stop or restart Devicegroup Mirroring (maintenance purposes, etc.)

##### **Testing Graceful DR Failover--activate:**

- Ensure MirrorView links and IP connections are up between Control Stations
- Verify devicegroup is Active & Synchronized or Consistent using nas\_devicegroup –info
- Target Servers should be plugged into same Network as Source Servers
- Target side needs to support DNS, Usermapper, & any other CIFS environmentals

##### **MirrorView Failover:**

- Use when disaster on Source side, or maintenance to be done, or to test DR
- Log in as “dradmin” account on Destination Control Station, then root to initiate failover or restore

##### **What happens during a MirrorView Activate (Failover):**

###### **#/nas/sbin/nas\_mvview –activate**

**Note:** During actual –activate, backends are sync'd, User prompted for Source shutdown, Source CS & DM's are shutdown, storage devices are failed over to Secondary Image & promoted to Primary status RW, Servers NBS environment reconfigured (cs.nbs.rw for RDF7 & RDF10), Target environment activated, Servers taken offline and failover completed as replace procedure is done. If the actual CLARiiON backend remains connected and available during the failover, changes will continue to be sync back to the original Source, making it easier to run the restore process later

- a. Destination CS shuts down Source CS & DM's gracefully, if possible
- b. Mirrored devices are failed over to their respective Secondary Images, then promoted as RW Primary images, becoming visible to Destination RDF Servers
- c. Primary Image Source LUNs are taken offline and no longer available on Source side(if source mirrors still available, synchronization will continue in reverse direction)
- d. The remote LUN 9 NAS\_DB is mounted RW on the destination Control Station (mirrored copy of Source LUN 4 NASDB). Configuration information is read from /nas/rdf/500/server/slot\_x to the local RDF Standbys using build\_config, and a compiled boot.cfg is copied to the local /nas/dos/slot\_x directory since the Data Mover must boot from a local LUN 0.
- e. After replaying the NASDB configuration information to the RDF Standbys, the Servers are rebooted and come up as the DR replacement Server for the original Source system

**Note:** During –activate, build\_config produces a new boot.cfg file based on the contents of LUN 9 (NBSNAS) and is copied to local DOS LUN 0 so that if a destination server is rebooted it will be able to boot. This file is copied to the local CS LUN 0, or /nas/dos/slot\_2 directory, NOT the /nas/rdf/500 directory. After failover is completed, there are (2) NAS databases on the Destination side; (1) for nasadmin for any local storage, & (1) DB for DR Admin for MirrorView management

##### **AMPLIFYING NOTE ON CONTROL LUNS:**

It may be useful to remember that during normal ‘Steady-state’ operation, Source LUNs 0, 1, 4 [DOS, ufslog, NBSNAS] are mirrored to Target LUNs 6, 7, & 9, but that after failover, only LUNs 7 & 9 remain mounted, taking on the role of ufslog and NBSNAS, respectively.

With DOS, during failover a new boot.cfg file is created from NBSNAS (LUN 9) using build\_config, and then copied to the local system's /nas/dos partition, which is LUN 0—this anomaly occurs because the Server must boot from a LUN 0. Hence, after failover is completed, LUN 6 is not actively used. In any case, no NASDB changes should be made while in a failed over state.

\$ ls –la /nas/rdf/500 → DOS is a link to local /nas/dos

lrwxrwxrwx 1 root root 8 Jul 19 10:23 dos -&gt; /nas/dos

**CONDUCTING MIRRORVIEW FAILOVER:****1. On Secondary Celerra, login with mvadmin account, then root user and initiate failover:****\$su – mvadmin****Note:** Once logged in as mvadmin, su to root user**#su**

&lt;root passwd&gt;

**#/nas/sbin/nas\_mview –activate**

Sync with CLARiiON backend ..... done

Is source site ns1 ready for complete shut down (power OFF)? [yes or no] yes

Shutting down remote site ns1 ..... done

Sync with CLARiiON backend ..... done

STARTING an MV ‘FAILOVER’ operation

Device group: MirrorView-DR ..... done

The MV ‘FAILOVER’ operation SUCCEEDED.

Failing over Devices ... done

Reconfiguring NBS access for server\_2 ..... done

Reconfiguring NBS access for server\_3 ..... done

Activating the target environment ... done

server\_2 : going offline

rdf : going active

replace in progress ... done

failover activity complete

server\_3 : going offline

rdf : going active

replace in progress ... done

failover activity complete

commit in progress (not interruptible) ...done

commit in progress (not interruptible) ...done

**Note:** If the failover hangs, and you've determined that there isn't anything wrong with the Destination Clariion or Celerra, clears the locks in /nas/lock/db and rerun the –activate command. Additionally, may need to reboot DM before Shares will be accessible.**2. Verify devicegroup status after Failover:****# nas\_devicegroup –list | –info <id=>****# nas\_devicegroup –list**ID name owner storage ID acl type  
1 SVT MirrorView DR 50001 APM00030600872 0 MVIEW**# nas\_fs -l****# nas\_devicegroup –info id=1**Sync with CLARiiON backend ..... done  
name = SVT MirrorView DR  
description =  
uid = 50:6:1:60:90:60:7:BB:0:0:0:0:0:0:0  
state = Consistent  
role = Primary  
condition = Active  
recovery policy = Automatic  
number of mirrors = 9  
mode = SYNC  
owner = 50001  
mirrored disks = root\_disk,root\_ldisk,d5,d7,d8,d9,d10,d11,d12,  
local clarid = APM00030600872  
remote clarid = APM00023700172  
mirror direction = local -> remote**Note:** Navisphere should show that Primary side luns are now ‘Secondary Copy’ and Secondary luns are ‘Mirrored’**# nas\_disk –l**id inuse sizeMB storageID-devID type name servers  
1 y 11263 APM00030600872-0006 CMSTD root\_disk 1,2  
2 y 11263 APM00030600872-0007 CMSTD root\_ldisk 1,2

```

3 y 2047 APM00023700172-0002 CLSTD d3 1,2
4 y 2047 APM00023700172-0003 CLSTD d4 1,2
5 y 2047 APM00030600872-0009 CMSTD d5 1,2
6 y 2047 APM00023700172-0005 CLSTD d6 1,2
7 y 10239 APM00030600872-005B CMSTD d7 1,2
8 y 10239 APM00030600872-005D CMSTD d8 1,2
9 y 10239 APM00030600872-005F CMSTD d9 1,2
10 y 10239 APM00030600872-005A CMSTD d10 1,2
11 y 10239 APM00030600872-005C CMSTD d11 1,2
12 y 10239 APM00030600872-005E CMSTD d12 1,2

```

**Note:** MirrorView devices will be seen as CMSTD after failover to Secondary side and Control & Data Luns should be visible

### Excerpts from /nas/log/dr\_log.al for –activate failover:

```

2006-09-06 07:13:14.754 db:0:4455: Thread id: 0x917BAA8 Start
2006-09-06 07:13:15.039 db:0:4455: /nas/sbin/nas_mview -activate -nasuid 500 -preop
2006-09-06 07:13:15.185 db:0:4455: dr session created
2006-09-06 07:13:50.641 db:0:4455: validating mg config data
2006-09-06 07:29:20.342 db:0:4455: Running: /nasmcd/nas_mcd -e 10.241.168.50 haltstatus 2>&1
2006-09-06 07:29:20.873 db:0:4455: exit code=1 Error: Operation not permitted →Error of no concern—loop status
2006-09-06 07:29:20.875 db:0:4455: shutting down remote site
2006-09-06 07:31:38.443 db:0:4455: starting activate target
2006-09-06 07:31:38.444 db:0:4455: starting device activate
2006-09-06 07:33:02.527 db:0:4455: failover complete for dev grp: 172_AP_MView (Consistency Group)
2006-09-06 07:33:02.528 db:0:4455: activating env on target site
2006-09-06 07:33:02.548 db:0:4455: configuring R2 devices
2006-09-06 07:33:02.599 db:0:4455: set nbs access nbsid=7 dart=server_2
2006-09-06 07:33:07.950 db:0:4455: set nbs access nbsid=10 dart=server_2
2006-09-06 07:33:08.457 db:0:4455: platform svc: /nasmcd/sbin/dr_helper -nbsadd server_2 10 /dev/ndj1 2>&1
2006-09-06 07:33:30.097 db:0:4455: MV device group 172_AP_MView: failed over
2006-09-06 07:33:30.127 db:0:4455: umounting dos: /dev/ndj1
2006-09-06 07:33:30.160 db:0:4455: umounting nas: /dev/ndj1
2006-09-06 07:33:30.185 db:0:4455: /sbin/fsck -t ext3 -y -f /dev/ndj1 (routine)
2006-09-06 07:33:38.868 db:0:4455: adding nas entry /dev/ndj1 /nas/rdf/500 ext3rw,sync to /nas/rdf/vfstab
2006-09-06 07:33:38.910 db:0:4455: mounting nas: /dev/ndj1
2006-09-06 07:33:38.982 db:0:4455: rw mounting dos
2006-09-06 07:33:38.984 db:0:4455: platform svc: /bin/mount -t msdos -o rw /dev/ndg1 /nas/rdf/500/dos 2>&1
2006-09-06 07:33:41.247 db:0:15353: Thread id: 0x925E4E0 Start
2006-09-06 07:33:41.535 db:0:15353: /nas/sbin/nas_mview -activate -nasuid 500 -postop
2006-09-06 07:33:41.934 db:0:15353: starting server activate
2006-09-06 07:33:42.495 db:0:15353: activating nas db
2006-09-06 07:33:42.497 db:0:15353: backup src servers
2006-09-06 07:33:42.880 db:0:15353: copy_all_files: /bin/cp -d /nas/rdf/500/server/server_1/* /nas/rdf/500/server/server_1/rdf 2>&1
2006-09-06 07:33:44.363 db:0:15353: umounting dos: /nas/rdf/500/dos
2006-09-06 07:33:44.394 db:0:15353: prepare gatekeepers
2006-09-06 07:33:44.406 db:0:15353: copy_all_files: /bin/cp -d /nas/rdf/500/dev///* /nas/rdf/500/dev/rdf 2>&1
2006-09-06 07:33:44.602 db:0:15353: Running: /nas/rdf/500/sbin/build_config -local /nas/server/slot_2
/nas/rdf/500/server/server_1 2>&1
2006-09-06 07:34:16.081 db:0:15353: /nas/sbin/nas_mview -activate -nasuid 500 -postop... end

```

### # df -h

| Filesystem       | Size        | Used        | Avail        | Use%       | Mounted on            |
|------------------|-------------|-------------|--------------|------------|-----------------------|
| /dev/hda3        | 2.0G        | 826M        | 1.0G         | 44%        | /                     |
| /dev/hda1        | 30M         | 2.7M        | 26M          | 10%        | /boot                 |
| none             | 251M        | 0           | 251M         | 0%         | /dev/shm              |
| /dev/nde1        | 1.7G        | 669M        | 1016M        | 40%        | /nbsnas               |
| /dev/nda1        | 133M        | 36M         | 97M          | 27%        | /nas/dos              |
| /dev/ndf1        | 1.7G        | 86M         | 1.5G         | 6%         | /nas/var              |
| /dev/hda5        | 2.0G        | 560M        | 1.3G         | 30%        | /nas                  |
| <b>/dev/ndj1</b> | <b>1.7G</b> | <b>664M</b> | <b>1021M</b> | <b>40%</b> | <b>/nas/rdf/50001</b> |

### # cat /nas/rdf/vfstab

```
/dev/ndj1 /nas/rdf/500 ext3 rw,sync
```

[root@nyip2 server]# cat server\_setup

1:0:1:1:2,1000,2,: →Server\_2:no standby: n/a: manual failover: slot\_2 on destination ACL 1000 State 2

[root@nyip2 slot\_2]# pwd

/nas/rdf/500/server/slot\_2

[root@nyip2 slot\_2]# ls -la nbs\*

lrwxrwxrwx 1 root root 30 Aug 29 12:51 nbs.cs -> /nas/server/server\_1/nbs.cs.rw

-rwxrwxr-x 1 nasadmin nasadmin 1019 Aug 29 12:51 nbs.cs.rw

[root@nyip2 slot\_2]# cat nbs.cs.rw

volume disk NBS1 c0t0l0

volume disk NBS5 c0t0l4

volume disk NBS6 c0t0l5

volume disk RDF7 c0t0l6

-----abridged-----

nbs add nbsid=1 vol=NBS1 rw=192.168.1.100:192.168.1.101:192.168.2.100:192.168.2.101 exclusive raw share

nbs add nbsid=5 vol=NBS5 rw=192.168.1.100:192.168.1.101:192.168.2.100:192.168.2.101 exclusive raw share

nbs add nbsid=6 vol=NBS6 rw=192.168.1.100:192.168.1.101:192.168.2.100:192.168.2.101 exclusive raw share

nbs add nbsid=7 vol=RDF7 exclusive raw rw=192.168.1.100:192.168.1.101:192.168.2.100:192.168.2.101  
100:192.168.2.101 exclusive raw rw=192.168.1.100:192.168.1.101:192.168.2.100:192.168.2.101

[root@nyip2 slot\_2]# cat boot.cfg

volume disk 1 c0t0l6 disk\_id=1 size=11263

volume disk 2 c0t0l7 disk\_id=2 size=11263

volume disk 90 c16t5l14 disk\_id=12 size=10239

volume disk 89 c16t5l12 disk\_id=11 size=10239

volume disk 88 c16t5l10 disk\_id=10 size=10239

volume disk 87 c0t5l15 disk\_id=9 size=10239

volume disk 86 c0t5l13 disk\_id=8 size=10239

volume disk 85 c0t5l11 disk\_id=7 size=10239

volume disk 2 c0t0l7 disk\_id=2 size=11263

#### How Server\_2 on Destination now sees Secondary LUN Mirrors:

[root@nyip2 slot\_2]# server\_devconfig server\_2 -p -s -a |grep -v no

server\_2 :

SCSI devices :

chain= 16, scsi-16

stor\_id= APM00030600872 celerra\_id= APM00030600872001D

-----abridged-----

tid/lun= 0/6 type= disk sz= 0 val= -5 info= DGC RAID 5 02190600060006NI

tid/lun= 0/7 type= disk sz= 0 val= -5 info= DGC RAID 5 02190700070007NI

tid/lun= 0/9 type= disk sz= 0 val= -5 info= DGC RAID 5 02190900090009NI

chain= 48, scsi-48

stor\_id= APM00030600872 celerra\_id= APM00030600872001D

-----abridged-----

tid/lun= 0/6 type= disk sz= 11263 val= 1 info= DGC RAID 5 02190600060006NI diskerr= mismatch:tid

tid/lun= 0/7 type= disk sz= 11263 val= 2 info= DGC RAID 5 02190700070007NI diskerr= mismatch:tid

tid/lun= 0/9 type= disk sz= 2047 val= 5 info= DGC RAID 5 02190900090009NI diskerr= mismatch:tid

-----abridged-----

tid/lun= 5/11 type= disk sz= 10239 val= 7 info= DGC RAID 5 02195B005B005BNI diskerr= mismatch:tid

[root@nyip2 server]# nas\_disk -l (Example of Mirrored Control Luns & one Data LUN)

| id | inuse | sizeMB | storageID-devID     | type  | name       | servers |
|----|-------|--------|---------------------|-------|------------|---------|
| 1  | y     | 11263  | APM00030600872-0006 | CMSTD | root_disk  | 1,2     |
| 2  | y     | 11263  | APM00030600872-0007 | CMSTD | root_ldisk | 1,2     |
| 5  | y     | 2047   | APM00030600872-0009 | CMSTD | d5         | 1,2     |
| 7  | y     | 10239  | APM00030600872-005B | CMSTD | d7         | 1,2     |

| id | inuse | sizeMB | storageID-devID     | type  | name       | servers |
|----|-------|--------|---------------------|-------|------------|---------|
| 1  | y     | 11263  | APM00030600872-0006 | CMSTD | root_disk  | 1,2     |
| 2  | y     | 11263  | APM00030600872-0007 | CMSTD | root_ldisk | 1,2     |
| 5  | y     | 2047   | APM00030600872-0009 | CMSTD | d5         | 1,2     |
| 7  | y     | 10239  | APM00030600872-005B | CMSTD | d7         | 1,2     |

| # | /nas/sbin/nas_mview -info |
|---|---------------------------|
|---|---------------------------|

\*\*\*\*\* Device Group Configuration \*\*\*\*\*

name = SVT MirrorView DR

description =

uid = 50:6:1:60:90:60:7:BB:0:0:0:0:0:0:0

state = Consistent

role = Primary

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!

```

condition      = Active
recovery policy = Automatic
number of mirrors = 9
mode          = SYNC
owner         = 50001
mirrored disks = root_disk,root_ldisk,d5,d7,d8,d9,d10,d11,d12,
local clarid   = APM00030600872
remote clarid  = APM00023700172
mirror direction = local -> remote
***** Servers configured with RDFstandby *****
id      = 1
name    = server_2
acl     = 0
type    = nas
slot    = 2
member_of =
standby = server_3, policy=auto
RDFstandby= slot=2
status   :
defined = enabled
actual = online, active
id      = 2
name    = server_3
acl     = 0
type    = standby
slot    = 3
member_of =
standbyfor= server_2
RDFstandby= slot=3
status   :
defined = enabled
actual = online, ready
***** Servers configured as standby *****
No servers configured as standby

```

## **OVERVIEW OF MIRRORVIEW RESTORE:**

Backends sync'ed, mirror configuration validated, prompted for Source side shutdown, prompted for network restoration, destination Data Movers rebooted as RDF Standbys, backend sync'ed again, MV Resume is performed to make LUNs consistent & devicegroup updated, NBS access shifted back to cs.nbs.ro on failed over side, Failback operation completes, and original Source Data Movers are brought back online with NASDB

## **CONDUCTING MIRRORVIEW RESTORE:**

### **1. Login as mvadmin account and conduct –restore operation:**

**\$ su – mvadmin (login as mvadmin)**

**# su (login as root)**

**# nas\_storage –sync –all** (run prior to restore to sync up storage db with Celerra)

**# /nas/sbin/nas\_mview –restore** (Initiate restore on Destination side)

Sync with CLARiiON backend ..... done

Validating mirror group configuration ..... done

Contacting source site nyip1, please wait... done

Running restore requires shutting down source site nyip1.

Do you wish to continue? [yes or no] yes

Shutting down remote site nyip1 ..... done

Is source site nyip1 ready for storage restoration ? [yes or no] yes

Sync with CLARiiON backend ..... done

STARTING an MV 'RESUME' operation.

Device group: SVT MirrorView DR ..... done

The MV 'RESUME' operation SUCCEEDED.

Synchronized: 100% (\*\*\*\*\*)

**Note:** Once devices become consistent, IO access is stopped

Updating device group ... done  
Is source site ready for network restoration ? [yes or no] yes  
Restoring servers ..... done  
Waiting for servers to reboot ..... done  
Removing NBS access for server\_2 .. done  
Removing NBS access for server\_3 .. done  
Waiting for device group ready to failback ... done  
Sync with CLARiiON backend ..... done  
STARTING an MV 'FAILBACK' operation.  
Device group: SVT MirrorView DR ..... done  
The MV 'FAILBACK' operation SUCCEEDED.  
Restoring remote site nyip1, please wait... done (this action can take 10 minutes or more to complete)  
done

**Note:** For Warning 5026....Please run restore on primary side

**2. If restore fails, complete by running on Source side as local nasadmin/root User and specifying /nasmcd/sbin/nas\_mview -restore:**

# **/nasmcd/sbin/nas\_mview -restore** (do not use /nas/sbin/nas\_mview for restore)  
Powering on servers ( please wait ) ..... done  
Sync with CLARiiON backend ..... done  
STARTING an MV 'SUSPEND' operation.  
Device group: SVT MirrorView DR ..... done  
The MV 'SUSPEND' operation SUCCEEDED.  
server\_2 : going standby  
rdf : going active  
replace in progress ...done  
failover activity complete  
server\_3 : going standby  
rdf : going active  
replace in progress ...done  
failover activity complete  
commit in progress (not interruptible)...done  
commit in progress (not interruptible)...done  
Sync with CLARiiON backend ..... done  
STARTING an MV 'RESUME' operation.  
Device group: SVT MirrorView DR ..... done  
The MV 'RESUME' operation SUCCEEDED.  
commit in progress (not interruptible)...done  
commit in progress (not interruptible)...done

**Note:** Reboots CS after the restore succeeds. Make sure to use the /nasmcd/sbin path and not the /nas/sbin path, or else NAS may not properly restart and a lock will be left on the Control Station: /nasmcd/lock/dr.lck. Clearing the lock file manually will result in an automatic reboot of the CS, which will also resolve the issue.

→Synchronizes source devices to 99%, then I/O shutoff to Servers, remote Servers rebooted as remote standbys, mirrored devices are failed back, source Servers restored

→/nas is synchronized with changes from /nas/rdf/500 side, and servers are rebooted and brought up on Source, NAS Services started

**3. Verify that failback was successful:**

# **nas\_devicegroup -list | -info <id= >**

# **df -h**

# **nas\_disk -l**

# **server\_mount server\_export server\_devconfig**

**Note:** From Navisphere, verify that Primary side luns say 'Mirrored' and Secondary luns say 'Secondary Copy'

**SUSPENDING MIRRORVIEW SYNCHRONIZATION ON DEVICEGROUP:**

Used to temp halt mirroring from Source to Destination, as for maintenance purposes, etc.

# **nas\_devicegroup -suspend "SVT MirrorView DR"**

Sync with CLARiiON backend ..... done  
STARTING an MV 'SUSPEND' operation.  
Device group: SVT MirrorView DR ..... done  
The MV 'SUSPEND' operation SUCCEEDED.  
Done

**Navisphere Display:**

On Source side, Remote Mirrors will display as “F” for faulted, and then “AdminFractured” for each Secondary Image

**Note:** Use to administratively fracture and halt operation of consistency group by halting the mirroring between Source & Destination—same net effect as in suspending the MirrorView link. If there are spaces in the Consistency Group name, must use “ “ as in above example. A –suspend must be followed by a –resume to restart mirroring & synchronizing destination LUNs with source LUNs. Both –suspend and –resume commands are run from the Active side.

**PERFORMING RESTART & RESYNC OF MIRROR PAIRS:**

When would this command be used?

1. After issuing manual Administrative fracture of devicegroups via -suspend
2. Recovering from a destination-side failure, link down, or SP failure, if the devicegroup recovery policy is set to manual
3. Documentation says this is only used on Source side, but later states that if this command is run on destination side after failing over to the destination side, then a full synchronization is performed to reconstruct devicegroups and mirrors (on source), i.e., after a disaster has occurred on the Source.

**# nas\_devicegroup -resume "SVT MirrorView DR"**

Sync with CLARIION backend ..... done  
 STARTING an MV 'RESUME' operation.  
 Device group: SVT MirrorView DR ..... done  
 The MV 'RESUME' operation SUCCEEDED.

Done

**Note:** Used to resume after a suspend of the Consistency Group. Or, if recovery policy of a MirrorView Consistency Group is set to manual and the mirror fractures, then the nas\_devicegroup –resume command would be needed to recover.

**/nas/log/cmd\_log or nas\_log.al**

2006-05-03 20:01:02.021 db:0:18228:S: nas\_devicegroup -suspend SVT MirrorView DR  
 2006-05-03 20:02:29.255 db:0:18228:E: nas\_devicegroup -suspend SVT MirrorView DR

**SYSTEM vs. ADMIN FRACTURE OF SECONDARY MIRROR IMAGE/DEVICEGROUP:**

A fracture is when the Primary Image on the Source can no longer access the Secondary Image on the Destination. A system fracture can occur when a single SP is down (or bad cable) on either the Source or Destination sides, or when the MirrorView link between the two systems is down. The Recovery Policy in place determines if the Primary will begin resynchronizing with the Secondary Image automatically, or manually. An Admin Fracture is performed using nas\_mvview –suspend to suspend the synchronization between Primary & Secondary Images.

**OUTPUT FROM RUNNING –INIT FROM DESTINATION SIDE:**

**Note:** Must run the init as the local nasadmin and root account, not the DR account

**# /nas/sbin/nas\_mvview -init id=1**

Celerra with MirrorView/Synchronous Disaster Recovery  
 Initializing nyip1 --> nyip2  
 Contacting nyip1 for remote storage info  
 Local storage system: APM00030600872  
 Remote storage system: APM00023700172  
 Discovering storage on nyip1 (may take several minutes)  
 Setting security information for APM00030600872  
 Discovering storage APM00023700172 (may take several minutes)  
 Discovering storage (may take several minutes)  
 Contacting nyip1 for remote storage info  
 Gathering server information...  
 Contacting nyip1 for server capabilities...  
 Analyzing server information...  
 Source servers available to be configured for remote DR

- 
1. server\_2:nyip1 [ remote standby is server\_2:nyip2 ]
  2. server\_3:nyip1 [ remote standby is server\_3:nyip2 ]
  - v. Verify standby server configuration
  - q. Quit initialization process
  - c. Continue initialization

Select a nyip1 server: c

Standby configuration validated OK  
 Using administrative user "mvadmin"  
 Initializing Active-->Passive (nyip1-->nyip2)  
 Do you wish to continue? [yes or no] yes

Updating MirrorView configuration cache

Setting up server\_3 on nyip1

Setting up server\_2 on nyip1

Setting acl for server\_3 on nyip2

Setting acl for server\_2 on nyip2

Creating device group SVT MirrorView DR on nyip1

done

## **WHAT DESTINATION SIDE OF ACTIVE/PASSIVE MIRRORVIEW LOOKS LIKE:**

**# /nas/sbin/nas\_mvinfo**

\*\*\*\*\* Device Group Configuration \*\*\*\*\*

No device group configured

\*\*\*\*\* Servers configured with RDFstandby \*\*\*\*\*

No servers configured with RDFstandby

\*\*\*\*\* Servers configured as standby \*\*\*\*\*

id = 1

name = server\_2

acl = 2000, owner=mvadmin, ID=50001

type = standby

slot = 2

member\_of =

standbyfor =

status :

defined = enabled

actual = online, ready

id = 2

name = server\_3

acl = 2000, owner=mvadmin, ID=50001

type = standby

slot = 3

member\_of =

standbyfor = server\_2

status :

defined = enabled

actual = online, ready

**# nas\_storage -l**

id acl name serial\_number

1 0 APM00030600872 APM00030600872

2 0 APM00023700172 APM00023700172

**# nas\_devicegroup -l**

ID name owner storage ID acl type

1 SVT MirrorView DR 50001 APM00030600872 0 MVIEW

**# nas\_server -l**

id type acl slot groupID state name

1 4 2000 2 0 server\_2

2 4 2000 3 0 server\_3

**# nas\_disk -l**

id inuse sizeMB storageID-devID type name servers

1 y 11263 APM00030600872-000F CLSTD root\_disk 1,2

2 y 11263 APM00030600872-0010 CLSTD root\_ldisk 1,2

3 y 2047 APM00030600872-0018 CLSTD d3 1,2

4 y 2047 APM00030600872-0019 CLSTD d4 1,2

5 y 2047 APM00030600872-001A CLSTD d5 1,2

6 y 2047 APM00030600872-001B CLSTD d6 1,2

**# server\_devconfig server\_2 -p -s -a |grep -v no**

server\_2 :

SCSI devices :

chain= 0, scsi-0

stor\_id= APM00030600872 celerra\_id= APM00030600872000F

tid/lun= 0/0 type= disk sz= 11263 val= 1 info= DGC RAID 5 02190F0000000FNI  
 tid/lun= 0/1 type= disk sz= 11263 val= 2 info= DGC RAID 5 02191000010010NI  
 tid/lun= 0/2 type= disk sz= 2047 val= 3 info= DGC RAID 5 02191800020018NI  
 tid/lun= 0/3 type= disk sz= 2047 val= 4 info= DGC RAID 5 02191900030019NI  
 tid/lun= 0/4 type= disk sz= 2047 val= 5 info= DGC RAID 5 02191A0004001ANI  
 tid/lun= 0/5 type= disk sz= 2047 val= 6 info= DGC RAID 5 02191B0005001BNI  
 tid/lun= 5/10 type= disk sz= 0 val= -5 info= DGC RAID 5 02195A005A005ANI  
 chain= 16, scsi-16  
 stor\_id= celerra\_id=  
 tid/lun= 0/0 type= disk sz= 0 val= -5 info= DGC RAID 5 02190F0000000FNI  
 tid/lun= 0/1 type= disk sz= 0 val= -5 info= DGC RAID 5 02191000010010NI  
 tid/lun= 0/2 type= disk sz= 0 val= -5 info= DGC RAID 5 02191800020018NI  
 tid/lun= 0/3 type= disk sz= 0 val= -5 info= DGC RAID 5 02191900030019NI  
 tid/lun= 0/4 type= disk sz= 0 val= -5 info= DGC RAID 5 02191A0004001ANI  
 tid/lun= 0/5 type= disk sz= 0 val= -5 info= DGC RAID 5 02191B0005001BNI  
 tid/lun= 0/6 type= disk sz= 0 val= -5 info= DGC RAID 5 02190600060006NI  
 tid/lun= 0/7 type= disk sz= 0 val= -5 info= DGC RAID 5 02190700070007NI  
 tid/lun= 0/9 type= disk sz= 0 val= -5 info= DGC RAID 5 02190900090009NI  
 tid/lun= 5/11 type= disk sz= 0 val= -5 info= DGC RAID 5 02195B005B005BNI  
 tid/lun= 5/15 type= disk sz= 0 val= -5 info= DGC RAID 5 02195F005F005FNI (output abridged)

### # /nas/sbin/navicli -h 10.241.168.52 storagegroup -list

Storage Group Name: Celerra\_nyip2

Storage Group UID: 7D:C4:52:4C:8E:BB:DA:11:80:2C:F1:96:72:E5:A1:96

HBA/SP Pairs:

| HBA UID                                         | SP Name | SPPort |
|-------------------------------------------------|---------|--------|
| 50:06:01:60:90:60:02:78:50:06:01:60:10:60:02:78 | SP A    | 2      |
| 50:06:01:60:90:60:02:78:50:06:01:60:10:60:02:78 | SP B    | 2      |
| 50:06:01:60:90:60:02:78:50:06:01:61:10:60:02:78 | SP A    | 3      |
| 50:06:01:60:90:60:02:78:50:06:01:61:10:60:02:78 | SP B    | 3      |
| 50:06:01:60:90:60:02:78:50:06:01:68:10:60:02:78 | SP A    | 2      |
| 50:06:01:60:90:60:02:78:50:06:01:68:10:60:02:78 | SP B    | 2      |
| 50:06:01:60:90:60:02:78:50:06:01:69:10:60:02:78 | SP A    | 3      |
| 50:06:01:60:90:60:02:78:50:06:01:69:10:60:02:78 | SP B    | 3      |

HLU/ALU Pairs:

HLU Number ALU Number

| HLU Number | ALU Number |
|------------|------------|
| 0          | 15         |
| 1          | 16         |
| 2          | 24         |
| 3          | 25         |
| 4          | 26         |
| 5          | 27         |
| 93         | 93         |
| 92         | 92         |
| 91         | 91         |
| 7          | 7          |
| 90         | 90         |
| 6          | 6          |
| 95         | 95         |
| 94         | 94         |
| 9          | 9          |

### **CLARIION COMPONENTS:**

Clariion Management authentication using Global user with username/password

Consistency Group for remote mirrors, Control & User luns

Mirrors—primary and secondary image

Storage Groups—source and destination image

### **(5) MIRROR DATA STATES:**

Out of Sync—full sync needed

Synchronized (In Sync)—Primary & Secondary LUNs are in full sync, normal state for luns receiving I/O

Consistent—actual state of most systems. If write does not occur on Primary LUN for 60 secs, system declares an ‘In Sync’ status  
Synchronizing—mirror sync in progress

Rolling Back—Only applies to MirrorView/A operations, which is rolling Secondary Image back to a known good state

#### **MIRRORVIEW/A (Asynchronous):**

→Not supported with Celerra

→Much greater distances involved

→Uses SnapCache LUNs on Primary & Secondary images to maintain Tracking Log and Transfer Log

**Note:** Changes are tracked with Tracking Log until update interval occurs, when role of Transfer Log and Tracking Log are reversed so that tracking log becomes the transfer log for the marked Session

#### **MIRRORVIEW/S (SYNC):**

→Maintains synchronous copies of primary images at secondary site

→Changes on primary side are synchronously made to secondary

→Host writes to local storage LUN, then sends to secondary storage before local storage acknowledges write back to Host

→(1) 128MB private Lun for each SP dedicated for MirrorView software

→Bitmap indicating portions of primary image that may be different from secondary image

→Write Intent Log (WIL) on Primary array is used to facilitate sync between arrays and to prevent need for full sync if a crash occurs, otherwise, is not referenced during normal operations

→Fracture Log is a bitmap of 64kb chunks that is only put into use if the link between Primary & Secondary goes down for over 10 seconds, which is the heartbeat interval—changes are automatically tracked in this bitmap once a ‘fracture’ occurs—generally, you would provision the MirrorView link to automatically recover when link is restored, as opposed to a manual resync

#### **MIRRORVIEW/S CONSISTENCY GROUP:**

→Serves as a collection of synchronous mirrors that keeps all mirrored Control Lun & Data Lun members consistent (i.e. synchronized)

→All LUNs in the CG must be from a single Storage system, and also needs to be the boot CLARiiON for the Celerra involved

→Each device group member [aka mirror image LUN] must have a corresponding mirrored destination LUN (mirror pair)

→MView cannot be failed over (promoted) if in “synchronizing” or “out-of-sync” state

→Failover is an all or nothing state for devicegroup members (no partial or individual mirror failovers)

→Ufslog LUN 1/7 is one reason for need to include Control Luns in the Consistency Group—another is the NASDB, LUNs 4/9

→Max. of (16) mirrors in CX600/700 Consistency Group [3 Control Luns (Luns 0, 1, 4) + 13 Data Luns]

→(3) System LUNs and (13) Data LUNs for CX600/700 and (3) System & (5) Data LUNs for CX400/500

→Celerra manages MirrorView via the Devicegroup (nas\_devicegroup)

→Cannot be failed over (promoted) if in “synchronizing” or “out-of-sync” state

→Ufslog is one reason for need to include Control Luns in the consistency group

→Max. of (16) mirrors in CX600/700 Consistency Group [3 Control Luns (Luns 0, 1, 4) + 13 Data Luns]

→(3) System LUNs and (13) Data LUNs for CX600/700 and (3) System & (5) Data LUNs for CX400/500

#### **CONSISTENCY GROUP CONDITIONS:**

Active—normal state

Inactive—not accepting IO

Admin Fractured—admin or media failure fracture, must be manually synchronized

System Fractured—Link down, SP is down, etc. (will recover if policy for devicegroup recovery is set to Auto)

Waiting on Sync—automatic recovery occurring, or waiting for manual recovery of group by Administrator

Unknown—transient state where group is partially fractured

Image Recovery Policy for Fractured Mirrors:

Auto—restart synchronization automatically on recovery

Manual—User must start resync using –resume

Waiting on admin—waiting for Admin to resync

Invalid

#### **CONSISTENCY GROUP STATES:**

Synchronized—all Secondary images synchronized with primary images (-activate, -restore, -suspend can occur)

Consistent—all Secondary images in Synchronized or Consistent state (-activate, -restore, -suspend can occur), I/O occurring (transitions to synchronized if no I/O in 60 sec)

Synchronizing—at least one mirror in group synchronizing (cannot activate, restore, or suspend in this state)

Out-of-Sync—some or all mirrors out-of-sync; group fractured, waiting for synchronization (cannot activate, restore, or suspend)

Scrambled—mix. of primary & secondary images

Local Only—primary images only

Incomplete—some members have only primary images after failed promote

Empty—no mirror members

#### **SECONDARY IMAGE STATES:**

In Sync—Primary & Secondary images have byte-for-byte copy

Consistent—IO's occurring, but system consistent (normal state)

Synchronizing—Secondary Image in process of sync with Primary

Out of Sync—loss of sync, must do full synchronization between mirrors

### **VERIFYING STATE OF CONSISTENCY GROUP FROM NAVISPHERE:**

Expand 'Consistency Groups' section under the Array, rightclick and view properties of Consistency Group:

**State:** Consistent (Consistent, Synchronizing, Synchronized, Out-of-Sync, etc)

**Condition:** Active (Admin Fractured, System Fractured, Waiting on Admin, Inactive)

**Role:** Primary

**Note:** Above example shows normal operational state of Consistency Group

### **VERIFYING STATE FROM CLI:**

#### **# nas\_devicegroup -list**

```
ID name      owner storage ID  acl type  
1 SVT MirrorView DR 0  APM00023700172 0  MVVIEW
```

#### **# nas\_devicegroup -info id=1 | -all**

```
Sync with CLARiiON backend ..... done  
name          = SVT MirrorView DR  
description   =  
uid           = 50:6:1:60:80:60:5:35:0:0:0:0:0:0:0:0  
state         = Consistent  
role          = Primary  
condition     = Active  
recovery policy = Automatic  
number of mirrors = 9  
mode          = SYNC  
owner         = 0  
mirrored disks = root_disk,root_ldisk,d5,d7,d8,d9,d10,d11,d12,  
local clarid   = APM00023700172  
remote clarid  = APM00030600872  
mirror direction = local -> remote
```

#### **# nas\_devicegroup -info id=1 -sync no**

**Note:** Info without a full sync

### **TROUBLESHOOTING MIRRORVIEW FROM CLI:**

#### **USING NAS DEVICEGROUP SPECIAL USAGE COMMANDS:**

##### **# export NAS\_DB\_DEBUG=1**

```
# env  
NAS_DB_DEBUG=1
```

##### **# nas\_devicegroup**

###### **special usage:**

```
| -list [-backend]  
| -info {<name>} lid=<id> | -all } [-backend] [-sync [yes|no]]  
| -suspend {<name>} lid=<id> } [-backend]  
| -resume {<name>} lid=<id> } [-backend]  
| -add <name>  
| -delete {<name>} lid=<id> }  
| -failover {<name>} lid=<id> } [-backend] [-Force]  
| -fallback {<name>} lid=<id> } [-backend]
```

#### **USING NAVICLI MIRRORVIEW, NAVICLI.JAR, or NAVISECCLI COMMANDS:**

**Note:** In order to run the navicli mirrorview commandset, must first setup security for the user in the Control Station shell context to which you are logged into. Alternatively, run the navisecccli commands and specify –password and –user for each command.

##### **1. Create Security Context for Nasadmin User in order to run MirrorView commands (root is not privileged to run cmds):**

```
$ java -jar /nas/opt/Navisphere/bin/navicli.jar -AddUserSecurity -password nasadmin -scope 0
```

##### **2. MirrorView commands:**

```
$ /nas/sbin/navicli -h 10.241.168.57 mirrorview
```

usage: mirrorview

```
-info  <-cancreate> <-mirroredluns> <-mirrorableluns>  
      <-logs> <-systems> <-maxmirrors>
```

```

-list   <-name <name>> <-mirroruid <mirroruid>> <-description>
      <-lun> <-state> <-faulted> <-transition>
      <-qthresh> <-requiredimages>
      <-imagesize> <-imagecount> <-usewriteintentlog> <-images>
-enablepath pathSPHost
-create -name name -lun lun_number
  <-description description>
  <-qthresh threshold>
  <-requiredimages num_images> <-usewriteintentlog> <-o>
-disablepath pathSPHost <-o>
-destroy -name namel-mirroruid mirroruid <-orphan> <-o>
-activate -name namel-mirroruid mirroruid
-deactivate -name namel-mirroruid mirroruid <-o>
-change -name namel-mirroruid mirroruid <-newname name>
  <-description description> <-qthresh threshold>
  <-requiredimages num_images> <-usewriteintentlog 0|1> <-o>
-addimage -name namel-mirroruid mirroruid <-arrayhost hostname
  -lun lunnumberl-arrayhost hostname -lunuid lunuidl-arrayuid
  arrayuid -lunuid lunuid <-recoverypolicy manuallauto>
  <-issyncrequired 0|1> <-syncrate high|medium|low>
-removeimage -name namel-mirroruid mirroruid -imageuid imageuid <-o>
-changeimage -name namel-mirroruid mirroruid -imageuid imageuid
  <-recoverypolicy manuallauto> <-syncrate high|medium|low>
  <-o>
-promoteimage -name namel-mirroruid mirroruid -imageuid imageuid <-o>
-fractureimage -name namel-mirroruid mirroruid -imageuid imageuid <-o>
-syncimage -name namel-mirroruid mirroruid -imageuid imageuid <-o>
-allocatelog -spA lun_number -spB lun_number <-unbind>
-deallocatelog <-o> <-unbind>
-listlog
-listsyncprogress <-name name | -mirroruid mirroruid>

```

**\$ /nas/sbin/navicli -h 10.241.168.57 mirrorview –info** (verbose output, see optional switches below)

Can a mirror be created on this system: YES

Logical Units that are mirrored in Primary Images: 93 91 95 1 94 92 4 0 90

Logical Units that are mirrored in Secondary Images:

Logical Units that can be mirrored: 54 70 36 19 6 47 5 73 39 55 7

5 71 33 49 42 7 58 44 74 12 60 76 52 68 10 34 50 79 17 45 3 61 37 53 69 11 32 20

48 2 80 40 56 72 38 66 8 13 41 57 77 15 43 59 35 21 51 67 9 18 46 14 78 16

Is Write Intent Log Used: YES

Remote systems that can be enabled for mirroring:

Remote systems that are enabled for mirroring:

| UID of the array | Status |
|------------------|--------|
| -----            | -----  |

50:06:01:60:90:60:07:BB Enabled on both SPs

Maximum number of possible Mirrors: 100

**\$ /nas/sbin/navicli -h 10.241.168.57 mirrorview -info -logs**

Is Write Intent Log Used: YES

**\$ /nas/sbin/navicli -h 10.241.168.57 mirrorview -info -mirroredluns**

Logical Units that are mirrored in Primary Images: 93 91 95 1 94 92 4 0 90

Logical Units that are mirrored in Secondary Images:

**\$ /nas/sbin/navicli -h 10.241.168.52 mirrorview -info -mirroredluns**

Logical Units that are mirrored in Primary Images:

Logical Units that are mirrored in Secondary Images: 14 6 9 94 95 7 90 91 92 93

Note: Must run command on Secondary side to obtain list of Secondary Images

**\$ /nas/sbin/navicli -h 10.241.168.57 mirrorview –list | -list –lun –state | -list -faulted | -list –transition**

MirrorView Name: Lun 0 Primary

MirrorView Description:

MirrorView UID: 50:06:01:60:80:60:05:35:35:00:00:00:00:00:00:00

Logical Unit Numbers: 0

Remote Mirror Status: Mirrored  
 MirrorView State: Active  
 MirrorView Faulted: NO  
 MirrorView Transitioning: NO  
 Quiesce Threshold: 60  
 Minimum number of images required: 0  
 Image Size: 23068672  
 Image Count: 2  
 Write Intent Log Used: YES  
 Images:  
     Image UID: 50:06:01:60:80:60:05:35  
     Is Image Primary: YES  
     Logical Unit UID: 60:06:01:60:2A:F0:0A:00:37:E8:D8:FB:88:BB:DA:11  
     Image Condition: Primary Image  
     Preferred SP: A  
     Image UID: 50:06:01:60:90:60:07:BB  
     Is Image Primary: NO  
     Logical Unit UID: 60:06:01:60:7E:B5:09:00:04:E2:09:58:32:D6:DA:11  
     Image State: Consistent  
     Image Condition: Normal  
     Recovery Policy: Manual  
     Preferred SP: A  
     Synchronization Rate: Medium  
**MirrorView Name:** **LUN 90 Primary**  
 MirrorView Description:  
 MirrorView UID: 50:06:01:60:80:60:05:35:31:00:00:00:00:00:00:00  
 Logical Unit Numbers: 90  
 Remote Mirror Status: Mirrored  
 MirrorView State: Active  
 MirrorView Faulted: NO  
 MirrorView Transitioning: NO  
 Quiesce Threshold: 60  
 Minimum number of images required: 0  
 Image Size: 20971520  
 Image Count: 2  
 Write Intent Log Used: YES  
 Images:  
     Image UID: 50:06:01:60:80:60:05:35  
     Is Image Primary: YES  
     Logical Unit UID: 60:06:01:60:2A:F0:0A:00:42:E8:D8:FB:88:BB:DA:11  
     Image Condition: Primary Image  
     Preferred SP: A  
     Image UID: 50:06:01:60:90:60:07:BB  
     Is Image Primary: NO  
     Logical Unit UID: 60:06:01:60:7E:B5:09:00:07:E2:09:58:32:D6:DA:11  
     Image State: Synchronized  
     Image Condition: Normal  
     Recovery Policy: Manual  
     Preferred SP: A  
     Synchronization Rate: Medium

### \$ /nas/sbin/navicli -h 10.241.168.57 mirrorview -list -lun -state

MirrorView Name: Lun 93 Primary  
 Logical Unit Numbers: 93  
 Remote Mirror Status: Mirrored  
 MirrorView State: Active

### \$ /nas/sbin/navicli -h 10.241.168.57 mirrorview -list -lun -faulted

MirrorView Name: Lun 93 Primary  
 Logical Unit Numbers: 93  
 Remote Mirror Status: Mirrored  
 MirrorView Faulted: YES

**\$ /nas/sbin/navicli -h 10.241.168.57 mirrorview –listlog** (displays Write Intent Log LUNs)

Storage Processor: SP A

Lun Number: 100

Storage Processor: SP B

Lun Number: 101

**\$ /nas/sbin/navicli -h 10.241.168.57 mirrorview –listsyncprogress** (shows % sync progress of each mirrored lun)

MirrorView Name: Lun 93 Primary

Has Secondary Images: YES

Image UID: 50:06:01:60:90:60:07:BB

Image State: Synchronized

Synchronizing Progress(%): 100

MirrorView Name: Lun 0 Primary

Has Secondary Images: YES

Image UID: 50:06:01:60:90:60:07:BB

Image State: Consistent

Synchronizing Progress(%): N/A

**Note:** By definition, if an I/O has not occurred for 60 seconds or more, the Image State will be “Synchronized”. If an I/O has occurred within the 60 sec. window, then the Image State will be “Consistent”. These are the two desired LUN states.

**\$ java -jar /nas/opt/Navisphere/bin/navicli.jar -h 10.241.168.57 mirror -sync -listgroups**

Maximum Number of Groups Allowed: 16

Maximum Number of Mirrors per Group: 16

Group Name: SVT MirrorView DR

Group ID: 50:06:01:60:80:60:05:35:00:00:00:00

Description:

State: Consistent

Role: Primary

Condition: Active

Recovery Policy: Automatic

Mirror Name: Lun 91 Primary

**Note:** The –list commands show information about the remote mirrors

## **TROUBLESHOOTING MIRRORVIEW USING NAVISECCLI:**

**Note:** If a security file does not already exist for nasadmin, create one using the following—same file works for Navicli or Naviseccli

**\$ /nas/opt/Navisphere/bin/navisecccli -h 10.241.168.57 -AddUserSecurity -password mypass -scope 0 -user nasadmin**

**\$ /nas/opt/Navisphere/bin/navisecccli -h 10.241.168.57 mirror -sync -info** (info on mirror installation)

**Note:** This command is equivalent to the navicli -h <spa> mirrorview –info

## **OTHER NAVISECCLI COMMANDS:**

***Caution:*** Do not use any of the following commands unless instructed to do so by Engineering

**\$ /nas/opt/Navisphere/bin/navisecccli -h 10.241.168.57 mirror -sync -enablepath <sp>**

**\$ /nas/opt/Navisphere/bin/navisecccli -h 10.241.168.57 mirror -sync -disablepath <sp>**

**\$ /nas/opt/Navisphere/bin/navisecccli -h 10.241.168.57 mirror -sync -allocatelog -spA 101 -spB 102**

**\$ /nas/opt/Navisphere/bin/navisecccli -h 10.241.168.57 mirror -sync -deallocatecatalog**

**\$ /nas/opt/Navisphere/bin/navisecccli -h 10.241.168.57 mirror -sync -destroy** (destroy mirror & images)

**\$ /nas/opt/Navisphere/bin/navisecccli -h 10.241.168.57 mirror -sync -create** (create new mirror)

**\$ /nas/opt/Navisphere/bin/navisecccli -h 10.241.168.57 mirror -sync -creategroup** (create Consist. Group)

**\$ /nas/opt/Navisphere/bin/navisecccli -h 10.241.168.57 mirror -sync -destroygroup <CG\_name>**

**\$ /nas/opt/Navisphere/bin/navisecccli -h 10.241.168.57 mirror -sync -addtogroup | -removefromgroup**

**\$ /nas/opt/Navisphere/bin/navisecccli -h 10.241.168.57 mirror -sync -changegroup | -syncgroup | -fracturegroup | -promotegroup**

**\$ /nas/opt/Navisphere/bin/navisecccli -h 10.241.168.57 mirror -sync -addimage** (add sec. lun to mirror)

**\$ /nas/opt/Navisphere/bin/navisecccli -h 10.241.168.57 mirror -sync -removeimage** (remove second. from mirror)

**\$ /nas/opt/Navisphere/bin/navisecccli -h 10.241.168.57 mirror -sync -syncimage** (starts sync of secondary)

**\$ /nas/opt/Navisphere/bin/navisecccli -h 10.241.168.57 mirror -sync -promoteimage** (promote secondary)

**\$ /nas/opt/Navisphere/bin/navisecccli -h 10.241.168.57 mirror -sync -change** (change mirror properties)

**\$ /nas/opt/Navisphere/bin/navisecccli -h 10.241.168.57 mirror -sync -changeimage** (change second. image)

**\$ /nas/opt/Navisphere/bin/naviseccli -h 10.241.168.57 mirror –sync –fractureimage** (fracture image from mirror)

## **TROUBLESHOOTING MIRRORVIEW/S:**

- An expectation that NAS folks learn the CLARiiON platform & Navisphere to resolve backend configuration issues
- Identify Error messages, review Logs, set debug logging, search Primus
- For True DR situations, must escalate to EE when customer ready to rebuild Source side for Restoration [EE & Development must sponsor this activity]

### **After failover, the active NASDB on Destination side will be located:**

**/nas/rdf/500/** [Control Station mounts as /dev/ndj1—mirrored LUN 9]

→ For ‘True Disaster’ event where Source Celerra & CLARiiON arrays are down, the –activate will indicate that a full resync will be required to perform a restore, & promotes the Secondary LUN images to ‘Local-Only’ Primary mirrors so that failover can occur even without the Source online

→ NAS Upgrades require no special preparation or ‘downed’ links—following normal practice for ‘paired’ Celerras, Upgrade Target side first, then the Source

→ FLARE Upgrades should follow COLU NDU where applicable, and pay special attention to the CLARiiON Procedure Generator, as there are some special conditions around MirrorView fixes with FLARE. In general, Flare 19 Patch 034 and above will be o.k. and has been successfully tested

## **MIRRORVIEW LOGS:**

**/nas/log/dr\_log.al** → Normal location of DR logs after –activate, -init, or –restore has completed; located in local /nas/dos of the local Celerra system

**/tmp/dr\_log.al** → Logs are temporarily written to this directory on the system when –init, –activate, or –restore operations are in progress—logs are moved to /nas/log after the MirrorView threads are done running. It should be noted that even after failures, the logs are moved to this location.

→ Point of this information is that the applicable DR logs do not reside in /nas/rdf/500/log with respect to –init, -activate, or –restore

**\$ ls -la /nas/log |grep dr\***

**-rwxrwxr-x 1 nasadmin nasadmin 38860 May 18 13:54 dr\_log.al** [Main MView log of interest]

```
-rw-r--r-- 1 root root 1329 May 11 07:26 dr_log.al.err
-rw-r--r-- 1 root root 16821 May 10 15:14 dr_log.al.trace
```

**/nas/var/messages** [Good source of information for NBS issues]

**/nas/log/symapi.log** [some info may be found here related to backend issues]

**/nas/tools/collect\_support\_materials** script used for escalations, also contains dr\_logs

**/nas/rdf/500/log** [or UID in the path could be 501, 50001, etc.]

**Note:** After –activate is complete, this is where NBSNAS is located and where logs pertaining to the Servers while operating in a failed over state, are stored.

## **DEBUG LOGGING AND ERROR MESSAGES:**

Set SYMAPI or NASDB Debug Logging if needed:

**# export CLARAPI\_DEBUG=1** (Clariion API debug logging—outputs to screen & dumps to /nas/log/symapi.log)

**# export NAS\_DB\_DEBUG=1** (NAS debug logging--/nas/log)

### **New CCMD format MirrorView Error Messages:**

Many error messages will be in new CCMD format and documented, either in MView engineering documentation, or Primus (Check these sources first)

**Error 13421904897 : The device group is System Fractured.**

(seen during nas\_mvview –activate → refer to emc136398)

Other error messages are still generic, such as Error 5008, and may be seen across different features and for different situations

## **GENERIC ERROR MESSAGES STILL SEEN WITH NAS 5.5:**

Error 5008: 139:cmd failed: /bin/umount /dev/ndj1 2>&1: (# /nas/sbin/nas\_mvview –restore)

Error 5008: -1:Cannot restore nyip1. Please run restore on site nyip1. (# /nas/sbin/nas\_mvview –restore)

Error 5023: APM00023700172 unable to connect to storage (# /nas/sbin/setup\_slot -i 2

Error 5005: failed to complete command (# /nas/sbin/setup\_slot -i 2)

Validating mirror group configuration ..... Error 2237: Execution failed: Segmentation violation: Operating system signal.  
(# /nas/sbin/nas\_mvview –activate)

Error 2221: Operation not permitted (# /nas/sbin/nas\_mvview –activate)

Error 2237: server\_2 : Execution failed: Read-only file system: Routine failure. [RAW\_FILE.file\_open] (# server\_mount server\_2)

# /nas/sbin/nas\_mvview -activate fails: Error 5017: storage health check failed (APM00023700172 SPA is failed over)

Error 2237: Execution failed: No such file or directory: Routine failure. (# nas\_server -l)

Error 2237: Execution failed: is\_commitable: Precondition violated. [NAS\_DB.commit]

## **IMPORTANT RULE-of-THUMB ON ACTIVATE OR RESTORE FAILURES:**

**Try to determine point of failure before proceeding:**

From lab testing, errors will either prevent you from performing an action (figure out issue and carryon), or will generally allow the MirrorView devices to failover or failback—most problems seem to surface at the end of the process, where the Servers are not completely failed over or failed back, leaving NAS in a state of limbo and Servers in an inconsistent state

**Retry the –activate or –restore if first attempt fails:**

Check for obvious issues, then retry the –activate or –restore command—in the case of restores, if the 2nd attempt fails, then try to complete the restore on the Source side logged in as nasadmin, then root, using /nasmd/sbin/nas\_mview –restore directory only!

**ADVANCED MIRRORVIEW TROUBLESHOOTING:**

**MIRRORVIEW REFERENCES:**

**CLARiiON backend setup guide:**

Celerra MirrorView/Synchronous Setup on CLARiiON Backends NAS 5.5 (NAPA III release 5.5.23.x)

See sections for setting up Active/Passive & Active/Active MView

MirrorView not intended for customer setup

Most config & setup calls will come from PS

**Celerra Configuration and Troubleshooting Guide:**

Using MirrorView/Synchronous with Celerra for Disaster Recovery NAS 5.5 (NAPA III release 5.5.23.x)

Concepts, System requirements, Initializing, activating failover, conducting restore, Managing, Troubleshooting, Error messages

No real Celerra CLI for configuring MirrorView Backend

CLARiiON migrating to new navisecli command set

**RUNNING MIRRORVIEW COMMANDS FROM CONTROL STATION:**

**Note:** Setup the security context, then run navicli cmd as nasadmin

a.) \$ java -jar /nas/opt/Navisphere/bin/navicli.jar -AddUserSecurity -password nasadmin -scope 0

b.) \$ /nas/sbin/navicli -h 10.241.168.52 mirrorview –info | -list

Or

a.) \$ /nas/opt/Navisphere/bin/navisecli -h 10.241.168.57 -AddUserSecurity -password mypass -scope 0 -user nasadmin

b.) \$ /nas/opt/Navisphere/bin/navisecli -h 10.241.168.57 mirror -sync -info (info on MirrorView installation)

**Note:** Navisecli will replace navicli in the near future

**USING SERVER STANDBY TO RESTORE or ACTIVATE SERVERS:**

→Using server\_standby to complete activate or restore for an operation that completes everything but leaves Server in faulted state

In certain situations, if the –activate succeeds in failing over MirrorView devices on CLARiiON backend, and NASDB is mounted on /nas/500/rdf, but the destination Server is in a faulted status, may need to run following to complete failover

**# nas\_server -l**

```
id    type acl slot groupID state name
1     4   0   2       2   server_2.faulted.server_3
2     1   0   3       0   server_2
```

**# server\_standby server\_2.faulted.rdf –activate rdf**

Same situation applies when failing back to Source using –restore—may need to use server\_standby on Source DM

**# server\_standby server\_2.faulted.rdf –restore rdf**

**Note:** In general, do not run the server\_standby command without TS or Eng approval. Many things should be verified before this command is run.

**IMPORTANCE OF SERVER SETUP FILE: /nas/server/server\_setup**

The server\_setup file is crucial in defining the role of Local Data Mover standbys on a Celerra, and the Remote RDF Standby relationships that may be in place for MirrorView DR. This file is dynamically updated during –activate & -restore operations, hence is susceptible to corruption or damage if the operation does not complete successfully. Examining this file is useful in understanding the current state of the Servers.

→Example of the Server\_setup file for Active/Passive configuration

**EXAMPLE 1:** Active side with Server\_2 Primary & local Standby defined as Server\_3. Local Standby Server\_3 has RDF failover relationship on slot\_3 on the Remote Celerra & Local Primary Server\_2 has RDF Standby (failover) relationship on slot\_2 on Remote Celerra

**# nas\_server -l** (Active Source side)

```
id    type acl slot groupID state name
1     1   1000 2      0   server_2
2     4   1000 3      0   server_3
```

**# cat /nas/server/server\_setup**

**2:0:1:1:3,1000,0,: -->Server\_3: no local bkup server:1=n/a:1=manual fail:3,1000,0=RDF Stdby slot\_3, ACL 1000, State 0**

**1:2:1:3:2,1000,0,: -->Server\_2:local bkup server\_3:1=n/a:3=auto fail:2,1000,0=RDF Stdby slot\_2, ACL 1000 State 0**

**Server\_setup file Fields Defined:**

**Primary\_id: Backup\_id: Component n/a: Failover Policy: RDF Standby Slot#, ACL, State:**

**EXAMPLE 2:** Target side showing both Servers as RDF Standbys to the Source Celerra

# nas\_server -l (Passive Target side)

```
id type acl slot groupID state name
1 4 1000 2      0 server_2
2 4 1000 3      0 server_3
```

# cat server\_setup

1:2:1:3:: -->Server\_2:local backup server\_3:1=n/a:3=auto failover:no RDF relationships defined on this side

-->Server\_3 not referenced since it has no local Standby or RDF relationship defined

#### Data Mover Failover Policies in /nas/server/server\_setup file:

1=Manual failover (Default Celerra policy; all RDF setups are manual failover policy)

2=Retry (Panic will cause reboot to clear condition, if unsuccessful, failover occurs)

3=Auto (Data Mover will failover automatically after panicking)

#### IMPORTANCE OF NBS CONFIGURATION FILES:

NS systems use the Data Mover as NBS Server with backend visibility that the Control Station then uses as an NBS Client to see and write to the NAS DB. The NBS configuration files are especially important in the role of DR Failover, whether for SRDF or for MirrorView.

#### Normal Data Mover NBS configuration—Source Side Data Mover for MirrorView:

# ls -la /nas/server/slot\_2/nbs\*

```
lrwxrwxrwx 1 root root 28 Jul 7 15:36 nbs.cs -> /nas/server/slot_2/nbs.cs.ro
```

# cat /nas/server/slot\_2/nbs.cs

```
nbs add nbsid=1 vol=NBS1 rw=192.168.1.100:192.168.1.101:...2.101 exclusive raw share (NBS1 LUN 0 /dev/nda1 DOS)
```

```
nbs add nbsid=5 vol=NBS5 rw=192.168.1.100:192.168.1.101:...2.101 exclusive raw share (NBS5 LUN 4 /dev/nde1 NBSNAS)
```

```
nbs add nbsid=6 vol=NBS6 rw=192.168.1.100:192.168.1.101:...2.101 exclusive raw share (NBS6 LUN 5 /dev/ndf1 VAR)
```

```
nbs add nbsid=7 vol=RDF7 exclusive raw ro=192.168.1.100:192.168.1.101:...2.101 (NBS7 LUN 6 /dev/ndg1 Remote)
```

```
nbs add nbsid=10 vol=RDF10 exclusive raw ro=192.168.1.100:192.168.1.101:...2.101 (NBS10 LUN 9 /dev/ndj1 Remote)
```

**Note:** LUN 6 is remote equivalent to LUN 0 DOS boot, LUN 9 equivalent to LUN 4 NBSNAS, which becomes “rw” on destination side after nas\_mview –activate. During activate, build\_config is run for Data Mover and boot.cfg copied to local /nas/dos/slot\_x directory so that DM can reboot with configuration. After activate, DOS partition is unmounted from CS. LUNs 1 & 4 failover to promoted mirror LUNs 7 & 9 and are used by Destination Data Movers.

#### NBS FILES AFTER ACTIVATE:

# ls -la nbs\* /nas/rdf/500/server/slot\_2

```
lrwxrwxrwx 1 root root 30 Aug 29 12:51 nbs.cs -> /nas/server/server_1/nbs.cs.rw
```

# cat nbs.cs.rw

```
nbs add nbsid=7 vol=RDF7 exclusive raw rw=192.168.1.100:192.168.1.101:192.168.2.100:192.168.2.101 (NBS7 LUN 6 /dev/ndg1 Remote)
```

```
nbs add nbsid=10 vol=RDF10 exclusive raw rw=192.168.1.100:192.168.1.101:192.168.2.100:192.168.2.101 (NBS10 LUN 9 /dev/ndj1 Remote)
```

**Note:** Example shows where LUNs 6 & 9 on the Target Celerra become RW to the local Servers after failover

#### TARGET DEVICES AFTER FAILOVER:

[root@nyip2 mvadmin]# nas\_disk -l

| id | inuse | sizeMB | storageID           | devID | type       | name | servers        |
|----|-------|--------|---------------------|-------|------------|------|----------------|
| 1  | y     | 11263  | APM00030600872-0006 | CMSTD | root_disk  | 1,2  | (LUN 6)        |
| 2  | y     | 11263  | APM00030600872-0007 | CMSTD | root_ldisk | 1,2  | (LUN 7)        |
| 3  | y     | 2047   | APM00023700172-0002 | CLSTD | d3         | 1,2  |                |
| 4  | y     | 2047   | APM00023700172-0003 | CLSTD | d4         | 1,2  |                |
| 5  | y     | 2047   | APM00030600872-0009 | CMSTD | d5         | 1,2  | (LUN 9)        |
| 6  | y     | 2047   | APM00023700172-0005 | CLSTD | d6         | 1,2  |                |
| 7  | y     | 10239  | APM00030600872-005B | CMSTD | d7         | 1,2  | (1st Data LUN) |

**Note:** Unlike SRDF setups, the Remote Celerra will not see its Secondary Mirrored luns until they are ‘promoted’ to Primary status during successful failover—see CMSTD devices

[root@nyip2 mvadmin]# server\_devconfig server\_2 -p -s -a |grep -v no

server\_2 :

SCSI devices :

chain= 0, scsi-0

stor\_id= APM00030600872 celerra\_id= APM00030600872001D

tid/lun= 0/0 type= disk sz= 11263 val= 1 info= DGC RAID 5 02191D0000001DNI (**Local boot LUN 0**)

tid/lun= 0/1 type= disk sz= 11263 val= 2 info= DGC RAID 5 02191E0001001ENI

tid/lun= 0/2 type= disk sz= 2047 val= 3 info= DGC RAID 5 02191F0002001FNI

```
tid/lun= 0/3 type= disk sz= 2047 val= 4 info= DGC RAID 5 02192100030021NI  
tid/lun= 0/4 type= disk sz= 2047 val= 5 info= DGC RAID 5 02192200040022NI  
tid/lun= 0/5 type= disk sz= 2047 val= 6 info= DGC RAID 5 02192300050023NI  
tid/lun= 0/6 type= disk sz= 11263 val= 1 info= DGC RAID 5 02190600060006NI diskerr= mismatch:tid (DOS LUN 6 after failover)  
tid/lun= 0/7 type= disk sz= 11263 val= 2 info= DGC RAID 5 02190700070007NI diskerr= mismatch:tid (UFSLOG LUN 7 after failover)  
tid/lun= 0/9 type= disk sz= 2047 val= 5 info= DGC RAID 5 02190900090009NI diskerr= mismatch:tid (NBSNAS LUN 9 after failover)
```

### **CONSISTENCY GROUP/DEVICEGROUP CONFIGURATION FILE:**

**Note:** Each Active Celerra configuration will have a master file that essentially defines the properties of the mirrored luns found in the Consistency Group/Devicegroup. Upon successful failover, the existing Source MirrorView configuration is disassembled and the Target Celerra LUNs are promoted from Secondary Image status to Primary Images. This file is especially important during failback and is used to reconstruct and rebuild the original Mirror relationship between Source & Target sides.

#### **# cat /nas/site/.clariion\_mg\_config**

```
Lun 0 872 Primary Side ALU 29:Celerra_nyip2:APM00030600872:0029:0000:Celerra_nyip1:APM00023700172:0201:0006  
LUN 1 872 Primary Side HLU 30:Celerra_nyip2:APM00030600872:0030:0001:Celerra_nyip1:APM00023700172:0202:0007  
Lun 4 872 Primary Side HLU 34:Celerra_nyip2:APM00030600872:0034:0004:Celerra_nyip1:APM00023700172:0203:0009  
Lun 17 872 Primary Side HLU 50:Celerra_nyip2:APM00030600872:0017:0050:Celerra_nyip1:APM00023700172:0110:0110  
Lun 19 872 Primary Side HLU 52:Celerra_nyip2:APM00030600872:0019:0052:Celerra_nyip1:APM00023700172:0112:0112  
Lun 18 872 Primary Side HLU 51:Celerra_nyip2:APM00030600872:0018:0051:Celerra_nyip1:APM00023700172:0111:0111  
Lun 28 872 Primary Side HLU 53:Celerra_nyip2:APM00030600872:0028:0053:Celerra_nyip1:APM00023700172:0113:0113
```

### **DEBUGGING MIRRORVIEW ISSUES USING XML DEBUG:**

There is a method for querying XML data and using the data for debugging MirrorView failures.

1. Find the query string relevant to a recent failure in the /nas/http/logs/access\_log file
2. Search the logfile from the end to find the last occurrence of "?UseWriteIntentLog"
3. After matching up time and date with problem being debugged, locate and copy all of the information from the indicated start and finish locations, then copy the paragraph string to notepad:  
Mode=Pre&Command=nas\_storage&arg=-query:Id==1&arg=-fields:.....abridged..... :\*&arg=-  
fields:Alu,Hlu&arg=-format:%25s:%25s;
4. Run the following command and paste the above paragraph string:  
# export QUERY\_STRING=' Mode=Pre&Command..... Alu,Hlu&arg=-format:%25s:%25s;'
5. If successfully cut and pasted, the export Query will produce debug information viewable with the following:

```
6. # nas_cmd @cs_cgi  
Content-type: text/html  
<HTML>  
<HEAD><SCRIPT language='javascript1.1'>  
var command_exit_status=0;  
var command_error_code=0;  
var command_error_msg="";  
var command_message_code='0';  
</SCRIPT></HEAD>  
<BODY ><PRE>  
remote_storage_id=APM00023700172,remote_storage_type=Clariion,remote_storageAuthorized=True,remote_storage_SPA=10.241  
.168.57,remote_storage_SPB=10.241.168.58,,remote_vdisk_mirrors=LUN 0  
Primary:APM00023700172:0000:APM00030600872:0006:True:6::LUN 1  
Primary:APM00023700172:0001:APM00030600872:0007:True:7::LUN 4 Primary:A  
PM00023700172:0004:APM00030600872:0009:True:1::LUN 91 Primary:APM00023700172:005  
B:APM00030600872:005B:True:6::Lun 93 Primary:APM00023700172:005D:APM00030600872:  
005D:True:1::LUN 95 Primary:APM00023700172:005F:APM00030600872:005F:True:6::Lun  
90 Primary:APM00023700172:005A:APM00030600872:005A:True:1::Lun 92 Primary:APM00  
23700172:005C:APM00030600872:005C:True:1::Lun 94 Primary:APM00023700172:005E:APM  
00030600872:005E:True:1::,remote_device_group=172_AP_MView,remote_devicegroup_mi  
rrors=LUN 0 Primary:APM00023700172:0000:APM00030600872:0006:True:6::LUN 1 Primar  
y:APM00023700172:0001:APM00030600872:0007:True:7::LUN 4 Primary:APM00023700172:0  
004:APM00030600872:0009:True:1::LUN 91 Primary:APM00023700172:005B:APM0003060087  
2:005B:True:6::Lun 93 Primary:APM00023700172:005D:APM00030600872:005D:True:1::LU  
N 95 Primary:APM00023700172:005F:APM00030600872:005F:True:6::Lun 90 Primary:APM0  
0023700172:005A:APM00030600872:005A:True:1::Lun 92 Primary:APM00023700172:005C:A  
PM00030600872:005C:True:1::Lun 94 Primary:APM00023700172:005E:APM00030600872:005
```

```
E:True:1:;,remote_storage_group=Celerra_nyip1,remote_storage_hlu_mapping=0000:0  
000:0001:0001;0002:0002;0003:0003:0004:0004;0005:0005;0095:0055;0094:0054;0093:0  
053;0092:0052;0091:0051;0090:0050;0203:0009;0202:0007;0113:0113;0201:0006;0112:0  
112;0111:0111;0110:0110;,</PRE></BODY>  
</HTML>
```

**Note:** The semi-colons separate the output by operations. Look for an operation that has failed.

## **NAS 5.5 NDMP & LEGATO NETWORKER:**

- NDMP drive sharing supported only for SAN networks
- Dynamic Device Sharing allowed between Data Movers in the same TLU
- Ports 10001 – 10004 are used for 3-way NDMP (otherwise just port 10000)
- PAX is Portable Archive Interchange, a protocol designed to work with UNIX tape formats for file-level backups/restores
- TAR format in PAX that traverses file tree in depth-first order, then across
- DUMP format in PAX that traverses file tree in mixed width-first and depth-first order
- NAS 5.5 supports File Filtering, Integrated Checkpoints, Volume Backups (VBB), and DDAR for backup solutions
- Legato requires HIST & DIRECT=y for DAR backup/recoveries
- Legato cannot scan for robot with NDMP FC unless TLU is attached to Data Mover NDMP Host

## **LEGATO DATA MANAGER FOR CELERRA (aka FILE SYSTEM MANAGER FOR NAS):**

- Strictly speaking, File System Manager for NAS version 2.0 replaces DM for Celerra version 1.0 & is not directly tied to 5.5 release
- Will support Solaris 9 & 10 & Win2k/2k3, using Java & C, and iAnywhere database
- Contains better job control & monitoring, improved user interface, integration with IIM, Fulltime AutoStart, VisualSRM
- All paths for Jobs will be absolute for Unix and UNC for Windows
- Migrated files now use file versioning (i.e., every time file is recalled and re-migrated, new version created on Secondary)
- Can use either CLI or GUI for configuring
- Fulltime AutoStart will monitor the DMC and restart and stop services as appropriate

### **Migration Policies:**

Migration → Source, Dest, Mode Filemover/Symlink, Readback method, File Match criteria

Orphan Files → Delete policy, new with v2.0, File Match criteria

Source Scan

### **Future Version of FSM for NAS:**

- .Next will be released mid 2006 and will contain HTTP and BAR support
- HTTP support direct to disk or direct to Centera
- Bulk metadata retrieval for faster file scanning
- Better migration rules via Visual SRM
- Possible support for Linux Control Station as policy engine

## **FILEMOVER METADATA RETRIEVAL (replaces Bulk Attributes Retrieval):**

- Uses XML over HTTP using port 5080 on DM
- Authentication: Digest, Null authentication, IP Address
- BAR is an extension of DHSM
- Scans file system to retrieve attributes and uses basic filtering criteria, producing XML reports, based on a targeted scans using an enhanced FileMover API and scan criteria
- Can perform full or incremental metadata retrieval using FileMover API
- When retrieving files from Secondary storage, an HTTP GET will be issued using URL for OFFLINE\_PATH
- HTTP connections will be defined as ‘persistent’ and multiple requests can go over the same connection, though all requests must be returned in the order in which they were received
- FileMover will still support up to 1024 connections to the Secondary Storage system
- Max of (4) concurrent metadata retrieval jobs at one time

DHSM\_GET\_BULK\_ATTRS

DHSM\_QUERY\_BULK\_RETRIEVAL

DHSM\_ABORT\_BULK\_RETRIEVAL

→ Two requests used: DHSM\_SET\_OFFLINE\_ATTRS & DHSM\_GET\_ATTRS

**\$server\_config server\_2 -v “bar list | help | info | abort” “printstats bar” “printstats bar full”**

## **FILEMOVER HTTP SUPPORT (aka DHSM):**

- Introduces HTTPClient v1.1 support for Secondary Storage so that other products can utilize FileMover (KVS, Documentum, Opentext, IXOS, Sharepoint)

**Note:** Up to this point, secondary storage was accessed only via NFS or CIFS interface—can now access using HTTP, so applications such as KVS can access secondary storage

→ max 32 simultaneous tcp connections for FileMover to increase HTTP Read performance

**Note:** Previously, could only use single TCP connection to conduct HTTP Client reads from Secondary Storage

→Uses mtime or entity tags for validation

→Allows for 32k reads from HTTP Client for full and partial read policies, and 8k for passthrough

**Note:** Previously, only allowed 8k Reads, creating excessive overhead for HTTP Server

→Support for CGI (Common Gateway Interface) & Non-CGI HTTP connections

→Can use either basic or digest authentication challenges

#### **CREATING CGI HTTP CONNECTION:**

**\$ fs\_dhsm -connection pfs -create -type http -secondary http://apache.njqa.us.dg.com/cgi-bin/access.sh**

**Note:** Uses default port 80, hence this feature is designed for LAN secure networks

#### **CONVERTING FILES TO STUB FILES:**

**\$ set\_attributes -v 1123623275 10.5.8.224 nfs:/pfs/foo http://apache.njqa.us.dg.com/cgi-bin/access.sh/dm2/foo**

**Note:** Offline Attributes are set via API for OFFLINE\_MTIME & OFFLINE\_ETAG (latter entity tag is new attribute for http)

#### **CREATING NON-CGI HTTP CONNECTION:**

**\$ fs\_dhsm -connection pfs -create -type http -secondary http://10.5.8.74/dir1 -cgi n**

#### **CONVERTING NON-CGI FILES TO STUB FILES:**

1. Telnet to Secondary Storage to obtain entity tag from specific file

Etag: "80f7b39566cc51:d6a"

2. Set file offline by converting to Stub:

**\$ set\_attributes -e "80f7b39566cc51:d6a" 10.5.8.224 nfs:/pfs/bar http://10.5.8.74/dir1/dir2/bar**

#### **TROUBLESHOOTING:**

##### **Non-CGI Connection Failure:**

**\$ fs\_dhsm -connection svr4fs1 -create -type http -secondary http://dhsm-w2k/dir1**

Error 5005: failed to complete command

Server\_log:

2005-08-26 09:05:46: DHSM: 3: HTTP verifyServerHealthInternal(): Unexpected HTTP status 302 : 302 Found

2005-08-26 09:05:46: DHSM: 3: If a non-cgi connection is desired "-cgi n" should be used during connection creation

Solution: Use -cgi n when creating connection

##### **Web Server Down When Trying to Create Connection:**

**\$ fs\_dhsm -connection svr4fs1 -create -type http -secondary http://apache.njqa.us.dg.com/cgi-bin/access.sh**

Error 5005: failed to complete command

Server log:

2005-08-26 10:34:42: DHSM: 3: HTTP createDartHTTPConnections(): could not connect/talk to server apache.njqa.us.dg.com

Solution: Restart Web Server

##### **Missing User Credentials will cause Connection Failure:**

**\$ fs\_dhsm -connection svr4fs1 -create -type http -secondary http://dhsm-w2k/dir1 -cgi n**

Error 5005: failed to complete command

Server log:

2005-08-26 11:44:27: DHSM: 3: HTTP: Couldn't get http response - 1001 User and password required

2005-08-26 11:44:27: DHSM: 3: HTTP createDartHTTPConnections(): could not connect/talk to server dhsm-w2k

Solution: Retry using username and password

##### **Wrong CGI Format being used by Application when accessing Stub Files:**

2005-08-26 14:15:39: DHSM: 3: HTTP: Expecting a content type of "application/octet-stream"

2005-08-26 14:15:39: DHSM: 3: response had "text/html; charset=ISO-8859-1" content type

Solution: Modify CGI application has to conform

#### **NDMP DDAR 3-WAY SUPPORT:**

→Directory DAR restore of given directory and all subdirectories and files

→Two variables required for DDAR: RECURSIVE (default=yes) & DIRECT (default=no)—if both variables set to yes, DDAR will be used for directory restores

→Incremental DUMP backup of files that have changed, and directories leading to the files, vs. backing up all dirs in past

→Depth-first file tree order traversal for merging Dump & TAR formats (traverse straight down before moving across tree)

→If DAR & RECURSIVE variables are set, recoveries of all subfolders and files under a specified directory can occur

→Incremental backups will only backup directories leading to the changed files (as opposed to the current PAX DUMP format, which uses width-first file tree traversal)

→Veritas Backup Selection window SET type=dump /ddar\_fs1

→Enable DDAR for Veritas NetBackup Server:

C:\Program Files\VERITAS\NetBackup\db\config\ndmp.cfg (add NDMP\_DAR\_DIRECTORY\_ENABLED to new file)

## **NDMP FILE FILTERING FOR BACKUPS:**

- Celerra NDMP with File Filtering allows exclusion of files and/or directories when performing NDMP backups
- Implemented in the FTS (File Traversal System—submodule used by PAX) in DART, using NASS threads
- Excludes files based on prefix or suffix set on NDMP client
- Initial support for PAX filtering only, using EMC\_EDIR0[1-5 directory filters] & EMC\_EFILE0[1-5 file filters]

### **Uses two types of filters:**

EXCLUDEDIR & EXCLUDEFILE, represented as follows:

#### **EMC\_EDIR**

#### **EMC\_EFILE**

- Filters can be pattern match or wild-cards \* and ? at beginning or end for files, and only at end for directories

EMC\_EDIR01=/fs1/dir1/temp

EMC\_EFILE01=/fs1/\*.tmp

**Note:** Only (5) directory filters and (5) file filters can be specified

## **PATTERN MATCHING AND FILTER RULES:**

### **DIRECTORIES:**

- Each filter must include path “/” with mountpoint to filesystem
- Wildcards \* or ? can only be at the end of a directory filter

EMC\_EDIR01=/EMC/dir1/temp

### **FILES:**

- No fs mountpoint path required
- Wildcard ? or \* can be at beginning or end of a filename, not in the middle

EMC\_EFILE01=\*\tmp`

## **CELLERRA NDMP FILE FILTERING RESTRICTIONS:**

- Cannot be used with NDMP VBB, only regular NDMP PAX backups

## **FILE FILTERING PARAMETERS FOR PAX:**

**param PAX filter.numDirFilter** (0=disabled; 5=default; 50=max number of directory exclusion filters)

**param PAX filter.numFileFilter** (0=disabled; 5=default; 50=max number of file exclusion filters)

**param PAX filter.caseSensitive** (0=disabled; 1=default, case sensitive)

## **CREATING FILE FILTERING JOB WITH LEGATO 7.2:**

- Enter server name in Legato Server's \Window\system32\drivers\etc\hosts file
- Legato NetWorker>NetWorker Administrator>Rightclick Windows Server>Connect to This Server>Networker Groups>Manage Groups>Groups>Create>Filtering Group
- Legator NetWorker>Client Operations>Manage Clients>Create>Name: srv2 Save set: /fs1 Group: Filtering Group
- Preferences Tab: Aliases: srv2 Storage nodes: srv2 Clone storage nodes: srv2 Ndmp: Yes
- Remote Tab: Remote access: "@" Remote user: ndmp Password: \*\*\*\*\* Backup command: nsrndmp\_save -T tar Application information: DIRECT=y; EMC\_EFILE01=\*\doc; EMC\_EFILE02=\*\tmp; HIST=y; UPDATE=y
- Create c:\bootstrap folder on Legato Server
- Legato NetWorker>Media Management>Devices>rightclick c:\bootstrap>Operations>Label: Check the 'Mount after Labeling' box
- Run manual backup: NetWorker Groups>Manage Groups>rightclick Filtering Group>Start

## **TROUBLESHOOTING FILE FILTERING ISSUES:**

EDIR or EFILE messages will be logged in Server Log

## **CELLERRA NDMP – VOLUME BACKUP [aka, VBB & RESTORES(Volume Based Backup):]**

Not a PAX-type backup. Conducts backup of used blocks at volume level, if History is set, will process metadata first. Note that this feature requires that the file system be mounted RO. 3-way VBB backups not yet supported. To conduct a VBB restore to NMFS or VDM file systems:

### **EXAMPLE:**

1. Create temp file system for the restore
  2. Create VBB restore path: \$.server\_config server\_2 -v "param vbb tempDir=/ndmpfs\_1/.vbbtemp"
  3. Allows restore to NMFS file system path /ndmpnest/abc
- NDMP, VTLU, Backup with Integrated Checkpoints, all supported
  - Idea for VBB is to increase backup performance, especially for many small files
  - NDMP VBB processes metadata first, then transfers used blocks via volume level
  - VBB uses new integrated checkpoint functionality
  - VBB performs DUMP-TAR merge using 8 threads
  - Idea is to traverse file system as fast as possible to read metadata, send file history to client while writing data to tape
  - Retrieve metadata from tape, create files and write data, set file attributes
  - VBB Full Restore overwrites everything in file system with restored data (fs must be converted to rawfs for this operation)
  - Restore of files can be done if appropriate HIST=y flag set during the Backup
  - Restores can be done at file system or individual file/directory level, but are slower than regular NDMP

→VBB Backup file system must be mounted RO or using checkpoint, or Backup with Integrated Checkpoints feature

→VBB backups up entire file system, not partials

→Does not work with DDAR or File Filtering for Backups

→VBB process based on inode order to avoid fragmentation & is multi-threaded for performance

→VBB performs Full Destructive Restore (Volume level) or File level Restore (FS level)

#### **HOW VBB BACKUP WORKS:**

→File system traversed to collect metadata

→Metadata sent to NDMP client and written to tape

→File History sent to Backup Host

→File data is written to tape in block order

#### **FULL DESTRUCTIVE or FILE LEVEL RESTORES SUPPORTED:**

→If HIST=y is done for backup, then file restore can occur

→Full restores are destructive by design, and requires that the file system be of equal size or greater, and in rawfs format

**#nas\_fs -n vbb\_bak -type rawfs -c samesize=pfs1 -pool=clar\_r5\_performance -o slice=y**

#### **FULL DESTRUCTIVE RESTORE:**

1. Have rawfs file system created & mounted

2. NetWorker User>Operation>Save Set Recover>Source Client: select Data Mover

3. Highlight Save Set to recover>Recover Options>Relocate recovered data to this path: **/celerra\_vol\_45**

**Note:** Default recover syntax is /celerra\_vol\_xx with the file system Volume ID number as the last part of the string

#### **Server Log:**

Performing a full destructive restore of filesystem ...

#### **ENABLING VBB BACKUP ON NDMP HOST:**

##### **LEGATO Backup Syntax:**

**nsrndmp\_save -Type vbb**

OTHER VARIABLES: SNAPSURE=y; VBB=y; DIRECT=y; HIST=y; UPDATE=y

##### **VERITAS & OTHERS:**

**VBB=y or set type =vbb**

→Set HIST=y if file history is desired

#### **CELLERA NDMP WITH VBB LIMITATIONS/RESTRICTIONS:**

→Backups are conducted on the entire file system only, not partial

→File system must be Read-Only (Use Checkpoint or mount file system RO)

→File Filtering and DAR/DDAR restores are not supported

→3-Way restores will not be supported

#### **NDMP BACKUP WITH INTEGRATED CHECKPOINTS (SnapSure):**

With this feature set, the Backup job will automatically create a Checkpoint, conduct the backup, and then delete the Checkpoint

→Point-in-time backups of live file systems

→Celerra NDMP Backups with Integrated Checkpoints supports regular NDMP, 3-Way NDMP, & NDMP VBB, and is based on backups of SnapSure checkpoints

→Both VBB & Restартable Backups rely on Integrated Checkpoint feature

→Set variable SNAPSURE=y to allow for creation of Checkpoints (NDMP\_CREATE\_AND\_MOUNT\_CHECKPOINT\_EVT)

→Compatible with PAX & VBB backups, and NDMP Versions 2, 3, 4, File Filtering, DDAR

→Checkpoints are created before the backup job and automatically mounted under rootfs in a directory called /.automaticNDMPCKpts (or Snaps)

→Drawbacks? Older checkpoints can create a read performance hit, use no more than (4) checkpoints for pfs

→/.etc/backupSnapDB will be the database for checkpoints for NDMP

→Checkpoints will be deleted after the successful backup (DART will send

NDMP\_UNMOUNT\_AND\_DELETE\_CHECKPOINT\_EVT to CS)

→NDMP Integrated Checkpoints reside in the PFS

automaticTempNDMPCKpt6412 [example of temporary NDMP Checkpoint name]

→NAS 5.5 maintenance release introduces Snapsure environmental variable for those sites where the Backup software does not yet have a variable to set

**# server\_param server\_3 -facility NDMP -info snapsure -v**

server\_3 :

name = snapsure

facility\_name = NDMP

default\_value = 0

current\_value = 0

configured\_value =

user\_action = none

change\_effective = immediate

range = (0,1)

description = Use SnapSure file system for backup

**Note:** Set value to 1 to allow Backup software, such as TSM, to be able to conduct Integrated Backups using Checkpoints

### **NDMP RESTARTABLE BACKUP:**

→ Purpose would be to allow for a restart of a failed backup, as long as the point prior to failure is known to be good

→ Works by using a backup context based on a checkpoint of pfs prior to backup, and deletes it after completing—DMA will restart by accessing NDMP\_MOVER\_GET\_STATE values in the checkpoint, via the BRE (Backup Restart Extension) interface

### **iSCSI LUN REPLICATION:**

→ Providing for asynchronous remote point in time copies of LUNs—copies difference between two local snaps to remote Celerra

→ Snaps managed by Hosts, as hosts own iSCSI LUNs, and snap functionality exported to host via NBS protocol.

→ Host runs the CBMAPI (Celerra Block Management API) NBS client and provides API to ERM

→ Architecture requires Host and two data movers

→ ERM is user interface to manage iSCSI snaps and replication of snaps

→ CBMAPI is extension to SNAPAPI

→ Only CLI can be used for iSCSI replication in this release

## **CELERRA REPLICATOR FOR iSCSI (iSCSI-COPY with RM/SE 3.1.0.1):**

**Note:** GA NAS 5.5, Celerra Replication Manager/SE

→ Application consistent point-in-time replication of iSCSI LUNs using VSS technology

→ Replication of LUNs based on snaps (snaps are consistent point-in-time copies), asynchronous using RM/SE for Celerra

→ iSCSI Luns replicated to local, loopback, or remote Celerra in support of Disaster Recovery solution

→ Replication is asynchronous and based on iSCSI snaps managed by RM/SE for Celerra (config, jobs, policies)

→ Interface used will be the Windows RM/SE for Celerra to ensure applications such as SQL, Exchange, are consistent

→ Replicator will perform Baseline copy or Differential copy when replicating from Source to Target

→ Can replicate from single source LUN to multiple target LUNs

→ Production LUN remains available during replication copy session

→ Maximum number of concurrent sessions per DM is 128

### **APPLICATIONS SUPPORTED FOR iSCSI REPLICATION:**

Exchange 2000, 2003 via VSS, SQL Server 2000, 2005, NTFS

### **SUPPORTED PLATFORMS FOR RM/SE:**

Windows 2000 and Windows 2003

### **CELERRA REPLICATOR FOR iSCSI LIMITATIONS/RESTRICTIONS:**

→ Only Windows 2000/2003 support

→ No Celerra Manager configuration for destination LUN

→ Support for up to 128 active concurrent sessions

→ Manual failover & failback

→ Manually extend Source Lun, destination Lun extended using DP Mgr

### **KNOWN ISSUES:**

If iSCSI LUNs are being managed by RM/SE, and the system is upgraded from NAS 5.3/5.4 to 5.5, a code problem causes the iSCSI LUNs opaque file (attributes) to grow greater than 8k, resulting in loss of access during the Upgrade iSCSI conversion process on snap file systems. See emc158502 AR93043 for more details.

VCS: 3: Attribute file format error, fsid=60, path=fs60\_T4\_LUN1\_000187910099\_008D.vs

name=fs60\_T4\_LUN1\_000187910099\_008D.ckpt044.opaque offset=8196 status=OK

### **COMPONENTS THAT COMPRIZE CELERRA REPLICATOR FOR iSCSI:**

RM/SE for Celerra (ERM); CBMAPI (SNAPAPI); NBS Server; DP Manager; File Version Delta; RCP Transport; DART Interconnect (DIC); VersionSet/Version (VCS); Control Station

**CBMAPI**→Windows API that exports Celerra Block Management for Snapshot, Replication, OSL, runs as NBS Client on CS and is used by RM/SE. From CLI, SNAPTEST.exe is used to connect to API.

**NBS Server**→Enhanced to support Replication & Snapshots, processes request from CBMAPI NBS Client, and calls DP Service. Packets are in XML format with 128-byte header

**DP Manager**→DpInit—Starts DP service; DpConfig—XML encoding/decoding; DpContext—keeps track of replications;

DpService—interfaces for managing replication; DpPolicy—policy engine; DpTaskManager—handles failover and restores from remote snap; DpTunnel—tunneling service for NBS; DpCopier—sends and receives data

### **/nas/server/slot\_\*/eof**

dpinit [dpinit service is started after the httpd secmac service]. Dpinit always starts during failover operations, but if commented out in the eof file, will not start during system reboots.

**RCP**→data transport service

**DIC**→Dart Interconnect—XML over HTTP uses digest authentication

**Note:** The Dart Interconnect layer will use HTTP client and server protocols to exchange XML packets between Darts—Phase I Burgundy, Phase II Cognac, at which time DART will run an HTTP Server.

VCS→Version Control System—interface for snapshot operations, maintains on-disk and in-memory databases

### **THREE TYPES OF REPLICATION:**

Local→replication between two data movers on local Celerra system

Loopback→replication on same data mover

Remote→replication between data mover on Source CS to data mover on Target CS

### **Replication Control Management:**

RM/SE for Celerra; SnapApi; Control Station

→Used by RM/SE & Host applications to communicate with NBS Server on DART

→Windows API exports block management function for Snapshot, Replication

### **Server Components:**

NBS Server; DP Manager;FileVersion Delta; RCP Transport; DART Interconnect (DIC); VCS (Version Control System)

→NBS Server is enhanced for snapshot functionality

→Management records require upgrading, and are stored in /nasmcd/rootfs/slot\_x/.etc/nbsdb

→DP Manager is the iSCSI replication engine, runs on both Source & Target servers

→dpinit is included in boot.cfg to recover replication sessions during server reboots

→Databases reconstructed into memory after reboots using on-disk databases /nasmcd/rootfs/slot\_x/.etc/dp

→Only one snap can be running at a time, but can have 128 active sessions

→File versioning based on ufs extents (extent tables and data versions)

→RCP (Replication Control Protocol) transport used as data transfer service over TCP/IP (same service as used in fs\_copy & IP Replicator)

### **RCP Transport Parameters:**

#### **\$server\_param server\_2 -f RCP -m tcpwindow -v 1048576**

**Note:** Default tcp window size is 128kb—above example shows setting param to 1MB. Window\_Size = Round Trip Delay \* Desired Rate. This setting is the HWM value.

#### **param RCP allowipX=<ip address>**

**Note:** Use above param to set 1-8 IP addresses (X) for access control

#### **\$server\_param server\_2 -f RCP -m tcpwindowlowat -v 0 | 614400** (600kb)

**Note:** Default value 0 = 64k. Low\_Watermark = Window\_size - #Bytes ACK'd when TCP transmit stops

DIC→Dart Interconnect uses XML over HTTP for communication. Digest authentication using shared secret & MD5 hash. Nas\_cel is used to configure password/passphrase on both Source/Target

VCS→Manages snapshot operations via File Version & File Version Set objects

→File Version is abstraction for storage object file

→File Version set is a collection of objects stored in directory

Control Station→New command nas\_replicate. Other commands modified, nas\_cel and server\_iscsi

### **iSCSI COMMAND SET:**

→nas\_replicate -list | -info | -abort id=<sessid> -local | -failover id=<sessid>

→nas\_cel -name eng1 -create 192.168.2.30 -passphrase xtratree

→server\_iscsi server\_2 -lun -n 10 -create mytarget -size 100M -fs fs1 -readonly (creating RO destination LUN)

### **CONFIGURING iSCSI LUN REPLICATION:**

#### **Overview:**

Setup iSCSI source Lun

Create RO destination Lun of same size as source

Configure authentication Source & Target

Login to source target from iSCSI initiator

Create iSCSI-Copy job with destination IP address and authentication name

Create ERM schedule (RM/SE) to create local snap and mark snap for replication

#### **STEPS:**

1. Setup communication link and trust relationship between Local & Remote Control Stations:

# nas\_cel -name cs\_remote 10.241.168.52 -passphrase replication (repeat on Remote CS for cs\_local)

2. Create samesize destination iSCSI LUN for replication (Source LUN already exists):

# server\_iscsi server\_2 -lun -number 50 -create t1 -size 40G -fs remote\_fs -readonly yes

# server\_iscsi server\_2 -lun -number 50 -create t1 -size 40G -fs remote\_fs -vp yes -readonly yes [virtually provisioned fs]

3. Configure iSCSI LUN Replication Session:

From Windows host, log into Celerra iSCSI target for Source iSCSI LUN, create iSCSI Copy job using RM/SE and specify the destination Data Mover IP address. Create an RM/SE job for snapshot creation and mark snapshot for replication

### **VERIFYING REPLICATION ROLE FOR LUNS AND REPLICATION SESSIONS:**

**\$ server\_iscsi server\_2 -lun -info -all**

**\$ nas\_replicate -list**

**\$ nas\_replicate -info -all**

**SETTING RCP PORT NUMBER:**

**\$ .server\_config server\_2 -v "param RCP port=20000"**

**Note:** Default port 8888

**CHANGING TCP WINDOW SIZE:**

**\$ .server\_config server\_2 -v "param RCP tcpwindow=65536"**

**Note:** Default value 128kb

**STEPS FOR PERFORMING iSCSI SNAP REPLICATION:**

→data written to iSCSI lun

→iSCSI snap is taken from RM/SE console on Source side (S1)

→Snap is marked for replication

→Baseline copy of Source snap replicated to Target side

→Snap is made of Target side iSCSI lun (D1)

→Source & Target snaps become the ‘common’ base

→as data changes on Source, the process is repeated with creation of new Snap on source, replication to destination iSCSI Lun, new Destination Snap created, and new ‘common’ base becomes S2 and D2

**Note:** Destination lun is RO

**FAILING OVER iSCSI REPLICATION SESSION:**

→Failover initiated: **\$nas\_replicate -failover -id=<sessid>** (source becomes RO and destination RW)

**Note:** Destination iSCSI LUN takes over as PLU—initiated from remote CS only. See Using Celerra Replicator for iSCSI Tech Module--there are many steps involved in actually performing a complete failover.

**ABORTING AN iSCSI REPLICATION SESSION:**

**\$ nas\_replicate -abort id=<sessionID> | -local**

**Note:** Might need to do this when other side is unavailable or RM/SE is unavailable. This command is run on the local CS to abort either the local session only or both sides of the session.

**RESTORING FROM DESTINATION TO SOURCE WITH COMMON BASE SNAPS:**

→Differential snap taken on Destination and delta between D2 and D3 is copied back to Source, which becomes RO

**RESTORING FROM DESTINATION TO SOURCE WITHOUT COMMON BASE SNAPS:**

→Baseline copy of destination snap is copied back to Source, which is RO

**Troubleshooting:**

\$nas\_replicate -info shows configured parameters, replication states, and control/data states

Replication Log Events (V2RPL & DPSVC)—Lun full, network down, db access error, other errors

Disable replication by removing dpinit and eof from boot.cfg file

Replication database found in /etc/dp/replica, /etc/dp/policy, /etc/dp/version, /etc/dp/info, /etc/dp/taskManager

**Control Station commands to use in troubleshooting iSCSI snaps, etc.:**

**# /usr/local/bin/nbs-snap**

Usage: nbs-snap -a action [-i PluNbsId] [-n snapId] -s serverName [-v]

nbs-snap -a listPLU -s serverName

nbs-snap -a getPLUAttr -i PluNbsId -s serverName

nbs-snap -a listSnap -i PluNbsId -s serverName

nbs-snap -a create -i PluNbsId -s serverName

nbs-snap -a delete -i PluNbsId -n snapId -s serverName

nbs-snap -a getSnapAttr -i PluNbsId -n snapId -s serverName

nbs-snap -a getAppData -i PluNbsId -n snapId -s serverName

nbs-snap -a createTWS -i PluNbsId -n snapId -s serverName

nbs-snap -a deleteTWS -i PluNbsId -n snapId -s serverName

nbs-snap -a promote -i PluNbsId -n snapId -s serverName

nbs-snap -a demote -i PluNbsId -n snapId -s serverName

nbs-snap -a restore -i PluNbsId -n snapId -s serverName

nbs-snap -a commitRestore -i PluNbsId -n snapId -s serverName

nbs-snap -a undoRestore -i PluNbsId -n snapId -s serverName

nbs-snap -a refresh -i PluNbsId -n snapId -s serverName

nbs-snap -a setAppData -i PluNbsId -n snapId -s serverName -d <data>

nbs-snap -a deleteALL -i PluNbsId -s serverName

nbs-snap -a getSN -i NbsId -s serverName

nbs-snap -a getNBSIdbySN -i SN -s serverName

nbs-snap -a asyncCreate -i PluNbsId -s serverName

```
nbs-snap -a asyncQuery      -i infoId -s serverName  
nbs-snap -a deleteCookie    -i infoId -s serverName
```

# **/usr/local/bin/nbs-snap -a listPLU -s server\_2**

```
fs35_T15_LUN0_APM00042403638_0000  
fs35_T15_LUN1_APM00042403638_0000
```

**Note:** Listing Production LUNs on the Data Mover

# **/usr/local/bin/nbs-snap -a listSnap -i fs35\_T15\_LUN0\_APM00042403638\_0000 -s server\_2**

```
fs35_T15_LUN0_APM00042403638_0000.ckpt000
```

**Note:** Listing specific snap for the PLU

# **/usr/local/bin/nbs-snap -a delete -i fs35\_T15\_LUN0\_APM00042403638\_0000 -n**

```
fs35_T15_LUN0_APM00042403638_0000.ckpt000 -s server_2
```

Delete snapshot succeed.

**Note:** Deleting a snapshot for the PLU

### **MORE EXAMPLES:**

\$ **server\_iscsi server\_2 -lun -info 4**

server\_2 :

Logical Unit 4 on target GHKEXP\_iscsi0:

(Production) fsid=54 size=200000MB alloc=1814MB dense read-write

```
path=/Enertia/fs54_T1_LUN4_APM00083804182_0000/fs54_T1_LUN4_APM00083804182_0000 (snapped)
```

replication=source

max\_extension\_size=128049MB

Healthy

# **/usr/local/bin/nbs-snap -a listSnap -i fs54\_T1\_LUN4\_APM00083804182\_0000 -s server\_2**

```
fs54_T1_LUN4_APM00083804182_0000.ckpt000_809208850002168
```

```
fs54_T1_LUN4_APM00083804182_0000.ckpt000_809208850077915
```

# **/usr/local/bin/nbs-snap -a getSnapAttr -i fs54\_T1\_LUN4\_APM00083804182\_0000 -n**

```
fs54_T1_LUN4_APM00083804182_0000.ckpt000_809208850002168 -s server_2
```

nbsSnapGetAttrRply:

BlkSize -> 512

NumBlks -> 409600000

TWSName -> NONE

dr -> 1

LUN -> -1

CreateTime -> Sat Dec 20 19:40:36 2008

target ->

AppLabel -> ReplicationV2

\$ **server\_iscsi server\_2 -snap -list -target GHKEXP\_iscsi0 -lun 4**

server\_2:

| Snap Name                                                 | Lun Number | Target        | Create Time                  |  |
|-----------------------------------------------------------|------------|---------------|------------------------------|--|
| fs54_T1_LUN4_APM00083804182_0000.ckpt000_60627073788771 4 |            | GHKEXP_iscsi0 | Wed Dec 17 00:29:00 CST 2008 |  |
| fs54_T1_LUN4_APM00083804182_0000.ckpt000_60627073714949 4 |            | GHKEXP_iscsi0 | Thu Dec 18 22:41:50 CST 2008 |  |

## **CELERRA iSCSI DR SOLUTION:**

→Configuring Disaster Recovery using iSCSI and Celerra Replicator

→Failing over to DR site

→Preparing Failback

→Failing back to Source site

### **Requirements:**

NAS 5.5, Windows 2000/2003 Servers running hotfixes KB891957/KB898790, 2.0 Initiators, Solutions Enabler 6.2, and Replication Manager/SE 3.1.0.2

### **CONFIGURING DR:**

#### **Setup Source & Remote Side Windows 2000 Servers:**

→Windows 2000/2003 Server with hotfixes, KB891957 & KB898790, install MS Initiator 2.0, install Solution Enabler v6.2, install Replication Manager/SE 3.1, install applicable NTFS, SQL, or Exchange application

→Setup Replication relationship & sessions between Local & Remote Celerras

### **CONDUCTING FAILOVER:**

- Failover using nas\_replicate –failover, log into destination Windows Server--mask luns for access and mounting, clearing the readonly attributes using bat file against device and drive number: #RmMountVolume.bat 2 R:
- Change RMSE database location on the remote Windows Server to point to the location of db on production system
- Update SQL or Exchange db hostname entries
- Enable DR mode on Host Windows system and restart RM/SE
- Recover replicas from RM/SE Host by selecting Restore for applications

#### **FAILING BACK:**

- Assuming true disaster, rebuild Windows 2000 Server on original Source & setup Replication sessions for applications
- Take remote Windows Server offline and prepare applications for failback
- Record current replication session information and perform failback using nas\_replicate –failover
- Make LUNs visible to original Source Server, mount volumes, clear readonly flag from device & drive letter
- Change RMSE database location to source Windows Server and update application db pointers
- Enable DR mode on Host and recover replicas, then recover applications

## **CELLERRA MANAGER ENHANCEMENTS:**

### **iSCSI WIZARD:**

iSCSI Target Wizard → Creating new iSCSI target LUN using Target name and portals

iSCSI LUN Wizard → Creating new iSCSI LUN or multiple LUNs of same size by selecting or creating target, selecting or creating file system, configuring LUN masks, and other settings if necessary

**Note:** Enable multiple access for clustered Servers during LUN Masking setup page. You might use this feature for Microsoft Cluster Services for Exchange, for example, in which each Cluster Server must share the external storage. MSCS nodes will have access to the LUN in case of failover if multiple initiators are listed here.

### **MaxRequestHoldTime:**

The default HoldTime for the MS Initiator is 60 seconds, and is defined as the time in seconds that iSCSI requests will be queued if a connection to the Target is lost, and the connection is being retried. After the Hold period expires, further requests fail with “error no device” and the disk is removed from the Windows host list. Increasing the MaxRequestHoldTime value to 600 seconds, or 10 minutes, is recommended for VMWare and Exchange/SQL iSCSI environments where requests may need more time:

**HKLM>System>CurrentControlSet>Control>Class>4D36E97B-E325-11CE-BFC1-**

**08002BE10318>0000>Parameters>MaxRequestHoldTime = 600** → Increased value from 1 to 10 minutes

### **CREATING AN iSCSI LUN:**

1. Create Target first [unique alias name & portal]
2. Select File System to hold LUN
3. Enter LUN information and configure LUN Masking (initiators added to grant access to LUNs)
4. Configure CHAP access on initiators
5. Set iSCSI/iSNS settings

### **MS iSCSI INITIATOR VERSIONS:**

Version 1.06=Build 302

Version 2.02=Build 1895

Version 2.03=Build 3099

Version 2.04=Build 3273

Version 2.05=Build 3392

Version 2.06=Build 3497

Version 2.07=Build 3640

Version 2.08=Build 3825 (latest—Vista & Windows Server 2008 comes pre-installed with the iSCSI Initiator software)

### **REPORT LUNS:**

→New SCSI-3 specification used only with Symm & Clariion arrays to speed fibre channel discovery of luns, channel failover, and channel failback

### **NAS VPL (VIRTUALLY PROVISIONED LUNS--THIN PROVISIONING):**

A marketing concept that calls for sparse luns for regular NAS file systems and iSCSI file systems—also known as thin provisioning, over-provisioning, etc., with claims of lowering TCO & yielding greater availability for applications, though the reality will most likely become a support millstone. Concept was introduced in order to compete with NetApps, though NetApps can logically shrink file systems to reclaim space, whereas EMC’s solution cannot.

→Client systems see max logical size of file system, which is not the actual size from DM perspective

→Requires that auto-extension feature be used and a Max size be specified

### **iSCSI VIRTUAL PROVISIONING LUN (aka sparse lun):**

**Note:** Purpose of VPL is to create largest LUN possible, but underlying file system can be much smaller, and extended when needed. Purpose is to not use disk space until required. Auto-fs extend fits nicely into this feature, but IS NOT a guarantee against data loss. It is very possible to fill a file system faster than the auto-extend feature can extend, since VPL does not use PBR to reserve blocks.

#### **BEST PRACTICES FOR iSCSI VPL LUNS:**

- Use auto-fs extend feature
- Monitor fs capacity and logs
- Hosts do not need to rescan storage when underlying file system is extended, no disruption to service
- Formerly called iSCSI Sparse Lun, basically means you can create a Lun >in size than the actual file system underneath
- Max size for VPL LUN is still 2TB less 1MB
- As best practice, always use auto-fs extend feature with VPL
- Do not mix DENSE and VPL luns on the same file system
- LUNs have space allocated as blocks are used, and file system blocks are allocated when a chosen HWM is reached
- Main drawback, however, is that allowing a VPL file system to fill to 100% will cause data corruption/loss (no PBR used)
- Can only create the VPL lun via CLI, but can list, delete, extend via GUI
- Can list, delete, extend VPL via Celerra Manager
- Sparse luns do not need to reserve blocks when extended

#### **CREATING iSCSI VPL FROM CLI:**

```
# server_iscsi server_2 -lun -number 0 -create target1 -size 10G -fs sparsefs1 -virtually_provisioned yes  
$server_iscsi server_2 -lun -create t1 -0 -size 1000M -fs fs_iscsi -vp yes  
$server_iscsi server_2 -lun -info 0 -target t1  
    max_extension_size=537794MB
```

→ Can create virtually provisioned temporary writable snaps (TWS) for regular Luns only (aka Dense Luns), used for Snaps

#### **Set following param:**

**param nbs sparseTWS=1**

#### **DYNAMIC iSCSI LUN EXTENSION:**

- Supports dynamic extension of regular iSCSI luns and virtually provisioned iSCSI (PLU) LUNs by increasing visible size to Hosts
- Total LUN size cannot exceed 1MB < than 2TB
- Celerra iSCSI target allows lun to be created in dense or sparse mode
- Do not use both dense & sparse LUNs in same file system
- Dense luns required block reservations when extending
- Extend using CLI or Celerra Manager
- Celerra Replicator for iSCSI LUNs will extend destination side first, then source side
- Cannot be used for extending promoted Snaps or Destination IP Replication iSCSI luns

#### **DYNAMIC iSCSI LUN EXTENSION LIMITATIONS/RESTRICTIONS:**

- All dynamic iSCSI lun extensions require specific host-side actions to actually extend the volume and file system for the Host
- Cannot increase LUN size past that which the file system can support
- Max extension size of VPL lun is max LUN Size (2 TB) less current LUN size
- Max extension size of DENSE LUN limited to free space in the file system itself

#### **Extension Rules:**

- With iSCSI replication, destination LUN will be extended after the Source when the Data Protection service detects LUN change
- Note:** Destination LUN cannot be extended from Control Station for iSCSI replication (Only the DP service)
- Can extend regular iSCSI Lun via CLI or GUI, but for active iSCSI Sessions, the actual rediscover should be done offline for both Windows and Linux
- Promoted Luns cannot be extended
- Cannot shrink Lun size once extended
- Cannot restore from old snapshots once Lun has been extended

#### **Querving iSCSI LUN to Determine Extension Size:**

```
$ server_iscsi server_2 -lun -info 0
```

Logical Unit 0 on target t1:

(Production) fsid=40 **size=1000MB** alloc=0MB dense

**max\_extension\_size=537794MB**

#### **Extending iSCSI LUN from CLI:**

```
$ server_iscsi server_2 -lun -extend 0 -target t1 -size 1000 (MB)
```

#### **EXTENDING iSCSI LUN FROM WINDOWS HOST:**

1. Extend iSCSI Lun from Control Station
2. Run rescan for extended Lun size using Disk Manager or CLI diskpart “rescan” command
3. Unmount file system on Data Mover
4. Extend volume using CLI diskpart “extend” command
5. Remount file system

### **Windows Diskpart Utility:**

- Diskpart extends partitions by restriping, without downtime
- Does not support Dynamic partitions or Extended partitions, only NTFS basic primary partitions.
- Supports Windows 2000/2003

### **USING DISKPART TO EXTEND CELERRA iSCSI LUNs:**

1. Download diskpart\_setup.exe from Microsoft and install on Windows host
2. c:\>diskpart.exe
  - DISKPART>list disk
  - DISKPART>rescan
  - DISKPART>select disk x
  - DISKPART>list partition
  - DISKPART>list disk
  - DISKPART>rescan
  - DISKPART>select partition x
  - DISKPART>extend

**Note:** iSCSI LUNs for Exchange Server databases should have an offset of 64, which equals 32KB

### **EXTENDING iSCSI LUN FROM LINUX HOST:**

1. Extend iSCSI Lun from Control Station
2. Allow Linux to discover extended Lun by restarting the iSCSI service
3. Unmount file system on Data Mover
4. Use fdisk to recreate the newly extended partition
5. Extend the file system using #resize2fs command
6. Remount file system

### **CELLERRA iSCSI OSL API SUPPORT(Open Storage Library):**

Collection of block storage API's used to manage heterogeneous storage arrays while presenting a common storage information model and programming interfaces, Open Storage APIs v 5.4. Celerra will use Celerra Block Management API (CBMAPI)

- Already supported by Clariion and Symmetrix backends
- Supports Microsoft VDS (Virtual Disk Service) v.1.1 & VSS functionality
- Storage APIs will support storage discovery, LUN provisioning, LUN masking
- Various enclosures are DAEs, SPEs, DPEs, DMEs; Klondike DPEs are SATA disk drives; Katana for FC, and Katina for 4GB FC
- Debug Log module for STORAPI is called LOG\_OSL

### **RESTRICTIONS:**

- Only supports iSCSI objects

### **SNAP MANAGEMENT FOR OSL:**

- NbsStorSnapDisk—snapshot for production Celerra iscsi drive
- NbsStorDeleteSnap—delete a snapshot
- NbsStorQuerySnaps—list of existing snapshots
- NbsStorPromoteSnap—promote a snapshot as RW and assigns a LUN
- NbsStorDemoteSnap—demotes a snapshot & unassigns the LUN
- NbsStorRestoreDisk—restore a snapshot
- NbsStorUndoRestoreDisk—undo a restoration, step 1
- NbsStorCommitRestoreDisk—undo a restoration, step 2
- NbsStorDeviceCreate—create LUN and associate with target
- NbsStorDeviceDelete—delete LUN
- NbsStorDeviceExplan—expand a LUN
- NbsStorDeviceSizeInfo—LUN inquiry

### **DYNAMIC VOLUME CREATION:**

In current implementation, CS is used to issue command to create a new volume to DART, and DART responds by creating volume with CTL. In new implementation, DART will automatically create new basic volumes when it bus scans and finds new devices on the backend. Purpose is to provide for better path failover, better handle configuration changes, and to avoid mismatches.

- CS will still be required to ‘discover’ the volume and matchup to a CTL (must know either dvolume name or CTL for CS command)

→Bus scan is done on Server reboot, link-up event, or by server\_devconfig command

### **AUTOMATIC FILE SYSTEM AUTO EXTENSION:**

- Defined by HWM from 50-99%, default HWM is 90% (nas\_fs -hwm <50-99%>)
  - Must define the max extension limit of the fs using nas\_fs -max\_size <integer>[T|G|M] (default=10GB)
- Note:** Generally, once HWM reached, fs will extend by 5% of fs size if greater than 10GB, or 10GB, whatever is larger. Extensions of 5GB will occur again if fs usage is less than 3% below HWM.
- Auto-extension is only supported for file systems created or modified using AVM (uxfs, mgfs, FileMover)

→Supports auto-extension for iSCSI fs using Virtually Provisioned LUNs, but NOT regular iSCSI LUNs

→Supports auto extension of destination, then source file system for IP Replication or FS Copy

→Supports auto extension of Timefinder PFS, then mirrored snaps

→DART will generate the event at the HWM, Control Station will catch with event handler to extend file system

→Virtual provisioning allocates storage but it is not actually used by file system until needed

**\$ nas\_fs -name auto\_noprovision\_fs1 -c size=10G pool=<pool\_type> -auto\_extend yes -hwm 90% - max\_size 2G -o slice=y**

**Note:** Creating auto extend file system without virtual provisioning and max extend size of 2GB

**\$ nas\_fs -name auto\_provision\_fs1 -c size=10G pool=<pool\_type> -auto\_extend yes -vp yes -hwm 90% - max\_size 2G -o slice=y**

**Note:** Auto extend file system with virtual provisioning set and max extend size of 2GB

**\$ nas\_fs -modify fs1 -auto\_extend yes -hwm 50%**

**\$ nas\_fs -i fs1**

auto\_ext = hwm=50%,virtual\_provision=no

**Note:** Default behavior is to not auto-extend unless “-auto\_extend yes” is specified (set by options in nas\_fs)

→Feature interoperates for replication, timefinder, filermove, snapsure, but not MirrorView, SRDF, VDM

## **AUTO FS EXTENSION RESTRICTIONS/LIMITATIONS:**

→Does NOT support auto-extension of regular dense iSCSI lun file systems

→Does NOT support auto-extension for checkpoint file systems, nmfs file systems, or raw file systems

→Does NOT support auto-extension of file systems that are used with SRDF or MirrorView

→Once allocated, unused space cannot be reclaimed

## **HTTP CLIENT FOR DART:**

→HTTP Client 1.1 is necessary to support FileMover and DART Interconnect features (Dart-to-Dart comms)

→Typical exchange is that Client opens TCP connection to Server with request; Server sends back response; connection closed

→Multiple requests can be sent for same TCP connection, Client & Server can authenticate using digest or basic authentication

→Client will not be a Web Browser (which interprets body data), but will allow for sending/receiving messages, and connection parameters

→Choices of Security: No security; Basic authentication; Digest authentication

## **NFS VERSION 4:**

See NFS Protocol section of TechGuide

## **NAS 5.5 POST GA RELEASES--NAPA:**

NAPA is used to indicate various maintenance releases for NAS 5.5 that include special features, fixes, etc.

### **NAPA 1 GA 9 June 2006 NAS 5.5.21.4:**

**NAPA 1 SUPPORTED ITEMS:** Main purpose of this release is to support the CX3 Clariion hardware/flare versions

→4Gb Fibre Channel UltraPoint (Stiletto) DAE3P 3U drive enclosures (5.25”)

→Support in Celerra Manager for new CX3 hardware components

→10Gb Ethernet & 4Gb FC support

→FLARE 22 Atlas support, needed to support new CX3 arrays, 4Gb Stiletto DAE3Ps, with NAS NaviCLI & Widesky updates

→Support for new 4Gb Stiletto DAEs, Jackhammer, Sledgehammer, & Hammerhead arrays, for CX3-20, CX3-40, & CX3-80 series

→Automatic Log & Dump Collection and Transfer tool (Introduced NAS 5.5.21.4 & 5.4.25.1)

A feature that automatically collects and gzips logs or dumpfiles to /nas/var/log or /nas/var/dump, respectively, or if configured, can transfer the dumpfile or logs directly to a specified FTP server. Only auto-collect enabled, transfer piece must be enabled separately. See emc135846.

→New UFS ACL DB Cache mechanism

**Note:** The “old” ACL cache will still be used regardless of the initialization of the “new” ACL cache. If the “new” cache is initialized, there will be two levels of ACL caches. The “old” ACL cache will always be checked first for lookups in the write path. If there is a cache miss in the “old” cache, and the new cache is initialized, then the lookup will be done with the new ACL cache.

→Default AVM stripe size changing from 8kb to 32768 for Clariion Storage Pools (clar\_r5\_performance; clar\_r5\_economy; clar\_r1)

**Note:** This means that a file system that is extended may use the 32k stripe depth if new volumes are added to the pool—according to Eng., this is not a problem

### **DAE3P Disk Enclosure:**

(15) Low profile 2/4Gbps FC Drives using new Frumon-based spinup, and (2) 4Gbps LCC Cards

**Note:** DAE3P can use either FC or SATA drives

### **Atlas 22 FLARE Support:**

New Bus architecture, Full SW replication, more reserved space on Vault drives, SP reboot and shutdown from GUI, 2Gb/4Gb speed changes, Auto Dump analysis

**NAPA 2:** NAS 5.5.22.2 GA August 11, 2006

→Hammerhead NSX Generation II hardware (H2G2) XBlade-65 Data Movers with 10GbE and 4GB FC support

→Ultrapoint Stiletto 4Gb DAE fibre channel drive support with Atlas flare support [3.22.020.5.0xx]

→NS40 introduction & support for Sledgehammer arrays

→Single port 10GbE Neterion Xframe Optical NIC I/O Module support

→4Gb Fibre Channel Support

→DUDL thread deadlock detection for NFS & CIFS

→500GB LCFC drive support introduced for Gateway models (NS500, 700, NSX, NS40, & NSX2)

**Note 1:** LCFC drives are not sold with NS20 systems, therefore are not supported on that platform

**Note 2:** Though documentation and Release Notes only talk of Clariion support, DMX-3 symmetrix also supports LCFC drives. This is seen in the following cryptic clip from eLab Navigator support matrix:

15. Symm DMX-3

\* All shipping drives are supported

#### **Symmetrix DMX-3 ...Low-Cost Fibre Channel Drives:**

Technical Note p/n 300-004-737 discusses performance characteristics related to use of 500GB LCFC drives

→LCFC drives can be used for performance-insensitive, sequential read/write (such as Backup-to-disk), or low access density applications

→LCFC should not be used as R2 or BCV in write-intensive environments, where random IO performance is important, or disk response is important, etc.

#### **NAS NS40 GATEWAY & INTEGRATED SERIES:**

→Min. NAS 5.5.22.2 code and CLARiiON CX3-40 array with Flare 22

→Gateway & Integrated models available

→Available as factory racked (factory-install, configuration wizard setup) or field mounted (fresh install)

→Does not contain UPS

→Dual 2.8GHz P4 CPUs, 4-8 GB memory, 1, 2, or 4Gb Fibre Channel speeds, (4) GigE Ethernet ports

→Single or dual-Data Mover configurations available based on XBlade technology, but only Single Control Station allowed

→Based on Sledgehammer hardware, DM blades are dual Intel 2.8Ghz Pentium IV processors with 4GB memory & 800MHz fsb

→Capacity per Blade is 16TB fibre channel drives, 32TB mix fibre channel & ATA

→SAN Gateway versions can connect to most array backends, while Integrated model connects only to Sledgehammer (CX3-40)

→Projected replacement for NS500 & NS700 models based on Clariion CX3-40 platform

**Note:** An NS40 with (2) Data Movers is considered equivalent to NS700 with (2) DMs. If an NS700 has (4) DMs and requires larger disk capacities, then the NS80 would be the replacement for that level of NS700

→1U Enclosures, 1 or 2 blades only (Data Movers), single Chivas Control Station (2GB memory), (4) Power Supplies, Internal Management Switch, (2) serial ports for DART console—COM1 & COM2, 40GB ATA drive, CD-ROM, Floppy

→Chivas Pentium IV 2.8GHz Control Station with 2GB memory; (1) 10/100/1000 GbE controller; (2) 10/00/1000 Mb/s Ethernet LAN ports; (1) 10/100 Mb/s Ethernet port; 80Gb harddrive; (2) Serial ports, one for CallHome modem and one for Laptop connection

→NS40 Blade with NAS personality will have dual 2.8GHz Nocona P4 processors, 4GB memory, (2) serial console ports on lower left side, (2) RJ45 connectors to Internal Management switch, (4) Broadcom 5704 Gigabit Ethernet Optical ports (cge0-cge3) [or (2) Copper fge ports & (2) Optical cge ports], (2) Agilent 1/2/4Gb Aux Fibre Ports, and (2) Agilent 1/2/4Gb BE Fibre Ports

→SAN blades for SP's can have either (2) or (4) Fibre ports each, while NAS blades have only (2) Fibre ports

→Internal Management switch is built into Motherboard, as opposed to NSX, where Management Switches are FRU's, with (2) RJ45 ports for access—switch monitors Fan modules, peer management switch function, Data Mover status registers, power & reset states of DMs

mgmt\_2\_3 192.168.1.50 Enclosure 0 management switch A

mgmt\_2\_3b 192.168.2.50 Enclosure 0 management switch B

spa 192.168.1.200 Integrated BE SPA on Network A Gateway 192.168.1.100 [to primary CS on primary network]

spb 192.168.2.201 Integrated BE SPB on Network B Gateway 192.168.2.100 [to primary CS on secondary network]

#### **DAE3P ENCLOSURE ADDRESSING FOR CX3-40 ARRAY:**

| Type, Loop Address | Description          | Enclosure Address (aka EA or Enclosure ID) |
|--------------------|----------------------|--------------------------------------------|
| DAE3P; 0,0         | First Encl           | 0                                          |
| DAE3P; 1,0         | 2 <sup>nd</sup> Encl | 0                                          |
| DAE3P; 0,1         | 3 <sup>rd</sup> Encl | 1                                          |
| DAE3P; 1,1         | 4 <sup>th</sup> Encl | 1                                          |
| DAE3P; 0,2         | 5 <sup>th</sup> Encl | 2                                          |
| DAE3P; 1,2         | 6 <sup>th</sup> Encl | 2                                          |
| DAE3P; 0,3         | 7 <sup>th</sup> Encl | 3                                          |
| DAE3P; 1,3         | 8 <sup>th</sup> Encl | 3                                          |

#### **Other NAPA 2 Items—5.5.22.2:**

→Six new checks added to PUHC, Error, Warning, Info

→**500GB capacity, 2GB speed, 7200RPM Low Cost (LCFC) Drives**, which will be supported with NAPA 2 5.5.22.x maintenance release, will be treated as CLATA drives by Celerra and will use the clarata\_r3 storage pool. They will use RAID3 4+1(single lun RG) and 8+1 (dual-lun RG) templates. Available only for Gateway NS500, 700, 40, NSX models at this point, but only with CX3 array hardware(?)

#### Server Thread Utility:

→Thread deadlock detection for CIFS & NFS. If one NFSD thread, or all CIFS threads are blocked for 6 minutes or more, the system will panic if the parameter has been set, since by default the system will not panic for deadlocked threads:

**# .server\_config server\_2 -v "param kernel threads.maxBlockedTime"**

kernel.threads.maxBlockedTime INT 0x014772a4 **360 360** (8,4294967295) FALSE NONE '

Time in seconds defining the global time before we consider a pool is in deadlock'

**Note:** Default deadlock detection value is 360 seconds, 6 minutes—warnings are logged in sys\_log

**# .server\_config server\_2 -v "param kernel threads.panicIfHung"**

kernel.threads.panicIfHung INT 0x014772a8 **0 0** (0,1) FALSE NONE 'Parameter to force panic if a pool is considered in deadlock'

**Note:** Default is not to panic when threads are hung

#### Server Thread Commands:

**\$ server\_thread server\_2 -report | -service | -list | -pool | -stats**

**Note:** In order to use the server\_thread command, you must either create a link to server\_mgr or use the nas\_cmd facility to directly execute the command, as follows:

**\$ /nas/bin/nas\_cmd @server\_thread server\_2 -list** [please note there is a space before the @ symbol]

server\_2 : done

GENERAL REPORT:

Status : OK

Blocking threshold : 22

Blocked count : 0

Total count : 329

**\$ ln -s /nas/bin/server\_mngr /nas/sbin/server\_thread**

**# .server\_config server\_2 -v "THREAD help"**

1153759825: KERNEL: 4: THREAD

THREAD list [service=<service\_name> | pool=<pool\_name> ] [all]

THREAD report [verbose [service=<service\_name> | pool=<pool\_name>]]

THREAD stats

1153759825: KERNEL: 4: ThreadsServicesSupervisor dumpOptions

DUMP\_SERVICES =0x00000001

DUMP\_POOLS =0x00000002

DUMP\_THREAD =0x00000004

DUMP\_BLOCKED\_THREADS =0x00000008

PROCESS\_BLOCKED\_STATE =0x00000010

DUMP\_STACK =0x00000020

WATCHDOG\_THRD =0x02000000

NOT\_XML =0x04000000

XML =0x08000000

REPORT =0x10000000

LIST =0x20000000

STATS =0x40000000

INTERNAL\_COMMAND =0x80000000

current dumpOptions=0x40000000

**# .server\_config server\_2 -v "THREAD list"**

1179315567: KERNEL: 4:

General Report:

1179315567: KERNEL: 4: Status : OK

1179315567: KERNEL: 4: Blocking threshold : 22s

1179315567: KERNEL: 4: Blocked count : 0

1179315567: KERNEL: 4: Total count : 329

**\$ ln -s /nas/bin/server\_mngr /nas/sbin/server\_thread**

**\$ /nas/sbin/server\_thread ALL -stats**

server\_2 : done

STATISTICS:

Panic if hang : 0

Max blocked time : 360

|             | THREADS    | ALERTS   | PANICS   | SERVICE  | BLOCKED  | DETECTED | SOLVED | DETECTED | SOLVED |
|-------------|------------|----------|----------|----------|----------|----------|--------|----------|--------|
| <b>NFSD</b> | <b>118</b> | <b>4</b> | <b>4</b> | <b>2</b> | <b>2</b> |          |        |          |        |
| LOCKD       | 5          | 1        | 1        | 0        | 0        |          |        |          |        |
| MAC         | 0          | 0        | 0        | 0        | 0        |          |        |          |        |
| HTTPD       | 0          | 0        | 0        | 0        | 0        |          |        |          |        |

**Server Log Messages:**

2007-02-26 17:00:19: KERNEL: 4: System Version-T5.5.25.2

2007-02-28 11:26:16: KERNEL: 3: 1: ThreadsServicesSupervisor: Service:NFSD Pool:NFSD\_Exec PENDING for 181 seconds

2007-02-28 11:29:15: KERNEL: 3: 2: ThreadsServicesSupervisor: Service:NFSD Pool:NFSD\_Exec BLOCKED for 360 seconds:  
Server operations may be impacted

2007-02-28 11:34:15: KERNEL: 4: ThreadsServicesSupervisor: Pending threads resolved: threads= was:31; now:0 Pools= was:1; now:0 (seen for 660 seconds)

**# /nas/sbin/server\_thread server\_2 -list -service CIFS -pool SMB1 -all**

server\_2 : done

SERVICE : CIFS

POOL : SMB1

Threads : 253

Time to panic : 360s

Options : 0

Critical tasks : 0

THREAD NAME STATE TIME (s) LAST ACTION

1SMB010 idle 2268.988

1SMB013 idle 2268.988

-----abridged-----

**GENERAL REPORT:**

Status : OK

Blocking threshold : 22

Blocked count : 0

Total count : 253

**GEN 2 (H2G2) HARDWARE SUPPORT:**

→Hammerhead Gen 2 (H2G2) hardware called Hammerhead, Sledgehammer, &amp; Jackhammer, SAN-attached arrays for UltraScales family, CX3-80, CX3-40, &amp; CX3-20, respectively

2<sup>nd</sup> Generation for NSX platform, Server blades are called XB-65. 3.6GHz Irwindale Xeon processor, 4GB RAM, 64-bit capable, Wildfire motherboard, with PLX PCI-Express switch to communicate to I/O Annex via mid-plane connection. Replaces 1<sup>st</sup> Generation Tsunami (4) Port 2Gbps FC I/O Module with new Twister (4) Port 4Gbps FC I/O Module based on Agilent Tachyon chipset. New FC I/O modules will also present Blue/Green LEDs to indicate 4Gbps/1 or 2Gbps link speeds. Blizzard SFP Modules for GbE can have copper SFP (8-ports for NAS), which would be hard-coded to 4G, or optical SFP ports, which would autonegotiate 1, 2, or 4G. An optional Xtreme E 10GbE NIC will be available (fxg0). For backward compatibility, the XB-65's can be put into an existing NSX frame, provided the NAS software has been upgraded first to a version that supports the H2G2 hardware (5.5.22.2+).

→10GbE Neterion single port I/O Module

→CX3 arrays still use first 5 drives for O/S boot, system configuration, &amp; as cache vault for normal operation

→First 3 drives are used for FLARE configuration database and PSM

**Hammerhead CX3-80 (Used for NS80 [not supported until NAPA III], the replacement series for NS700):**

Dual 3.6GHz P4 Xeon processors per SP; 8GB RAM per SP for 16GB total, max. write cache size 3GB; (8)@4Gb FC SAN ports; (4)@4GB FC BE ports per SP; (8)@GbE NAS ports; LH side Management Module B for SPA &amp; RH side Management Module A for SPB; 16Gb cache; 480 drives; 213TB total; up to 2048 LUNs per system; 256 HA Hosts; 73Gb or 146Gb 15K rpm FC 4Gbps Drives; 73Gb, 146Gb, or 300Gb 10K rpm FC 2Gbps Drives; 500Gb 7200rpm 2Gbps LC-FC drives; (4) FC FE Ports per SP &amp; (4) FC BE Ports per SP

**Sledgehammer CX3-40 (Used for NS40, the replacement series for NS500):**

Dual 2.8GHz P4 Xeon processors per SP; 4GB RAM per SP for 8GB total, with max. of 2500MB write cache; (4)@4Gb FC Frontend ports; (2) or (4)@4Gb FC Backend ports per SP; 8GB cache; 240 drives; 119TB total; up to 1024 LUNs per system; 128 HA Hosts; 73Gb or 146Gb 15K rpm FC 4Gbps Drives; 73Gb, 146Gb, or 300Gb 10K rpm FC 2Gbps Drives; 500Gb 7200rpm 2Gbps LC-FC drives; (2) FC FE Ports per SP &amp; (2) FC BE Ports per SP

**Jackhammer CX3-20:**

The CX3-20 uses the same chassis as the CX3-40. Single 2.8GHz P4 Xeon CPU per SP; 2GB RAM per SP; (4)@4Gb FC FE ports; (1)@4Gb FC BE loop per SP; 4Gb system memory, max. write cache size 1053MB; 120 drives; 59TB total; 512 LUNs per system;

128 HA Hosts; 73Gb or 146Gb 15K rpm FC 4Gbps Drives; 73Gb, 146Gb, or 300Gb 10K rpm FC 2Gbps Drives; 500Gb 7200rpm 2Gbps LC-FC drives; (2) FE FC ports per SP; (1) BE FC port per SP

→Piranha-2 & Bigfoot Clariion AX arrays, NS500G/NS700G

→Stiletto UltraPoint 4GB FC Drive DAE3P Enclosure support, with either 4Gbps or 2Gbps drives using (2) 4Gbps LCCs (Link Control Cards). Drives to use Frumon vs. Address-based spinup (Can mix 2GB & 4GB drives on same array, but not recommended on the same loop, as the loop speed would downgrade to the slowest drive)

→Blinking Blue & Green LED's on SP's indicate link failure

→Northstar paddle card to plug SATA drives into Stiletto DAE, supported by Vulcan FLARE

→NS350 (can upgrade to NS500G), new Integrated-only mid-tier entry level platform to be released for NAS & iSCSI, in single or dual DM versions, as well as an Integrated NS704 model. NS350 can have 30 drives for max. storage of 6TB. Each SP will have (4) Cu Gbe Ethernet connections. New 'Ease of Use' packaging for installation.

→Minimum DM hardware 510

→Direct Upgrades from 5.2, 5.3, 5.4 (5.1 or lower must first upgrade to 5.3, then to 5.5)

→Max. mixed FC/ATA capacity 16TB across all platforms except for NS350, 10TB

→Max. FC Capacity 16TB for NSX & NS700, 6TB for NS600 & NS350, and 10TB for NS500, 4TB for 510 DM, 8TB for 514 DM

→Max number FSIDs 4096 for NSX only, 2048 for all other platforms

→Max number Replication Sessions 64 for NSX only, 32 for all other platforms

→Max Snaps per file system 96 for all platforms except 64-only for 510 DM Hardware

## **NAPA 3:**

**NAS 5.5.23.2 Oct 16, 2006**

→MirrorView release with RPQ only

→Vulcan Flare 22 support for Hammers, Bigfoot, and Northstar SATA II drives

Control Station will need new Vulcan NaviCli and Widesky support

→Northstar SATA II Drives with DAE paddle card support for UltraPoint 4GB enclosures, supported by Vulcan Flare

[3.22.020.5.5xx]

Northstar 7200RPM SATA II Drives are for use with CX3 Arrays and NS40 Celerra (SATA technology plugged into fc backplane) SATA II drives run at 3Gb link speeds but must go into 4Gb Enclosures

Celerra will see Northstar SATA II drives as ATA CLATA disk type

Drives cannot be located in 1<sup>st</sup> shelf [i.e., boot enclosure]

Flare 3.22.0x0.5.5xx Vulcan required

Cannot mix SATA II drives with Klondike ATA drives in same RG or any other drive type within same enclosure

SATA II can be used as a HotSpare for FC drives but not for Klondike drives

NS40 Integrated systems require use of setup\_clariion -init and User-defined configuration for shelves (array templates disabled)

Only RAID3 4+1 and/or 8+1 RGs allowed (CLATA or CMATA)

Northstar SATA II and LCFC drives can be mixed in the same FS or clarata\_r3 Storage pool

RAID 3 4+1 will have Single LUN per RG while RAID 3 8+1 will have 2 LUNs per RG

New Shelf Template used: HS=Slot0, then RAID3 4+1 and/or 8+1 for up to 15 drives per shelf

→Support for NS80G Gateway Celerra for Hammerhead

Touted as the replacement platform for Gateway NS704G's—no support yet for Integrated models

No UPS offered; 2-4 Data Movers; Single CS but 2<sup>nd</sup> could be added; Capacity increases to 20TB/blade with fibre channel drive

→LCFC Blizzard Drives for NS40 Celerra & CX3 arrays (Sledgehammer)

500Gb drives at 2Gb bus speed running 7200RPM

Supports only RAID3 4+1 and/or 8+1 RGs (CLATA or CMATA)

Also used on Symmetrix platforms

Introduces support for Integrated Celerras, cannot mix LCFC and FC drives within same Enclosure (Gateway support in NAPA II)

But can mix LCFC, SATA II, or ATA drives in the same file system since all are recognized as clarata\_r3 Storage Pool types

Integrated systems require use of setup\_clariion -init and User-defined configuration for shelves (array templates disabled)

New Shelf Template used: HS=Slot0, then RAID3 4+1 and/or 8+1 for up to 15 drives per shelf

Celerra treats LCFC drives as ATA CLATA disk type. See AR87112 for issue with setup\_clariion script and disk.pm script that looks for specific TLA number, and if it doesn't find 005048596, assumes the drives are to be setup for Raid 5.

→Multi-level NFS Export capability

DART will have the ability to NFS export multiple directories within a single file system

For both NFSv2 & v3, if toplevel fs is exported RO, and nested exports RW, then the effective access for all clients will be RO

Conversely, if nested exports are given RW access, then the parent and other directories exported RO, effective access for all clients will be RW across all directories.

For NFSv4 clients, access permissions are calculated every time a client crosses a directory, therefore, RW and RO access will be enforced.

NMFS is not compatible with this feature

Nested NFS Exports will be enabled by default—to disable, enter param and reboot

## **param nfs multiLevelExports=0**

A known issue is that Clients could possibly have two different versions of a file cached if using multiple exports

### **→PUHC [Pre Upgrade HealthCheck] Checks Added for Upgrades (8 Errors and 1 warning)**

Upgrade will fail if Autoassign or Autotrespass are not properly set; Read Cache on SPs must be enabled; Backend connectivity to SPs checked; /etc/nas\_device.map checked for correct root partition OS1DSK for CS; Unique IDs checked in /nas/server/slot\_x/start file; Control Reserve LUNs validated; /etc/resolv.conf checked for valid DNS names

### **→Clariion Sector Error handling Phase I**

DUDL fix to reduce downtime of file systems. Uncorrectables issue on CLARiiON will still trigger panic, but Data Mover will reboot and come up with all other file systems mounted and will have pfs, checkpoint, replication file systems unmounted until fixed by revector and fsck process.

DART-only fix for Phase I, meaning that potentially, you could drop a patch on the Data Mover to perform the Revector repair File System corruption and recovery applies only to user data and not metadata sectors

Scenarios could involve Production File Systems (Server panic & reboot with affected fs <unmounted>), Checkpoints (Inactive), Replication Sessions (Inactive), or FS Copy operations (Aborted)

Requires use of new sense codes from Flare 16 & 19 for 0311 media errors

Revector tool works better than previous method since only the affected sector is checked and repaired vs. whole lun

### **PFS RECOVERY: →Data Mover Panicks with bad Clariion sector**

#### **Server\_log:**

2006-07-24 09:40:55: CAM:3: I/O Error: c96t117 Irp Oxbb711104 CamStatus Ox84 SesiStatus 0x02 Sense **0x3/0x11/0x00**

#### **Dump Header:**

PANIC in file ..BVolumeIrp.cxxx at line: 323: IO failure on Vol: 271, blkNo:36048, SKey: 0x311

271 is volume number; 36048 is physical sector triggering anic, and 0x311 is the Clariion Sense Error

#### **Sys\_log:**

UFS: 3:7:Slot\_2 Mount failed, Please start Revectoring using “revector start vol=449”

**Note:** Approximately 4 minutes total for the timeouts to trigger Data Mover panic

Step 1. Must run the Revector Utility against the affected volume referenced in the sys\_log—remaps bad sectors & 0’s out

### **\$ .server\_config server\_2 “revector start vol=xxx”**

**Note:** Tool also creates a textfile report placed in rootfs, later used by MapBlock utility. Revector command should work, but affected volume must be available to Data Mover. If command fails, use \$ server\_mount server\_x -v to establish db link to fs for DM Step 2. Monitor Revector work until completed:

### **\$ .server\_config server\_2 “revector display vol=xxx”**

**Note:** Output will display progress and then “No Revectoring session found...” when process completes

Step 3. Remount file system on Server to initiate FSCK & MapBlock Report

\$ server\_mount server\_2 fs1 /fs1

**Note:** Remounting the file system will kick off an auto-fsck and the MapBlock utility using information found from revector report to complete work and output list of files affected

### **CHECKPOINT, REPLICATION, FS COPY RECOVERY:**

Replication & Checkpoints will become Inactive on affected production file systems, and FS\_Copy will abort

1. Use revector tool against pfs & recover file system
2. Create new pfs checkpoint and then refresh any others (fs\_ckpt fs1 –list should show checkpoints going from INACTIVE to %)
3. Abort and then restart fs\_copy
4. Abort Replication and Playback services (Source/Target sides) and restart replication

→Printstats command for sysstat data—(19) Data Mover statistics can be collected via snapshots—not a public tool

### **NAPA 4:**

NAS 5.5.24.2, November 2006

→NS & NSX support for CX3-20c & CX3-40c CLARiiON iSCSI platforms with 4-Front-End IO ports (Flare 22 with special SW Enabler)

→CIFS Troubleshooting tool: server\_cifssupport

→Large Major & Minor Device numbers (32-bit up to 4G values)

→NS80 Celerra Hammerhead Integrated Platform using direct optical fibre connections to CX3-80 array (Blades & SP’s have built-in SFP connectors), replacement for NS702/704; 2-4 DM’s or 1-2 CS’s; Chivas CS (same as NSX); either XB-60 or XB-65 Data Mover blades; SP’s will use a Nor’easter Management Module with each SP on a separate network and vlan (192.168.1.200/192.168.2.201), and two Ethernet ports per module, one for Service and one for Customer LAN; the NAS Management Modules are built into the Typhoon board

**Note:** The SP Resume PROM personality should be set to “SAN”. This is referred to as an Integrated Celerra that runs in SAN mode vs. AUX, mainly due to the use of LC-to-LC fibre optic cables between DM & SPs.

### **SERVER CIFSSUPPORT TOOL:**

The server\_cifssupport tool was developed as a troubleshooting aid for EMC Support and its Customers. Its primary focus with the 5.5.24.2 release is to assist in troubleshooting various CIFS issues. The tool is comprised of several different segments, and depending on the nature of the issue being investigated, different switches and associated options may be selected. The four major sections are -pingdc, -accessright, -cred, and -secmap. When the command is run, certain output goes to screen and to the Server Log using the CIFSSUPPORT topic & other appropriate headers, such as SMB.

Specifically, the server\_cifssupport is a multi-purpose tool that can validate connectivity between CIFS Server & Domain Controllers, manage & verify access rights for Users, produce User credentials & validate mappings to Celerra, and manage Secure Mapping cache (SecMap)

#### Documentation:

Server\_cifssupport not documented in Release Notes, Doc CD, or Man Pages

“Using the Celerra server\_cifssupport Command Rev A01” (Powerlink Customer Service document)

Tool not published for customers, but there is no reason that customers cannot be made aware of this tool for their own troubleshooting, etc.

#### Example Server Log Heading:

2006-10-25 09:36:38: CIFSSUPPORT: 4: pingDC compname=sector DC=GEORGE Step='Lookup trusted domains' OK

#### I. \$ server\_cifssupport -pingdc -compname or -netbios -dc -verbose

The -pingdc switch is used to validate the current status of the CIFS server in the declared Windows Domain, and processes a series of checks each time the command is run. Warnings and Errors are outputted to screen for the appropriate check when an Error or Warning condition occurs. When using the -verbose switch, output of the tests, are logged to the Server Log under the CIFSSUPPORT heading. Pingdc verifies CIFS and DC connectivity. The following actions are tested whenever the pingdc command is run, for either the compname or netbios name indicated, and to the specific DC if stipulated, with additional information written to the Server Log if the -verbose switch is used:

#### Actions Taken by Command—from TTT Documentation:

- DC availability from WINS or DNS
- DC connectivity via CIFS ports and ICMP ping
- Compname Join status with the domain
- Server authentication with a DC
- Access to the IPC\$ share
- Access to NETLOGON pipe for NTLM authentication
- Access to LSA pipe for mapping SIDs to Users and Groups
- Access to SAMR pipe for merging UNIX & Windows groups when using NTCredential option
- Status of trusted domain queries
- Status of privilege name queries such as those used with I18N in Usrmgr

**Note:** The test will stop upon the first Error, but will continue if only Warnings are received

#### Actions Taken as Shown by Server Log:

2006-10-25 09:43:32: CIFSSUPPORT: 4: Starting test Pre-conditions  
2006-10-25 09:43:32: CIFSSUPPORT: 4: Starting test Test-setup  
2006-10-25 09:43:32: CIFSSUPPORT: 4: Starting test Check Password File  
2006-10-25 09:43:32: CIFSSUPPORT: 4: Starting test Check DNS  
2006-10-25 09:43:32: CIFSSUPPORT: 4: Starting test TCP connectivity  
2006-10-25 09:43:32: CIFSSUPPORT: 4: Starting test Check CIFS Ports  
2006-10-25 09:43:32: CIFSSUPPORT: 4: Starting test Logon IPC\$  
2006-10-25 09:43:32: CIFSSUPPORT: 4: Starting test Request Domain Sid  
2006-10-25 09:43:32: CIFSSUPPORT: 4: Starting test Open NETLOGON Secure Channel  
2006-10-25 09:43:32: CIFSSUPPORT: 4: Starting test Lookup privileges  
2006-10-25 09:43:32: CIFSSUPPORT: 4: Starting test LSA lookup mappings  
2006-10-25 09:43:32: CIFSSUPPORT: 4: Starting test SAMR lookups  
2006-10-25 09:43:32: CIFSSUPPORT: 4: Starting test Lookup trusted domains

#### **# server\_cifssupport server\_2 -pingdc -compname sector -verbose**

server\_2 : done

PINGDC GENERAL INFORMATIONS

DC SERVER:

Netbios name : GEORGE

CIFS SERVER :

Compname : sector

Domain : 2k3.pvt.dns

Warning 17455906883: server\_2 : compname sector DC=GEORGE Trusted domain='w2k.pvt.dns' status='There are currently no logon servers available to service the logon request'

**Note:** In the above example, one of the trusted domains is not available and a warning is generated. This could be a problem if a User from the Trusted Domain was trying to log into the Celerra and could not.

**SERVER LOG OUTPUT OF COMPLETED PINGDC:**

**2006-10-25 09:43:32: CIFSSUPPORT: 4: Starting test Pre-conditions**

2006-10-25 09:43:32: CIFSSUPPORT: 4: End of test Pre-conditions

**2006-10-25 09:43:32: CIFSSUPPORT: 4: Starting test Test-setup**

2006-10-25 09:43:32: CIFSSUPPORT: 4: End of test Test-setup

**2006-10-25 09:43:32: CIFSSUPPORT: 4: Starting test Check Password File**

2006-10-25 09:43:32: CIFSSUPPORT: 4: End of test Check Password File

**2006-10-25 09:43:32: CIFSSUPPORT: 4: Starting test Check DNS**

2006-10-25 09:43:32: CIFSSUPPORT: 4: pingDC compname=sector DC=GEORGE Step='Check DNS' OK

2006-10-25 09:43:32: CIFSSUPPORT: 4: End of test Check DNS

**2006-10-25 09:43:32: CIFSSUPPORT: 4: Starting test TCP connectivity**

2006-10-25 09:43:32: CIFSSUPPORT: 4: End of test TCP connectivity

**2006-10-25 09:43:32: CIFSSUPPORT: 4: Starting test Check CIFS Ports**

2006-10-25 09:43:32: CIFSSUPPORT: 4: pingDC compname=sector DC=GEORGE Step='Check CIFS Ports' OK

2006-10-25 09:43:32: CIFSSUPPORT: 4: End of test Check CIFS Ports

**2006-10-25 09:43:32: CIFSSUPPORT: 4: Starting test Logon IPC\$**

2006-10-25 09:43:32: CIFSSUPPORT: 4: pingDC compname=sector DC=GEORGE Step='Logon IPC\$' OK

2006-10-25 09:43:32: CIFSSUPPORT: 4: End of test Logon IPC\$

**2006-10-25 09:43:32: CIFSSUPPORT: 4: Starting test Request Domain Sid**

2006-10-25 09:43:32: CIFSSUPPORT: 4: pingDC compname=sector DC=GEORGE Step='Request Domain Sid' OK

2006-10-25 09:43:32: CIFSSUPPORT: 4: End of test Request Domain Sid

**2006-10-25 09:43:32: CIFSSUPPORT: 4: Starting test Open NETLOGON Secure Channel**

2006-10-25 09:43:32: CIFSSUPPORT: 4: pingDC compname=sector DC=GEORGE Step='Open NETLOGON Secure Channel' OK

2006-10-25 09:43:32: CIFSSUPPORT: 4: End of test Open NETLOGON Secure Channel

**2006-10-25 09:43:32: CIFSSUPPORT: 4: Starting test Lookup privileges**

2006-10-25 09:43:32: CIFSSUPPORT: 4: pingDC compname=sector DC=GEORGE Step='Lookup privileges' OK

2006-10-25 09:43:32: CIFSSUPPORT: 4: End of test Lookup privileges

**2006-10-25 09:43:32: CIFSSUPPORT: 4: Starting test LSA lookup mappings**

2006-10-25 09:43:32: CIFSSUPPORT: 4: pingDC compname=sector DC=GEORGE Step='LSA lookup mappings' OK

2006-10-25 09:43:32: CIFSSUPPORT: 4: End of test LSA lookup mappings

**2006-10-25 09:43:32: CIFSSUPPORT: 4: Starting test SAMR lookups**

2006-10-25 09:43:32: CIFSSUPPORT: 4: pingDC compname=sector DC=GEORGE Step='SAMR lookups' OK

2006-10-25 09:43:32: CIFSSUPPORT: 4: End of test SAMR lookups

**2006-10-25 09:43:32: CIFSSUPPORT: 4: Starting test Lookup trusted domains**

2006-10-25 09:43:32: CIFSSUPPORT: 4: compname sector DC=GEORGE Trusted domain='NTDOMAIN' status='There are currently no logon servers available to service the logon request'

2006-10-25 09:43:32: CIFSSUPPORT: 4: pingDC compname=sector DC=GEORGE Step='Lookup trusted domains' OK

2006-10-25 09:43:32: CIFSSUPPORT: 4: End of test Lookup trusted domains

**Example of Error seen when Administrator cannot be Mapped:**

# server\_cifssupport server\_2 -pingdc -compname sector -verbose

server\_2 : done

Error 13160939577: server\_2 : pingdc failed due to NT error INVALID\_ID\_AUTHORITY at step LSA lookup mappings

**Note:** Would mean that there is a mapping issue. To recreate the problem, stopped the Usermapper service and removed the SecMap cache entry for Administrator, as this is the default account that is used by pingdc for the LSA Lookup test.

**Example of Error seen when Compname has not been Joined:**

# server\_cifssupport server\_2 -pingdc -compname nobody -verbose

Error 13160939589: server\_2 : The server has not been joined to domain Active Directory.

**II & III. \$ server\_cifssupport -cred -name -domain -sid -uname -uid -build [-admin] -netbios -compname -standalone & -accessright**

The -cred switch allows the Administrator or troubleshooter to build a Windows or Unix credential for a particular User, useful for determining how the Celerra has resolved the User's Security Access Token, Group Memberships, Unix permissions, and GPO privileges, into a mapping of credential information. This tool also uses the -build option to create a credential when the User is not physically connected to the Celerra.

**The following actions are tested when using the -cred switch:**

**\$ server\_cifssupport -cred -name mcdougal -domain 2k3 -build -admin administrator**

→Build either a Windows or UNIX User credential to see if cred build & resolution are correct (No requirement for User or Client logon). Use the -admin switch with the -build option, if you know an Administrative User account—this will allow ability to obtain output even if the Compname is invalid. Credential is printed for review. -ldap switch used to resolve any Universal Groups found.

**Note:** It should be noted that the -build will force a mapping if a mapping does not already exist, and that if using local passwd/group files or NIS for the source of mappings, the CIFS Resolver param should be set to 1 to ensure that both “user & “user.domain” are

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
queried—if not used, it would be possible that the desired account would not be queried by name and if Usermapper were running, would result in an unintended mapping. For example, NIS has user defined as “user” but with CIFS resolver set to default 0, DART will only look for “user.domain”, NIS mapping would fail, and Usermapper would create a new mapping for “user.domain”.

**Example of Windows User:**

**# server\_cifssupport server\_2 -cred -name backup -domain 2k3 -build**

server\_2 : done

ACCOUNT GENERAL INFORMATIONS

|             |   |                                         |
|-------------|---|-----------------------------------------|
| Name        | : | backup                                  |
| Domain      | : | 2K3                                     |
| Server      | : | SECTOR                                  |
| Primary SID | : | S-1-5-15-bfc56af5-d6cf8701-5f67b1a3-4da |
| UID         | : | 32769                                   |
| GID         | : | 32768                                   |

ACCOUNT GROUPS INFORMATIONS

| Type | UNIX ID    | Name             | Domain       | SID                                     |
|------|------------|------------------|--------------|-----------------------------------------|
| NT   | 32772      | Domain Users     | 2K3          | S-1-5-15-bfc56af5-d6cf8701-5f67b1a3-201 |
| NT   | 32773      | Domain Guests    | 2K3          | S-1-5-15-bfc56af5-d6cf8701-5f67b1a3-202 |
| NT   | 32770      | Domain Admins    | 2K3          | S-1-5-15-bfc56af5-d6cf8701-5f67b1a3-200 |
| NT   | 4294967294 | Everyone         |              | S-1-1-0                                 |
| NT   | 4294967294 | NETWORK          | NT AUTHORITY | S-1-5-2                                 |
| NT   | 4294967294 | ANONYMOUS LOGON  | NT AUTHORITY | S-1-5-7                                 |
| NT   | 2151678496 | Administrators   | BUILTIN      | S-1-5-20-220                            |
| NT   | 2151678497 | Users            | BUILTIN      | S-1-5-20-221                            |
| NT   | 2151678498 | Guests           | BUILTIN      | S-1-5-20-222                            |
| NT   | 1          | UNIX GID=0x1 &ap |              | S-1-5-12-2-1                            |
| UNIX | 32772      |                  |              |                                         |
| UNIX | 32773      |                  |              |                                         |
| UNIX | 32770      |                  |              |                                         |

**# server\_cifssupport server\_2 -cred -uname tom -build**

server\_2 :

Error 5005: server\_2 : The user tom doesn't exist.

**Note:** Example of error that would be seen when the User cannot be found in NIS or local passwd file

**# server\_cifssupport server\_2 -cred -uname tom**

server\_2 :

Error 5005: server\_2 : The UNIX user tom has no credential in the cache

**Note:** User is found but is not connected/or does not have an NTCredential in the cache

**# server\_cifssupport server\_2 -cred -uname tom -build**

server\_2 : done

**Note:** User is found and a UNIX Credential is created for the UNIX user using the –build option. Keep in mind that eventhough the User was found in the local passwd file and a credential was created, the user is not added to Secmap cache as a result of this mapping.

**User access to files or shares are tested using the –accessright switch. Accessrights can be checked against all the various file system accesspolicies:**

**\$ server\_cifssupport -accessright**

The –accessright switch can be used to verify the effective access rights for Users to specific Shares, or specific Directories and Files, both from a Windows and Unix context. When using the -build option, this tool can evaluate access rights even when the User is not connected to the Celerra. When using the -policy option, this tool allows Administrators to evaluate the effects on access rights that may vary depending on the Celerra AccessPolicy in effect on the file system [e.g., NT, UNIX, SECURE, NATIVE, MIXED, MIXED\_COMPAT].

**ACCESS MASK VALUES:**

Whenever the –accessright command is run for Windows Users, an Allowed Mask is generated that essentially summarizes what DART will provide as the effective accessright to the Share or Files provided in the –path statement. DART compares the access mask to the masks contained in the ACE entries, looking for bits that match, when granting or restricting access:

**How do we translate the mask into accessrights?**

- 1.) From visual inspection of Windows permissions via Explorer interface
- 2.) Each value is outputted by the command in human readable form too—see example below
- 3.) Can also translate Mask into Windows permissions (if you know how)

**MASK EXAMPLE:**

**# server\_cifssupport server\_2 -accessright -name thomas –domain 2k3 -path /sector/thomas -build**

Path : /sector/thomas

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!

Allowed mask : **0x301ff** →32-bit binary access mask value in hex, which translates to 11000000011111111 binary

Action : List Folder / Read data  
 Action : Create Files / Write data  
 Action : Create Folders / Append Data  
 Action : Read Extended Attributes  
 Action : Write Extended Attributes----output abridged-----

### **BITS 0-31 DEFINED:**

Bits 0-15 are considered low-order bits used for object specific access rights

Bits 16-22 are the bits used to define Standard Access Rights

DELETE is used to delete an object

READ\_CONTROL is the right to read an object's SD information

SYNCHRONIZE allows the object to use synchronization and allows threads to wait for the correct state

WRITE\_DAC is the right to modify DACL in the SD of the object

WRITE\_OWNER is the right to change the owner in the SD of the object

Bit 23 →Identifies the object's SACL, or System Access Control List, used for auditing purposes only—sets or gets SACL

Bits 24-27 are reserved & not used

Bits 28-31 →Used to define Generic Access Rights & are considered the high-order bits in the access mask

GENERIC\_ALL is used to indicate Read, Write, & Execute access

GENERIC\_EXECUTE is used to indicate Execute access

GENERIC\_READ indicates read access

GENERIC\_WRITE indicates Write Access

```
$ server_cifssupport -accessright -name -domain -sid -uname [uid] -path [-share] -policy [mixed, native, secure, nt, unix] -build [-admin] -compname -standalone
$ server_cifssupport -accessright -name jamoke -domain 2k3 -share public -build
$ server_cifssupport -accessright -sid S-1-5-15-597e341a-ce8e74d1-46b -path /fs1/homedirs/joe -build
$ server_cifssupport -accessright -name harry -domain 2k3 -path /fs1/homedirs/harry -build -policy unix
# server_cifssupport server_2 -accessright -name james -domain 2k3 -path /sector/tmatta -build
```

server\_2 : done

#### ACCOUNT GENERAL INFORMATIONS

Name : james  
 Domain : 2k3  
 Path : /sector/tmatta

Allowed mask : **0x0** →This User does not have any access to the indicated folder—no explicit ACE to allow access

```
# server_cifssupport server_2 -accessright -name percy -domain 2k3 -path /sector/tmatta -build
```

server\_2 : done

#### ACCOUNT GENERAL INFORMATIONS

Name : percy  
 Domain : 2k3  
 Path : /sector/tmatta  
 Allowed mask : **0x20089** →This User has been granted READ access only to the indicated folder

Action : List Folder / Read data  
 Action : Read Extended Attributes  
 Action : Read Attributes  
 Action : Read Permissions

```
# server_cifssupport server_2 -accessright -name tmatta -domain 2k3 -path /sector/tmatta -build
```

server\_2 : done

#### ACCOUNT GENERAL INFORMATIONS

Name : tmatta  
 Domain : 2k3  
 Path : /sector/tmatta  
 Allowed mask : **0x301ff** →This User has Full Control access to this directory

Action : List Folder / Read data  
 Action : Create Files / Write data  
 Action : Create Folders / Append Data  
 Action : Read Extended Attributes  
 Action : Write Extended Attributes  
 Action : Traverse Folder / Execute File  
 Action : Delete Subfolders and Files  
 Action : Read Attributes  
 Action : Write Attributes

Action : Delete

Action : Read Permissions

→Accessrights to files, directories, & shares using Windows mask

#### **IV. \$ server\_cifssupport -secmap -report -list [-name -domain -sid -uid -gid] -create [-name -domain -sid] -verify [-name -domain -sid] -update [-name -domain -sid] -delete [-name -domain -sid] -export [-file <filename>] -import**

The -secmap switch allows Administrators to verify and export the contents of the SecMap cache database, useful when trying to resolve UID/GID mapping or permissions issues. The tool also allows the ability to delete database entries, modify entries, add entries, etc. Please keep in mind that the SecMap cache is a persistent repository of the first time that a SID-to-UID/GID mapping occurred on the Data Mover or VDM Server, and records where the mapping entry originated from [e.g., NIS lookup, Usermapper lookup, etc Passwd file lookup, etc]

##### **EXAMPLE OUTPUT:**

**# server\_cifssupport server\_2 -secmap -list**

server\_2 : done

SECMAP USER MAPPING TABLE

| UID | Origin | Date | Name |
|-----|--------|------|------|
|-----|--------|------|------|

SID

|                                         |            |                          |            |
|-----------------------------------------|------------|--------------------------|------------|
| 32769                                   | usermapper | Wed Oct 18 13:14:41 2006 | 2K3\backup |
| S-1-5-15-bfc56af5-d6cf8701-5f67b1a3-4da |            |                          |            |

SECMAP GROUP MAPPING TABLE

| GID | Origin | Date | Name |
|-----|--------|------|------|
|-----|--------|------|------|

SID

|                                         |            |                          |                  |
|-----------------------------------------|------------|--------------------------|------------------|
| 32772                                   | usermapper | Wed Oct 18 13:14:42 2006 | 2K3\Domain Users |
| S-1-5-15-bfc56af5-d6cf8701-5f67b1a3-201 |            |                          |                  |

**Note:** Please note that times are listed in the timezone that the Data Mover has been set to, but that GMT time is the default timezone

**# server\_cifssupport server\_2 -secmap -list -name thomas -domain 2k3.pvt.dns**

server\_2 : done

SECMAP USER MAPPING TABLE

| UID | Origin | Date | Name |
|-----|--------|------|------|
|-----|--------|------|------|

SID

|                                         |            |                          |            |
|-----------------------------------------|------------|--------------------------|------------|
| 32770                                   | usermapper | Fri Oct 20 17:44:02 2006 | 2K3\thomas |
| S-1-5-15-bfc56af5-d6cf8701-5f67b1a3-4b4 |            |                          |            |

**# server\_cifssupport server\_2 -secmap -report**

server\_2 : done

SECMAP GENERAL INFORMATIONS

|      |   |          |
|------|---|----------|
| Name | : | server_2 |
|------|---|----------|

|       |   |         |
|-------|---|---------|
| State | : | Enabled |
|-------|---|---------|

|    |   |   |
|----|---|---|
| Fs | : | / |
|----|---|---|

|            |   |    |
|------------|---|----|
| Used nodes | : | 15 |
|------------|---|----|

|             |   |   |
|-------------|---|---|
| Used blocks | : | 0 |
|-------------|---|---|

SECMAP MAPPED DOMAIN

|      |     |
|------|-----|
| Name | SID |
|------|-----|

|     |                                              |
|-----|----------------------------------------------|
| 2K3 | S-1-5-15-bfc56af5-d6cf8701-5f67b1a3-ffffffff |
|-----|----------------------------------------------|

**# server\_cifssupport server\_2 -secmap -verify -name thomas -domain 2k3.pvt.dns**

server\_2 :

Error 5005: server\_2 : Cannot get mapping for S-1-5-15-bfc56af5-d6cf8701-5f67b1a3-4b4

**Note:** Above output indicates that the User is not yet mapped in the database. Use the -create option with -name and -domain to force a new mapping into SecMap cache, or use the cifssupport -cred to force mapping for a User that has not yet been mapped:

**# server\_cifssupport server\_2 -cred -name thomas -domain 2k3 -build**

server\_2 : done

ACCOUNT GENERAL INFORMATIONS

|      |   |        |
|------|---|--------|
| Name | : | thomas |
|------|---|--------|

|        |   |     |
|--------|---|-----|
| Domain | : | 2K3 |
|--------|---|-----|

|        |   |        |
|--------|---|--------|
| Server | : | SECTOR |
|--------|---|--------|

|             |   |                                         |
|-------------|---|-----------------------------------------|
| Primary SID | : | S-1-5-15-bfc56af5-d6cf8701-5f67b1a3-4b4 |
|-------------|---|-----------------------------------------|

|     |   |       |
|-----|---|-------|
| UID | : | 32770 |
|-----|---|-------|

|     |   |       |
|-----|---|-------|
| GID | : | 32768 |
|-----|---|-------|

**# server\_cifssupport server\_2 -secmap -create -name percy -domain 2k3**

server\_2 : done

#### SECMAP USER MAPPING TABLE

| UID                                     | Origin     | Date                     | Name      |
|-----------------------------------------|------------|--------------------------|-----------|
| SID                                     |            |                          |           |
| 32771                                   | usermapper | Fri Oct 20 17:54:58 2006 | 2K3\percy |
| S-1-5-15-bfc56af5-d6cf8701-5f67b1a3-4c5 |            |                          |           |

**# server\_cifssupport server\_2 -secmap -verify -name thomas -domain 2k3.pvt.dns**

server\_2 : done

**Note:** Apparently, all the verify command does is tell you whether the name or sid is in the database, but doesn't give you any details.

**# server\_cifssupport server\_2 -secmap -update -name percy -domain 2k3**

server\_2 : done

**Note:** Update command can be used to update entries in secmap with values stored in the mapping source, then performs ACL update on file systems(?). Also used to update records with specified domain and SID entries.

**# server\_cifssupport server\_2 -secmap -delete -name percy -domain 2k3**

server\_2 : done

**# server\_cifssupport server\_2 -secmap -delete -name "Domain Controllers" -domain 2k3**

server\_2 : done

**Note:** To delete names that have spaces, use quotations

**# server\_cifssupport server\_2 -secmap -export -file sec\_dmp**

server\_2 : done

**# cat sec\_dmp**

S-1-5-15-bfc56af5-d6cf8701-5f67b1a3-200:2:96:8002:8002:2K3\Domain Admins

S-1-5-15-bfc56af5-d6cf8701-5f67b1a3-4da:1:96:8001:8000:2K3\backup

**Note:** Unfortunately, the dump does not separate into Users and Groups like the list function does.

**# server\_cifssupport server\_2 -secmap -import -file sec\_dmp**

server\_2 :

Error 5005: server\_2 : Unexpected MAPPING\_RECORD element in Import request

**Note:** Could not get the import function to work at all—this is a known bug

#### REMOVING ENTIRE SECMAP CACHE WITHOUT STOPPING CIFS:

```
$ server_cifssupport server_x -secmap -list |grep "S-1" | sed 's/S-1@/ /' | awk -F '@'{printf "S-1%s\n", $2}' |xargs -i  
$ server_cifssupport server_x -secmap -delete -sid {}
```

#### USING SERVER PARAM AND DISABLING SECMAP CACHE:

**# server\_param server\_2 -facility cifs -info secmap.enable**

server\_2 :

|                  |                                           |
|------------------|-------------------------------------------|
| name             | = secmap.enable                           |
| facility_name    | = cifs                                    |
| default_value    | = 1                                       |
| current_value    | = 1                                       |
| configured_value | =                                         |
| user_action      | = restart Service                         |
| change_effective | = restart Service                         |
| range            | = (0,1)                                   |
| description      | = Control CIFS Secure Mapping cache state |

**# server\_param server\_2 -facility cifs -modify secmap.enable -value 0**

server\_2 : done

Warning 17716815753: server\_2 : You must stop and start cifs for secmap.enable changes to take effect

#### NAPA 5:

NAS 5.5.26.4 ETR Jan 25, 2007

→ RAID5 4+1 & 6+1 Support for LCFC & SATA II drives, but only for CX3 arrays

→ Support for new Dewars Control Station (Possibly released with GNapa instead, hard to determine exactly)

→ Flare 24 Mars support, and support for CX3-20F/CX3-40F arrays in Gateway configurations

→ Support added for new Headhunter IO module on backend arrays CX3-20F & CX3-40F for Gateway systems--adds (4) FC front-end ports on Jackhammer CX3-20F arrays [total of (12) front-end fibre ports and (6) backend fibre ports], and Sledgehammer CX3-

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
40F arrays [total of (4) front-end fibre and (4) backend fibre ports]. Headhunter is a Quad 4Gb controller using PCI express bus and FPGA for Diplexing.

**Note:** With this NAS release, CX Fish arrays are only supported for Flare 19 and higher, while CX3 Hammer arrays are only supported for Flare 22 and higher.

## **GRANDNAPA:**

A special TLC release (NAS 5.5.27.5) to port serviceability items into the product, GA 14 March 2007

### **CONTROL STATION SESSION TIMEOUT GNAPA:**

→Applies to CLI shell sessions only, with csh & tcsh governed by variable called “autologout” in minutes and bash/ksh governed by variable called “TMOUT in seconds

**EXAMPLE:** /etc/profile

```
TMOUT=3600
```

```
export TMOUT
```

→Default value will be 60 minutes and is enabled by default

#### **# /nas/sbin/nas\_config -sessiontimeout**

Session timeout value is 60 minutes

```
done
```

→Celerra Manager is not setup with a session timeout by default (480 minutes)

→Values are set in /etc/profile & /etc/csh.cshrc

**# nas\_config -sessiontimeout <minutes>** | 0 or off disables timeout

**Note:** Timeout values will be between 5-240 minutes--translated into seconds for /etc/profile & in minutes for /etc/csh.cshrc and updates appropriate config files—log off and back on for the values to go into effect

### **CONFIGURABLE LOGIN BANNER GNAPA:**

→Both login banner and MOTD (Message of the Day) will be displayed on Console, SSH, or Celerra Mgr login  
/etc/issue & /etc/motd, respectively

→NAS Upgrades will not overwrite or replace these files

→Banner is not enabled by default for SSH, nor does Celerra Mgr support any kind of Login Banner or MOTD

### **FIREWALL FRIENDLINESS GNAPA:**

Network Service management added to Celerra Manager based on services outlined in white paper, “The Celerra Network Server on the Enterprise Network”

Feature allows an Admin to Enable/Disable ‘non-essential’ services on Data Movers & Control Stations [DM will indicate reboot needed for certain services]

Changes will be logged in /nas/log/sys\_log, facility is ADMIN

### **AUTOMATIC LOG COLLECTION & TRANSFER PHASE II GNAPA:**

Adds GUI functionality for Enabling/Disabling transfer feature, not the collection portion of the feature

Adds GUI capability to collect logs, but does NOT collect dumps [GrandNapa /nas/var/log, but puts materials in /nas/var/emcsupport in Cognac NAS 5.6]

Events & Alerts posted for the LogCollect facility [nas\_event -l -f LogCollect]

### **MPFSi ENHANCEMENTS GNAPA:**

#### **Large File fix:**

→Reduces metadata logging required when writing many blocks at once, translating to less CPU overhead and less ufslog contention, as well as improved scalability

#### **Hierachical Volume Management:**

→Improves Client caching and offloads more volume tree traversals to the Client, reducing Server processing required & RPC overhead

→Enhances FMP protocol—Clients request space allocation and Server used to provide multiple extents across LUNs—improvement will be to provide only a single extent back to the Client, and the Client will determine which LUN to write to

→This enhancement requires updated Client software—LinRoad 5.0, Windows will be post-5.0

#### **CIFS Fixes:**

→More detailed info returned to FMP Client

→FMP client will use multiple interfaces for CIFS, FMP, iSCSI whereas previous behavior was to use single interface for all protocols

→Multiple CIFS hosts can be defined

### **MPFSi LINUX CLIENT CHANGES GNAPA:**

→install-mpfs script for the 4.3 Client, which will support RHEL 3 & 4, SuSE, et.al.

→New mpfsinfo validates accessibility of each file system, lists LUNs

→mpfsinq modified to show more CLARiiON information, such as active & passive paths to LUNs

/proc/mpfs/devices (devices seen by kernel)

/proc/mpfs/params (mpfs parameter values)

/proc/hrdp (HighRoad protection status)

→Persistent parameters /etc/mpfs.conf

globPrefetchSize default 128, but can increase to 2048—this determines how many blocks are returned by Server to Client allocation requests. Values changed in mpfs.conf file become persistent during reboots; use mpfscctl to update kernel memory until next reboot.

→mpfsdiscover daemon uses /etc/sysconfig/EMCmpfs file for parameters & is used to detect new devices and trespassed LUNs (every 15 minutes)

→mpfsi enhanced to be as robust as NFS and can now replay uncommitted writes after system crash—dirty pages are maintained until committed via FMP flush; replay does AllocSpace reissue, rewrites data, then does FMP flush

## **CELLERRA iSCSI SNAP/REPLICATION ON LINUX GNAPA:**

→PLU is a version file with blocks allocated on-demand

→Snaps are point-in-time version files that do not change, unless they are a TWS (one Temporary Writeable Snap), which can be modified

→Snaps are created instantaneously. New writes to PLU are written to new location. Snaps can only be promoted once, becomes RW. Newer snaps are deleted as part of the restore procedure. Snaps can be deleted in any order except for promoted & restore.

→iSCSI Replication is LUN-based vs. File System-based, with Source & Destination pair, and allows one-to-many sessions

→iSCSI Replication uses baseline copy to create common base, then incremental transfers

→iSCSI destination snaps can be promoted for Replication sessions

### **I. iSCSI SNAP CREATION & DATA OPERATION:**

→PLU LUN with data blocks B1, B2, etc

→Snap S1 created, assumes ownership of data blocks, though PLU still points to original data blocks

→New data written to PLU Lun block B1

→Modified block B1' is created and owned by PLU

### **II. iSCSI SNAP PROMOTION:**

→PLU LUN has several Snaps, of which one will be promoted

→Writeable snap for S2 created, called S2'

→S2' is given a LUN identity and promoted

### **III. iSCSI SNAP RESTORE:**

→Writeable snap for S2 created, called S2'

→S2' becomes new PLU & newer snaps are deleted

### **iSCSI REPLICATION:**

→Create Snap

→Mark Snap for replication

→Perform baseline copy of Snap to Destination LUN

→D1 snap created on destination & becomes commonbase with S1 source snap

### **Linux Platform:**

→iSCSI devices presented as regular SCSI devices on Linux, supporting RHEL 3 U5, RHEL 4 U2, SuSE SLES 9 Sp2

→iSCSI initiators 3.6.x for Linux 2.4.x & 4.0.x for Linux 2.6.x

→Install using rpm –Uvh iscsi-initiator-utils-x.x.x.rpm

→Configure initiator name /etc/initiator.name, target info /etc/iscsi.conf, service info /etc/init.d/iscsi start, setup the auto-start on boot service using chkconfig –a iscsi, session & connection status using iscsi-ls command, mount devices during boot using /etc/fstab.iscsi

### **Host Utilities (CBM=Celerra Block Management):**

# cbm\_security –c cbmadmin\* –s secret →used to configured authorized user credential

# cbm\_iscsi →List LUNs on host, detailed info, perform snapshot operations

# cbm\_replicate used for replication operations

# lun\_discover used for dynamic LUN discovery

→Configuration file /etc/cbm/cbm.conf

### **Celerra Commands:**

# server\_security server\_x –a –p chap –n cbmadmin –s secret

#server\_iscsi server\_x –lun –n 10 –create mytarget –size 100M –fs fs1 –readonly yes

# nas\_replicate

# nas\_cel –name eng001 –create 10.241.168.23 –passphrase cbmcmbcm (Dart Interconnect Authentication for Remote Replication)

**Note:** DIC password entries in DM passwd file are used only for Local or Loopback Replication authentication

### **Troubleshooting:**

/nas/var/messages

# iscsi-ls; /proc/partitions; /proc/scsi/scsi

Log Facility /etc/cbm/cbm.conf with different levels

Use server\_iscsi on CS for service, configuration, and mask information

Use nas\_replicate to verify Replication sessions

Check Server Log, looking for NBS, ISCSI, DP, VCS facilities

## **XLT CORRUPTION DETECTION & RECOVERY GNAPA:**

- CIFS issues can occur affecting Data Mover Joins, case sensitivity to Shares and pathnames, etc.
- Feature will detect XLT corruption on-disk, not DART in-memory corruption, and take automatic recovery steps
- (3) types of XLT files that DART is concerned with: Stored in rootfs etc\_common on DART & /nas/site/locale on CS  
  unidata.txt>contains upper & lowercase mappings in UTF-16 format  
  xlt.cfg>contains applicable translation information for Host/Network encodings  
  specific code pages have mappings between client locale and UTF-16
- Mechanism will be that any time that XLT files are updated (such as uc\_config –update), the files will be verified against CRC32 Checksums and stored in xlt.crc
- When corruption is detected, CS will log an event, CallHome, and copy source files from /nas/site/locale to Data Mover
- If etc\_common requires fsck, there is no easy method except to take all DMs offline

## **CLI CHECKPOINT SCHEDULING GNAPA:**

- Similar functionality as in Celerra Manager using the # nas\_ckpt\_schedule command, using 24-hour clock times
- # **nas\_ckpt\_schedule –list** [outputs all Checkpoint schedules] | **-info id=x** [outputs Checkpoint schedule for PFS] **-modify**
- Note:** -create used to create single use or recurrent checkpoints; -modify used to change active & pending schedules

### **Creating Schedule:**

# **nas\_ckpt\_schedule –create ckpt\_sched –filesystem id=x –description “my new schedule” –recurrence once –ckpt\_name <ckpt\_name> -start\_on 2007-3-17 –runtime 13:15**

**Note:** Schedules can also be created for daily, weekly, or monthly periodicity

Daily Schedule:

# **nas\_ckpt\_schedule –create –recurrence daily –end\_on 2007-12-31 –runtimes 08:00,15:00,23:00 –keep 21**

Weekly Schedule:

# **nas\_ckpt\_schedule –create –recurrence weekly –days\_of\_week Fri,Thu,Wed,Tues,Mon –runtimes 17:30 –keep 20**

Monthly Schedule:

# **nas\_ckpt\_schedule –create –recurrence monthly –days\_of\_month 1,15 –runtimes 3:30 –keep 6**

Other Options:

-delete to remove a schedule

-pause will skip the next scheduled checkpoint -resume will restart the schedule

## **CELERRA PAHC--PROACTIVE HEALTH CHECK NAS 5.5.27.5:**

**Note:** Pre & Up = PUHC; Pro = PAHC

**PAHC AUTO-CHECKUP runs every two weeks!**

### **PAHC CRON JOB:**

# **cat /nas/site/cron.d/nas\_sys|grep auto**

31 3 \* \* 7 root /nas/tools/up > /dev/null 2>&1

→In the above example, the PAHC cron fires at 3:31 a.m. every Sunday, but, the auto\_checkup script checks to see when the script last ran, and will not run unless it's been two weeks. So, it really runs the PAHC checks every two weeks, not every week as the CRON job suggests

### **/nas/tools/auto\_checkup**

```
touch /nas/site/.checkup.autorun
      /nas/bin/nas_checkup -auto > /dev/null 2>&1
    fi
    # if it is > 2 wks old, touch it and run pahc
    # otherwise do nothing
```

### **SCRIPT USED TO RUN PAHC:**

### **/nas/bin/nas\_checkup**

#### **REVIEWING CURRENT PAHC CHECK ERROR NUMBERS & MESSAGES:**

### **/nas/tools/upgrade\_check\_tools/messages.txt**

**Note:** Review the messages.txt file for verbose output of PAHC checks and messages

### **Example:**

```
~dm::check_for_excessive_memory_utilization~
--abridged----
```

3 | external | Cannot determine the memory usage status. | Run "%s" manually to investigate further. The output from the above command should be in the following format: \* "type,slot,histogram"

"cached,2,0:0:0:1629:0:0:0:0:0:0:0:0:0:0:0:2:3:4"

### **/nas/tools/upgrade\_check\_tools/checks.txt**

**Note:** Review this file to when, what NAS version, and on what platforms a check is run on

**Example:**

[\*\*/nas/tools/upgrade\\_check\\_tools/checks.txt file\*\*](#)

```
#check           | modes | start | remove | nbsacc | prim | sec | platform ~
dm::check_for_excessive_memory_utilization | pre+pro | 5.6 | | yes | no | ns+ns40+nsx+cfs ~
```

**LOGS CREATED WHEN PAHC RUNS:**

```
-rw-rw-r-- 1 nasadmin nasadmin 8803 Feb 22 03:32 checkup.090222-033102.log -->PAHC scheduled checkup
-rw-rw-r-- 1 nasadmin nasadmin 8803 Mar 8 03:32 checkup.090308-033101.log -->PAHC scheduled checkup
-rw-rw-r-- 1 nasadmin nasadmin 27403 Feb 22 03:32 .checkup_support.090222-033102.log -->PAHC scheduled checkup
-rw-rw-r-- 1 nasadmin nasadmin 27403 Mar 8 03:32 .checkup_support.090308-033101.log -->PAHC scheduled checkup
Note: When nas_checkup is run manually or as a result of a NAS Install, the word "run" is included in the filename
-rw-rw-r-- 1 nasadmin nasadmin 8803 Mar 13 16:29 checkup-run.090313-162800.log -->Manual /nas/bin/nas_checkup
-rw-rw-r-- 1 nasadmin nasadmin 27403 Mar 13 16:29 .checkup_support-run.090313-162800.log -->manual /nas/bin/nas_checkup
```

→CLI version of PAHC tool runs On-demand or at end of NAS Install from /nas/bin/

→Runs as a Cron job or from CLI on-demand [CRON=auto\_checkup; On-demand=nas\_checkup, but really calls /nas/tools/check\_nas\_upgrade using "pro" switch]

→Reports are collected by SYR

→The scheduled Cron is called the "auto\_checkup", alerts will be seen in Celerra Manager, nas\_checkup will create two logs each time it is run (one log in normal line output & one log in XML output for CallHome--system will CallHome when appropriate event)

**PAHC LOGGING:**

[\*\*# nas\\_logviewer -t /nas/log/sys\\_log |grep Scheduled\*\*](#)

Mar 8 03:32:57 2009:87519264880::Scheduled nas\_checkup has posted a warning. See /nas/log/checkup.090308-033101.log for more details.

[\*\*# strings /nas/log/webui/alert\\_log|grep checkup\*\*](#)

```
/nas/log/checkup.090208-033102.log
/nas/log/checkup.090222-033102.log
/nas/log/checkup.090308-033101.log
```

**Celerra Manager:**

Scheduled nas\_checkup posted a warning. EMC Customer Service will contact you. Please see /nas/log/checkup.Mar-11-3:34:02.log.

**Manual On-Demand Run of Nas Checkup & Logs generated (or by NAS Install script at end of a new installation):**

```
-rw-rw-r-- 1 nasadmin nasadmin 3408 Mar 22 22:38 checkup-run.Mar-22-22:36:16.log →Results of each of the 41+ checks
-rw-rw-r-- 1 nasadmin nasadmin 13808 Mar 22 22:38 .checkup_support-run.Mar-22-22:36:16.log →XML CallHome format
```

[\*\*# head checkup-run.Mar-22-22:36:16.log\*\*](#)

Check Version: 5.5.27.5

Check Command: /nas/bin/nas\_checkup

Check Option: pro

nas\_version:5.5

nas\_release:27-5 slot:slot\_0 platform:NS subplatform:NS-SLEDGEHAMMER

-----Checks-----

2007-03-22T22:36:23-0400 cs::check\_if\_minimum\_free\_space\_exists\_ns: (Pass Pass Pass Pass Pass Pass)

2007-03-22T22:36:23-0400 cs::check\_if\_enough\_free\_space\_exists\_ns: (Pass Pass Pass Pass Pass) -----output abridged-----

[\*\*# head -20 .checkup\\_support-run.Mar-22-22:36:16.log\*\*](#)

```
<CHECK>
<START_TIME>2007-03-22T22:36:23-0400</START_TIME>
<CATEGORY>Control Station</CATEGORY>
<CHECK_NAME>check_if_minimum_free_space_exists_ns</CHECK_NAME>
<CHECK_SEVERITY>Error</CHECK_SEVERITY>
<END_TIME>2007-03-22T22:36:23-0400</END_TIME>
<CHECK_RESULT>Pass</CHECK_RESULT>
```

**Auto Checkup Run & Logs generated:**

[\*\*# cat checkup.Mar-11-03:45:02.log\*\*](#)

-----Warnings-----

Storage System: Check if FLARE is supported

Symptom: Backend Storage Requirements Check Failed:

REQUIREMENT: CX600\_Software >= 2.19 AND CX600\_Software < 2.25

ID: APM00030600863 FOUND: 02.16.600.5.024 INSTRUCTIONS: Upgrade Flare.

\*Action : If you are setup for Call-Home: Your EMC Service Provider will be in contact with you in order to resolve this issue. If you are not setup for Call-Home: Contact EMC Customer Service and refer to EMC Knowledgebase emc146016. Include this log with your support request.

-rw-rw-r-- 1 nasadmin nasadmin 3408 Mar 23 16:51 checkup.Mar-23-16:50:01.log  
-rw-rw-r-- 1 nasadmin nasadmin 13808 Mar 23 16:51 .checkup\_support.Mar-23-16:50:01.log

### **MANUALLY RUNNING PAHC HEALTHCHECK:**

# **/nas/bin/nas\_checkup | -drinkmoreovaltine | -debug | -auto** (don't use this from CLI) | **-help | -version**

**Note:** Run –drinkmoreovaltine switch when troubleshooting output issues, will provide suggested fixes

# **/nas/bin/nas\_checkup** [Performs checks on CS, DM, and Storage system]

#### **Control Station Checks:**

- Check if minimum free space exists
- Check if minimum free space exists ns
- Check if enough free space exists
- Check if enough free space exists ns
- Check if NAS Storage API is installed correctly
- Check if NAS Storage APIs match
- Check if NBS clients are started
- Check if NBS configuration exists
- Check if NBS devices are accessible
- Check if NBS service is started
- Check if standby is up
- Check if Symapi data is present
- Check if Symapi is synced with Storage System
- Check integrity of NASDB
- Check if primary is active
- Check all callhome files delivered
- Check if NAS partitions are mounted

#### **Data Mover Checks:**

- Check boot files
- Check if hardware is supported
- Check if primary is active
- Check if root filesystem has enough free space
- Check if using standard DART image
- Check MAC address
- Check network connectivity
- Check status

#### **Storage System Checks:**

- Check disk emulation type
- Check disk high availability access
- Check disks read cache enabled
- Check disks and storage processors write cache enabled
- Check if access logix is enabled
- Check if FLARE is committed
- Check if FLARE is supported
- Check if microcode is supported
- Check no disks or storage processors are failed over
- Check that no disks or storage processors are faulted
- Check that no hot spares are in use
- Check that no hot spares are rebuilding
- Check control lun size
- Check if storage processors are read cache enabled

→Default schedule every other Sunday at 3:30-4:00 a.m. /nas/site/cron.d/nas\_sys

**Note:** The actual trigger time in minutes is randomly assigned at install time between 3:30-4:00 a.m., so each system will have slightly different times

# cat /nas/site/cron.d/nas\_sys

**33 3 \* \* 7 root /nas/tools/auto\_checkup > /dev/null 2>&1**

### **CHANGING THE PRESET PAHC CRON SCHEDULE:**

1. vi edit the entry in /nas/site/cron.d/nas\_sys to the desired time & day
2. Remove following file:  
# rm -f /nas/site/.checkup.autorun →File exists only if scheduled job has run at least once
3. Restart cron daemon:  
#/sbin/service crond restart or # touch /etc/crontab

4. Confirm change: /nas/log/sys\_log or /nas/log/nas\_checkup.log

### **NAS EVENTS FOR NAS CHECKUP:**

#### **# nas\_event -l -a callhome |grep Checkup**

```
Checkup    221 Scheduled nas_checkup posted information. EMC Customer Service has been contacted. See nas_checkup log.
Checkup    222 Scheduled nas_checkup posted a warning. EMC Customer Service has been contacted. See nas_checkup log.
Checkup    223 Scheduled nas_checkup discovered an error. EMC CustomerService has been contacted. See nas_checkup log.
```

#### **# nas\_event -list -a callhome |grep -i checkup**

```
|-> Checkup(142)
223     ERROR(3)     Scheduled nas_checkup has discovered an error. The file ${src_file_path},8,%s} with the extension
${dest_extension},8,%s} is attached.
```

#### **\$ nas\_event -list -f Checkup (NAS 5.5)**

#### **# nas\_event -list -c CS\_PLATFORM -f Checkup (NAS 5.6 syntax)**

```
id      description
10     nas_checkup found no problems.
11     nas_checkup posted information. See /nas/log/nas_checkup-run.*.log.
12     nas_checkup posted a warning. See /nas/log/nas_checkup-run.*.log.
13     nas_checkup discovered an error. See /nas/log/nas_checkup-run.*.log.
19     nas_checkup had to exit before performing any checks.
21     nas_checkup posted information. Contact EMC Customer Service and refer to EMC Knowledgebase emc146016.
22     nas_checkup posted a warning. Contact EMC Customer Service and refer to EMC Knowledgebase emc146016.
23     nas_checkup discovered an error. Contact EMC Customer Service and refer to EMC Knowledgebase emc146016.
110    Scheduled nas_checkup found no problems.
111    Scheduled nas_checkup posted information. See /nas/log/nas_checkup.*.log.
112    Scheduled nas_checkup posted a warning. See /nas/log/nas_checkup.*.log.
113    Scheduled nas_checkup discovered an error. See /nas/log/nas_checkup.*.log.
119    nas_checkup had to exit before performing any checks.
121    Scheduled nas_checkup posted information. EMC Customer Service will contact you. See nas_checkup log.
122    Scheduled nas_checkup posted a warning. EMC Customer Service will contact you. See nas_checkup log.
123    Scheduled nas_checkup discovered an error. EMC Customer Service will contact you. See nas_checkup log.
221    Scheduled nas_checkup posted information. EMC Customer Service has been contacted. See nas_checkup log.
222    Scheduled nas_checkup posted a warning. EMC Customer Service has been contacted. See nas_checkup log.
223    Scheduled nas_checkup discovered an error. EMC Customer Service has been contacted. See nas_checkup log.
```

#### **# cat /nas/sys/nas\_eventlog.cfg**

```
# CS_PLATFORM:Checkup
#
facilitypolicy 6:142, 7
    disposition range=1-199, logfile "/nas/log/sys_log"
    disposition range=112-113, mail user
    disposition range=122-123, mail user
    disposition range=200-300 severity=3-3, callhome binary
    disposition range=301-301 severity=6-6, callhome immediate
## need to check later
```

### **Adding email notification for PAHC for Customers:**

- Add email notifications via Celerra Manager from new facility called ‘Checkup’, specify severity and mail action

- Using CLI

- check to see if checkup\_eventlog.cfg file is loaded (# nas\_event -L -i)
- If loaded, unload the checkup\_eventlog.cfg file before editing (# /nas/bin/nas\_event -Unload /nas/site/checkup\_eventlog.cfg)
- vi edit the /nas/site/checkup\_eventlog.cfg file and add email addresses [format?]
- Reload the /nas/site/checkup\_eventlog.cfg file (#/nas/bin/nas\_event -Load /nas/site/checkup\_eventlog.cfg)
- Verify that file loaded using # nas\_event -L -i
- Confirm change by posting an event for 112 or 113 (# /nas/sbin/postevent -f 142 -i 112 -s 4)

### **SAMPLE CHECKUP\_EVENTLOG.CFG FILE AFTER NAS UPGRADE:**

#### **# cat /nas/site/checkup\_eventlog.cfg**

```
# Checkup
# This file provides you an example of how to receive email notification for aut
omatic nas_checkup
# You can change the email address 'root@localhost' to your desired address
# To activate this feature, you need to run the following command
# /nas/bin/nas_event -L /nas/site/checkup_eventlog.cfg
```

```
# To disable it, you need to run the following command  
# /nas/bin/nas_event -U /nas/site/checkup_eventlog.cfg  
facilitypolicy 142,6  
    disposition range=112-113 severity=3-4, mail "root@localhost"  
    disposition range=122-123 severity=3-4, mail "root@localhost"
```

### **SYR REPORTING:**

→For NAS 5.6, search in SYR for “CS\_PLATFORM\_CHECKUP:223”

### **CELERRA SETUP WIZARD IMPROVEMENTS GNAPA:**

→Mainly just minor usability improvements, and CIFS service will start automatically after successful Join

### **XML API v2 SERVER GNAPA:**

→Ongoing effort to join the XHMP JServer interface with the APL schema, compatible with active & passive management of Celerra

→XML API v2 will be built on top of APL Middleware and JServer, used specifically for monitoring & managing Celerra

→Introduced for 3<sup>rd</sup> Party API integrations, may later replace JServer XML API V1 (aka XHMP)

→HTTP(S) will be transport between XML client and Celerra

→XML API Server talks to both APL & XHMP (JServer), all of which are run on the Control Station

→XML API will receive indications from APL objects and JServer

→XML API talks to APL for config info and for active management of Celerra components

→XML queries will be synchronous, but XML operations will be asynchronous—XML API Client will submit operation request and internal task is created on Celerra. Completed tasks will have task indication returned to XML client as an indication

→Multiple requests can be put into XML API packets:

Query—query for instances of Celerra objects

QueryStats—query for history of statistics

Task—management action requested, such as creating new file system

ScheduleTasks—scheduling of tasks

Subscribe—subscribe for indications, which will be for Stats and Tasks for Burgundy release, Clients will register and receive updates

Unsubscribe—Unsubscribe for indications

→Burgundy release of XML API will be limited and cannot do Replication, iSCSI, VTLU, Licensing, Usermapper, CAVA, HighRoad, FileMover, or CDMS APIs

→Provides for configuration info, statistics, management, and indications of changes for Celerra

→GrandNapa improvements are in iSCSI Management, support for File System auto-extension, Symmetrix queries, Clariion queries, and additional statistics

→New public API interface on Control Station using XML data carried over HTTPS & Interfaces defined using XML schema

→Protocol between Client and XML API in two parts: Request/Reply (HTTP POST) & Indications (HTTP GET)

**Note:** Currently, the XML API v2 interface is not enabled on the Celerra—will be enabled in GrandNapa

### **XML API SERVLET:**

→CelerraManagementServices servlet runs on Control Station and will listen for https requests from clients, forward to API Server, and then return results from API Server to clients

### **STARTING XML API V2 SERVER:**

**# /nas/sbin/start\_xml\_api\_server**

**# cat /nas/sys/nas\_mcd.cfg**

```
daemon "XML API Server"  
executable  "/nas/sbin/start_xml_api_server"  
optional    yes  
canexit     yes  
autorestart yes  
ioaccess    no
```

### **STOPPING XML API V2 SERVER:**

**Note:** Comment out the nas\_mcd.cfg entries and use following command to stop

**# /nas/sbin/hup\_api** [Also use to restart the server without commenting out entries]

### **XML API V2 CONFIGURATION FILE:**

#### **CONFIG FILE:**

**# cat /nas/sys/xml\_api.conf**

```
jsserver.req.server.port = 8020  
xml.api.req.server.port = 8050  
xml.api.ind.server.port = 8051  
xml.api.server.log = log/cel_api.log  
xml.api.server.log.generations = 5  
xml.api.servlet.log = log/webui/cel_api.log
```

```
xml.api.servlet.elog = log/webui/cel_api_error.log
xml.api.servlet.logmask = 0
apl.ind.server.port = 8886
xml.api.user.debug.flag = false
xml.api.user.request.validation.flag = true
xml.api.trace.apl.calls = true
xml.api.enable.indications.ext = true
xml.api.trace.apl.indications = true
xml.api.trace.user.requests = true
xml.api.quota.poll.offset = 150
```

### **/nas/sys/xml\_api.conf**

→Config file contains information on Ports used by various applications (JServer & XML API Server) & configuration params  
→Any changes to the config would require restart of the XML API v2 Server  
→Port 8020 is JServer port used by Celerra Monitor, Celerra Manager Tomcat servlet, and XML API Server  
→Port 8050 is XML Server port used in socket creation for sending Tomcat servlet requests, & Indications from JServer, to XML Server  
→Port 8051 is XML Server port used in socket creation for receiving Tomcat servlet indications from XML API Server

### **# cat /nas/sys/xml\_api.conf**

```
# This file defines properties that are used by Tomcat servlet component,
jsrver.req.server.port = 8020
xml.api.req.server.port = 8050
xml.api.ind.server.port = 8051
xml.api.server.log = log/cel_api.log
xml.api.server.log.generations = 5
xml.api.servlet.log = log/webui/cel_api.log
xml.api.servlet.elog = log/webui/cel_api_error.log
xml.api.servlet.logmask = 0
apl.req.server.port = 9824
apl.ind.server.port = 8886
xml.api.user.debug.flag = false
xml.api.user.request.validation.flag = true
xml.api.trace.apl.calls = true
xml.api.enable.indications.ext = true
xml.api.trace.user.requests = true
xml.api.quota.poll.offset = 150
```

### **XML API V2 Log Files:**

/nas/log/webui/cel\_api.log (Servlet log file)  
/nbsnas/log/webui/cel\_api\_error.log (Servlet log file)  
/nas/log/cel\_api.log (Server log file, logs requests from XML users, requests & replies to APL, & APL indications received)

### **XML API V2 JAR FILE:**

### **/nas/api**

```
# ls -la
-rw-rw-r-- 1 root    sys      308232 Dec 19 21:37 api_server.jar
-rw-r--r-- 1 root    root        0 Jan 14 02:21 jvm.err
-rw-r--r-- 1 root    root        0 Jan 14 02:21 jvm.out
```

**Note:** If XML API V2 Server fails to restart after 3 tries, an “api\_retry” file is created in /nas/api—retries every 20 minutes

### **XML API V2 Server Processes:**

### **# ps fax|grep api**

```
2220 ? S 0:00 \_ /bin/sh /nas/sbin/start_xml_api_server
```

### **# ps fax |grep java |grep -i xmx180**

```
2240 ? S 0:07 \_ /usr/java/bin/java -server -Xmx180m -Xss10485
2337 ? S 0:00 \_ /usr/java/bin/java -server -Xmx180m -Xss1
2339 ? S 37:10 \_ /usr/java/bin/java -server -Xmx180m -
2342 ? S 0:00 \_ /usr/java/bin/java -server -Xmx180m - [many more processes running]
```

### **Special Debug Flags for xml\_api.conf that are not enabled by default:**

xml.api.user.debug.flag →Java stack traces  
xml.api.trace.apl.calls →logs requests & responses to APL  
xml.api.trace.apl.indications →logs indications from APL  
xml.api.trace.user.requests →logs XML API user requests

## **CLARIION SECTOR ERROR PHASE I PRIME GNAPA:**

- Flare 19 Patch 030 minimum
- Data Mover will panic for 03/11 error but not failover, and will reboot and recover—only affected fs will be <unmounted>
- Key point is that the list of BadBlocks is obtained directly from CLARiiON, and the cse\_recover uses CLARiiON to repair
- Repair using “cse\_recover server\_x bvid=x phyblkno=yyy” derived from panic string, sys\_log, or “mepanic dump” output
- Note:** cse\_recover will require Username & Password to activate if Clariion security is set
- Run fsck (includes MapBlock—maps logical blocks to actual objects in file system)
- Remount file system
- Create new Checkpoint and then refresh others (if Checkpoints are involved)
- Abort replication session on each side, then restart from scratch
- Provide Report.txt to customer

### **# /nas/sbin/cse\_recover**

```
*****/nas/sbin/cse_recover version 2.0 *****  
* Usage: /nas/sbin/cse_recover  
    <movername>  
    bvid=<Basic_LUN_number>  
    phyblkno=<Bad_block_number>  
*****
```

## **CELERRA SUPPORT FOR SECURE CLI:**

- 5.5.27.5 supports navisecli on Gateway Celerras (when used by the Clariion environment)
- 5.5.30.4 supports navisecli on Integrated Celerras (sets up Array security domain on new installs)
- Beginning with 5.5.30, /nas/sbin/navicli has become a link which calls /nas/sbin/nas\_navi. The nas\_navi executable checks for local security files on the Control Station (/nas/site/.clar\_security or /etc/.clar\_security), and if finding the username and password entries, will pass the appropriate credentials to the array via /nas/opt/Navisphere/bin/navisecli. If the local CS security cache files are not present, the nas\_navi command will route the command to /nas/sbin/classic\_navicli.

### **Debugging Issues:**

```
# export SYMAPI_DEBUG=-1  
# export SYMAPI_DEBUG_FILENAME=/tmp/symapi_debug.log  
# /nas/symcli/bin/symcli
```

Symmetrix Command Line Interface (SYMCLI) Version V6.4.2.17 (Edit Level: 845)  
built with SYMAPI Version V6.4.2.17 (Edit Level: 845)

```
# /nas/sbin/navicli -help |head -3
```

@(#)Navisphere CLI Revision 6.26.0.5.1 for Linux on 8/20/2007 8:12:01

Copyright (c) 2007 EMC Corporation. All Rights Reserved

Release (output) level: 26 (17):

## **SECURE NAVICLI SUPPORT GNAPA:**

- GNapa 5.5.27 Gateway systems support Secure CLI, Secure Navi, navisecli, etc.—Integrateds will support with Napa 8 release (5.5.30)

**Note:** /nas/sbin/navicli is replaced after GNapa upgrade with nas\_navi executable. nas\_navi will be used to check for cached credentials on the Celerra, read the username & password, fill in the credentials, then send Secure CLI commands to array (/nas/sbin/navisecli)

```
-rwxr-xr-x 1 root root 1691996 Jan 16 10:20 nas_navi
```

```
lrwxrwxrwx 1 root root 8 Feb 2 15:14 navicli -> nas_navi
```

- Flare 19 introduced the Secure Navi (navisecli) commandset based on the Navisphere 6.x security model (role-based management; auditing; SSL protection; centralized account management)—provides for connection security to arrays using navisecli

→Replacement for the less secure Classic NaviCLI

→Celerra will cache the Global username and password on the Control Station during install using encryption

→Use # nas\_storage –modify –security to setup or change authentication to Clariion backends

→Integrated System would not normally require the use of CLARiiON username and password—will use the Classic NaviCLI

→Navisecli is only used or required if the Backend has security initialized

→Navisecli will be located in following directories: /nas/opt/Navisphere/bin/navisecli; /nas/sin/navisecli; /nasmcd/sbin/navisecli

→For installs, only the boot CLARiiON will require security—additional backends will have security setup afterwards

### **Using /nas/sbin/navisecli (Host-based) Commands:**

Storage systems will not allow the use of Host-based Secure CLI (navisecli) commands unless a security file is created on the array with either a Global or Local context (0, 1 respectively). For Global security, the username, password, & scope must match that of the Domain security account. Use the -AddUserSecurity to create the security file for the user account on the Host running the SecureCLI

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
commands. When running naviseccli commands from the Celerra Control Station, with or without an array security model in place, the user would have to add the CS Host to the array security file using –AddUserSecurity.

## **FLARE 28 BEHAVIOR:**

There are situations where the security information on the array has been damaged or otherwise removed, that may subsequently cause navicli commands to output the following Error, particularly if changing the IP Addresses on the SPs, or using the nas\_raid –s script to perform a backend cleanup:

**# ./navicli -h 128.221.252.200 networkadmin -get**

Error 12031: CLARiiON 128.221.252.200 running FLARE R28 requires security for "navicli" commands.

See emc211557 and AR141973. The solution should be to run nas\_storage –modify id=x –security to recreate and synchronize the array security information in the Celerra security cache files (.clar\_security) and update the storageAPI password in /nas/symapi/config/emcpwddb.dat file. Alternatively, running **/tftpboot/bin/setup\_clariion\_security <spa\_ip> <spb\_ip> nasadmin nasadmin -initialize** should also work.

## **REINITIALIZING ARRAY SECURITY DOMAIN:**

**Note:** In some cases the security domain may need to be destroyed on the array, then re-created, before the Celerra can authenticate with the Array

1. Open a web browser and go to the setup page for SP A:  
[http://<Current\\_IP\\_Address\\_of\\_SP\\_A>/setup](http://<Current_IP_Address_of_SP_A>/setup)
2. Scroll to the bottom of the screen and click Destroy Security and Domain Information.
3. A warning screen will appear. Select Yes and click Submit to reset the security and domain information for SP A.
4. Wait for the management server to stop and then restart.
5. Go back to the setup page for SP A:  
[http://<Current\\_IP\\_Address\\_of\\_SP\\_A>/setup](http://<Current_IP_Address_of_SP_A>/setup)
6. Scroll down and click Restart Management Server.
7. A warning screen will appear. Select Yes and click Submit to restart the management server.
8. Repeat the previous step to reset the security and domain information for SP B.
9. After the management server has been restarted on SP B, the array's security and domain information will be in an un-initialized state, similar to the way it is shipped from the factory.
10. Change IPs on the SPs (if required)
11. Use Navisphere Manager to connect to SP A and re-create the Domain Security domain
  - a. The “Confirm: Navisphere Security” dialog box opens, indicating that global security is not initialized. Select Yes to reinitialize the array's security settings.
  - b. Enter the global user name and password that was originally in use on the array.
  - c. Log in to the array with the user name and password you just entered.
  - d. From the Navisphere Manager main menu, select File > Set Up Domain > Select Master. The Select Master dialog box appears.
  - e. Highlight the IP address of SP A and click OK.

## **IMPLEMENTING THE CLARiiON DOMAIN SECURITY MODEL ON CELERRA**

### **INTEGRATEDS & GATEWAY SYSTEMS:**

Beginning with NAS 5.5.30, all NS20, NS40, NS350, NS80 integrateds will install and setup the Clariion Domain Security model, using a default username “nasadmin” and password “nasadmin”. Additionally, the Clariion User account and Password information is cached on the Celerra Control Station in the form of a /has/site/.clar\_security file, which is then used internally by the Control Station to authenticate to the array in support of the Secure NaviCLI model.

**# /tftpboot/bin/setup\_clariion\_security <spa\_ip> <spb\_ip> nasadmin (user) nasadmin (password) –initialize**

**Note:** -initialize found with 5.5.32, and 5.6.36 and above

**# nas\_storage –modify <array\_ID> –security –username nasadmin –password nasadmin**

In addition, any password changes to the backend that may be made as a result of the CSA configuration process, are synchronized on the Celerra Control Station using nas\_storage –modify –security. An unresolved issue is that the Clariion security model has not been implemented on any Integrated system installed prior to 5.5.30, unless someone has consciously setup Clariion Domain security after the initial installation.

**Typical Symptoms when Array security has not been established on an Integrated System:**

**nas\_diskmark, server\_devconfig, or Celerra Manager "Rescan" disk discovery operations fail:**

**ERROR: 3501: Storage API Code=3593: sysmapi\_C\_clariion\_load\_error**

\$ nas\_storage -c -a

Discovering storage (may take several minutes)

**Error 3501: Storage API code=3593: SYMAPI\_C\_CLARIION\_LOAD\_ERROR**

An error occurred while data was being loaded from a Clarion

**# /nas/sbin/navicli -h 192.168.1.200 getcrus**

SPE3 Invalid Enclosure

```
Invalid SP State: Present
Invalid SP State: Present
Invalid Power State: Present
Invalid Power State: Present
Invalid Power State: Present
Invalid Power State: Present
Invalid SPS State: Present
Invalid SPS State: Present
Invalid SPS Cabling State: Valid
Invalid SPS Cabling State: Valid
```

**Note:** Please see emc193775 (AR122612) for guidance on permanently implementing Array security after upgrading to NAS 5.5.32.4 or higher. See emc194888 for an interim procedure for setting up Navisecccli security for use with the SYMAPI on systems prior to 5.5.32. See emc203143 if implementing CLARiiON security for the first time on Celerra Integrateds running NAS 5.5.37+ or 5.6.40+ (AR122612).

**UPDATING/CHANGING/SYNCHRONIZING INTEGRATED ARRAY SECURITY INFORMATION:**

The preferred method to use for changing the array security account password on the Integrated platform is to use the following syntax, since this will automatically change the password on the array while simultaneously updating the Control Station security cache files (.clar\_security) and symapi database password (emcpwddb.dat):

**# nas\_storage -modify id=1 -security -username nasadmin -password nasadmin -newpassword <passwd>**

If you simply want to synchronize the Celerra security cache and symapi db with the password on the array, without changing the array password, use the following command:

**# nas\_storage -modify id=1 -security** (enter username and password at the prompt)

If the User wants to use Navisecccli commands directly via Control Station CLI, you must either specify the array security account name and password as part of the command syntax, or perform a one-time add of the local host information to the array security file, using the following commands, respectively:

**# /nas/sbin/navisecccli -h 10.241.168.183 -user nasadmin -password nasadmin -scope 0 <command>**

**# /nas/sbin/navisecccli -h 10.241.168.179 -AddUserSecurity -scope 0 -user nasadmin**

**I. IMPLEMENTING THE CLARiiON DOMAIN SECURITY MODEL ON PRE-5.5.32.4****INTEGRATEDS (emc194888):**

1. Enable CLARiiON Navisecccli authentication for the SYMAPI using the user nasadmin with password nasadmin:

```
# /usr/symcli/bin/symcfg authorization add -hostname 192.168.1.200 -username nasadmin -password nasadmin
# /usr/symcli/bin/symcfg authorization add -hostname 192.168.2.201 -username nasadmin -password nasadmin
```

2. Set correct permissions and ownership on the SYMAPI password file:

```
# chmod 664 /nas/symapi/config/emcpwddb.dat
# chown nasadmin:nasadmin /nas/symapi/config/emcpwddb.dat
```

```
# ls -la emc*
```

```
-rw-rw-r-- 1 nasadmin nasadmin 230 Aug 12 12:35 emcpwddb.dat
```

3. Confirm the effectiveness of the workaround:

**# nas\_storage -check -all**

Discovering storage (may take several minutes)

done

4. Verify the SYMAPI security credentials:

**# /usr/symcli/bin/symcfg authentication list**

| Hostname      | Username | Namespace | Port |
|---------------|----------|-----------|------|
| 192.168.1.200 | nasadmin |           |      |
| 192.168.2.201 | nasadmin |           |      |

5. At the customer's convenience, schedule a NAS upgrade to 5.5.32 or later, then implement the fix to permanently setup the Array Domain Security Model using emc193775 (applies to NAS 5.5.32-5.5.36 or 5.6.36-5.6.39) or emc203143 (applies to NAS 5.5.37+ or 5.6.40+).

**Note:** One of the byproducts of this issue is that the getcrus command may still show “Invalid” states for components. If this is seen, enable IP Forwarding:

```
# /nas/sbin/navicli -h 192.168.1.200 getcrus
SPE3 Invalid Enclosure
Invalid SP State: Present
# echo 1 > /proc/sys/net/ipv4/ip_forward
# cat /proc/sys/net/ipv4/ip_forward
1
```

## **II. IMPLEMENTING THE CLARiiON DOMAIN SECURITY MODEL ON 5.5.32-5.5.36, or-5.6.36-**

### **5.6.39 INTEGRATEDS (emc193775):**

1. Login to the Control Station as nasadmin, then su to Root.
2. Check for the existence of the /tftpboot directory on the Control Station:

```
# cd /
# ls -->look for the tftpboot directory in the ls output
tftpboot
```

3. If the tftpboot directory is not present, create the directory by using the following utility (otherwise continue to Step 4):

```
# /nasmcd/sbin/t2pxe -tftp start
```

```
=====
EMC NAS PXE-BOOT Setup
=====
```

Starting TFTP Service...

Done.

4. Enable CLARiiON security using the following example (SP IP Addresses may be different--verify IP addresses in the /etc/hosts file):

```
# /tftpboot/bin/setup_clarion_security 192.168.1.200 192.168.2.201 nasadmin nasadmin -initialize
```

**Note:** No output returned by command. The -initialize option is new with 5.5.32, and sets up the Array account, then creates the .clar\_security cache files on the Control Station. It does not, however, update the emcpwddb.dat cache, which is the reason why nas\_storage -m -s needs to be run in Step 5. Beginning with NAS 5.6.40, the nas\_storage -m -s command will be able to create an Array account for Integrated and FC-Enabled models, and also update the .clar\_security and /nas/sympapi/config/emcpwddb.dat files

5. Synchronize the CLARiiON Security credentials with the Celerra Control Station cache files:

```
# nas_storage -modify <array_id or name> -security -username nasadmin -password nasadmin
```

Setting security information for SL7E1080700121

done

**Note:** Use nas\_storage -l to identify Array ID or Name

6. If t2pxe was used to create the tftpboot directory, stop PXE services now:

```
# /nasmcd/sbin/t2pxe -tftp stop
```

```
=====
EMC NAS PXE-BOOT Setup
=====
```

Stopping TFTP Service...

Done.

7. Verify that the array security credentials have been properly established and run nas\_storage -check -all:

```
# /nas/sbin/navicli -h 192.1.4.231 security -list
```

Username: nasadmin

Role: administrator

Scope: global

```
# nas_storage -info id=1 |grep -C1 username
```

authenticated = True

username = nasadmin

```
# nas_storage -check -all
```

Discovering storage (may take several minutes)

Done

## **III. IMPLEMENTING THE CLARiiON DOMAIN SECURITY MODEL ON NAS 5.6.40+ or 5.5.37+**

### **INTEGRATEDS (emc203143, AR122612):**

1. Setup the array security model on the Integrated Celerra

```
# nas_storage -modify id=x -security -username nasadmin -password nasadmin
```

**Note:** Beginning with NAS 5.6.40.0 and 5.5.37.0, the nas\_storage -m -s command will create the Array account (if it doesn't exist) for Integrated and FC-Enabled models, and also updates the appropriate Celerra cache files--clar\_security and /nas/sympapi/config/emcpwddb.dat

## 2. Verify

```
# /nas/sbin/navicli -h 192.1.4.231 security -list
```

Username: nasadmin

Role: administrator

Scope: global

## **IV. IMPLEMENTING THE CLARiiON DOMAIN SECURITY MODEL ON GATEWAY CELERRAS (emc195868):**

### **Option 1:**

With NAS 5.5.27.5 and later, Celerra Gateways support the CLARiiON security model. Use the following procedures to establish security between the Celerra and the array if not already setup.

- Verify that Global Array security has already been setup on the array(s) using Navisphere or CLI (or setup now)
- Use the following command to update or create the security cache files used by Celerra to authenticate internally with the array:

```
# nas_storage -modify <storage_id> -security -username <array_admin_account_name> -password <array_admin_account_password>
```

### **Example when entering incorrect array password:**

```
# nas_storage -modify id=1 -security -username nasadmin -password nasadmin
```

Setting security information for APM00083201184

Error 3502: APM00083201184: Storage API code=4651: SYMAPI\_C\_CLAR\_NOT\_PRIVILEGED

Operation denied by Clariion array - you are not privileged to perform the requested operation

### **Example when using correct array password:**

```
# nas_storage -modify id=1 -security -username nasadmin -password nasadmin1
```

Setting security information for APM00083201184

done

### **Security and Password files updated on Celerra:**

```
# ls -la /nas/site/.clar*          -rw-rw-r-- 1 nasadmin nasadmin 276 Dec  5 12:48 /nas/site/.clar_security
# ls -la /nas/symapi/config/emcpwd* -rw-rw-r-- 1 nasadmin nasadmin 237 Dec  5 12:48 /nas/symapi/config/emcpwddb.dat
```

### **Option 2:**

For NAS 5.5.26.x or earlier, the steps to establish authentication between the Celerra and the array are different and are considered an interim fix only:

### **Interim Fix:**

- Enable CLARiiON authentication for the SYMAPI on the Celerra using the user nasadmin with password nasadmin:

```
# /usr/symcli/bin/symcfg authorization add -hostname 192.168.1.200 -username nasadmin -password nasadmin
```

```
# /usr/symcli/bin/symcfg authorization add -hostname 192.168.2.201 -username nasadmin -password nasadmin
```

- Set correct permissions and ownership on the SYMAPI password file:

```
# chmod 664 /nas/symapi/config/emcpwddb.dat
```

```
# chown nasadmin:nasadmin /nas/symapi/config/emcpwddb.dat
```

```
# ls -la emc*
```

-rw-rw-r-- 1 nasadmin nasadmin 230 Aug 12 12:35 emcpwddb.dat

- Verify communications with the Array from Celerra:

```
# nas_storage -check -all
```

Discovering storage (may take several minutes)

done

- Verify the SYMAPI security credentials:

```
# /usr/symcli/bin/symcfg authentication list
```

| Hostname     | Username | Namespace | Port |
|--------------|----------|-----------|------|
| 10.150.90.10 | nasadmin |           |      |
| 10.150.90.11 | nasadmin |           |      |

- As a permanent fix, once the system has been upgraded to 5.5.27 or later, apply the nas\_storage command to update the Celerra Security cache files. Repeat the command whenever User or Password changes are made to the array's Global Administrator account:

```
# nas_storage -modify <storage_id> -security -username <array_admin_account_name> -password <array_admin_password>
```

## **ENABLING/DISABLING CLASSIC CLI WITH FLARE 26 (emc173093):**

### **Disable ClassicCLI from Navisphere or from Naviseccli command:**

- Navisphere>arrayname>rightclick and select “Enable/Disable Classic CLI” on the fly

- From CLI:

```
# /nas/sbin/naviseccli -h 192.1.4.214 classiccli -disable
```

Do you really want to disable Classic cli(y/n) y

ClassicCLI: Disabled Succesfully

# /nas/sbin/navisecli -h 192.1.4.214 classiccli -status

ClassicCLI ENABLED: No

# /nas/sbin/navisecli -h 192.1.4.214 classiccli -enable

Do you really want to enable Classic cli?(y/n) y

ClassicCLI: Enabled Succesfully

# /nas/sbin/navisecli -h 192.1.4.214 classiccli -disable -o [Use to bypass prompt]

ClassicCLI: Disabled Succesfully

## **SYMPTOM SEEN IF CLASSIC CLI DISABLED OR IF CLARIION SECURITY CACHE FILES**

### **MISSING ON CELERRA:**

# /nas/sbin/classic\_navicli -h 192.1.4.214 getagent

Error returned from Agent

Agent denied request -- Caller not privileged.

# /nas/sbin/navicli -h 192.1.4.214 getagent

Error returned from Agent

Agent denied request -- Caller not privileged.

### **SECURITY CACHE FILES MISSING ON CELERRA CONTROL STATION:**

# nas\_storage -modify <array\_ID> -security -username <array\_admin\_name> -password <array\_password>

Setting security information for SL7E1080700121

done

**Note:** Above command recreates the /nas/site and /etc/.clar\_security files with the proper Username and Passwords for use by the Celerra when authenticating to the Backend. See emc173093 for more information.

### **NAVISECCLI EXAMPLE:**

# /nas/sbin/navisecli -h 192.168.1.200 getagent

Security file does not exist. Use navisecli -AddUserSecurity to create a security file. You must also have a valid user account on the storage system to issue this command. If you do not have a user account, use navisecli -address <IPAddress | NetworkName> security -adduser to add a user account.

**Note:** The above output means that array security is set on the backend and that the navisecli files on the array have not been updated to allow the Secure CLI to run from the Control Station. Use the following command to add the correct password to the navisecli file on the array for the Celerra, specifying the Scope, Username, and Password.

# /nas/sbin/navisecli -h 192.168.1.200 -AddUserSecurity -Scope 0 –user nasadmin

Enter password:

# /nas/sbin/navisecli -h 192.168.1.200 getagent (navisecli now succeeds)

Agent Rev: 6.22.21 (4.0)

Name: K10

Desc:

Node: A-APM00063303725

-----output abridged-----

# /nas/sbin/navisecli -h 192.168.1.200 -user emc -password emc -scope 0 getlog -h -10

**Note:** Scope 0 indicates Global security, while Scope 1 indicates Local array security—running the command for both Scopes is o.k. if unsure.

### **Dummy Celerra Issue:**

# /nas/sbin/navisecli -h 172.24.159.122 -user emc -password emc -scope 0 getlog -h -10

Date Time CRU EvtCd Event Message SnsKey XCd1 XCd2 Source

02/27/2007 18:36:46 (4640)Access is denied to Navisphere CLI called by 'dummy\_celerra' from '172.24.164.13'.

**Note:** If the Celerra needs to authenticate with the array, and a valid user and password are not cached on the CS (.clar\_security files), then the communicate will fail and result in the “dummy\_celerra” message in the SP logs. AR91272 and 5.5.28.0 addresses an issue seen with log flooding of dummy\_celerra messages.

### **Celerra Security Files for Clariion:**

→Kept in two locations: /nas/site/.clar\_security (primary source) & /etc/.clar\_security (used when /nas is unavailable)

-rw-rw-r-- 1 nasadmin nasadmin 552 Feb 2 15:33 .clar\_security

-rw-rw-r-- 1 nasadmin nasadmin 0 Feb 2 15:33 .clar\_security.lck

→The security files are synchronized at Install time, CS failover/failback, and when using nas\_storage –modify –security command

→If security files become corrupted, remove .clar\_security from both locations and run nas\_storage –modify to add back

**Note:** Celerra Manager uses the cached security credentials to log into the Storage systems that are displayed—would see “Unauthenticated” icon if security was set on Clariion but not yet updated on Celerra: Celerras>Storage>Systems><array>

## **BEHAVIOR WHEN USING AN INVALID ARRAY USER ACCOUNT FROM CELERRA:**

### **# nas\_storage -modify id=1 -security**

Enter the Global CLARiiON account information

Username: mark

Password: \*\*\*\*\* Retype your response to validate

Password: \*\*\*\*\*

Setting security information for APM00063303725

Error 3502: APM00063303725: Storage API code=4651: SYMAPI\_C\_CLAR\_NOT\_PRIVILEGED

Operation denied by Clariion array - you are not privileged to perform the requested operation

**Note:** The above error is generated when trying to modify security cache files on the Celerra when the actual user account "admin" does not exist on the array. Corrective action would be to re-issue the command using a valid array account & password, which would then update the /etc.clar\_security files on the Celerra.

### **Log Errors:**

#### **# /nas/sbin/naviseccli -h 192.168.1.200 -user nasadmin -password nasadmin -scope 0 getlog -h -5**

08/23/2007 08:34:50 (4640)Access is denied to Navisphere CLI called by 'dummy\_celerra' from '192.168.1.100'.

#### **cmd\_log**

2007-08-23 11:38:14.193 db:0:2881:S: nas\_storage -modify id=1 –security →Command starts, but never completes

#### **sympapi\_log**

2007-08-23 11:44:18.789 6694 1026 EMC:Celerra CS User Generated

Error returned from the Management Server on 192.168.1.200

Authentication failed. Possible reasons for failure are invalid security file, invalid username, password and/or scope.

2007-08-23 11:44:18.850 6694 1026 EMC:Celerra CS User Generated

/nasmcd/sbin/naviseccli -User 'mark' -Password -Scope 0 -Address 192.168.1.200 -Timeout 30 security -list -user 'mark' 2>&1

## **.clar\_security File Observations:**

### **# /nas/sbin/navicli -h 192.168.1.200 security -list**

/nas/sbin/classic\_navicli [-p] [-vlq] [-m] [-np] [-t timeout] [-h hostname] [-d device] [-help] [-f filename] CMD <optional-args>

**Note:** Above output abridged. The navicli security -list command will only work if there are properly updated local .clar\_security cache files on the Control Station, since this command requires the CS to use naviseccli in the background. Removing both sets of .clar\_security files from /etc and /nas/site directories will reproduce the above output.

### **# /nas/sbin/navicli -h 192.168.1.200 security -list**

Username: nasadmin

Role: administrator

Scope: global

### **Missing .clar\_security files will not affect the ability to run the following commands:**

#### **# nas\_storage -check -all**

Discovering storage (may take several minutes)

done

#### **# nas\_storage -sync -all**

done

## **Three Methods for changing Global Array Security Password/Role:**

1. Navisphere>Tools>Security>Change Password

### **# /nas/sbin/naviseccli -h 10.241.168.179 -user nasadmin -password Engpasswd2 -scope 0 security -changeuserinfo -user nasadmin -scope 0 -newpassword nasadmin**

WARNING: You are about to change user: nasadmin (global)

Proceed?(y/n) y

#### **Alternate Syntax:**

##### **a) # /nas/sbin/naviseccli -h 10.241.168.179 -AddUserSecurity -scope 0 -user nasadmin**

Enter password: <nasadmin>

##### **b) # /nas/sbin/naviseccli -h 10.241.168.179 security -changeuserinfo -user nasadmin -scope 0 -newpassword nasadmin1**

WARNING: You are about to change user: nasadmin (global)

Proceed?(y/n) y

### **3. # nas\_storage -modify id=1 -security -username nasadmin -password nasadmin -newpassword <passwd>**

Changing password on APM00071600514

done

**Note:** Of the (3) methods mentioned above, only the nas\_storage –modify command is recommended since it updates the password on the array, and the local /etc/.clar\_security cache files, whereas the other two methods only change the password on the array itself. The ability to change array security passwords using nas\_storage is limited to the new NS20/20FC/40/40FC models, but an NS80I can have Array security in place and use the nas\_storage –m –security to update the .clar\_security cache files.

## **ADDING OR REMOVING SECURITY INFORMATION FOR NAVISECCLI ON THE ARRAY:**

**# /nas/sbin/navisecccli -h 192.168.1.200 security -list**

Security file not found. Already removed or check -secfilepath option.

**Note:** Above message indicates that the navisecccli security is not set for the Celerra Control Station

**Resolution:**

**Adding User account(s) on the array for the Celerra CS, using Scope & Password Options:**

**# /nas/sbin/navisecccli -h 192.168.1.200 –AddUserSecurity -scope 0 -user nasadmin**

Enter password:

**# /nas/sbin/navisecccli -h 192.168.1.200 security -list**

Username: nasadmin

Role: administrator

Scope: global

**Removing User accounts from Navisecccli files on the array for Celerra:**

**# /nas/sbin/navisecccli -h 192.168.1.200 -removeusersecurity -scope 0**

**Note:** Above command will remove all Global account information from the Navisecccli security files stored on the array

**# /nas/sbin/navisecccli -h 192.1.4.214 security -list**

Security file not found. Already removed or check -secfilepath option.

## **CREATING GLOBAL ARRAY SECURITY ACCOUNT IF NONE EXISTS ON ARRAY:**

**CLI:**

**1. Create User & Initialize security using NAVISECCLI:**

**# /nas/sbin/navisecccli -h 10.241.168.179 security -adduser -user nasadmin -password nasadmin -scope global -role administrator**

WARNING: You are about to add user: nasadmin

Proceed?(y/n) y

**# /nas/sbin/navisecccli -h 10.241.168.179 domain -list**

Node: APM00083201184

IP Address: 10.241.168.179 (Master)

Name: SPA

Port: 80

Secure Port: 443

IP Address: 10.241.168.180

Name: SPB

Port: 80

Secure Port: 443

**Note:** With Flare 28, the –adduser command adds the admin user, initializes Domain security, & sets SPA as the Master

**2. Create User and Initialize security using GUI:**

Navisphere Manager>File>Initialize Security

**3. Create User and Initialize security using Celerra CLI nas\_storage command:**

**CREATE SECURITY DOMAIN, MASTER, & CELERRA CACHE FILES NAS 5.5.37+ or NAS 5.6.40+**

**# nas\_storage -modify id=1 -security -username nasadmin -password nasadmin**

Setting security information for APM00083201184

done

Info 26306750341: Security has been initialized on CLARiiON backend APM00083201184.

**Note:** Command initializes the security domain with SPA as the domain master, creates the array Global Administrator nasadmin, and updates the symapi password and security cache files on the Celerra

# ls -la /nas/site/.clar\*

-rw-rw-r-- 1 nasadmin nasadmin 276 Dec 5 13:24 /nas/site/.clar\_security

-rw-rw-r-- 1 nasadmin nasadmin 0 Dec 5 13:24 /nas/site/.clar\_security.lck

# ls -la /nas/symapi/config/emcpwd\*

-rw-rw-r-- 1 nasadmin nasadmin 234 Dec 5 13:24 /nas/symapi/config/emcpwddb.dat

## SETTING THE DOMAIN MASTER:

```
# /nas/sbin/navisecli -h 192.168.1.200 domain –setmaster 192.168.1.200
```

## ADDING AN SP TO DOMAIN SECURITY:

```
# /nas/sbin/navisecli -h 192.168.1.200 domain –add 192.168.1.200
```

## ADDING ANOTHER GLOBAL USER SECURITY ACCOUNT TO THE ARRAY:

```
# /nas/sbin/navisecli -h 192.168.1.200 –user nasadmin –password nasadmin –scope 0 security -adduser -user second_admin -password second_admin -scope global -role administrator
```

## LISTING DOMAINS:

```
# /nas/sbin/navisecli -h 192.168.1.200 domain -list
```

## DESTROYING DOMAIN SECURITY:

**Note:** Flare 28 doesn't let you remove the only Administrator account on the array, so use the following to destroy domain security & its associated Users:

### EXAMPLE:

```
# /tftpboot/bin/navisecli -h 128.221.252.200 security -messner -rmuser -user nasadmin -scope 0
```

WARNING: You are about to remove user: nasadmin (global)

Proceed?(y/n) y

Delete user operation failed. The selected user is the only global administrator.

## USE TO REMOVE DOMAIN & ALL SECURITY ACCOUNTS:

```
# /nas/sbin/navisecli -h 10.241.168.179 domain -remove 10.241.168.179
```

WARNING: You are about to remove the master node or its peer node. THIS WILL DESTROY THE DOMAIN. Proceed? (y/n) y

**Note:** Using either SP address destroys the global security domain

### Verify:

```
# /nas/sbin/navisecli -h 10.241.168.179 domain -list
```

Domain operation failed. User does not exist.

```
# /nas/sbin/navisecli -h 10.241.168.180 domain -list
```

Domain operation failed. User does not exist.

### GUI:

Optionally can destroy domain security using Navisphere's /setup program>Destroy Security and Domain Information

## ACCESSING SETUP PROGRAM FROM CELERRA CS USING LYNX:

→Use the lynx http client to connect directly to the SP's /setup program over the internal or external network

**Note:** From minimal testing, best results came from using an SSH session to the Control Station using Putty, as it tends to set the right environmental screen variables to handle the following program [i.e., sets TERM=xterm, sets LS\_COLORS=xxxxxxxx, etc.]

```
# lynx http://128.221.252.200/setup
```

→Using the interface is tricky. To log into a system with security initialized, use the down arrow key to get to the user line and enter "nasadmin", then arrow down to password and enter the password, then arrow down to "submit" line, press enter. Arrow down to Host and IP address lines and edit if needed. Arrow down to "Restart Management Server" or "Destroy Security and Domain Information" if you need to destroy the Global security setup on the array, etc.

## REMOVING USER ACCOUNTS FROM THE DOMAIN:

```
# /nas/sbin/navisecli -h 192.168.1.200 security –rmuser –user nasadmin –scope global
```

**Note:** Works only if there are other Domain Administrative accounts available

## UPDATING or RECREATING THE CLARiiON SECURITY CACHE FILES ON CELERRA:

```
# nas_storage -modify APM00030600872 -security
```

Enter the Global CLARiiON account information

Username: nasadmin

Password: \*\*\*\*\* Retype your response to validate

Password: \*\*\*\*\*

Setting security information for APM00030600872

done

```
# ls -la /etc.clar*
```

```
-rw-rw-r-- 1 nasadmin nasadmin 276 Aug 23 08:34 .clar_security  
-rw-rw-r-- 1 nasadmin nasadmin 0 Aug 23 08:34 .clar_security.lck
```

```
# ls -la /nas/site.clar*
```

```
-rw-rw-r-- 1 nasadmin nasadmin 276 Aug 23 08:34 /nas/site.clar_security  
-rw-rw-r-- 1 nasadmin nasadmin 0 Aug 23 08:34 /nas/site.clar_security.lck
```

```
/nas/symapi/config
```

```
-rw-rw-r-- 1 nasadmin nasadmin 216 Aug 23 08:34 emcpwddb.dat
```

## **CHANGING ARRAY SECURITY ACCOUNT PASSWORD:**

**# nas\_storage -modify id=1 -security -username nasadmin -password nasadmin -newpassword nasadmin**

Changing password on APM00063303725

done

**Note:** Above syntax is available in later NAS 5.5 code & updates array password for account named, and updates the local .clar\_security cache files on the Celerra

## **VIEWING CLARIION SECURITY ACCOUNT INFO ON BACKEND:**

**# nas\_storage -info id=1 |grep -C1 username**

authenticated = True

username = nasadmin

**# /nas/sbin/nas\_navi -h 10.241.168.52 remoteconfig -getconfig -users**

Users: nasadmin@10.241.168.63

**Note:** This command reads the list of Privileged Users, as seen in Navisphere>SPA>Properties>Agent tab, which is located locally on the Clariion SP in the c:>Emc\Navisphere\common\agent.config file. Privileged Users are used for Host authentication, such as to support Unix systems.

**\$ /nas/sbin/navisecli -h 10.241.168.52 security -list**

Username: nasadmin

Role: administrator

Scope: global

**# /nas/sbin/navicli -h 192.168.1.200 security -list**

/nas/sbin/classic\_navicli [-p] [-vlq] [-m] [-np] [-t timeout] [-h hostname] [-d device] [-help] [-f filename] CMD <optional-args>

**Note:** Above output abridged. The navicli security -list command will only work if there are properly updated local .clar\_security cache files on the Control Station, since this command requires the CS to use navisecli in the background. Removing both sets of .clar\_security files from /etc and /nas/site directories will reproduce the above output.

**# /nas/sbin/navicli -h 192.168.1.200 security -list**

Username: nasadmin

Role: administrator

Scope: global

**# /usr/symcli/bin/symcfg authentication list**

| Hostname      | Username | Namespace | Port |
|---------------|----------|-----------|------|
| 10.241.168.55 | nasadmin |           |      |
| 10.241.168.56 | nasadmin |           |      |

## **EXAMPLES OF CLARiiON SECURITY PREVENTING NAS COMMANDS FROM WORKING:**

**# nas\_diskmark -m -a**

Discovering storage (may take several minutes)

Error 5008: root\_disk root\_ldisk root\_ldisk root\_ldisk : volume is unreachable by server\_2

**# server\_devconfig server\_3 -c -s -a**

Discovering storage (may take several minutes)

server\_3 :

Error 5008: server\_3 : root\_disk root\_ldisk d10 d9 root\_ldisk root\_ldisk : volume is unreachable by server\_3

**# /nas/sbin/setup\_slot -i 3**

Initializing server in slot 3 as server\_3

Error 5008: server\_3 : root\_disk root\_ldisk d10 d9 root\_ldisk root\_ldisk : volume is unreachable by server\_3

or

Error 5002: server\_2: APM00055003968

    SMAPI\_C\_CLAR\_NOT\_PRIVILEGED

Operation denied by Clariion array – you are not privileged to perform the requested operation

**Note:** Operations such as setup\_slot, devconfig -create, or nas\_diskmark will fail if CLARiiON security is set and the Celerra does not have a correctly populated /etc/.clar\_security file in order to authenticate to the backend.

## **MUST CORRECT ARRAY USER WHEN TRYING TO MODIFY CELERRA CACHE:**

**# nas\_storage -modify APM00030600872 -security -username root -password nasadmin**

Setting security information for APM00030600872

Error 3502: APM00030600872: Storage API code=4651: SYMAPIC\_C\_CLAR\_NOT\_PRIVILEGED

Operation denied by Clariion array - you are not privileged to perform the requested operation

**Note:** You would use the above command syntax to add the Clariion Global Security account username & password to the local Celerra cache for all Clariion backends that are attached and using Global security

**Conditions where Secure CLI may apply:**

1. For new NAS Gateway installations, where CLARiiON security has been set on the Celerra's boot array, the User will be prompted to enter the CLARiiON username and password in order to authenticate to the CLARiiON backend. Beginning with NAS 5.5.27.5 and higher, if the array(s) has been configured with global domain security, use the # nas\_storage -modify <array\_name> -security command against each array listed in the nas\_storage -l output, then follow the prompts to enter the correct username and password for the Clariion security account. This will then produce an /etc/.clar\_security cache on the local Celerra system, which will then allow the Celerra, or Users on the Celerra, to use navisecccli commands without username & password prompting, and will allow devconfig, setup\_slot, and other commands that need to reconfigure the backend, to run.

**Note:** Anytime that the Clariion security account is changed, or password is changed, the corresponding nas\_storage -modify -security should also be run on the Celerra to update the .clar\_security cache file.

2. If a CLARiiON system is using Classic CLI prior to a NAS Upgrade, the upgrade will proceed using Classic Navi commands. Just be aware that if CLARiiON security is subsequently set on the array, the nas\_storage -modify command will need to be run in order to build the authentication cache on the Celerra Control Station [/nas/site/.clar\_security file]

3. If a CLARiiON system is using Secure CLI, and then a NAS Upgrade is performed, the User will be prompted to enter the CLARiiON security account and password, or the Upgrade will fail with a PUHC [Pre Upgrade Health Check] error.

4. If a new array is added, and CLARiiON security is set, then the User will need to enter the CLARiiON Username and password when prompted prior to complete discovery of the new CLARiiON.

5. For Integrated Celerra arrays prior to NAS 5.5.30.x [Integrated systems do not normally have any Clariion security attributes], use of the navisecccli command set can be accomplished by adding a password to the Clariion cache for the Control Station Host:

**# /nas/sbin/navisecccli -h 192.168.1.200 -AddUserSecurity -Scope 0**

Enter password:

**Note:** Run the nas\_storage -modify <array\_name> -security if CLARiiON security is set and the setup\_slot, devconfig -create, nas\_diskmark, or other backend configuration commands are failing, and security credentials are required on the local Celerra Control Station. Celerra Users will also have the ability to create or update the .clar\_security credential cache via the Celerra Manager interface: Celerras>Storage>Systems:<array>

6. Beginning with NAS 5.5.30.x, Integrated NAS system installs will automatically create a Global Scope 0 security account on the array, with SPA as the Domain Master, using nasadmin as the user account and “nasadmin” as the password, until otherwise reconfigured. Please note that any changes to the Clariion security account on the backend array must then be updated in the .clar\_security cache file on the Celerra Control Station using the nas\_storage -modify -security command.

**Two Steps Required on NAS Integrated Systems when adding or changing Clariion security account or password:**

**/nas/sbin/navisecccli -h 192.168.1.200 -AddUserSecurity -Scope 0** (Scope 0 is Global, Scope 1 is local to array)

**/nas\_storage -modify APM00071600514 -security**

**Locations of Celerra Credential Cache for CLARiiON Security:**

**/nas/site**

|             |   |          |          |                  |                |                                                                                           |
|-------------|---|----------|----------|------------------|----------------|-------------------------------------------------------------------------------------------|
| -rw-rw-r--  | 1 | nasadmin | nasadmin | 552 Feb 2 15:33  | .clar_security | (Primary copy used to authenticate Celerra to CLARiiON)                                   |
| <b>/etc</b> |   |          |          |                  |                |                                                                                           |
| -rw-rw-r--  | 1 | nasadmin | nasadmin | 552 Feb 19 10:22 | .clar_security | (Backup copy, used when /nas partitions are not available, or if manual restore required) |

**Example Showing that Clariion security set, account name known, and Celerra has been authenticated:**

**# nas\_storage -info id=1 |grep -C1 username**

authenticated = True

username = nasadmin

**Note:** True indicates that the Celerra has authenticated to the backend, and username shows a valid user account was found in the .clar\_security file on the Control Station. False and/or blank username field would indicate that nas\_storage -modify would need to be run in order to populate the local Celerra /etc/.clar\_security cache file if CLARiiON security is set.

**Example of CLARiiON user account not found in Celerra .clar\_security cache:**

**# nas\_storage -info id=1 |grep -C1 username**

authenticated = True

username = →Missing username and/or security cache files--run nas\_storage -modify <array> -security to update or create security cache files

**Note:** Security cache files exist with 5.5.27 Gateways and 5.5.30 Integrateds, and later

**NS40/NS40 NEBS (Network Equipment Building Systems) GNAPA:**

→Special platform to conform to NEBS standards. Blades increase to 2U in height with DC Power supplies; CS remains 1U in size

→Supported arrays: CX700, CX3-20, CX3-40, CX3-80

→NS40 has dual 2.8GHz processors, 4GB memory, 16TB capacity (1-16 DAE's of 15 drives each, 6-240 disks total)

→Lightning Surge Protector—Firefly—for Ethernet ports

→Product being shipped only with single Data Mover (Standby to be available later)

## **DATA LUN MIGRATION (RG0) & CONTROL LUN EXTENSION (5.5.32.x min.):**

Purpose is to ensure that LUN0 (root\_disk) and LUN 1 (root\_ldisk) are 11GB in size prior to upgrading to Cognac and allows for migration of User LUNs from RAID Group 0, if required, and extension of LUNs 0 & 1 (root\_disk & root\_ldisk, respectively) to 11GB using /nasmcd/sbin/nas\_extlun –check l -init. All systems must be at 11GB for LUN 0 & 1 prior to upgrading to NAS 5.6. One of the technical challenges addressed by the script is to preserve and move the diskmark information on the 128<sup>th</sup> sector from the end of the LUN, as well as the first 8kb information for each volume. Though the script guides the User through some checking and prompts, the actual LUN extension is a manual task, and varies depending on Symmetrix (metavolume aggregation) or CLARiiON.

## **LUN 0/1 EXTENSION/MIGRATION REQUIREMENTS:**

- Gateway or Integrated NAS systems, Clarion or Symmetrix backends
- Minimum NAS 5.5.32.x before LUN extension process can be run.
- Offline procedure [Data Movers are halted by –init script & rebooted]
- LUN 0/1 can only be migrated to same RG Type, Protection, FC Drive types [Raid 5 4+1], same RG as other Control LUNs, and same block count in target LUN
- Flare 19 minimum on Clariion array
- Must have diskspace on the backend to migrate User luns if not enough space in Control LUN Raid Group to expand LUNs 0/1

### **Clarion Behavior:**

- Uses Clariion LUN migration tool to migrate smaller size lun to 11GB size. Migrates from original LUN name to target LUN and then renames target to original name
- Use nas\_extlun tool on MirrorView Source side only
- Clariion uses block counts for volume size info
- nas\_extlun –check checks for MB & Block counts for disk space requirements
- Clariion platforms will have Backout Procedure based on Clones

### **Symmetrix Behavior:**

- Smaller lun is appended with new metavolume
- Use nas\_extlun on RDF Source side only
- Symmetrix uses cylinder sizes for volume size info
- nas\_extlun –check checks for MB & cylinders for disk space requirements
- Special consideration required for SRDF environments

## **SYSTEMS INELIGIBLE FOR COGNAC UPGRADE:**

1. Any system that cannot perform the LUN extension to 11GB for Luns 0 & 1
2. NS600 arrays
3. CNS or CFS if they have any 507 or 510 Data Movers [Only 514 DMs acceptable]
4. Data Mover rootfs extended using root\_disk\_reserve region of LUN 0 & starting offset for partial dump 0 matches up

**Synopsis of LUN Extension script:** Version 1.2 with NAS 5.5.32.x

Run nas\_extlun –check first to verify whether extension is necessary. The -init checks Backend & Data Mover state, builds minimum boot.cfg, reboots all Data Movers, Stops NAS Services, Prompts User to extend LUNs, DART moves diskmark & first 8kb info (used by MPFSi Solaris clients), NAS starts, Diskmarks updated, new boot.cfgs created and all Servers rebooted.

### **# /nasmcd/sbin/nas\_extlun -check**

NAS Control LUN Extension Utility - Version 1.1

Fri Apr 20 09:54:10 2007

Enter Service Password [or type "ABORT" to exit]:

Validating Primary Control Station ... OK

Mounting system mountpoints ... OK

Fri Apr 20 09:58:33 2007

Verifying Platform type ... OK

Verifying NAS version ... OK

Selecting an active slot ... OK (Slot\_2)

Collect NAS system Information:

Reading NAS system Database info ... OK

Reading storage system info ... OK

Verifying no MirrorView setup on this system ... OK

\*\*\*\*\* Summary of Control LUN info \*\*\*\*\*

Storage System: Clariion 600

Storage ID: APM00023801040

root\_disk: ALU 0, Current size = 4096 MB (8388608 blocks);

Proposed New size = 11264 MB (23068672 blocks).

root\_ldisk: ALU 1, Current size = 4096 MB (8388608 blocks);

Proposed New size = 11264 MB (23068672 blocks).

\*\*\*\*\*

<<DIRECTION>>

Must run "nas\_extlun -init" to initiate the above LUN extension operation.

Unmounting system mountpoints ... OK

Operation Succeeded.

### **CCA POLICY FOR CONTROL LUN EXTENSION:**

→See NAS CCA Guide for requirements

→Upon CCA approval for CLARiiON LUN Extension, a link to the procedure is provided

→A link to the procedure is provided with the RPQ approval for the Symmetrix LUN Extension

### **SYMMETRIX BOOT ARRAY CONTROL LUN EXTENSION (Non-RDF):**

→Overall, procedure is similar to CLARiiON LUN Extension and uses nas\_extlun script

→Requires NAS 5.5.32 or higher

→Supports SYMM-5, DMX, DMX-2, DMX-3, DMX-4

→LUN Extension requires Server downtime

→Uses Symmetrix MetaVolume method for extension

→Cylinder count differs depending on SYMM model

DMX-3 or newer systems would use 12394 cylinders to achieve 11619 MB LUN size

DMX-2 or older systems would use 12788 cylinders to achieve 11619 MB LUN size

1. Run LUN Extension HealthCheck:

**# /nasmcd/sbin/nas\_extlun -check**

2. Run extension script:

**# nas\_extlun -init**

→prompts for password, does healthcheck, stops IO to Data Movers, stops NAS services, reboots Data Movers to minimum boot.cfg mode, waits for User to extend LUNs by appending previously created meta volumes

→When nas\_ext script detects that LUNs are now correct size, will move diskmarks on LUNs 0 & 1, relocates first 8k block for LUN1 & create root\_ldisk\_reserve\_2\_slice (used to extend Data Mover rootfs when upgrading to NAS 5.6), updates database, restores boot.cfg files, and restart Data Movers, then NAS services

→Extension progress recorded in /var/log/nas\_extlun.log

### **CLARiiON CONTROL LUN EXTENSION/MIGRATION (and Data LUN Migration):**

1. Determine whether an extension is required:

```
# nas_disk -l
id inuse sizeMB storageID-devID type name servers
1 y 4095 APM00023801040-0000 CLSTD root_disk 1,2
2 y 4095 APM00023801040-0001 CLSTD root_ldisk 1,2
```

2. Run LUN Extension HealthCheck:

The nas\_extlun -check script should be run to provide the proposed new LUN size & block count for LUN 0 & 1.

**# /nasmcd/sbin/nas\_extlun -check**

NAS Control LUN Extension Utility - Version 1.1

-----abridged-----

\*\*\*\*\* Summary of Control LUN info \*\*\*\*\*

```
Storage System: Clariion 600
Storage ID: APM00023801040
root_disk: ALU 0, Current size = 4096 MB (8388608 blocks);
           Proposed New size = 11264 MB (23068672 blocks).
root_ldisk: ALU 1, Current size = 4096 MB (8388608 blocks);
           Proposed New size = 11264 MB (23068672 blocks).
```

3. Determine size/block count of LUNs requiring migration and availability of R5 (4+1) storage:

In order to migrate User LUNs from RaidGroup 0, available storage will be required, most likely requiring R5 (4+1) raid groups. If storage cannot be found or added to the system, this procedure cannot go forward.

**# /nas/sbin/navicli -h 192.168.1.200 getrg 0 -lunlist**

RaidGroup ID: 0

List of luns: 0 1 2 3 4 5 16 17

**# /nas/sbin/navicli -h 192.168.1.200 getlun 16 -capacity**

LUN Capacity(Megabytes): 115491

LUN Capacity(Blocks): 236527008

# /nas/sbin/navicli -h 192.168.1.200 getlun 17 -capacity

LUN Capacity(Megabytes): 115491

LUN Capacity(Blocks): 236527008

4. Determine target LUNs & make sure Raid Types and Storage Pool disktypes are identical:

# nas\_disk -i

9 n 272815 APM00023801040-0012 CLSTD d9 1,2 -->LUN 18

19 n 272815 APM00023801040-0013 CLSTD d19 1,2 -->LUN 19

# nas\_disk -info d9

id = 9

name = d9

acl = 0

in\_use = True

size (MB) = 272815

type = CLSTD

protection= RAID5(4+1)

5. Delete d9 and d19 from Celerra database, unbind from CLARiiON backend, then rebind luns as the same size as each of the Data LUNs that are being migrated:

# nas\_disk -d d19

id = 19

name = d19

acl = 0

in\_use = False

size (MB) = 272815

type = CLSTD

protection= RAID5(4+1)

# /nas/sbin/navicli -h 192.168.1.200 unbind 18

Unbinding a LUN will cause all data stored on that LUN to be lost.

Unbind LUN 18 (y/n)? y

# /nas/sbin/navicli -h 192.168.1.200 unbind 19

Unbinding a LUN will cause all data stored on that LUN to be lost.

Unbind LUN 19 (y/n)? y

#### Removing RG if desired:

# /nas/sbin/navicli -h 192.168.1.200 removerg 8

# /nas/sbin/navicli -h 192.168.1.200 bind r5 18 -rg 12 -sq bc -cap

236527008

# /nas/sbin/navicli -h 192.168.1.200 bind r5 19 -rg 12 -sq bc -cap

236527008

# /nas/sbin/navicli -h 192.168.1.200 getlun 18 -state -bind

State: Binding

Prct Bound: 22

# /nas/sbin/navicli -h 192.168.1.200 getlun 18 -capacity

LUN Capacity(Megabytes): 115491

LUN Capacity(Blocks): 236527008

6. Perform Data LUN Migration:

# /nas/sbin/naviseccli -AddUserSecurity -scope 1

Enter password:

# /nas/sbin/naviseccli -h 192.168.1.200 migrate -start -source 16 -dest 18 -rate asap

# /nas/sbin/naviseccli -h 192.168.1.200 migrate -start -source 17 -dest 19 -rate asap

# /nas/sbin/naviseccli -h 192.168.1.200 migrate -list

Source LU Name: LUN 17

Source LU ID: 17

Dest LU Name: LUN 19

Dest LU ID: 19

Migration Rate: ASAP

Current State: MIGRATING

Percent Complete: 3

Time Remaining: 49 minute(s)

**Note:** Please note that the original LUN numbers 16 & 17 in the above example, which were located in RG0, are migrated to LUNs 18 & 19 in RG12. The target LUNs, however, are replaced by the original LUN numbers 16 & 17, and original Clariion device numbers, and therefore retain their original d volume number.

7. Update Diskmarks on Celerra volumes:

# **nas\_diskmark -m -a**

8. Create new 11GB LUNs in RG0 for Control LUN 0 & 1 expansion:

# **/nas/sbin/navisecccli -h 192.168.1.200 bind r5 50 -rg 0 -sqbc -cap 23068672**

# **/nas/sbin/navisecccli -h 192.168.1.200 bind r5 51 -rg 0 -sqbc -cap 23068672**

9. Run the nas\_extlun –init script to perform the LUN migration/extension:

# **/nasmcd/sbin/nas\_extlun –init**

**Note:** Script will require a password to continue with the –init migration process. Also, the script will notify the user when an SSH session to the CS is required in order to manually extend LUNs 0 & 1

“Waiting for user to finish the backend LUN extension tasks on the above LUN(s)...

10. Connect to CS using SSH, mount /nas, and run migration on LUN0/1 to LUN 50/51:

# **mount /nas**

# **/nas/sbin/navisecccli -h 192.168.1.200 migrate -start -source 1 -dest 51 -rate asap**

WARNING: The destination LUN is larger than the source. Performing this operation will increase the size of your LUN. Please verify that any hosts attached to this volume can handle dynamic changes to volume size.

Do you really want to perform the action (y/n)? y

**Note:** A LUN migration from smaller to larger LUN is also an automatic LUN Extension. At the end of the migration, the nas\_extlun script will resume and complete the process, and restart NAS Services.

“Fri Apr 27 11:39:33 2007

Initiating NAS service start ... OK

Waiting for NAS service to start ..... OK

Operation Succeeded.”

## **TROUBLESHOOTING:**

**/var/log/nas\_extlun.log**

### **SERVER CHECKUP TOOL GNAPA:**

**Purpose:** This is a new troubleshooting tool that has the ability to check a single Component (CIFS) and (26) CIFS dependencies, and outputs CCMD Warnings or Errors if a potential problem is discovered. See –info –all output below for listing of all dependencies. This tool is meant to check configuration, status of resources, verify some best practice configurations, report anomalies, and provide recommended actions to resolve detected issues. In future releases, this tool may be tied into the Celerra PAHC tool.

→The only ‘Component’ being checked with this release is CIFS—others to be added later are, NFS, HTTPS, DHSM, REPLICATION, iSCSI, etc.

→Tool is not designed to automatically fix any issues found

→Check man page for further syntax information

**\$ server\_checkup server\_2 -list | -info <component> -all | -test <component> -subtest <dependency> -quiet [run checks & output status only] -full [runs more extensive tests]**

**# server\_checkup server\_3 -list**

server\_3 :

CIFS →CIFS is the only ‘component’ currently incorporated into this utility

**# server\_checkup server\_3 -quiet |grep “\*”**

#### **Run Full Test if any Asterisks are seen in the Output:**

ACL : Checking the number of ACL per file system.....\*Pass

DC : Checking the connectivity and configuration of the DCs.....\*Fail

SmbList : Checking the range availability of SMB ID.....\*Pass

UM\_Server : Checking the consistency of usermapper server database.....\*Pass

NB: a result with a '\*' means that some tests were not executed. use -full to run them

#### **Testing Specific Dependencies:**

**# server\_checkup server\_3 -test CIFS -subtest dns**

#### **Example of Domain Guests Group Anomaly:**

**# server\_checkup server\_3 -test CIFS -subtest localgrp**

-----Checks-----

Component CIFS :

LocalGrp : Checking the local groups database configuration..... Fail

-----CIFS : LocalGrp Warnings-----

Warning 17451974726: server\_3 : The local group 'Guests' of server 'MVVIEW\_DM3' contains an unmapped member: S-1-5-15-bfc56af5-d6cf8701-5f67b1a3-202. The access to some network resources may be refused (access right checking in Secure or Unix

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
mode). --> According the configured resolver of your system (NIS, etc config files, usermapper, LDAP...),add the missing members. Refer to the commands reference manual for more details.

**Note:** The above anomaly is actually normal in W2k/W2k3 domains. The domain “guest” account is disabled by default, meaning that the default Domain Guest group (RID 202) and its members are not mapped by Celerra since the account is disabled and not normally used. If lsarpc were run against the default user and domain group account, a mapping would be generated, the Celerra LocalGroups file would map the “Domain Guests” group and this warning would go away.

# .server\_config server\_3 -v "lsarpc if=s2 user='domain guests'"

# .server\_config server\_3 -v "lg list" |grep guests

1172156923: LGDB: 4: group 2K3\domain guests 8005 S-1-5-15-bfc56af5-d6cf8701-5f67b1a3-202 \*

**Running Checks in Quiet Mode:**

# server\_checkup server\_3 -quiet -full

-----Checks-----

Component CIFS :

ACL : Checking the number of ACL per file system..... Pass  
Connection: Checking the load of TCP connections of CIFS..... Pass  
Credential: Checking the validity of credentials..... Pass  
DC : Checking the connectivity and configuration of the DCs..... Fail  
DFS : Checking the DFS configuration files and DFS registry..... Pass  
DNS : Checking the DNS configuration and connectivity to DNS servers. Fail  
EventLog : Checking the configuration of Windows Event Logs..... Pass  
FS\_Type : Checking if all file systems are all DIR3 format..... Pass  
GPO : Checking the GPO configuration..... Pass  
HomeDir : Checking the configuration of home directory share..... Pass  
I18N : Checking the I18N mode and the Unicode=UTF8 translation tables. Pass  
Kerberos : Checking machine password update for Kerberos..... Pass  
LocalGrp : Checking the local groups database configuration..... Pass  
NIS : Checking the connectivity to the NIS servers, if defined..... Pass  
NTP : Checking the connectivity to the NTP servers, if defined..... Pass  
Security : Checking the CIFS security settings..... Pass  
Server : Checking the CIFS files servers configuration..... Pass  
Share : Checking the network shares database..... Pass  
SmbList : Checking the range availability of SMB ID..... Pass  
Threads : Checking for CIFS blocked threads..... Pass  
UM\_Client : Checking for the connectivity to usermapper servers, if any.... Pass  
UM\_Server : Checking the consistency of usermapper server database..... Pass  
UnsupOS : Checking for unsupported client network OS..... Pass  
UnsupProto: Checking for unsupported client network protocols..... Pass  
VC : Checking the configuration to Virus Checker servers..... Pass  
WINS : Checking for the connectivity to WINS servers, if defined..... Pass

# server\_checkup server\_3 -info -all

**DEPENDENCY : ACL**

DESCRIPTION : Number of ACL per file system.

TESTS :

In full mode, check if the number of ACL per file system doesn't exceed 90% of the maximum limit.

**DEPENDENCY : Connection**

DESCRIPTION : TCP connection number

TESTS :

Check if the number of CIFS TCP connections doesn't exceed 80% of the maximum number.

**DEPENDENCY : Credential**

DESCRIPTION : Users and groups not mapped

TESTS :

Check if all credentials in memory are mapped to a valid SID.

**DEPENDENCY : DC**

DESCRIPTION : Connectivity to the domain controllers

TESTS :

Check the connectivity to the favorite DC (DCPing), In full mode, check the connectivity to all DC of the domain, Check if DNS site information are defined for each computer name, Check if the site of each computer name has an available DC, Check if trusted domain of each computer name can be reached, Check the ds.useDCLdapPing parameter is enabled, Check the ds.useADSite parameter is enabled.

**DEPENDENCY : DFS**

DESCRIPTION : DFS service configuration on computer names

TESTS :

Check the DFS service is enabled in registry if DFS metadata exists, Check the DFS metadata of each share with DFS flag are correct, Check if share names in DFS metadata are valid and have the DFS flag, Check if each DFS link is valid and loaded, Check in the registry if the WideLink key is enabled and corresponds to a valid share name.

**DEPENDENCY : DNS**

DESCRIPTION : DNS domain configuration

TESTS :

Check if each DNS domain has at least 2 defined servers,

Check the connectivity to each DNS server of each DNS domain, Check if each DNS server of each DNS domain supports really the DNS service, Check the ds.useDSFile parameter (automatic discovery of DC), Check the ds.useDSFile parameter is enabled if the directoryservice file exists.

**DEPENDENCY : EventLog**

DESCRIPTION : Event Logs parameters on servers

TESTS :

Check if the pathnames of each event logs files are valid (application, system and security), Check if the maximum file size of each event logs file doesn't exceed 1GB, Check if the retention time of each event logs file doesn't exceed 1 month.

**DEPENDENCY : FS\_Type**

DESCRIPTION : DIR3 format of filesystems

TESTS :

Check if each file system is configured in the DIR3 format

**DEPENDENCY : GPO**

DESCRIPTION : GPO configuration on Win2K servers

TESTS :

Check if the size of the GPO cache file doesn't exceed 10% of the total size of the root file system, Check the last modification date of the GPO cache file is up-to-date, Check the cifs.gpo and cifs.gpoCache parameters have not been changed,

**DEPENDENCY : HomeDir**

DESCRIPTION : Home directory shares configuration

TESTS :

Check if the home directory shares configuration file exists, the feature is enabled, Check if the home directory shares configuration file is optimized(40 lines maximum), Check the syntax of the home directory shares configuration file.

**DEPENDENCY : I18N**

DESCRIPTION : Internationalization and translation tables

TESTS :

Check if computer name exists, the I18N mode is enabled, Check the .etc\_common file system is correctly mounted, Check the syntax of the definition file of the Unicode characters, Check the uppercase/lowercase conversion table of Unicode character is valid.

**DEPENDENCY : Kerberos**

DESCRIPTION : Kerberos configuration

TESTS :

Check the machine password update is enabled and up-to-date.

**DEPENDENCY : LocalGrp**

DESCRIPTION : Local groups and local users

TESTS :

Check the local group database doesn't contain more than 80% of the maximum number of servers, Check if the servers in the local group database are all valid servers, Check the state of the local group database (initialized and writable), Check if the members of built-in local groups are all resolved in the domain, Check the number of built-in local groups and built-in local users, Check if the number of defined local users doesn't exceed 90% of the maximum number.

**DEPENDENCY : NIS**

DESCRIPTION : Network Information System (NIS) configuration

TESTS :

If NIS is configured, check at least 2 NIS servers are defined (redundancy check), Check if each NIS server can be contacted on the network, Check if each NIS server really supports the NIS service.

**DEPENDENCY : NTP**

DESCRIPTION : Network Time Protocol (NTP) configuration

TESTS :

If NTP is configured, check at least 2 NTP servers are defined (redundancy check), Check if each NTP server can be contacted on the network, If computer names exist, check if NTP is configured and is running.

**DEPENDENCY : Security**

DESCRIPTION : Security settings

TESTS :

If the I18N mode is enabled, check the share/unix security setting is not in use, Discourage to use the share/unix security setting, Check the cifs.checkAcl parameter is enabled if the security setting is set to NT.

**DEPENDENCY : Server**

DESCRIPTION : Files servers

TESTS :

Check if each CIFS server is configured with a valid IP interface, Check if each computer name has joined its domain, Check if each computer name is correctly registered in their DNS servers, Check if the DNS servers have the valid IP addresses of each computer name, Check if a DNS domain exists if at least one computer name exists,

**DEPENDENCY : Share**

DESCRIPTION : Network shares

TESTS :

Check the available size and i-nodes on the root file system are at least 10% of the total size, Check the size of the share database doesn't exceed 30% of the total size of the root file system, Check if the pathname of each share is valid and is available, Check if each server in the share database really exists, Check if the I18N mode is enabled, all the share names are UTF-8 compatible, Check the list of ACL of each share contains some ACE, Check the length of each share name doesn't exceed 80 Unicode characters.

**DEPENDENCY : SmbList**

DESCRIPTION : 64k UID, TID and FID limits

TESTS :

In full mode, check the 3 SMB ID lists (UID, FID and TID) don't exceed 90% of the maximum ID number.

**DEPENDENCY : Threads**

DESCRIPTION : Blocked threads and overload

TESTS :

Check CIFS threads blocked more than 5 and 30 seconds, Check the maximum number of CIFS threads in use in the later 5 minutes doesn't exceed 90% of the total number, Check the number of threads reserved for Virus Checker doesn't exceed 20% of the total number of CIFS threads.

**DEPENDENCY : UM\_Client**

DESCRIPTION : Connectivity to the usermapper server

TESTS :

If usermapper servers are defined, check each server can be contacted, Check if usermapper servers are defined, NIS is not simultaneously activated.

**DEPENDENCY : UM\_Server**

DESCRIPTION : Usermapper server

TESTS :

If a usermapper database is defined locally, check its size doesn't exceed 30% of the total size, Check if configuration file is in use, the filling rate of the ranges doesn't exceed 90%, Check if configuration file is in use, 2 ranges do not overlap, Check if secmap is enabled, In full mode, check the SID/UID and SID/GID mappings and reverses are correct and coherent.

**DEPENDENCY : UnsupOS**

DESCRIPTION : Client OS not supported

TESTS :

Check for unsupported client network OS.

**DEPENDENCY : UnsupProto**

DESCRIPTION : Unsupported protocol commands detected

TESTS :

Check for unsupported client network protocol commands.

**DEPENDENCY : VC**

DESCRIPTION : Virus checker configuration

TESTS :

If VC is enabled, check the syntax of the VC configuration file, Check if the VC 'enable' file and the VC configuration are compatible, Check the number of VC servers. Make sure at least 2 servers are defined, for redundancy, Check if there are offline VC servers, Check if the VC high watermark has not been reached, Check the connection of VC servers to the Data Mover.

**DEPENDENCY : WINS**

DESCRIPTION : WINS servers.

TESTS :

If NetBIOS names are defined, check if at least one WINS server is defined, Check the number of WINS servers. check if two servers are defined for redundancy, Check if each WINS server can be contacted on the network, Check these servers are really WINS servers, Check if the NetBIOS are correctly registered on the servers.

**INFORMATION WRITTEN TO SERVER LOG:**

2007-02-23 11:27:23: ACLUPD: 4:11: check acl on 14 ended successfully

2007-02-23 11:27:23: SMB: 4: >tmpDC=GEORGE(192.1.4.217) R=6 T=1000 ms S=0,1/-1

2007-02-23 11:27:23: CIFSSUPPORT: 4: compname mview\_dm3 DC=GEORGE Trusted domain='NTDOMAIN' status='There are currently no logon servers available to service the logon request'  
2007-02-23 11:27:23: LIB: 3: 1: DNS: unable to connect to name server 10.241.168.113: I/O error  
2007-02-23 11:27:27: LIB: 3: 1: DNS: unable to connect to name server 10.241.168.113: Connection timed out  
2007-02-23 11:27:27: SMB: 4: Unknown OS logons: 0 Policy=GrantAll  
**Note:** Looks as if server\_checkup also calls the server\_cifssupport tool for certain functions

## **DUMP COMPRESSION & ENHANCEMENTS ON BACKEND GNAPA:**

Dumps will be automatically compressed when written to the backend slots—more useful information is also supposed to be provided. Following example is a dumpfile automatically extracted and zipped by the automatic collect utility—if automatic collect is not configured or running, the dumpfile remains on LUN0 on the backend, which can be checked using the dump\_slot command:

1. Checking for dumpfile and compression on Backend dumpslot:

**# /nas/sbin/dump\_slot -v -F**

Checking LBA 0xf3800 for Full Dump...

Dump info FOUND:

Dart Slot 2

Full Dump incomplete!

Do you wish to save this dump anyway? [y or n] y

COMPRESSED = YES

2. Checking dumpfile if already extracted by automatic collection script, which gzips the dump—unzip file:

**# gzip -d dump\_ML2803000397.070123\_0723.dump.gz**

3. Verifying whether new compression technique has been applied to the dumpfile:

**# /nas/sbin/nas\_crash /nas/var/dump/dump\_ML2803000397.070123\_0723.dump**

DART time of dump: Tue Jan 23 07:39:40 2007 UTC

Product: EMC Celerra File Server

Uptime: 006 days, 18:33:55

Version: T5.5.80.0

IP Address: 192.168.1.2

Host Name: server\_2

COMPRESSED = YES

4. Use the dump\_slot utility to uncompress the dump

**# /nas/sbin/dump\_slot -i dump\_ML2803000397.070123\_0723.dump -o**

**<provide\_name\_for\_uncompressed\_dump>**

Dump size is 704203264 bytes (671 MB) . A dump file of this size will be created.

Done

5. Verifying:

**# /nas/sbin/nas\_crash /nas/var/dump/dump\_ML2803000397.070123\_0723.uncompressed**

COMPRESSED = NO

6. Help usage for dump\_slot command:

**# /nas/sbin/dump\_slot -h**

usage: dump\_slot [-f device\_name] [-d dump\_output\_file ] [-Dump\_number n]

[-x sysinfo only] [-w upload one dump only] [-verbose] [-help]

<[-Fulldump] | slot\_id>

| -i compressed\_dump\_file -o uncompressed\_dump\_file\_to\_create

## **OTHER GNAPA ENHANCEMENTS:**

→Upgrades will produce a nas\_prestage.tar.gz tarball in /var/tmp, which will reduce time by not requiring multiple extraction calls during the upgrade and allows for quicker restarts after failures

→Mars Flare support, Dewars CS support for NS40/NS80, iSCSI Snaps and replication for Linux, MPFSi Phase 4, limited support for CLARiiON AX150 arrays, Celerra Manager support for IE 7.0, 750GB SATA drives on CLARiiON

→NFSv4 statistics and other calls are now supported on the Data Mover

→Support for CLARiiON Gateway AX150 arrays, using ax1mpfs license key

## **NAPA 6:**

NAS 5.5.28.x ETR April 2007

→Gateway NAS support for new CX3-10C Tackhammer arrays [Arrays could have either all FC drives, all ATA drives including vault, or a combination of FC (vault) with ATA drives, min. Flare 24 Patch 008]

**Note:** The iSCSI combo cards will be used to support Hosts other than the Celerra in this configuration. All ATA configuration would mean use of ATA drives for both Clariion Vault drives and NAS Control Luns

→Some iSCSI Replication performance enhancements

→Support for Vista O/S clients and Celerra, but in W2k or W2k3 domains only (Not Longhorn)

→Firefox 1.5 Browser support

→Psuedo IP Interface support for Morgan Stanley only

## **NAPA 7:**

NAS 5.5.29.x ETR June 2007

→Introduces support for Enginuity microcode 5772 on Symmetrix for RAID 6 configurations (Raid 6 6+2).

**Note:** Symmetrix supports Control Luns only on Mirrored devices, not RAID of any kind, nor is Celerra AVM able to use RAID with Symmetrix devices.

## **NAPA 8:**

NAS 5.5.30.x ETR Aug 2007

→Tiered Services Model for Customers on NS20 platform initially (Enhanced, & Premium Support)

→NS20 & NS40 Factory or Field install models will ship with pre-installed NAS & FLARE software within a Celerra StorageGroup and Access Logix enabled, & pre-configured LUNs that marry the Backend to the Celerra

→An NS20 could be configured with all ATA drives (meaning vault drives & Control LUNs)

→Second SPS on array is optional

→Customer-Installable NS20/20FC WagonWheel install approach, Peel-off CS info label, Read Me First, Placemat, Powerlink Landing Page, CSA tool with Registration capability

→Landing Page for specific platform troubleshooting, CRU/FRU procedures (Pwr Supplies, SFPs, Disks, etc), support information

→CSA Celerra Startup Assistant for cable check and basic system configuration (replaces InitWizard, but uses iwd process)

→New NS20 Integrated & NS20FC-Enabled models (latter model is not called a Gateway) using CX3-10 & CX3-10F, respectively

→New NS40 Integrated & NS40FC-Enabled models (the FC is not a Gateway, but a variant of the Integrated), using CX3-40F for both models

→CSA Celerra Startup Assistant, RegWiz, & NAT-enabled NST utilities to be provided on Apps & Tools CD

→NS20 & NS40 Integrated (no more old style Integrateds) SPs will operate in SAN mode, with direct fibre optic connections between DM blades and SPs, use of Storagegroups & AccessLogix for Integrated & FC Enabled models

→“All-in-One” NS20F & NS40F Fibre Channel Enabled Models (additional Hosts can connect to array directly or via switch)

→Network access to SPs using Flare 26 version of NST tool (20/20FC/40/40FC) for standard troubleshooting & CRU wizards

→Network access to SPs using Flare 24 version of Navisphere OffArray Mgr tool (20FC/40FC only) for storage management of arrays

→CX3-10, CX3-10F, CX3-40F Arrays, Flare 24 running CelerraService pkg on backend to identify Celerra-owned arrays

→Celerra NAT Service (IPTables), Internal Network changes, & Public IP address for SPA & SPB built on CS aliases to allow network connection from NST & Navisphere

→Dewars Control Station for NS20 & NS40 models

→CAVA version 3.5.8 to be introduced, which supports Automatic Pattern Update with Trend Micro ServerProtect 5.5.8 Build 1176+

**Note:** ServerProtect will be able to tell CAVA when a new pattern update is available, and CAVA will inform DART through the use of the heartbeat mechanism—this will essentially invalidate the “scan on first read” cache to that files will be rescanned when accessed.

→Introduces ability to create a RW iSCSI LUN clone from an iSCSI Snap using server\_iscsi –modify option

### **CREATING iSCSI LUN CLONE:**

**Note:** Essentially, would need to use either Linux or Windows RM or RM/SE application (cmd\_cli) to do the following:

a) Create iSCSI copy job Source to Destination LUN using RM or RM/SE

b) Expire the last replica/snap to destroy iSCSI replication sessions

c) Use # server\_iscsi –modify to change iSCSI replication destination LUN from RO to R/W

**Note:** LUN attribute cannot be changed if LUN replication in use, LUN being accessed by Host, LUN has been promoted

d) Create LUN mask of clone to Windows Host and login via iSCSI

e) Remove VSS R/O and hidden bits on the clone LUN using diskpart utility, or RmMountVolume.bat utility with RM or RM/SE

f) Resulting clone can be Snapped, Replicated, or Extended

→New installs will enforce use of Secure Navi [Naviseccli], while upgrades to 5.5.30.x will support Naviseccli, but not require it

**Note:** New installs will apply default User “nasadmin” and Password “nasadmin” to setup security with SPA as the Domain Master

→Celerra will support Enginuity 5772 and Symmetrix RAID 6

## **NS20/40 INTEGRATED, NS20FC/40FC ENABLED CELERRA:**

GA August 2007 NAS 5.5.30.x

**What's different about the new SAN Mode on the SPs vs. the typical Integrated SAN AUX Mode:**

- CSA utility replaces InitWizard for initial system setup (Windows utility, Apps CD or Powerlink—csa\_install.exe)
- Celerra implements the IPTables Service, the Linux version of NAT (Network Address Translation), to translate public SP IP's to internal SP IP's (actually, CS acts as a Router and rewrites source/destination addresses in IP packets to allow SP hosts access)
- Control Station is configured with (2) aliases on eth3 for Public IP addresses to SPs, and (2) internal network aliases (used for internal SP-to-SP communication when dual CS platforms are involved), & IP Forwarding set to 1 by default
- DMs connect to upper LAN Mgmt port on SPs for the internal Celerra network, as opposed to the lower LAN Service port

**Note:** Original NS40 connected to Service Ethernet ports on SPs and SPs ran in SAN-AUX mode

→No more SAN AUX mode or copper cables between SPs and DMs

→DMs connect to SPs using Fibre Optic LC-to-LC cables (SFP connectors) instead of copper HSSDC2 cables

→DMs will connect to Ports 2 & 3 fibre for CX3-20F & CX3-40F arrays (NS20FC, NS40, NS40FC), & Ports 0 & 1 Fibre for NS20

→SPs run in SAN mode, similar to what NS80 Integrated already does, not SAN-AUX, & means that the two LAN ports on the SP are now separated by VLAN security

**Note:** VLAN security means that Service & Mgmt ports are separated. Can no longer access Internal Celerra network from Service port, but could use the Service port to connect to the SPs default CLARiiON IP network (128.221.1.250/251). With the original NS40 Integrated model running in SAN AUX mode, you could connect to the Internal Celerra network over either the Service or Mgmt LAN port since there was no VLAN security

→Both models support the Celerra Integrated NST utility, but the NS20FC will also support the Navi OffArray Manager

**Note:** “CelerraService” pkg is installed only on dedicated backends (NS20, NS20FC), and is used as a CelerraIdentifier for the NST, via a cimom API call to the SPs after a User first logs in via the NAT Service on the Control Station.

→NS20 (aka Sledgetack) Integrated is the replacement for NS350, in either an EMC rack or customer-provided racks, no Gateway model to be offered

→NS20 introduces first Tiered Service Support for Celerra—Enhanced & Premium Support with 5x9 or 7x24 onsite support, and 4-hour vs. 2-hour remote Support Center support, respectively

→NAS software upgrades are not gratuitous and should be covered under a separate SW maintenance contract

→Upgrades of DAEs & Disks via Clariion tools, Upgrades from Single to Dual Blades (EMC), Upgrades from NS2-AUX to NS2-AUXF supported

→SRDF & MPFS will not be supported on the NS20

#### **FRONTEND FIBRE CHANNEL PORTS ON NS20/NS20FC/NS40/NS40FC SPs:**

→CX3-10F configurations (NS20FC) will have a total of (6) frontend fibre channel ports per SP, (2) for Celerra connectivity, and (4) for other Hosts to connect via FC optic cable directly or via Switched Fabric

→CX3-10 configurations (NS20) will have a total of (2) frontend fibre channel ports per SP, both for Celerra connectivity to backend

→CX3-40F configurations (NS40 & NS40FC) will have a total of (4) frontend fibre channels per SP--(2) ports are unused for NS40, while (2) ports are used for Celerra connectivity for NS40; (2) ports are available for Windows Hosts for the NS40FC shared array model

→CX3-40 configurations (original NS40 Integrated) have only (2) frontend fibre channel ports per SP, both for Celerra connectivity

#### **CELLERRA TIERED SERVICES:**

##### **I. Enhanced Support (CallHome)**

5x9 Next Business Day Onsite support, 7x24 Call Center support with 4-hour response times, 3-year HW warranty

Customer to replace CRUs (fans, pwr supplies, SFPs, drives via DRU tool)

Onsite support for FRUs failures

CallHome to Celerra Support and/or Partners

NAS & Flare Upgrades

System installation included by default but can be deselected on Sales Order for customer installation

Software updates require separate maintenance contract

##### **II. Premium Support (CallHome)**

7x24 Onsite Support with 4-hour response & 7x24 CallCenter Support with 2-hour response, 3-year HW warranty

CRUs & FRUs replaced by EMC, but customers can replace some parts

CallHomes to Celerra Support and/or Partners

Software updates require separate maintenance contract

#### **NS20/20FC, NS40/40FC INSTALLATIONS:**

→Preloaded NAS/Flare & Control Luns on all Integrated & FC Enabled systems

→All models will run AccessLogix and have a Celerra Storage Group for both Integrated & FC Enabled models

**Note:** NAS Upgrades (old NS40 Integrated) will not implement Storage Group or use Access Logix

→New installs implement Navisecccli using SPA as Domain Master with Global security enabled (Scope 0), default user “nasadmin” and password “nasadmin”, while upgrades to Napa 8 will keep existing security (Classic or Secure CLI)

→Control Station will be pre-configured with two new internal Gateway IP address aliases (eth0:1, eth2:1—strictly speaking only used for SP-to-SP communications for dual CS environments), as well as POSTROUTING entries in /etc/sysconfig/iptables for the Internal IP addresses of SPA & SPB—the CS acts as a router and rewrites Host and Destination IP addresses in IP packets

→Physical installation of system guided by WagonWheel Steps and Placemat document, along with Read Me First summary of the installation process

→Control Station will have a peel-off label documenting TLA serial #, Model #, NAS & Flare versions, Unit Tested date, Pre-Configured (yes), MAC address of CS, SP gateway addresses

→System cabling to be color-coded for ease of hookup

→On system powerup, Users will use CSA wizard to configure the system and Register the Celerra, along with Powerlink Landing Page for Information, Troubleshooting, Documentation, Procedures, etc.

→After CSA configuration, the Control Station will have a new Public IP address for both SPA & SPB, respectively (aliases eth3:1, eth3:2), along with IPTables entries for NAT translation services that perform a passthrough service between Public and Private addresses on the SPs--all designed to allow connectivity from the public network using either the NAT-enabled NST or Navi OffArray Mgr tools

→Intent is that “reinstalls” & “fresh installs” will not be required

→/etc/rc3.d/S95cable\_check script prevents casual “reinstall” over factory system; Runs on factory system prior to S95nas startup script; Starts up the IWD nas\_ipinit daemon on CS so that CSA can be used to configure the system; Removes itself as a startup script after successful CSA configuration; Starts NAS, Web, & Tomcat services after successful application of Pre-Configuration

→Use the CSA to configure Control Station IP, new Public IPs for SPA/SPB, perform Cablecheck, and complete customer-configurable setup and testing for NFS or CIFS or iSCSI using single Export/Share/File System, as well as to Register system with EMC/CSI database for Callhome/Service Support

→Clariion SPs will reboot after configuration of the new alias IP's for SP-to-SP communications [192.168.1.104 & 2.104], but this is part of the factory install and should not normally be required unless doing a complete reinstall

#### **Example Command to Change SP Gateway IPs:**

# /nas/sbin/navicli -h 192.168.1.200 networkadmin -set -gateway 192.168.1.104

# /nas/sbin/navicli -h 192.168.2.201 networkadmin -set -gateway 192.168.2.104

→AccessLogix will be enabled & Storage Group for Celerra created [info stored in /etc/be\_sg\_info & be\_rg\_info]

#### **Example of NAS Upgrade on ‘old’ NS40 Integrated (to NAS 5.5.30+)**

→Upgrade on old NS40 Integrated systems does not enable AccessLogix [Data Access control: DISABLED], though AccessLogix service on array is “Active”

→Upgrade does not configure Secure CLI if using Classic CLI

→Upgrade checks for and then installs the CelerraService pkg on the array

→Upgrade does not change the existing Gateway addresses on SPs, but recommends changing at the earliest convenience, though neglects to mention that this change requires SP reboot:

#### **Upgrade Log:**

SPA's gateway is not 192.168.1.104. Please change it at the earliest convenience

SPB's gateway is not 192.168.2.104. Please change it at the earliest convenience

→Upgrade creates the alias files on the Control Station to support internal network SP-to-SP communications

# cat ifcfg-eth0:1

DEVICE=eth0:1

IPADDR=192.168.1.104

# cat ifcfg-eth2:1

DEVICE=eth2:1

IPADDR=192.168.2.104

→Aliases did not appear in CS memory until after running following:

# /sbin/ifconfig (no entries for eth0:1 or eth2:1)

# /nasmcd/sbin/cs\_dhcpd\_monitor (run this command or reboot CS to bring eth0:1 & eth2:1 UP)

# /sbin/ifconfig

eth0:1 Link encap:Ethernet HWaddr 00:00:F0:9F:B3:74

inet addr:192.168.1.104 Bcast:192.168.1.255 Mask:255.255.255.0

UP BROADCAST RUNNING MULTICAST MTU:1500 Metric:1

eth2:1 Link encap:Ethernet HWaddr 00:00:F0:9F:53:07

inet addr:192.168.2.104 Bcast:192.168.2.255 Mask:255.255.255.0

UP BROADCAST RUNNING MULTICAST MTU:1500 Metric:1

Base address:0xd880 Memory:fbea0000-fbec0000

→Upgrade starts the IPTables service, flushes IPTables, and creates new entries to support internal network SP-to-SP communications

→Upgrade permanently enables IP Forwarding

# cat /proc/sys/net/ipv4/ip\_forward

1

## **STORAGE MANAGEMENT OF ARRAY:**

→NS20 & NS40 Integrated models will require use of Celerra Manager to configure backend storage

→NS20FC/40FC will require use of Navisphere OffArray Manager to configure backend storage

**Note:** Celerra Manager & CLI are disabled. Setup\_clariion script recognizes the FC Enabled system and will display message: “Additional configuration of this shared storage array may be required by the SAN administrator.”

## **Control & Data LUNs on NS20 FC or NS40 FC Enabled Systems**

- All storage to be managed by Navisphere OffArray Manager
- Celerra Manager>Tools>Navisphere is greyed out and Storage>Systems>array\_ID>Configure button is not displayed
- Note:** If trying to use /nas/sbin/setup\_clariion to configure storage, script will provide message: “Additional configuration of this shared storage may be required by the SAN administrator.”
- Single Data LUN is created with Install image, at 25% of the available storage space left in RG0 (allows other Hosts to use the remaining 75% from RG0)

# nas\_disk -l

```
id inuse sizeMB storageID-devID type name servers
1 y 11263 APM00071600514-0000 CLSTD root_disk 1,2
2 y 11263 APM00071600514-0001 CLSTD root_ldisk 1,2
3 y 2047 APM00071600514-0002 CLSTD d3 1,2
4 y 2047 APM00071600514-0003 CLSTD d4 1,2
5 y 2047 APM00071600514-0004 CLSTD d5 1,2
6 y 2047 APM00071600514-0005 CLSTD d6 1,2
7 n 233373 APM00071600514-0010 CLSTD d7 1,2
```

## **Control & Data LUNs on NS20 or NS40 Integrated**

- All storage on array to be managed via Celerra Manager or CLI setup\_clariion
- Celerra Mgr>Tools>Navisphere tab is greyed out
- Use Celerra>Storage>Systems>array\_ID>Configure button
- Integrated installs configure all remaining space in RG0 with (2) Data Luns of equal size

# nas\_disk -l

```
id inuse sizeMB storageID-devID type name servers
1 y 11263 APM00071600514-0000 CLSTD root_disk 1,2
2 y 11263 APM00071600514-0001 CLSTD root_ldisk 1,2
3 y 2047 APM00071600514-0002 CLSTD d3 1,2
4 y 2047 APM00071600514-0003 CLSTD d4 1,2
5 y 2047 APM00071600514-0004 CLSTD d5 1,2
6 y 2047 APM00071600514-0005 CLSTD d6 1,2
7 n 466747 APM00071600514-0010 CLSTD d7 1,2
8 n 466747 APM00071600514-0011 CLSTD d8 1,2
```

## **CONFIGURING STORAGE USING CELERRA MANAGER:**

Celerra Mgr>Storage>Systems>array\_ID>Configure, brings up (3) options

- o CX\_All\_4Plus1\_Raid\_5 (default selection)
- o CX\_Standard\_Raid\_5
- o User Defined (this option does not work in Celerra Manager—must use CLI setup\_clariion)

## **CONFIGURING STORAGE USING SETUP CLARIION CLI SCRIPT:**

# /nas/setup/setup\_clariion

Clariion Array: APM00071600514 Model: CX3-40f Memory: 4096

The following 4 template(s) available:

1. CX\_All\_4Plus1\_Raid\_5
2. CX\_Standard\_Raid\_5
3. User Defined → Selected the User Defined template
4. None

Configuration for APM00071600514

1. FC\_RAID5\_4+1\_HS\_4+1\_HS\_HS\_HS\_HS
2. FC\_RAID5\_4+1\_HS\_R1\_R1\_R1\_R1\_HS → Selected this template to complete shelf configuration
3. FC\_RAID5\_4+1\_HS\_8+1
4. FC\_RAID5\_4+1\_HS\_R1\_R1\_4+1
5. None

Do you want to continue and configure as shown [yes or no]?: yes

Enclosure 0\_0.

Created disk group 8, luns 18,19

Created disk group 9, luns 20,21

Created disk group 10, luns 22,23

Created disk group 11, luns 24,25

Created spare 201

Binding complete.

All luns are created successfully!

Enclosure info:

---

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14

---

0\_0: 300 300 300 300 300 300 300 300 300 300 300 300 300 300 300

FC 0 0 0 0 HS 8 8 9 9 10 10 11 11 HS MIX

---

Configuration completed!

Setup of CLARiiON APM00071600514 storage device complete.

Discovering storage (may take several minutes)

### **RECONFIGURING STORAGE ON CELERRA INTEGRATED ARRAYS:**

#### **/nas/site/ APM00071600514.cfg**

If you had a system where you were rearranging LUNs on the backend, you would need to rename the above file in order to be able to run the setup\_clariion script successfully, as the saved configuration file does not allow the Celerra to reconfigure storage. Rename the file and then proceed with /nas/sbin/setup\_clariion –init to reconfigure storage.

### **ARRAY EVENT MANAGEMENT/CALLHOME:**

NS20, NS40, NS20FC, NS40FC models all use the Celerra as the Event Monitoring and Reporting system for the Array (there is no separate CLARAlert agent running on a Windows system since the Celerra “owns” the array)

#### **NAS EVENT MONITORING FOR CLARIION ARRAY**

/nas/sys/nas\_mcd.cfg →Basic configuration file  
daemon “Navi Event Monitor” →Basic monitoring process on Celerra  
/nas/sbin/navilog\_mon →Serves as event collector for backend from Navi Agent and uses postevent to log Information, Warning, Error, & Critical events to the Celerra Eventlog  
/nas/log/agent.log [Navisphere agent log]  
/nas/log/navimon.log [Navisphere monitor log]

#### **Control Station Processes:**

```
# ps -eafl |grep navi
000 S root 5816 2941 0 69 0 - 816 pipe_w Aug14 ? 00:00:00 /usr/bin/perl /nas/sbin/navilog_mon
100 S root 5819 2941 0 69 0 - 3729 do_sel Aug14 ? 00:00:02 /nas/opt/Navisphere/bin/naviagent -d -f
/nas/etc/Navisphere/agent.config -r /nas
```

#### **# nas\_event -list -a callhome |grep Navi**

```
NaviEventMonitor 4 Navi Event with severity CRITICAL was received
NaviEventMonitor 100 A control lun has been cache compromised
```

#### **# nas\_event -l -f NaviEventMonitor**

|     |                                                    |
|-----|----------------------------------------------------|
| id  | description                                        |
| 1   | Navi Event with severity INFORMATION was received  |
| 2   | Navi Event with severity WARNING was received      |
| 3   | Navi Event with severity ERROR was received        |
| 4   | Navi Event with severity CRITICAL was received     |
| 5   | Unknown Navi Event was received                    |
| 10  | Navi Event Monitor has terminated                  |
| 100 | A control lun has been cache compromised           |
| 101 | A control lun has been trespassed                  |
| 200 | Device group has changed status                    |
| 201 | A SFP on the CLARiiON is reporting an error status |

### **EXAMPLE SYS LOG EVENT ERROR & EVENT ID:**

Aug 27 11:42:51 2007 NaviEventMonitor:3:3 Backend Event Number 0x904 Host →Navi Severity 3=ERROR, Event ID 3

Aug 27 12:09:54 2007 NaviEventMonitor:4:2 Backend Event Number 0x850 Host →Navi Severity 4=CRITICAL, Event ID 2

### **Required O/S Versions & packages for Napa 8 Platforms**

→Celerra O/S NAS 5.5.30.x (Napa 8)  
→Clariion O/S Flare Version 24  
→Clariion Array AccessLogix package  
→Clariion Array CelerraService package

### **Necessary Celerra Utilities & Versions for Napa 8 Platforms**

→NST (Navisphere Service Taskbar)

**Note:** NST\_setup\_6.26.50.1.14.exe (47MB), Flare 26 version required, Apps & Tools CD, or Clariion sw site

Windows application allowing Customers to access the Array on Integrated & FC-Enabled platforms for limited hardware troubleshooting, healthchecks, and hardware maintenance functions

→Navisphere Manager (OffarrayUI)

**Note:** Windows\_UIs\_6.24.2.5.0.exe (25MB), Flare 24 version required, download from CLARiiON SW site & being shipped with all NS20FC & NS40FC models or available from <http://www.cs.isus.emc.com/csweb2/dgweb/>

Windows application required for managing all Array Storage on the FC-Enabled platforms for Celerra and/or other Hosts—not for use with pure NS20 or NS40 Integrateds

→CSA (Celerra Startup Assistant)

**Note:** csa\_install.exe (38MB), Apps & Tools CD or Powerlink>Celerra Tools site

Windows application required for initializing and minimally configuring new Celerra installations, with built-in Registration wizard

→Stand-alone Registration Wizard (RegWiz)

**Note:** regwiz\_install.exe (34MB), Apps & Tools CD or Powerlink>Celerra Tools site

Windows application allowing for stand-alone Registration of the Celerra system after CSA configuration, required in order to obtain proper EMC Support

→Celerra Manager

**Note:** Built into NAS code, requires compatible Client Web Browser & min. JRE version (no downloads, etc.)

NAS application used for configuring, monitoring, and managing the Celerra; Required for the management of all Array Storage for NS20/NS40 Integrated Celerras

## **NS20 INTEGRATED:**

Uses NS40 blades connecting to CX3-10 array (called the NS2-AUX)

**Note:** CX3-10 is stripped down version of CX3-20—no Combo card, SP's have only 1GB memory, default configuration is for single SPS

(2) Front-end FC Ports BE0 & BE1 for Celerra backend connectivity, to SP ports 0 & 1 Fibre (No Headhunter IO card)

(1) Front-end FC Port for Tape connectivity on DM Blade

## **NS20FC ENABLED (“All-in-One”):**

Uses NS40 blades connecting to CX3-10F array (called the NS2-AUXF)

Ships with Headhunter FC IO Module to add (4) FC Ports on each SP, Blades connecting to Ports 2 & 3 Fibre

(2) Front-end FC Ports for Celerra backend connectivity to SPs, ports 2 & 3 Fibre

(1) Front-end FC Port for Tape connectivity on DM Blade

(4) Front-end FC ports on SPs for other Hosts direct or FC Switch-attached [Ports 0, 1, 4, & 5 Fibre; Mirrorview Port 1]

## **ARRAY MODELS—minimum Flare 24:**

NS2-AUX (CX3-10) & NS2-AUXF (CX3-10F) Models—Integrated Model only, no Gateway systems

**Note:** The CX3-10 & CX3-10F are Celerra array-variants only and not offered elsewhere. The name “SledgeTack” is based on the fact that the DM hardware is an NS40 (CX3-40 array is called Sledgehammer) & the Backend array is Tackhammer (CX3-10 variants)

## **DETERMINING NS20/20FC BACKEND ARRAY TYPE:**

**/nas/sbin/navicli -h 192.168.1.200 getagent**

Model: CX3-10 or CX3-10f

→Output will show whether system is CX3-10 or CX3-10F, the latter which would mean that FC is Enabled for other Hosts

**/nas/sbin/navicli -h 192.168.1.200 storagegroup -list**

**Note:** Both models will use a Celerra Storage Group

50:06:01:60:C1:E0:24:00:50:06:01:69:41:E0:24:00 SP B 3 →Ports 3 & 2 are used for NS20FC or NS40FC

50:06:01:60:C1:E0:24:00:50:06:01:68:41:E0:24:00 SP A 3 →Ports 0 & 1 are used for NS20

**# /nas/sbin/model**

NS20 or NS20FC

**Note:** Command calls /nas/bin/nas\_xml –info and obtains model name from the 2<sup>nd</sup> entry in the PRODUCT\_NAME field. However, the model command also needs to run server\_sysconfig to determine blade motherboard type and CPU information in order to determine model type. If the model command cannot run server\_sysconfig for some reason, the blade model will return as “UnknownFC”, “UnknownG”, etc.

**# nas\_xml -info head**

```
<CELERRA SRC='controlstation'>
<CELERRA_CABINET NAME='APM000738018380000' TYPE='Celerra SledgeHammer'
PRODUCT_NAME='Celerra NS20FC' SERIAL_NO='APM00073701085'
```

**Celerra Manager>Celerra Home>System Info screen will show model number**

## **BASIC SPECS FOR NS20 MODELS:**

→1-2 Data Mover 1U blades, Dual Intel 2.8GHz processors, (4) GB Memory, 800MHz FSB speeds, (4) 10/100/1000 GbE ports, (2) FC ports to the array, & (1) FC port for tape backups—FC port speeds support 2-4Gbps, 16TB capacity per blade FC

**Note:** Data Movers will connect to the SPs on the Internal network to the upper Management port on the SPs vs. the lower Service port used on the NS40—this is because the NS20 will run in SAN mode. NS20 will use NS40 frontend blades, and either CX3-10 (2 FE FC ports) or CX3-10F (6 FE FC ports) backends.

→Single Dewars 1U Control Station with 2GB memory & 2.4GHz CPU, 1 serial RS-232 port for modem, 1 serial RS-232 port on the front of the CS for serial access, 1 External 10/100 Ethernet port for customer's network [eth3—upper left LAN port], & (2) 10/100/1000 Ethernet ports for Internal Network [eth2 is labeled #2, located upper right LAN port; eth0 is labeled 10/100, located lower right LAN port]. Control Station will be labeled with Flare & NAS versions, date, & CS MAC address

→SPE enclosure contains (2) SPs, (1) Standby Power Supply (SPS), (1) or (2) NS20 Blades, and (1) Control Station

→DAE Enclosures have 3U footprint, max. of 60 FC and/or ATA drives allowed in same cabinet, but not same DAE3

## **DEWARS CS ETHERNET PORTS & LABELS:**

eth3 eth2 [labeled #2, connects to lower lan port on DM3]

eth1 eth0 [labeled 10/100, connects to lower lan port on DM2]

### **Raid Configurations:**

1+1 R1, 4+1 R5, 8+1 R5 for Fibre Channel drives

4+1 R5, 6+1R5, 4+1R3, 8+1R3 for SATA drives

**Drives Supported:** Max capacity 35TB

FC 73, 146GB, & 300GB 15K rpm drives

FC 146 & 300GB 10K rpm drives

SATA II Northstar 500 & 750GB 7200rpm drives

FC & SATA II drives can share same array but not same DAE

Minimum configuration of (6) SATA II or FC drives in 6+1 R5

**Note:** FC & SATA II drives can share the same array, but not the same shelf—system drives still require Fibre Channel Drives could be all SATA II, all FC, or mix of FC/ATA on the array, with max of (15) drives per DAE

## **IMPORTANT CHANGES WITH NS20/NS20FC/NS40/NS40FC MODELS:**

**Tiered Service Support model (NS20/20FC only), Enhanced Installability (preloaded code, Placemat, Landing Page, CSA), All-in-One FC Enabled model, new Integrated model, Clariion Mgmt access to Integrated backend SPs (Via new Control Station NAT Service, Clariion NST & OffArrayMgr tools)**

→SPs run in SAN Mode, DMs connect to SPs using Fibre Optic LC-to-LC cables, DMs connect to upper LAN Mgmt port on SP for internal network, Access Logix & Storage Groups used for all models, NST tool connects to new public IPs on the CS, in order to access private IP addresses for SPs on backend (via NAT service)

→NS20FC model adds additional FC ports FE ports for customer SAN Hosts, and BE ports 2/3 for additional DAE buses

→Mirrorview setups will now use Port 1 vs. the highest array Port number, meaning that new installs of NS20FC/NS40FC will have the DMs cabled to SP ports 2 & 3 Fibre vs. ports 0 & 1 Fibre

→All models should be shipped with pre-installed software and will use a colored Placemat contained in the Open Me First package, a Landing Page website for product information and support, and Celerra Startup Assistant (CSA) to configure the box on initial startup

→Based on CX3-10 or CX3-10F array (Celerra variants only), which are really stripped down CX3-20 arrays, less memory, use of all ATA drives is allowed, etc.

**Note:** Only CLARiiON-sold version is the CX3-10C combo

## **CHANGES ON NS80 INTEGRATED FOR NAPA 8:**

--NS80 already uses fibre optic DM-to-SP cabling & topology, as well as SPs that use the “SAN” RESUME personality

--NS80 will be factory-installed with AccessLogix & Celerra Storage Group on the Backend [same as what is being done with NS20]

--Storage & RAID Group info resides in /etc/be\_sg\_info, /etc/be\_rg\_info, resp.

--NS80 will incorporate SPA—SPB communication gateway aliases on CS, meaning that IPTables and NAT Service will be used for SP-to-SP communication when dual Control Stations are configured

--IP Forwarding will be set to 1 with upgrade or from new installs: /proc/sys/net/ipv4/ip\_forward

--The array will have security setup as a Global Domain with user “nasadmin” and default password “nasadmin” assigned, and the Celerra will have array security credentials cached in /nas/site/.clar\_security, for new installs of NS80I

--NS80 will not be configured to support the NST or OffArray Navi Mgr

**Note:** When adding blades to an existing NS80I configuration running NAS 5.5.30 or higher, use the following syntax in order to add the Blade FC ports to the default Storage group, which is a change from the traditional syntax for Integrateds

**# /nas/sbin/setup\_slot -i -g 4** [i.e., -gateway\_auto\_config]

### **Workaround:**

1.) Add Data Mover HBA WWNs manually to both SPs using navicli:

**# /nas/sbin/navicli -h <sp> storagegroup -setpath -o -gname Celerra\_emcnas\_i0 -hbauid <hbauid> -spa | -spb -spport 0 | -spport 1**

2.) Then use setup\_slot -i 4 to complete setup

## **CELERRA DHCP:**

**Two processes are running on the Control Station:**

# ps -eafl |grep dhcp

```
4 S root    3165  1 0 76 0- 779 wait 14:42 ? 00:00:02 /bin/sh /nasmcd/sbin/cs_dhcpd_monitor
1 S root    4473  1 0 75 0- 1015 - 14:43 ? 00:00:00 /usr/sbin/dhcpd -cf /etc/encl_dhcpd.conf eth0 eth2
```

/etc/encl\_dhcpd.conf

--dhcp is used to manage and enforce rules for IP address allocation for Primary (128.221.252.0) & Secondary (128.221.253.0) internal networks

--dhcp is used to assign an IP address to Data Movers when PXE booting for recovery purposes

--dhcp is used to assign APC\_UPS ranges when applicable [via static entries]

--dhcp is used to assign Primary & Secondary mgmt switch IP addresses on mgmt\_2\_3 & mgmt\_2\_3b (internal switches)

--dhcp is used for Enclosure Management for Blades when blades are set to NAS Resume Prom personality

/nas/sbin/cs\_dhcpd\_monitor

--CS monitors dhcp process (/var/run/dhcpd.pid) & automatically restarts if the daemon is stopped for some reason

--dhcp daemon only runs on the Control Station which is running as the Primary

--dhcp monitor service is used to bring up the alias interfaces eth0:1 & eth2:1 that are used for gateway IP addresses for SPA & SPB communications

--dhcp monitor service is used to monitor NAT/Proxy ARP service and restart if required using clarion\_mgmt –start\_service (checks that Proxy RP service is up every 30 secs)

--dhcp monitor feature is only used with IPDART systems (Hammer, Sledgehammer)

--dhcp monitor service ensures ip\_forwarding is always set to 1

--mounts /ftpboot directory

/etc/rc.d/init.d/dhcpd

--startup script which starts the dhcpcd daemon during Control Station reboots

/nas/site/encl\_dhcpd

--starts up dhcpcd dameon using /nas/sbin/setup\_enclosure –dhcpcd start &> /dev/null, and is used to restart daemon when necessary by **DOES CS ETH3 RUN DHCP Client?**

No, not normally. However, a person could run /usr/sbin/netconfig and select [ ] Use dynamic IP configuration (BOOTP/DHCP), and change the eth3 interface to run using DHCP. This would not be a supported function for Celerra.

**CELLERRA NAT (Linux IPTables, aka NAT Network Address Translation):**

**Purpose:**

→Beginning with new installations for Napa 8 (NAS 5.5.30.x), NS systems will implement NAT translation using the IPTables service on the CS.

→NAS Upgrades to Napa 8 also implement NAT & will recommend that the new gateway IPs be configured separately on the SPs

→Main purpose of NAT is to support the Clariion CRU & Storage Management tools [NST & Navisphere OffArray Manager]

→From a Linux perspective, IPTables are used in setting up IP forwarding, packet filtering rules, and Routing rules in tables, with the Linux system acting as a Gateway between Public and Private networks, usually in conjunction with NAT (Network Address Translation) or masquerading. Each table can consist of built-in chains and user-defined chains. Some default tables are “filter”, “nat”, “mangle”, etc., and have built-in chains, such as INPUT, FORWARD, OUTPUT, PREROUTING, POSTROUTING, etc.

**Important IPTables Commands/Files, etc.**

/sbin/service iptables start | stop | restart | status

/sbin/chkconfig iptables –list | -on

# /sbin/chkconfig --list iptables

```
iptables      0:off 1:off 2:on 3:on 4:on 5:on 6:off
```

# rpm -q iptables

iptables-1.2.5-3

/etc/sysconfig/iptables

/sbin/iptables-save >/home/nasadmin/ibak.doc (command will save the /etc/sysconfig/iptables)

/etc/init.d/iptables →initial install script to setup iptables

/nasmcd/sbin/cs\_dhcpd\_monitor →Used to bring up CS aliases if down

Monitors for DHCP daemon and starts; checks that CS aliases for gateway IPs to array are up; checks NAT/Proxy service is started

# /nasmcd/sbin/clarion\_mgmt –info | -stop | -start (Use to verify or recreate public SP IP addresses, NAT entries, and Aliases on Control Station)

**CONFIGURING NAT ON NAS 5.6.37:**

# /nasmcd/sbin/clarion\_mgmt -start -spa\_ip 192.1.4.220 -spb\_ip 192.1.4.221 -use\_nat

Checking if running as root...yes

Checking if model is supported...yes

Checking for integrated system...yes  
Checking if interface eth3 is configured...yes  
Checking if interface eth3:1 is configured...no  
Checking if interface eth3:2 is configured...no  
Checking if SP (128.221.252.200) is up...yes  
Checking if SP (128.221.253.201) is up...yes  
Removing network alias eth3:1...done  
Removing network alias eth3:2...done  
Creating network alias eth3:1...done  
Creating network alias eth3:2...done  
done

**Note:** You should not normally have a system running NAT—NAT was discontinued after 5.5.30, though allowed

### # /nasmcd/sbin/clariion\_mgmt -info

Public IP address for SPA: 192.1.4.220

Public IP address for SPB: 192.1.4.221

Start on boot : yes

Current implementation : NAT

Status : Started

done

## **DEFAULT FACTORY INSTALL CONFIGURATION 5.5.30.4:**

### **Pre-CSA Configuration on NS41 Integrated:**

#### **/etc/hosts**

|               |                                    |
|---------------|------------------------------------|
| 192.168.1.200 | A_APM00071600514 SPA # CLARiiON SP |
| 192.168.2.201 | B_APM00071600514 SPB # CLARiiON SP |

#### **# /nas/sbin/model**

NS40

**# cat /etc/sysconfig/iptables** (Pre-CSA configuration)

# Generated by iptables-save v1.2.5 on Tue Nov 13 14:38:24 2007

\*nat

:PREROUTING ACCEPT [0:0]

:POSTROUTING ACCEPT [0:0]

:OUTPUT ACCEPT [0:0]

-A POSTROUTING -s 192.168.1.200 -d 192.168.2.201 -p tcp -j SNAT --to-source 192.168.1.200

-A POSTROUTING -s 192.168.2.201 -d 192.168.1.200 -p tcp -j SNAT --to-source 192.168.2.201

-A POSTROUTING -s 192.168.1.200 -d 192.168.2.201 -p udp -j SNAT --to-source 192.168.1.200

-A POSTROUTING -s 192.168.2.201 -d 192.168.1.200 -p udp -j SNAT --to-source 192.168.2.201

COMMIT

# Completed on Tue Nov 13 14:38:24 2007

**# /nasmcd/sbin/clariion\_mgmt -info** →No /etc/clariion\_mgmt.cfg file

Error 12: Not configured

#### **# cat /nas/site/sp\_info**

|               |                                    |
|---------------|------------------------------------|
| 192.168.1.200 | A_APM00071600514 SPA # CLARiiON SP |
|---------------|------------------------------------|

|               |                                    |
|---------------|------------------------------------|
| 192.168.2.201 | B_APM00071600514 SPB # CLARiiON SP |
|---------------|------------------------------------|

**# cat /etc/be\_sg\_info** →Pure NS40 Integrated file—add FC\_ENABLED=YES to file & /nas/sbin/model –c to change to FC model

SHARED\_BE\_SYS\_RAID\_GROUP\_ID=0

STORAGE\_GROUP\_NAME=Celerra\_emcnas\_i0

STORAGE\_GROUP\_INIT\_REC=50:06:01:60:c1:e0:53:35:50:06:01:60:41:e0:53:35;emcnas\_i0\_dm2\_p0

STORAGE\_GROUP\_INIT\_REC=50:06:01:60:c1:e0:53:35:50:06:01:61:41:e0:53:35;emcnas\_i0\_dm2\_p1

STORAGE\_GROUP\_ALUS=0 1 2 3 4 5

#### **# nas\_disk -l**

| id | inuse | sizeMB | storageID-devID     | type  | name       | servers |
|----|-------|--------|---------------------|-------|------------|---------|
| 1  | y     | 11263  | APM00071600514-0000 | CLSTD | root_disk  | 1       |
| 2  | y     | 11263  | APM00071600514-0001 | CLSTD | root_ldisk | 1       |
| 3  | y     | 2047   | APM00071600514-0002 | CLSTD | d3         | 1       |
| 4  | y     | 2047   | APM00071600514-0003 | CLSTD | d4         | 1       |
| 5  | y     | 2047   | APM00071600514-0004 | CLSTD | d5         | 1       |
| 6  | y     | 2047   | APM00071600514-0005 | CLSTD | d6         | 1       |
| 7  | n     | 466747 | APM00071600514-0010 | CLSTD | d7         | 1       |
| 8  | n     | 466747 | APM00071600514-0011 | CLSTD | d8         | 1       |

**Note:** FC model would only have a single Data LUN based on 25% of total available space in RG0

# cat network → Prior to CS being assigned IP & Hostname

NETWORKING=yes

FORWARD\_IPV4=false

HOSTNAME=localhost.localdomain

# /sbin/ifconfig

```
eth0    Link encap:Ethernet HWaddr 00:04:23:DC:8C:E3  →Standard Primary Internal Network
        inet addr:192.168.1.100 Bcast:192.168.1.255 Mask:255.255.255.0
              UP BROADCAST NOTRAILERS RUNNING MULTICAST MTU:1500 Metric:1
eth0:0   Link encap:Ethernet HWaddr 00:04:23:DC:8C:E3  →Secondary Backup Internal Network
        inet addr:192.168.2.100 Bcast:192.168.2.255 Mask:255.255.255.0
              UP BROADCAST NOTRAILERS RUNNING MULTICAST MTU:1500 Metric:1
eth0:1   Link encap:Ethernet HWaddr 00:04:23:DC:8C:E3  →Gateway alias to SPA
        inet addr:192.168.1.104 Bcast:192.168.1.255 Mask:255.255.255.0
              UP BROADCAST NOTRAILERS RUNNING MULTICAST MTU:1500 Metric:1
eth2    Link encap:Ethernet HWaddr 00:0E:0C:CF:C7:2B  →Gateway alias to SPB
        inet addr:192.168.2.102 Bcast:192.168.2.255 Mask:255.255.255.254
              UP BROADCAST NOTRAILERS RUNNING MULTICAST MTU:1500 Metric:1
eth2:1   Link encap:Ethernet HWaddr 00:0E:0C:CF:C7:2B  →Special route between CS & SPB, single blade system
        inet addr:192.168.2.104 Bcast:192.168.2.255 Mask:255.255.255.254
              UP BROADCAST NOTRAILERS RUNNING MULTICAST MTU:1500 Metric:1
              Base address:0xcc00 Memory:fc5e0000-fc600000
eth3    Link encap:Ethernet HWaddr 00:0E:0C:CF:C7:2A  →Not yet configured until after CSA
        UP BROADCAST NOTRAILERS RUNNING MULTICAST MTU:1500 Metric:1
```

# cat ifcfg-eth2

DEVICE=eth2

IPADDR=192.168.2.102

NETMASK=255.255.255.54 → Netmask is really 254, benign file issue

NETWORK=192.168.2.201

BROADCAST=192.168.2.255

ONBOOT=yes

# cat ifcfg-eth2:1

DEVICE=eth2:1

IPADDR=192.168.2.104

NETMASK=255.255.255.54 → Netmask is really 254

NETWORK=192.168.2.201

BROADCAST=192.168.2.255

ONBOOT=no

# /nas/sbin/navicli -h 192.168.1.200 security -list

Username: nasadmin

Role: administrator

Scope: global

# ls -la /etc.clar\*

```
-rw-rw-r-- 1 nasadmin nasadmin 276 Nov 13 16:05 /etc.clar_security
-rw-rw-r-- 1 nasadmin nasadmin 0 Nov 13 16:05 /etc.clar_security.lck
```

#### Post-CSA NAT Configuration on NS41 Integrated [5.5.30 code]:

# /nasmcd/sbin/clariion\_mgmt -info

Public alias IP address for SPA: 192.1.4.214

Public alias IP address for SPB: 192.1.4.215

Start on boot : yes

Status : Started

Done

#### 5.5.32.4 Output:

# /nasmcd/sbin/clariion\_mgmt -info

Public IP address for SPA: 192.1.4.220

Public IP address for SPB: 192.1.4.221

Start on boot : yes

Current implementation : NAT

Status : Started

done

# cat /etc/clariion\_mgmt.cfg

SPA\_PUBLIC\_IP=192.1.4.214

SPB\_PUBLIC\_IP=192.1.4.215

ONBOOT=yes

**Note:** With NAT, this file shows the pseudo public SP IP addresses that were applied during CSA

# cat /nas/site/sp\_info

192.168.1.200 A\_APM00071600514 SPA # CLARiiON SP

192.168.2.201 B\_APM00071600514 SPB # CLARiiON SP

# /nas/sbin/navicli -h 192.168.1.200 networkadmin -get

Storage Processor: SP A

Storage Processor Network Name: spa

Storage Processor IP Address: 192.168.1.200

Storage Processor Subnet Mask: 255.255.255.0

# tail /etc/hosts

192.168.1.200 A\_APM00071600514 SPA # CLARiiON SP

192.168.2.201 B\_APM00071600514 SPB # CLARiiON SP

192.1.4.218 sludge\_4.w2k.pvt.dns sludge\_4

# cat /etc/sysconfig/iptables (NAT IPTables setup after CSA)

# Generated by iptables-save v1.2.5 on Tue Nov 13 21:28:35 2007

\*nat

:PREROUTING ACCEPT [369:30551]

:POSTROUTING ACCEPT [3345:229190]

:OUTPUT ACCEPT [3331:228742]

-A PREROUTING -d 192.1.4.214 -p tcp -j DNAT --to-destination 192.168.1.200

-A PREROUTING -d 192.1.4.214 -p udp -j DNAT --to-destination 192.168.1.200

-A PREROUTING -d 192.1.4.214 -p icmp -j DNAT --to-destination 192.168.1.200

-A PREROUTING -d 192.1.4.215 -p tcp -j DNAT --to-destination 192.168.2.201

-A PREROUTING -d 192.1.4.215 -p udp -j DNAT --to-destination 192.168.2.201

-A PREROUTING -d 192.1.4.215 -p icmp -j DNAT --to-destination 192.168.2.201

-A POSTROUTING -s 192.168.1.200 -d 192.168.2.201 -p tcp -j SNAT --to-source 192.168.1.200

-A POSTROUTING -s 192.168.2.201 -d 192.168.1.200 -p tcp -j SNAT --to-source 192.168.2.201

-A POSTROUTING -s 192.168.1.200 -d 192.168.2.201 -p udp -j SNAT --to-source 192.168.1.200

-A POSTROUTING -s 192.168.2.201 -d 192.168.1.200 -p udp -j SNAT --to-source 192.168.2.201

-A POSTROUTING -s 192.168.1.200 -p tcp -j SNAT --to-source 192.1.4.214

-A POSTROUTING -s 192.168.1.200 -p udp -j SNAT --to-source 192.1.4.214

-A POSTROUTING -s 192.168.1.200 -p icmp -j SNAT --to-source 192.1.4.214

-A POSTROUTING -s 192.168.2.201 -p tcp -j SNAT --to-source 192.1.4.215

-A POSTROUTING -s 192.168.2.201 -p udp -j SNAT --to-source 192.1.4.215

-A POSTROUTING -s 192.168.2.201 -p icmp -j SNAT --to-source 192.1.4.215

COMMIT

# Completed on Tue Nov 13 21:28:35 2007

# Generated by iptables-save v1.2.5 on Tue Nov 13 21:28:35 2007

\*filter

:INPUT ACCEPT [177992:128951926]

:FORWARD ACCEPT [43:4261]

:OUTPUT ACCEPT [196395:194804957]

COMMIT

# Completed on Tue Nov 13 21:28:35 2007

# /sbin/iptables -t nat -L -n

Chain PREROUTING (policy ACCEPT)

target prot opt source destination

DNAT tcp -- 0.0.0.0/0 192.1.4.214 to:192.168.1.200

DNAT udp -- 0.0.0.0/0 192.1.4.214 to:192.168.1.200

DNAT icmp -- 0.0.0.0/0 192.1.4.214 to:192.168.1.200

DNAT tcp -- 0.0.0.0/0 192.1.4.215 to:192.168.2.201

DNAT udp -- 0.0.0.0/0 192.1.4.215 to:192.168.2.201

DNAT icmp -- 0.0.0.0/0 192.1.4.215 to:192.168.2.201

Chain POSTROUTING (policy ACCEPT)

target prot opt source destination

```

SNAT    tcp -- 192.168.1.200    192.168.2.201    to:192.168.1.200
SNAT    tcp -- 192.168.2.201    192.168.1.200    to:192.168.2.201
SNAT    udp -- 192.168.1.200    192.168.2.201    to:192.168.1.200
SNAT    udp -- 192.168.2.201    192.168.1.200    to:192.168.2.201
SNAT    tcp -- 192.168.1.200    0.0.0.0/0      to:192.1.4.214
SNAT    udp -- 192.168.1.200    0.0.0.0/0      to:192.1.4.214
SNAT    icmp -- 192.168.1.200   0.0.0.0/0      to:192.1.4.214
SNAT    tcp -- 192.168.2.201   0.0.0.0/0      to:192.1.4.215
SNAT    udp -- 192.168.2.201   0.0.0.0/0      to:192.1.4.215
SNAT    icmp -- 192.168.2.201   0.0.0.0/0      to:192.1.4.215

```

### # cat /etc/sysconfig/network

```

NETWORKING=yes
GATEWAY=192.1.4.254
GATEWAYDEV=eth3
HOSTNAME=sludge_4
DOMAINNAME=w2k.pvt.dns
# cat static-routes →Special route between CS & SPB due to single Blade configuration
eth2 host 192.168.2.201 dev

```

# /sbin/ifconfig (eth3 now has external IP address and new eth3:1 & eth3:2 aliases created for SP IP addresses)

```

eth3    Link encap:Ethernet HWaddr 00:0E:0C:CF:C7:2A
        inet addr:192.1.4.218 Bcast:192.1.4.255 Mask:255.255.255.0
        UP BROADCAST NOTRAILERS RUNNING MULTICAST MTU:1500 Metric:1
eth3:1   Link encap:Ethernet HWaddr 00:0E:0C:CF:C7:2A →NAT alias for SPA
        inet addr:192.1.4.214 Bcast:192.1.4.255 Mask:255.255.255.0
        UP BROADCAST NOTRAILERS RUNNING MULTICAST MTU:1500 Metric:1
eth3:2   Link encap:Ethernet HWaddr 00:0E:0C:CF:C7:2A →NAT alias for SPB
        inet addr:192.1.4.215 Bcast:192.1.4.255 Mask:255.255.255.0
        UP BROADCAST NOTRAILERS RUNNING MULTICAST MTU:1500 Metric:1
eth0:1   Link encap:Ethernet HWaddr 00:04:23:DC:8C:E3 →Internal Gateway alias on CS to SPA
        inet addr:192.168.1.104 Bcast:192.168.1.255 Mask:255.255.255.0
eth2:1   Link encap:Ethernet HWaddr 00:0E:0C:CF:C7:2B →Internal Gateway alias on CS to SPB
        inet addr:192.168.2.104 Bcast:192.168.2.255 Mask:255.255.255.0

```

### # /sbin/arp -n -i eth3

| Address     | HWtype | HWaddress         | Flags | Mask | Iface |
|-------------|--------|-------------------|-------|------|-------|
| 192.1.4.254 | ether  | 00:D0:97:10:07:38 | C     |      | eth3  |
| 192.1.4.248 | ether  | 00:0B:DB:90:0A:21 | C     |      | eth3  |

### # cat /etc/sysconfig/network-scripts/ifcfg-eth3

```

DEVICE=eth3
IPADDR=192.1.4.218
NETMASK=255.255.255.0
NETWORK=192.1.4.0
BROADCAST=192.1.4.255
ONBOOT=yes

```

### FACTORY INSTALL IFCFG-ETH3 FILE:

#### # cat ifcfg-eth3

```

DEVICE=eth3
USERCTL=no
ONBOOT=no
BOOTPROTO=none

```

#### # /sbin/route -n

Kernel IP routing table

| Destination   | Gateway     | Genmask         | Flags | Metric | Ref | Use | Iface |
|---------------|-------------|-----------------|-------|--------|-----|-----|-------|
| 192.168.2.201 | 0.0.0.0     | 255.255.255.255 | UH    | 0      | 0   | 0   | eth2  |
| 192.168.2.102 | 0.0.0.0     | 255.255.255.254 | U     | 0      | 0   | 0   | eth2  |
| 192.1.4.0     | 0.0.0.0     | 255.255.255.0   | U     | 0      | 0   | 0   | eth3  |
| 192.168.2.0   | 0.0.0.0     | 255.255.255.0   | U     | 0      | 0   | 0   | eth0  |
| 192.168.2.0   | 0.0.0.0     | 255.255.255.0   | U     | 0      | 0   | 0   | eth2  |
| 192.168.1.0   | 0.0.0.0     | 255.255.255.0   | U     | 0      | 0   | 0   | eth0  |
| 127.0.0.0     | 0.0.0.0     | 255.0.0.0       | U     | 0      | 0   | 0   | lo    |
| 0.0.0.0       | 192.1.4.254 | 0.0.0.0         | UG    | 0      | 0   | 0   | eth3  |

```
# /sbin/sysctl net.ipv4.conf.eth0.proxy_arp
net.ipv4.conf.eth0.proxy_arp = 0
# /sbin/sysctl net.ipv4.conf.eth2.proxy_arp
net.ipv4.conf.eth2.proxy_arp = 0
# /sbin/sysctl net.ipv4.conf.eth3.proxy_arp
net.ipv4.conf.eth3.proxy_arp = 0
```

**Note:** Proxy ARP is not yet setup—pure NAT implementation with 5.5.30 code & CSA

**# /sbin/iptables -t nat -L** (verify internal gateway aliases & check NAT rules)

Chain PREROUTING (policy ACCEPT)

| target | prot | opt | source   | destination                     |
|--------|------|-----|----------|---------------------------------|
| DNAT   | tcp  | --  | anywhere | 10.241.168.250 to:192.168.1.200 |
| DNAT   | udp  | --  | anywhere | 10.241.168.250 to:192.168.1.200 |
| DNAT   | icmp | --  | anywhere | 10.241.168.250 to:192.168.1.200 |
| DNAT   | tcp  | --  | anywhere | 10.241.168.251 to:192.168.2.201 |
| DNAT   | udp  | --  | anywhere | 10.241.168.251 to:192.168.2.201 |
| DNAT   | icmp | --  | anywhere | 10.241.168.251 to:192.168.2.201 |

Chain POSTROUTING (policy ACCEPT)

| target | prot | opt | source           | destination                       |
|--------|------|-----|------------------|-----------------------------------|
| SNAT   | tcp  | --  | A_APM00071600514 | B_APM00071600514 to:192.168.1.200 |
| SNAT   | tcp  | --  | B_APM00071600514 | A_APM00071600514 to:192.168.2.201 |
| SNAT   | udp  | --  | A_APM00071600514 | B_APM00071600514 to:192.168.1.200 |
| SNAT   | udp  | --  | B_APM00071600514 | A_APM00071600514 to:192.168.2.201 |

Chain OUTPUT (policy ACCEPT)

| target | prot | opt | source | destination |
|--------|------|-----|--------|-------------|
|--------|------|-----|--------|-------------|

**# /sbin/iptables -t nat -L POSTROUTING**

### **Removing CS aliases for gateway IPs, NAT entries for SPs, & gateway IPs on SPs (if required):**

1. Remove existing gateway aliases for eth0:1 & eth2:1

```
# /sbin/ifdown eth0:1
# /sbin/ifdown eth2:1
# rm -f /etc/sysconfig/network-scripts/ifcfg-eth0:1
# rm -f /etc/sysconfig/network-scripts/ifcfg-eth2:1
```

2. Remove POSTROUTING entries in /etc/sysconfig/iptables related to SPA-to-SPB

```
/sbin/iptables -t nat -D POSTROUTING -p tcp -s 192.168.1.200 -d 192.168.2.201 -j SNAT --to-source 192.168.1.200
/sbin/iptables -t nat -D POSTROUTING -p tcp -s 192.168.2.201 -d 192.168.1.200 -j SNAT --to-source 192.168.2.201
/sbin/iptables -t nat -D POSTROUTING -p udp -s 192.168.1.200 -d 192.168.2.201 -j SNAT --to-source 192.168.1.200
/sbin/iptables -t nat -D POSTROUTING -p udp -s 192.168.2.201 -d 192.168.1.200 -j SNAT --to-source 192.168.2.201
```

3. If necessary, delete gateway addresses on SPs (requires SP reboot)

```
# /nas/sbin/navicli -h 192.168.1.200 networkadmin -set -gateway 0.0.0.0
# /nas/sbin/navicli -h 192.168.1.200 networkadmin -get
```

Storage Processor: SP A

Storage Processor Network Name: SPA

Storage Processor IP Address: 192.168.1.200

Storage Processor Subnet Mask: 255.255.255.0

Storage Processor Gateway Address: 192.168.1.104

**Note:** NAS 5.6.45 introduces the /nas/sbin/spa\_spb\_comm script, which maintains rules for SP-to-Control Station communications.

Among the rules are checks to ensure that the original default internal IP address entries are maintained in the iptables file, that the alias interface files are maintained with a default gateway for SPA of 128.221.252.104, and SPB 128.221.253.104, even though Proxy ARP may be configured.

### **Rebuilding Gateway addresses & NAT entries for SPA/SPB on CS & SPs:**

1. Recreate CS aliases, IPTTables entries, and Gateway addresses on SPs:

- a. Vi edit to create alias files /etc/sysconfig/network-scripts/ifcfg-eth0:1, ifcfg-eth2:1
 

```
# vi ifcfg-eth0:1
DEVICE=eth0:1
IPADDR=192.168.1.104
# vi ifcfg-eth2:1
DEVICE=eth2:1
IPADDR=192.168.2.104
```
- b. Bring alias interfaces up:

```
/nasmcd/sbin/cs_dhcpd_monitor
c. Rebuild the proper gateway SP-to-SP routing rules using /sbin/iptables:
/sbin/iptables -t nat -A POSTROUTING -p tcp -s 192.168.1.200 -d 192.168.2.201 -j SNAT --to-source 192.168.1.200
/sbin/iptables -t nat -A POSTROUTING -p tcp -s 192.168.2.201 -d 192.168.1.200 -j SNAT --to-source 192.168.2.201
/sbin/iptables -t nat -A POSTROUTING -p udp -s 192.168.1.200 -d 192.168.2.201 -j SNAT --to-source 192.168.1.200
/sbin/iptables -t nat -A POSTROUTING -p udp -s 192.168.2.201 -d 192.168.1.200 -j SNAT --to-source 192.168.2.201
# /sbin/iptables-save >/etc/sysconfig/iptables →This takes the new entries from memory and writes to the iptables file
d. Rebuild SP gateway addresses (only if required, since SPs will need reboot)
# /nas/sbin/navicli -h 192.168.1.200 networkadmin -set -gateway 192.168.1.104
# /nas/sbin/navicli -h 192.168.2.201 networkadmin -set -gateway 192.168.2.104
```

## **CONTROL STATION AND SP COMMUNICATIONS:**

Beginning with NAS 5.6.43, a new script /nas/sbin/spa\_spb\_comm was created to enforce certain rules for the maintenance of a framework for communication between the Control Station and SPs. The rules themselves are checked whenever the PUHC or PAHC scripts are run, or when spa\_spb\_comm –status or –setup is run. The main rules check to ensure that the interface aliases for eth0:1 and eth2:1 exist with the correct gateway entry for the respective SP, that the /etc/hosts have entries for each SP, and that the internal IP address routes for the SPs exist in the /etc/sysconfig/iptables file.

### **Examples:**

#### **/nas/bin/nas\_checkup**

Storage System : Checking SPA SPB communication..... Fail  
-----Errors-----

Storage System: Check SPA SPB communication

Symptom: At least one iptables rule is missing.

Action : Run "spa\_spb\_comm -setup" to fix the problem.

### **Fix:**

#### **# /nas/sbin/spa\_spb\_comm -setup**

Setting up SPA-SPB communication

Flushing firewall rules: [ OK ]  
Setting chains to policy ACCEPT: nat filter [ OK ]  
Unloading iptables modules: [ OK ]  
Applying iptables firewall rules: [ OK ]

Setup completed OK

### **Status Examples:**

#### **# /nas/sbin/spa\_spb\_comm -status**

Checking Gateway of SPA

Gateway for SPB not found

**# /nas/sbin/spa\_spb\_comm -status** -->Example of normal output with no problems detected

Checking Gateway of SPA

Checking Gateway of SPB

### **Example interface alias file is created if missing:**

#### **# /nas/sbin/spa\_spb\_comm -setup**

Setting up SPA-SPB communication

Flushing firewall rules: [ OK ]  
Setting chains to policy ACCEPT: filter nat [ OK ]  
Unloading iptables modules: [ OK ]  
Applying iptables firewall rules: [ OK ]

Checking Gateway of SPA

SPA gateway is already set to 128.221.252.104

SPB's gateway is not 128.221.253.104. Please change it at the earliest convenience.

**Note:** Though the message says to change SPB's gateway at the earliest convenience, the interface alias file is in fact created if it is missing, though it may not be loaded into Control Station memory. Run /sbin/ifup <alias\_name> followed by /sbin/ifconfig to verify whether the alias is up. The User should verify that the SP in question is setup with the correct Network and Gateway IP addresses, and change accordingly using Navisphere or Navicli.

#### **/nas/sbin/spa\_spb\_comm -setup force**

**Note:** This syntax will apply the default internal IP address gateway address to the SP's, and will reboot the SPs. The “force” switch also recreates the gateway alias eth0:1 or eth2:1 files if none exists, or creates a new alias file if the existing one has something other than the default .104 gateway. Additionally, the “force” switch will load the newly created alias file into Control Station memory. See emc210468 for additional information.

#### **# /nas/sbin/spa\_spb\_comm -setup force**

Setting up SPA-SPB communication

Flushing firewall rules: [ OK ]  
Setting chains to policy ACCEPT: filter nat [ OK ]  
Unloading iptables modules: [ OK ]  
Applying iptables firewall rules: [ OK ]

Checking Gateway of SPA

SPA gateway is already set to 128.221.252.104

Changing the gateway of SPB → In this example, had 128.221.253.10 as gw—force changes back to 128.221.253.104

## **GLOBAL ARRAY SECURITY FOR NEW PLATFORMS, & NAVISECCLI:**

- A factory-installed Celerra system will setup Global Domain Security on SPA using a default User “nasadmin” with Password “nasadmin”, with Global Scope 0 and Role of “administrator”
- Existing system upgrades will not modify Global Array Security settings or add Global Security settings—it will allow continued use of Classic Navi
- Array security & local Celerra /etc/.clar\_security cache files are in sync at install to support Classic Navicli commands

## **CREATING ARRAY SECURITY FILE FOR USING NAVISECCLI COMMANDS:**

- Navisecccli security files are stored separately on the Clariion from the Global Array security files, though the username and password are the same
  - In order to use Navisecccli commands, an additional step needs to be taken to create the Navisecccli security file, using the same Username and password as that created by the install for Clariion Global security
- a.) # /nas/sbin/navisecccli -h 192.168.1.200 getagent** [no security file for Secure CLI]  
Security file does not exist. Use navisecccli -AddUserSecurity to create a security file. You must also have a valid user account on the storage system to issue this command. If you do not have a user account, use navisecccli -address <IPAddress | NetworkName> security -adduser to add a user account.

## **b.) # /nas/sbin/navisecccli -h 192.168.1.200 –AddUserSecurity –Scope 0 –user nasadmin**

- Enter password: [enter same password as used for Global Array security acnt]  
→ Storage systems do not allow the use of Secure CLI (navisecccli) commands until the User accountname, password, and scope are defined on the array

## **Viewing Global Array Security from GUI or CLI:**

### **Navisphere>Tools>Security>User Management**

- shows Username ‘nasadmin’ with Administrator Role and Global scope

### **# /nas/sbin/navisecccli -h 192.168.1.200 security -list**

Username: nasadmin  
Role: administrator  
Scope: global

### **# nas\_storage -info id=1 |grep -C1 username**

authenticated = True  
username = nasadmin

## **Three Methods for changing Global Array Security Password/Role:**

1. Navisphere>Tools>Security>Change Password
2. # /nas/sbin/navisecccli -h 192.168.1.200 security -changeuserinfo -user nasadmin -scope global -newpassword nasadmin  
WARNING: You are about to change user: nasadmin (global)  
Proceed?(y/n) y
3. # nas\_storage -modify id=1 -security -username nasadmin -password nasadmin -newpassword ns20isforme  
Changing password on APM00071600514  
done

**Note:** Of the above methods, only nas\_storage -modify command is recommended since it updates the password on the array & the local /etc/.clar\_security cache file on the Celerra—if the other two methods were used, would need to run nas\_storage -modify –security to update .clar\_cache files.

## **USING NAVISECCLI COMMANDS FROM CELERRA:**

**Note:** We still require the following step in order to update the security file for Navisecccli in order to use navisecccli commands

### **# /nas/sbin/navisecccli -h 192.168.1.200 –AddUserSecurity –scope 0 –user nasadmin**

Enter password:

## **CREATING A GLOBAL ARRAY SECURITY ACCOUNT IF NONE EXISTS ON ARRAY:**

### **# /nas/sbin/navisecccli -h 192.168.1.200 security -adduser -user nasadmin -password nasadmin -scope global -role administrator**

WARNING: You are about to add user: nasadmin

Proceed?(y/n) y

This storage system is not in a domain. It is highly recommended you create a domain.

**Note:** If this is a Celerra-owned array (e.g.,NS20/20FC), log into the array using Navisphere and then set the Domain Master to SPA

### **ADDING ANOTHER GLOBAL USER SECURITY ACCOUNT TO THE ARRAY:**

# /nas/sbin/navisecli -h 192.168.1.200 –user nasadmin –password nasadmin –scope 0 security -adduser -user second\_admin -password second\_admin -scope global -role administrator

## **NS20 FIBRE CHANNEL & NETWORK CABLING:**

### **DM-to-SP Connectivity:**

DM Blades connect to (2) FE FC ports on the SPs (Fibre Optic cables & SFP connectors)

Total of (2) FE Fibre Channel ports on the array

--DM2 Blade BE0 connected to SPA port 0 Fibre

--DM2 Blade BE1 connected to SPB port 0 Fibre

--DM3 Blade BE0 connected to SPA port 1 Fibre

--DM3 Blade BE1 connected to SPB port 1 Fibre

### **DM, CS, & SP LAN Connectivity:**

--DM2 blade lower LAN port to CS 10/100 port

--DM2 blade upper LAN port to SPA upper Management LAN port

--DM3 blade lower LAN port to CS #2 Port

--DM3 blade upper LAN port to SPB upper Management LAN port

### **SP-to-DAE Connectivity:**

--SPA FC Port BE0 to LCCA PRI port on DAE3P

--SPB FC Port BE0 to LCCB PRI port on DAE3P

## **NS20FC ENABLED FIBRE CHANNEL & NETWORK CABLING:**

### **DM-to-SP Connectivity:**

DM Blades connect to (2) FE FC ports on the SPs (Fibre Optic cables & SFP connectors)

Additionally, SPs have (4) extra FE FC optical ports for direct-connect or Switched Fabric Host connections

Total of (6) FE Fibre Channel ports on the array

--DM2 Blade BE0 connected to SPA port 2 Fibre

--DM2 Blade BE1 connected to SPB port 2 Fibre

--DM3 Blade BE0 connected to SPA port 3 Fibre

--DM3 Blade BE1 connected to SPB port 3 Fibre

**Note:** SP FC ports 0, 1, 4, 5 available for other Hosts, with Port 1 recommended for Mirrorview

### **DM, CS, & SP LAN Connectivity:**

--DM2 blade lower LAN port to CS 10/100 port

--DM2 blade upper LAN port to SPA upper Management LAN port

--DM3 blade lower LAN port to CS #2 Port

--DM3 blade upper LAN port to SPB upper Management LAN port

### **SP-to-DAE Connectivity:**

--SPA FC Port BE0 to LCCA PRI port on DAE3P

--SPB FC Port BE0 to LCCB PRI port on DAE3P

## **USING NAT-ENABLED TOOLS TO ACCESS CLARIION BACKENDS:**

→NST can be used on NS20/40 & NS20FC/40FC models, selecting NAT option on connection screen

→Navisphere OffArray Manager can only be used on NS20FC/40FC models, selecting NAT option

**Note:** OffArray restriction has been lifted with the advent of the Proxy ARP solution, which became the standard CSA configuration with 5.5.31

## **NAVISPHERE SERVICE TASKBAR (NST):**

→NST Version 6.26.50.1.14 or higher is required to support the Celerra Integrated functionality that the NST tool provides

→JRE 1.5.0\_06

**Note:** When properly connected to both SP's on an Integrated Celerra using the NAT option with the NST logon screen, certain functionality is restricted and a popup will be displayed when a User clicks on one of the following items:

Download and Install Hot Fix

Prepare for Installation

Register Storage

Software Assistant>Install Software

Software Assistant>Help menu

**“This feature is not supported on your Celerra integrated system.”**

→For the new NS20/20FC and the new NS40/40FC Celerras, the Navisphere Service Taskbar will be available to Customers for performing limited troubleshooting tasks and hardware maintenance, and is considered an off-array tool.

→The NST & OffArray Navisphere will be allowed for the NS20FC/NS40FC models (Navisphere Web Manager is not NAT aware and will not work).

→NST will be capable of auto-updates, a new feature in Flare 26

**Note:** Since the SPs are not normally accessible via the public network for Integrated units, a new strategy has been devised to allow for two new public IP addresses on the Control Station, and a “Celerra Identifier” flag will be set by the “Celerra Service Enabler” API in Flare that will allow the NST to access either SPA or SPB from a Windows host by using NAT translation to pass requests from the Windows Host to the SPs via the Control Station.

**# /nas/opt/Navisphere/bin/navicli -h 192.168.1.200 ndu -list**

Name of the software package: -CelerraService

Revision of the software package: -

Commit Required: NO

Revert Possible: NO

Active State: YES

Is installation completed: YES

Is this System Software: NO

## **HOW NST WORKS**

→User must select Options from NST screen and enter the public IPs for both SPA & SPB in the NAT connection dialogue box in order to connect to the internal IP addresses on the SPs

→User then logs into Clariion with username and password

→After logon, a CIMOM API call is made to the SP in order to retrieve the CelerraIdentifier flag (if the flag exists with the CelerraService pkg running on the backend, that indicates an array dedicated to Celerra)

→Once CelerraIdentifier is found, the JnfxRunEnvironment sets the bCelerraEnabled flag to true, at which point certain NST functions will become “disabled” and made known to the User via the use of system PopUps

## **NST/NAT COMPONENTS:**

→CelerraService pkg running on Array

→NST API to check CelerraIdentifier flag on backend

→NAT (Network Address Translation) service running on Control Station, performs pass-through routing to SPs from the NST using Linux IPTables that will be created initially from the CSA interface, or manually by using clarion\_mgmt, and creates the NAT Service and IPTables, Control Station aliases eth3:1 & eth3:2 will be configured for the two new IP addresses (SPA, SPB) that will allow the Public-to-Private NAT pass-through service from the NST and/or OffArrayMgr utility, and Control Station aliases eth0:1 & eth2:1 will be created to provide for the SP-to-SP communications that are needed in order to route packets between the 192.168.1 and 192.168.2 networks (192.168.1.104 SPA 192.168.2.104 SPB), with the CS acting as the Router between the SPs.

→The Windows host would open the NST application, select the “Network Address Translation (NAT) Connection” option, then enter the public IP addresses (Hosted on Control Station) for both SPA & SPB. Once the connection is made, the User logs in using the username and password configured on the Clariion backend for global security. An API in the NST tool contacts the SPs and checks for the existence of the CelerraService. If found, the system is considered a Celerra backend configuration and certain tasks will be disabled by the NST tool.

## **Functions that are disabled on the NST utility for Celerra Integrateds:**

NST>Tools>Software Maintenance Status →Unsupported Feature popup

Hardware Registration>Register Storage System →Unsupported Feature popup

Software Assistant: →All wizards disabled, Unsupported Feature popup

Download and Install Hot Fix

Prepare for Installation (Step-1)

Install Software (Step-2)

## **Functions that are Enabled for the NST Wizard:**

Hardware Installation> Install Disk Array Enclosure & Install Disk

Hardware Maintenance

Verify Storage System →Array Healthcheck tool that does System check, captures config data, and reports on issues found. XML & .zip captures are written to c:\EMC\repository\APM00071600514\HCK directory

Replace Disk →This is commonly known as the “DRU Tool”, copies of analysis summary kept in c:\EMC\repository\APM00071600514\DRU directory

Capture SPCollect →Gathers SPCollects and writes to c:\EMC\repository\APM00071600514\HCK directory, though offers ability to Browse to document Repository or use the “Save as” to save to new location

Engineering> Upgrade Disk Firmware (for disks, accessed in Eng Mode only)

Engineering Mode Offers one additional wizard

Ctrl + shift + f12 Password: SIR

Engineering>Upgrade Disk Firmware

## **Testing “Replace Disk” DRU Tool using Navicli:**

1. Pick drive to disable (must be part of a RAID Group—will not show up otherwise in getcrus list)
2. Disable drive using navicli and verify:

```
# /nas/sbin/navicli -h 192.168.1.200 cru_on_off -messner 0_0_12 0 (this places disk 12 in Bus 0 Enclosure 0 down)
# /nas/sbin/navicli -h 192.168.1.200 getcrus
```

DAE3P Bus 0 Enclosure 0 \*FAULT\*

(Bus 0 Enclosure 0 : Faulted; Bus 0 Enclosure 0 Disk 12 : Removed)

3. Perform Replace Disk procedure using NST Hardware Maintenance>Replace Disk Wizard

4. At last screen after replacing disk, you will need to issue following command to let the tool find the drive and complete task (since we manually disabled, need to manually re-enable)

```
# /nas/sbin/navicli -h 192.168.1.200 cru_on_off -messner 0_0_12 1 [places drive back up]
```

**Note:** If there are two faulted drives in the same Raid Group, the “Disk Analysis Summary” says that the “Storage system is not a candidate for customer disk replacement” and the Details button shows the “Override Issues” option as greyed out, meaning that the issue needs to be escalated to EMC support.

## **How NST connects to backend SPs over public network:**

→Start NST from Windows, click radial button for “Network Address Translation” (NAT), select “both” SPA & SPB’s public IP address, enter username & password, and connection is established if the CelerraEnabler was discovered via the Clariion Identifier API, and IP requests are forwarded to the SP via the NAT Service & IPTables

## **CONTROL STATION, SPA & SPB, COMMUNICATION PATHS:**

### **Public Passthrough eth3:1 or eth3:2:**

→NST User connects to public IP for SPA, which resides on CS alias eth3:1, and passes traffic to Internal CS eth0 via NAT, to SPA  
→NST User connects to public IP for SPB, which resides on CS alias eth3:2, and passes traffic to Internal CS eth2 via NAT, to SPB

**Note:** The SP’s public IP addresses & aliases, and IPTables NAT entries are configured on the Control Station, either during CSA setup or separately by the clarion\_mgmt CLI tool

### **SP-to-SP Communications over Internal Network eth0:1 & eth2:1:**

→Communications between SPA & SPB are routed by CS between Internal gateway aliases eth0:1 192.168.1.104 (SPA) and eth2:1 192.168.2.104 (SPB) in situations where Dual Control Stations are setup (e.g., NS80 platform)

**Note:** Secondary CS is setup with the same ‘Floating IP’ configuration & is not enabled until failover—the CS running NAS Services also runs the Floating IP scheme.

→SP-to-SP communication paths are setup during system factory install, as are the requisite NAT rules

## **CELLERRA CONFIGURATION CHANGES FOR NST SUPPORT:**

→During initial setup using the CSA (Celerra Startup Assistant), two public IP Addresses will be configured as aliases eth3:1 & eth3:2 for SPA & SPB on the Control Station

→A new NAT translation service on the Control Station will serve as a passthrough service for the NST & OffArray Utilities (NS20FC/NS40FC only)

→The NST or OffArray Navi Manager runs on a Windows system on the customer’s public network

→The OffArray Navi Manager software & FC Enablement license will not be licensed for non-FC models

→For the Internal Celerra network, the Control Station will host two new aliases, eth0:1 & eth2:1, that will use the 192.168.1.104 (SPA) & 192.168.2.104 (SPB) IPs as gateways for SP-to-SP access

→CSA calls the following script to configure the NST interfaces, or could be used manually:

```
/nasmed/sbin/clarion_mgmt -stop | -start -spa_ip xxx.xxx.xxx -spb_ip xxx.xxx.xxx
```

→LUNs 16 & 17 will be created on the exclusive Celerra RAIDGroup 0 for Integrateds, and for the All-in-One if more than 6 disk drives are detected—otherwise will reserve some LUN 0 space

### **Maintenance Scenarios:**

Disk to be replaced by EMC using DRU (Clariion Disk Replacement Utility—part of the NST)

Pwr Supplies, Fans, SPS to be replaced by EMC using Landing Page link

SFP/Cabling, Blade, SP, LCC, CS, & Enclosure/Midplane will be replaced by EMC using Landing Page

## **NAVISPHERE OFFARRAY MANAGER:**

→Storage management application to be used with NS20FC & NS40FC systems for configuring additional storage for Celerra and for other Hosts on the array

→Flare 24 version of the application is used with the new Celerra FC platforms [version 6.24.2.5.0] and is NAT capable  
Download from CLARIION Software site: <http://www.cs.isus.emc.com/csweb2/dgweb/software/index.asp#CX3%20-%20ARRAY%20BASED%20SOFTWARE>

Navisphere OffArray UI’s (Windows) 6.24.2.5.0

Install the “Windows\_UIs\_6.24.2.5.0.exe” on the Windows system

Installation places a “Navisphere Manager” shortcut on the Desktop

- a.) Launching “Navisphere Manager” brings up a Web Browser using the following URL, which is just the path to, C:\Program Files\EMC\ManagementUI\6.24.2.5.0\WebContent\start.html, followed by a Navisphere Connection screen
- b.) Select Options>Network Address Translation (NAT) Connection & enter the public IP addresses of SPA and SPB
- c.) Once connected, use the righthand pane to conduct traditional storage-management activities, such as configuring new Raid Groups, Binding Luns, adding Luns to Storage Groups, assigning proper HLU assignments when Celerra Luns are involved, etc.
- d.) Use the lefthand pane to access various Wizards for “Storage Management”, “Monitoring”, “Replication”, “Reporting”, & “Service”

**Note 1:** When managing/adding storage for Celerra NS20FC/NS40FC systems, the wizards are not particularly useful. The “Storage Management” wizard can Allocate storage by creating new RGs, new LUNs, and assigning to Servers, but since Celerra doesn’t run an agent, it cannot make use of this feature. Luns could be created for Celerra using the “Create LUN without Host assignment” option in the wizard, but the Assign wizard could not be used to assign to the Celerra Storagegroup for the reason already mentioned—logged as a bug and fixed with 5.6.39.x release.

**Note 2:** Beginning with Napa 9 and the Proxy ARP implementation, either Navisphere Offarray or Onarray Manager utilities are acceptable. Onarray Navisphere runs on the SPs and is more easily accessed since there is no separate application that needs to run on the client PC.

### **Using the Storage Management>Allocate wizard & Navisphere GUI to add Storage on the Celerra:**

- a.) Create new RAID Group (or use existing RG) using Allocate wizard
- b.) Select newly created RG and bind desired number of LUNs (since Celerra is not seen on the assignment list, select “Create LUN without Host assignment”)
- c.) In the Navisphere pane (righthand side) of GUI, highlight Celerra\_emcnas\_i0 Storage Group>Select LUNs
  - Available LUNs
    - Expand list on each SP and click the checkbox for the desired LUNs
  - Selected LUNs
    - Locate LUNs that are being added and assign a valid HLU number (>than 16) in the far right of the screen
- d.) Add the new LUNs to the Celerra database using either the GUI Rescan function or the CLI nas\_diskmark option  
Celerra Manager>Celerras>Storage>Systems>Rescan

# nas\_diskmark -m -a

### **Using the Navisphere GUI to add Storage on the Celerra:**

- a.) Create new RAID Group by rightclicking RAID Groups>Create RAID Group (select number of disks, RAID protection, etc.)
- b.) Rightclick new RAID Group>Bind LUN (create desired luns, SPs to use (auto), size of LUNs, ALU number, etc.)
- c.) Rightclick Storage Groups>Celerra\_emcnas\_i0>Select LUNs
  - Available LUNs
    - Expand list on each SP and click the checkbox for the desired LUNs
  - Selected LUNs
    - Locate LUNs that are being added and assign a valid HLU number (>than 16) in the far right of the screen
- d.) Add the new LUNs to the Celerra database using either the GUI Rescan function or the CLI nas\_diskmark option  
Celerra Manager>Celerras>Storage>Systems>Rescan

# nas\_diskmark -m -a

### **UPGRADING FLARE USING NAVISPHERE OFFARRAY MANAGER:**

- When in Non-Engineering mode and using Software Operations, a popup says that NST is the tool to use for SW updates
- In Engineering mode>Array>Software Operations>Software Installation Wizard>it is possible to perform NDU

### **SOFTWARE UPGRADES TO NS20/40 INTEGRATED SYSTEMS (NAS/FLARE):**

- Software upgrades to Napa 8 are allowed for Celerra—use CD-ROM setup script
  - Only thing to know is that it will change the configuration on certain models (NS40I, NS80I)
    - Adds CelerraService pkg to backend array
    - Permanently enables IP Forwarding
    - Starts IPTables NAT Service and adds aliases eth0:1 & eth2:1 to support new backend gateway SP-to-SP communication scheme (though doesn’t actually change the SP gateway IP on its own—advises User to do this when the SPs can be rebooted)
  - Flare upgrades are also allowed, but not via NST or Navisphere tools
  - Though not yet revised, the Flare NDU procedure from NAS ProcGen is adequate to perform the NDU using CLI from the Celerra
- Note:** Officially, you are not allowed to use the Navisphere GUI for NDU, but this does work if using Engineering Mode and the SIW
1. Open Navisphere, select Options>NAT and enter IP’s for both SP’s
  2. Log into array using security account and password
  3. Do ctrl + shift + F12 and password to enter Engineering Mode
  4. Perform NDU: **Navisphere Manager>Array>Properties>Software Operations>Software Installation Wizard**
- Note:** Either GUI or CLI methods requires that Statistics Logging be disabled on both SPs for the NDU process
5. CLI commandline example:  
**#/nas/sbin/navicli -h 192.168.1.200 ndu -install /home/nasadmin/flare/CX3-40-Bundle-03.24.040.5.014.pbu -delay 360**

## **CELERRA AND PROXY ARP IMPLEMENTATION**

→Beginning with Napa 9 release, Proxy ARP will automatically be setup during the CSA Pre-Configuration “Apply” screen phase for all NS20 & NS40 models—we will no longer be using the NAT implementation that shipped with Napa 8.

→SPs will be assigned a public IP address and will reboot during the CSA setup phase

→A status bar on the righthand side of the CSA screen will inform user of the progress made for the (12) steps that are part of the Proxy\_ARP upgrade process

### **CSA Example:**

“Setting up backend IP…

Backend IP Setup | Step 11/12: Waiting for CLARiiON software to start on SPB.”

### **Troubleshooting:**

**/sbin/arp -n -i eth3**

**/nasmed/sbin/clariion\_mgmt -start\_service -v** [Fixes ARP entries, adds Host routes, enables Proxy ARP on interfaces]

**Note:** This option is no longer functional in NAS 5.6

**/sbin/sysctl net.ipv4.conf.eth0.proxy\_arp net.ipv4.conf.eth2.proxy\_arp net.ipv4.conf.eth3.proxy\_arp** [Enables Proxy ARP svc]

## **MANUALLY SETTING PROXY ARP TO 1:**

```
# echo 1 >/proc/sys/net/ipv4/conf/eth0/proxy_arp  
# echo 1 >/proc/sys/net/ipv4/conf/eth2/proxy_arp  
# echo 1 >/proc/sys/net/ipv4/conf/eth3/proxy_arp
```

## **PROXY ARP CONFIGURATION ON 5.5.31.6—Dual Blade NS40FC:**

**# /nasmed/sbin/clariion\_mgmt -info**

Public IP address for SPA: 192.1.4.220

Public IP address for SPB: 192.1.4.221

Start on boot : yes

Current implementation : Proxy-ARP

Status : Started

done

**# cat /etc/sysconfig/static-routes** →These routes are created during Proxy ARP setup during CSA configuration

eth0 host 192.1.4.220 dev

eth2 host 192.1.4.221 dev

**Note:** The static-routes file is used for single blade systems to define a path between the CS, using eth2, and SPB, using an internal IP address of 128.221.253.102, but after Proxy ARP is setup, the public IP addresses are used as the routes to SPA & SPB.

**# cat static-routes**

eth2 host 128.221.253.201 dev

eth0 host 192.1.4.220 dev

eth2 host 192.1.4.221 dev

**# cat /etc/sysconfig/network-scripts/ifcfg-eth3** →Normal external interface configuration file after Proxy ARP setup

DEVICE=eth3

IPADDR=192.1.4.218

NETMASK=255.255.255.0

NETWORK=192.1.4.0

BROADCAST=192.1.4.255

ONBOOT=yes

### **Only Following CS Aliases maintained:**

**# cat ifcfg-eth0:1**

DEVICE=eth0:1

IPADDR=128.221.200.104

NETMASK=255.255.255.0

NETWORK=128.221.200.0

BROADCAST=128.221.200.255

ONBOOT=no

**# cat ifcfg-eth2:1**

DEVICE=eth2:1

IPADDR=128.221.201.104

NETMASK=255.255.255.0

NETWORK=128.221.201.0

BROADCAST=128.221.201.255

ONBOOT=no

# cat iptables →Only POSTROUTING entries maintained, and old internal structure kept

```
-A POSTROUTING -s 128.221.252.200 -d 128.221.201.201 -p tcp -j SNAT --to-source 128.221.252.200
-A POSTROUTING -s 128.221.253.201 -d 128.221.200.200 -p tcp -j SNAT --to-source 128.221.253.201
-A POSTROUTING -s 128.221.252.200 -d 128.221.201.201 -p udp -j SNAT --to-source 128.221.252.200
-A POSTROUTING -s 128.221.253.201 -d 128.221.200.200 -p udp -j SNAT --to-source 128.221.253.201
-A POSTROUTING -s 192.1.4.220 -d 0.0.0.0 -p tcp -j SNAT --to-source 192.1.4.220
-A POSTROUTING -s 192.1.4.220 -d 0.0.0.0 -p udp -j SNAT --to-source 192.1.4.220
-A POSTROUTING -s 192.1.4.220 -d 0.0.0.0 -p icmp -j SNAT --to-source 192.1.4.220
-A POSTROUTING -s 192.1.4.221 -d 0.0.0.0 -p tcp -j SNAT --to-source 192.1.4.221
-A POSTROUTING -s 192.1.4.221 -d 0.0.0.0 -p udp -j SNAT --to-source 192.1.4.221
-A POSTROUTING -s 192.1.4.221 -d 0.0.0.0 -p icmp -j SNAT --to-source 192.1.4.221
```

# /sbin/iptables -t nat -L

Chain PREROUTING (policy ACCEPT)

target prot opt source destination

Chain POSTROUTING (policy ACCEPT)

target prot opt source destination

```
SNAT  tcp -- 128.221.200.200  128.221.201.201  to:128.221.200.200
SNAT  tcp -- 128.221.201.201  128.221.200.200  to:128.221.201.201
SNAT  udp -- 128.221.200.200  128.221.201.201  to:128.221.200.200
SNAT  udp -- 128.221.201.201  128.221.200.200  to:128.221.201.201
SNAT  tcp -- A_APM00071600514  0.0.0.0          to:192.1.4.220
SNAT  udp -- A_APM00071600514  0.0.0.0          to:192.1.4.220
SNAT  icmp -- A_APM00071600514  0.0.0.0          to:192.1.4.220
SNAT  tcp -- B_APM00071600514  0.0.0.0          to:192.1.4.221
SNAT  udp -- B_APM00071600514  0.0.0.0          to:192.1.4.221
SNAT  icmp -- B_APM00071600514  0.0.0.0          to:192.1.4.221
```

# tail /etc/hosts

192.1.4.220 A\_APM00071600514 SPA # CLARiiON SP

192.1.4.221 B\_APM00071600514 SPB # CLARiiON SP

# cat /nas/site/sp\_info

192.1.4.220 A\_APM00071600514 SPA # CLARiiON SP

192.1.4.221 B\_APM00071600514 SPB # CLARiiON SP

# cat clarion\_mgmt.cfg

SPA\_PUBLIC\_IP=192.1.4.220

SPB\_PUBLIC\_IP=192.1.4.221

ONBOOT=yes

CURRENT\_IMPL=Proxy-ARP

# /nas/sbin/navicli -h 192.1.4.220 security -list

Username: nasadmin

Role: administrator

Scope: global

# ls -la /nas/site.clar\* →Security cache files updated during Proxy ARP configuration process, and validated by security -list

-rw-rw-r-- 1 nasadmin nasadmin 276 Apr 1 08:39 /nas/site.clar\_security

# /sbin/sysctl net.ipv4.conf.eth0.proxy\_arp

net.ipv4.conf.eth0.proxy\_arp = 1

# /sbin/sysctl net.ipv4.conf.eth2.proxy\_arp

net.ipv4.conf.eth2.proxy\_arp = 1

# /sbin/sysctl net.ipv4.conf.eth3.proxy\_arp

net.ipv4.conf.eth3.proxy\_arp = 1

# /nas/sbin/navicli -h 192.1.4.220 networkadmin -get

Storage Processor: SP A

Storage Processor Network Name: spa

Storage Processor IP Address: 192.1.4.220

Storage Processor Subnet Mask: 255.255.255.0

Storage Processor Gateway Address: 192.1.4.254

# /nas/sbin/navicli -h 192.1.4.221 networkadmin -get

Storage Processor: SP B

Storage Processor Network Name: spb

Storage Processor IP Address: 192.1.4.221

Storage Processor Subnet Mask: 255.255.255.0

Storage Processor Gateway Address: 192.1.4.254

# /sbin/arp -n |egrep "192.1.4.220|192.1.4.221"

|             |                           |      |
|-------------|---------------------------|------|
| 192.1.4.220 | ether 00:60:16:0F:4D:4D C | eth0 |
| 192.1.4.221 | ether 00:60:16:0F:4E:68 C | eth2 |
| 192.1.4.220 | * * MP                    | eth3 |
| 192.1.4.221 | * * MP                    | eth3 |

# cat /proc/sys/net/ipv4/ip\_forward

1

## **ADDING A WINDOWS HOST TO FC ENABLED ARRAY & IMPLEMENTING PROXY ARP:**

Powerlink>Celerra Tools>NS20 Integrated>under the Install section Step 5 provides link to procedure to add a Direct-connect or Fabric Switch-connected Windows Host to the FC-Enabled array

### **Host Requirements:**

Windows 2000/2003 with QLogic or Emulex Fibre Channel HBA (check eLab)

Navisphere\_Server\_Utility\_6.24.1.4.0.exe installed on Windows client, which uses a Registration Service Feature to register Host HBA's with the array

Install PowerPath on the Windows Server

### **Configuring Host:**

Connect to Ports 0, 1, 4, or 5 fibre on NS20FC array [Ports 0 & 1 Fibre NS40FC]

Create zone entries on Fabric Switch

Register the Windows Host with the array

Upgrade Celerra NAT to Proxy ARP

Connect to SPs using either Navisphere Manager or NST

Allocate Storage & create LUNs for the new Host

Verify PowerPath sees all paths & write disk signature to LUNs

### **What is Proxy ARP?**

→Proxy ARP is a Linux feature that bridges broadcast traffic using host routing for defined interfaces

→Celerra Control Station becomes a Layer 2 Bridge, and SPs are converted to Public IP Addresses, yet still connect to private network

→SNAT (Static Network Address Translation) works by using a one-to-one mapping of a public IP address to a masqueraded private IP address (on the Server), whereas Proxy ARP allows the public IP address to be routed to the protected side of the Router (i.e., Control Station)

→In practice, this means that the SPs will be configured with public IP address & Gateway addresses and the private IP addressing scheme is dropped

→Please note that the Proxy ARP conversion requires that SPs be rebooted!

→Proxy ARP is the default implementation model for all Celerra Integrated installs of NS20/20FC, NS40/40FC, NX4/NX4FC

→Proxy ARP is needed because of Clariion issues related to multiple Domain management and Layered applications, in which NAT for the NST & Navisphere Mgr tools may not otherwise always work correctly—Proxy ARP solves this

### **Clariion mgmt Script NAS Version Rules:**

#### **Rule 1:**

For NAS 5.5.30.5 systems, the clariion\_mgmt.zip download is no longer available from Powerlink—if running 5.5.30 with NAT and needing to convert to Proxy ARP, the recommendation would be to upgrade to the latest NAS code, then convert using clariion\_mgmt

#### **Rule 2:**

For NAS 5.5.31.6 systems, or higher, use the clariion\_mgmt script that is resident on the system with the respective NAS version

**Note 1:** Last-minute changes were made to the original clariion\_mgmt script in order to convert a Celerra from NAT to ProxyARP, and is available on Powerlink. Download & unzip the following file, place the file onto portable media, mount the media to the CS, and copy the clariion\_mgmt.exe script to /nasmcd/sbin and /nas/sbin, overwriting the existing clariion\_mgmt script.

[http://powerlink.emc.com/km/live1/en\\_US/Offering\\_Technical\\_Documentation/clariion\\_mgmt.zip](http://powerlink.emc.com/km/live1/en_US/Offering_Technical_Documentation/clariion_mgmt.zip)

**Note 2:** The above instruction only applies when upgrading to Proxy ARP for NAS 5.5.30 systems! Beginning with Napa 9, all new NS20 or NS40 installs will implement Proxy ARP during CSA configuration.

### **Basic implementation of Proxy ARP on Celerra involves the following:**

- I. Enabling proxy arp in the kernel for eth0, eth2, eth3: proxy\_arp=1
- II. Adding Routes on CS to SPs using static-routes file
- III. Changing IP & gateway addresses on SPs to public
- IV. Updating /etc/hosts with new SP addresses
- V. Upgrading CS NAT to Proxy ARP implementation [clariion\_mgmt.cfg]
- VI. Updating Clariion Domain Security cache files on Celerra
- VII. Automatically fails Celerra luns back after SPs have rebooted

### **Things to know before running clariion mgmt script:**

- Do not upgrade from NAT to Proxy ARP if running 5.5.30 (Upgrade latest 5.5 first, then convert to Proxy ARP)
- Be aware that script will require that both SPs be rebooted in order to change the network & gateway IP addresses
- Be aware that the script reapplies the Clarion array security settings on the Celerra, so the User must know the username & password [a default FC Enabled Celerra will have username of “nasadmin” and password “nasadmin”, though customer may have changed the password—the username remains the same]
- Be aware that there are hidden options in clariion\_mgmt to assist in upgrading in the face of problems, as well as to revert to the original NAT configuration
- If a problem cannot be surmounted during the conversion to Proxy ARP, open a case with EMC Support

### **Upgrading from NAT to Proxy ARP with NAS 5.5.30\*:**

\*The following procedure is no longer recommended since the clariion\_mgmt.zip file is no longer maintained on Powerlink. The recommendation would be to upgrade to latest 5.5 code, then convert to Proxy ARP by first doing clariion\_mgmt –stop to cleanup NAT, then clariion\_mgmt –start –use\_proxy\_arp to convert.

1. Verify current NAT configuration:

```
# /nasmcd/sbin/clariion_mgmt -info
Public IP address for SPA: 10.241.168.151
Public IP address for SPB: 10.241.168.152
Start on boot      : yes
Current implementation : NAT
Status           : Started
```

2. Download latest clariion\_mgmt script from Powerlink, unzip, and copy to the following locations on the Control Station:

[http://powerlink.emc.com/km/live1/en\\_US/Offering\\_Technical/Technical\\_Documentation/clariion\\_mgmt.zip](http://powerlink.emc.com/km/live1/en_US/Offering_Technical/Technical_Documentation/clariion_mgmt.zip)

```
# cp clariion_mgmt /nasmcd/sbin
# cp clariion_mgmt /nas/sbin
```

**Note:** If the system is running NAT and has been upgraded to 5.5.31 or higher, you should not use the Powerlink “clariion\_mgmt.zip” file—the correct clariion\_mgmt script is already in /nasmcd/sbin.

3. Execute the Proxy Upgrade:

```
# /nasmcd/sbin/clariion_mgmt -upgrade_to_proxy_arp
```

4. Verify connectivity with SPs after upgrade completes:

```
# /nas/sbin/navicli -h 10.241.168.151 getagent [repeat for SPB]
```

5. Verify results of Proxy ARP Upgrade:

```
# /sbin/sysctl net.ipv4.conf.eth0.proxy_arp
```

net.ipv4.conf.eth0.proxy\_arp = 1

```
# /sbin/sysctl net.ipv4.conf.eth2.proxy_arp
```

net.ipv4.conf.eth2.proxy\_arp = 1

```
# /sbin/sysctl net.ipv4.conf.eth3.proxy_arp
```

net.ipv4.conf.eth3.proxy\_arp = 1

```
# /nasmcd/sbin/clariion_mgmt -info
```

Public IP address for SPA: 10.241.168.151

Public IP address for SPB: 10.241.168.152

Start on boot : yes

Current implementation : Proxy-ARP

Status : Started

```
# cat static-routes
```

eth0 host 10.241.168.151 dev

eth2 host 10.241.168.152 dev

```
/etc/sysconfig/network-scripts
```

```
# cat route-eth0
```

scope host 10.241.168.151 dev eth0

```
# cat route-eth2
```

scope host 10.241.168.152 dev eth2

```
# /sbin/route -n
```

Kernel IP routing table

| Destination    | Gateway | Genmask         | Flags | Metric | Ref | Use | Iface |
|----------------|---------|-----------------|-------|--------|-----|-----|-------|
| 10.241.168.151 | 0.0.0.0 | 255.255.255.255 | UH    | 0      | 0   | 0   | eth0  |
| 10.241.168.152 | 0.0.0.0 | 255.255.255.255 | UH    | 0      | 0   | 0   | eth2  |

```
# cat /etc/clariion_mgmt.cfg
```

SPA\_PUBLIC\_IP=10.241.168.151

SPB\_PUBLIC\_IP=10.241.168.152

ONBOOT=yes

CURRENT\_IMPL=Proxy-ARP

```
# /sbin/arp -n |egrep "151|152"
```

|                |       |                   |    |      |
|----------------|-------|-------------------|----|------|
| 10.241.168.152 | ether | 00:60:16:0F:4E:68 | C  | eth2 |
| 10.241.168.151 | ether | 00:60:16:0F:4D:4D | C  | eth0 |
| 10.241.168.152 | *     | *                 | MP | eth3 |
| 10.241.168.151 | *     | *                 | MP | eth3 |

```
# cat /etc/sysconfig/iptables
```

```
-A POSTROUTING -s 10.241.168.151 -d 0.0.0.0 -p tcp -j SNAT --to-source 10.241.168.151
-A POSTROUTING -s 10.241.168.151 -d 0.0.0.0 -p udp -j SNAT --to-source 10.241.168.151
-A POSTROUTING -s 10.241.168.151 -d 0.0.0.0 -p icmp -j SNAT --to-source 10.241.168.151
-A POSTROUTING -s 10.241.168.152 -d 0.0.0.0 -p tcp -j SNAT --to-source 10.241.168.152
-A POSTROUTING -s 10.241.168.152 -d 0.0.0.0 -p udp -j SNAT --to-source 10.241.168.152
-A POSTROUTING -s 10.241.168.152 -d 0.0.0.0 -p icmp -j SNAT --to-source 10.241.168.152
```

```
# /sbin/iptables -t nat -L -n
```

Chain POSTROUTING (policy ACCEPT)

| target | prot | opt | source         | destination |
|--------|------|-----|----------------|-------------|
| SNAT   | tcp  | --  | 10.241.168.151 | 0.0.0.0     |
| SNAT   | udp  | --  | 10.241.168.151 | 0.0.0.0     |
| SNAT   | icmp | --  | 10.241.168.151 | 0.0.0.0     |
| SNAT   | tcp  | --  | 10.241.168.152 | 0.0.0.0     |
| SNAT   | udp  | --  | 10.241.168.152 | 0.0.0.0     |
| SNAT   | icmp | --  | 10.241.168.152 | 0.0.0.0     |

```
# tail /etc/hosts
```

```
10.241.168.151 A_APM00071600514 SPA # CLARiiON SP
10.241.168.152 B_APM00071600514 SPB # CLARiiON SP
```

```
# cat /proc/sys/net/ipv4/ip_forward
```

```
1
```

```
# /sbin/arp -n |egrep "192.1.4.251|192.1.4.252" →Output of Proxy ARP config with two Blade system
```

|             |       |                   |    |                                                                  |
|-------------|-------|-------------------|----|------------------------------------------------------------------|
| 192.1.4.251 | ether | 00:60:16:19:BB:5E | C  | eth0 →SPA IP & MAC Address tied to CS eth0 as complete ARP entry |
| 192.1.4.252 | ether | 00:60:16:19:A3:91 | C  | eth2 →SPB IP & MAC Address tied to CS eth2 as complete ARP entry |
| 192.1.4.252 | *     | *                 | MP | eth3 →SPB permanent & published ARP entry                        |
| 192.1.4.251 | *     | *                 | MP | eth3 →SPA permanent & published ARP entry                        |

### Abridged output from /nasmcd/sbin/clariion\_mgmt -upgrade to proxy arp:

Removing network alias eth3:1...done

Removing network alias eth3:2...done

Adding rules to allow outbound traffic from SPA

Changing SPA IP from 192.168.1.200 to 10.241.168.151 (subnetmask 255.255.255.0, gateway 10.241.168.128)

Waiting for SPA to go down...done

Waiting for SPA to come back up.....done

Adding host specific route for SPB

Adding rules to allow outbound traffic from SPB

Adding ARP entry for SPB

Updating /etc/hosts entry for SPB

Changing SPB IP from 192.168.2.201 to 10.241.168.152 (subnetmask 255.255.255.0, gateway 10.241.168.128)

Waiting for SPB to go down...done

Waiting for SPB to come back up.....done

Updating NAS database with new CLARiiON IP addresses

Updating NAS cache of CLARiiON password

Enter the Global CLARiiON account information

Username: nasadmin

Password: \*\*\*\*\*

Retype your response to validate

Password: \*\*\*\*\*

Setting security information for APM00071600514

Done

### UPGRADING TO PROXY ARP WITH NAS 5.5.31:

**Note:** Original documentation says that to convert from NAT to Proxy ARP only required the –upgrade\_to\_proxy\_arp switch for clariion\_mgmt. This method does not work and must be modified. See emc175337.

#### Workaround:

##### 1. Breakdown NAT configuration:

```
# /nasmcd/sbin/clariion_mgmt -stop (removes eth3:1 & eth3:2 aliases and IPTables entries)
```

Checking if running as root...yes  
Checking if model is supported...yes  
Checking for integrated system...yes  
Checking if interface eth3 is configured...yes  
Checking if interface eth3:1 is configured...yes  
Checking if interface eth3:2 is configured...yes  
Checking if SP (128.221.110.200) is up...yes  
Checking if SP (128.221.111.201) is up...yes  
Removing network alias eth3:1...done  
Removing network alias eth3:2...done  
done

**# /nasmcd/sbin/clariion\_mgmt -info**

Error 12: Not configured

**2. Convert directly to Proxy ARP from an Unconfigured state:**

**# /nasmcd/sbin/clariion\_mgmt -start -spa\_ip 192.1.4.220 -spb\_ip 192.1.4.221 -use\_proxy\_arp**

Checking if running as root...yes  
Checking if model is supported...yes  
Checking for integrated system...yes  
Checking if interface eth3 is configured...yes  
Checking if interface eth3:1 is configured...no  
Checking if interface eth3:2 is configured...no  
Checking if SP (128.221.110.200) is up...yes  
Checking if SP (128.221.111.201) is up...yes  
Checking if a gateway is setup for eth3...yes

Step 1/12: Setting up Proxy ARP for SPA on Control Station

Adding host specific route for SPA  
Adding rules to allow outbound traffic from SPA

Adding ARP entry for SPA

Updating /etc/hosts entry for SPA

Step 2/12: Changing SPA IP address.

Changing SPA IP from 128.221.110.200 to 192.1.4.220 (subnetmask 255.255.255.0, gateway 192.1.4.254)

Step 3/12: Waiting for SPA to reboot.

Waiting for SPA to go down...done

Step 4/12: Waiting for SPA to boot up.

Waiting for SPA to come back up...done

Step 5/12: Waiting for CLARiiON software to start on SPA.

Waiting for CLARiiON software to start on SPA.....done

Step 6/12: Updating NAS database with SPA IP address

Updating SYMAPI database with new CLARiiON IP addresses...done

Step 7/12: Setting up Proxy ARP for SPB on Control Station

Adding host specific route for SPB

Adding rules to allow outbound traffic from SPB

Adding ARP entry for SPB

Updating /etc/hosts entry for SPB

Step 8/12: Changing SPB IP address.

Changing SPB IP from 128.221.111.201 to 192.1.4.221 (subnetmask 255.255.255.0, gateway 192.1.4.254)

Step 9/12: Waiting for SPB to reboot.

Waiting for SPB to go down.....done

Step 10/12: Waiting for SPB to boot up.

Waiting for SPB to come back up...done

Step 11/12: Waiting for CLARiiON software to start on SPB.

Waiting for CLARiiON software to start on SPB.....done

**MANUALLY RECOVERING OR SETTING UP PROXY ARP WITH NAS 5.6:**

**Note:** Both SP's could either be on the Celerra internal network (128.221.252.200/128.221.253.201), or on a Public IP address, but the following example is shown with both SPs beginning on the Celerra internal networks. The following example presumes that the Control Station Hostname and external IP address are already configured.

**2) Edit the /etc/hosts and /nas/site/sp\_info files so that both SPA & SPB reflect their public IP addresses:**

**# vi /etc/hosts # vi /nas/site/sp\_info**

**3) Enable Proxy ARP on the eth0, eth2, and eth3 Control Station interfaces:**

```
# /sbin/sysctl net.ipv4.conf.eth0.proxy_arp
```

net.ipv4.conf.eth0.proxy\_arp = 0 → 0 indicates that Proxy ARP is disabled

```
# /sbin/sysctl net.ipv4.conf.eth0.proxy_arp=1
```

```
# /sbin/sysctl net.ipv4.conf.eth2.proxy_arp=1
```

```
# /sbin/sysctl net.ipv4.conf.eth3.proxy_arp=1
```

**4) Build the IP Tables entries for the SP Public IP addresses for TCP, UDP, & ICMP:**

```
# /sbin/iptables -t nat -A POSTROUTING -p tcp -s 10.241.168.179 -d 0.0.0.0 -j SNAT --to-source 10.241.168.179
# /sbin/iptables -t nat -A POSTROUTING -p udp -s 10.241.168.179 -d 0.0.0.0 -j SNAT --to-source 10.241.168.179
# /sbin/iptables -t nat -A POSTROUTING -p icmp -s 10.241.168.179 -d 0.0.0.0 -j SNAT --to-source 10.241.168.179
# /sbin/iptables -t nat -A POSTROUTING -p tcp -s 10.241.168.180 -d 0.0.0.0 -j SNAT --to-source 10.241.168.180
# /sbin/iptables -t nat -A POSTROUTING -p udp -s 10.241.168.180 -d 0.0.0.0 -j SNAT --to-source 10.241.168.180
# /sbin/iptables -t nat -A POSTROUTING -p icmp -s 10.241.168.180 -d 0.0.0.0 -j SNAT --to-source 10.241.168.180
```

**5) Save the IP Tables entries from memory to the IP Tables file:**

```
# /sbin/iptables-save >/etc/sysconfig/iptables
```

**6) Verify IP Tables entries:**

```
# cat /etc/sysconfig/iptables
```

-----abridged-----

```
:POSTROUTING ACCEPT [2984081:210282993]
:OUTPUT ACCEPT [0:0]
-A POSTROUTING -s 128.221.252.200 -d 128.221.253.201 -p tcp -j SNAT --to-source 128.221.252.200
-A POSTROUTING -s 128.221.253.201 -d 128.221.252.200 -p tcp -j SNAT --to-source 128.221.253.201
-A POSTROUTING -s 128.221.252.200 -d 128.221.253.201 -p udp -j SNAT --to-source 128.221.252.200
-A POSTROUTING -s 128.221.253.201 -d 128.221.252.200 -p udp -j SNAT --to-source 128.221.253.201
-A POSTROUTING -s 10.241.168.179 -d 0.0.0.0 -p tcp -j SNAT --to-source 10.241.168.179
-A POSTROUTING -s 10.241.168.179 -d 0.0.0.0 -p udp -j SNAT --to-source 10.241.168.179
-A POSTROUTING -s 10.241.168.179 -d 0.0.0.0 -p icmp -j SNAT --to-source 10.241.168.179
-A POSTROUTING -s 10.241.168.180 -d 0.0.0.0 -p tcp -j SNAT --to-source 10.241.168.180
-A POSTROUTING -s 10.241.168.180 -d 0.0.0.0 -p udp -j SNAT --to-source 10.241.168.180
-A POSTROUTING -s 10.241.168.180 -d 0.0.0.0 -p icmp -j SNAT --to-source 10.241.168.180
```

**7) Change the IP addresses on the SPs using Secure Navicli:**

```
# /nas/sbin/navisecli -h 128.221.252.200 -user nasadmin -password nasadmin -scope 0 networkadmin -set
-address 10.241.168.179 -subnetmask 255.255.255.0 -gateway 10.241.168.128
```

**Note:** Wait approximately 5 minutes to ensure that SP has completely rebooted. You may need to use ctrl + c to return to the command prompt after issuing the IP address change command.

```
# /nas/sbin/navisecli -h 128.221.253.201 -user nasadmin -password nasadmin -scope 0 networkadmin -set
-address 10.241.168.180 -subnetmask 255.255.255.0 -gateway 10.241.168.128
```

**Note:** SPs will need to reboot prior to Flare 29. Check and restore luns after both SPs have been rebooted.

**8) Add the route entries for SPA and SPB to the Control Station:**

```
# /sbin/route add 10.241.168.179 eth0 -->For SPA
```

```
# /sbin/route add 10.241.168.180 eth2 -->For SPB
```

**Note:** After the routes have been added, you should be able to ping the SPs by IP address or name. If you cannot, verify the route entries exist for eth0 and eth2 for the SPs. Verify that Arp Table entries also exist. Re-add the routes if necessary, and stop and restart the interface aliases if required to refresh the arp entries.

```
# /sbin/route -n |egrep "eth0|eth2"
```

```
10.241.168.179 0.0.0.0      255.255.255.255 UH  0    0      0 eth2
```

```
10.241.168.180 0.0.0.0      255.255.255.255 UH  0    0      0 eth0
```

```
# /sbin/arp -n
```

```
# /sbin/ifdown eth0:1
```

```
# /sbin/ifup eth0:1
```

```
# /sbin/ifdown eth2:1
```

```
# /sbin/ifup eth2:1
```

**9) Verify that you can ping the SPs by IP address and issue navicli commands:**

```
# ping spa
```

```
# ping spb
```

```
# /nas/sbin/navisecli -h 10.241.168.179 -user nasadmin -password nasadmin -scope 0 getagent
```

# /nas/sbin/naviseclli -h 10.241.168.180 -user nasadmin -password nasadmin -scope 0 getagent

Note: Check for trespassed luns using getlun -trespass, and trespass respective luns back to their owner SPs

**10) At this point, the security credentials and symapi database needs to be updated with the new SP IP addresses:**

# /tftpboot/bin/setup\_clariion\_security 10.241.168.179 10.241.168.180 nasadmin nasadmin –initialize

Note: Syntax calls for SPA IP, then SPB, followed by <admin\_account> and <admin\_password>. There is no output returned after issuing this command

**11) Perform a storage health check:**

# nas\_storage -c -a

Discovering storage (may take several minutes)

done

**12) Use clariion\_mgmt to start the Proxy ARP service:**

# /nasmcd/sbin/clariion\_mgmt –start –spa\_ip 10.241.168.179 –spb\_ip 10.241.168.180 –skip\_rules

**13) Verify the Proxy Arp service:**

# /nasmcd/sbin/clariion\_mgmt –info

Public IP address for SPA: 10.241.168.179

Public IP address for SPB: 10.241.168.180

Start on boot : yes

Current implementation : Proxy-ARP

Status : Started

## **TROUBLESHOOTING PROXY ARP SETUP/CONVERSIONS:**

**I. Proxy ARP will not upgrade if NAT is not completely setup, or if NAT is corrupt**

Use following to upgrade directly to Proxy ARP without a prior NAT config:

# /nasmcd/sbin/clariion\_mgmt –info

Error 12: Not configured

# /nasmcd/sbin/clariion\_mgmt –start –spa\_ip 10.241.168.151 –spb\_ip 10.241.168.152 –use\_proxy\_arp

Use following to fix NAT, then rerun –upgrade\_to\_proxy\_arp command:

# /nasmcd/sbin/clariion\_mgmt –start\_service

# /nasmcd/sbin/clariion\_mgmt –info →Output from 5.5.30 for NAT configuration

Public alias IP address for SPA: 192.1.4.214

Public alias IP address for SPB: 192.1.4.215

Start on boot : yes

Status : Started

Done

# /nasmcd/sbin/clariion\_mgmt –info →5.5.31 adds “Current implementation” field to output

Public IP address for SPA: 10.241.168.151

Public IP address for SPB: 10.241.168.152

Start on boot : yes

Current implementation : NAT

Status : Started

done

# /nasmcd/sbin/clariion\_mgmt –upgrade\_to\_proxy\_arp

**II. Proxy ARP Upgrades has one SP on public but other on private network**

Stop NAT and then run conversion to Proxy ARP:

# /nasmcd/sbin/clariion\_mgmt –stop –skip\_rules

# /nasmcd/sbin/clariion\_mgmt –info

Error 12: Not configured

# /nasmcd/sbin/clariion\_mgmt –start –spa\_ip 10.241.168.151 –spb\_ip 10.241.168.152 –use\_proxy\_arp

Or, try reverting back to NAT and starting again:

# /nasmcd/sbin/clariion\_mgmt –revert\_to\_nat

# /nasmcd/sbin/clariion\_mgmt –upgrade\_to\_proxy\_arp

**III. Run following if arp or route entries are not seen for eth0, eth2, and eth3, or if Proxy ARP is not enabled on the interfaces**

# /nasmcd/sbin/clariion\_mgmt –start\_service | -v

Note: The above option is no longer valid with NAS 5.6.

**III. Inspect the clariion\_mgmt Log to determine issue when failures occur:**

# ls -la /var/log/clariion\_mgmt.log

-rw-r--r-- 1 root root 246144 Mar 16 21:06 /var/log/clariion\_mgmt.log

→All clariion\_mgmt actions are recorded in this log

## **SUPPORT GUIDANCE—5.5.30 & 5.5.31:**

### **PUTTING SYSTEM BACK TO UNCONFIGURED STATE (i.e., no NAT or Proxy-ARP configuration)**

**Note:** When a system is in a misconfigured state, the recommended action is to revert back to a completely unconfigured state, as shown below, and then perform the Proxy ARP setup using one of the applicable scenarios outlined in Step 9.

**Step 1. Verify whether System is reporting a NAT, Proxy ARP, Error: Not configured state, or some other condition:**

**# /nasmcd/sbin/clariion\_mgmt -info**

Error 12: Not configured -->This output shows neither NAT or Proxy\_ARP configured and represents the correct state for an "unconfigured" system

#### **Recommended Actions:**

a.) If any output other than "Error 12: Not configured" is displayed, return the system to an unconfigured state as outlined in the following steps.

**Step 2. Verify the following files and edit if required:**

**/etc/hosts /nas/site/sp\_info /etc/sysconfig/iptables** (iptables used only with NAT--editing not required)

**Note:** SPs should have internal IP addresses for an unconfigured or NAT state

SPA: 192.168.1.200 or 128.221.252.200

SPB: 192.168.2.201 or 128.221.253.201

#### **Recommended Actions:**

a.) Edit the /etc/host and /nas/site/sp\_info files to the default internal IP addresses for both SPs.

**Step 3. Verify actual IP addresses on SPs using ping or navicli networkadmin -get, etc.**

#### **Recommended Actions:**

a.) Use "navicli networkadmin -get" to verify SP IP addresses, or use "ping spa" or "ping spb" from the Control Station  
b.) If required, restore SPs to the default internal IP & gateway address scheme using navicli -h <spa or spb> networkadmin -set -address 192.168.x.x -gateway 192.168.x.x. If the SPs cannot be reached via navicli from the Control Station, connect directly to the SP's bottom LAN port (service port), configure a 128.221.1.254 IP address on the Client PC, open a browser and enter either 128.221.1.251/setup if connected to SPA's LAN port, or 128.221.1.250/setup if connected to SPB's LAN port. Use the /setup program to enter the internal IP address and Gateway on each SP.

b.) Wait until the SP reboots and is completely up before repeating step b. to set the internal IP & gateway address on the peer SP.

c.) Use "navicli getlun -trespass" to check for trespassed luns and use "trespass lun x" to restore the lun to the proper SP Owner

**Step 4. Check for the existence of /nas/site/clariion\_mgmt.cfg & /etc/clariion\_mgmt.cfg files**

#### **Recommended Actions:**

a.) Delete the clariion\_mgmt.cfg files if they exist

**Note:** The clariiom\_mgmt.cfg file should exist only when NAT or Proxy-ARP are configured on the Celerra. The SP IP addresses in this file always represent the Public IP addresses set when using the CSA wizard to configure the system (true for either the 5.5.30 NAT implementation or 5.5.31 Proxy-ARP implementation).

**Step 5. Check for existence of interface alias files on the Control Station /etc/sysconfig/network-scripts/ifcfg-eth3:1 & eth3:2**

#### **Recommended Actions:**

a.) Delete the alias files eth3:1 and eth3:2 if they exist

b.) Reboot the Control Station to clear any memory remnants that may confuse the state of the Celerra

**Note:** The eth3:1 & eth3:2 alias files should only exist for a NAT configuration. These files are deleted when the system is converted to Proxy-ARP and do not exist at all in an Unconfigured state. NAT is only configured by default with NAS 5.5.30 installs.

Beginning with NAS 5.5.31, the CSA wizard configures the system for Proxy-ARP and does not use NAT.

**Step 6. Verify state of system after Control Station Reboot:**

#### **Recommended Actions:**

a.) Verify SP addresses in /etc/hosts & /nas/site/sp\_info are now reflecting the default internal IP addresses

b.) Verify that eth3:1 & eth3:2 aliases are gone and not in memory (/sbin/ifconfig)

c.) Verify that clariion\_mgmt.cfg files are NOT present

d.) Verify that system is in an unconfigured state:

**# /nasmcd/sbin/clariion\_mgmt -info**

Error 12: Not configured

**Note:** If system was properly torn down to an unconfigured state, command should show the above output

**Step 7. Verify SYMAPI db Integrity:**

#### **Recommended Actions:**

a.) Run **# nas\_storage -check -all**

**Note:** Should return without any errors

b.) If errors are seen, use nas\_storage -modify id=x -security to renew local security cache files on Celerra, or use nas\_storage -modify id=x -network -spa xxx -spb xxx to update the SYMAPI database with the correct IP address for the SPs if required, or an error is returned from nas\_storage -check -all that seems to indicate an SP IP address discrepancy in the database.

**Step 8. Verify Clariion Security Cache on Celerra:**

**Recommended Actions:**

a.) Use the following to check for the integrity of the Clariion security cache files, and then update the local cache files if required:

**/nas/sbin/navicli -h 192.168.1.200 security -list**

Username: nasadmin

Role: administrator

Scope: global

**Note:** If the command does not return the above output, the security cache files are corrupt and need to be refreshed

b.) Refresh Celerra's Clariion security cache files (/nas/site/.clar\_security) by entering the proper Clariion Username & Password at the appropriate prompts:

**# nas\_storage -modify id=x -security**

**Step 9. Proceed with NAT or Proxy-ARP reconfiguration by selecting one of the following scenarios:**

**Setting up NAT from unconfigured state (NAS 5.5.30 only):**

**/nasmcd/sbin/clariion\_mgmt -start -spa\_ip 192.1.4.214 -spb\_ip 192.1.4.215**

**Setting up Proxy\_ARP from unconfigured state (NAS 5.5.31 only):**

**/nasmcd/sbin/clariion\_mgmt -start -spa\_ip 192.1.4.214 -spb\_ip 192.1.4.215 -use\_proxy\_arp**

**Converting from NAT to Proxy-ARP (NAS 5.5.30 only):**

Download the clariion\_mgmt.zip script from Powerlink, unzip, and upload the new clariion\_mgmt script to the Control Station, and overwrite the existing clariion\_mgmt script in /nasmcd/sbin and /nas/sbin with the uploaded version.

[http://powerlink.emc.com/km/live1/en\\_US/Offering\\_Technical/Technical\\_Documentation/clariion\\_mgmt.zip](http://powerlink.emc.com/km/live1/en_US/Offering_Technical/Technical_Documentation/clariion_mgmt.zip)

**/nasmcd/sbin/clariion\_mgmt -start -spa\_ip 192.1.4.214 -spb\_ip 192.1.4.215 -upgrade\_to\_proxy\_arp**

**Converting from NAT to Proxy-ARP (NAS 5.5.31 only):**

**/nasmcd/sbin/clariion\_mgmt -stop** [Removes existing NAT or Proxy ARP configuration]

**/nasmcd/sbin/clariion\_mgmt -start -spa\_ip 192.1.4.214 -spb\_ip 192.1.4.215 -use\_proxy\_arp**

**Converting directly to Proxy-ARP (NAS 5.5.31 only):**

**/nasmcd/sbin/clariion\_mgmt -start -spa\_ip 192.1.4.251 -spb\_ip 192.1.4.252 -use\_proxy\_arp**

**Differences between NAT vs. Proxy-ARP implementations on the Celerra and array:**

1. For NAS 5.5.30 installations, NAT is setup by default on all NS20/20FC/NS40/40FC systems during CSA configuration
2. For NAS 5.5.30 systems, NAT to Proxy-ARP conversions are allowed only on the NS20FC & NS40FC platforms
3. For NAS 5.5.31 and higher, all NS20 & NS40 systems can be converted to Proxy-ARP implementations, which is the new standard configuration
4. For NAS 5.5.31 and higher installations, Proxy-ARP is setup by default on all NS20/20FC/NS40/40FC systems during CSA configuration
5. During NAS 5.5.30 CSA configuration, NAT creates two aliases on the Control Station, which are used as Routing Table entries by the Control Station that represent the Publicly assigned IP addresses for SPA (eth3:1) and SPB (eth3:2)
6. During NAS 5.5.30 CSA configuration, NAT creates the proper entries in /etc/sysconfig/iptables that allow for the communication by external Hosts from the Public network, to the backend SPs on the private internal network (with NAT, SPs retain their default internal IP addresses), with the Control Station acting as a Router
7. For Proxy-ARP implementations, eth3 aliases and NAT are not used, though NAT entries are purposely retained in the /etc/sysconfig/iptables file if the system was converted from NAT
8. For NAT implementations, the /etc/hosts & /nas/site/sp\_info files should contain Internal IP addresses for both SPs
9. For Proxy-ARP implementations, the /etc/hosts & /nas/site/sp\_info files should contain the actual Public IP addresses of the SPs
10. For NAT implementations, the SPs use the default Celerra Internal addresses [192.168.1.200/192.168.2.201 or 128.221.252.200/128.221.253.201]
11. For Proxy-ARP implementations, the SPs are assigned a true public IP address & gateway (SPs no longer using internal IP addresses or gateways)
12. For either NAT or Proxy-ARP implementations, the /nas/site/clariion\_mgmt.cfg file will exist, and the IP addresses in this file always represent the Publicly assigned IP addresses for the SPs (whether NAT alias IPs or true Public IPs for Proxy ARP), and is created during the CSA configuration process
13. For Proxy-ARP implementations, the Control Station serves as a bridge between the Public network and the private Celerra network, eventhough the SPs themselves have true Public IP address assignments (network packets bound for SPA or SPB must still pass through the Control Station whenever Users connect to the SPs using NST or Navisphere)
14. For NAT implementations, external Clients must use the Offarray Navisphere or NST with the NAT option in order to properly communicate to the backends.
15. For Proxy ARP implementations, external Clients can use the Onarray or Offarray Navisphere and NST without specifying the NAT option during the connection process.

## Unsuccessful Proxy ARP conversion on NAS 5.5.31:

### Summary of System State:

--No navicli communication with SPA  
--SPB still on internal IP address and can be communicated with  
--The eth3:1 & eth3:2 aliases were never deleted, and all other files reflect a Proxy-ARP implementation

### 1. Attempted to convert a 5.5.31 system from NAT to Proxy ARP using an Engineering workaround:

#### **# /nasmd/sbin/clariion\_mgmt -upgrade\_to\_proxy\_arp -skip\_rules**

**Note:** This left the system in a NAT state according to clariion\_mgmt -info, but updated most files for Proxy ARP, left SPA configured with the external IP address, and left SPB on the default internal IP address

### 2. The recovery steps involved examining the following files and taking appropriate actions:

#### **# /nasmd/sbin/clariion\_mgmt -info** -->Output should show "Proxy-ARP" and not "NAT"

Public IP address for SPA: 192.1.4.251

Public IP address for SPB: 192.1.4.252

Start on boot : yes

Current implementation : NAT

Status : Started

Done

#### **# cat /nas/site/sp\_info** -->File was updated for Proxy-ARP with public IP addresses for SPs

192.1.4.251 A\_APM00073801838 SPA # CLARiiON SP

192.1.4.252 B\_APM00073801838 SPB # CLARiiON SP

#### # tail /etc/hosts -->File was updated for Proxy-ARP with public IP addresses for SPs

192.1.4.251 A\_APM00073801838 SPA # CLARiiON SP

192.1.4.252 B\_APM00073801838 SPB # CLARiiON SP

/etc/sysconfig/network-scripts/ifcfg-eth3:1 & eth3:2 files were present and in Control Station memory (used only for NAT)

#### **/sbin/ifconfig** -->eth3:1 & eth3:2 aliases still exist and should have been removed during the Proxy-ARP conversion

eth3:1 Link encap:Ethernet HWaddr 00:1B:21:05:C6:68

inet addr:192.1.4.251 Bcast:192.1.4.255 Mask:255.255.255.0

eth3:2 Link encap:Ethernet HWaddr 00:1B:21:05:C6:68

inet addr:192.1.4.252 Bcast:192.1.4.255 Mask:255.255.255.0

#### **# cat /nas/site/clariion\_mgmt.cfg**

SPA\_PUBLIC\_IP=192.1.4.251

SPB\_PUBLIC\_IP=192.1.4.252

ONBOOT=yes

**Note:** File should show the Public IP addresses assigned to the SPs after CSA configuration for either NAT or Proxy-ARP. This file should not exist when the system is in an unconfigured state.

### 3. Overview of Recovery steps:

I. a.) Used clariion\_mgmt -stop -skip\_rules, which managed to restore communications with SPA [still on external IP though] and removed eth3:1/eth3:2 aliases

b.) Used navicli -h 192.1.4.251 networkadmin -set -address 192.168.1.200 -gateway 192.168.1.104 to change SPA back to an Internal IP address configuration

c.) Edited /etc/hosts & /nas/site/sp\_info files to restore default internal IP addresses for both SPs

d.) Rebooted Control Station

e.) Verified that files remained unchanged

f.) Verified communication with both SPs on internal IP address via Navicli

g.) Failed back trespassed luns

h.) Used nas\_storage -check -all successfully

i.) Fixed broken Clariion security cache files on Celerra by using nas\_storage -modify id=x -security to update cache files

### 4. Results should show unconfigured state

#### **# /nasmd/sbin/clariion\_mgmt -info**

Error 12: Not configured

**Note:** System unconfigured and ready to be converted directly to Proxy-ARP using clariion\_mgmt -start -spa\_ip xxx -spb\_ip xxx -use\_proxy\_arp

## CELLERRA STARTUP ASSISTANT (CSA):

→The CSA is a stand-alone GUI-based wizard with NS20 & NS40's that replaces the NS350 InitWizard (but uses same iwd process and InitWizard function on Control Station side) and assists in pre-configuring Celerra installations

**Note:** CSA is currently available on all shipping Integrated systems

→All NS20 & NS40 factory & field-install models are expected to have NAS, Flare, & Control LUNs ready for system configuration on power-up, complete with physical label on CS indicating a Factory install, plus an “electronic seal” that will warn a User that the system has been pre-installed if a boot floppy & CD-ROM are used to boot the system (on such a bootup, the system will check for the

CableCheck script and if found, will consider it a factory install and notify the User). In order to perform a true fresh install, the system would require that the Backend be cleaned of Celerra LUNs and RGs before proceeding.

→S95cable\_check script first checks to see if eth3 files exist on CS in /etc/nas\_device.map (fails if IP already created), runs a /nas/tools/.factory\_check script, starts the iwd daemon using /sbin/service nas\_ipinit, mounts /nas, verifies the NAS Model, untars /nas/tools/tftpboot.tar.gz to /tftpboot, creates /tmp/.factory\_check\_successful file, runs .factory\_check again, waits for User to run CSA

→Beginning with Napa 9 release, Proxy ARP will automatically be setup during CSA Pre-Configuration “Apply” screen phase for all NS20 & 40 integrated installations in a series of (12) steps that will be seen on the progress bar, and involves rebooting the SPs for the change from Internal to External IP addressing

→CSA support has been extended to NX4, NS-120, NS-480, & NS-960 platforms

## **CSA VERSIONS & NAS SUPPORT (CSA Compatibility Guide—emc186947):**

### **NAS 5.5:**

CSA 5.5.30 supports NAS 5.5.30 only, and sets up NAT by default during the configuration process

CSA 5.5.31 supports NAS 5.5.31 only, and sets up Proxy ARP by default during the configuration process

CSA 5.5.32 supports NAS 5.5.32 and 5.5.31, setting up Proxy ARP by default

CSA 5.5.33 supports NAS 5.5.33, 5.5.32, and 5.5.31, failing with version not compatible message on NAS versions > than 5.5.33

**Note:** There is no CSA version for NAS versions 5.5.34, 5.5.35, or 5.5.36

CSA 5.5.37 supports NAS 5.5.37, 5.5.36, 5.5.35, and 5.5.34

### **NAS 5.6:**

CSA 5.6.36 supports NAS 5.6.36 only

CSA 5.6.37 supports both NAS 5.6.36 and 5.6.37 (Due to issues, use only 5.6.38 CSA with NAS 5.6.37 installs)

CSA 5.6.38 supports NAS 5.6.36, 5.6.37, and 5.6.38 (But use only the 5.6.38 CSA posted on Powerlink, not the Apps & Tools CD)

CSA 5.6.39 supports NAS 5.6.39 only (introduces ConnectHome and Email Notification changes within CSA)

CSA 5.6.40 is compatible with NAS versions 5.6.40 and 5.6.39 (CSA is backwards compatible with 5.6.39 NAS version)

CSA 5.6.41 is compatible with NAS versions 5.6.41, 5.6.40, and 5.6.39 (CSA is backwards compatible with 5.6.40 & 5.6.39 NAS versions)

CSA 5.6.42 is compatible with NAS versions 5.6.42, 5.6.41, 5.6.40, and 5.6.39 (CSA is backwards compatible with 5.6.41, 5.6.40, & 5.6.39 NAS versions)

CSA 5.6.43 is compatible with NAS versions 5.6.40, 5.6.41, 5.6.42, and 5.6.43, but NOT with NAS 5.6.39 systems

CSA 5.6.43.84 is compatible with NAS versions 5.6.39, 5.6.40, 5.6.41, 5.6.42, & 5.6.43 Installs. This version also introduces the new Celerra Provisioning Wizard functionality, presented as the ‘Configure’ option when first opening the CSA application. This version of the CSA supports provisioning on NX4, NS20, NX40, NS-120, & NS-480 systems, with NAS versions 5.6.40, 5.6.41, 5.6.42, or 5.6.43. This version is also only available for EMC employees and Partners for the initial release, and is downloadable from Powerlink.

**Note:** With NAS 5.6, the CSA version is displayed on the righthand side in the Welcome screen. Additionally, the complete NAS version can be seen on the last line of the c:\Program Files\EMC\CSA\startup\config\startup.ini file.

### **Example:**

CSA\_Version = 5.6.44.4

## **HOW THE WINDOWS CSA & CONTROL STATION IWD DISCOVERY PROCESS WORKS:**

→At initial bootup, the Control Station console should display the following: “Waiting for Celerra Startup Assistant (CSA), to continue...”

→At this point, the “iwd” process should be running on the Control Station and listening on Port 60260 for UDP broadcast traffic

→CSA application is started on the Windows client and on “Welcome” screen, selecting ‘Next’ begins a UDP broadcast from the CSA client to 255.255.255.255 to Port 60260

→The iwd process on the Control Station receives the broadcast packet and replies with a UDP packet containing the MAC address and NAS version to the CSA client

→CSA client then displays the MAC address in the ‘Select a MAC address’ dropdown bar & pre-populates various network & dns parameters in the CSA screen by parsing the Client’s ipconfig/all output

→User enters Control Station IP, Hostname, Gateway, and other information on the CSA “Initialization” screen, and a UDP packet is sent to the MAC address

→The iwd process parses the UDP packet, compares that Client & CS IP Address are on same subnet, and sets the CS IP address, netmask, Hostname, and Gateway

→CSA presents logon screen (with option to change default passwords), and then logs into Control Station using IP (all further communications between CSA and CS are via IP)

## **PREREQUISITES FOR USING CSA:**

→Windows Host with csa\_install.exe application installed (Apps & Tools CD or powerlink)

→Windows host using CSA must be on same physical subnet as Control Station in order to receive the UDP MAC address broadcasts from Control Station

→Control Station must not have any eth3 IP entries. If the CS has an external IP address already configured, the CSA will error out on the Initialization screen past the Welcome screen:

**ERROR: Celerra not found**

→Control Station must have S95cable\_check script running in order to start the IWD daemon, which then listens for udp broadcasts on Port 60260 and responds to CSA broadcasts to physical network 255.255.255.255

→Remove /tmp/.factory\_check\_successful file if CSA is not discovering the Control Station [presence of this file may not prevent CSA from discovering the Celerra, but will prevent the console ‘Waiting for Celerra Startup Assistant’ message from being displayed, which is one of the key indicators that the CSA is ready for discovering the Control Station]

→If the Client Windows system has already been used to configure other NS20/20FC/40/40FC systems, the user will need to remove the previous XML configuration file in order to run the CSA wizard:

Program Files>EMC>CSA>startup>test>CSA\_APM000e0ccfc72a.xml

**Note:** This issue was resolved with AR100493 NAS 5.5.32, which will now allow the User to select a “New Configuration” or existing MAC address from a Dropdown menu [Welcome page > Select configuration profile]

**JAVA VERSION OF CLIENT:**

C:>java -version

java version "1.5.0\_05"

**CONTROL STATION JAVA VERSION 5.6.43:**

# cd /nbsnas/http/webui

# java -version

java version "1.5.0\_11"

Java(TM) 2 Runtime Environment, Standard Edition (build 1.5.0\_11-b03)

Java HotSpot(TM) Client VM (build 1.5.0\_11-b03, mixed mode, sharing)

**JAVA VERSION OF CSA:**

CSA comes bundled with its own jre version, such as 1.5.0\_09, so it does not matter what version is on the client

**STANDARD WORKAROUND FOR "Celerra not found" CSA DISCOVERY ISSUES:**

**Note:** When there is any doubt as to whether the Windows Client is on the same physical network as the Celerra Control Station, use the following workaround procedure to directly connect a straight-through (or crossover) Ethernet cable between the Windows Client and the Control Station:

1. Configure the Windows Client with a static IP address on the same logical network being used by the Control Station, using a netmask of 255.255.255.0, and no default gateway address.
  2. Use a straight-through Ethernet cable to connect the CSA client directly to the external port on the Control Station (labeled as “eth3”).
  3. Reboot the Control Station and wait for the reboot to complete.
  4. Verify that the CSA Client does not have a pre-existing CSA configuration XML file (check Program Files>EMC>CSA>startup>test>CSA\_000e0ccfc72a.xml), and remove the file if present.
5. Restart the CSA wizard and click Next at the Welcome screen to initiate the broadcast and discovery process, then continue with the CSA configuration.

**REASONS AND RESOLUTIONS FOR THE "Celerra not found" ERROR REPORTED BY CSA**

**WIZARD DURING CONTROL STATION DISCOVERY/INITIALIZATION PROCESS:**

**Note:** Please refer to emc173495 for additional information

**1. The Control Station external interface (eth3) is already configured.**

**Explanation:**

Once the external interface has been configured, NAS Services are started and the IWD daemon, which is the Linux process that provides the interface MAC address and NAS version to the CSA, is no longer needed and is shutdown.

**Resolution:**

The existence of a configured interface would mean that the Celerra was previously configured, perhaps because of earlier customer or installer testing. Whatever the reason, if the desire is to completely “reconfigure” the Celerra system, refer to solution emc166265 to restore the Celerra to a pre-CSA configuration state. In other situations it may be best to conduct a complete Backend Cleanup and fresh NAS factory installation in order to return a Celerra system to a state where it is ready for CSA configuration - see emc171121 for the link to an internal EMC procedure to cleanup and reinstall an Integrated Celerra system.

**2. The IWD Daemon is not running.**

**Explanation:**

Missing /etc/rc3.d/S95cable\_check script or the iwd process was killed.

**Resolution:**

Reboot the Control Station and log into the system as root user from the serial console. Check to see whether the IWD process is running. If the process has not started, try to manually start it by running /sbin/service nas\_ipinit start. Check the /var/log/messages for additional information on the “InitWizard” to determine why the iwd process may not have started:

```
# ps -eafl |grep -i iwd  
/nasmcd/sbin/iwd 040 R root 3429 3428 0 76 0 - 348 - 11:59 pts/0 00:00:00
```

**Note:** For further troubleshooting, refer to solution emc166265 for advice on how to restore the system to a state where the CSA configuration can begin.

### **3. The Windows client is not on the same physical network (or does not have an IP on the same logical network) as the Control Station.**

The Windows client running the CSA application must be on the same physical broadcast network as the Control Station in order for the MAC address CSA discovery feature to find the unconfigured Control Station. It must also be using an IP address from the same subnet that is applied to the Control Station

#### **Explanation:**

By definition, the CSA client and the Control Station must be co-located on the same physical broadcast domain, and likewise be configured to use the same logical network in order for the CSA "discovery" process to find the unconfigured Control Station.

#### **Resolution:**

When there is any doubt as to whether the Windows client is on the same physical network as the Celerra Control Station, use the "Standard Workaround for 'Celerra not found' CSA discovery issues" above.

### **4. The Celerra Control Station may not be properly powered up.**

#### **Explanation:**

The Control Station may not have been properly powered on or has hung during bootup.

#### **Resolution:**

- a. Manually reboot the Control Station and verify that the system powers up.
- b. For additional troubleshooting, follow the steps in "Accessing the Linux Serial Console Directly to log into the Celerra" above.
- c. The Windows client running the CSA may be running any combination of VMWare, using multiple physical NICs, firewall software, or VPN software.

#### **Explanation:**

VMWare and multi-homed Windows clients running the CSA can negatively affect the Control Station MAC address discovery process. Additionally, a Windows client running the CSA with firewall and VPN applications can also impact Control Station discovery and initialization using the CSA Wizard.

#### **Resolution:**

Manually disable all firewall and VPN applications on the Windows client. Manually disable all virtual or physical NICs on the Windows client, with the exception of the LAN port. See solution emc171539 for more information on the effects of VMWare and multi-homed hosts on the CSA configuration process.

### **6. The Cisco Security Agent (also known as CSA, CSAgent, etc.) security application blocks JRE requests to discover the Celerra Control Station on clients running the Celerra CSA Wizard.**

#### **Explanation:**

The Cisco Security Agent blocks the Celerra CSA JRE application when it attempts to send UDP packets to port 60260, which is used to discover the Celerra MAC address. The Event Viewer>Application Log clearly shows the Cisco Security Agent (CSA) preventing the Celerra CSA process from succeeding, with "The process 'C:\Program Files\EMC\CSA\jre1.5.0\_09\bin\javaw.exe'...attempted to accept a connection as a server on UDP port 60260...The operation was denied."

#### **Resolution:**

The permanent resolution is for the Cisco Security Agent to allow the Celerra CSA Wizard to use UDP port 60260 so that the Control Station 'discover' process can occur. The current status of this issue is that the problem is resolved. Any EMC user that experiences this problem should make sure that they first log into the Corporate network to update their logon profiles, which will then allow the Celerra CSA wizard to function as designed. For troubleshooting purposes, review the Event Viewer>Application log for more information on what applications are and are not affected by the Cisco Security Agent/Firewall.

### **CONTROL STATION CONSOLE DISPLAY ON INITIAL POWERUP OF FACTORY SYSTEM:**

"Starting cable\_check: This is an NS40FC system waiting to be configured by Celerra Startup Assistant (CSA). Use CSA to verify that the system has been cabled properly. Once the check is complete, CSA will instruct the Cable Check Utility to continue and NAS services will be started and a login prompt provided.

If you want to break out from this and login to the Control Station (you must really know what you are doing) press "L".

Waiting for Celerra Startup Assistant (CSA), to continue....."

→NAS Services & Box Monitor are not running at this stage (S95cable\_check script is set to run prior to S95nas)

→/etc/rc3.d/S95cable\_check script runs, does cable check, starts up iwd daemon

→You can get to login prompt by pressing shift + L"

### **CSA BYPASS PROCEDURE emc166266 NS20/NS40:**

**Warning:** You should not normally need to bypass the CSA process, but, in the event the CSA has encountered a bug or some other undetermined reason for not being able to complete a factory configuration, then it might be possible to skip the CSA altogether and use the following steps to setup the system—please consider this a last-resort measure as the CSA is not intended to be circumvented.

1. Press shift + L keys and log into the Control Station over the serial connection as root user
2. Rename the /etc/rc3.d/S95cable\_check file (to prevent it from running on reboot)
3. Rename the /etc/rc3.d/ccS95nas link to S95nas (link to NAS startup script /etc/rc.d/init.d/nas)

**Note:** The S95cable\_check script renames S95nas to ccS95nas when it starts up to prevent NAS from starting, then renames the file back at a certain point in the CSA process, which allows NAS services to start.

4. Set Hyperterm session emulation to VT100

5. Edit the following files using the vi editor:

a.) Add line to /etc/hosts for Control Station IP & Name

10.241.168.150 sludge4 #Control Station

b.) Edit hostname & add Default Route in /etc/sysconfig/network

Hostname=sludge4

Gateway=10.241.168.128

c.) Edit /etc/sysconfig/network-scripts/ifcfg-eth3 file to configure public IP address

DEVICE=eth3

IPADDR=10.241.168.150

NETMASK=255.255.255.0

NETWORK=10.241.168.0

BROADCAST=10.241.168.255

ONBOOT=yes

BOOTPROTO=none

d.) Ensure all ifcfg\* files have root ownership and proper permissions

# chmod 755 /etc/sysconfig/network-scripts/ifcfg\*

6. Reboot Control Station and wait for NAS Services to fully start

7. Verify that eth3 and Hostname of CS are correct

8. Run clarion\_mgmt script to build NAT [builds alias & IPTables entries] or Proxy ARP configurations:

a. First verify that neither NAT or Proxy ARP are configured:

# /nasmcd/sbin/clarion\_mgmt –info

Error 12: Not configured

b. Rebuild NAT or Proxy ARP configuration:

→NAS 5.5.30.5 only, for NAT: # /nasmcd/sbin/clarion\_mgmt –start –spa\_ip 10.241.168.251 –spb\_ip 10.241.168.252

→NAS 5.5.31.6+ only, for Proxy ARP: # /nasmcd/sbin/clarion\_mgmt –start –spa\_ip 10.241.168.251 –spb\_ip 10.241.168.252

#### **use\_proxy\_arp**

**Note:** To rebuild NAT on 5.5.30 only, must use 5.5.30 version of clarion\_mgmt script, using –start syntax. If the same command were issued using NAS 5.5.31.6 or higher, the default would be to setup Proxy ARP, not NAT.

9. Use Celerra Manager to configure Celerra

### **FACTORY INSTALL PROTECTION MECHANISM AGAINST ACCIDENTAL REINSTALL:**

If a system has been factory-installed, the /etc/rc3.d/S95cable\_check script will provide the following console warning if a User attempts to perform a fresh install by using a boot floppy, cd-rom, and uses serialinstall method:

“EMC Celerra Control Station Linux

```
+-----+ Installation Exists +-----+
| This system has been pre-installed in the factory. It is ready to be configured using |
| Celerra Startup Assistant. |
| See solutions emc159893, emc163261 and emc163262 for more details.|
| Press OK to reboot the Control Station. |
| Press <Alt-F2> to enter a command shell. |
| |
| +---+ |
| | OK | |
| | “ |
```

**Note:** The purpose of this mechanism is to prevent casual reinstallation of the entire system when minimal configuration is required to get the system up and running using the CSA utility

At this point, reboot CS without install media, and once logged in, remove the /etc/rc3.d/S95cable\_check script, then determine whether Backend Cleanup procedure is needed before attempting NAS reinstallation

### **CSA BYPASS PROCEDURE FOR NS-120/NS-480 SYSTEMS:**

1. Connect a null modem serial cable between a service laptop and the Serial Console port on the back of the Control Station.

**Note:** Serial Console port is located on the right-hand side of Control Station when viewed from the back of the system, indicated by a "wrench" symbol.

2. Open a HyperTerminal session on the Client Workstation using the following settings:

19200 bits per second, 8 data bits, parity None, 1 stop bit, flow control None, and terminal emulation Auto Detect/ANSI.

3. Click the phone icon to establish the serial output in HyperTerminal

4. Press the shift + L keys and log into the Control Station over the serial connection as nasadmin, then su to root. Or, if the login screen is already displayed, login as nasadmin, then su to root.

5. Rename the /etc/rc3.d/S95cable\_check file to prevent it from running on reboot.

# mv S95cable\_check s95cable\_check ori

6. Verify that /etc/rc3.d/S95nas shows up as a link to /etc/rc.d/init.d/nas:

```
# ls -la /etc/rc3.d/S95nas
```

```
lrwxrwxrwx 1 root root 20 Apr 29 11:00 /etc/rc3.d/S95nas -> /etc/rc.d/init.d/nas
```

**Note:** Create the link if it does not already exist

```
# ln -s /etc/rc.d/init.d/nas /etc/rc3.d/S95nas
```

7. Verify that the /etc/sysconfig/network-scripts/ifcfg-eth3 file exists with the following entries. Vi edit to create the file if necessary:

```
# cat ifcfg-eth3
```

```
DEVICE=eth3
```

```
USERCTL=no
```

```
ONBOOT=no
```

```
BOOTPROTO=none
```

8. Use the following steps to temporarily assign an IP and Gateway Address to the Control Station:

```
# /sbin/ifconfig eth3 10.241.168.150 netmask 255.255.255.0 broadcast 10.241.168.255
```

```
# /sbin/route add default gw 10.241.168.254
```

**Note:** The address and gateway are not yet written to the Control Station configuration files

9. Use the assigned CS IP address to open and log into Celerra Manager, which will be used to assign the permanent Hostname, IP Address, Gateway Address, and other network information:

- Log into Celerra Manager as Root

- Go to Celerra Home>Control Station Properties

- Set Hostname, IP Address, Netmask, Gateway, and any other relevant network information, then "Apply", then "o.k."

10. Reboot the Control Station, login as nasadmin, then su to root, and wait for the NAS partitions to mount (wait 5-10 minutes):

```
# df -h
```

|                                         |      |      |      |     |             |
|-----------------------------------------|------|------|------|-----|-------------|
| /dev/nude1                              | 1.7G | 752M | 861M | 47% | /nbsnas     |
| /dev/hda5                               | 2.0G | 589M | 1.3G | 31% | /nas        |
| /dev/ndal1                              | 134M | 54M  | 80M  | 41% | /nbsnas/dos |
| /dev/mapper/emc_vg_lun_5-emc_lv_nas_var | 97M  | 5.6M | 87M  | 7%  | /nbsnas/var |

11. Verify Control Station Hostname & External IP information:

```
# hostname
```

```
sleet-120
```

```
# /sbin/ifconfig eth3
```

|                                                     |                                              |
|-----------------------------------------------------|----------------------------------------------|
| eth3                                                | Link encap:Ethernet HWaddr 00:15:17:67:32:F4 |
| inet addr:10.241.168.178                            | Bcast:10.241.168.255 Mask:255.255.255.0      |
| inet6 addr: 3ffe:80c0:22c:139:215:17ff:fe67:32f4/64 | Scope:Global                                 |
| inet6 addr: fe80::215:17ff:fe67:32f4/64             | Scope:Link                                   |
| UP BROADCAST RUNNING MULTICAST MTU:1500 Metric:1    |                                              |

12. Run the following healthcheck:

```
# nas_storage -check -all
```

Discovering storage (may take several minutes)

done

13. Use the clariion\_mgmt script to setup the Proxy ARP configuration:

- Verify the state of the current configuration (should be in a 'Not configured' state):

```
# /nasmcd/sbin/clariion_mgmt -info
```

Error 12: Not configured

- Setup Proxy ARP (SPs must reboot during this process):

```
# /nasmcd/sbin/clariion_mgmt -start -spa_ip 10.241.168.251 -spb_ip 10.241.168.252 -use_proxy_arp
```

Checking if running as root...yes

Checking if model is supported...yes

Checking for integrated system...yes

Checking if interface eth3 is configured...yes

Checking if interface eth3:1 is configured...no

Checking if interface eth3:2 is configured...no

Checking if IP (10.241.168.179) is available...yes

Checking if IP (10.241.168.180) is available...yes

Checking if SP (128.221.252.200) is up...yes

Checking if SP (128.221.253.201) is up...yes

Checking if a gateway is setup for eth3...yes

Step 1/12: Setting up Proxy ARP for SPA on Control Station

Adding host specific route for SPA

Adding rules to allow outbound traffic from SPA

Adding ARP entry for SPA

Updating /etc/hosts entry for SPA

Step 2/12: Changing SPA IP address.

Changing SPA IP from 128.221.252.200 to 10.241.168.179 (subnetmask 255.255.255.0, gateway 10.241.168.128)

Step 3/12: Waiting for SPA to reboot.

Waiting for SPA to go down.....done (93 secs)

Step 4/12: Waiting for SPA to boot up.

Waiting for SPA to come back up.....done (167 secs)

Step 5/12: Waiting for CLARiiON software to start on SPA.

Waiting for CLARiiON software to start on SPA....done (88 secs)

Step 6/12: Updating NAS database with SPA IP address.

Updating SYMAPI database with new CLARiiON IP addresses...Ignoring signal in critical section

.Ignoring signal in critical section

.done (107 secs)

Step 7/12: Setting up Proxy ARP for SPB on Control Station

Adding host specific route for SPB

Adding rules to allow outbound traffic from SPB

Adding ARP entry for SPB

Updating /etc/hosts entry for SPB

Step 8/12: Changing SPB IP address.

Changing SPB IP from 128.221.253.201 to 10.241.168.180 (subnetmask 255.255.255.0, gateway 10.241.168.128)

Step 9/12: Waiting for SPB to reboot.

Waiting for SPB to go down.....done (77 secs)

Step 10/12: Waiting for SPB to boot up.

Waiting for SPB to come back up.....done (162 secs)

Step 11/12: Waiting for CLARiiON software to start on SPB.

Waiting for CLARiiON software to start on SPB....done (93 secs)

Step 12/12: Updating NAS database with SPB IP address.

Updating SYMAPI database with new CLARiiON IP addresses...done (81 secs)

FINISH: Operation took a total time of 15 minutes 12 seconds to complete.

14. Verify the Proxy ARP configuration and healthcheck the backend and Celerra systems:

**# nas\_storage -c -a**

Discovering storage (may take several minutes)

done

**# /nas/bin/nas\_checkup**

**# /nasmcd/sbin/clarion\_mgmt -info**

Public IP address for SPA: 10.241.168.179

Public IP address for SPB: 10.241.168.180

Start on boot : yes

Current implementation : Proxy-ARP

Status : Started

15. Verify access to the SPs using ping and log into the array using the NST tool, etc.

**# ping <sp\_ip>**

16. Configure ConnectHome and Email User Notification features:

a) Open Celerra Manager>Support>ConnectHome and EmailUser

17. Run the Registration Wizard and submit the registration file to EMC

## **BACKEND CLEANUP PROCEDURE & FACTORY REINSTALL:**

**Note:** Occasionally, a system will require a complete “factory” reinstallation of NAS, especially for systems that require the use of an Internal Celerra IP address scheme that is different than the 192.168.x defaults. NS systems cannot be reinstalled, however, until the backend Luns, Raid Groups, & Storage Groups are removed. The following procedure can be used on NS20, NS40, & NS80 Integrated systems. Also important to note is that the SPs need to be on the same Internal networks as the Primary & Secondary internals on the Celerra for the install to complete.

**Warning:** The nas\_raid -s cleanup command will remove the default Celerra\_emcnas\_i0 Storage group and all associated Raid Groups and LUNs, including the Hot Spares, and destroys array security and disables AccessLogix. If there are other Storage Groups on the backend, then the nas\_raid cleanup command assumes that the array is shared and will not remove any other Storage Groups and associated RGs & LUNs, nor will it remove Hot Spares, leaves Domain security enabled, and keeps AccessLogix Enabled.

1. Perform the following steps to cleanup a system in preparation of a “factory reinstallation” of NAS on an Integrated platform:

**Note:** If the system is already a factory shipped installation, and a fresh installation is required, you may need to remove any floppy or CD-ROM media, reboot the Control Station, login at the prompt as root, then rename the /etc/rc3.d/S95cable\_check file before

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
proceeding. If the system hangs and does not come up with a login prompt, reboot the system again, and press “Shift + I” keys during bootup when the screen suggests Interactive Startup mode—then, start all services except for the “cable\_check” & “nas\_ipinit” services to complete bootup to the command prompt, then rename the S95cable\_check file.

a.) Login and rename the /etc/rc3.d/S95cable\_check file

# mv /etc/rc3.d/S95cable\_check /etc/rc3.d/cable\_check.ori

b.) Disable NAT or Proxy ARP array management (may or may not be applicable):

# /nasmed/sbin/clariion\_mgmt -stop

**Note:** If the above fails, run again and add the –skip\_rules switch

c.) If the /tftpboot directory does not exist, extract from /nas/tools

# cd /nas/tools

# gzip -d tftpboot.tar.gz

# cd /

# tar -xvf /nas/tools/tftpboot.tar

d.) Unset NAS\_DB, stop NAS Services, and run nas\_raid cleanup script from /tftpboot/setup\_backend, then verify

# unset NAS\_DB

# /sbin/service nas stop

# cd /tftpboot/setup\_backend

# ./nas\_raid -n ./bin/navicli -a 192.168.1.200 -b 192.168.2.201 -s cleanup

Log will be created in the current directory →/tftpboot/setup\_backend/nas\_raid.log

Do you want to clean up the system [yes or no]?: yes

Cleaning Storage Group "Celerra\_emcnas\_i0"

Removing LUN .....

Removing diskgroup ..

Removing initiators ...

Removing storage group "Celerra\_emcnas\_i0"

Removing spares

Cleanup (200) ..

Access Logix disabled

Security domain removed

Done

# navicli -h 192.168.1.200 getrg -lunlist

# navicli -h 192.168.1.200 storagegroup -list | -status

**Note:** Script removes Celerra Storage Group, LUNs, and RAID Groups, Host initiator records from Storage Group, Deletes Storage Group, Deletes Hot Spare & RAID Group 200, Disables Access Logix, and removes Domain Security. However, if this is a ‘Shared’ backend with other SG’s created, then Access Logix, other SG’s & luns, and Hot Spares are not touched. Also, ThinPools and ThinLuns are not deleted.

e.) If additional cleanup is desired to achieve a complete factory “reinstall” of the backend components, use Navisphere to connect to SPA and perform the following steps:

1.) Navisphere>Engineering Mode [ctrl + shift + F12]>Array>Properties>Software>CelerraService>uninstall>yes

2.) After CelerraService package is removed, log into each SPs /setup program and reset SP IP & Gateway addresses to the old defaults: 192.168.1.200/2.201 & 192.168.1.100/2.100

**Note:** Will need to use 128.221.1.250/setup & 128.221.1.251/setup & wait as SP reboots before updating the other SP. You would connect Ethernet cable from Client to SPB’s Service LAN port to connect to SPA, and vice versa.

3.) Optionally, deconfigure Write & Read cache on the array from the GUI.

Array>Properties>Cache: uncheck Write and SPA/SPB Read Cache boxes

Array>Properties>Memory: toggle write and read cache values to 0

f.) Perform ‘fresh’ install using floppy boot & CD, specify “serialinstall”, follow prompts, & DO NOT setup external LAN on Control Station

**Note:** Fresh install creates Backend gateway addresses, reboots SPs, sets up Domain Security on array, installs the CelerraService pkg on backend, provides option to configure system as “Fibre Channel Enabled” or “Integrated” [On NS40 systems only, since both models ship with Headhunter IO card on SPs]

g.) Use CSA wizard after factory reinstall to complete the Customer setup/Registration, followed by further configuration of Backend on Integrateds using Celerra Mgr/CLI, or Navi OffArray Mgr for further Backend configuration on FC models.

### **What to do if nas raid -s cleanup cannot be run?**

→Use straight ethernet cable & connect directly to SPB’s service LAN port (btm), configure workstation with IP address on 128.221.1 network, open web browser, enter default CLARiiON IP address for SPA for Navisphere: 128.221.1.250

(1) Rightclick Celerra Storage Group>Connect Hosts>Select Hosts and remove

(2) Rightclick Celerra Storage Group>Destroy

- (3) Navisphere>Expand SPA and highlight LUNs 0-5>Unbind
- (4) Navisphere>Unowned LUNS>LUN 200>Unbind
- (5) Highlight RAID Group 0>Destroy
- (6) Highlight RAID Group 200>Destroy
- (7) Destroy ThinLuns and ThinPools, if required
- (8) ctrl + shift + f12 messner for Engineering Mode:  
Rightclick Array>Select Engineering Mode>Disable Access Logix
- (9) Navisphere (Engineering Mode)>Array>Properties>Software>select CelerraService>Uninstall
- (10) Enter Clariion Setup Program (128.221.1.250/setup), change Gateway IP if necessary, on SPA, which requires reboot (then connect to SPA's service port to access SPB and change Gateway IP address)
- (11) Reconnect to SPA /setup program, choose the option “Reset all domain information and restart the Management Server” to “Destroy Security and Domain Information” on array, if required—no reboot necessary

### **CLEANING UP BACKEND VIA CLI:**

- a) Remove all HS luns & delete RGs 200-204
- b) Remove SG MPFS\_Client\_0
- c) Remove initiator records (check for HBA records using navicli -h <sp\_ip> port -list)

# **/tftpboot/bin/navicli -h 128.221.252.200 unbind 200**

Unbinding a LUN will cause all data stored on that LUN to be lost.

# **/tftpboot/bin/navicli -h 128.221.252.200 removevg 200**

# **/tftpboot/bin/navicli -h 128.221.252.200 storagegroup -destroy -gname MPFS\_Client\_0**

Destroy Storage Group MPFS\_Client\_0 (y/n)? y

# ./navicli -h 128.221.252.200 port -list -hba |grep Logged

Logged In: YES

Logged In: YES

# /nasmcd/sbin/t2reset reboot -s 2

# **/tftpboot/bin/navicli -h 128.221.252.200 port -removehba -all | -host <host\_name>**

Remove the following initiator records for all attached hosts:

Only logged out but registered initiator records will be removed. Do you want to continue (y/n)? y

**Note:** There have been reports of Initiator Records that have not been removed by the Backend Cleanup script. If the navicli port –removehba –all command fails because the Host Initiator is logged in, reboot the blades or unseat them from the midplane, then run the removehba –all command.

# **/tftpboot/bin/navicli -h 128.221.252.200 thinlun -destroy -l <lun\_number>**

# **/tftpboot/bin/navicli -h 128.221.252.200 storagepool -type <pool\_type> -destroy -id <poolID> -name <pool\_name>**

### **FC ENABLED SYSTEMS MAY NOT BE TOTALLY CLEANED UP USING NAS RAID:**

If other Hosts are sharing the array, their Raid Groups & LUNs will most likely be configured under different Storage Groups than the default Celerra SG. Please be aware that the nas\_raid cleanup script will only destroy RGs & LUNs that are grouped under the Celerra\_emcnas\_i01” Storage Group. If other Storage Groups exist, their respective luns & RGs will remain intact. Additionally, Array Security is left intact and Access Logic is not disabled. The default Hot Spare 200 and RG 200 are also left alone and not deleted. Use Navisphere to remove luns from all Storagegroups, unbind all luns, then destroy all remaining Raid Groups and Storage Groups, then run the /setup program to destroy domain security on the array.

### **FACTORY REINSTALL ON OLD NS40 INTEGRATED USING NAS 5.5.31.x CODE:**

**Note:** NAS 5.5.30.x and above will install without issue on the ‘first generation’ NS40 Integrateds, and you could choose not to setup the external CS LAN IP address during the install, and you could use the CSA to assign the Hostname & external IP Address, but the CSA will thereafter fail with a NAS Model error, meaning that there is nothing more than can be configured from that point. So, the more logical approach would be to just reinstall old systems that are not qualified to use the CSA with the external LAN IP address and CS hostname during the NAS install and not use the CSA at all.

### **WHAT NAS 5.5.30.x & HIGHER CONFIGURES ON OLD NS40 HARDWARE DURING INSTALL:**

- AccessLogix is enabled
- LUNs are added to a storage group that is created during the installation, and is based on the Control Station’s assigned hostname Example: Celerra\_sludge3
- /nas/sbin/model reports NS40 (no change)
- (2) Data LUNs are created that consume all available space in RG0
- CelerraService package is installed on the array (at least if already run Flare 24 on the backend)
- eth0:1 & eth2:1 aliases are created on the Control Station for the gateway IPs on the SPs [192.168.1.104 & 192.168.2.104]
- IPTables service is started and NAT entries are created for SPs
- A POSTROUTING -s 192.168.1.200 -d 192.168.2.201 -p tcp -j SNAT --to-source 192.168.1.200
- A POSTROUTING -s 192.168.2.201 -d 192.168.1.200 -p tcp -j SNAT --to-source 192.168.2.201

-A POSTROUTING -s 192.168.1.200 -d 192.168.2.201 -p udp -j SNAT --to-source 192.168.1.200

-A POSTROUTING -s 192.168.2.201 -d 192.168.1.200 -p udp -j SNAT --to-source 192.168.2.201

→Sets up IP Forwarding with a value of 1 [/proc/sys/net/ipv4/ip\_forward]

→Sets up Domain array security using the NAS installation default account name of ‘nasadmin’ with password ‘nasadmin’

→Deletes the /etc/rc3.d/S95cable\_check script after the first reboot of the CS

## **VARIOUS PHASES DURING CSA CONFIGURATION:**

### **Welcome screen & Initialization Phase**

→After Welcome screen, enter Username and Password

→Select the Control Station that was discovered during the MAC broadcast in the dropdown menu; plugin the CS Hostname and IP information, then Next

→Setting up blade page, enter DNS, NTP Server, and Gateway information, and Public IP Addresses for SPA and SPB (calls clarion\_mgmt script to configure CS aliases and NAT entries to support the public SP addresses)

→Collecting license and site information page, add Site ID and Modem information, apply licenses that apply, select “Call EMC” or “Email Home” option (calls nas\_support script to configure NAS Events, Sendmail, and other files)

→During initialization, the CSA application runs on a Windows Host over the same physical network as that of the external network interface of the Control Station. The application locates the CS via MAC address broadcasts, and checks that the Electronic Seal for the system is intact [i.e., checks to see if an IP address has been already configured for the external CS interface], at which point the configuration wizard would start. If the CS was previously configured with an external IP address, the MAC address broadcast mechanism will not work and the CSA installer will fail to find a configurable Celerra. The functional spec advises that a fresh install of the system is required at this point, though there may be workarounds. For example, if a network connection was terminated in the middle of initial preconfiguration, the CSA should be able to continue from where it left off

### **Health Check Wizard Phase**

→After Initialization is completed, a system CableCheck is conducted (using a modified setup\_enclosure script) and the state of the software is checked

### **Pre-Configuration Phase**

→After the Healthcheck phase, all configuration changes are presented for review. After applying changes, the configuration is saved to the local Windows Host in XML format, making the CSA wizard restartable from last point of exit, and NAS Services, Web, Apache, & Tomcat services are started

→Click ‘Next’, then ‘Finish’ to move to the Wizards page, where CIFS Shares, NFS Exports, iSCSI LUNs, and Reg. can be done

### **Configuration Wizards**

→Various servlets that are derived from Celerra Manager code, are called to assist in setting up a file system, Share, Export, iSCSI LUNs, & CIFS Server

→NFS, iSCSI, and CIFS icons are presented

### **Registration Phase**

As a final step, the Celerra can be registered with EMC using the Celerra Platform Registration Wizard

→Enter IP address (or Hostname) and User and Password for Control Station, select Installer identity, then enter Customer Information and Address

→Wizard then runs /nas/sbin/log\_config script on CS to gather up necessary data, then gives options on how to transmit the information to EMC (Automatic, Manual, Email)

### **Finish Phase**

The Finish page gives User the option of connecting to the product “Landing Page” (on Powerlink) or starting a Celerra Manager session.

## **SCRIPTS USED BY CSA DURING CONFIGURATION:**

1. Initial bootup starts /etc/rc3.d/S95cable\_check before S95nas

2. Initial bootup also calls /nas/sbin/nas\_ipinit to startup the IWD daemon in order to broadcast MAC address of CS

3. Upon CSA Initialization screen, a **/nas/sbin/setup\_enclosure –factoryCableCheck** script is run to validate system cabling and integrity of the communication network—broadcast determines if all Mgmt Switches are present, in their correct slots, and that enclosures are ordered correctly

**Note:** CableCheck will compare MAC addresses found against what is stored in /etc/nas\_enclosure.map, checks that all mgmt switches are present and in their correct slot, and verify that enclosure order is correct

### **CABLE CHECK CODES & ACTIONS:**

| CODE       | DESCRIPTION                       | ACTION Required                                             |
|------------|-----------------------------------|-------------------------------------------------------------|
| 1001000000 | CableCheck successful, no errors  |                                                             |
| 1001010200 | Mgmt switch for Slot_2 missing    | Check that lower mgmt port on DM2 is cabled to CS port eth0 |
| 1001010300 | Mgmt switch for Slot_3 missing    | Check that lower mgmt port on DM3 is cabled to CS port eth2 |
| 1001030200 | Mgmt switch for Slot_2 not cabled | Check that lower mgmt port on DM2 is cabled to CS port eth0 |
| 1001030300 | Mgmt switch for Slot 3 not cabled | Check that lower mgmt port on DM3 is cabled to CS port eth2 |
| 1001999999 | Unrecoverable error               | Contact EMC Support                                         |

4. CSA calls clariion\_mgmt to configure public IP Address aliases to support NST/Navisphere access to SPs from a Windows Host, and configures NAT IPTable entries

#### **# /nasmcd/sbin/clariion\_mgmt -start -spa\_ip 10.241.168.250 -spb\_ip 10.241.168.251**

```

Checking if running as root...yes
Checking if model is supported...yes
Checking for integrated system...yes
Checking network interface eth3...configured
Checking network interface eth3:1...no configuration file
Checking network interface eth3:2...no configuration file
Checking if IP forwarding is enabled...yes
Removing network alias :1...done
Removing network alias :2...done
Creating network alias eth3:1...RTNETLINK answers: File exists
done
Creating network alias eth3:2...RTNETLINK answers: File exists

```

5. Beginning with Napa 9 release, the Proxy ARP script is executed during the Pre-Configuration phase of the CSA to setup Proxy ARP on all NS20 & NS40 integrated models during configuration/installation.

#### **# cat /etc/sysconfig/iptables**

```

-A PREROUTING -d 10.241.168.250 -p tcp -j DNAT --to-destination 192.168.1.200
-A PREROUTING -d 10.241.168.250 -p udp -j DNAT --to-destination 192.168.1.200
-A PREROUTING -d 10.241.168.250 -p icmp -j DNAT --to-destination 192.168.1.200
-A PREROUTING -d 10.241.168.251 -p tcp -j DNAT --to-destination 192.168.2.201
-A PREROUTING -d 10.241.168.251 -p udp -j DNAT --to-destination 192.168.2.201
-A PREROUTING -d 10.241.168.251 -p icmp -j DNAT --to-destination 192.168.2.201

```

**/nas/site/clariion\_mgmt.cfg** (Configuration stored here)

**/etc/clariion\_mgmt.cfg** (Copy of configuration stored here)

#### **# cat /nas/site/clariion\_mgmt**

SPA\_PUBLIC\_IP=10.6.4.158

SPB\_PUBLIC\_IP=10.6.4.187

ONBOOT=yes

### **RECONFIGURING USING CLARIION MGMT:**

The /nasmcd/sbin/clariion\_mgmt script is invoked by CSA during system initialization, but can also be run manually to reconfigure the public IP addresses and aliases that are used on the Control Station for NST & Navi OffArray Mgr connectivity—there are no reboots required when using this tool if running a NAT configuration. For Proxy\_ARP, the SP's will need to reboot if their IP Addresses change.

#### **# /nasmcd/sbin/clariion\_mgmt -info**

```

Public alias IP address for SPA: 10.241.168.250
Public alias IP address for SPB: 10.241.168.251
Start on boot      : yes
Status           : Started
done

```

#### **# /sbin/ifconfig**

```

eth3:1  Link encap:Ethernet HWaddr 00:0E:0C:CF:C7:2A
        inet addr:10.241.168.250 Bcast:10.241.168.255 Mask:255.255.255.0
eth3:2  Link encap:Ethernet HWaddr 00:0E:0C:CF:C7:2A
        inet addr:10.241.168.251 Bcast:10.241.168.255 Mask:255.255.255.0

```

**Note:** Above entries represent public IP's and aliases for SPA, SPB, resp.

#### **# /nasmcd/sbin/clariion\_mgmt -stop**

```

Checking if running as root...yes
Checking if model is supported...yes
Checking for integrated system...yes
Checking network interface eth3...configured
Checking network interface eth3:1...configured
Checking network interface eth3:2...configured
Checking if IP forwarding is enabled...yes
Removing network alias eth3:1...done
Removing network alias eth3:2...done

```

Done

### What does running clariion\_mgmt -stop do?

- Removes Public SP IP's from Control Station aliases on eth3
  - Removes PREROUTING entries in /etc/sysconfig/iptables for translation of public IP's to internal IP addresses for SPA/SPB
  - Verifies that IP\_Forwarding is set on the Control station
- Note:** Beginning with Napa 9, the clariion\_mgmt -stop removes the Proxy ARP configuration, and resets the public IP addresses on the SPs to the default 192.168.x.x addresses, which requires SP reboots. Please note that this command does NOT prompt the User for confirmation before rebooting SPs.

# cat /proc/sys/net/ipv4/ip\_forward

1

## TROUBLESHOOTING CSA CABLE CHECK FAILURES:

- CableCheck failure messages must be resolved before the CSA setup process can continue. It's important to know that the serial Console to the Control Station is disabled until the CableCheck can be successfully run
- Serial console will be disabled, and NAS Services will not start until a successful CableCheck is run
- Users can connect to the serial console via Hyperterm or SSH for further troubleshooting, but may need to use Shift + L in order to get to a login prompt [new console locking feature]
- Use the downloaded documentation or link to Powerlink to debug CableCheck and other CSA error popups

## TROUBLESHOOTING CSA ISSUES:

- CSA Popups will display information on what to troubleshoot, but may not always be accurate
  - Powerlink access to troubleshooting information via Landing Page
- Note:** There is a downloadable 3MB NS20/40 Cable Check Troubleshooting Bundle called “InstallationTroubleshooting.zip”, available from Powerlink>Celerra Tools>NS20 Integrated>Celerra Troubleshooting>Installation Troubleshooting. Unzip the bundle in the Program Files>EMC>CSA folder on the system from which the CSA will be run & overwrite any existing files with the newer files. This will provide the proper Cable Check troubleshooting popups when running the Cable Check portion of the CSA wizard during system configuration.
- Windows workstation logs some CSA information in a startup\_wiz.log file Program Files>EMC>CSA>Startup>startup\_wiz.log
  - No CSA logs are created on the Control Station
- Note:** Beginning with NAS 5.6.39.5, the clariion\_mgmt script creates a logfile whenever run. For CSA configurations, the log is created after the Cable Check when the backend IPs are being configured.

### /var/log/clariion\_mgmt.log

- CSA Application resides in Program Files>EMC>CSA directory
- Keep in mind that once the CS IP Address & Hostname info is entered into the CSA, a “CSA...xml” file is created if the application exits prior to completion, the purpose of which is to allow restart from point of last operation:  
Program Files>EMC>CSA>startup>test>CSA\_000e0ccfc72a.xml
- Key troubleshooting tip: If stuck, exit CSA application, then reopen CSA, logon with Username & Password, and see if you get beyond last problem
- Search for Primus solutions for CSA

## RUNNING CSA CABLE CHECK SCRIPT MANUALLY:

# /nbsnas/tools/.factory\_check verify\_cable -list →This produces component list (NX4 NAS 5.6.39)

- 1,Control Station 0 MGMT A,CS Network
- 2,Control Station 0 MGMT B,CS Network
- 3,MGMT Switch,MGMT Switch
- 4,SPA,SP
- 5,SPB,SP
- 6,Blade 2,Reason Code
- 7,Blade 2 Primary Network,DM Network
- 8,Blade 2 Secondary Network,DM Network
- 9,Blade 2 BE 0,DM Fiber
- 10,Blade 2 BE 1,DM Fiber
- 11,Blade 3,Reason Code
- 12,Blade 3 Primary Network,DM Network
- 13,Blade 3 Secondary Network,DM Network
- 14,Blade 3 BE 0,DM Fiber
- 15,Blade 3 BE 1,DM Fiber
- 16,SPS A,SPS
- 17,SPS B,SPS

# /nbsnas/tools/.factory\_check verify\_cable -comp 16

1006019999

**Note:** Normal exit code for a component would be 0. The 1006019999 output reflects a cable check error, which can be found on the CSA client by looking for the matching error code html file in Program Files>EMC>CSA>help>NS20FC>cabling>1006019999.html.

# ls -la /tmp

-rw-r--r-- 1 root root 293 Aug 13 10:06 component\_list.csv →File created when script list is run

## **RETURNING TO PRE-CSA CONFIGURATION STATE:**

- a.) Run # /nasmcd/sbin/clariion\_mgmt –stop to remove Public IP aliases for the SPs from the Control Station & associated Network Address Translation (NAT) entries from IPTables
- b.) Copy the cable\_check script to /etc/rc3.d/S95cable\_check [the challenge here will be that a completed CSA configuration deletes this file—if a copy has been saved ahead of time, restore to /etc/rc3.d—if a copy cannot be found, contact EMC Support and reference the NAS version involved]
- c.) Remove /tmp/.factory\_check\_successful file if it exists
- d.) Rename the /etc/sysconfig/network-scripts/ifcfg-eth3 file to remove the Control Station's external interface configuration file
- e.) Remove any extra/unwanted entries for the Control Station Hostname and IP address from the /etc/hosts using the vi editor, etc.
- f.) Connect to the serial port on the front of the Control Station, use HyperTerminal and configure as follows:  
Connect using: COM1 or COM2; Bits per second: 19200 Data bits: 8 Parity: None Stop bits: 1 Flow Control: None (verify that Autodetect & ANSI are selected for terminal emulation)
- g.) Reboot the Control Station and wait for the following console screen message before continuing with the next step

### **Original CSA prompt screen and message after entering shift + L:**

*Starting cable\_check: This is an NS40FC system waiting to be configured by Celerra Startup Assistant (CSA). Use CSA to verify that the system has been cabled properly. Once the check is complete, CSA will instruct the Cable Check Utility to continue and NAS services will be started and a login prompt provided. If you want to break out from this and login to the Control Station (you must really know what you are doing) press "L".*

*Waiting for Celerra Startup Assistant (CSA), to continue.....*

*Terminating...*

*Renaming /etc/rc3.d/S95nas to /etc/rc3.d/ccS95nas*

*Remember to rename it back so NAS services can start when CS reboots.*

### **New CSA prompt screen and message after entering shift + L NAS 5.6.39:**

*Starting cable\_check: This is an NX4FC system waiting to be configured by Celerra Startup Assistant (CSA). Use CSA to verify that the system has been cabled properly. Once the check is complete, CSA will instruct the Cable Check Utility to continue and NAS services will be started and a login prompt provided.*

*If you want to break out from this and login to the Control Station (you must really know what you are doing) press "L".*

*Waiting for Celerra Startup Assistant (CSA), to continue.....L*

*Terminating...*

*NAS services are skipped since CSA health check is not completed yet. But if you are fully aware and would like to start the NAS services, type "service nas start" on the command line after logging in as root onto the control station.*

h.) On the Windows Client running the CSA wizard, remove the existing CSA...xml file from the following location:

Program Files>EMC>CSA>startup>test>CSA\_000e0ccfc72a.xml (example of filename)

i.) Startup the CSA wizard from the Windows Client: Welcome page>Next, should then be able to pickup the MAC address broadcast and continue with remaining configuration steps using CSA

## **STANDALONE REGISTRATION WIZARD [RegWiz]:**

→Though Registration is possible through the use of the CSA, there is also a Standalone application that can be used to perform the Registration after the system has been configured

→Download from “Celerra Tools” on Powerlink, or from Apps & Tools CD

Celerra Registration Wizard – v1.0

Windows application called “regwiz\_install.exe”

Installation creates a desktop shortcut called “Launch Regwiz”

Program Files>EMC>RegWiz

→Preference is to conduct Registration over live Internet connection, but when selecting Manual upload method, payload file can be saved locally and transferred to EMC Website later without requiring direct Internet connection

### **Temp Files Produced by Registration Wizard:**

Documents and Settings\Administrator\Local Settings\Temp\temp.xml & Celerra\_Registration.xml (Files are deleted after a successful upload)

### **Actual payload file when manually saving to Workstation:**

Celerra\_Registration\_APM00071601776.config

## **RegWiz Registration Options**

- o I want to register this product using Automatic Registration
- o I want to manually upload this information to the EMC Registration Web Site
- o I want to attach this information to an email and transmit it to EMC

**Note:** If selecting the Email method, there are two options:

1. Save to local drive and send .config payload file to email address referenced
2. Connect to email client, enter SMTP address for mail server, and send directly from internet-connected client

**Production Website for Automatic or Manual upload of the payload.config file:**

<http://iio.emc.com/colu/aru/>

**Email Registration Address:**

[b2b\\_product\\_registrations@emc.com](mailto:b2b_product_registrations@emc.com)

**What happens to the Registration payload once it is uploaded?**

1. Payload submitted automatically from application and HTTP, or via email attachment, or saved locally and sent via email
2. Encrypted payload goes to [B2BRegistrations@emc.com](mailto:B2BRegistrations@emc.com)
3. Files go to ProcessReg Queue & get pushed out to SYR DB
4. Files parsed & stored in SYR DB
5. SYR generates xml file format, opens SR and assigns to IBG, where IBG reviews and updates SR. Email sent to sender after getting payload and then after processing payload for registration

**NS40 INTEGRATED, NS40FC ENABLED (All-in-One) CELERRAs**

→ Enhanced Installability (preloaded code, Placemat (Setting up the EMC Celerra...document in Shipping envelope), Landing Page(Powerlink Celerra Tools page), CSA, All-in-One FC Enabled model, new Integrated model, Clariion Mgmt access to Integrated backend SPs (Via Control Station NAT Service, using Clariion NST & OffArrayMgr tools)

→ NS40 Integrated (CX3-40) and NS40FC (FC Enabled) All-in-One (CX3-40F)

→ Physically, the HeadHunter (4) Port FC IO card will be shipped with NS40 Integrateds beginning in August, but the additional Front-end FC Ports on the SPs themselves will be “disabled” and not available for use by other Hosts

→ Original NS40 Integrated to be replaced by the “new” NS40 Integrated

→ Both models will support the Celerra Integrated NST utility, and the NS40FC will also support the Navi OffArray Manager

→ Single or dual Blades, single Chivas or Dewars Control Station

**Note:** DC NS40F will use only the Chivas CS

→ SPs run in SAN mode, using AccessLogix & Storagegroups (vs. SAN AUX mode), meaning that network cables will be connected to upper LAN management ports on the SPs, not Service ports (original NS40 connected to Service port on SPs and SPs ran in SAN-AUX mode)

→ DMs connect to SPs using Fibre Optic LC-to-LC cables

→ DMs connect to upper LAN Mgmt port on SP for internal network

→ NST or OffArray Navi will connect to SPs via new NAT Service on Control Station (IPTables, new aliases on CS)

→ Models will ship with pre-installed software for Factory or Field-Installed cabinets, and will use a Placemat, Landing Page website, and Celerra Startup Assistant (CSA) to configure the box on initial startup

**Note:** During install, if FC option is selected, the management of the array will be done via OffArray Navisphere

→ Celerra will handle Back-End errors and CallHomes for array issues

→ Mirrorview uses Port 1 vs. the highest array Port number, and DMs will be cabled to SP Front-end FC ports 2 & 3 Fibre

→ Non-disruptive upgrades from NS41F to NS42F, and upgrades from integrated to gateway available

**NS40 INTEGRATED:**

→ Integrated Model

→ CX3-40F Array running in SAN mode, AccessLogix/Storagegroups

→ NS40 blades will use (2) FC Ports to SPs and (1) FC Port for tape device via optical fibre

→ Blade ports BE0 & BE1 will connect to SP Ports 2 & 3 Fibre [4-port FC IO Headhunter card shipped but not enabled]

**NS40F FC ENABLED (“All-in-One”):**

→ Fibre Channel Enabled Model (Not a Gateway, allows other Hosts to connect to array)

→ CX3-40F Array running in SAN mode, AccessLogix/Storagegroups

→ NS40 blades will use (2) FC fibre optic ports with SFP modules to connect to SPs and (1) FC Port for tape device via optical fibre

→ NS40FC Blades will use either (4) Copper 10/100/1000 Ethernet ports or (2) copper 10/100/1000 & (2) GbE optical ports

→ CX3-40F array adds (2) additional FC FE ports 2/3 for customer SAN Hosts, and (2) BE Ports 2/3 for additional DAE buses, for total of (4) BE DAE ports and (4) FE Host ports [Other Hosts to use Ports 0, & 1; Mirrorview to use Port 1]

**Note:** Quad IO card called HeadHunter

→ Blades will connect to SP Ports 2 & 3 Fibre [4-port FC IO Headhunter card shipped with this model & all ports FC enabled]

**DETERMINING NS40 MODELS & BACKEND ARRAY TYPE:**

[Celerra Manager>Celerra Home>System Info screen will show model number](#)

**New NS40 Integrated:**

\$ /nas/sbin/model

NS40

\$ /nas/sbin/navicli -h 192.168.1.200 storagegroup -status

Data Access control: ENABLED

\$ nas\_storage -i APM00071600514

model\_num = CX3-40f

\$ **/nas/sbin/navicli -h 192.168.1.200 getagent**

Model: CX3-40f

# **/nas/sbin/setup\_backend/get\_array\_version**

APM00071600514 CX3-40f 03.24.040.5.006 1 1

#### **New NS40FC:**

\$ **/nas/sbin/model**

NS40FC

**Note:** Both models use the same CX3-40f backend

→Examine the /etc/be\_sg\_info file and see if FC\_ENABLED=YES is set, which is for new NS40FC system, otherwise, if AccessLogix installed by FC\_ENABLED=NO, then the model would be the new NS40 Integrated

\$ **cat /etc/be\_sg\_info**

SHARED\_BE\_SYS\_RAID\_GROUP\_ID=0

STORAGE\_GROUP\_NAME=Celerra\_emcnas\_i0

STORAGE\_GROUP\_INIT\_REC=50:06:01:60:c1:e0:53:35:50:06:01:60:41:e0:53:35:emcnas\_i0\_dm2\_p0

STORAGE\_GROUP\_INIT\_REC=50:06:01:60:c1:e0:53:35:50:06:01:61:41:e0:53:35:emcnas\_i0\_dm2\_p1

STORAGE\_GROUP\_INIT\_REC=50:06:01:60:c1:e0:53:35:50:06:01:68:41:e0:53:35:emcnas\_i0\_dm3\_p0

STORAGE\_GROUP\_INIT\_REC=50:06:01:60:c1:e0:53:35:50:06:01:69:41:e0:53:35:emcnas\_i0\_dm3\_p1

STORAGE\_GROUP\_ALUS=0 1 2 3 4 5

FC\_ENABLED=YES

#### **Old NS40 Integrated:**

\$ **/nas/sbin/model**

NS40

\$ **/nas/sbin/navicli -h 192.168.1.200 storagegroup -status**

Data Access control: DISABLED [i.e., Access Logix]

\$ **nas\_storage -i APM00071600514 [Uses CX3-40 array]**

model\_num = CX3-40

#### **NS40 Integrated for MPFSi:**

\$ **/nas/sbin/model**

NS40

\$ **/nas/sbin/navicli -h 192.168.1.200 storagegroup -status**

Data Access control: DISABLED [i.e., Access Logix]

\$ **nas\_storage -i APM00071600514 [Uses CX3-40C array]**

model\_num = CX3-40C

**Note:** Supposed to use copper cables between DM & SPs; SPs run in SAN-AUX mode; Internal IP cabling plugs into SPs bottom service port, not mgmt port

#### **NS40FC Blade Front-End Connectivity:**

(2) FC optical fibre ports connecting DMs to SPs

(1) FC optical port for tape drive connectivity

Either (4) copper 10/100/1000 Ethernet ports per Blade or (2) copper 10/100/1000 & (2) GbE optical ports per Blade

#### **CLARiiON SPs and Back-End Connectivity:**

Total of (8) FC ports

(2) FE FC optical Ports for NS40F Blades-to-SP connectivity using Fibre Optic cables and SAN personality for SPs

(2) FE FC optical Ports for other Hosts direct-connected or FC Switch-attached

(4) BE FC arbitrated loop copper Ports to connect SPs to DAEs

#### **BASIC SPECS FOR NS40/NS40FC:**

→Basic hardware specs of new NS40s are similar to the original NS40 model

→1-2 Data Mover 1U blades, Dual Intel 2.8GHz processors, (4) GB Memory, 800MHz FSB speeds, (4) 10/100/1000 GbE ports,

(2) FC ports to the array, & (1) FC port for tape backups—FC port speeds support 2-4Gbps

→16TB per blade for FC drives, 32TB capacity per blade for mixed FC/ATA drives

→Allows for mixed FC, LCFC, and SATA drives in a shelf (like disks will be bound by RaidGroups)

→Recommended Hot Spares: 1/30 fibre channel drives; 1/15 ATA drives

#### **NS40FC RESTRICTIONS:**

**Note:** Not considered a Gateway, eventhough FC option is enabled for additional Hosts

→No Fabric connectivity between Blades and SPs (just direct-connect FC)

→Cannot have multiple Backends

→Can only have the CX3-40F backend

#### **NS40-to-NS40C HARDWARE UPGRADE PROCEDURE—Napa 9+:**

**Note:** The new NS40 Integrateds [with CX3-40F backend and new private network & fibre channel cabling scheme] can be upgraded after install by running the upgrade procedure

→Upgrade changes /nas/sbin/model output from NS40 to NS40FC

→Should see following line added to /etc/be\_sg\_info file

FC\_ENABLED=YES

1. Run Hardware Upgrade Script:

**# /nas/sbin/nas\_hw\_upgrade -fc\_option -enable**

2. Verify:

# /nas/sbin/model

NS40FC

# /nas/bin/nas\_checkup

**Note:** Above option does not work for 5.5.31 or 32 code—see ARs105630/117708. Must use the following workaround:

**# echo FC\_ENABLED=YES >>/etc/be\_sg\_info**

**# /nas/sbin/model -c**

NS40FC

## **NAPA 9:**

NAS 5.5.31.6 Oct 15, 2007

→Flare 26 Leo support--introduces Clariion support for RAID 6 with 4+2, 6+2, & 12+2 configurations (Can withstand loss of two drives in same RAID group without data loss)

→SCSI Asymmetrical Logical Unit Access (ALUA—multipath access to LUNs), Proactive Sparing with Hard Media errors, Vault drive load enhancements, WIL Cache improvements. Flare 26 only supported on CX3 arrays and CX300, 500, 700.

→Introduction of /nas/tools/nas\_summary and /nas/tools/sccslist utilities

**# /nas/tools/nas\_summary**

nas\_summary v0.1

#####

System type: NS20FC (Gateway) →AR110565, NS20FC is not considered a Gateway system but an Integrated

Version: 5.5.31-6

#####

Number of data movers: 2

server\_2: NS20 5.5.31.6 State:5 Primary

server\_3: NS20 5.5.31.6 State:5 (Standby)

#####

Number of arrays: 1

APM00073801838: Clariion CX3-10f

SPA: Flare 24 Read Cache: 32 Write Cache: 278 OPERATIONAL

SPB: Flare 24 Read Cache: 32 Write Cache: 278 OPERATIONAL

#####

Filesystems: UXFS:1 Checkpoint:0 Total:24

Number of disk vols (LUNs): 7

#####

Features in use: CIFS: \*True\* iSCSI: False NFS: \*True\*

Rep: False SRDF: False DHSM: False

#####

DBCHK State: No errors found

## **Symm DMX-4 support:**

Includes support for SATA 4Gb speed ATA drives on the Symm. This will require new Storage API Solutions Enabler v6.4.2; (6) new disk types for Symm: ATA, R1ATA, R2ATA, BCVA, R1BCA, R2BCA; (6) new system-defined storage pools: symm\_ata, symm\_ata\_bcv, symm\_ata\_rdf\_src, symm\_ata\_rdf\_tgt, symm\_ata\_bcv\_src, symm\_ata\_bcv\_tgt; (6) new volume profiles. Customer could use STD disks on PFS and ATA BCVs on the destination system when using Timefinder/FS and RDF. Control Volumes will not be allowed on Symm ATA drives.

## **PROXY ARP & CSA CHANGES:**

Beginning with the Napa 9 release, Proxy ARP will automatically be setup during the CSA Pre-Configuration “Apply” screen phase for all NS20 integrated, NS20FC, NS40 integrated, & NS40FC installations—the ‘Backend IP setup’ status field will inform users of the progress as different tasks are being completed during the Proxy ARP setup. Napa 8 systems that have the original NAT implementation will not be automatically upgraded to Proxy ARP. There is a hidden switch with /nasmcd/sbin/clariion\_mgmt –upgrade\_to\_proxy\_arp that can be used if the customer or support person wants a system to run with Proxy ARP after upgrading to Napa 9+.

## **FIXING ARP ENTRIES FOR CS INTERFACES, RECREATING HOST ROUTES FOR SP's, ENABLING PROXY ARP SERVICE:**

# /nasmcd/sbin/clariion\_mgmt –start\_service -v

## **SETTING UP, MODIFYING, or REMOVING PROXY ARP:**

**clariion\_mgmt –start** → to setup Proxy ARP for public access to SPs [requires SP reboots]

**clariion\_mgmt –modify –spa\_ip xxx.xxx.xxx –spb\_ip xxx.xxx.xxx** → change IP addresses for SPs on CS and SPs [requires SP reboots]

**clariion\_mgmt –stop –skip\_rules** → removes Proxy ARP setup [requires SP reboots]

**Note:** With Proxy ARP implementation in Napa 9 (NAS 5.5.31) during CSA configuration, it is more difficult to breakdown a system to a pre-CSA configuration state. May need to manually vi edit /etc/hosts, /nas/site/sp\_info, /nas/site/clariion\_mgmt.cfg files to give correct IP addresses to SPs. Additionally, may need to manually update the SP IP addresses via navicli.

### **How does Proxy ARP work?**

--Client wants to communicate with the SPA and does an ARP broadcast to locate SPs at Ethernet layer

--Control Station receives broadcast, has ARP entries for both SPs, and does ARP broadcast in reply

--Client sends TCP/IP packets to CS

--CS forwards packets to both SP's, and SPB drops packet while SPA processes

## **RAID 6 CLARIION-Flare 26:**

→uses 128 block element size, with (2) disks worth of parity per RAID 6 group spread across all spindles

→not supported for vault disks

→uses naviseccli only, no ClassicCLI commands—must configure system to use Secure CLI

→Clariion supports only even number of spindles: 2+2, 4+2, 6+2, 8+2, 10+2, 12+2, 14+2

→Supported configurations are 4+2, 6+2, & 12+2 for all disk types in clar\_r6, clarata\_r6, cm\_r6, & cmata\_r6 system pools

→Use setup\_clariion for shelf-by-shelf setup of FC , LCFC, or ATA drives

→Binds (2) luns per RG for FC drives using 4+2 or 6+2 Raid6

→Binds (4) luns per RG for FC drives using 12+2 configuration

→Binds (1) lun per RG for ATA drives using 4+2 or 6+2

→Binds (2) luns per RG for ATA drives using 12+2

## **CELLERA CONNECTHOME FEATURE (Napa 9):**

→ConnectEMC Linux 2.0 support, replaces Celerra CallHome process with Celerra ConnectHome

**Note:** ConnectEMC 2.0.27-b118 is the Linux version on the Control Station

→Supports EmailHome to EMC or to the ESRS Gateway, FTPCallHome to ESRS Gateway only, or CallHome via modem to EMC.

**Note:** Multiple callhome transports can be configured for Primary, Secondary, and Tertiary priority.

→ConnectHome maintains Modem support, and adds Email & FTP as new transport methods

→Each transport mechanism (Modem, Email, FTP) supports a Primary and optional Secondary failover destination—with multiple destinations, the connectemc daemon stops after the first successful transfer

→ConnectEMC is part of the EMA (Event Messaging Architecture) developed by the RAPiD software group for Windows & Linux

→Site ID is not required to process ConnectHomes in SYR or CSI

→System Serial Number, and one that is properly recorded in the Install Base Group database in CSI, is required for a ‘known’ callhome. Systems without serial number or unknown serial numbers would become unknown callhomes, unknown SR queue, etc.

## **PURPOSE OF CONNECTHOME:**

1. Primary purpose is to provide a method for transporting event files via email, ftp, or modem (outbound service)

2. Secondary functions include ability to set or modify the dial-in number, the System Serial number, Site ID number, and whether to Enable or Disable dial-in capability (inbound service)

## **FRESH INSTALLATIONS & CONNECTHOME:**

→Installs do not automatically configure the transport portion of ConnectHome (the default is not to configure anything)

→Installs of 5.5.31.6 and later implements the Celerra ConnectHome feature

→To make ConnectHome operational, must configure with either a selected Primary, and/or Secondary, and/or Tertiary transport mechanism, using either email (or email to ESRS gateway), FTP using ESRS gateway, Modem, or a combination of the three

→A valid Celerra Serial Number must also exist to ensure proper delivery & handling by EMC

→Configure ConnectHome using the Celerra Manager>Celerras>Support>ConnectHome tab, CLI using /nas/sbin/nas\_connecthome, or with NAS 6.0, Unisphere > System > Service Tasks: Manage Connect Home

→/nas/sys/connecthome.config replaces /nas/sys/callhome.config

→The Callhome Email message body is not human friendly. It's BASE64 encoded, irrespective of the encryption setting.

**Note:** If “Enable Encryption” is turned off for ConnectHome, the resulting CallHome payload can be decoded with a simple Base64 decoding tool, or via CLI on the Control Station:

# perl -MMIME::Base64 -ne 'print decode\_base64(\$\_)' <base\_64encoded\_file\_name>

## **BASE64 ENCODED CALLHOME TEST EVENT:**

PD94bWwgdmVyc2lvbj0iMS4wIiB1bmNvZGlubz0iVVRGLTgiPz4KPENvbm5lY3R1b21lIHht  
bG5zOnhzat0iaHR0cDovL3d3dy53My5vcmcvMjAwMS9YTUxTY2hbWEtaW5zdGFuY2UiIFNj  
aGVtYVZ1cnNpb249IjEuMC44Ij4KCTxUcmFuc1R5cGU+MDAxMDwvVHJhbnNUeXB1PgoJPFRy

YW5zVH1wZUR1c2M+RXZ1bnRYTUw8L1RyYW5zVH1wZUR1c2M+Cgk8VHJhbnNIRD4wPC9UcmFu  
c01EPgoJPE5vZGUGsUQ9IkFQTTAwMDczNzAxMDg1IiBTdGF0ZT0iT25saW51IiBTdGF0dXM9  
Ik9LIj4KCTxJZGVudGlmaWVypgoJCTxDbGFyaWZ5SUQ+QVBNNMDAwNzM3MDEwODU8L0NsYXJp  
Zn1JRD4KCQk8U210ZU5hbWU+czI0MDwvU210ZU5hbWU+CgkJPFW1bmRvcj5FTUM8L1Z1bmRv  
cj4KCQk8RGV2aWN1VH1wZT5DRUxFU1JBPC9EZXPY2VUeXB1PgoJCTxNb2R1bD5OUzIwRkM8  
L01vZGVsPgoJCTxTZXJpYWxOdW1iZXI+QVBNNMDAwNzM3MDEwODU8L1N1cmhbE51bWJlcj4K  
CQk8V1dOPjwvV1dOPgoJCTxQbGF0Zm9ybT48L1BsYXRmb3JtPgoJCTxPUz5EYXJ0PC9PUz4K  
CQk8T1NFVkvVSPjUuNi4zNi0yPC9PU19WRVI+CgkJPFW1b2R1VmVypjwvVWNvZGVWZXI+CgkJ  
PEVtYmVktGv2ZWw+MDwvRW1iZWRMZxZ1bD4KCQk8SW50ZXJuYWxNYXhTaXp1PjA8L01udGVy  
bmFsTWF4U216ZT4KCQk8Q29tbWVud48L0NvbW1lbnQ+Cgk8L01kZW50aWZpZXI+Cgk8Q29u

**BASE64 DECODED CALLHOME TEST EVENT:**

```
<?xml version="1.0" encoding="UTF-8"?>
<ConnectHome xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" SchemaVersion="1.0.8">
    <TransType>0010</TransType>
    <TransTypeDesc>EventXML</TransTypeDesc>
    <TransID>0</TransID>
    <Node ID="APM00073701085" State="Online" Status="OK">
        <Identifier>
            <ClarifyID>APM00073701085</ClarifyID>
            <SiteName>s240</SiteName>
            <Vendor>EMC</Vendor>
            <DeviceType>CELRRA</DeviceType>
            <Model>NS20FC</Model>
            <SerialNumber>APM00073701085</SerialNumber>
            <WWN></WWN>
            <Platform></Platform>
            <OS>Dart</OS>
            <OS_Ver>5.6.36-2</OS_Ver>
            <UcodeVer></UcodeVer>
            <EmbedLevel>0</EmbedLevel>
            <InternalMaxSize>0</InternalMaxSize>
            <Comment></Comment>
        </Identifier>
        <Con
```

**SETTING COUNTRY CODE ON MULTITECH MT5634ZBA MODEMS:**

**Note:** If the system ships with an MT5634ZBA modem, it may require the use of a utility or CLI procedure to correct country/regional code settings. This information is taken from the NS Setup Guides.

**Windows ZBAWizard.exe Utility:**

1. Download utility to laptop
2. Connect serial cable between modem and laptop
3. Execute ZBAWizard.exe & enter decryption password “EMC”
4. GlobalWz application queries for modem
5. Select appropriate country/region settings and click finish

**Manually Updating Country/Region settings:**

1. Connect serial cable between modem and laptop
2. Open HyperTerminal session and configure 19200 bits per second, 8 data bits, set parity to None, 1 for stop bit, flow control None, terminal emulation ANSI
3. Verify current country/region setting:  
ATI9 (returns country value in decimal)
4. Set desired country code & verify to make sure the setting was updated:  
AT%T19,0,10 (last digit is modem value in hex, which in decimal is 16 for Japan)  
ATI9

**Note:** Find the applicable country/region code from the appropriate Celerra Setup Guide for MT5634ZMA

**NEW EMC MODEM:**

Multitech MT9234ZBA, 100-602-107 →Replaces the MT5634ZBA 100-602-106

ECO 69503

**TROUBLESHOOTING MODEM ERRORS:**

**NO DIALTONE**

--telephone line might not be connected or phone line cable is bad

--modem might not recognize dial tone if PBX system (can change modem to ignore dial tones)

**BUSY**

--number being dialed is busy, or missing the 9 prefix to access an outside line when dialing

**NO ANSWER**

--answering system will not go “off-hook” (i.e., does not answer), or the number dialed is wrong

**NO CARRIER**

--system answered but did not establish a connection (wrong number, voice number vs. modem, or computer or software was off or faulty, or poor line conditions are causing connection issues)

**General Modem Issues:**

--lack of line connection or bad cable/connector

--wrong dial tone

--busy signal

--wrong number or no modem at the other end

--faulty computer, modem, or software at the other end

--modem incompatibility

--poor line conditions

**CHANGING MODEM RINGCOUNT ANSWERING:**

→Default is to have the modem answer immediately without ringing. To change the default, vi edit the /nasmed/sbin/ch\_mgetty\_monitor script and change following line:

**/sbin/mgetty -n 4 -x 0 \$(basename \$serialdev) > /dev/null 2>&1**

**DISABLING CONNECTHOME FOR NAS UPGRADES, MAINTENANCE, etc:**

**# /nas/sbin/nas\_connecthome -service stop**

Delivery of pending events has been paused.

**# /nas/sbin/nas\_connecthome -service start**

No pending events have been cleared and delivery has been resumed.

**# /nas/sbin/nas\_connecthome -service start -clear**

Any pending events have been cleared and delivery has been resumed.

**Note:** The ability to stop, restart, or clear ConnectHome events is currently undocumented: Not in Release Notes, Cmd Ref Manual, Platform Setup Guides, or Man Pages. AR112733 is getting this documented, and emc173789 shows the recommended way to stop and restart the ConnectHome service. There is also no way to tell if the delivery service has been paused or not—see AR115618. See AR115610 to add capability to pause and restart delivery service in Celerra Manager.

**NAS UPGRADES AND CONNECTHOME:**

→NAS Upgrades will convert whatever is in use to the new ConnectHome format—any errors will be noted in the NAS Upgrade log, verify using /nas/sbin/nas\_connecthome –info; Modem numbers are carried over; Email SMTP server is carried over if set in /etc/sendmail.cf; ESRS Gateway (FTP) IP Address, Username, and Password are carried over—rsc5\*\_nas\_event.cfg file is unloaded and FTP transport enabled

**Note:** NAS Upgrades should be from any prior version to any version above 5.5.31.6, which does not properly preserve custom modem settings, such as the ignore-dialtone, pulse, or tone settings--these must be reset after an upgrade.

→If /nas/sbin/ch\_dd was used to setup pre-5.5.31 systems, the dial-out prefix, and primary/backup phone numbers, will be preserved

→CNS, CFS, & older NS models will need to have Serial Number manually updated after Upgrading to ConnectHome feature

→NAS Upgrades from 5.5.30.x or earlier, to 5.5.31.6 will convert existing CallHome configurations to the new ConnectHome feature. In general, all existing callhome configuration values are preserved if upgrading to 5.5.32 or higher. This is especially significant for those sites outside of the U.S. where default EMC modem settings are usually not applicable. For example, pulse, tone, and ignore-dialtone detection settings are not preserved and will require resetting in the /nas/opt/connectemc/login.cmds file (5.5.31.6 only) and the /nas/opt/connectemc/modem.cfg file (5.5.32.x and later will use this file if there are specific non-default modem settings). Additionally, other modem-specific settings may be impacted. It is generally recommended that the callhome process be tested after a NAS upgrade to ensure the viability of the transport mechanism in use. Please use the platform-specific Celerra Network Server Setup Guide (on Powerlink) as your primary reference for resetting modem settings. Configuration information is also contained in the Setup Guides and Celerra Command Reference Manual. Please note that the Man Pages for nas\_connecthome are missing for 5.5.31.6 and will be present with the 5.5.32.x release

→NAS Upgrades from 5.5.30.x or earlier, to 5.5.31.6, will preserve only the ignore-dialtone detection setting. All other custom modem settings will need to be redefined

→NAS Upgrades from 5.5.31.6 or earlier, to 5.5.32.4, or later, will permanently preserve any custom modem settings in the /nas/opt/connectemc/modem.cfg file, instead of the /nas/opt/connectemc/login.cmds file

→Original files should be preserved in /var/sadm/pkg/emcnas/save directory after the Upgrade

**CONNECTHOME REFERENCES:**

**man pages** →Beginning with NAS 5.5.32 and higher

**Platform-specific Celerra Network Server Setup Guide** →Powerlink & Doc CD

**Celerra Network Server Systems Operations Technical Module** →Powerlink & Doc CD

**Celerra Command Reference Manual** →Powerlink & Doc CD

**Release Notes** →NAS 5.5.31.6

**Powerlink Celerra Tools**>NX4/NX4FC Integrated>"Install" quadrant page>Step 5: Configure for Production>Perform Common

**Post-CSA tasks**>Configure CallHome or ConnectHome telephone numbers

**Resetting ignore-dialtone, Pulse, or Tone dialing for 5.5.31.6:**

**TYPICAL LOGIN.CMDS FILE AS SEEN WITH VI EDITOR:**

```
44    set dial method auto  
45    set dial timeout 300  
46    set speed 9600
```

1. Edit the line 44 in the file /nas/opt/connectemc/login.cmds & add the ignore-dialtone line if required:

set dial method auto →change word "auto" to "tone" or "pulse" to modify [ATDT=Tone; ATDP=Pulse]

set dial timeout 300

set dial ignore-dialtone on [ATX1DT—add this line to the file]

**Resetting ignore-dialtone, Pulse, or Tone dialing for 5.5.32.x:**

1. Modify or create the file /nas/opt/connectemc/modem.cfg:

set dial ignore-dialtone on (add this line if the system does not use a dial-tone)

set dial method tone - or - set dial method pulse (add one of these lines if the modem uses a tone or pulse dial-tone)

**Note:** The login.cmds file is the default file in /nas/opt/connectemc. If custom modem settings existed in prior NAS versions, the modem.cfg file would get created during the upgrade to preserve the special settings. If custom modem settings need to be created for 5.5.32 or higher, the modem.cfg file should be edited to add the appropriate custom settings

**set dial method tone | pulse**

**set dial timeout 300**

**set dial ignore-dialtone on**

**LOCATION OF ORIGINAL MODEM FILE AFTER UPGRADE:**

**/var/sadm/pkg/emcnas/save**

**MODEM SETTINGS FILE:**

**/nas/sys/callhome.modem** (original Celerra CallHome modem settings file)

**/nas/opt/connectemc/login.cmds** (Beginning with NAS 5.5.31.6, this file handles the the modem connection options and provides vax username & password automatically when calling into the EMC vax system.)

**/nas/opt/connectemc/modem.cfg** (Custom modem settings file for NAS 5.5.32.x and above; file is created during upgrades from earlier NAS versions if there was a callhome.modem file, but has to be manually created for new installations if custom settings are required)

**CONFIGURING CELERRA CONNECTHOME (root user only):**

1. CLI interface

2. Celerra Manager>Celerras>Support>Connect Home>select primary delivery mechanism--Email; FTP; or Modem: Set Secondary or Tertiary delivery methods if desired. Can modify and test using Celerra Mgr, as well as disable or enable the Transport mechanism, but cannot clear, stop, or start the service from GUI—must use CLI for this.

**Note:** Must login as Root user to edit the "Connect Home" page

**CONNECTHOME CONFIGURATION FILE:**

→Whenever ConnectHome is changed or configured, either by the GUI or by the nas\_connecthome –modify command, the changes are written to /nbsnas/sys/connecthome.config, then synchronized via cron job to the local IDE drive on the Control Station, which runs every 5 minutes, to /nasmcd/CHomeFiles/connecthome.config

→Since ConnectHome actually uses the local IDE version of the connecthome.config file, there is a potential gap between the time someone changes the config file until the change is actually seen in the ConnectHome

→This issue will be resolved with CMR10, so that both files are updated simultaneously whenever –modify is run

**EXAMPLE:**

**# /nas/sbin/nas\_connecthome -modify -service\_info -email\_to emailalert@emc.com -email\_from**

**default@emc.com**

**# ls -la /nbsnas/sys/connecth\***

**-rw-r--r-- 1 root root 686 Jul 28 15:14 /nbsnas/sys/connecthome.config**

**# ls -la /nasmcd/CHomeFiles/connecth\***

**-rw-r--r-- 1 root root 696 Jul 28 15:09 /nasmcd/CHomeFiles/connecthome.config**

→Running pstree against NAS MCD process shows hierarchy of processes. "Dirsync" is a process designed to sync any file changes on the local /nas partition with the backend /nbsnas partition, every 3 minutes. It also syncs any changes from /nbsnas/sys to the local /nasmcd/CHomeFiles directory, every 5 minutes.

**# pstree -pau 25137 | grep dirsync**

dirsync,30078 /nas/sbin/dirsync /nas /nbsnas 180

dirsync,30118 /nas/sbin/dirsync -c /nas/sys /nasmcd/CHomeFiles 300

## **SOME CONNECTHOME CONFIGURATION RULES:**

- Can have all transport mechanisms enabled at the same time, or none
- To enable or disable transport mechanisms, we are actually assigning a priority using “-modify email\_priority 1, 2, or 3”, or removing the priority setting altogether using the “-modify email\_priority Disabled” syntax
- Only root user can make transport mechanism modifications (CLI, Celerra Mgr, Unisphere)
- Cannot Enable or change a transport method [priority] without first having a valid email address, FTP IP address, or Modem number, and at least one of the mechanisms set to a priority of 1 for the Primary entry

## **NEED PRIMARY EMAIL, IPADDRESS, or MODEM NUMBER SET FIRST BEFORE TRANSPORT CAN BE ENABLED:**

### **1. Modem Number Not Set (Example):**

```
# /nas/sbin/nas_connecthome -modify -modem_priority 2
```

Error 14504231241: The Modem transport cannot be enabled because the Primary Modem Number is not configured.

```
# /nas/sbin/nas_connecthome -infotail
```

Modem :

Priority = Disabled

Primary :

Phone Number =

### **2. Setting Modem Number:**

```
# /nas/sbin/nas_connecthome -modify -modem_number 918005270941
```

### **3. Setting Modem Priority & Verifying:**

```
# /nas/sbin/nas_connecthome -modify -modem_priority 2
```

Ok

```
# /nas/sbin/nas_connecthome -infotail
```

Modem :

Priority = 2

## **DISABLING TRANSPORT MECHANISMS:**

```
# /nas/sbin/nas_connecthome -modify -modem_priority Disabled -ftp_priority Disabled -email_priority Disabled
```

Ok

## **PRIMARY CONNECTHOME APPLICATIONS & FILES:**

→ Main ConnectHome application is /nas/sbin/connectemc, which monitors the /nas/log/ConnectHome directory, delivering all files via the selected Transport mechanism whenever an RSC...xml file is detected

→ General configuration file for ConnectHome is /nas/opt/connectemc/ConnectEMC.ini

→ Transport configuration file is /nas/opt/connectemc/ConnectEMC\_config.xml

## **CONNECTHOME CLI:**

```
# /nas/sbin/nas_connecthome
```

**Note:** ConnectHome consolidates legacy CallHome scripts into one—ch\_dd, ch\_cfg, ch\_dialin\_enable, ch\_dialin\_disable, ch\_test] -info

```
| -test { -email_1 | -email_2 | -ftp_1 | -ftp_2 | -modem_1 | -modem_2 }
```

```
| -modify [ -modem_priority {Disabled | 1 | 2 | 3} ]
```

|                                              |                                                |
|----------------------------------------------|------------------------------------------------|
| [ -modem_number <phone_number> ]             | [ -modem_bt_tymnet { yes   no } ]              |
| [ -modem_number_2 <phone_number> ]           | [ -modem_bt_tymnet_2 { yes   no } ]            |
| [ -ftp_priority {Disabled   1   2   3} ]     | [ -ftp_ipport <ip_addr>[:<port>] ]             |
| [ -ftp_user <username> ]                     | [ -ftp_passwd [passwd] ]                       |
| [ -ftp_folder <path> ]                       | [ -ftp_mode {active   passive} ]               |
| [ -ftp_ipport_2 <ip_addr>[:<port>] ]         | [ -ftp_user_2 <username> ]                     |
| [ -ftp_passwd_2 [passwd] ]                   | [ -ftp_folder_2 <path> ]                       |
| [ -ftp_mode_2 {active   passive} ]           | [ -email_priority {Disabled   1   2   3} ]     |
| [ -email_to {<email_addr> ,<email_addr>} ]   | [ -email_subject <email_subject> ]             |
| [ -email_server { <hostname>   <ip_addr> } ] | [ -email_server_2 { <hostname>   <ip_addr> } ] |
| [ -dial_in_number <phone_number> ]           | [ -serial_number <serial_number> ]             |
| [ -site_id <site_id> ]                       | [ -encryption_enabled { yes   no } ]           |
| [ -dial_in_enabled { yes   no } ]            | [ -help ]                                      |

```
# /nas/sbin/nas_connecthome -service → Switches are currently undocumented and do not display with -help, etc.
```

usage: nas\_connecthome

```
-service { start [-clear] | stop | clear }
```

```
# /nas/sbin/nas_connecthome -info
```

ConnectHome Configuration:

Encryption Enabled = yes

Dial In :

Enabled = yes  
Modem phone number =  
Site ID =  
Serial number = APM00064001123

Email :

Priority = 1  
Recipient Address(es) = szg30@sun1.hosts.pvt.dns  
Subject = CallHome Alert sludge3

Primary :

Email Server = sun1.hosts.pvt.dns

Secondary :

Email Server =

FTP :

Priority = Disabled

Primary :

FTP Server =  
FTP Port = 21  
FTP User Name = onalert  
FTP Password = \*\*\*\*\*  
FTP Remote Folder = incoming  
FTP Transfer Mode = active

Secondary :

FTP Server =  
FTP Port = 21  
FTP User Name = onalert  
FTP Password = \*\*\*\*\*  
FTP Remote Folder = incoming  
FTP Transfer Mode = active

Modem :

Priority = Disabled

Primary :

Phone Number = 918005270941  
BT Tymnet = no

Secondary :

Phone Number = 918006262452  
BT Tymnet = no

## **CALLHOME DIRECTORY REPOSITORY:**

### **/nas/log/ConnectHome**

RSC\_APM00064001123\_102107\_034449696.xml

**Note:** Up to 100 files can be pending in the event directory, processed single-threaded FIFO manner

### **DOES CS KEEP LOCAL COPY OF SUCCESSFUL & FAILED CALLHOME XML FILES?**

#### **ANSWER: YES**

The Control Station keeps a copy of both Successful and Failed CallHome XML files.

### **Location for Successful CallHome Files:**

#### **/nas/opt/connectemc/logs**

**ConnectEMC\_email.log** → Provides timestamp and xml filename of Callhomes delivered (via email)

**ConnectEMC** → Provides timestamp when CallHome was transferred and file appended to

#### **/nas/opt/connectemc/logs/archive**

**ConnectEMC.archive** → Keeps actual copy of successful XML Callhome files

**Note:** Open the file and scroll down to the end to see the latest XML callhome content. View the numbered.archive files to find older Callhome output.

### **Location for Failed CallHomes:**

#### **/nas/opt/connectemc/logs/failed** → Individual XML callhome files that fail are stored in this directory

#### **Connect Home Log File:**

# nas\_logviewer /nas/log/connecthome\_log |tail

Feb 8 11:40:03 2010:CS\_PLATFORM:ConnectHome:INFO:25:::1265647203:ConnectHome event file(s) successfully transferred

**Note:** Useful file as quick reference to see if Callhomes are being transferred, but does not give XML filename, as the ConnectEMC\_email.log does

Other CallHome Information & Logs:

**/nas/opt/connectemc; ls -la poll**

```
lrwxrwxrwx 1 root root 20 Oct 16 15:51 poll -> /nas/log/ConnectHome
```

**# /nas/sbin/nas\_connecthome -service stop | start (-clear) | clear**

**# /nas/sbin/nas\_connecthome -service stop** → Pauses the delivery service for any pending events

Delivery of pending events has been paused.

**Note:** Process takes about 1 minute to stop, and XML callhome files in /nas/log/ConnectHome are removed

**# ps -eafl |grep connect** (processes that run during –service stop)

**/nas/sbin/nas\_connecthome -service stop**

```
000 S root 32690 32689 0 68 0 - 1128 nanos! 13:19 pts/0 00:00:00
```

**/usr/bin/perl /nas/opt/connectemc/.connectemc\_suspend -q -m stop -c suspend**

**# /nas/sbin/nas\_connecthome -service start** → Prompts to clear any pending events and resumes the delivery service

Do you want to clear all the pending event files?(yes/no) :yes

Any pending events have been cleared and delivery has been resumed.

**Note:** If no events are pending, just resumes delivery service

**# /nas/sbin/nas\_connecthome -service start -clear** → Clears pending events from /nas/log/ConnectHome without prompting, & resumes the delivery service [maybe there is a delay, as running this command does not immediately remove the callhome file?]

Any pending events have been cleared and delivery has been resumed.

**# /nas/sbin/nas\_connecthome -service clear** [Pauses, clears pending events, and resumes Callhome delivery service]

None exist. No pending events have been cleared.

**# /nas/sbin/nas\_connecthome -service clear**

Delivery of pending events has been paused.

Any pending events have been cleared and delivery has been resumed.

**/nas/sbin**

**nas\_connecthome** → CLI tool for the ConnectEMC feature

**connectemc** → daemon executable that monitors and delivers event files using the configured transport mechanism(s)

**Note:** Primary file is /nas/sbin/connectemc [/opt/connectemc/connectemc is a link to /nas/sbin/connectemc]

**kermit** → modem control executable

**connectemc\_starter** → process that runs everytime the connecthome service is started

**connecthome\_upgrade** → Script that performs ConnectHome upgrade during NAS Upgrades

**/nas/sys**

**connecthome.config** → Replaces callhome.config

**Note:** The /nas/sys/connecthome.config file contains Dialin modem number, Site number, Serial number, information about all the Transport methods [email, modem, FTP], & encryption status information, and replaces the /nas/sys/callhome.config. A backup copy of this file is stored in /nasmcd/CHomeFiles/connecthome.config.

**# cat connecthome.config** → Use /nas/sbin/nas\_connecthome -info for output

dialInEnabled:yes

modemPhone:5083085000

siteID:s323

celerraSerial:APM00064001123

emailPriority:1

emailFrom:connectemc@emc.com

emailTo:szg30@sun1.hosts.pvt.dns

emailSubject:CallHome Alert

primaryEmailServer:sun1.hosts.pvt.dns

secondaryEmailServer:

ftpPriority:Disabled

primaryFtpAddress:

primaryFtpPort:21

primaryFtpUserName:onalert

primaryFtpPassword:EMCCONNECT

primaryFtpFolder:incoming

primaryFtpTransferMode:active

secondaryFtpAddress:

secondaryFtpPort:21

secondaryFtpUserName:onalert

secondaryFtpPassword:EMCCONNECT

secondaryFtpFolder:incoming  
secondaryFtpTransferMode:active  
modemPriority:Disabled  
primaryModemPhone:918005270941  
primaryModemTymnet:no  
secondaryModemPhone:918006262452  
secondaryModemTymnet:no  
encryptionEnabled:no  
**emailhome.config** → EmailHome header file

### /nas/log

/nas/log/connecthome\_log → Log that records successful CallHome transfers & CallHome tests

# nas\_logviewer -v connecthome\_log |tail

```
logged time      = Dec 3 12:38:26 2008
creation time    = Dec 3 12:38:26 2008
slot id         =
id              = 96108609561
severity        = INFO
component       = CS_PLATFORM
facility        = ConnectHome
baseid          = 25
type            = EVENT
brief description = ConnectHome event file(s) successfully transferred
full description  = The ConnectEMC daemon succeeded in transferring the ConnectHome event file(s) to the configuration destination(s).
recommended action = For details, here are the log files:
- General : /opt/connectemc/logs/ConnectEMC
- Email : /opt/connectemc/logs/ConnectEMC_email.log
- FTP : /opt/connectemc/logs/ConnectEMC_ftp.log
- Modem : /opt/connectemc/logs/ConnectEMC_dial.log
```

**Note:** View this log using nas\_logviewer /nas/log/connecthome\_log -v option to provide verbose output

/nas/log/sys\_log → Control Station System Log showing CallHome attempts and success or failure

**Note:** Use nas\_logviewer for sys\_log and connecthome\_log with -v option to view expanded entries on CallHome events

/nas/log/ConnectHome → ConnectHome RSC xml event files

/nas/log/connectemc → symbolic link to /nas/opt/connectemc/logs [archive ConnectEMC ConnectEMC\_email.log ConnectEMC\_dial.log failed output]

### /nas/opt/connectemc

callhome.ini → Modem initialization file

connectemc.pem → RSA private encryption key

login.cmds → Modem settings in 5.5.31.6 only [changes to modem.cfg file with 5.5.32.x]

connectemc → Symbolic link to /nas/sbin/connectemc executable

connectemc.sh → Sets path to libraries [/opt/connectemc/lib]

logs → Various ConnectHome logs, see next section

rsapub1.key → Encryption key file

connectemc.cer → Certificate

curl-ca-bundle.crt

naslogger\_conduit → Logging

poll → directory containing ConnectHome xml files, but is a symbolic link to /nas/log/ConnectHome

ConnectEMC\_config.xml → Shows the type of transport service that is in effect [Email, FTP, Modem]

<ConnectConfig Type="Email">

kermit → Symbolic link to /nas/sbin/kermit executable

ConnectEMC.ini → General configuration file on log format, log names, log sizes, log archiving, log levels, polling, transport service configuration, archive options, encryption, recycle parameters, etc.

lib

Queue

recycle (directory of failed callhomes that will be retried)

recycle.log (log of retried callhomes with projected retry times)

### /nas/opt/connectemc/logs

archive (directory of ConnectEMC.archive files that captures actual content of XML Callhomes)

Example:

# view /nas/opt/connectemc/logs/archive/ConnectEMC.archive

```
-- Time : 21-10-2007 03:45:53.
-- Filename : /opt/connectemc/poll/RSC_APM00064001123_102107_034449696.xml.
-----output abridged--followed by contents of the RSC xml file---
failed (Archive directory of failed callhome files)
# ls -al /nas/opt/connectemc/logs/failed
RSC_APM00083103123_112608_141558641_ConnectEMC_config.xml
RSC_APM00083103123_112608_141558641.end
RSC_APM00083103123_112608_141558641.xml
ConnectEMC (daemon log)
ConnectEMC_email.log (record of last email attempt)
ConnectEMC_ftp.log (record of last FTP attempt)
ConnectEMC_dial.log (record of last modem attempt)
output
```

### **EXAMPLE UNSUCCESSFUL MODEM CALLHOME:**

# /nas/sbin/nas\_connecthome -test -modem\_1

```
ConnectEMC ConnectEMC 2.0.28-bl75 Tue Oct 16 10:03:40 EDT 2007
EMA API: EMAPI 1.0.8-bl71 Tue Oct 16 10:03:33 EDT 2007
Copyright (C) EMC Corporation 2003-2007, all rights reserved.
```

```
-----Reading configuration file: ConnectEMC.ini.
Run Service begin...
Dial failed for phone number {{82066097}}. Dial Error: ATD82066097(8).
Error 14504231224: Test failed for the Primary Modem.
# tail /nas/opt/connectemc/logs/ConnectEMC
2008-05-16T21:42:44 206 Info ConnectEMC 2.0.28-bl75 rscServiceFramework() Transaction file found.
/opt/connectemc/poll/RSC_APM00071601776_051608_214232274.xml
2008-05-16T21:42:44 201 Info EMAAPI 1.0.8-bl71 RSCSerivceTransactionFound() Routine Entrance.
/opt/connectemc/poll/RSC_APM00071601776_051608_214232274.xml
2008-05-16T21:42:44 508 Info EMAAPI 1.0.8-bl71 ConnectEMCReadConfigFile() Reading ConnectEMC config
file. /opt/connectemc/ConnectEMC_config.xml
2008-05-16T21:42:44 533 Info EMAAPI 1.0.8-bl71 EncryptFiles() Encryption library msg.
encryption complete
2008-05-16T21:42:44 533 Info EMAAPI 1.0.8-bl71 EncryptFiles() Encryption library msg.
2008-05-16T21:42:44 510 Info EMAAPI 1.0.8-bl71 RSCSericeTransactionFound() Attempting to transfer file(s).
TxType: Dial, Retry count: 0, Phone #: 82066097
2008-05-16T21:42:44 201 Info rscApi major:1-0 ConnectEMCUnixDial() Routine Entrance.
2008-05-16T21:48:05 556 Info ConnectEMC 2.0.28-bl75 ConnectEMCUnixDial() Kermit Error Code: 8
```

# tail ConnectEMC\_dial.log

- . SET MODEM HANGUP-METHOD RS232 and try again.
- . If that doesn't work, try again with SET DIAL HANGUP OFF.
- . Use SET DIAL DISPLAY ON to watch the dialog between Kermit and modem.
- . SHOW COMMUNICATIONS, SHOW MODEM, SHOW DIAL to see current settings.
- . HELP SET MODEM, HELP SET DIAL, and HELP DIAL for more information.

(Use SET HINTS OFF to suppress future hints.)

\*\*\*\*\*

Dial failed for phone number {{82066097}}. Dial Error: ATD82066097(8).

# nas\_logviewer -t /nas/log/sys\_log |tail

```
May 16 21:42:32 2008:96108609636::Test event from /nas/sbin/nas_connecthome -test.
May 16 21:42:44 2008:96108609586::Connecthome Test of the Primary Modem initiated
May 16 21:48:06 2008:83223707702::Connecthome Test of the Primary Modem failed
```

# nas\_logviewer /nas/log/connecthome\_log

```
May 16 21:42:32 2008:CS_PLATFORM:ConnectHome:INFO:100:::1210988552:Test event from /nas/sbin/nas_connecthome -test.
May 16 21:42:44 2008:CS_PLATFORM:ConnectHome:INFO:50:::1210988564:Connecthome Test of the Primary Modem initiated
May 16 21:48:06 2008:CS_PLATFORM:ConnectHome:ERROR:54:::1210988886:Connecthome Test of the Primary Modem failed
# tail /nas/opt/connectemc/logs/archive/ConnectEMC.archive [XML CallHome file contents]
```

**/nas/opt/connectemc/logs/failed** [Directory where unsuccessful XML callhomes are stored]

### **EXAMPLE SUCCESSFUL MODEM CALLHOME:**

```
# /nas/sbin/postevent -c 6 -f 133 -i 100 -s 6 →Creates the Test CallHome event
# tail /nas/opt/connectemc/logs/ConnectEMC →Details on connectemc daemon handling of CallHome event
2008-05-16T20:52:03 206 Info ConnectEMC 2.0.28-bl75 rscServiceFramework() Transaction file found.
/opt/connectemc/poll/RSC_APM00071601776_051608_205200844.xml
2008-05-16T20:52:03 201 Info EMAAPI 1.0.8-bl71 RSCSerivceTransactionFound() Routine Entrance.
/opt/connectemc/poll/RSC_APM00071601776_051608_205200844.xml
2008-05-16T20:52:03 508 Info EMAAPI 1.0.8-bl71 ConnectEMCReadConfigFile() Reading ConnectEMC config
file. /opt/connectemc/ConnectEMC_config.xml
2008-05-16T20:52:03 533 Info EMAAPI 1.0.8-bl71 EncryptFiles() Encryption library msg.
2008-05-16T20:52:03 encryption complete
2008-05-16T20:52:03 533 Info EMAAPI 1.0.8-bl71 EncryptFiles() Encryption library msg.
2008-05-16T20:52:03 encryption complete
2008-05-16T20:52:03 510 Info EMAAPI 1.0.8-bl71 RSCSericeTransactionFound() Attempting to transfer file(s).
TxType: Dial, Retry count: 0, Phone #: 82066097
2008-05-16T20:52:03 201 Info rscApi major:1-0 ConnectEMCUinxDial() Routine Entrance.
2008-05-16T20:57:26 556 Info ConnectEMC 2.0.28-bl75 ConnectEMCUinxDial() Kermit Error Code: 8
2008-05-16T20:57:26 511 Warning EMAAPI 1.0.8-bl71 RSCSericeTransactionFound() Transport was unsuccessfull
in sending all files. Status: 1, File: /opt/connectemc/poll/RSC_APM00071601776_051608_205200844.xml.
2008-05-16T20:57:26 510 Info EMAAPI 1.0.8-bl71 RSCSericeTransactionFound() Attempting to transfer file(s).
TxType: Dial, Retry count: 1, Phone #: 82066097
2008-05-16T20:57:26 201 Info rscApi major:1-0 ConnectEMCUinxDial() Routine Entrance.
2008-05-16T20:58:40 512 Info EMAAPI 1.0.8-bl71 RSCSericeTransactionFound() Transport was successfull in
sending all files.
2008-05-16T20:58:40 516 Info EMAAPI 1.0.8-bl71 RSCSericeTransactionFound() Appending TXed files to
archive. Source File(s): /opt/connectemc/poll/RSC_APM00071601776_051608_205200844.xml, Archive File:
ConnectEMC.archive
2008-05-16T20:58:40 119 Info EMAAPI 1.0.8-bl71 RSCAppendFilesToArchive() Parameters passed.
Archive option:S.
```

**# tail /nas/opt/connectemc/logs/ConnectEMC\_dial.log** →Modem connection and results log for CallHome event

AT&F0

OK

ATD82066097

CONNECT 31200 V42bis

Trying callhome (10.15.54.97)... Open

VAX/VMS 5.5-2H4 on FWVAXA - A member of the Fieldwatch VAX Cluster

Authorized Users Only

Username:

Username:

Username:

Username: nas

Password: rsupport

Last interactive login on Wednesday, 16-JUL-1997 09:55

Last non-interactive login on Wednesday, 16-JUL-1997 13:33

FIELDWATCH

FIELDWATCH

EMC Corporation

1. DISPATCH/LOGISTICS

2. REPORTS-PLUS(TM)

3. VOICE-PLUS(TM)

4. FOCUS(TM)

9. FENNER

99. Logoff System

Choice: GOXFER

ENTER PHONE: 5082498610

\*\*B0100000027fed4

Sending: RSC\_APM00071601776\_051608\_205200844.xml

Bytes Sent: 1024/ 2638 BPS:603 ETA 00:02

Bytes Sent: 2048/ 2638 BPS:820 ETA 00:00

Bytes Sent: 2638 BPS:641

Transfer complete

+++

ATQ0H0

Modem hangup OK

# **nas\_logviewer -t /nas/log/sys\_log |tail** →Records ConnectHome events and whether transferred successfully

May 16 20:52:00 2008:96108609636::Test event from /nas/sbin/nas\_connecthome -test.

May 16 20:58:40 2008:96108609561::ConnectHome event file(s) successfully transferred

# **tail /nas/opt/connectemc/logs/archive/ConnectEMC.archive** →Files contain complete copy of successfully transferred XML

```
-- Time : 16-05-2008 20:58:40.
-- Filename : /opt/connectemc/poll/RSC_APM00071601776_051608_205200844.xml.

<?xml version="1.0" encoding="UTF-8"?>
<ConnectHome xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" SchemaVersion="1.0.8">
    <TransType>0010</TransType>
    <TransTypeDesc>EventXML</TransTypeDesc>
    <TransID>0</TransID>
    <Node ID="APM00071601776" State="Online" Status="OK">
        <Identifier>
            <ClarifyID>APM00071601776</ClarifyID>
            <SiteName>999</SiteName>
            <Vendor>EMC</Vendor>
            <DeviceType>CELERRA</DeviceType>
            <Model>NS40FC</Model>
            <SerialNumber>APM00071601776</SerialNumber>
            <WWN></WWN>
            <Platform></Platform>
            <OS>Dart</OS>
            <OS_VER>5.6.37-6</OS_VER>
            <UcodeVer></UcodeVer>
            <EmbedLevel>0</EmbedLevel>
            <InternalMaxSize>0</InternalMaxSize>
            <Comment></Comment>
        </Identifier>
        <Connection>
            <ConnectType>Dial</ConnectType>
            <AccessType>Telnet</AccessType>
            <Version></Version>
            <InPolicy></InPolicy>
            <OutPolicy></OutPolicy>
            <RouterIP></RouterIP>
            <IPAddress>192.1.4.218</IPAddress>
            <IPName></IPName>
            <NAT_IP></NAT_IP>
            <EMC_IP></EMC_IP>
            <State></State>
            <Time></Time>
            <UserID></UserID>
            <AppName></AppName>
            <UdpSocket></UdpSocket>
            <ConnectNum>5082498610</ConnectNum>
            <Port>ttyS0</Port>
        </Connection>
        <HeartBeat>
        </HeartBeat>
        <InternalData>
            <EventList>
                <Event>
```

```

        <SymptomCode>100</SymptomCode>
        <Category>Status</Category>
        <Severity>Info</Severity>
        <Status>OK</Status>
        <Component></Component>
        <ComponentID></ComponentID>
        <SubComponent></SubComponent>
        <SubComponentID></SubComponentID>
        <CallHome>Yes</CallHome>
        <FirstTime>2008-05-16T20:52:00</FirstTime>
        <LastTime>2008-05-16T20:52:00</LastTime>
        <Count>1</Count>
        <EventData>
```

&lt;![CDATA[CCMD ID: 96108609636

Brief Description: Test event from /nas/sbin/nas\_connecthome -test.

Full Description: This event is used to test the current ConnectHome configuration.

Recommended Actions: No action is required.

]]&gt;

```

        </EventData>
        <Description>
```

&lt;![CDATA[COMPONENT: CS\_PLATFORM , FACILITY: ConnectHome

]]&gt;

```

        </Description>
        </Event>
```

&lt;Event&gt;

```

        <SymptomCode>100</SymptomCode>
        <Category>Configuration</Category>
        <Severity>Info</Severity>
        <Status>OK</Status>
        <Component></Component>
        <ComponentID></ComponentID>
        <SubComponent></SubComponent>
        <SubComponentID></SubComponentID>
        <CallHome></CallHome>
        <FirstTime>2008-05-16T20:52:00</FirstTime>
        <LastTime>2008-05-16T20:52:00</LastTime>
        <Count>1</Count>
        <EventData>
```

&lt;![CDATA[APM00071600514

]]&gt;

```

        </EventData>
        <Description>
```

&lt;![CDATA[Backend Storage Serial Number(s)

]]&gt;

```

        </Description>
        </Event>
```

&lt;/EventList&gt;

&lt;/InternalData&gt;

&lt;ExternalFiles&gt;

&lt;/ExternalFiles&gt;

&lt;/Node&gt;

&lt;/ConnectHome&gt;

&lt;/ConnectHome&gt;

# nas\_logviewer -t /nas/log/connecthome\_logtail → Use nas\_logviewer to extract details about ConnectHome events

May 16 20:52:00 2008:96108609636::Test event from /nas/sbin/nas\_connecthome -test.

May 16 20:58:40 2008:96108609561::ConnectHome event file(s) successfully transferred

# nas\_logviewer /nas/log/connecthome\_logtail -2

May 16 20:52:00 2008:CS\_PLATFORM:ConnectHome:INFO:100:::1210985520:Test event from /nas/sbin/nas\_connecthome -test.

May 16 20:58:40 2008:CS\_PLATFORM:ConnectHome:INFO:25:::1210985920:ConnectHome event file(s) successfully transferred

# nas\_logviewer -v /nas/log/connecthome\_log |tail

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!

logged time = May 16 20:58:40 2008  
creation time = May 16 20:58:40 2008  
slot id =  
id = 96108609561  
severity = INFO  
component = CS\_PLATFORM  
facility = ConnectHome  
baseid = 25  
type = EVENT  
brief description = ConnectHome event file(s) successfully transferred  
full description = The ConnectEMC daemon succeeded in transferring the ConnectHome event file(s) to the configuration destination(s).  
recommended action = For details, here are the log files:  
- General : /opt/connectemc/logs/ConnectEMC  
- Email : /opt/connectemc/logs/ConnectEMC\_email.log  
- FTP : /opt/connectemc/logs/ConnectEMC\_ftp.log  
- Modem : /opt/connectemc/logs/ConnectEMC\_dial.log

```
# cat /nas/opt/connectemc/logs/kermitcmds.ksc
take /opt/connectemc/login.cmds
do init
set line /dev/callhome
do predial
dial {{82066097}}
do direct_login
add send-list /opt/connectemc/poll/RSC_APM00071601776_051608_205200844.xml
send /protocol:zmodem\13
pause 5\13
close\13
```

## **VERIFYING MODEM PORT ON CONTROL STATION:**

```
# vi /nas/sbin/modemtst
define DEBUG 1
#/nas/sbin/modemtst → Set to DEBUG mode 1
NAS_DB = /nas
NUM_PORTS = 1
\&@[]: Dimension = 2
0. /nas/sbin/kermit
1. /nas/sbin/modemtst
PENDING_CALLHOME = 0
modemtst: version 1.0
modemtst: readlink found: /dev/callhome --> /dev/ttyS0
modemtst: Current callhome port: /dev/callhome --> /dev/ttyS0
isValidPort: arg = 1
ch_port = 1 port = 1
isValidPort: arg = 1
modemtst: Modem /dev/ttyS0 OK
Port /dev/ttyS0 is valid
isValidPort: arg = 1
portIsOkay: arg = 1 port_status = 0
Port is OKAY. Good exit.
isValidPort: arg = 1
isValidPort: arg = 1
portIsOkay: arg = 1 port_status = 0
```

## **TROUBLESHOOTING MODEM USING MINICOM:**

1. Configure minicom profile for CS modem

# minicom -s

Serial port setup>A>Edit /dev/ttyS1 to read /dev/ttyS0 for the modem serial port  
Accept defaults, 38400, 8, N 1, and HW Flow Control Yes  
Save Setup as <name>

2. Connect to modem and test dial out:

# minicom <name>

atz

OK

atdt82066097

CONNECT 31200 V42bis

Trying callhome (10.15.54.97)... Open

VAX/VMS 5.5-2H4 on FWVAXA - A member of the Fieldwatch VAX Cluster

Authorized Users Only

Password: rsupport

Last interactive login on Wednesday, 16-JUL-1997 09:55

Last non-interactive login on Wednesday, 16-JUL-1997 13:33

FIELDWATCH

FIELDWATCH

EMC Corporation

1. DISPATCH/LOGISTICS

2. REPORTS-PLUS(TM)

3. VOICE-PLUS(TM)

4. FOCUS(TM)

9. FENNER

99. Logoff System

Choice: 99

ANATH logged out at 05/16/08 14:19:18

[Connection to callhome closed by foreign host]

OK

atz

OK

CTRL-A Z for help | 38400 8N1 | NOR | Minicom 2.00.0 | VT102 | Offline

## **CONTROL STATION PROCESSES FOR CONNECTHOME:**

# ps -eafl |grep connect

```
140 S root 1463 1 0 69 0 - 1284 do_sel Aug27 ? 00:00:00 sendmail: accepting connections
000 S root 3531 1689 0 69 0 - 373 nanosl Aug27 ? 00:00:00 /nas/sbin/log_trimmer -n /nas/log/connecthome_log 2000 1 6 h t 10 n
000 S root 4757 1689 0 69 0 - 566 wait4 Aug27 ? 00:00:14 /bin/bash /nas/sbin/connectemc_starter
000 S root 5243 4757 0 69 0 - 1838 nanosl Aug27 ? 00:00:00 /nas/opt/connectemc/connectemc
```

**Note:** The connectemc process replaces the old ch\_monitor, while the kermit process replaces the ch\_app process

## **CONFIGURING CONNECTHOME--CELERRA MANAGER/UNISPHERE:**

**Note:** /nas/log/cel\_api.log records fact that a modification was done, and what the host IP was, but does record the actual modification

**Celerras>Support>ConnectHome tab** (Celerra Mgr)

**Unisphere > System > Service Tasks: Manage Connect Home** (Root user Scope local only!)

### **MISCELLANEOUS ITEMS:**

→Can modify Site ID, Celerra Serial Number, Enable/Disable Dial-in number, Enter/Modify Dial-in modem number, Enable or Disable Encryption, and can Test any of the ‘configured’ transport methods [needs to be configured to showup on the list]

**Note:** Missing serial number will result in UNKNOWN callhomes. Serial number is automatically detected for later NS models, but manual updates will be required for CNS, CFS, & older NS models.

### **MISC. DEFAULTS:**

Dial-in Enabled & Encryption are enabled by default—all other values blank until set

### **EMAIL ITEMS:**

→Can modify Priority (Disabled, Primary, Secondary, Tertiary) of the transport method, specify Primary & Secondary SMTP server, Subject line of email, specify multiple Recipient Address(es) with comma separation

### **EMAIL DEFAULTS:**

Default Sender is [connectemc@emc.com](mailto:connectemc@emc.com), Default Recipient is [emailalert@emc.com](mailto:emailalert@emc.com), Default Subject Line is “CallHome Alert”—all other values blank until set

### **FTP ITEMS:**

→Can modify Priority, specify Primary & Secondary FTP Server, Port to use for FTP, Username, Password, default Remote Location path, Transfer Mode (Active or Passive)

**FTP DEFAULTS:**

Default FTP user account is “onalert”, Default FTP password is “EMCCONNECT”, Default FTP destination path is “incoming”, and Default transfer mode is Active—all other values blank until set

**MODEM ITEMS:**

→Can modify Priority, specify Primary & Secondary Modem Number, Enable or Disable BT Tymnet

**MODEM DEFAULTS:**

Default setting for BT Tymnet is no—all other values blank until set

**CONFIGURING CONNECTHOME FROM CLI:**

**SETTING DIAL-IN MODEM NUMBER, SITE ID, SERIAL NUMBER, or ENCRYPTION(payload):**

# /nas/sbin/nas\_connecthome -modify -dial\_in\_number 15082498600 | -serial\_number | -site\_id | -encryption\_enabled yes | no

Ok

# /nas/sbin/nas\_connecthome -info |head

ConnectHome Configuration:

Encryption Enabled = no

Dial In :

Enabled = yes

Modem phone number = 15082498600

Site ID = s323

Serial number = APM00064001123

**ENABLING or DISABLING DIAL-IN MODEM NUMBER:**

# /nas/sbin/nas\_connecthome -modify -dial\_in\_enabled yes | no

Ok

**SETTING TRANSPORT DESTINATIONS:**

**Configuring Primary & Secondary Email Destinations:**

# /nas/sbin/nas\_connecthome -modify -email\_to

emailalert@emc.com,szg30@sun1.hosts.pvt.dns,szg30@sun2.hosts.pvt.dns -email\_from admin@emc.com

-email\_subject 'Email Callhome Alert' -email\_server sun1.hosts.pvt.dns -email\_server\_2

sun2.hosts.pvt.dns

Ok

# /nas/sbin/nas\_connecthome -info |grep -C3 Email

Email :

Priority = 1

Sender Address = [admin@emc.com](mailto:admin@emc.com) →Set using CLI only (later NAS versions can set in GUI)

Recipient Address(es) = emailalert@emc.com,szg30@sun1.hosts.pvt.dns,szg30@sun2.hosts.pvt.dns

Subject = Email Callhome Alert

Primary :

Email Server = sun1.hosts.pvt.dns

Secondary :

Email Server = sun2.hosts.pvt.dns

**Note:** Multiple email recipients allowed using comma separation between addresses, no spaces, via CLI only—EMC recommendation to keep [emailalert@emc.com](mailto:emailalert@emc.com) as the default destination address. The Callhome Email message body is Base64 encoded & is not human readable. Default Email Subject line is ‘Callhome Alert’. Default Email Sender is ‘connectemc@emc.com’. Default action is to send all email in Base 64, and encrypted. Disable encryption to decode email Base 64 messages for testing.

**Configuring Primary & Secondary FTP Server Destinations (to ESRS Gateway):** →Sets Primary FTP server & port (if specified), and Secondary FTP server & port

# /nas/sbin/nas\_connecthome -modify -ftp\_ipport 168.159.216.19 -ftp\_ipport\_2 192.1.4.248

Ok

# /nas/sbin/nas\_connecthome -info |grep -i ftp

FTP :

FTP Server = 168.159.216.19

FTP Port = 21

FTP User Name = onalert

FTP Password = \*\*\*\*\*

FTP Remote Folder = incoming

FTP Transfer Mode = active

FTP Server = 192.1.4.248

FTP Port = 21  
FTP User Name = onalert  
FTP Password = \*\*\*\*\*  
FTP Remote Folder = incoming  
FTP Transfer Mode = active

### **Modifying FTP User & Password:**

# /nas/sbin/nas\_connecthome -modify -ftp\_user anonymous -ftp\_passwd onalert@emc.com

Ok

**Note:** Default ftp user is “onalert” and default ftp password is “EMCCONNECT”

### **Configuring Primary & Secondary Modem Destinations:**

# /nas/sbin/nas\_connecthome -modify modem\_number 9,18885552222 -modem\_number\_2 9,18885553333

Ok

### **Example of EMC Destination Modem Numbers for ConnectHome Service by Country:**

|       | <b>AT&amp;T 1-800</b> | <b>AT&amp;T Local</b> | <b>Verizon</b>              | <b>BT 1-800</b> |
|-------|-----------------------|-----------------------|-----------------------------|-----------------|
| U.S.  | 866 285 3258          | 508 435 0987          | 800 626 2452 & 800 527 0941 | n/a             |
| U.K.  | 0800 028 7702         | 0203 024 4653         | 0800 169 6651               | n/a             |
| Japan | 00531 11 5188         | 035 767 9355          | 00531 12 0466               | n/a             |
| Ger   | 0800 182 7792         | 0695 170 9349         | 0800 100 9325               | n/a             |

<http://www.cs.isus.emc.com/csweb2/udh/countrytollfree.htm>

### **Setting Transport Method Priority:**

# /nas/sbin/nas\_connecthome -modify -email\_priority 1 -ftp\_priority 2 -modem\_priority 3

# /nas/sbin/nas\_connecthome -modify -ftp\_priority 2

Ok

# /nas/sbin/nas\_connecthome -info |grep -C1 FTP

FTP :

Priority = 2

# /nas/sbin/nas\_connecthome -modify -modem\_priority 2

Ok

### **Disabling Transport Method Priority:**

# /nas/sbin/nas\_connecthome -modify -ftp\_priority Disabled

Ok

# /nas/sbin/nas\_connecthome -info |grep -C1 FTP

FTP :

Priority = Disabled

# /nas/sbin/nas\_connecthome -modify -modem\_priority Disabled

### **Removing a Transport Method from the Configuration:**

# /nas/sbin/nas\_connecthome -modify -ftp\_ipport "" -ftp\_ipport\_2 ""

Ok

# /nas/sbin/nas\_connecthome -modify -ftp\_ipport ""

Error 14504231237: The Primary FTP Server cannot be unconfigured because the FTP transport is enabled.

**Note:** Must disable the transport method using -modify -ftp\_priority Disabled before the transport can be removed

# /nas/sbin/nas\_connecthome -modify -email\_server\_2 ""

# /nas/sbin/nas\_connecthome -modify -modem\_number ""

Ok

**Note:** Be aware of syntax difference between the Primary Modem number and the Secondary

# /nas/sbin/nas\_connecthome -modify -modem\_number\_2 ""

**Note:** Use double quotations in place of the FTP IP address, Email address, or Modem number that is to be removed from the configuration

### **Testing a Transport Service (Testing Celerra ConnectHome):**

# /nas/sbin/nas\_connecthome -test -email\_1 [or -ftp\_1, -modem\_1, etc.]

-----  
Reading configuration file: ConnectEMC.ini.

Run Service begin...

Test succeeded for Primary Email.

### **Logs Showing Successful ConnectHome Email Transfer:**

**# tail /nas/opt/connectemc/logs/ConnectEMC\_email.log**

# tail ConnectEMC\_email.log

[2008-05-22 13:06:29] \*\*\*\*\* New SMTP Session started \*\*\*\*\*

[2008-05-22 13:06:29] Sender Name: connectemc@emc.com, Recipient Name: whatsa@sun1.hosts.pvt.dns.

[2008-05-22 13:06:29] Email Server: 192.1.4.210, File count: 1.

[2008-05-22 13:06:29] Subject: CallHome Alert.

[2008-05-22 13:06:29] Sending /opt/connectemc/poll/RSC\_APM00071601776\_052208\_130617084.xml.en2

[2008-05-22 13:06:29] Email sent to mail server successfully

**# tail /nas/opt/connectemc/logs/ConnectEMC**

2008-05-22T13:06:29 510 Info EMAPI 1.0.8-bl71 RSCSericeTransactionFound() Attempting to transfer file(s).

TxType: Email, Retry count: 0, Addr:whatsa@sun1.hosts.pvt.dns, Server:192.1.4.210

2008-05-22T13:06:29 658 Info Email ConnectEMCEmail() Email send success

whatsa@sun1.hosts.pvt.dns

2008-05-22T13:06:29 658 Info Email ConnectEMCEmail() Email send success

220 sun1.hosts.pvt.dns ESMTP Sendmail 8.11.7p1+Sun/8.11.7; Thu, 22 May 2008 12:55:37 -0400 (EDT)

HELO localhost

250 sun1.hosts.pvt.dns Hello [192.1.4.218], pleased to meet you

MAIL FROM:<connectemc@emc.com>

250 2.1.0 <connectemc@emc.com>... Sender ok

RCPT TO:<whatsa@sun1.hosts.pvt.dns>

250 2.1.5 <whatsa@sun1.hosts.pvt.dns>... Recipient ok

DATA: Sending...

354 Enter mail, end with "." on a line by itself

250 2.0.0 m4MGtb111491 Message accepted for delivery

**# nas\_logviewer /nas/log/sys\_log |tail**

May 22 13:06:16 2008:CS\_PLATFORM:ConnectHome:INFO:100:::1211475976:Test event from /nas/sbin/nas\_connecthome -test.

May 22 13:06:28 2008:CS\_PLATFORM:ConnectHome:INFO:50:::1211475988:Connecthome Test of the Primary Email initiated

May 22 13:06:30 2008:CS\_PLATFORM:ConnectHome:NOTICE:51:::1211475990:Connecthome Test of the Primary Email

succeeded

**# nas\_logviewer -v /nas/log/connecthome\_log|tail**

argument value = Primary Email

argument type = 8

brief description = Connecthome Test of the Primary Email succeeded

full description = The /nas/sbin/nas\_connecthome CLI Test succeeded. The ConnectEMC daemon succeeded in transferring the ConnectHome test event file to its configured destination.

recommended action = For details, see the following:

- General : /nas/log/connectemc/ConnectEMC
- Email : /nas/log/connectemc/ConnectEMC\_email.log
- FTP : /nas/log/connectemc/ConnectEMC\_ftp.log
- Modem : /nas/log/connectemc/ConnectEMC\_dial.log

### **Example of Successful ConnectHome FTP Transfer:**

**# /nas/sbin/nas\_connecthome -test -ftp\_1**

-----  
ConnectEMC 2.0.27-bl18 Wed Aug 22 10:24:32 EDT 2007

RSC API Version: 2.0.27-bl18

Copyright (C) EMC Corporation 2003-2007, all rights reserved.

-----  
Reading configuration file: ConnectEMC.ini.

Run Service begin...

\* About to connect() to 168.159.216.19 port 21

\* Trying 168.159.216.19... \* connected

\* Connected to 168.159.216.19 (168.159.216.19) port 21 -----output abridged-----

USER anonymous

< 331 Enter your email address as your password

```
> PASS onalert@emc.com
< 230 Password accepted
> PWD
< 257 "/" is your current location
* Entry path is '/'
> CWD incoming
< 250 OK. Current directory is /incoming -----output abridged-----
> STOR RSC_APM00064001123_111207_161635666.xml
< 150 Accepted data connection
* Remembering we are in dir incoming/
< 226-61952.1 Mbytes free disk space
< 226-File successfully transferred
< 226 0.002 seconds (measured here), 0.99 Mbytes per second
* Connection #0 to host 168.159.216.19 left intact
> QUIT
```

### **Testing Modem Dialout:**

1. First make sure the TTY port is functioning correctly:

```
# /nas/sbin/modemst
modemst: Modem /dev/ttys0 OK
```

2. Then test ConnectHome modem configuration:

```
# /nas/sbin/nas_connecthome -test -modem_1
```

Wed May 21 07:47:42 EDT 2008

```
-----  
ConnectEMC ConnectEMC 2.0.28-bl75 Tue Oct 16 10:03:40 EDT 2007  
EMA API: EMAPI 1.0.8-bl71 Tue Oct 16 10:03:33 EDT 2007  
Copyright (C) EMC Corporation 2003-2007, all rights reserved.
```

-----  
Reading configuration file: ConnectEMC.ini.

Run Service begin...

Test succeeded for Primary Modem.

## **CELERRA PLATFORMS BASED ON FLEET CLARIION BACKENDS:**

**NAS FOXGLOVE (NS-960, NS-960FC, NS-G8, NS-960iS):** High-end enterprise product, GA March 2009

- Supported with NAS 5.6.43 and Flare 28++ Mira release
- Foxglove consists of the NS-G8 Gateway, the NS-960, NS-960FC, & NS-960iS Integrated models
- Foxglove is not customer installable
- Replacement for NS80, NSX, and NS80G platforms
- First NAS system to be built on both Backend and Frontend CX4-based hardware
- Integrated systems now referred to as the “Unified Storage” marketing solution
- NS-960 Integrated ships in EMC-racked configurations only, no Customer-racked systems, with 2-8 Blades, 1-2 Control Stations
- The NS-G8 Gateway ships in EMC racks or can be racked in Customer racks, 2-8 Blades, 1-2 CS
- The NS-960iS ships with 2-4 Data Movers, supporting maximum of 480 drives (not 960 drives)
- NS-960 integrated model supports 2-8 Blades, 960 drives maximum
- NS-960FC model supports 2-6 Blades, 960 drives for use with FC Hosts

**Note:** With NAS 5.6.45 and the approved upgrade to add a 5<sup>th</sup> FC IO Module to each SP, the NS-960FC can support 8 Blades and external FC Hosts

→ The Blade Wildcat-S 4U CPU Chassis, is the first CPU chassis to be used by CLARiiON, SYMMETRIX, & NAS [Symmetrix Tigon; Clariion Dreadnought; NAS Foxglove] as part of the Common SPO convergence project

→ NS-G8 & NS-960 will ship only with Wildcat-S XBlades and always with the Tornado IO SLIC annex

**Note:** The Wildcat-S Clovertown CPU chassis is being replaced by the Wildcat-S VE Harpertown board in Fall 2009 timeframe (CMR10) for both Celerra and Clariion. Board is supposed to be 100% compatible with Wildcat-S for mixed systems.

→ Foxglove as an Xblade can be used in NS80, NSX, and NS960 systems

**Note:** Xblade types cannot be mixed within the same Enclosure for NSX/NS80, just same overall Cabinet

→ For NSX, NS80/NS80G, use of Foxglove Xblade requires swap out of Enclosures and all blades to support Foxglove

→ Foxglove Blades are called X-Blades (no name assigned, though previous materials called them X-Blade90)

→ Based on CLARiiON Dreadnought hardware--follow-on product to XBlade-65 architecture based on Common SPO “Wildcat-S”

→ Replacement for NS80 & NSX models, total of 8 blades (NS-G8 replaces NSX)

→Use of IO Modules, SLICs (Small I/O Cards), with (4) SLIC slots per Blade, and an additional (2) SLIC slots per Blade in an IO Annex slot called the Tornado I/O module carrier

**Note:** At GA, only one of the two slots in the IO Annex is used for Foxglove Blades (other slot will be used later on). In each Blade enclosure, the Annex slots (4 & 5) on the lefthand side, when viewed from the rear of the cabinet, are owned by the Upper blade (blades 3, 5, 7, 9), while the Annex slots (4 & 5) on the righthand side, are owned by the Lower blade (blades 2, 4, 6, 8).

→Earthquake management switch module uses Coldfire firmware that will be upgraded by NAS, and to be backwards compatible with Typhoon and Scorpion management switches

→Foxglove NS-G8 to support up to 4-backend arrays, and supports CX300/CX400/CX500/CX600/CX700, CX3 series, CX4 series, Symm 5, DMX1-3, Tigon arrays

#### **Titan Cabinet:**

→NS-RACK-60U cabinet for single or dual 24-AMP service

→Each rack holds (4) PDU's (Power Distribution Units—similar to mounted power strips in the rack) and (4) PDP's (Power Distribution Panels—external power attaches to PDP for PDUs)

→Power cord North America PW40U-60-US (Single phase L6-30 Plug) 208V 24A

→Power cord Europe PW40U-60-IEC3 (IEC 309 332 Plug) 230V 30A

→If calculated cabinet power requirements are under 24A, connect upper PDU's to lower PDP's

→If calculated cabinet power is >than 24A, do not disconnect upper PDU's

#### **Unified Storage Configurations (NS-960; NS-960FC; NS-960iS):**

→8-Blades for NAS-only configuration (requires no external FC ports)

→6-Blades for NAS and FC configuration where (2) FC ports/SP are dedicated for external Hosts

→4 Blades for NAS and FC configuration where (4) FC ports/SP are used for external Hosts

→2 Blades for NAS and FC configuration where (6) FC ports/SP are used for external Hosts

#### **SUPPORTED STORAGE CAPACITIES:**

→896TB IP systems (7 blades \* 128TB/blade for NS-960 & NS-G8)

**Note:** Foxglove is technically limited to about 480TB support with 2TB LUNs on the Celerra. CMR7 increases LUN size support to 16TB. Currently, the CX4 platform is limited to 256 LUNs in a single Storage Group, making it possible for a system shipping with 5.6.43 to reach the LUN limit and hence not be able to use additional disks on the backend. Maximum storage capacity of the CX4-960 is currently about 760TB.

→Supports 960 drives for MPFS over FC

→Supports 480 drives for MPFS over iSCSI

#### **SUPPORTED DRIVES FOR NS-960:**

146GB 15K FC drives

300GB 15K FC drives

400GB 10K FC drives

450GB 15K FC drives

1TB 7200rpm SATA drives [1 HS per 15 drives]

1TB 5400rpm SATA drives

73GB EFD drives (max. of 30/system, 1HS for every 30 drives, no mixing of drive types, RAID5 4+1 or 8+1 for Celerra, 5.6.43)

#### **SUPPORTED RAID PROTECTION:**

##### **Fibre Channel:**

RAID1/0, 2 luns, clar\_r10

RAID5 4+1 or 8+1, 2 luns, clar\_r5\_performance & clar\_r5\_economy, respectively

RAID6 4+2 or 6+2, 2 luns, clar\_r6

RAID 6 12+2, 4 luns, clar\_r6

##### **SATA:**

RAID1/0, 2 luns, clarata\_r10

RAID5 4+1 or 6+1, 1 lun, clarata\_archive

RAID5 8+1, 2 luns, clarata\_archive

RAID6 4+2 or 6+2, 2 luns, clarata\_r6

RAID 6 12+2, 4 luns, clarata\_r6

##### **EFD:**

RAID5 4+1 or 8+1, 2 luns, clarssd\_r5

#### **FOXGLOVE BLADE SPECS:**

→Dual socket, four core Intel 2.33GHz processors with Dual 1.33GHz FSB buses

→Blades capable of 4/8/16/32GB memory, Foxglove will ship with 8GB physical RAM (but only 4GB useable by pre-Barossa NAS)

\$ server\_sysconfig server\_2 -P

Total main memory (MB) = 4090

→(4) SLIC IO slots onboard, with another two per IO Annex (only one of the two slots used at GA for Blades) Lower left Annex is for the Top blade in an enclosure (3, 5, 7, 9), while lower right Annex is for Bottom blade in the enclosure (2, 4, 6, 8).

→128TB per blade for Cognac, 256TB/Blade Barossa

→Fall 2009 the DIMM chipsets are changing from 512Mbit DDR to 1Gbit DDR, meaning that for DIMM replacements, the 1Gbit memory sticks must be replaced in pairs (replace all DIMMS with 4 pairs prior to NAS 5.6.xxx, and after 5.6.xxx, only need to replace a single pair of DIMMs)

#### **Blade Replacement Comment:**

→There are two FRU replacement parts for the Blades. The DIMMs (100-562-959), and the Blade Assembly with CPU Motherboard (103-800-002C. The latter does not come with DIMMs or SLICs. So, you would order the Blade Assembly for a CPU Module failure, then transfer DIMMs and SLIC IO modules from one Blade assembly to the other.

# /nas/sbin/enclosure\_status -e 2 -v

Failed FRU CPU Module Failed

#### **CX4-960 SP SPECS:**

→Dual quad-core Xeon 2.66GHz processors per SP using 16GB memory

→Default Read cache setting 512MB, default Write cache approx. 4.5GB per SP

→Max 4096 LUNs and 1024 Storage Groups per Array

#### **CSA COMPATIBILITY:**

→Support for Foxglove in CSA

→Support for single or dual Control Station systems

**Note:** Cable check has additional checks for secondary CS, additional Blades, additional management switches, and checks to make sure that BE and FE IO ports on the SPs match required configurations

#### **BLADE/SP SLIC IO MODULES:**

--Tomahawk [103-054-100c] is a 4-port 1/2/4Gb FC IO Module (PMC Tachyon QE4) capable of CU & Optical SFP modules

**Note:** Copper SFPs would be used for connectivity to the Backend DAEs

--Harpoon [103-053-100a] is a 2-port 1Gb iSCSI IO Module using 10/100/1000 Mbps Ethernet RJ45 connectors & HW iSCSI offload

--Thunderchild is a 4-port 10/100/1000 BaseT GbE copper

--Thunderbolt is a 4-port GbE card with (2) Optical and (2) Copper GbE ports

--Firestorm is a single port Neterion 10GbE IO card

--SLICs cannot be hot-swapped or added to the Array without SP reboots (alternating reboots per NST)

--There are a total of (5) different SLIC IO modules between the SPs and Celerra

--Tomahawk Fibre Channel module is used by both SPs and Blades

--iSCSI Harpoon module is used only by the SPs

--Thunderchild, Thunderbolt, and Firestorm Ethernet modules are used by Blades only

#### **(8) SUPPORTED SLIC CONFIGURATIONS FOR FOXGLOVE BLADE ENCLOSURES:**

→(8) Basic IO module configurations are allowed

→All Blades in the cabinet must have the same identical SLIC configuration

→Each Blade can support up to (4) IO modules in the Data Mover assembly, and an additional (2) slots per Blade in a separate IO Annex (Up to 6 total SLIC cards per Blade)

→IO Annex A & B—Tornado Annex allows two PCI-Express slots for four additional IO modules (though only one slot used at GA)

**Note:** Lower left Annex is for the Top blade in an enclosure (3, 5, 7, 9), while lower right Annex is for Bottom blade in the enclosure (2, 4, 6, 8).

→SLIC Slot\_0 will always have the 4-port FC IO card for booting the blades

→SLICs will not be recognized without a system reboot

A →Single 4-port Tomahawk card, Two 4-port Thunderchild copper GbE cards (NS960DM-8A)

A1 →Single 4-port Tomahawk card, Four 4-port Thunderchild copper GbE cards (NS960DM-8A1)

B →Single 4-port Tomahawk card, Two 4-port Thunderbolt GbE cards (4 ports optical & 4 copper) (NS960DM-8B)

B1 →Single 4-port Tomahawk card, Four 4-port Thunderbolt GbE cards (8 ports optical & 8 copper) (NS960DM-8B1)

C →Single 4-port Tomahawk card, Two single port Firestorm 10GbE cards (NS960DM-8C)

C1 →Single 4-port Tomahawk card, Four single port Firestorm 10GbE cards (NS960DM-8C1)

C3 →Single 4-port Tomahawk card, Two single port Firestorm 10GbE cards, Two Thunderbolt 4-port GbE cards (NS960DM-8C3)

C5 →Single 4-port Tomahawk card (slot\_0), Two single port Firestorm 10GbE cards (slot\_1, slot\_2), Two Thunderchild 4-port GbE cards (slot\_3, slot\_4) (NS960DM-8C5)

**Note:** Model number convention for NS-G8 the same, just substitute NSG8 for each of the 8 IO configurations

#### **BLADE IO MODULE MODEL, PART NUMBER, NAME:**

NS960DM-8A1, 103-055-100B, Thunderchild 4-port copper GbE

#### **SLIC PORT NAMING CONVENTION:**

→<port\_type> + <SLIC\_slot\_position from 1-to-r> + <port\_number\_within\_the\_slot>

→Blade SLIC slots are numbered from left-to-right: 0, 1, 2, 3, 4

#### **Examples:**

#### **SLIC Configuration “C” (Single Tomahawk (Slot 0) and Two Firestorm cards (Slots 1 & 2)):**

BE-0-0; BE-0-1; AUX-0-0; AUX-0-1 (Slot 0)

FXG-1-0 (Slot 1)

FXG-2-0 (Slot 2)

**SLIC Configuration “A” (Single Tomahawk (Slot 0) & Two Thunderchild (Slots 1 & 2)):**

BE-0-0; BE-0-1; AUX-0-0; AUX-0-1 (Slot 0)  
CGE-1-0; CGE-1-1; CGE-1-2; CGE-1-3 (Slot 1)  
CGE-2-0; CGE-2-1; CGE-2-2; CGE-2-3 (Slot 2)

**SLIC SIGNATURE DURING NAS INSTALL OR UPGRADES:**

→setup\_slot –init generates IO module signature for the Blade, stored in following files:  
/nas/server/slot\_x/device  
/nas/dos/slot\_x/boot.cfg

**Example:**

```
# cat /nas/server/slot_3/devicelgrep slic  
slicsignature validate=131585
```

**FOXGLOVE ENCLOSURES (Chassis):**

**DME0 ENCLOSURE:**

→The NAS DME Enclosure(s) will use the Wildcat-S CPU, Chassis, & midplane, (2) Power Supplies, (4) Blowers, (2) Earthquake Management switch modules, (2) Blades with various SLIC IO module configurations, and (2) Tornado Annexes  
DME0 is 100-520-839 as the FRU  
DME1 is 100-520-840 as the FRU

**Note:** Enclosure replacements will provide only the midplane chassis—all existing DME hardware will be transferred to the new chassis

→Lower lefthand LAN port on Side A Management Switch always plugs into CS Port A for ETH0 primary Internal Network LAN  
→Lower lefthand LAN port on Side B Management Switch always plugs into CS Port B for ETH2 secondary Internal Network LAN  
→Upper lefthand LAN port on DME0 Side A Management Switch is used to cross-connect to additional DMEs (lower lefthand LAN port on the A side), and then to the Upper Mgmt LAN port on SPA’s Management Switch from the last DME in the loop  
→Upper lefthand LAN port on DME0 Side B Management Switch is used to cross-connect to additional DMEs (lower lefthand LAN port on the B side), and then to the Upper Mgmt LAN port on SPB’s Management Switch from the last DME in the loop  
→Lower righthand LAN port on DME0 Side A Management Switch is used to connect to the primary internal LAN port A on CS1  
→Lower righthand LAN port on DME0 Side B Management Switch is used to connect to the secondary internal LAN port B on CS1

**NAS EARTHQUAKE MANAGEMENT MODULES:**

→Firmware upgraded during NAS upgrades, and to be backwards compatible with earlier switches  
→Enclosure ID controlled by hardware

**(3) LAN Management ports**

Lower righthand LAN port (DME0 only) would be used to connect to either A (Primary) or B (Secondary) network port on CS1  
Lower lefthand LAN port (DME0 only) would be used to connect to either A (Primary) or B (Secondary) network port on CS0  
Upper lefthand LAN port would be used to daisychain to lower lefthand port on each additional Data Mover enclosure, and then to connect from last DME to the upper LAN management port on the SPs (depending on SPA or SPB side of cabinet)

**Single Serial Maintenance port**

(1) serial mini-db9 port for maintenance

**LEDs**

Fault LED, Power LED, 7-segment display showing enclosure ID

**SPE ENCLOSURE:**

→Solar Flare Management Module: Top RJ45 LAN port is the Management LAN port; Bottom RJ45 LAN port is the Service LAN  
→Bottom mini-db9 serial port is for connectivity to SPS A or B, while top mini-db9 serial port is the Maintenance serial port  
→The CLARiiON SPE Enclosure is similar (NS-G8), but will use different SLIC IO module configurations, and a Solar Management Module for each SP

Clariion SPE is 100-562-918 Dreadnought

→The CLARiiON SPE AUX chassis (NS-960, 960FC, 960iS) is similar to the regular SPE chassis, except that it is configured for use with the Integrated Celerra systems--the RESUME prom is backed up to the PSM database, and upon Enclosure replacement, FLARE will copy the original RESUME information over the replacement chassis’s RESUME

Clariion AUX SPE is 100-520-841

**CLARiiON SOLAR FLARE MANAGEMENT MODULES:**

**LAN Ports**

(2) LAN ports: Top LAN port is the Management LAN, Bottom LAN port is the Service LAN port (laptop connection)

**Other Ports**

Bottom serial mini-db9 port used to connect to SPS A or SPS B

Top serial mini-db9 port used for maintenance

(1) USB port

**LEDs**

Fault & Power LED

## **BACKEND FC Port Rules:**

- Each Backend Loop supports up to 8 DAEs (8@15=120 drives/loop) and uses 1 FC port per SP per Loop
- NS-960 supports AUX or AUXF arrays, the latter only when no FC Enabled license is ordered
- NS-960FC supports the AUXF array, with up to 8 Backend Loops [8 DAEs/Loop =64 DAEs total, yielding 960 drives over 8 SP FC Ports], and up to (6) Blades to allow for external FC Host connectivity
- NS-960iS supports the AUXI array, with up to 4 Backend Loops @8 DAE's each=32DAE's@15drives=480 (4 FC ports/SP for 4 Loops) and 4 Blades
- Each Blade uses 1 FC port per SP (8 Blades would require 8 FC ports/SP) and plug into FC ports 2 & 3 on the SPs

## **FOXGLOVE ARRAY MODELS:**

### **ORDER ENTRY NICKNAMES:**

AUX

AUXF

AUXI

### **TRUE ARRAY SYSTEM MODEL:**

CX4-960

- There are (3) Array configurations allowed with Foxglove Integrateds: AUX, AUXF, and AUXI
- From a CLI or GUI system perspective, the array will always be seen as a “CX4-960”
- The AUX SP configuration uses (2) 4-port FC IO SLIC modules/SP, for a total of 8 FC ports/SP
- The AUXF SP configuration uses (4) 4-port FC IO SLIC modules/SP, for a total of 16 FC ports/SP
- The AUXI SP configuration uses (2) 4-port FC IO SLIC modules/SP, and (4) 2-port iSCSI IO modules/SP (two modules onboard the SP and two modules in each SP’s IO Annex), for a total of 8 FC ports/SP and 8 iSCSI ports/SP
- IO Annex slot position is 4 & 5 for each SP [An SP with all SLIC slots populated are numbered 0-5, from left-to-right]

## **FOXGLOVE PSN (PRODUCT SERIAL NUMBER):**

- The Foxglove platform uses the new PSN as the Celerra Serial Number
  - Earlier platforms used the EMC\_SERIAL\_NUMBER field from the Data Mover or Enclosure\_0 Resume to determine Celerra Serial Number
  - During factory configuration, the PSN is written to Enclosure\_0’s Resume Prom as  
EMC\_PRODUCT\_SERIAL\_NUMBER="FNM00083800203" and also to the /etc/product\_numbers.db file
- # cat /etc/product\_numbers.db  
#Product Numbers Database file  
#This file contains the Product Serial Number(PSN) and Product Part Number(PPN).  
Version=1.0

**PSN=FNM00083800203**

PPN=900-525-001

# /nas/sbin/serial

FNM00083800203

**Note:** If the PSN value was missing from the product\_numbers.db file, the serial command would use the EMC\_SERIAL\_NUMBER field, which is the physical serial number of Enclosure\_0

### **NAS INSTALLATION OF NS-960:**

- During “serialinstall” or “serialkickstart” the PSN (Product Serial Number) prompt will appear, and if the RESUME prom is properly populated from the factory, the correct PSN number will appear. If so, just accept that value as the default. If the PSN number does not correspond to the PSN from the PSNT tag, then input the correct PSN number.

## **FOXGLOVE PRODUCT PART NUMBERS (PPN, see product\_numbers.db file):**

### **FOXGLOVE PSNT (Product Serial Number Tags):**

900-525-001 (NS-960 or NS-960FC)

900-525-002 (NS-G8)

900-525-003 (NS-960iS)

→Beginning with the NS-960 release, Celerra systems will ship with a “dogtag” called the PSNT

→The PSNT consists of the Product Serial Number, Product Part Number, and Revision

→The PSNT will be used to replace the traditional 100-xxx-xxx TLA number formerly used as a placeholder for the product in CSI  
→PSNT attaches to lower righthand side of DME0’s chassis [100-520-839] on the power supply

## **FOXGLOVE NAS SYSTEM MODELS:**

NS-960

NS-960FC

NS-960iS

NS-G8

**SYSTEM MODEL COMPARISONS:**

| Celerra Mgr/<br><u>CLI Model</u> | Array<br><u>Type</u> | Navisphere/<br><u>CLI Model</u> | FC Ports/<br><u>SP</u> | Max.<br><u>Blades</u> | Max. Loops/<br><u>Max. Drives</u> | Ext. Hosts | MPFS | FC Lic. | Navi Lic. |
|----------------------------------|----------------------|---------------------------------|------------------------|-----------------------|-----------------------------------|------------|------|---------|-----------|
| NS-960                           | AUX                  | CX4-960                         | 8                      | 2-4                   | 4/480                             | no         | no   | no      | no        |
|                                  | AUXF                 | CX4-960                         | 16                     | 2-8                   | 8/960                             | no         | no   | no      | no        |
| NS-960FC                         | AUXF                 | CX4-960                         | 16                     | 2-6                   | 8/960                             | yes        | yes  | yes     | yes       |
| NS-960iS                         | AUXI                 | CX4-960                         | 8                      | 2-4                   | 4/480                             | no         | yes  | no      | no        |

**Note:** Number of Blades & Backend Loops supported are dictated by the AUX array model, which determines the number of FC BE & FE ports available per SP. Each Blade consumes one FE FC port/SP. Each BE loop consumes one BE FC port/SP. Only the NS-960FC comes with an FC and Navisphere license. The NS-960 ships with an AUXF array if configured with >than 5 Blades, and/or if configured with >than 480 Drives. Additionally, all models can be optionally configured with a 2<sup>nd</sup> Control Station. Blades can be configured with a choice of one, of the eight different IO module configurations available, but when ordered, all blades within the Celerra system must use the same IO Module configuration. Each Blade has a two-slot IO Annex for additional IO modules, though the second slot in the Annex is not currently used for any IO configurations.

**I. NS-960 INTEGRATED (AUX or AUXF array)**

- Supports 8 or 16 FC Ports/SP, depending on backend configuration
- Supports AUX or AUXF backend configuration
- Supports basic NAS IP host connectivity via NFS, CIFS, or iSCSI protocols
- Supports 1 or 2 Control Stations
- Supports 2, 3, or 4 Blades, and up to 480 drives with AUX backend
- Supports 5-8 Blades, or 481+ drives with AUXF backend
- AUX array is configured with (2) Tomahawk IO cards/SP, for a total of 8 FC ports/SP, slots 0 & 1
- AUXF array is configured with (4) Tomahawk IO cards/SP, for a total of 16 FC ports/SP, slots 0, 1, 2, 3
- Does not ship with FC License
- Does not ship with Navisphere license
- Not upgradeable to FC Enabled system at GA

**II. NS-960FC INTEGRATED (AUXF array):**

- Supports use of all 960 drives for FC Hosts
  - Supports 1 or 2 Control Stations
  - Supports 2-6 Blades, and up to 960 drives
- Note:** NAS 5.6.45 and the procedure to add a 5<sup>th</sup> FC IO Module lifts this restriction. NS-960FC can support 8 Blades and external FC Hosts if the SPs contain the 5<sup>th</sup> FC IO Module.
- Supports AUXF backend configuration only
  - Ships with four Tomahawk SLICs per SP, slots 0, 1, 2, 3
  - A 6-Blade NS-960FC would have 2 FC ports/SP available for FC Hosts
  - A 4-Blade NS-960FC would have 4 FC ports/SP available for FC Hosts
  - A 2-Blade NS-960FC would have 6 FC ports/SP available for FC Hosts
  - Ships only with the AUXF array, which means the SPs have (4) Tomahawk IO cards, or 16 total ports/SP
  - Ships only when ordered with FC License
  - Navisphere licensing levels split into 480 drives or less and 481 drives or more

**SPECIAL RULES FOR FC:**

- FC license required & is bundled with Navisphere license for 480 drives or less
- FC license required & is bundled with Navisphere license for 481-960 drives

**Note:** FLARE 28.7 NDU automatically adds the Drive Expansion enabler on the backend (though this may have started with Mira FLARE 28.5)

# /nas/sbin/navicli -h 10.241.168.151 ndu -list

Name of the software package: -960driveExpansion

Revision of the software package: -

Commit Required: NO

Revert Possible: YES

Active State: YES

Is installation completed: YES

Is this System Software: NO

**III. NS-960iS iSCSI (AUXI array):**

- Supports 1 or 2 Control Stations
- Supports 2, 3, or 4 Blades, and max 480 drives
- Supports AUXI backend configuration only
- System can be used for NAS IP Hosts and MPFS over iSCSI Hosts

→SPs ship with (2) Tomahawk FC SLICs for a total of 8 ports/SP, slots 0 & 1

→SPs ship with (2) Harpoon 2-port iSCSI SLICs, and another (2) Harpoons in the IO Annex for each SP, slots 2, 3, 4, 5

**Note:** Each IO Annex supports two SLIC modules (leftside Annex for SPB; rightside Annex for SPA)

## **IV. NS-G8 GATEWAY**

→Supports 2-8 Blades, 1-2 Control Stations

→Supports CX, CX3, CX4 arrays; Symm5, DMX1, DMX2, DMX3, & Tigon Symmetrix

→Maximum of (4) attached arrays per Celerra

## **STORAGE MANAGEMENT OPTIONS FOR FOXGLOVE:**

→NS-960 & NS-960iS systems will use CLI setup\_clariion script or Celerra Manager

→NS-960FC will use Navisphere tools

→NS-G8 uses Clariion-based tools, like Navisphere

## **FRUs/CRUs/UPGRADES:**

### **Upgrades at GA:**

Add Blades

Add Control Station

### **POST-GA Upgrades:**

Add SLIC to Blades, Dec 2009

Change FC port usages on Integrateds

Upgrading DIMMs (DIMMs or motherboard changeout)

Replace XB60 or XB65 blades with XB90 (includes swapping out Enclosure)

NS80/80G upgrade to XB90 requires all blades & enclosures be updated to XB90

NS80, NS-480 conversion to NS-960

NSX, NS80G conversion to NS-G8

NS-960 to NS-960FC Upgrade, adds (2) FC Modules to each SP, adds FC Enable license (May 2009)

NS-960 to NS-960iS Upgrade, adds (4) iSCSI Modules to each SP (May 2009)

NS-960FC AUXF Upgrade, adds a 5<sup>th</sup> FC IO module to each SP (May 2009), allowing expansion to 8 Blades/system

NS-960FC AUXF Upgrade, adds single iSCSI IO Module to each SP for MirrorView or SANCopy (May 2009)

### **CRU LIST:**

Replace Blade Enclosure Power Supply

Replace Blade SFPs

Replace Disks

Replace SPs

**Note:** Actually, you replace the SP assembly and reuse CPU, SLICs, DIMMs as needed, or replace faulty CPU, SLICs, DIMMs)

Replace SP Enclosure

Replace SP Enclosure Power Supply

Replace SLIC

Replace SPS

Replace SFP in SPE

Replace LCC in DAE

### **FRU LIST, same as above, plus:**

Replace Blades

Replace Blade enclosure

Replace Control Stations

Replace Management modules

Power Supplies DAE

Blade DIMMs

### **PUHC/PAHC CHECKS:**

→Master dog tag cache file missing (PSNT)

→Missing SLIC signature in device file

→SPA-SPB communicate check failure

→Proxy-ARP setup check on NS40, NS80, NS-960

## **FOXGLOVE ISSUES:**

### **MANAGEMENT SWITCH FIRMWARE:**

→The Coldfire firmware on the NAS “Earthquake” Management Switches in Enclosure 0 is susceptible to I2C Bus Errors

→emc215163 was created to originally track and handle the issue

→emc219555 was created to document the process for upgrading firmware to 2.98 and to resolve any current Bus Errors

### **sys log:**

CS\_PLATFORM:BoxMonitor:CRITICAL:536::::1243324284:Enclosure 0 fault occurred.

CS\_PLATFORM:BoxMonitor:CRITICAL:515::::1243324320:Enclosure 0 blade B I2C PSA bus error.

**DIMMs CHANGES (Dual In-line Memory Module):**

→Foxglove systems will be converting from an older style 1GB DIMM (8 total slots, part #100-562-959), to a newer style 1GB DIMM that will be packaged in pairs under a new part number. Until NAS firmware can support, all (8) DIMMs will need to be replaced for any DIMMs issue. With a future NAS version, only a single DIMMs pair will need to be replaced.

**DAE CABLING KITS:**

→Direct Express order entry system does not properly configure enough DAE backend cables for systems with (6) DAEs or more. To be corrected by Sept 2009

**CELERRA PLATFORM BASED ON MAMBA CLARIION BACKENDS**

**CELERRA NX4 (Wormwood):**

**FIRMWARE & SOFTWARE VERSIONS:**

NAS 5.6.39.5

FLARE 02.23.050.5.703 [ECO 67366 Flare 02.23.052.5.705; ECO 70011 Nov 2009 for patch 707]

NST 6.28.50.2.29 on Apps & Tools CD [Clariion auto-update provides 6.28.50.2.35]

FLARE 28 Navisphere and NaviExpress versions

**NX4 FACTS:**

→GA Aug 2008 NAS 5.6.39.5, lower cost model than NS20, using the Mamba AX4-5f8 array (Flare 23 Vega 02.23.050.5.703), NS40 Blades, & Maynard Control Station (no floppy) #100-520-665

→Mamba SPs can have 2-port (Clariion) and 4-port FC I/O Modules, the latter 4-port solution is the BOA card for NAS only

**Note:** NaviExpress has a feature that propagates Host data to all available FC ports—this functionality is disabled on a physical port basis. Ports 0 & 1 Fibre are for Hosts, Ports 2 & 3 AUX are for NAS Hosts only.

→A Customer installable & implementable release

→Will not support SRDF, TimeFinder, or MPFSi

→All-in-One Model only, with FC license optional--(4) FC SP ports (BOA) available for sharing array with other Hosts [two per SP]

→Optional FCOEL (FC Option Enablement License) + NaviExpress or Navisphere license

→Tiered Services--Enhanced, & Premium Support

→Support for NaviExpress (Vega Flare), RAID 1/0, SAS with new AVM profiles for R5 & R1/0, new templates for 12 drive enclosure

→New Email User notification feature configured during CSA configuration and/or Celerra Manager>Support>Email User tab

→To be shipped as EMC 40U racked or Customer-racked

→Mamba is a 2U enclosure containing (2) SPs with 4-port FC Boa (NAS-only) IO module (2 ports for Celerra; 2 ports for Windows or Unix Hosts), and (12) SAS or SATA drives

→2U Mamba DPE array with up to 12 SAS/SATA drives; (2) SPs; 4-60 disks totaling (4) expansion slots of 12 drives each

**NX4 RESTRICTIONS/LIMITATIONS:**

→MPFS/MPFSi; Timefinder; SRDF are not supported on NX4

→Celerra Manager storage related statistics will not be available since Mamba AX4-5 array does not support Java

**Note:** If customers had full Navisphere, though, and NaviAnalyzer, they could produce array statistics. Current Celerra Manager Storage System Properties displays the following message for “Front-End I/O Requests”: Data not available on this storage system model.

**FACTORY LUN CONFIGURATION ON RG0:**

→A change from the current implementation, RG0 will have all remaining space allocated to the creation of two Data LUNs

→This change will apply to NX4 and all NS20, NS40 models as of August GA

**CELERRA NX4 FOOTPRINT:**

(4) expansion DAEs of 12 SATA and/or SAS drives each [4@2U]

(1) DPE with SPs & 12 SATA drives [2U]

(1) SPS [1U] with 2<sup>nd</sup> SPS optional

(1) Control Station [1U]

(1) Blade enclosure with 2 Data Movers [1U]

**Max. footprint:** 13U

**NX4 SYSTEM HARDWARE:**

**Sledgehammer Data Movers (NS20/NS40 Blades)**

→(4) 10/100/1000 Front-End Ethernet copper ports for Blades (Broadcom cge0-cge3) as one offering, and (4) Ethernet ports with (2) Copper & (2) Optical Ports (fge0-fge1) as a different offering, with (4) FC ports (Agilent QX4) to BE array & tape drives

→Single or Dual-Blade systems, either FC-Enabled model NX4FC or Non FC-Enabled model NX4

→2.8GHz Intel P4 processors, 4GB memory, 800MHz FSB, CMB-Sledgehammer motherboard (NS40), (2) Pwr Supplies per blade

→Total storage capacity 16TB/blade for IP configurations

### **Maynard Control Station (100-520-665)**

#### **CLARiiON AX4-5F8 “Mamba” array with 4-port BOA FC IO card**

### **NX4/AX4-5F8 SYSTEM INFORMATION:**

→NX4 Platform information located on Powerlink Celerra Tools section

→Celerra does not offer Gateway support for the AX4-5 backend

→EMC Celerra NX4 System Single & Dual Blade Installation Guides (Replaces ReadMeFirst & Placemat sheet)

**Note:** Contains Wagonwheel chart, Overview of Installation steps, detailed installation section, Cabling diagrams, CSA Worksheet & information

→All NX4 systems will be factory installed and require the use of the CSA to configure during initial startup

→Minimum NAS version is 5.6.39.5, Flare version 02.23.050.5.703

→NX4 Integrated will have same physical backend (AX4-5F8), but will not be licensed as an FC model unless ordered that way

→NX4 Integrated will manage & configure additional Backend storage from Celerra Manager GUI or via Celerra CLI using setup\_backend script

→NX4FC model will use NaviExpress or Navisphere to configure and manage additional Backend storage

→NaviExpress is the default CLARiiON management tool. System can be upgraded to Navisphere, but this is a one-way upgrade.

→Flare 23 is the base flare for the AX4-5 system. Flare upgrades are not currently revertable—they are permanent—emc184109

→Default is to ship only a single Standby Power Supply (SPS A), with a 2<sup>nd</sup> SPS B optional

→All NAS systems will ship with Dual SPs to support Write Cache

→SPs and O/S first shelf drives ship within an SPE enclosure. Additional drive shelves are shipped as DAEs, up to four total

→The SP has only a single LAN Management Port, which is used to connect the Internal Management networks between SPA & DM2, and SPB & DM3 for dual blade systems. There is no 2<sup>nd</sup> LAN Service port.

→The SP has two mini-db Serial ports, but only the top port is used as the Service Port (and is marked with Wrench symbol)

→Each SP will have (4) FC Front End ports [Port 2 AUX, Port 3 AUX, Port 0 Fibre, Port 1 Fibre], and Celerra is dedicated to using the AUX ports, while additional Hosts will use only the 0 & 1 Fibre ports.

### **CONNECTING TO SP SERIAL SERVICE PORT AX4-5 (Wrench symbol):**

1. Connect using null modem serial cable between Windows workstation (db9 female) and SP top service port (mini-db9 male)  
Part #038-003-084

2. Configure HyperTerminal session [9600 baud, 8 N 1 None] using ANSIW emulation

**Note:** This connection would allow troubleshooting SP bootups, access to BIOS, or access to POST menu when using “ctrl + c” (or Esc key if that does not work) combination to let POST complete, followed by password SHIP\_it. POST model shows up as “BOOMSLANG: SAN”

### **CONNECTING DIRECTLY TO AX4-5F SP LAN PORT:**

For a factory installed AX4-5F, configure the laptop with an address of 1.1.1.2 at 255.255.255.0 mask. Connect the LAN cable from laptop to SP LAN port, open a web browser, and access the setup program using 1.1.1.1/setup. Once the setup program is available, you can then set the IP address of the SP(s) to the Celerra internal addresses of 128.221.252.200/128.221.253.201 and gateways 128.221.252.104/128.221.253.104. Repeat procedure for alternate SP, which will have the same ‘floating’ system IP of 1.1.1.1.

**Note:** Actually, for a factory fresh AX4-5 system, you may need to run a Storage System Initialization Wizard before you can assign the IP addresses for the SPs. After initialization, connect to SPA using 1.1.1.1 (SPB shows up as 0.0.0.0), assign the IP address, then connect to SPB using 1.1.1.1, and assign its IP address. With a Hub or Switch connecting to both SP LAN ports at once, you can assign both IP addresses for the SPs during the initialization process. You may also want to NDU the array to the proper flare version before doing a NAS installation.

### **NX4 CRU LIST:**

#### **NAS HEAD:**

Blades

Blade Power/Cooling Modules

SFPs on Blades

#### **ARRAY:**

SP/DAE Power/Cooling Modules

SPs

SP I/O modules

SP Memory modules

SPS

SFPs in SP

Power/Cooling Modules DAE

Drives

LCCs

### **NX4 FRU LIST:**

#### **NAS HEAD:**

Power/Cooling Module Blade

Blade

SFP module Blade

Control Station [Special FRU—fault isolation checklist]

Blade Enclosure

**ARRAY:**

Power/Cooling Module SP

SP

SP I/O Module

SP Memory Module

SPS

SFP module on SPs

DPE SP/Midplane enclosure

Power/Cooling Module DPE/DAE

LCC Module

Disk

DAE Enclosure

**NX4 HW UPGRADES:**

Add Disk

Add Blade

Add SPS [Missing from documentation list]

Add DAEs

Enable FC Option

Upgrade Flare

Upgrade from NaviExpress to Navisphere via license and procedure to add the Navisphere Manager enabler [Procedure available from: <http://www.emc.com/microsites/clarion-support/ax45-install.htm>]>Upgrade to Navisphere Manager: select ‘Upgrade to Manager guide’ and download]

**NST SUPPORT NX4 or NX4FC:**

6.28.50.2.29 is the current NST version on Apps & Tools CD, Auto-update version from CLARiiON website is 6.28.50.2.35

**Hardware Registration:**

Register Storage System → Popup reports that “This feature is not supported on your Celerra integrated system.”

**Note:** Clarion perspective, this is supported on AX4 array

**Hardware Installation:**

Install Disk Array Enclosure

Install Disk

Install I/O modules and/or SFPs

→ all of these wizards report that the Storage System is not supported with these tools

**Software Assistant:**

Download and Install Hot Fix

Prepare for Installation

Install Software

Help

→ All of these wizards are supported

**Hardware Maintenance:**

Verify Storage System—supported and works

Replace Disk—not supported on AX4-5F8 system

Upgrade Disk Firmware—appears to be supported, but would use Eng. mode and Engineering>Upgrade Disk Firmware wizard

Capture SPCollect—supported and works

**NAVISPHERE EXPRESS MANAGEMENT TOOL (NaviExpress):**

→ New NaviExpress Clariion array management GUI

→ NaviExpress is an HTML web-based user interface, with HTML generated by Backend via NaviJrProvider. Client Browser sends requests via HTTP/URL to Cimom agent, which forwards to NaviJrProvider, which processes requests for HTML content, or uses get/set storage object wrappers to communicate or take action on the array

→ NaviExpress GUI consists of Main Menu frame, Header frame, and Content frame

→ All AX4-5 arrays will ship with NaviExpress installed on the array by default, though NX4 is not licensed to use it

→ A NaviExpress license comes as part of the FC Enabled license upgrade [aka NX4FC]

→ NaviExpress can be upgraded to a full Navisphere license, but no returning to NaviExpress

→ NaviExpress will be restricted to not pass Host Blob information on the NAS FC Ports [Celerra uses its own method to propagate this information]

→ NaviJr.Provider provides User interface via HTML Web Pages, Web Page classes that plug into the NaviJr Provider for NaviExpress GUI & monitors array health

**New Terminology or behavior seen in NaviExpress GUI:**

**Disk Pools** equate to Raid Groups

**Virtual Disks** equate to LUNs

Storage Groups are not visible with NaviExpress

Supports only Naviseccli

**Determining Management Tool for NX4:**

# /nas/sbin/navicli -h 192.1.4.220 managedby

Managed By: NaviExpress

**UPGRADING FROM NAVIEXPRESS TO NAVISPHERE:**

→Would need to purchase a Navisphere license

→Then download the CLARiiON procedure for upgrading

www.emc.com/ax45-support>Install>Upgrade to Navisphere: select ‘Upgrade to Manager guide’ to download

**Note:** This is a one-way upgrade only

→Download the ‘NavisphereManagerAXEnabler.ena’ file from Powerlink:

Navigator Dropdown > EMC Services Partner Web > select CLARiiON > AX4-5 Series > Softward Downloads > find the enabler

→Log into the array via NaviExpress and run the software upgrade wizard to install the enabler

**CREATING CELERRA LUNS USING NAVIEXPRESS:**

Navisphere Express>Manage>Virtual Disks (i.e., LUNs)

1. Select an available disk pool (i.e., Raid Group)
2. Enter a custom name (or keep the default name: “Virtual Disk 201”) and capacity for the new LUN, as well as how many ‘Virtual Disks’ (luns) to create
3. Assign the Virtual Disk to a Server [locate the default name used for Celerra: “Celerra\_<CS\_hostname>” in the dropdown menu]
4. Click apply to create [Flare 23 does not have Fast Bind capability]

**Comment:** LUNs created via NaviExpress automatically select the next lowest available HLU number, which in the case of Celerra, would most likely be in the Reserved HLU LUN range of 6-15 decimal. Please see emc191681 for more details, but essentially beginning with NAS 5.6.28 and above, the NAS diskmark code will automatically reassign any Reserved HLU LUN number to a valid HLU LUN number, thereby preventing assignment of Reserved LUN numbers. For Navisphere Express managed systems, the diskmark protection will be silent and will not inform the User of the change. For Reserved HLU assignments created and added to StorageGroups using Navisphere Manager, an informational message will be displayed during diskmarking to show what is being done, and the diskmark protection code will prevent an HLU from being assigned to a Celerra volume:

Info 26306752254: APM00072303347 reassigned LUN 0009 in storage group 'Celerra\_aviator' from host id 0007 to 0017

**BACKEND EVENT MONITORING:**

→Backend Event monitoring based on Storage API (Solution Enabler API) rather than Navi Agent EVMonitor [NaviExpress does not have support for Naviagent event reporting]

→To be implemented on all NAS integrated systems beginning with NX4

→Backend is polled every 5 minutes, meaning that event reporting could be delayed with NX4

→Implemented within Master Control Daemon (nas\_mcd.cfg) using a Management Daemon (mgmtd)

# ps -ef |grep mgmtd

```
root  32427 3466 0 Jul17 ?    00:00:00 /bin/sh /nas/sbin/start_sys_mgmtd
root  14956 14644 0 Jan13 ?    00:00:50 /nas/sbin/mgmtd
```

# view /nas/sys/nas\_mcd.cfg

```
daemon "System Management Daemon"
  executable  "/nas/sbin/start_sys_mgmtd"
  optional    yes
  autorestart yes
  ioaccess   no
```

**MANAGEMENT DAEMON LOGS:**

/nas/log

-rwxrwxr-x 1 nasadmin nasadmin 211083 Aug 29 10:04 nas\_log.al.mgmtd

-rw-rw-r-- 1 nasadmin nasadmin 6921 Aug 29 10:04 nas\_log.al.mgmtd.crash (Records stops and starts)

**Management Daemon Core Dumps:**

/nas/var/dump

**MINIMUM AND MAXIMUM CONFIGURATIONS:**

**Entry Configuration 5U .4TB:**

→Single DPE with dual SPs & 5@146GB SAS drives

→Single SPS

→Single Maynard Control Station

→Single Blade

**Dual Blade Configuration 13U 45TB:**

- Single DPE with dual SPs and 12@1TB SATA drives
- Single SPS
- Single Maynard Control Station
- Dual Blades
- Up to four expansion DAEs with 12@1TB SATA drives each

### **TIERED SERVICES MODEL:**

- Enhanced and Premium services (no basic)

#### **Enhanced Support:**

- Call Center 7x24, 4-hour support
- Onsite 5x9 next business day support
- CRUs to be replaced by Customer using Email Notification CCMD messaging. CRU parts shipped for next day delivery.
- FRUs will be next day EMC/Partner support
- Flare NDU can be done by customers using NST
- NAS Upgrades must be done by EMC/Partner and require software maintenance contract
- Customer installable with optional installation service

#### **Premium Support:**

- Call Center 7x24 2-hour response time support
- Onsite 7x24 4-hour response time support
- All CRU/FRUs serviced by EMC/Partner
- Flare & NAS Upgrades performed by EMC/Partner and requires software maintenance contract
- Customer installable with optional installation service

### **SELECTED NX4 HW/SW MODEL NUMBERS:**

|             |                                                                    |
|-------------|--------------------------------------------------------------------|
| NX4-1C-A    | NX4 NAS ENCLOSURE W/ ONE BLADE - FACTORY INSTALL TLA #100-520-800  |
| NX4-1C-A-FD | NX4 NAS ENCLOSURE W/ ONE BLADE - FIELD INSTALL TLA #100-520-800    |
| NX4-2C-A    | NX4 NAS ENCLOSURE W/ TWO BLADES - FACTORY INSTALL TLA #100-520-801 |
| NX4-2C-A-FD | NX4 NAS ENCLOSURE W/ TWO BLADES - FIELD INSTALL TLA #100-520-801   |
| NX4-AUXF    | NX4 CAPTIVE ARRAY - FACTORY INSTALLED TLA #100-520-802             |
| NX4-AUXFCR  | NX4 CAPTIVE ARRAY - FIELD INSTALLED TLA #100-520-802               |
| NX4-CS      | NX4 CONTROL STATION - FACTORY INSTALLED TLA #100-520-665           |
| NX4-CS-FD   | NX4 CONTROL STATION - FIELD INSTALLED TLA #100-520-665             |
| NX4-EXPAN   | NX4 EXPANSION ENABLER SAS/SATA DRIVES                              |
| NX4-FCOPT-L | NX4 FC PORT OPTION ENABLEMENT LICENSE (includes NaviExpress)       |

### **CLARiiON ARRAY SUPPORT:**

#### **AX4-5F8 Array for Wormwood NX4:**

- To use only the 4-port Fibre Channel FE BOA I/O card on array model sold for NAS NX4 platform, with Ports 0 & 1 Fibre dedicated for shared Hosts, and Ports 2 & 3 AUX for NAS Hosts
- The “8” in AX4-5F8 represents the total of (8) FC FE ports on the array
- SP’s will not support iSCSI
- Although AX4-5 arrays allow single SP configurations, only dual SP configurations will be sold with NX4 integrateds
- Single or Dual SPS configurations provided
- Array will use only SAS or SATA drives, between 4-60 drives, with 12-drive DAE’s offered as expansion units after DPE
- Minimum configuration 4 drives [Disks 0-3 contain Vault Drives Raid5 3+1, HS optional]
- Only 4 drives used for Vault
- Array offers only SecureCLI

**Note:** Control Station will use the Flare 28 version of Naviseccli

- Array supports RAID 1/0, Raid 3, Raid 5, and Raid 6, though NAS will not support RAID 3 on this platform
- SAS & SATA drives supported within same Enclosure, but not same Raid Group
- Array will not use Navigent, so array events will be delivered via a Storage API to the Control Station Management Daemon

### **AX4-5F FLARE SOFTWARE SUPPORT:**

- Celerra NX4 GA’ed with Flare 02.23.050.5.703 [Dec 2009 introduces patch .707 for additional support]
- Flare 23 (Zeta) to support BOA FC I/O cards, full Navisphere license, upgrade to dual SPs, layered applications, some NST tool support, and Windows Server 2008 support
- Vega Flare 23 supports NaviExpress, 4-port FC I/O BOA module on SPs for NAS-only, and RAID 6 support with NaviExpress, support for AX350 array
- Flare 23 does not support Fast Bind, so LUN bind operations will take a long time
- Uses Flare 28 version of NST, NaviExpress, Navisphere & NaviCLI

### **Typical FC Port Cabling between Blade & SPs:**

- Blade 2 port BE0 → port 2 AUX on SPA
- Blade 2 port BE1 → port 2 AUX on SPB
- Blade 3 port BE0 → port 3 AUX on SPA

Blade 3 port BE1 → port 3 AUX on SPB

**DISK DRIVE SUPPORT:**

146GB 15k SAS

300GB 15k SAS

400GB 10k SAS

1TB SATA

**DAE Modules:**

(4) expansion DAEs of 12 SATA and/or SAS drives each [4@2U]

**EXPANSION TIER ENABLER RULES:**

- a) First added DAE will require an Expansion Tier enabler, which would be installed using NST
- b) Expansion DAE enabler is also required on any system with more than 12 drives
- c) Required for any system with 10 or more Hosts
- d) Required for any system using Layered Applications

**Drive Types:**

--total of 5 shelves 60 drives, with an all-SATA configuration that includes SATA vault drives; a mixed SAS + SATA drives, just not in same Raid Group; and an all-SAS configuration

--400GB 10K SAS; 146GB 15K SAS; 750GB or 1TB 7200 RPM SATA drives [mixed SAS/SATA but not in same RG]

--Supports SAS and/or SATA drives within the same Enclosure, just not within the same RAID Groups

--Supports SATA vault drives

--For Hot Sparing, could use either Drive type as a Hot Spare for both types, provided the capacity was large enough, though Eng. recommendation is to use the largest ATA drive as the Hot Spare, since the ATA drives will generally be equal to or larger than SAS

--(1) Hot Spare for every 24 drives is the rule, as minimum

**RAID Types:**

--Celerra will support RAIDs 1/0, 5, &amp; 6 on the AX4-5 array, not Raid 3

--SAS/SATA drives can be in RAID 1/0, R5, or R6 configurations

--RG0 Control LUNs can be 3+1 R5 or 4+1 R5. We will allow the 3+1 R5 only for systems that are ordered with the minimum 4-disk (no HS) or 5-disk (has HS) configurations

**SUPPORTED RAID CONFIGURATIONS:**

| RAID Type | Drive | Default LUNs | AVM Pool        |
|-----------|-------|--------------|-----------------|
| 3+1 R5    | SATA  | 2 luns       | clarata_archive |
| 4+1 R5    | SATA  | " "          | " "             |
| 5+1 R5    | " "   | " "          | " "             |
| R1/0      | " "   | " "          | clarata_r10     |
| 4+2 R6    | " "   | " "          | clarata_r6      |
| 3+1 R5    | SAS   | " "          | clarsas_archive |
| 4+1 R5    | " "   | " "          | " "             |
| 5+1 R5    | " "   | " "          | " "             |
| R1/0      | " "   | " "          | clarsas_r10     |
| 4+2 R6    | " "   | " "          | clarsas_r6      |

**Templates:**

First tray OS R5: 3+1 R5; 6+1R5; Hot Spare

Additional Trays: 4+1R5 + 6+1R5

**Note:** Many other templates being finalized**BEST FIT LUN SIZE ALGORITHM:**

NX4 introduces a new algorithm which will attempt to generate LUNs that are the same size as already existing LUNs, and also of the same Raid Type and Disk technology, regardless of the number of disks in the RG or size of the spindles. After creation of all matching sized luns, the last LUN bound would consume the leftover space in the LUN to create a single odd size LUN. One of the reasons for the “Best Fit” algorithm is to spread same size LUNs over different RAID groups in order to maximize the underlying spindle count, hence increase performance.

**Example:**

If you had 4+1R5 raid group and 5+1R5 raid group, normal AVM process would be to create two luns of the same size within each of the RGs. Best Fit would instead create two luns of the same size as existing luns across both Raid Groups. Odd remaining space is allocated to a miscellaneous sized lun. If luns of the same size cannot be created, will revert to the traditional way and create two luns of same size in each RG.

**NEW DISK TYPES:**

CLSAS

CMSAS

**New AVM Storage Pools for NX4:**

clarata\_r10 → Clariion raid 1/0 on SATA

cmata\_r10 → Mirrored Clariion Raid 1/0 on SATA

clarsas\_r10 →Clariion raid 1/0 on SAS  
cmsas\_r10 →Mirrored Clariion Raid 1/0 on SAS  
clarsas\_archive →Clariion Raid 5 on SAS  
cmsas\_archive →Mirrored Clariion raid 5 on SAS  
clarsas\_r6 →Clariion raid 6 on SAS  
cmsas\_r6 →Mirrored Clariion raid 6 on SAS

#### **NX4 BACKEND CLEANUP PROCEDURE:** emc191982

1. Log into the Control Station as nasadmin, then su to root
2. Breakdown the Proxy ARP configuration to an "unconfigured" state (if applicable):

**# /nasmcd/sbin/clariion\_mgmt -stop**

Checking if running as root...yes

Checking if model is supported...yes

Checking for integrated system...yes

Checking if interface eth3 is configured...yes

Checking if SP (192.1.4.231) is up...yes

Checking if SP (192.1.4.232) is up...yes

Step 1/12: Changing SPA IP address.

Changing SPA IP from 192.1.4.231 to 128.221.252.200 (subnetmask 255.255.255.0, gateway 128.221.252.104)-----output abridged-----

**Note:** If the command fails, run again with the following additional syntax

**# /nasmcd/sbin/clariion\_mgmt -stop -skip\_rules**

**# /nasmcd/sbin/clariion\_mgmt -info**

Error 12: Not configured

3. Unset the NAS\_DB environmental variable, and stop NAS Services:

**# unset NAS\_DB**

**# /sbin/service nas stop**

4. Destroy the Celerra array configuration [deletes RaidGroups, LUNs, StorageGroup (Removes HBA Initiator records), array security, disables AccessLogix]

**Note:** For FC Enabled systems where additional storage has been configured for Other Hosts, the nas\_raid cleanup script will not touch other Storage Groups, will leave Array Security & AccessLogix intact and enabled

**# cd /**

**# /tftpboot/setup\_backend/nas\_raid -s cleanup**

Log will be created in the current directory

System 128.221.252.200 is up

System 128.221.253.201 is up

Detecting Data Movers...2 3 Done.

Clariion Array: SL7E1080700121 Model: AX4-5F8 Memory: 1023

!!! WARNING !!!

The CLARiiON array connected to this Celerra control station

contains an existing configuration. "Celerra\_nx4"

This could include user data.

If you continue by selecting 'y' to cleanup the system at the

next prompt all existing CLARiiON configuration for Celerra\_nx4 will be destroyed and the system will have to be reinstalled.

Proceed with caution!

!!! WARNING !!!

Do you want to clean up the system [yes or no]?: yes

Cleaning Storage Group "Celerra\_nx4"

Removing LUN .....

Removing diskgroup ...

Removing initiators ....

Removing storage group "Celerra\_nx4"

Removing spares

Security domain removed

Done

**Note:** If the tftpboot directory is not present, unzip and untar from the /nas/tools directory

**# tar xvfz /nas/tools/tftpboot.tar.gz**

5. Remove the S95cable\_check file (if applicable)

**# rm -f /etc/rc3.d/S95cable\_check**

6. Insert 5.6 CD-ROM and reboot Control Station, then begin the NAS reinstall by entering "boot:serialinstall" at the prompt. Do not configure the External LAN network for the Control Station during the installation process (this will be done using the CSA wizard after the NAS reinstall completes).

7. Use CSA wizard after the reinstall to configure the Control Station, public IP addresses for the SPs, Proxy ARP, and ConnectHome/Email User features.

**MANUALLY CLEANING UP NX4 ARRAY USING NAVIEXPRESS:**

1. Configure a laptop or workstation with a valid IP address on the primary Celerra internal network (for example, IP address of 128.221.252.190 and netmask of 255.255.255.0).

**Note:** If the SPs are on a public network, configure the laptop for the same network, then connect to NaviExpress using a web browser and the IP address of SPA

2. Connect an Ethernet cable directly between the workstation and the SPA LAN port (there is only one LAN port on an AX4-5 SP).

3. Launch a web browser and enter the internal IP address for SPA (for example, 128.221.252.200).

4. Log into Navisphere Express with the appropriate username and password.

5. In the Manage>Virtual Disks>menu, select a Celerra LUN name.

a) Click checkbox to select the host connection name from which to unassign the Virtual Disk (LUN)

Connection Name

Celerra\_emcnas\_i0

b) Click apply

c) Repeat this step for all other Celerra Virtual Disks (LUNs)

6. In the Manage>Virtual Disks menu, highlight a Celerra LUN, then select Destroy Virtual Disk. Repeat this task until all Celerra LUNs (Virtual Disks) are destroyed. You can only destroy one LUN at a time.

7. From the Manage>Disk Pools> menu, select the Celerra Disk Pool (for example, Disk Pool 1), verify that the Virtual Disks listed are for Celerra, and then click on the Destroy button. Repeat this step if there are any other Disk Pools associated with the Celerra installation.

**Note:** Disk Pools are equivalent to RAID Groups, and are named differently from RAID Groups typically seen on other NAS systems. For example, there is no RG0, only "Disk Pool 1", "Disk Pool 2", and the like.

8. Remove the Data Mover Blade Initiator records from the array. From the Manage>Connections menu, select the respective "HBA Port type: Fibre Channel" Initiator records under the Server listed for Celerra, and click on the Deregister button.

9. Reset SP IP addresses to default Internal IPs for NAS Reinstall, if required, using /setup [128.221.252.200/128.221.253.201]

10. Proceed with the Factory reinstallation using a NAS 5.6 bootable CD-ROM. Specify serialinstall at the boot: prompt, and continue with the NAS installation. At the configure external LAN prompt, do not setup the external IP address or hostname for the Control Station (this will be done after the factory installation using the CSA wizard).

11. Reboot the Control Station at the end of the factory installation, then configure and register the system using the CSA Wizard.

**1U MAYNARD CONTROL STATION (Part #100-520-665):**

**Phase I:**

→GA'ed in conjunction with NX4 release, CMR3 5.6.39.5 August 22, 2008

→2 GHz Celeron Intel CPU with 800MHz FSB, 2GB memory, (1) RS232 serial port for modem access (COM1 Serial Port A), (1) USB-to-RS232 null modem RS232 serial port for Service Laptop connection (COM2 Serial Port B), Keybd, Mouse, VGA connectors, 4 ethernet ports [1 for LAN, 2 for Private Celerra networks, & 1 for IPMI Dual CS environments], 250GB SATA drive, CD/DVD-ROM, no floppy, (3) USB ports, requires minimum 5.6 NAS

**Note:** With NAS 5.6.49.3, manufacturing may start using a new 500GB SATA drive from Seagate, ST3500514NS 500GB SATA Disk Drive as an alternate HDD. TEAC DV-28S-WZ3 is also an approved/alternate DVD Reader.

→Most ports are located on the back of the unit [1-VGA, 1-keybd, 1-mouse, 4-LAN, 1-Modem Serial, 1-Serial console port, 2-USB], front has (1) USB port, Power, LEDs, DVD Drive

**Note:** All (3) USB ports will support file transfer to/from Control Station

→No reset button. You would use the power-off button to powerdown, then press again to restart the CS.

→120/240VAC, 300Watt Power Supply, 120v power cord

→Maynard supported by NAS 5.6 and above only

**# cat /proc/cpuinfo**

**# cat /proc/meminfo**

**Ports Left-to-Right on Back of CS:**

Modem serial port, eth2 Secondary network (lower left), eth1 IPMI network (upper middle), eth0 Primary network (recessed PCI card, left LAN port), eth3 Public network (recessed PCI card, right LAN port), Serial console ttyS0 (recessed PCI card, far right port)

**Phase II:**

→Support to be added for dual IPMI support [i.e., dual Control Stations, NSX, NS80]

→Support added for NS20 & NS40 platforms with CMR4 5.6.40.x

**Known Issues:**

→eth0 and eth3 ethernet ports are so far recessed into the CS assembly that a User cannot remove the cables by hand—would need a flat screwdriver to depress the rj45 latch (AR122140)

→Early prototype models had incorrect Console Port Redirection to modem serial port (Serial Port A). Correct redirection setting is for the console serial port (Serial Port B). See emc192466 for guidance on BIOS settings for Maynard.

**Note:** The Control Station part number is not visible in the properties section of the Celerra Manager>Inventory>Control Station 0>properties. Use the following Linux utility to identify Control Station part number and other information as read from BIOS.

**# /usr/sbin/dmidecode|grep Version** (or “Product Name”)

```
Specification Version: 2.0
Version: A02
Version: FRU Ver 0.04
Version: 100-520-665
Version: S3200X38.86B.00.00.0033.112120071101
Version: Intel(R) Celeron(R) CPU        440 @ 2.00GHz
```

**DELL NX4 BRANDED PRODUCT (Blackwood Project):**

→Product launch Feb 2009, with 5.6.43 CMR6

→Phase I will add the Dell Service Tag info and Dell Vendor Name to RESUME PROM, which will become part of the Registration payload, and will also be present in ConnectHome payloads. In Phase I, EMC will forward ConnectHomes to Dell, and every 2 weeks will forward Registration payloads to Dell. Dell to provide service and support. Dell to stock FRUs.

→Phase II will send ConnectHomes directly to Dell using an OEM Email Address field in the RESUME Prom to key off of when ConnectHome events are being processed. We will send Registrations directly to Dell also, and will enforce this by the Dell OEM information in the Resume, which will restrict the registration option to FTP & Email only.

**/nas/sbin/get\_service\_info** →Wrapper for the t2vpd command

**/nas/sbin/t2vpd -oem**

```
SERVICE PROVIDER="EMC"
SERVICE TAG=" "
SERVICE EMAIL="US_CONNECTHOME@DELL.COM"
Apex: US\_ConnectHome@dell.com
Cork: EMEA\_ConnectHome@dell.com
```

**CHECKING & SETTING SERVICE PROVIDER NAME AND SERVICE TAG ID (5.6.43):**

1. Checking name of Service Provider in Enclosure 0 RESUME PROM:

**# /nas/sbin/t2net\_test -GetResumeProm 2 44 100 32**

t2net\_test: Return values:

```
0000: 0x45 0x4d 0x43 0x00 0x00 0x00 0x00 0x00 0x00 0x00  E M C . . . . .
```

or

**# /nasmcd/sbin/get\_service\_info**

```
SERVICE PROVIDER="EMC"
SERVICE TAG=" "
```

2. Updating Enclosure 0 midplane RESUME PROM with Service Provider name "DELL":

**# /nas/sbin/t2net\_test -SetResumeProm 2 44 100 5 44 45 4c 4c 00**

Resume cksum for this device has been updated. New Cksum = 0x6bdf847d

**Note:** 2 44 100 are fixed values for the Service Provider field; 5 is the length of the Dell name + 1; 44 45 4c 4c are Hex characters for Service Provider name in uppercase, “DELL”; and 00 is null value at end of string.

**# /nas/sbin/t2net\_test -GetResumeProm 2 44 100 32**

t2net\_test: Return values:

```
0000: 0x44 0x45 0x4c 0x4c 0x00 0x00 0x00 0x00 0x00 0x00  D E L L . . . . .
```

**# /nasmcd/sbin/get\_service\_info**

```
SERVICE PROVIDER="DELL"
```

3. Updating RESUME PROM with original Service Tag ID:

a. Use the following command to translate Dell's Alpha-numeric Service Tag ID into Hexadecimal

**# echo <Service\_Tag> | od -t x1**

Example:

**# echo JQDK481 | od -t x1**

**0000000 4a 51 44 4b 34 38 31 0a**

0000010

**Note:** Values between 0000000 - 0a represent the Service Tag ID in Hex notation

b. Set Service Tag ID in the replacement Enclosure 0 midplane RESUME PROM:

**# /nas/sbin/t2net\_test -SetResumeProm 2 44 0b0 8 4a 51 44 4b 34 38 31 00**

Resume cksum for this device has been updated. New Cksum = 0x27d08c7c

**Note:** 2 44 0b0 are fixed values indication slot\_2, midplane, and Resume field, respectively; 8 indicates the number of characters in the Service Tag ID + 1=8; 4a 51 44 4b 34 38 31 represent the Service Tag ID; and 00 is Null at the end of the string.

c. Verify that Service Tag ID was set correctly:

**# /nas/sbin/t2net\_test -GetResumeProm 2 44 0b0 32**

t2net\_test: Return values:

0000: 0x4a 0x51 0x44 0x4b 0x34 0x38 0x31 0x00 0x00 0x00 J Q D K 4 8 1 . . .

**# /nas/sbin/get\_service\_info**

SERVICE PROVIDER="DELL"

SERVICE TAG="JQDK481"

#### **NEW SWITCH TO ALLOW DELL TO MODIFY CONNECTHOME INFO (5.6.43):**

**# /nas/sbin/nas\_connecthome -modify -service\_info**

Service Information automatically detected

Ok

APEX SYSTEMS →email\_to US\_ConnectHome@dell.com

EMEA SYSTEMS →email\_to [EMEA\\_ConnectHome@dell.com](mailto:EMEA_ConnectHome@dell.com)

**# /nas/sbin/nas\_connecthome -modify -service\_info -dial\_in\_enabled no**

**# /nas/sbin/nas\_connecthome -modify -service\_info -encryption\_enabled no**

**# /nas/sbin/nas\_connecthome -modify -service\_info -email\_to US\_ConnectHome@dell.com -email\_from ConnectHome@dell.com**

#### **DEFAULT EMAIL SETTINGS:**

**/nas/sbin/nas\_connecthome -info**

Email :

Priority = Disabled

Sender Address = default@emc.com

Recipient Address(es) = emailalert@emc.com

Subject = CallHome Alert

#### **CHANGING TO DELL EMAIL:**

**# /nas/sbin/nas\_connecthome -modify -service\_info -email\_to US\_ConnectHome@dell.com -email\_from ConnectHome@dell.com**

**/nas/sbin/nas\_connecthome -info**

Email :

Priority = Disabled

Sender Address = ConnectHome@dell.com

Recipient Address(es) = US\_ConnectHome@dell.com

Subject = CallHome Alert

## ***CLARIION AX4-5 ARRAY (Mamba):*** GA January 2008

**Footprint:** 2U DPE; 2U DAE; 1U SPS

--Premium & Enhanced support, all AX4-5 arrays ship with Navisphere Express as the default, Flare 23 02.23.050.5.004 (Hercules)\*

\*Later Flare 23 releases are called Zeta [02.23.050.5.504], which offers full Navisphere support, and Vega .710

**Note:** Target Flare version for NAS products is Vega .703

--Low-end market single protocol [1GB iSCSI or 4GB FC models; Use CX3-10 when both protocols required]

--Support for Windows, Linux, NetWare, VMware, HP-UX, Sun, AIX hosts

--Scales from 4-60 drives, and supports 64 Hosts in an HA configuration, 1GB memory, single 1.66GHz processors on SPs [Single or Dual Controllers (SPs)] with 1GB memory, (4) Host FC ports per array, 45TB total storage, single SPS with dual optional, replacement for the AX150 & CX300 arrays

--Only SAS and/or SATA drives used for AX4-5 arrays

--AX4-5 supports RAID 3, 5, 1/0, & 6

--AX4-5 arrays only support SecureCLI [Classic Navi will not be available]

--Disk enclosures outside the DPE of 12 drives requires an Expansion Pack installation

--System disks are 0-3, and must be either all SAS or all SATA

--3Gbps 15K RPM SAS (Serial Attached SCSI) or SATA II disk drives, (12) drives per 2U DAE (SATA & SAS cannot be in same RG, but can share enclosure)

**Note:** With single SP systems, no SPS used. With Dual SP systems, an SP is hot-swappable

--2-port 4Gb FC FE module or (2) 1Gb/s iSCSI ports per SP, and 10/100/1000 Mgmt Ethernet port

**Note:** A special 4-port FC FE Boa I/O module is used on AX4-5f8 models used for NAS only

--AX4-5f8 array ships with SP's that use the 4-port Boa FC I/O Module [based on Tomahawk]

--Single Boomslang SP configuration allowed since 1GB flash memory available for writing dumps [but not for NAS Wormwood—require both SPs for NAS configurations]

--No Upgrade path

--Max LUNs 512, 30 RGs, 256 Luns per initiator, (4) arrays; max 64 hosts can connect via iSCSI or FC

### **SP BOOTUP INFORMATION:**

EMC BIOS Release 05.12

CPU = Intel(R) Celeron(R) CPU @ 1.66GHz

1023M System RAM Passed

1024 KB L2 Cache

-----abridged-----

Model: Boomslang: SAN

DiagName: Extended POST

DiagRev: Rev. 05.30

Build Date: Fri May 02 07:47:20 2008

StartTime: 06/05/2008 00:02:50

SaSerialNo: SL7B1080800740

### **DAEs:**

Automatic enclosure IDs

LED flashing means DAE offline

### **Power Supplies:**

Two per DPE and DAE, 550Watts

Single fan fault increases speed on others

Two fan faults will lead to array shutdown after two minutes

### **Shutting Down SPs:**

Quick push on power button leads to normal shutdown

4 sec. hold on button leads to forced shutdown

Navisphere is the preferred way to shutdown>System>Services>Shutdown

### **CLI Shutdown for AX4-5F8:**

# /nas/sbin/navicli -h 10.241.168.173 shutdown

Do you want to shutdown and power off the subsystem now (y/n)?

y

### **LED Indicators on Booting SPs:**

1/4 Hz blink = BIOS check

1 Hz blink = POST

4 Hz blink = FLARE

1-3-3-1 = DIMM failure

### **Fibre Channel Port Rules:**

Ports 0 & 1 Fibre are for non-Celerra hosts only

Ports 2 & 3 AUX on AX4-5f8 will be for NAS hosts only

### **Two Models:**

iSCSI Model with two copper 1Gbps iSCSI Front End ports/SP

FC Model with two optical 4Gbps FC Front End ports/SP

--Support for Layered Apps, AccessLogix, and Navisphere mgmt

### **Supported RAID Protection Types:**

RAID 3—five to nine disks using disk striping, single-disk parity, best suited for concurrent large I/O using 1-4 threads, read & write cache, and sequential I/O [Raid 3 NOT supported with Celerra]

RAID 1/0—groups of 2, 4, 6, 8, 10, 12, 14, or 16 disks of hardware mirror images, each image including 2-8 disks, using disk striping RAID 5—3-16 disks, disk striping, distributed parity across the disks, superb Read and good Write performance

RAID 6 (with Vega Flare)—4, 6, 8, 10, 12, 14, or 16 disk width pools allowed; RAID 6 uses two independent array positions for storing parity [Horizontal Redundancy & Diagonal Redundancy], meant to allow two adjacent data block, disk drive failures, or both

### **Supported Drive Types (GA+):**

750GB SATA II 7200rpm

400GB SAS 10K rpm

300GB SAS 15K rpm

146GB SAS 15K rpm

**Note:** SAS support requires a dual SP configuration. Drives can be mixed in same enclosure but not within the same RG. Vault drives can be all SAS or all SATA II. Hot Spare can be either SAS or SATA II.

### **Disk Shelf Rules:**

→Highest capacity drives to the left

- Then, highest speed drives to the left
- Don't mix drive types in same Raid Group

**Navisphere Express:**

All AX4-5 platforms initially ship with Navisphere Express, not Navisphere

NaviJr.Provider provides User interface via HTML Web Pages, Web Page classes that plug into the NaviJr Provider for NaviExpress GUI & monitors array health

NaviExpress will be restricted to not pass Host Blob information on the NAS FC Ports [Celerra uses its own method to propagate this information]

NaviExpress can be upgraded to full Navi with Flare 23 Zeta release

**New Terminology or behavior seen in NaviExpress GUI:**

Disk Pools equate to Raid Groups

Virtual Disks equate to LUNs

Storage Groups are not visible with NaviExpress

Supports only Naviseccli

**Determining Management Tool for NX4:**

# /nas/sbin/navicli -h 192.1.4.220 managedby

Managed By: NaviManager

# /nas/sbin/navicli -h 192.1.4.220 managedby

Managed By: NaviExpress

**March 2008 Zeta Flare Release:**

300GB SAS 15K rpm

1TB SATA II 7200 rpm

Full Navisphere support

**PERFORMING NDU FLARE UPGRADES ON AX4-5 SYSTEM:**

→There are several options for performing Flare upgrades, but note that presently, there is no way to revert back to a prior version of Flare on this platform, even if the NDU did not commit the new flare version

→Navisphere Express allows Flare NDU, and automatically commits the upgraded Flare version [Navisphere Express>System>Software>Upgrade Software]

→Secure CLI allows Flare NDU, and automatically commits the upgraded Flare version

→NST tool allows Flare NDU, with option to commit during the upgrade

**SECURE CLI NDU EXAMPLE:**

C:\Program Files\EMC\Navisphere CLI>NaviSECCLI.exe -h 128.221.252.200 ndu -install AX4-5-Bundle-02.23.050.3.728.pbu  
Running install rules...

=====

Version Compatibility : Rule passed.

Redundant SPs : Rule passed.

Acceptable Processor Utilization : Rule passed.

No Trespassed LUNs : Rule passed.

No Transitions : Rule passed.

No System Faults : Rule passed.

All Packages Committed : Rule passed.

Special Conditions : Rule has warning.

Statistics Logging Disabled : Rule passed.

Host Connectivity : Rule passed.

1 rule(s) have warnings.

Detailed rule results:

=====

RULE NAME: Special Conditions

RULE REVISION: 6.23.3.1.2.1

RULE STATUS: Rule has warning.

RULE DESCRIPTION: This rule is a warning to check for special conditions before installing software on the storage system.

RULE RESULT: Please check for the following conditions:

1. All attached servers are running failover software.
2. All attached VMWare ESX servers running pre v2.1.0 software have I/O stopped.
3. All attached AIX servers must meet the requirements described in EMC Knowledgebase Solution ID emc67186.

**RULE INSTRUCTION:**

Item number: 0

Name of the software package: FLARE-Operating-Environment

Revision of the software package: 02.23.050.3.728

Already Installed Revision 02.23.050.5.504

Installable YES

Disruptive upgrade: NO

Pre installation rules have been run to ensure the success of this software upgrade. One or more rules has a warning. The installation can continue with warnings.

The requested package(s) will be installed. Do you wish to proceed? : (y/n)?

y

#### **OTHER CLI OUTPUT FROM AX4-5:**

**C:\Program Files\EMC\Navisphere CLI>NaviSECCLI.exe -h 128.221.252.201 getagent**

Agent Rev: 6.23.3 (1.8)

Name: K10

Desc:

Node: B-SL7E1081700022

Physical Node: K10

Signature: 2162762

Peer Signature: 2162847

Revision: 2.23.50.5.504

SCSI Id: 0

Model: AX4-5

Model Type: Rackmount

Prom Rev: 4.06.00

SP Memory: 1023

Serial No: SL7E1081700022

SP Identifier: B

Cabinet: DPE4AX

**C:\Program Files\EMC\Navisphere CLI>NaviSECCLI.exe -h 128.221.252.201 getcache**

SP Read Cache State Enabled

SP Write Cache State Disabled

**Note:** Write Cache is disabled on Single SPS systems when rebooting SPA

**C:\Program Files\EMC\Navisphere CLI>NaviSECCLI.exe -h 128.221.252.201 getcrus**

DPE4AX Enclosure 0

SP A State: Present

SP B State: Present

Enclosure 0 Fan A State: N/A

Enclosure 0 Fan B State: N/A

Power Supply 1 State: Present

Power Supply 2 State: Present

Enclosure 0 SPS A State: Present

Enclosure 0 SPS B State: Empty

Enclosure 0 SPS A Cabling State: Valid

Enclosure 0 SPS B Cabling State: Cabling Status is unknown

**Note:** Cabling Status is unknown is reflected because this system has only a single SPS

**C:\Program Files\EMC\Navisphere CLI>NaviSECCLI.exe -h 128.221.252.200 getresume|more**

SP A

EMC Part Number: 100-562-107

EMC Artwork Revision: N/A

EMC Assembly Revision: A14

EMC Serial Number: SL7S1081600485

Vendor Part Number: N/A

Vendor Artwork Number: N/A

Vendor Assembly Number: N/A

Vendor Serial Number: N/A

Vendor Name: NEC

Location of Manufacture: JAPAN

Year of Manufacture: 08

Month of Manufacture: 4

Day of Manufacture: 24

Assembly Name: ASSY BOOMSLANG CANISTER

Programmable Name: BIOS:POST:LSIFW:EXBOOT:EXISTR

Programmable Revision: 4.06:4.25:4.31:3.0:3.05

I/O Module

EMC Part Number: 100-562-173  
EMC Artwork Revision: N/A  
EMC Assembly Revision: A07  
EMC Serial Number: SL7P1081600092  
Vendor Part Number: N/A  
Vendor Artwork Number: N/A  
Vendor Assembly Number: N/A  
Vendor Serial Number: N/A  
Vendor Name: NEC  
Location of Manufacture: JAPAN  
Year of Manufacture: 08  
Month of Manufacture: 4  
Day of Manufacture: 24  
Assembly Name: 2 PORT FC 4 GBIT DUAL PORT HBA  
Programmable Name: N/A  
Programmable Revision: N/A

**Note:** Output from Resume prior to changing identity to AX4-5F8 with BOA 4-port I/O module

**SHUTTING DOWN SPs FROM COMMAND LINE:**

**C:\Program Files\EMC\Navisphere CLI>NaviSECCLI.exe -h 128.221.252.200 shutdownpeersp**

This operation will shutdown and power off the SP.

Do you want to shutdown and power off the SP now? (y/n)?

y

**C:\Program Files\EMC\Navisphere CLI>NaviSECCLI.exe -h 128.221.252.200 shutdownsp**

This operation will shutdown and power off the SP.

**CONNECTING TO SP's VIA SERIAL MAINTENANCE PORT:**

1. Connect from Laptop to upper mini-db 9 serial port on SP (Maintenance port), using DB9 Female to mini-DB9 male cable [Part #038-003-084]
2. Configure HyperTerminal session using following settings:  
9600 bps; 8 data bits; None for parity; 1 Stop bits; None for Flow Control; Emulation: ANSIW
3. Use serial connection to observe bootup and to gain access to POST diagnostic menu:

Model: Boomslang: SAN

DiagName: Extended POST

*ABCab << Stopping after POST >> DEabcdegFGHabcdIJK*

- a) Press the “ctrl + c” keys here (or Esc key if that fails), which allows POST to complete, and then prints a Storage System Failure message:

\*\*\*\*\*  
\*\*\*\*\*

\* Extended POST Messages

\*\*\*\*\*  
\*\*\*\*\*

.... Storage System Failure - Contact your Service Representative ...

- b) Access POST by typing following password onscreen:

**SHIP\_it**

- c) POST Diagnostics:

Diagnostic Menu

POST 05.30

- 1) Reset Controller 11) System Test Sub-Menu
- 2) Enter Debugger 12) Image Sub-Menu
- 3) Display Warnings/Errors 13) Disk Sub-Menu
- 4) Boot OS 14) Resume PROM Sub-Menu
- 5) POST Sub-Menu 15) Power Sub-Menu
- 6) Display/Change Privilege 16) DMI Log Sub-Menu
- 7) Memory Sub-Menu 17) Information Display
- 8) Motherboard Sub-Menu 18) ICA Sub-Menu
- 9) I/O Module Sub-Menu 19) DDBS Service Sub-Menu
- 10) Enclosure Sub-Menu 20) Manufacturing Mode Sub-Menu

**SLEET NS-120/NS-480; NS-120FC/NS-480FC; NS-120iS/NS-480iS:**

→Mid-tier Integrated products, GA date December 10, 2008, NAS 5.6.41.2

- Dell will brand the NS-120 as a Dell OEM product with the 5.6.48 code release
  - Sleet 2 is the internal name for the NS-120 models, while Sleet 4 is the name for the NS-480 models
  - Blades are based on the CX3 NS40 hardware
  - Arrays are based on CX4 hardware using CX4-120 Nautilus (NS-120) and CX4-480 Trident (NS-480) backends
  - Increased system capacity and backend performance from CX4 arrays
  - Secure NaviCLI only, Flare 28
  - NST Version 28 with CRU support for Flare NDU and SLIC wizard CRU support
  - Use of Email User notification feature
  - Support for Customer provisioning and implementation, using built-in CPW in the CSA client (CPW has been delayed Q1 2009)
  - PowerPath license is included in NS-120FC model but not the NS-480FC
  - MirrorView must be part of the initial order and cannot be added after a system ships
  - EMC-racked or Customer-racked Factory installed configurations
  - Both Nautilus & Ironclad support FC & SATA drives in the same cabinet, but not within the same DAE
  - Both array types will support FC, iSCSI, and SATA technologies with the Akula Management Module
- Note:** No mixing of drive types within the same DAE. CX4-120 can use SATA in vault. Minimum configuration 5 drives + HS.
- NAS models will only support RAID 1/0, 5, & 6 [No RAID 1 or 3]
  - 4-Blade NS-484 option will be introduced in Q1 2009 [2 or 4 Blade NS-480's]
  - With the NS-480 more than 8 FC IO ports can be used, so the WWN port numbering scheme will be extended by two bits to support up to 32 FC ports
- FCP scsi-0: HBA 0: ALPA 0000ef **SP-a00:** 500601603b6002ab Class 3

## **MINIMUM SOFTWARE SUPPORT FOR SLEET:**

NAS 5.6.41.2  
FLARE 28, 04.28.000.5.009  
NST, Navisphere, and CLI based on Flare 28  
CSA 5.6.41.2

## **NAS SYSTEM MODELS:**

NS-120, NS-480 →plain Integrated model for basic NAS IP hosts  
NS-120FC, NS-480FC →Fibre Channel Enabled version supports NAS IP hosts or external FC Hosts, or MPFS over FC support  
NS-120iS, NS-480iS →iSCSI model for NAS IP hosts and external Hosts using MPFS over iSCSI

**Note:** Use the terminology “system model” to refer to the Hardware model name as indicated in Celerra Manager or from the CLI:

# /nas/sbin/model

NS-480iS

## **ARRAY MODELS:**

NS120-AUX →Factory racked NS-120, no FC or iSCSI ports for external Hosts, (1) 4-port FC card per SP  
NS120-AUXF\* →Factory racked NS-120FC, no iSCSI, (2) 4-port FC cards per SP  
NS120-AUXI →Factory racked NS-120iS, no FC, (1) 4-port FC card & (2) 2-port iSCSI per SP

**Note:** TLA for all three AUX variants is 100-520-673

NS-480-AUX →Factory racked NS-480, no FC or iSCSI ports for external Hosts, (2) 4-port FC cards per SP  
NS-480-AUXF\* →Factory racked NS-480FC, no iSCSI, configured with (3) 4-port FC cards per SP  
NS-480-AUXI →Factory racked NS-480iS, (2) 4-port FC cards & (3) 2-port iSCSI cards per SP

**Note:** TLA for all three AUX variants is 100-520-674

\*4 FC ports available for External Hosts

## **BLADE MODELS:**

→NS-120 ships with a 4-port copper Ethernet or 2-port 10GbE Optical + 2-port GbE copper Ethernet configuration  
NS121-A/FD 1-blade, 4-port Cu GbE, DME TLA part # 100-520-667  
NS121-C/FD 1-blade, 2-port 10GbE Optical/2-port 1GbE Cu Ethernet card (5.6.45)  
NS122-A/FD 2-blade, 4-port Cu GbE, DME TLA part # 100-520-668  
NS122-C/FD 2-blade, 2-port 10GbE Optical/2-port 1GbE Cu Ethernet card (5.6.45)  
NS120-AUX/CR, AUXF/CR, AUXI/CR, TLA part # 100-520-673 for all variants

**Note:** Blade FD or Array CR suffix indicates field racked system

→ NS-480 ships with either a 4-port copper or 2-port copper/2-port optical configuration  
NS482-A/FD 2-blade, 4-port Cu GbE, DME TLA part # 100-520-670  
NS482-B/FD 2-blade, 2-port Opt/2-port Cu GbE, DME TLA part #100-520-672  
NS482-C/FD 2-blade, 2-port 10GbE Optical/2-port 1GbE Cu Ethernet card (5.6.45)  
NS484-A/FD 4-blade, 4-port Cu GbE, DME TLA part # 100-520-688  
NS484-B/FD 4-blade, 2-port Opt/2-port Cu GbE, DME TLA part #100-520-689  
NS484-C/FD 4-blade, 2-port 10GbE Optical/2-port 1GbE Cu Ethernet card (5.6.45)  
NS480-AUX/CR, AUXF/CR, AUXI/CR, TLA part # 100-520-674 for all variants

**Note:** Blade FD or Array CR suffix indicates field racked system

**Product Configurations:**

The NS-120 & NS-480 will support three different configurations, which are really tied to what SLIC IO modules are resident on the SPs. The basic NAS model (NS-120/NS-480) is a traditional Integrated use for IP hosts. The NS-120FC/NS-140FC model supports IP hosts, and can support external Hosts with up to (4) Fibre Channel front-end ports that can be used for SAN-attached FC, direct-attached FC, and MPFS over FC. The NS-120iS/NS-480iS model supports IP hosts, and can support external Hosts for MPFS over iSCSI (4-ports for NS-120iS and 6-ports for NS-480-iS).

**SUPPORTED UPGRADES GA:**

DAE expansion

Adding 2<sup>nd</sup> SPS

Adding 2<sup>nd</sup> Data Mover

**POST-GA UPGRADES Q1 2009:**

FC Enable Upgrade for NS-120 & NS-480

Upgrading to a four blade Celerra NS-480 System

iSCSI Enable Upgrade (NS-120 to NS-120iS and NS-480 to NS-480iS)

Adding Blade Enclosure

Convert NS20, 120, or 40 to NS-480

Adding iSCSI modules to NS-120FC & NS-480FC for MirrorView/iSCSI support

**PRIVATE NETWORK CABLING:**

**Single Blade NS-120 (NS121-A):**

→eth0 on CS to Bottom internal LAN port on Blade 2 (mge0)

→Top internal LAN port on Blade 2 (mge1) to Top Management LAN port on SPA

→eth2 on CS to Top Management LAN port on SPB

**Dual Blade NS-120 (NS122-A) or NS-480 (NS482-A dual blade Cu Ethernet/NS482-B dual blade Cu/Optical Ethernet):**

→eth0 on CS to Bottom internal LAN port on Blade 2 (mge0)

→Top internal LAN port on Blade 2 (mge1) to Top Management LAN port on SPA

→eth2 on CS to Bottom internal LAN port on Blade 3

→Top internal LAN port on Blade 3 to Top Management LAN port on SPB

**4-Blade NS-484:** (Available with 5.6.43.8)

→eth0 on CS to Bottom internal LAN port on Blade 2 (mge0)

→Top internal LAN port on Blade 2 (mge1) to Bottom internal LAN port on Blade 4

→Top internal LAN port on Blade 4 to Top Management LAN port on SPA

→eth2 on CS to Bottom internal LAN port on Blade 3

→Top internal LAN port on Blade 3 to Bottom internal LAN port on Blade 5

→Top internal LAN port on Blade 5 to Top Management LAN port on SPB

→Initial shipments of NS-484 will share same TLA between Enclosure 0 & 1. The TLA's will be revised and become unique between enclosures beginning in April 2009

**Enclosure 0:**

100-520-670 4Cu GbE ports on blades

100-520-672 2Optical/2Cu GbE ports on blades

**Enclosure 1:**

100-520-688 4Cu GbE ports on blades

100-520-689 2Opt/2Cu ports on blades

**STORAGE MANAGEMENT:**

→Integrated and iSCSI systems will be managed completely by Celerra Manager

→FC systems will be managed by Navisphere

**DISK DRIVE SUPPORT ON CX4-120/480:**

FC (36, 73, 146, 300, 400 & 450GB)

SATA (250, 320, 500, 750GB & 1TB)

Low cost 500GB drives

**NST SUPPORT:**

→Both NS-120 & NS-480 will support all NST wizards except Register System and Post Activity SYR

**NS-120 “Sleet2” (CMR4+):**

→NS-120 integrated, NS-120FC, and NS-120iS MPFS iSCSI models (no gateway)

→Replacement for NS20 built on CX4-120 Fleet array

→“Unified Storage” marketing term for combined NAS, iSCSI, MPFS, FC solution

→Single CS only, 1-2 Data Mover Blades

→Maynard CS (100-520-665)

→NS40 Sledgehammer Data Movers, dual 2.8GHz processors, 4GB memory, 800MHz FSB, with (2) FE fibre channel ports, (1) FC port for tape connectivity, (1) management port, 4-port copper GbE Ethernet card, or 2-port 10GbE optical + 2-port copper GbE Ethernet card

→Field installed systems will be pre-configured with 8U ganged rails, which will be reused and mounted in the Customer rack

→Supports SATA & FC Drives, including SATA Vault drives

→2U Nautilus array with dual SPs 1.2GHz CPU & 3GB Memory, single SPS with optional 2<sup>nd</sup> SPS, (8) FC Ports per SP, 1-8 DAEs of 15 drives each, single BE loop for a total of 120 Fibre Channel or SATA drives

**Note:** With NAS 5.6.45, we now tolerate the array using 8Gbps FC IO Modules (Glacier)

→Nautilus 6U SPE contains SPs, DAE, and SPS

→For FC-Enabled model, up to 4 direct-attached Linux or Windows Hosts other than Celerra [FC or MPFS], or up to 120 SAN-attached Hosts [FC or MPFS]

→6-120 FC or SATA Drives [1-8 DAE's of 15 disks each]

→Max of 32TB IP storage/blade (64TB/system), or 96TB/blade for MPFS configurations

→Nautilus supports only a single BE Loop, and can have SATA drives for the Vault DAE

→Standard Integrated will ship with only a single Tomahawk FC IO module, while FC model ships with two Tomahawk FC IO modules

→No TimeFinder or SRDF support

#### **Program Purpose:**

→Introduce low-end Celerra product on CX4 Fleet array—"Technology Refresh"

→Increase storage capacity to 120 drives

→Increase performance

→Reduce service costs with enhanced user installation, implementation, and maintenance

→Use of CSA to install, factory pre-configured, EMC or Customer-racked configurations

#### **NS-120 HARDWARE CAPACITY/FOOTPRINT:**

→6-120 FC or SATA drives, maximum of 8 DAE's (cannot mix drives within same enclosure)

→32TB capacity per Blade FC or SATA, or 96TB/blade for MPFS configurations

→8U minimum footprint, max 30U footprint

→NS40 Blades built on CX4-120 Nautilus array

→Ships with single SPS (2<sup>nd</sup> SPS optional)

→NS-120 supports all 120 drives with MPFS configurations

### **THREE NS-120 MODEL OFFERINGS (NS-120; NS-120FC; NS-120iS):**

#### **NS-120 INTEGRATED CONFIGURATION (NS-120):**

→NAS IP Hosts-only configuration (CIFS, NFS, iSCSI)

→Single Maynard CS, 1-2 Data Movers, Fleet CX4-120 Nautilus

→SPs dual core CPU, 3GB memory, (1) Tomahawk 4-port FC IO modules for each SP

→Single or dual SPS

→1-8 DAEs, 6-120 FC or SATA drives (no mixing of drive types within a DAE)

→Navisphere not licensed for the backend, but NST can be used

→Maximum of 120 drives supported

→Ships with only a single 4-port FC Tomahawk IO card, no FC Enable model Upgrade at GA

#### **SP IO MODULE PORTS:**

##### **SLOT 0: Tomahawk 4-port FC IO module (A0=SPA; B0=SPB)**

Port 0 BE0→to DAE single loop

Port 1→not used

Port 2 FE0→to BE1 on DM2 SPB; to BE0 on DM2 SPA

Port 3 FE1→to BE1 on DM3 SPB; to BE0 on DM3 SPA

#### **NS-120FC CONFIGURATION:**

→NS120-FCOPT-L EMC CELERRA FIBRE CHANNEL PORT ENABLER CERTIFICATE

→Supports use of 120 drives for FC Hosts and Navisphere license

→NAS IP Hosts, external FC Hosts, or external hosts using MPFS over FC

→Two Tomahawk 4-port FC IO modules for each SP (all 4 ports 2<sup>nd</sup> Tomahawk available for MPFS or Other FC host)

→4 direct-attached FC MPFS Hosts or other FC Hosts

→128 SAN-attached FC MPFS Hosts or other FC Hosts

→Factory installed NS-120FC can support Mirrorview over FC, but FC Enabled Upgrades in the Field will not support the Mirrorview over FC option

→Navisphere licensed for this model, and required for configuring backend storage

#### **SP IO MODULE PORTS:**

##### **SLOT 0: 1<sup>st</sup> Tomahawk 4-port FC IO module (A0=SPA;B0=SPB)**

Port 0 BE0→to DAE single loop

Port 1→not used

Port 2 FE0→to BE1 on DM2 SPB; to BE0 on DM2 SPA

Port 3 FE1→to BE1 on DM3 SPB; to BE0 on DM3 SPA

#### **SLOT 2: 2<sup>nd</sup> Tomahawk 4-port FC IO module (A2=SPA; B2=SPB)**

All FC ports for Other Hosts on SPA/SPB (Port 3 would be used by Mirrorview)

#### **NS-120iS iSCSI MPFS CONFIGURATION:**

→NAS & MPFS over iSCSI configurations

→(1) 4-port FC Tomahawk I/O module on each SP (1-BE and 2-FE FC ports per SP)

→(2) 2-port Harpoon iSCSI I/O modules per SP for iSCSI MPFS hosts (all 4-ports for iSCSI MPFS Hosts)

**Note:** SP configuration based on the standard CLARiiON CX4-120 configuration

→Up to 4 direct-attached iSCSI MPFS Linux or Windows Hosts

→Up to 128 iSCSI-attached iSCSI MPFS Hosts

→During CSA configuration, a wizard walks through the setup of IP addresses for all iSCSI ports on the Array, offers to setup CHAP, and starts the MPFS service on Server\_2. Use the following command to view IP address assignments on the iSCSI ports:

**# /nas/sbin/naviseclli -h 10.241.168.182 connection -getport -sp a**

#### **SP IO MODULE PORTS:**

#### **SLOT 0: 4-port Tomahawk FC IO module (A0=SPA; B0=SPB)**

Port 0 BE0→to DAE single loop

Port 1 BE1→not used

Port 2 FE0→to BE1 on DM2 SPB; to BE0 on DM2 SPA

Port 3 FE1→to BE1 on DM3 SPB; to BE0 on DM3 SPA

#### **SLOT 1: 1<sup>st</sup> 2-port Harpoon iSCSI IO module (A1=SPA;B1=SPB)**

iSCSI 0 MPFS host

iSCSI 1 MPFS host

#### **SLOT 2: 2<sup>nd</sup> 2-port Harpoon iSCSI IO module (A2=SPA; B2=SPB)**

iSCSI 2 MPFS host

iSCSI 3 MPFS host

#### **NS-480 “Sleet4” (CMR4+):**

→Replacement for NS40 as a mid-range product, built on CX4-480 Trident Fleet array (2.2GHz 4GB memory)

→”Unified Storage” marketing term for combined NAS, iSCSI, MPFS, FC solution

→Maynard CS (100-520-665), NS40 Sledgehammer Data Movers

→Single CS only, 2 Blade system for December GA, and support for 2 or 4 Blade system with CMR6 Q1 2009 [Single or Dual DME]

**Note:** 2<sup>nd</sup> Data Mover Blade enclosure space is reserved by the configuration for expansion purposes, though the actual Enclosure itself may or may not be shipped

→NS40 Sledgehammer Data Movers, dual 2.8GHz processors, 4GB memory, 800MHz FSB, with (2) FE fibre channel ports, (2) FC port for tape connectivity, (1) management port, 4-port copper GbE Ethernet card, 2-port 1GbE optical + 2-port copper GbE card, or 2-port 10GbE optical + 2-port copper GbE Ethernet card

→Field installed systems will be pre-configured with 9U ganged rails, which will be reused and mounted in the Customer rack

→NS-480 integrated, NS-480FC, and NS-480iS MPFS iSCSI models (no gateway)

→NS-480 & NS-480FC will have a 4-blade configuration at a later release (CMR6 Q1 2009)

→2U Trident array with dual SPs, dual SPS, dual-core single 2.2GHz CPU & 4GB memory, (8) FC I/O ports per SP, 1-32 DAEs of 15 drives each, for total of 480 FC or SATA drives [FC & SATA must be in separate DAEs]

→For FC-Enabled model, up to 2 direct-attached Linux/Windows Hosts [FC MPFS or FC], or 60 SAN-attached MPFS or FC Hosts

→6-480 FC, SATA, & Flash Drives (no mixing of drives in same DAE; 1-32 DAE's of 15 disks each)

→Max IP storage 64TB/blade (128TB/system), or 128TB/Blade for MPFS configurations

**Note:** Increases to 192TB/system with Q1 2009 support for 4-Blade system

→Trident CX4-480 supports four BE Loops, and does not allow SATA drives for Vault DAE

→Standard Integrated will ship with both Tomahawk FC IO Modules, and will use Ports 0 & 1 to loop 0 & 1, resp., (dual loop), Ports 2 & 3 for DM2, DM3, respectively (2<sup>nd</sup> 4-port FC IO card will be used for Ports 0 & 1 on DM4, DM5, respectively, and other ports used for non-Celerra Hosts)

→No TimeFinder or SRDF support

#### **Program Purpose:**

→Introduce mid-range Celerra product on CX4 Fleet array—“Technology Refresh”

→Increase storage capacity to 480 drives

→Increase performance

→Introduce a 4-Blade configuration in Q1 2009

→Reduce service costs with enhanced user installation, implementation, and maintenance

→Use of CSA to install, factory pre-configured, EMC or Customer-racked configurations

## **NS-480 HARDWARE CAPACITY/FOOTPRINT:**

- 6-480 FC or SATA drives, maximum of 32 DAE's (cannot mix drives within same enclosure, nor use SATA for vault)
- 64TB capacity per Blade FC or SATA. 128TB/blade capacity for MPFS configurations.
- Note:** Max usable capacity for system would be 473TB (1<sup>st</sup> shelf 15@450GB FC drives; 465@1TB SATA drives)
- 9U minimum footprint, max 102U footprint
- NS40 Blades built on CX4-480 Trident array
- Ships with both SPS's
- Supports all 480 drives for MPFS configurations

## **THREE NS-480 MODEL OFFERINGS (NS-480; NS-480FC; NS-480iS):**

- Three NS-480 System Models
- (2) or (4) blades offered with CMR6 release

## **NS-480 INTEGRATED CONFIGURATION (NS-480):**

- NAS IP Hosts-only configuration (CIFS, NFS, iSCSI)
- Single Maynard Control Station, 2 or 4 Blades (latter offered in Q1 2009)
- Fleet Trident CX4-480 with dual core CPU, 4GB memory, and (2) Tomahawk 4-port FC IO modules for each SP
- Dual SPS
- 1-32 DAEs, 6-480 FC or SATA drives (no mixing of drives in same DAE)
- Each backend Loop represents up to 120 drives, so 4-Loops=480 drives, using 4-FC ports per SP
- Each blade uses 1-FC port per SP, 4-Blades use 4-FC ports per SP
- Ships with two Tomahawk SLICs per SP

## **SP IO MODULES for NS-480:**

### **SLOT 0: 1st 4-port Tomahawk FC IO module (A0=SPA; B0=SPB)**

- Port 0 BE0→to DAE loop 0
  - Port 1 BE1→to DAE loop 1
  - Port 2 FE0→to BE1 on DM2 SPB; to BE0 on DM2 SPA
  - Port 3 FE1→to BE1 on DM3 SPB; to BE0 on DM3 SPA
- ### **SLOT 1: 2<sup>nd</sup> 4-port Tomahawk FC IO module (A1=SPA; B1=SPB)**
- Port 0 BE2→to DAE loop 2 SPB & SPA
  - Port 1 BE3→to DAE loop 3 SPB & SPA
  - Port 2 FE2→to BE1 on DM4 SPB; to BE0 on DM4 SPA
  - Port 3 FE3→to BE1 on DM5 SPB; to BE0 on DM5 SPA

## **NS-480FC CONFIGURATION:**

- NS480-FCOPT-L EMC CELERRA FIBRE CHANNEL PORT ENABLER CERTIFICATE
- Supports use of 480 drives, but only 240 for external FC Hosts or MPFS-over-FC
- Comes with one Navisphere license for 240 or less drives, and another license for more than 240 drives
- NAS IP, FC Hosts, or MPFS over FC configurations
- (3) Tomahawk 4-port FC IO modules (3<sup>rd</sup> Tomahawk ports for MPFS or Other FC Hosts)
- (4) direct-attached FC MPFS Hosts or non-Celerra FC Hosts
- Supporting 256 SAN-attached FC MPFS or non-Celerra FC Hosts
- Factory installed NS-480FC supports Mirrorview over FC, but FC Enabled field upgrades will not support Mirrorview over FC

## **SP IO MODULES for NS-480FC:**

### **SLOT 0: 1st 4-port Tomahawk FC IO module (A0=SPA; B0=SPB)**

- Port 0 BE0→to DAE loop 0
- Port 1 BE1→to DAE loop 1
- Port 2 FE0→to BE1 on DM2 SPB; to BE0 on DM2 SPA
- Port 3 FE1→to BE1 on DM3 SPB; to BE0 on DM3 SPA

### **SLOT 1: 2<sup>nd</sup> 4-port Tomahawk FC IO module (A1=SPA; B1=SPB)**

- Port 0 BE2→to DAE loop 2 SPB & SPA
- Port 1 BE3→to DAE loop 3 SPB & SPA
- Port 2 FE2→to BE1 on DM4 SPB; to BE0 on DM4 SPA
- Port 3 FE3→to BE1 on DM5 SPB; to BE0 on DM5 SPA

### **SLOT 2: 3<sup>rd</sup> 4-port Tomahawk FC IO module (A2=SPA; B2=SPB)**

- Ports 0-3 SPA & SPB for FC Hosts (port 3 for Mirrorview if used)

## **NS-480iS iSCSI MPFS CONFIGURATION:**

- NAS or MPFS over iSCSI configurations
- (2) Tomahawk 4-port FC I/O ports
- (3) Harpoon 2-port iSCSI IO modules (all 6 ports for iSCSI MPFS Hosts)

**Note:** SP configuration based on the standard CLARiiON CX4-480 configuration

→2 Blades only for MPFS (Why? We have the FC ports available?)

→Supports up to 6 direct-attached iSCSI MPFS Hosts or 256 iSCSI-attached MPFS Hosts

→During CSA configuration, a wizard walks through the setup of IP addresses for all iSCSI ports on the Array, offers to setup CHAP, and starts the MPFS service on Server\_2. Use the following command to view IP address assignments on the iSCSI ports:

**# /nas/sbin/naviseccli -h 10.241.168.182 connection -getport -sp a**

**SP IO MODULES for NS-480iS iSCSI:**

**SLOT 0: 1st 4-port Tomahawk FC IO module (A0=SPA; B0=SPB)**

Port 0 BE0→to DAE loop 0 on SPA & SPB

Port 1 BE1→to DAE loop 1 on SPA & SPB

Port 2 FE0→to BE1 SPB on DM2; to BE0 SPA on DM2

Port 3 FE1→to BE1 SPB on DM3; to BE0 SPA on DM3

**SLOT 1: 2<sup>nd</sup> 4-port Tomahawk FC IO module (A1=SPA; B1=SPB)**

Port 0 BE2→to DAE loop 2 on SPA & SPB

Port 1 BE3→to DAE loop 3 on SPA & SPB

Port 2 FE2→to BE1 SPB on DM4; to BE0 SPA on DM4

Port 3 FE3→to BE1 SPB on DM5; to BE0 SPA on DM5

**SLOT 2: 1<sup>st</sup> Harpoon 2-port iSCSI IO module (A2=SPA; B2=SPB)**

iSCSI 0 iSCSI MPFS host

iSCSI 1 iSCSI MPFS host

**SLOT 3: 2<sup>nd</sup> Harpoon 2-port iSCSI IO module (A3=SPA; B3=SPB)**

iSCSI 2 iSCSI MPFS host

iSCSI 3 iSCSI MPFS host

**SLOT 4: 3<sup>rd</sup> Harpoon 2-port iSCSI IO module (A4=SPA; B4=SPB)**

iSCSI 4 iSCSI MPFS host

iSCSI 5 iSCSI MPFS host

**NS-120 & NS-480:**

**RAID SUPPORT:**

RAID 1/0, 5, or 6

**PRODUCT SUPPORT OPTIONS:**

Enhanced or Premium Support

**CRUs, FRUs, UPGRADES:**

**CRUs:**

**Celerra CRUs:**

Blade Power/Cooling Modules

Blades

Blade SFP Modules

**CLARiiON CRUs:**

SP Power/Cooling Modules

SP Management Modules

SP SFP Modules

SPS

Disks

**FRUs:**

**Celerra FRUs:**

Blade Power/Cooling Modules

Blades

Blade SFP Modules

Control Station

Blade Enclosure

**CLARiiON FRUs:**

SP Power/Cooling Modules

SP Management Modules

SP SFP Modules

SP Memory Modules

SP CPU Modules

SP I/O Modules

SPE Enclosure

SPS

DAE Power/Cooling Modules

LCC Modules

Disks

DAE Enclosure (Not posted on Powerlink, use CLARiiON procedure Generator)

**Upgrades:**

Add Blade (NS-120 only)

Upgrade to Mojito Blade (NS-120, NS-480, NS-G2, NX4), Q1 2010

Add SPS (NS-120 only)

Add Disk

Add DAE

FLARE upgrades

DART upgrades (EMC-only)

NS-480 to NS-480FC or NS-480iS, or adding iSCSI slic to NS-480FC Upgrade, March 2009

NS-120 to NS-120FC or NS-120iS, or adding iSCSI slic to NS-120FC Upgrade, March 2009

NS20 to NS-480 Upgrade/Conversion, Dec 2009

NS20FC to NS-480FC Upgrade/Conversion, Dec 2009

NS20 to NS-480iS Upgrade/Conversion, Dec 2009

NS20 to NS-480FC Upgrade/Conversion, Dec 2009

NS20 to NS-480iFC with iSCSI Upgrade/Conversion, Dec 2009

NS-120 to NS-480, Q4 2009

Upgrade to MPFS SW & HW support

Installing SW Enablers on Storage System

**NS-120 FACTORY INSTALL ISSUES:**

→Factory install failed with following console screen messages, even after hardware and cable swapouts

**Console Message during NAS Install:**

**Detecting Data Movers...2 3 Done.**

**Checking Data Mover to Backend fibre connectivity ...**

**Data Mover 2 Port 0 not logged in backend**

**Failed**

**Setting up Storage Groups.... Failed.**

**DEBUGGING FAILED NAS INSTALLS:**

→Above error is rather generic, so additional steps needed to isolate problem

1. Restart the Installation and configure the LAN port during the Linux Install (will need to connect via LAN later)
2. Pause at the following screen after the Linux install and reboot, and after LAN configuration and Data Mover discovery:

Celerra Install/Upgrades +-----+

```
|           |
| Please choose the Celerra System   |
| Configuration.                   |
|                               |
| +-----+ +-----+ |
| | Fibre Channel Enabled | | Integrated | |
| +-----+ +-----+ |
```

3. Use SSH and connect as root to the system

4. vi edit the following S95nas files by adding an -vx to have the Install script output verbose messages to console:

```
# /usr/bin/updatedb
```

```
# locate S95nas
```

```
# vi /etc/rc.d/rc3.d/S95nas
```

```
# vi /mnt/source/EMC/nas/S95nas
```

```
#!/bin/bash -vx
```

5. Reboot the system and set capture logging on terminal session

```
# reboot
```

6. At point of failure, review S95nas install script and find the point where it spits out the failure message, then find debug output failure and try to establish a reason for the failure.

**Root Cause:**

In this situation, when it says the Data Mover BE0 cannot log into the backend, it was true. The following CLARiiON CLI shows that the Celerra ports are not initialized, hence the install problem.

**Solution:**

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
The solution was to reinstall the CLARiiON FLARE on the array, which then allowed the IO ports to be properly seen by the CLARiiON.

### **VIEWING SLIC PORT INFORMATION ON CX4 ARRAYS:**

→Following output is from a CX4-120 array with only a single 4-port Tomahawk FC IO SLIC module. Note that ports 2 & 3 are dedicated for Celerra and are “Uninitialized”, an abnormal condition which will prevent successful NS-120 Factory NAS Install.

**# /tftpboot/bin/naviseclli -h 128.221.252.200 iportconfig -list**

SP ID: A

I/O Module Slot: 0

I/O Module Type: Fibre Channel

I/O Module State: Present

I/O Module Substate: Good

I/O Module Power state: On

I/O Carrier: No

Information about each port on this I/O module:

Physical Port ID: 0

Port Role: BE

Logical Port ID: 0

Port Usage: Normal

Port Type: Fibre Channel

Port State: Enabled

Port Substate: Good

Is Persisted: Yes

Physical Port ID: 1

Port Role: Unknown

Logical Port ID: N/A

Port Usage: Uninitialized

Port Type: Fibre Channel

Port State: Empty

Port Substate: The SFP is not present

Is Persisted: No

Physical Port ID: 2

Port Role: FE

Logical Port ID: N/A

Port Usage: Uninitialized

Port Type: Fibre Channel

Port State: Uninitialized

Port Substate: Uninitialized

Is Persisted: No

Physical Port ID: 3

Port Role: FE

Logical Port ID: N/A

Port Usage: Uninitialized

Port Type: Fibre Channel

Port State: Uninitialized

Port Substate: Uninitialized

Is Persisted: No

### **NORMAL OUTPUT FOR SP FC FE PORTS 2 & 3 for CELERRA:**

**# /nas/sbin/navicli -h 128.221.252.200 iportconfig -list**

**I/O Module Slot: 0**

Information about each port on this I/O module:

**Physical Port ID: 2**

Port Role: FE

Logical Port ID: 0

Port Usage: Normal

Port Type: Fibre Channel

Port State: Enabled

Port Substate: Good

Is Persisted: Yes

**Physical Port ID: 3**

Port Role: FE

Logical Port ID: 1

Port Usage: Special

Port Type: Fibre Channel

Port State: Enabled

Port Substate: Good

Is Persisted: Yes

## **VIEWING iSCSI PORT INFORMATION:**

**# /nas/sbin/navisecli -h 10.241.168.182 connection -getport -sp a**

SP: A

Port ID: 7

Port WWN: iqn.1992-04.com.emc:cx.apm00083202015.a7

iSCSI Alias: 2015.a7

IP Address: 10.241.168.185

Subnet Mask: 255.255.255.0

Gateway Address: 10.241.168.128

Initiator Authentication: false

-----output abridged-----

## **EQUIVALENT VIEW FROM NAVISPHERE:**

Rightclick Array name>Port Management

→Port information is listed, including IP addresses, Port number, type, speed, and WWN/IQN

## **ADDING ULTRAFLEX SLIC IO MODULES ON CX4:**

### **Adding IO Modules to SPs:**

→Must be added to same slot on each SP

→Only safe way to add IO Modules is to use NST wizard (does healthchecks and persists the ports)

→If added via navicli, must run ioportconfig –persist and allow SPs to reboot to persist the ports on the IO card (1/2 hour)

**Note:** SPB reboots first, twice, then SPA, twice, for a total of ½ hour

**Scenario:** Adding Harpoon iSCSI IO card to CX4-480 system using NST

**Note:** This activity requires use of the NST

1. Open NST>login>Hardware Installation>Install I/O modules and/or SFPs, and follow wizard instructions

**Comments:** The wizard reboots SPB a number of times to complete the discovery & initialization of the new IO module, then does the same thing for SPA.

2. Use Celerra Manager after the NST work is completed to failback any failed over Celerra LUNs

### **BEFORE:**

**/nas/sbin/navicli -h <spa\_IP> ioportconfig -list**

SP ID: A

I/O Module Slot: 2

I/O Module Type: Fibre Channel

I/O Module State: Present

I/O Module Substate: Good

I/O Module Power state: On

I/O Carrier: No

Information about each port on this I/O module:

Physical Port ID: 0

Port Role: FE

Logical Port ID: N/A

Port Usage: Uninitialized

Port Type: Fibre Channel

Port State: Uninitialized

Port Substate: Uninitialized →See emc215363, where IO Module was added to array prior to shipment as ‘Mode 2’ system

**Is Persisted: No**

### **AFTER:**

**# /nas/sbin/navicli -h 10.241.168.179 ioportconfig -list -iomodule 2 -sp a -pportid -portrole -porttype -portstate -isportpersisted**

SP ID: A

I/O Module Slot: 2

I/O Module Type: Fibre Channel

I/O Module State: Present

I/O Module Substate: Good

I/O Module Power state: On

I/O Carrier: No

Information about each port on this I/O module:

Physical Port ID: 0

Port Role: FE

Port Type: Fibre Channel

Port State: Enabled

**Is Persisted: Yes**

### **REPLACING FAILED SLIC IO MODULE PER CPG ON CX4:**

#### **Replacing IO Modules on SPs:**

→Basic steps are to shutdown the SP with the bad SLIC, using the shutdownpeerSP CLI command, replace the SLIC, then reboot the Peer SP from the active SP

1. Shutdown the SP that requires IO card replacement from its peer SP:

**# /nas/sbin/navisecl -h <sp\_ip> -user <username> -password <password> -scope 0 shutdownpeerSP**

2. Remove faulty IO module and replace with new module

3. Reboot SP where IO card was replaced from its peer SP:

**# /nas/sbin/navisecl -h <sp\_ip> -user <username> -password <password> -scope 0 rebootpeersp**

4. Verify system from Navisphere

#### **REPLACING SFP MODULES ON SLICs:**

→Basically, SFP replacement is online (remove bad SFP and insert new SFP)

→Important to know whether SFP ports are operating with correct BE or FE designation, and are “initialized”, etc.

#### **REMOVING SLIC IO MODULES:**

→Not a supported activity (see emc215415, module removed from array)

→For End of Quarter shipments, a SLIC IO Module may have been removed from the Array backend to change the model designation from an AUXF to an AUX

→Array logs message about IO Module missing and not inserted, and ioportconfig –list shows IO Module as “Missing”

→While not a harmful message, the only real cure is to destroy the array PSM and let it rebuild, following by a factory NAS install

1. Shutdown Celerra and array

2. Remove IO Modules A2/B2 from CX4-120 system

3. Power-up array, then the Celerra, examine for error messages

→SP Event Log and getlog output says “A configured IO module is not inserted.”, and Navisphere and ioportconfig –list shows module as “Missing”.

4. Reconfigure the array (destroys PSM database):

**# /tftpboot/bin/navicli -h <sp\_ip> ioportconfig –remove**

CAUTION: Both SPs will reboot and all customer data and array configuration WILL BE DESTROYED. Do you wish to continue (y/n)? y

Error: Target system has manually registered initiators. Remove all manual initiators before attempting the command.

**Note:** You must remove all Host Initiator records from the array, unbind all LUNs, destroy all RAID Groups, and destroy the Storage Group before the –remove command will proceed with the array reconfiguration

5. Set SPA/SPB IP & gateway addresses back to Celerra defaults (128.221.252.200 [.104]/128.221.253.201 [.104]

6. Perform Factory NAS install

## **NEW HARDWARE PLATFORMS**

### **PROBES SERIES:**

→Celerra models based on the Argonaut CPU module and Blackwidow Enclosure

### **ARGONAUT FRONT-END:**

→New Blade CPU Module based on Intel Nehalem/Westmere CPUs with 6@8GB DIMM slots & 3 Channels DDR3

## **UNIFIED VCX CELERRA/CLARIION MODELS**

→Initial release called VCX 1.0 [VCX 1.1 will support File Services upgrade]

→Convergence of Clariion and Celerra models for Celerra Integrateds

→Will maintain Celerra models for Gateway environments

→Control LUNs will be factory installed on Vault space (60-90GB) for block services for potential upgrades to file services, in RAID 6 (2+2)

→DART image to be upgraded whenever Flare is upgraded

→If upgrading to File Services, DART can be kept same, or upgraded, Flare may require update depending on NAS version

→Physical space will also be reserved for adding file services

→Reservation of (2) ports on SP's for adding Celerra Blades (file services) to the configuration

→Recoverpoint to be offered as the common File/Block replication solution

→ Celerra Replication to be supported for those requiring file system replication until Recoverpoint is enhanced

→ MVView to be offered in low-end where multi-sites exist

→ NAS & CLARiiON iSCSI to be maintained for now [If >than 8 snaps per LUN is required, Celerra iSCSI can offer up to 1024 snaps, whereas Clariion iSCSI is limited to 8 snaps/lun]

#### **Internal naming convention, from low-end to high-end:**

VCX-75; VCX-125; VCX-250; VCX-500; VCX-1000

#### **Columbus (to replace NS-120):**

→ 1-2 Maynard CS, 1-2 Blades, NAS, iSCSI, & FC protocols, paired with the Hellcat Sentry-Lite array (1.86GHz Nehalem 6GB memory), max 125 drives (SAS/SATA)

→ 2.13GHz Westmere CPU, 6GB memory, 48TB/Blade, (3) SLIC slots per Blade, (4) 6Gbps SAS backend ports per DPE and (4) 8Gbps FC connections

#### **Footprint:**

→ Minimum footprint is 8U (2U DME; 3U DPE; 1U SPS; 1U CS0; 1U CS1)

#### **Huygens (introduces an NS-240):**

→ 1-2 Maynard CS, 1-3 Blades, NAS, iSCSI, MPFS, & FC protocols, using Lightning Sentry array, max 250 drives (SAS/SATA)

→ 2.13GHz Westmere CPU, 12GB memory, 64TB/Blade, (4) SLIC slots per Blade, (2) 6Gbps SAS connections and (4) 8Gbps FC connections

#### **Footprint:**

→ Minimum footprint is 8U (2U DME; 3U DPE; 1U SPS; 1U CS0; 1U CS1)

#### **DAE's:**

→ Viper 6Gbps 3U SAS DAE holding (15) 3.5" drives or (15) 2.5" drives in the 3.5" holders

→ Derringer 6Gbps 2U SAS DAE holding (25) 2.5" drives

→ Voyager D-60 6Gbps 4U SAS DAE holding (60) 2.5" drives

#### **Mariner (replacement for NS-480):**

→ 1-2 Maynard CS, 2, 3, or 4 Blades, NAS, MPFS, iSCSI, & FC protocols, using Spitfire array, max 500 drives (SAS/SATA), 128TB/Blade

→ 2.4GHz 4-core Westmere CPU Argonaut, 12GB memory, 96TB/Blade, (5) SLIC slots per Blade

→ SP's also based on Argonaut, 2.4GHz Westmere with 18GB memory

→ Viper 3U SAS 6Gb/s DAE with [15@3.5"](#) or 2.5" drives; Derringer 2U SAS 6Gb/s DAE with [25@2.5"](#) drives; Voyager D-60 4U SAS 6Gb/s DAE with [60@2.5"](#) drives

→ Total 500 drives supported [108U using 33@15 drive DAEs; 31U using 20@25 drive DAEs; 35U using 8@60 drive DAEs]; max. of 4 BE Loops, each loop can have a max. of 125 drives; Vault drives can be SAS, SATA, EFD

→ Minimum footprint 7U

→ Raid 1/0, 5, 6 for file-based luns; Raid 0, 1, 3, 1/0, 5, 6 for block-based luns

#### **Cassini (replaces NS-960):**

→ 1-2 Maynard CS, 2-8 Blades, NAS, MPFS, iSCSI, & FC protocols, using Spitfire array, max 1000 drives (SAS/SATA)

→ 2.8GHz Westmere CPU, 24GB memory, 256TB/Blade, (5) SLIC slots per Blade

#### **Footprint:**

→ Minimum 10U, maximum 208U totaling 1000 drives

#### **Drive Types:**

EFD, SAS, or SATA, with any type as Vault

#### **DAE's:**

Viper, Derringer, & Voyager

#### **APOLLO--PROPOSED NX4 REPLACEMENT 2010/2011:**

→ Single Maynard CS, 1-2 Blades, using Hellcat-Lite array using max 75 3.5" SAS/SATA/EFD or 2.5" SAS/SATA drives

→ 1.86GHz Nehalem CPU, 8GB memory, 32TB/Blade, (4) SLICs per Blade

→ Footprint: 7U minimum and 19U maximum

## **CELERRA TLC TOOLS:**

**Note:** Several tools have been developed and are now deployed on current releases of 5.4 & 5.5

### **I. Automatic Log Collection & Transfer:**

This tool is especially powerful and triggered by certain NAS Events, such as CallHomes. Its primary purpose is to run the collect\_support\_materials script at the time of the triggering event, place the tar gzip or .zip file in /nas/var/log. It also automatically extracts and gzips panic dump files, putting the header file and dumpfile in /nas/var/dump. Additionally, this feature can be manually configured to automatically transfer the Logs & Dumpfiles to an FTP Server, as defined in the /nas/site/automaticcollection.cfg file. This tool is currently being deployed as part of the Pre-Upgrade Healthcheck Script and will be automatically installed on all NAS 5.1 and higher systems that do not already have the tool installed. See emc135846. With NAS 5.5.27.5 (GNapa), the Transfer feature can be Enabled or Disabled from Celerra Manager>Support>Log Collection tab, and a Log Collection can be triggered by the "Collect" option.

## **NAS 5.6 Manual Log Collection:**

# /nas/tools/automaticcollection -getlogs

Running Master Script

collect\_support\_materials[10342]: The collection script revision 2.9.2 has started.

Collecting output from server\_log

Collecting output from internal commands

Collecting event log configuration files

Collecting files from .etc dir of each DM

Collecting Mirrorview DR logs

Collecting Backend Monitor logs

Collecting /var logs

Collecting upgrade logs

Collecting files from /etc

Collecting files from /proc

Collecting /http/logs and /tomcat/logs

Collecting Celerra Manager tasks

Collecting cron files

Collecting Control Station process information and versions

**Note:** Logs are collected in /nas/var/emcsupport

# nas\_logviewer -t /nas/log/sys\_logtail

Jun 26 02:05:54 2008:96109133825::automaticcollection[10270]: The collection script revision 2.9.2 has started.

Jun 26 02:06:09 2008:96109133826::automaticcollection[10270]: The log collection script finished successfully. A support materials file has been created in /nas/var/emcsupport.

## **II. Celerra Log Parsing Tool:**

Windows-based java application that can be used to automatically extract and present some of the Celerra Server & CS logs in a GUI window from a collect\_support\_materials.tar.gz or .zip file, and merge the contents by chronological date and time. This tool is available for download from the NAS Support website as a zip file—unzip and install using the setup.exe program—review the documentation that is also provided from the zip file to use the Parsing Tool.

## **III. Automatic Change Monitor Tool (aka Backend Monitor—see emc152661):**

A preliminary script version of this tool has been deployed on 5.4 & 5.5 systems via the Pre-Upgrade Healthcheck script. The default configuration creates a nightly Cronjob that assesses backend status and changes related to the Data Movers, and logs changes in the /Celerra/backendmonitor directory. The sys\_log also logs any Errors found and when the backend\_change\_monitor tool starts and completes. The tool is run daily via /nas/site/cron.d/nas\_sys, after each CS reboot, and when Users run # /Celerra/backendmonitor/backend\_change\_monitor.

→Output of command is logged in /celerra/backendmonitor/Program\_Logs/FEATUREOUTPUT...

→Differences between queries are logged in /celerra/backendmonitor/Diff\_Logs/DIFF...

→Errors are logged in /celerra/backendmonitor/Error\_Logs/ALERTOUTPUT...

→Connectivity Checks are run between Data Mover, Backend HBAs, and Switches (if configured)

→Zoning information can be collected from Cisco & McData currently using SNMP with EF6000 MIBs

**Note:** 5.6.38.2 incorporates this feature in NAS, with changes (see emc195237)

## **NAS 5.6 COGNAC:** 5.6.36.4 RTM March 26, 2008; Current release 5.6.50.2

### **KNOWN ISSUES:**

1. /nas/tools/nas\_summary tool in NAS 5.5 but not ported to 5.6--AR95877 (restored in 5.6.38.2)
2. /nas/sbin/cse\_recover utility, used to correct sector errors for LUN/filesystem recovery, left out of 5.6--AR114835 (5.6.37)
3. Backend events not being handled as Events by Celerra /nas/sbin/navilog\_mon script due to parsing issue--AR117752 (5.6.38).
4. Modem ConnectHome failing on 5.6 because modem not being reset as ‘data-only’ but as ‘fax-id’--AR118928 (5.6.38).
5. Foxglove I2C Error mgmt switch firmware issues—see emc219555
6. Sledgehammer & Hammerhead mgmt switch firmware issues during upgrades—see emc223234
7. Disk discovery timeout issues during NAS Upgrades or device adds (see emc222370/AR146711 5.6.44 & 5.6.45)
8. Partimage upgrade failure during first reboot to new Linux on CS for upgrades to 5.6.46—needs primus
8. Add Blade procedure for Single blade systems causes two issues: [see emc185621 for workarounds]
  - a) upgrade\_to\_dual\_intnet script truncates the /etc/modprobe.conf file (In NAS 5.5, was /etc/modules.conf)--AR117560 (5.6.38).
  - b) Add Blade procedure also causes setup\_slot -init failure because of a “dpinit” entry in the /nas/server/slot\_3/eof file, resulting in rootfs not being mounted and then extended--AR117338, emc185621 (fixed 5.6.39).

# /nas/sbin/setup\_slot -x 3 →Extends rootfs from 16MB to 256MB

Mount root filesystem root\_fs\_3 on server\_3

```
server_3 : done
server_3 : done
Extending root_fs_3 from 16 MB to 256 MB ...
c) Add Blade procedure in Cognac also fails to write slot_2's uniqueID to the /nas/server/slot_2/start file:
# /nas/bin/nas_checkup
Check Version: 5.6.39.5
```

-----Errors-----

Data Movers: Check unique id

Symptom:

\* The uniqueid reported by the Data Mover server\_2 does not match the /nas/server/slot\_2/start file.

Action : Contact EMC Customer Service and refer to EMC Knowledgebase emc146016. Include this log with your support request.

#### **PROCEDURE TO ADD UNIQUEID TO SLOT:**

```
# cat /nas/server/slot_2/start
```

```
-serial swapserial
#verify_start
flashupg post=upgrade bios=upgrade
logsys add output disk=root_log_2 bufsz=256
logsys add output event bufsz=2048
ap boot
```

```
# /nas/sbin/setup_slot -u 2
```

Upgrading server in slot 2 ...

Ping server in slot 2 ...ok

server\_2 : done

```
# cat /nas/server/slot_2/start
```

```
-serial swapserial
#verify_start
flashupg post=upgrade bios=upgrade
logsys add output disk=root_log_2 bufsz=256
logsys add output event bufsz=2048
ap boot
```

**uniqueid validate=50060160:41e0c27b**

#### **Linux Kernel Upgrade from 2.4 to 2.6:**

```
# uname -r
```

2.6.9-42.5610.EMC

#### **CREATING BOOT FLOPPY FROM NAS CD-ROM ON NAS 5.6 CONTROL STATION:**

→Floppy is only required for CFS Eagle/Hawk systems

#### **NAS 5.6 Mountpoint Changes for Media on the Control Station:**

/media/floppy

/media/cdrom [for CD-ROM drives]

/media/cdrecorder [for CD-RW drives, such as the NSX]

#### **Symptoms if using wrong mountpoints:**

```
# mount -t msdos /dev/fd0 /mnt/floppy
```

mount: mount point /mnt/floppy does not exist

```
# mount /dev/cdrom /mnt/cdrom
```

mount: mount point /mnt/cdrom does not exist

#### **NAS 5.6 Floppy Image Name Change on NAS CD-ROM:**

/mnt/cdrom/images/boot.img (NAS 5.5 or below)

/media/cdrom/images/floppy.img (NAS 5.6 and above)

#### **NAS 5.6 Boot Floppy Image:**

Since the floppy image contains no recognizable file system that Linux or Windows can detect, the only way to verify the boot floppy image is to use cksum

#### **Boot Floppy Creation & Verification Procedure:**

##### **1. Mount CD-ROM & Floppy on Control Station:**

```
# mount -r /media/cdrom
```

```
# mount -t msdos /dev/fd0 /media/floppy
```

```
# mount
```

/dev/hdd on /media/cdrom type iso9660 (ro,nosuid,nodev)

/dev/fd0 on /media/floppy type msdos (rw)

**2. Create boot floppy system disk:**

```
# dd if=/media/cdrom/images/floppy.img of=/dev/fd0 bs=1024
```

1440+0 records in

1440+0 records out

**Note:** Eventhough the command prompt returns quickly, it takes a few minutes to complete the image write to the Floppy disk

**3. Use checksum to verify the image built to floppy vs. the image on the NAS CD Media:**

```
# dd if=/dev/fd0 of=/tmp/bootflop bs=1024
```

```
# ls -la /tmp/boot*
```

```
-rw-r--r-- 1 root root 1474560 Jan 31 14:36 /tmp/bootflop
```

```
# cksum /tmp/bootflop
```

```
1683387514 1474560 /tmp/bootflop
```

```
# cksum /media/cdrom/images/floppy.img
```

```
1683387514 1474560 /media/cdrom/images/floppy.img -->Checksums match on the file named "bootflop" vs. "floppy.img"
```

**4. Unmount Media from Control Station:**

```
# umount /media/floppy
```

```
# umount /media/cdrom
```

**INSTALLATION/UPGRADE CHANGES/REQUIREMENTS 5.6:**

→Normal upgrade or install mode requires use of bootable CD-ROM

→Recommended to use a serial connection or VGA connection to perform Upgrades to Cognac

→**/EMC/nas/install\_mgr -m upgrade** script replaces traditional /EMC/nas/setup script for upgrades, except for Golden Eagle

→Boot floppy media would be required for CFS-14 Eagle and CFS-SE Hawk systems

→Golden Eagle CNS-14 would run the setup.exe program from the CD mounted to a service Laptop to perform an upgrade or to restart an upgrade, not the install\_mgr script

→In-family upgrades can be run by deploying the package to the /celerra/upgrade directory and running the upgrade from this location [create the directory if needed]

→Can only upgrade directly from NAS 5.4 or higher

→90-day Upgrade moratorium, with exceptions managed by the NAS 5.6 SafeLaunch Team

→PUHC will fail if Flare version is not Flare 24 or higher, requiring Flare upgrade to minimum Flare 24 first, then NAS Upgrade, except that Flare 19 is supported for CX600 / CX400 backends only

→NS600 or 510 Data Mover hardware are not eligible for upgrading to 5.6

→CFS/CNS platforms must have 514 Data Movers to be able to upgrade to Cognac

→Eventhough NAS 5.4 is the minimum upgrade version, an interim upgrade to 5.5.32.4 would be required in order to extend LUNs 0 & 1 if they are not already at the minimize size of 11GB, then upgrade to 5.6

→Either log into Control Station directly as Root user to mount CD-ROM and run install\_mgr –mode upgrade script, or use the steps shown below if getting the following message when logged in as nasadmin & su'd to root:

Error : You have logged in as nasadmin and su'd to root from the /nas

directory.

Action: Retry the upgrade again with the following steps:

1. Exit from the current shell.

exit

2. Change the directory to the upgrade directory.

cd /media/cdrecorder/EMC/nas

3. Obtain root privileges.

su

4. Initiate the upgrade.

./install\_mgr -m upgrade

5. If the previous steps did not solve the problem, look at the following processes list to investigate the issue.

→Upgrades will enable Replication V2 if no V1 sessions or V1 license is detected

→Upgrades are restartable from point of last completed task [e.g., CS rebooted, etc.]

→For unattended installs, can use a DOS formatted floppy for the ksnas.cfg

**Note:** Use the CD-ROM to boot the system, then at the installation prompt, insert the floppy disk that contains the ksnas.cfg file, then enter the following:

**boot:serialkickstart a:\ksnas.cfg**

→A primary CS should be upgraded first, then the Secondary CS

**Upgrade Errata:**

→For lab purposes, there is an option to upgrade to Cognac with older hardware using the following option:

**/EMC/nas/install\_mgr -mode upgrade -allow\_upgrade\_on\_obsolete\_dm**

→Emulex firmware for CFS/CNS cabinets needs to be at a 3.90a7 version for 5.6 Upgrade

**Note:** Verify firmware version using # cat /proc/scsi/lpfc/0. CNS Golden Eagle firmware & CFS-14 Eagle firmware is automatically upgraded by Cognac upgrade process, so this is no longer an open issue for GA release.

→The Floppy disk image itself contains no file system image that Linux or Windows can recognize—it just contains a bootable image called Smart Boot Manager that can then boot the CD-ROM from the CD-Drive. The boot image has been renamed from /images/boot.img to /images/floppy.img for the Cognac NAS media.

→Cognac removes support for Telnet & RSH from Linux, and SSH from Celerra Manager tools

→Rootfs increases to 256MB in size from 128MB; New installs will default back to 1inode/8k block, but upgrades will preserve 1 inode/1k block if already in use. Upgrades to Cognac will automatically extend all production Servers' rootfs, but NOT the Standby server, or any Server with rootfs already extended. Instead, the Standby rootfs would be extended if firstconfigured to be a primary (server\_setup) server, followed by setup\_slot -init. VDM rootfs will remain 128MB.

→NS/NSX platforms will increase Linux swap file from 512 to 2GB on Chivas or 1GB on Falcon Control Stations

**Note:** New swap partition is /dev/hda6 & replaces /dev/hda2

→Linux Boot partition to increase to 128MB in size for NS/NSX platforms

→Work partition magic #: 0xbef0002

### **CONTROL LUN SIZING—LUNs 0, 1, 2, 3, & 4:**

→Install LUNs 0 & 1 must be 11GB, LUNs 2-4 must be 2 GB, and LUN 5 size will be set per the following:

### **CONTROL LUN SIZING—LUN 5:**

NS20/NS40/NS-G2/NX4/NS-120/NS-480=32GB

NS80/NSX/NS-960/NS-G8=64GB

NS350/NS500/70X/CNS/CFS=2GB

→For upgrades, LUNs 0-4 are same as for new install, but LUN5 can remain 2GB

**Note:** The NS-G2 Setup Guide instructs users to create a 64GB LUN5 during the install process

### **SYMMETRIX LUN SIZING ON 5.6 NEW INSTALLS DM3/DM4:**

LUNs 0 & 1 at 12394 cylinders, which equals 11GB

LUNs 2, 3, 4 at 2216 cylinders, 2GB

### **LUN 5 will vary:**

64GB cylinders would be 2@34956 volumes for DM4, or 4@17478 cylinders for DMX3

32GB would use 1@34956 cylinders for DMX3

2GB would use 1@2216 cylinders

### **UPGRADING FROM NAS 5.5 to 5.6:**

1. Log into Control Station as root [cannot be logged in as nasadmin]

2. Mount CD-ROM and run upgrade script

**# mount /dev/cdrom /mnt/cdrom**

**#cd /mnt/cdrom/EMC/nas; ./install\_mgr –mode upgrade** [or –mode=upgrade]

**Note:** Upgrade performs setup\_enclosure –checksystem and PUHC check\_nas\_upgrade healthchecks, then recommends that the upgrade be run from the Serial Console so as to observe the progress.

3. Message comes up advising of the need to do a system reboot after the pre-install checks are completed, along with important information that a User will need to note in order to proceed

a. System will be offline for about 35 minutes while the Linux kernel upgrades and then eventually returns to a login prompt

b. Login as Root, mount CD-ROM, exit to nasadmin, cd to CD-ROM directory, su to root, run upgrade command, upgrade begins

```
$ su  
# mount -r /media/cd*  
# exit  
$ cd /media/cd*/EMC/nas  
$ su  
# ./install_mgr -m upgrade
```

4. After the upgrade/install is completed, don't forget to log out and back in as nasadmin to re-establish NASDB environment in order to run nas and server commands

### **Upgrading to Cognac ‘Console Screen Message’ after Initial Upgrade Steps:**

The Control Station must be rebooted now. After the reboot, the Control Station will remain offline while EMC Celerra Control Station Linux is upgraded. The Linux upgrade could take up to 20 minutes to complete. You will not be able to log into the Control Station during this time. After the Linux upgrade completes, the Control Station will boot back up automatically. To continue with the NAS upgrade, do the following:

1. When the Control Station finishes booting, login as "root".

2. Mount the NAS CD and make it your working directory:

```
mount -r /media/cd*  
cd /media/cd*/EMC/nas
```

3. Continue the upgrade with the command: ./install\_mgr -m upgrade

### **NAS 5.6 UPGRADE/INSTALL LOGS:**

#### **Upgrade logs:**

/var/tmp/upgrade.log [temp log during the upgrade process]

/var/tmp/upgrade\_trace.log [temp trace log during upgrade process]  
/nas/log/upgrade.<nas\_version>.<timestamp>.log [permanent and completed upgraded log]  
/nas/log/upgrade\_trace.5.6.42-5.Mar-12.log  
/nas/log/install.<nas\_version>.<timestamp>.log

→Upgrade cookies are used to track state, copied to two locations on the Control Station, and deleted after the upgrade

#### **/var/log/install\_mgr\_info**

#### **/boot/install\_mgr\_info**

### **NAS 5.6 UPGRADE RECOVERY: Rescue Mode**

→NAS CD can boot into rescue mode to run linux commands—to be used for recovery. 5.5 to 5.6 upgrade can be rolled back to state just prior to upgrade using automatic backups created by the 5.6 script

→Not to be run on CNS-14 Golden Eagle or Secondary Control Station

#### **boot: serialrescue**

Accept ‘English’, answer ‘No’ for network interface setup, select ‘Skip’ for finding Linux installation, and ‘OK’ at /mnt/sysimage mounted message. Start recovery using /recover script.

#### **REVERTING ONCE UPGRADE WAS STARTED:**

→Do not use without consultation with Support and EE, as the suggested steps are only applicable in certain situations

1. Manually mount nas partitions
2. vi edit /nas/site/motd and /etc/motd.standby files, and remove all lines with “Warning!!”

3. Rename temp logs

```
# mv /var/tmp/upgrade.log /nas/log/upgrade.<code_date>.log  
# mv /var/tmp/upgrade_trace.log /nas/log/upgrade_trace.<code_date>.log
```

4. Remove .install\_mgr\_info file

```
# rm /boot/.install_mgr_info
```

```
# rm /var/log/.install_mgr_info
```

5. vi edit /var/sadm/pkg/emcnas/pkginfo

Change version line back to code at beginning of upgrade attempt

6. Run nas\_version and make sure no warning messages are reflected

### **CELERRA LOGICAL VOLUME MANAGER (Celerra LVM) 5.6:**

→Provides logical layer between physical disk/LUN and Celerra volumes/file systems, which allows physical devices to be put into pools or “Volume Groups”. This allows Logical Volumes to span multiple physical devices.

→Allows for dynamic management of Logical Volumes

#### **LVM HEALTHCHECK TOOL:**

→Checks mount status and state of LVMs [checks for unknown or unexpected LVMs, size mismatches, missing LVMs, etc.]

#### **# /nasmcd/sbin/lvm\_checkup**

Usage:

```
lvm_checkup -check <lv | all | error> | -recover <LVMTAB_file> | -help
```

Arguments:

-check lv - Only list the status of logical volume.

error - Only list the errors in LVM health check.

all - List all information in LVM health check.

-recover <LVMTAB\_file> Recover the /etc/lvmtab with LVMTAB\_file.

-help Display the contents of this help message.

#### **# /nasmcd/sbin/lvm\_checkup -check lv**

| LOGICAL VOLUME             | DIRECTORY              | SIZE(MB) | MOUNT | LVM STATUS |
|----------------------------|------------------------|----------|-------|------------|
| emc_lv_home                | /home                  | 600      | YES   | OK         |
| emc_lv_celerra_backup      | /celerra/backup        | 840      | YES   | OK         |
| emc_lv_nbsnas_jserver      | /nbsnas/jserver        | 1416     | YES   | OK         |
| emc_lv_nas_jserver         | /nas/jserver           | 1416     | YES   | OK         |
| emc_lv_nas_var             | /nas/var               | 100      | YES   | OK         |
| emc_lv_nas_var_dump        | /nas/var/dump          | 1692     | YES   | OK         |
| emc_lv_nas_var_auditing    | /nas/var/auditing      | 120      | YES   | OK         |
| emc_lv_nas_var_backup      | /nas/var/backup        | 840      | YES   | OK         |
| emc_lv_nas_var_emcsupport  | /nas/var/emcsupport    | 560      | YES   | OK         |
| emc_lv_nas_var_log         | /nas/var/log           | 212      | YES   | OK         |
| emc_lv_celerra_backendmo.. | /celerra/backendmoni.. | 8        | YES   | OK         |

#### **# /nasmcd/sbin/lvm\_checkup -check all**

LVM Version: 1.1

LVM Checksum: OK

NAS Platform: NS/NSX

CS Slot: 0

| LOGICAL VOLUME             | DIRECTORY              | SIZE(MB) | MOUNT | LVM STATUS |
|----------------------------|------------------------|----------|-------|------------|
| emc_lv_home                | /home                  | 600      | YES   | OK         |
| emc_lv_celerra_backup      | /celerra/backup        | 840      | YES   | OK         |
| emc_lv_nbsnas_jserver      | /nbsnas/jserver        | 1416     | YES   | OK         |
| emc_lv_nas_jserver         | /nas/jserver           | 1416     | YES   | OK         |
| emc_lv_nas_var             | /nas/var               | 100      | YES   | OK         |
| emc_lv_nas_var_dump        | /nas/var/dump          | 1692     | YES   | OK         |
| emc_lv_nas_var_auditing    | /nas/var/auditing      | 120      | YES   | OK         |
| emc_lv_nas_var_backup      | /nas/var/backup        | 840      | YES   | OK         |
| emc_lv_nas_var_emcsupport  | /nas/var/emcsupport    | 560      | YES   | OK         |
| emc_lv_nas_var_log         | /nas/var/log           | 212      | YES   | OK         |
| emc_lv_celerra_backendmo.. | /celerra/backendmoni.. | 8        | YES   | OK         |

Logical Volume Summary:

- \* emc\_lv\_home is ACTIVE, Mounted, OK
- \* emc\_lv\_celerra\_backup is ACTIVE, Mounted, OK
- \* emc\_lv\_nbsnas\_jserver is ACTIVE, Mounted, OK
- \* emc\_lv\_nas\_jserver is ACTIVE, Mounted, OK
- \* emc\_lv\_nas\_var is ACTIVE, Mounted, OK
- \* emc\_lv\_nas\_var\_dump is ACTIVE, Mounted, OK
- \* emc\_lv\_nas\_var\_auditing is ACTIVE, Mounted, OK
- \* emc\_lv\_nas\_var\_backup is ACTIVE, Mounted, OK
- \* emc\_lv\_nas\_var\_emcsupport is ACTIVE, Mounted, OK
- \* emc\_lv\_nas\_var\_log is ACTIVE, Mounted, OK
- \* emc\_lv\_celerra\_backendmonitor is ACTIVE, Mounted, OK

**# /nasmcd/sbin/lvm\_checkup -check error**

No error

**Troubleshooting LVMs:**

- LVM configurations should never be changed from the install or upgrade code version defaults—use lvm\_checkup to verify
- Recommendation would be to use /nas/sbin nas stop and start to unmount and remount LVMs
- LVMs are only mounted by the Control Station acting as the Primary
- NASDB backups do capture the new LVM files show below

**Updating LVM Cache:**

```
# mv /etc/lvm.cache /etc/lvm.cache_bak
# /usr/sbin/pvscan
PV /dev/ndf  VG emc_vg_lun_5   lvm2 [32.00 GB / 30.32 GB free]
PV /dev/hda7  VG emc_vg_pri_ide lvm2 [194.35 GB / 190.93 GB free]
PV /dev/nda3  VG emc_vg_lun_0   lvm2 [5.99 GB / 2.84 GB free]
Total: 3 [232.33 GB] / in use: 3 [232.33 GB] / in no VG: 0 [0 ]
```

**/etc/lvm**

-rw----- 1 root root 1744 Oct 5 03:31 .cache

**Troubleshooting LVM Issues:****/var/log/nas\_lvm.log****Various LVM Configuration Files & Locations:****# ls -la /etc/\*lvm\***

```
-r-xr-xr-x 1 root root 6366 Nov 13 19:07 /etc/lvmtab
-r-xr-xr-x 1 root root 6366 Dec 17 14:34 /etc/lvmtab.save
-rw-r--r-- 1 root root  83 Nov 28 12:10 /etc/nas_lvm_device.map
```

**/etc/lvm:**

```
drwx----- 2 root root 4096 Nov 27 06:03 archive
drwx----- 2 root root 4096 Nov 27 06:03 backup
-rw----- 1 root root 260 Dec 17 14:34 .cache
-rw-r--r-- 1 root root 10613 Nov 28 12:07 lvm.conf
```

**/etc/lvm/lvm.conf****/etc/nas\_lvm\_device.map****/etc/lvmtab****# cat nas\_lvm\_device.map**

LVM\_PRI\_IDE=/dev/hda7

LVM\_SEC\_IDE=/dev/hda7

LVM\_LUN\_0=/dev/ndaa3

LVM\_LUN\_5=/dev/ndf

**Note:** Do not edit LVM information in the above files

**Outputting Volume Group Information:**

# /usr/sbin/vgdisplay

--- Volume group ---

VG Name emc\_vg\_lun\_5

System ID

Format lvm2

Metadata Areas 1

Metadata Sequence No 5

VG Access read/write

VG Status resizable

MAX LV 0

Cur LV 4

Open LV 4

Max PV 0

Cur PV 1

Act PV 1

VG Size 64.00 GB

PE Size 4.00 MB

Total PE 16383

Alloc PE / Size 428 / 1.67 GB

Free PE / Size 15955 / 62.32 GB

VG UUID glOW53-QCES-kpvV-vckf-33Uz-6PMz-DOqkgx

-----abridged-----

**List of LVMs on Celerra Control Station:**

# /usr/sbin/lvdisplay |grep "LV Name"

|         |                                                   |
|---------|---------------------------------------------------|
| LV Name | /dev/emc_vg_lun_5/emc_lv_nas_var                  |
| LV Name | /dev/emc_vg_lun_5/emc_lv_nas_var_backup           |
| LV Name | /dev/emc_vg_lun_5/emc_lv_nas_var_log              |
| LV Name | /dev/emc_vg_lun_5/emc_lv_nas_var_emcsupport       |
| LV Name | /dev/emc_vg_pri_ide/emc_lv_home                   |
| LV Name | /dev/emc_vg_pri_ide/emc_lv_nas_jserver            |
| LV Name | /dev/emc_vg_pri_ide/emc_lv_celerra_backup         |
| LV Name | /dev/emc_vg_pri_ide/emc_lv_celerra_backendmonitor |
| LV Name | /dev/emc_vg_pri_ide/emc_lv_celerra_ccc            |
| LV Name | /dev/emc_vg_lun_0/emc_lv_nas_var_dump             |
| LV Name | /dev/emc_vg_lun_0/emc_lv_nbsnas_jserver           |
| LV Name | /dev/emc_vg_lun_0/emc_lv_nas_var_auditing         |

**Above List should match what is in /dev/mapper:**

# ls -l /dev/mapper

|            |                                                                             |
|------------|-----------------------------------------------------------------------------|
| brw-rw---- | 1 root disk 253, 7 Feb 8 11:35 emc_vg_lun_0-emc_lv_nas_var_auditing         |
| brw-rw---- | 1 root disk 253, 6 Feb 8 11:35 emc_vg_lun_0-emc_lv_nas_var_dump             |
| brw-rw---- | 1 root disk 253, 1 Feb 8 11:35 emc_vg_lun_0-emc_lv_nbsnas_jserver           |
| brw-rw---- | 1 root disk 253, 5 Feb 8 11:35 emc_vg_lun_5-emc_lv_nas_var                  |
| brw-rw---- | 1 root disk 253, 8 Feb 8 11:35 emc_vg_lun_5-emc_lv_nas_var_backup           |
| brw-rw---- | 1 root disk 253, 9 Feb 8 11:35 emc_vg_lun_5-emc_lv_nas_var_emcsupport       |
| brw-rw---- | 1 root disk 253, 10 Feb 8 11:35 emc_vg_lun_5-emc_lv_nas_var_log             |
| brw-rw---- | 1 root disk 253, 3 Jan 8 16:46 emc_vg_pri_ide-emc_lv_celerra_backendmonitor |
| brw-rw---- | 1 root disk 253, 2 Jan 8 16:46 emc_vg_pri_ide-emc_lv_celerra_backup         |
| brw-rw---- | 1 root disk 253, 11 Feb 8 11:35 emc_vg_pri_ide-emc_lv_celerra_ccc           |
| brw-rw---- | 1 root disk 253, 0 Jan 8 16:46 emc_vg_pri_ide-emc_lv_home                   |
| brw-rw---- | 1 root disk 253, 4 Feb 8 11:35 emc_vg_pri_ide-emc_lv_nas_jserver            |

**LVM's are further grouped under the following:**

# /usr/sbin/vgscan

Reading all physical volumes. This may take a while...

Found volume group "emc\_vg\_lun\_5" using metadata type lvm2

Found volume group "emc\_vg\_pri\_ide" using metadata type lvm2

Found volume group "emc\_vg\_lun\_0" using metadata type lvm2

**Verify that LVMs within a Volume Group are all available:**

# /usr/sbin/lvdisplay emc\_vg\_lun\_5

```
--- Logical volume ---
LV Name          /dev/emc_vg_lun_5/emc_lv_nas_var
VG Name          emc_vg_lun_5
LV UUID          knke62-eZ9h-hPzM-gCyj-vTms-Rcm3-PPBwJo
LV Write Access   read/write
LV Status        NOT available → This should read “available” and indicates a problem
```

→ See emc206823 for more information

# /usr/sbin/vgchange -a y emc\_vg\_lun\_5

# /nas/sbin/nas\_lvm\_tool --activate-mount

**Note:** Above commands should not be run without TechSupport and/or Eng. consultation. Vgchange can add a Volume Group back to /dev/mapper directory, and nas\_lvm\_tool can mount an LVM partition, though the normal method for mounting or unmounting LVMs is through the use of NAS services.

**5.6 LINUX CONTROL STATION PARTITION & CONTROL LUN LAYOUT--NS/NSX SYSTEM:**

# df -h | df -P -l | df -l

# df -lhP

| Filesystem                                               | Size  | Used | Avail | Use% | Mounted on              |                                                        |
|----------------------------------------------------------|-------|------|-------|------|-------------------------|--------------------------------------------------------|
| /dev/hda3                                                | 2.0G  | 877M | 1.1G  | 46%  | /                       | →Linux root partition 2GB                              |
| /dev/hda1                                                | 122M  | 4.0M | 112M  | 4%   | /boot                   | →Linux boot partition LUN 2 128MB (from 31MB)          |
| none                                                     | 1010M | 0    | 1010M | 0%   | /dev/shm                | →Linux shared memory partition (up to 2GB on newer CS) |
| /dev/nde1                                                | 1.8G  | 683M | 1003M | 41%  | /nbsnas                 | →LUN 4 2GB                                             |
| /dev/hda5                                                | 2.0G  | 570M | 1.4G  | 30%  | /nas                    | →2GB Linux partition LUN 2                             |
| /dev/mapper/emc_vg_pri_ide-emc_lv_home                   | 591M  | 27M  | 535M  | 5%   | /home                   | →Linux partition LUN 2 600MB                           |
| /dev/mapper/emc_vg_pri_ide-emc_lv_celerra_backup         | 827M  | 64M  | 722M  | 9%   | /celerra/backup         | →Linux partition LUN 2 840MB                           |
| /dev/mapper/emc_vg_pri_ide-emc_lv_celerra_backendmonitor | 7.8M  | 1.2M | 6.3M  | 16%  | /celerra/backendmonitor | →LUN 2 8MB                                             |
| /dev/mapper/emc_vg_lun_0-emc_lv_nbsnas_jserver           | 1.4G  | 45M  | 1.3G  | 4%   | /nbsnas/jserver         | →LUN 0 root_disk 1.38GB                                |
| /dev/mapper/emc_vg_pri_ide-emc_lv_nas_jserver            | 1.4G  | 45M  | 1.3G  | 4%   | /nas/jserver            | →Linux partition LUN 2 1.38GB                          |
| /dev/mapper/emc_vg_lun_5-emc_lv_nas_var                  | 97M   | 5.6M | 87M   | 7%   | /nbsnas/var             | →LUN 5 100MB                                           |
| /dev/mapper/emc_vg_lun_0-emc_lv_nas_var_dump             | 1.7G  | 35M  | 1.6G  | 3%   | /nbsnas/var/dump        | →LUN 0 root_disk 1.65GB                                |
| /dev/mapper/emc_vg_lun_0-emc_lv_nas_var_auditing         | 117M  | 5.6M | 105M  | 6%   | /nbsnas/var/auditing    | →LUN 0 root_disk 120MB                                 |
| /dev/mapper/emc_vg_lun_5-emc_lv_nas_var_backup           | 827M  | 65M  | 720M  | 9%   | /nbsnas/var/backup      | →LUN 5 840MB                                           |
| /dev/mapper/emc_vg_lun_5-emc_lv_nas_var_emcsupport       | 552M  | 17M  | 507M  | 4%   | /nbsnas/var/emcsupport  | →LUN 5 560MB                                           |
| /dev/mapper/emc_vg_lun_5-emc_lv_nas_var_log              | 206M  | 5.8M | 189M  | 3%   | /nbsnas/var/log         | →LUN 5 212MB                                           |
| /dev/ndal                                                | 134M  | 50M  | 84M   | 38%  | /nbsnas/dos             | →LUN 0 root_disk 134MB                                 |

**Note:** In dual CS environments, both CS will have access to all the LVM's, but only one CS can mount.

**OTHER LUN 0 root\_disk USAGES:**

Workpart: 1MB

Dart rootFS: 16MB/DM

Dart Server Log: 4MB/DM

Partial Dump: Single dump\_slot @2MB per DM

Full Dump\_slots: 3@1.5GB shared by all DMs

LVM Reserve Pool: 2.84GB

**LUN 1 root\_Idisk USAGES:**

DART UFSLog: 64MB/DM

DART rootFS ext: 112MB \* #of DMs

DART panic handler: 8MB \* # of DMs

DART rootFS ext2: 128MB \* # of DMs

Free Space: 6.6GB

**LUN 2 Linux CS:**

Swap file: 1 or 2GB

Reserved: 384MB

LVM Reserve Pool: 27-67GB(?)

Old /boot reserved: 31MB

**LUN 5:**

Overall size will vary: For NS20/NS40 LUN 5 size will be 32GB; For NSX/NS80 systems LUN 5 size will be 64GB.

LVM Reserve Pool

# /sbin/fdisk -l /dev/hda

Disk /dev/hda: 80.0 GB, 80000000000 bytes

255 heads, 63 sectors/track, 9726 cylinders

Units = cylinders of 16065 \* 512 = 8225280 bytes

| Device    | Boot | Start | End  | Blocks    | Id | System                               |
|-----------|------|-------|------|-----------|----|--------------------------------------|
| /dev/hda1 | *    | 1     | 16   | 128488+   | 83 | Linux                                |
| /dev/hda2 |      | 17    | 69   | 425722+   | 83 | Linux                                |
| /dev/hda3 |      | 70    | 330  | 2096482+  | 83 | Linux                                |
| /dev/hda4 |      | 331   | 9726 | 75473370  | f  | W95 Ext'd (LBA)                      |
| /dev/hda5 |      | 331   | 591  | 2096451   | 83 | Linux                                |
| /dev/hda6 |      | 592   | 852  | 2096451   | 82 | Linux swap                           |
| /dev/hda7 |      | 853   | 9726 | 71280373+ | 8e | Linux LVM → New partition for Cognac |

# df -h |grep "nbsnas"

|           |      |      |      |     |             |
|-----------|------|------|------|-----|-------------|
| /dev/nde1 | 1.8G | 786M | 901M | 47% | /nbsnas     |
| /dev/hda5 | 2.0G | 652M | 1.3G | 35% | /nas        |
| /dev/nda1 | 134M | 50M  | 84M  | 38% | /nbsnas/dos |

## NAS 5.6 COGNAC TRIVIA:

→root\_fs\_common [./etc\_common] uses root\_volume\_16 & root\_slice\_16

→Data Mover rootfs will be extended by 128MB to a new size of 256MB

Note: Upgrade will extend all Primary Server rootfs automatically, but not the Standby Server. If a Standby is converted to a Primary, its rootfs will be extended when running setup\_slot

→VDM rootfs's are not extended and remain at 128MB in size

→New install will use default 1k per 8k block inode density on rootfs

→JRE 1.5.0\_12 has been qualified for client use with Celerra manager

→Server Logs increasing to 4MB in size per DM—Block size 0x1FFF

→(3) full dump slots are 1.5GB in size—Block size 0x300004

→DM's reducing partial dump slots from (2) each to (1) partial dumpslot of 2MB for each DM—Block size 0xFFFF

→2<sup>nd</sup> copies of the NASDB backup have migrated from /home/nasadmin to /celerra/backup/

→Primary copies of NASDB Backups are still written to /nbsnas/var/backup/

→CD mountpoint now /media/cdrom

→Processes are displayed differently than the previous Linux version, as seen by NASMCD & Box Monitor processes:

# ps -ef |grep nas\_m

|      |       |   |           |          |                                                 |
|------|-------|---|-----------|----------|-------------------------------------------------|
| root | 29424 | 1 | 0 Feb21 ? | 00:00:05 | /nasmcd/nas_mcd -h /nasmcd /nas/sys/nas_mcd.cfg |
|------|-------|---|-----------|----------|-------------------------------------------------|

# ps -ef |grep boxm

|      |      |       |           |          |                                    |
|------|------|-------|-----------|----------|------------------------------------|
| root | 1970 | 29424 | 0 Feb21 ? | 00:00:27 | /nas/sbin/nas_boxmonitor /nas -i 8 |
|------|------|-------|-----------|----------|------------------------------------|

## Outputting NAS Events different in 5.6:

# nas\_event -list -c -info [Lists components]

# nas\_event -l -c DART [Lists facilities within a component]

# nas\_event -l -c DART -f DBMS [Lists events for a facility]

# nas\_event -list -a -info [List of possible actions for events]

# nas\_event -list -a callhome [List of events for a particular action]

## 5.6 Location Where Collect Support Materials & Other Log Collection Scripts are Written:

### Collect Support Materials:

→If transfer feature is not enabled, the collect\_support\_materials file is created and zipped to /nas/var/emcsupport

# pwd

/nas/var/emcsupport

# ls -la

-rw-rw-r-- 1 nasadmin nasadmin 8453115 Feb 27 08:43 support\_materials\_APM00073701085.080227\_0841.zip

# nas\_logviewer /nas/log/sys\_log |grep collect

Feb 27 08:43:52 2008:CS\_PLATFORM:LogCollect:INFO:2:::1204119832:automaticcollection[14927]: The log collection script finished successfully. A support materials file has been created in /nas/var/emcsupport.

### SPCollects:

/nas/var/log

→SPCollects gathered from /nas/tools/.get\_spcollect are written to this directory

## 5.6 NAS LOGVIEWER & SYS LOG:

### NEW FORMAT FOR SYS LOG:

The /nas/log/sys\_log is no longer a simple text file and must be viewed using a special nas\_logviewer utility

# tail /nas/log/sys\_log

120602527996108871986

120602887696108871980

120602887796108871985

120602888196108871986

# file /nas/log/eventstore/slot\_1/sys\_log →Actual location of sys\_log

/nas/log/eventstore/slot\_1/sys\_log: data

# nas\_logviewer -v /nas/log/sys\_log →Verbose mode, shows log messages in CCMD style

logged time = Mar 20 14:01:22 2008

creation time = Mar 20 14:01:22 2008

slot id =

id = 96108871986

severity = INFO

component = CS\_PLATFORM

facility = NASDB

baseid = 306

type = EVENT

brief description = nasdb\_backup: Celerra database backup done.

full description = The Celerra database backup process has created an archive of the current Celerra database files.

recommended action = No further action required.

# nas\_logviewer -t /nas/log/sys\_logmore → -t switch is terse format, more traditional sys\_log style, drops some facility and event severity info

Mar 17 14:11:27 2008:94763810819:2:The Usermapper service has been disabled.

Mar 17 14:11:32 2008:94763810820:2:The Usermapper database has been destroyed.

Mar 17 15:01:19 2008:96108871980::nasdb\_backup: NAS\_DB checkpoint is in progress.

Mar 17 15:01:20 2008:96108871985::nasdb\_backup: NAS\_DB checkpoint done.

# nas\_logviewer -f /nas/log/sys\_log →Opens log file and current events will be seen as they are entered in the log

# nas\_logviewer /nas/log/sys\_logmore →With no switches defined, adds facility & event severity information

Mar 17 14:11:27 2008:DART:USRMAP:INFO:3:Slot 2::1205777487:The Usermapper service has been disabled.

Mar 17 14:11:32 2008:DART:USRMAP:INFO:4:Slot 2::1205777492:The Usermapper database has been destroyed.

→Fortunately, the collect\_support\_materials script not only gathers up all the sys\_logs, but converts and outputs each sys\_log into a text file based on the nas\_logviewer -t terse output mode:

sys\_log.1.txt

sys\_log.2.txt

sys\_log.3.txt

sys\_log.4.txt

## LUN 0 WORKPART OVERVIEW:

Lun 0 serves as the partition table for the Workpart structure

DOS partition →11GB

Workpart → 1MB

RootFS →16@16MB

Server Log →16@4MB

Partial dumps →16 (2)MB slots, one partial per Server

Full dumps →3@1.5GB shared between all Servers

## LINUX LOCATE UTILITY:

With NAS 5.6 kernel update, the ‘locate’ utility is a symbolic link to ‘slocate’, which is more secure since it adds file permissions and ownership information to the slocate.db database, meaning that Users that do not have the permissions to access certain files will not see the files when running the “locate” utility.

### Locate and Updatedb are Symbolic Links to the Slocate Executable:

# ls -la /usr/bin/updatedb

lrwxrwxrwx 1 root slocate 7 May 6 09:10 /usr/bin/updatedb -> slocate

# ls -la /usr/bin/locate

lrwxrwxrwx 1 root slocate 7 May 6 09:10 /usr/bin/locate -> slocate

### Cognac does not create or update the slocate database automatically, as seen by a couple of different output messages:

# locate updatedb

warning: locate: warning: database /var/lib/slocate/slocate.db' is more than 8 days old

# locate pxenas

warning: locate: could not open database: /var/lib/slocate/slocate.db: No such file or directory

warning: You need to run the 'updatedb' command (as root) to create the database.

Please have a look at /etc/updatedb.conf to enable the daily cron job.

### Manually creating the slocate database:

/usr/bin/updatedb

### Location of slocate database:

/var/lib/slocate

**Location of slocate daily cron job:**

**/etc/cron.daily/slocate.cron**

**Use following workaround to setup and keep the slocate.db current:**

1. Verify the existence of the slocate database:

```
# ls -la /var/lib/slocate
```

```
-rw-r----- 1 root slocate 780388 May 19 11:56 slocate.db
```

2. If the slocate.db does not exist, create it using the following command:

```
# /usr/bin/updatedb (updatedb is a symbolic link for the slocate executable)
```

3. Edit the /etc/updatedb.conf file to activate the daily cron Job which will enable daily updates to the slocate.db database:

```
# vi /etc/updatedb.conf
```

```
# To enable the updatedb in cron, set DAILY_UPDATE to yes -->Edit this line and change no to yes to enable the daily cronjob to update the slocate.db:
```

```
DAILY_UPDATE=yes
```

```
PRUNEFS="sysfs selinuxfs usbdevfs devpts NFS nfs nfs4 afs fs proc smbfs cifs autofs auto iso9660 udf"
```

```
PRUNEPATHS="/tmp /usr/tmp /var/tmp /afs /net /fs /selinux /udev /mnt/floppy /media"
```

```
export PRUNEFS
```

```
export PRUNEPATHS
```

4. Reboot the Control Station to activate the revised cron schedule.

**Note:** /etc/cron.daily/slocate.cron job will run daily at 4:02 a.m. if the DAILY\_UPDATE is set to yes in the /etc/updatedb.conf file

**NAS 5.6 PAHC CHECKS:**

→LUN 0 & 1 11GB minimum, LUNs 2, 3, 4, 5 2GB

**# /nas/sbin/rootnas\_disk -i root\_disk**

```
id      = 1
```

```
name    = root_disk
```

```
acl     = 0
```

```
in_use  = True
```

```
size (MB) = 11263
```

**# /nas/sbin/rootnas\_volume -s root\_disk**

```
total = 11263 avail = 0 used = 11263 ( 100% ) (sizes in MB)
```

→Verify consistency of workpart

→Verify that root\_disk\_reserve is not in use by other slices and offset from partial dump matches

**# /nas/sbin/workpart -r lfgrep dump\_slices[0][0]**

```
dump_slices[0][0].lba = 0xd3800
```

**# /nas/sbin/rootnas\_volume -i root\_disk\_reserve |grep offset**

```
offset(MB) = 423
```

**Checking Offset value:** 0xd3800 / 2048 = 423

→LUNs owned by SPA

→LVM checks

**Location of CallHome Files:**

**/nas/opt/connectemc/recycle**

```
ls
```

```
RSC_APM00062405681_042007_100750260.xml
```

**Sys log viewer for new CCMD format:**

**# /nas/bin/nas\_logviewer -v /nas/log/sys\_log >system.log** (pulls all events out of the sys\_log and writes to a <filename>)

**NAS 5.6 CELERRA REPLICATOR V2 (IPREPV2):**

→RepV2 is an IP-based asynchronous remote replication utility based on Snapsure checkpoint technology

**Note:** Currently, V1-to-V2 upgrade required at 5.6.42 or higher—see CPG

→Replication can take place between File Systems, VDMs, or iSCSI LUNs

→Common base refers to base object on source or snap on destination side that matches, so that only changes need to be transmitted

→V2 replication based on Snapsure checkpoints, with each Replication session creating two internal checkpoints on the Source and on the Target systems

→RepV2 shares the same SavVol per file system as used by Snapsure for Checkpoints, etc.

→Since replication sessions flow directly between the Data Movers involved, there is no longer a “Playback Service”

**Some Terms:**

**Checkpoint**—Read-only logical point in time image of a file system—a snapshot or checkpoint

**Internal Checkpoints**—Read-only logical point in time image of a file system using a common base. Two internal checkpoints are created on the Source and two on the Destination for each replication session.

**Writable checkpoint**—Read-only checkpoint that is converted to Read-write

**Common Base**—the Internal Checkpoint common to both Source & Destination, used as a base for the next differential transfer of data from Source to Destination

**Delta**—Block changes of current internal checkpoint vs. the last replicated internal checkpoint. Replication transfers the delta over and refreshes the destination internal checkpoint.

**Differential**—difference between the common base and the changes accumulating on the source file system.

**nas\_copy**—command which copies the full checkpoint over to destination

#### **IMPORTANT REPLICATION PORTS:**

→Data Movers must have ports 8888 and 8887 available over the network

→Control Stations must have port 443 for HTTPS available over the network

#### **UPGRADING FROM V1 to V2 REPLICATION:**

##### **Reference the ‘Celerra Procedure Generator’:**

“Replicator V1 to V2 Upgrade Procedure” (See Standalone Procedures section)

##### **REP V1 to V2 UPGRADE SCRIPT:**

→Upgrades all Replication V1 file system and VDM sessions to V2, not iSCSI replication

→Can run script on Source or Destination side, but if restarting script, use same Control Station

→Must be running NAS 5.6.42 or higher

→All V1 sessions must have a valid restart checkpoint on source PFS

→Upgrade script will abort if fs\_copy is running

→Cannot run both V1 & V2 simultaneously

→Works for Local, Loopback, & Remote replication sessions, but the upgrade\_repv1\_to\_repv2 script must reside in /nas/sbin on each Control Station where RepV1 is running

**/nas/sbin/upgrade\_repv1\_to\_repv2**

**/nas/site/RepUpgradeV1V2.log**

**/nas/volume/RepUpgradeDB** →Replication database stored here, used if script needs to be restarted

##### **Restarting Script:**

If the Control Station was rebooted while the conversion was in progress, restart the script with the following option:

**/nas/sbin/upgrade\_repv1\_to\_repv2 -reboot\_recover**

##### **Basic Script Upgrade Steps:**

Basic steps are to Query for all V1 sessions first, do Verification check [e.g., ensures existence of restart checkpoints, consistent databases, minimum 5.6.38 code versions, all CS and DMs reachable, that no fs\_copy operations are running, that no Replication sessions are inactive, failed over, or suspended], verify nodes are reachable, abort the V1 sessions, deploy v2 and license, starts upgrade process, configures DART interconnects, starts Replication V2, cleans up RepV1 restart checkpoints

##### **Script can be run in Check Mode prior to maintenance window to assess replication health:**

**# /nas/sbin/upgrade\_repv1\_to\_repv2 -check**

The pre-upgrade check process has been started in background.

Script output is logged in "/nas/log/RepUpgradeV1V2.log"

**/nassbin/RepUpgradeV1V2.log**

**/nas/sbin/RepUpgradeV1V2\_Debug.log**

**/nas/sbin**

-rwxr-x--x 1 nasadmin nasadmin 101163 Jun 10 19:37 upgrade\_repv1\_to\_repv2

##### **STARTING THE V1 to V2 UPGRADE:**

**# /nas/sbin/upgrade\_repv1\_to\_repv2 -start**

**Note:** Use this command to start the V1 to V2 upgrade process

##### **RepV1 to V2 Logs:**

**# ls -la /nas/log/Rep\*** [or /nas/site directory]

-rw-r--r-- 1 root root 359 Jun 25 01:53 RepUpgradeV1V2\_Debug.log

-rw-r--r-- 1 root root 874 Jun 25 01:53 RepUpgradeV1V2.log

##### **Checking for Restart Checkpoints:**

**# nas\_fs -list |grep pfs**

pfs\_repl\_restart\_1

pfs\_repl\_restart\_2

##### **Creating Restart Checkpoints Manually:**

**Note:** Copy files to each replication file system so as to create a delta set. Create the first restart checkpoint, then copy additional files to each file system to create another delta set, then create the 2<sup>nd</sup> restart checkpoint.

**# fs\_ckpt pfs -name pfs\_repl\_restart\_1 -Create**

**# fs\_ckpt pfs -name pfs\_repl\_restart\_2 -Create**

**# fs\_ckpt virtual1 -name virtual1\_repl\_restart\_1 -Create**

operation in progress (not interruptible)...id = 89

name = virtual1

acl = 0

in\_use = True

```

type      = uxf5
worm     = off
volume   = v270
pool     = clarsas_r10
member_of = root_avm_fs_group_30
rw_servers= server_2
ro_servers=
rw_vdms =
ro_vdms =
auto_ext = hwm=90%,max_size=5000M,virtual_provision=yes
deduplication = Off
ckpts    = virtual1_ckpt1,virtual1_ckpt2,virtual1_repl_restart_1
ip_copies = virtual1_dest:nx4-2
stor_devs = SL7E1081700022-0013,SL7E1081700022-0010,SL7E1081700022-0011,SL7E1081700022-0012
disks    = d11,d8,d10,d9
disk=d11 stor_dev=SL7E1081700022-0013 addr=c16t113      server=server_2
disk=d11 stor_dev=SL7E1081700022-0013 addr=c0t113      server=server_2
disk=d8  stor_dev=SL7E1081700022-0010 addr=c0t110      server=server_2
disk=d8  stor_dev=SL7E1081700022-0010 addr=c16t110      server=server_2
disk=d10  stor_dev=SL7E1081700022-0011 addr=c16t111      server=server_2
disk=d10  stor_dev=SL7E1081700022-0011 addr=c0t111      server=server_2
disk=d9   stor_dev=SL7E1081700022-0012 addr=c0t112      server=server_2
disk=d9   stor_dev=SL7E1081700022-0012 addr=c16t112      server=server_2

id      = 311
name    = virtual1_repl_restart_1
acl     = 0
in_use  = True
type    = ckpt
worm    = off
volume  = vp328
pool    = clarsas_r10
member_of =
rw_servers=
ro_servers= server_2
rw_vdms =
ro_vdms =
checkpt_of= virtual1 Mon Jun 21 12:47:08 EDT 2010
deduplication = Off
used    = 4%
full(mark)= 90%
delta_number= 5
stor_devs = SL7E1081700022-0011,SL7E1081700022-0012
disks    = d10,d9
disk=d10 stor_dev=SL7E1081700022-0011 addr=c16t111      server=server_2
disk=d10 stor_dev=SL7E1081700022-0011 addr=c0t111      server=server_2
disk=d9  stor_dev=SL7E1081700022-0012 addr=c0t112      server=server_2
disk=d9  stor_dev=SL7E1081700022-0012 addr=c16t112      server=server_2

```

#### Refreshing Manually Created Restart Checkpoints:

- b) Generate new files on the production file system from a Host
- c) Run /nas/sbin/refresh\_ckpt to generate dataset transfer, or fs\_replicate –refresh <name> ?
- d) After several minutes, repeat steps a & b

#### Refreshing Internal Checkpoints on Destination:

# **fs\_ckpt id=78 -refresh -ALLOW\_REP\_INT\_CKPT\_OP**

#### V1 to V2 UPGRADE PROCESS:

1. Ensure that each pfs replication session has a time-out value greater > 0

\$ **fs\_replicate –info <name>**

# **fs\_replicate –modify <name> -option to=60**

**Note:** Record the current values of the TO settings

2. Each production file system must have two restart checkpoints—manually create them if necessary:

- a) Create file system activity on each pfs before creating the restart checkpoints
- b) Create first restart checkpoint on each pfs

**# fs\_ckpt pfs1 -name pfs1\_repl\_restart\_1 -Create**

- c) Create file system activity on each pfs again by copying files, etc.
- d) Create the second restart checkpoint on each pfs

**# fs\_ckpt pfs1 -name pfs1\_repl\_restart\_2 -Create**

### 3. Run the upgrade script in Check mode:

**# /nas/sbin/upgrade\_repv1\_to\_repv2 –check**

### 4. Run upgrade script:

**# /nas/sbin/upgrade\_repv1\_to\_repv2 –start**

**Note 1:** Script queries system to find all V1 sessions, then performs a verification check for overall state, database consistency, NAS version, connectivity to CS & DMs, that restart checkpoints are present, that no fs\_copy operations are running, that no Replication sessions are Inactive, and no suspended or failed-over sessions. Script aborts V1 sessions, turns off V1 license, configures authentication and interconnects, turns on V2 license, configures V2 sessions using restart checkpoints as common base, performing differential copy.

**Note 2:** Verify success of conversion in /nas/site/RepUpgradeV1V2.log

**Upgrade For V1 -> V2 Replication Completed on System.....**

### 5. Changing time-out policy after V2 upgrade:

**\$ nas\_replicate –modify <name> -manual\_refresh**

### 6. Verifying Replication sessions:

**# nas\_replicate -list**

| Name           | Type       | Local Mover | Interconnect               | Celerra | Status |
|----------------|------------|-------------|----------------------------|---------|--------|
| Rep4_FS82_FS30 | filesystem | server_2    | -->server_2_nx4-2_s+ nx4-2 |         | OK     |
| Rep3_FS81_FS29 | filesystem | server_2    | -->server_2_nx4-2_s+ nx4-2 |         | OK     |
| Rep2_FS77_FS27 | filesystem | server_2    | -->server_2_nx4-2_s+ nx4-2 |         | OK     |
| Rep1_FS76_FS26 | filesystem | server_2    | -->server_2_nx4-2_s+ nx4-2 |         | OK     |

**Note:** Replication sessions are renamed with the file system ID of Source first, then file system ID of Destination

### **TYPES OF REPLICATION SESSIONS:**

- Loopback [same DM used as source and destination]
- Local [from one DM to another in same cabinet]
- Remote [from one DM to another DM in a remote cabinet]
- One-to-many [source file system DM replicating to up to four destination DMs]
- Cascade [from source a to destination b, from b to c]

### **Restrictions:**

- Requires V2 licensing
  - Max replication or copy sessions per DM is 1024, but max active sessions are 256
  - Max initial copies in progress at one time is 16
  - one-to-many supports up to four destinations
  - Given Source file system or iSCSI LUN can support up to four replication Sessions
  - VDMs only support one replication session, no cascading, and no one-to-many configurations
- Note:** Cascading & one-to-many are supported with 5.6.42.5 + for VDMs
- Replication must be stopped before a VDM can be unloaded
  - Max of 96 snaps per file system, with 2048 iSCSI LUN snaps allowed

### **V1 to V2 Differences:**

- V2 requires use of DM-to-DM interconnects, which are not used with V1
- V2 supports one to many different destinations and also cascading sessions [A to B to C, etc]
- nas\_replicate replaces fs\_replicate
- nas\_copy used to make one-time copy of file system, checkpoint, or iSCSI LUN vs. fs\_copy
- V2 does not use restartable checkpoints
- Auto-extend SavVol feature and no playback service used with V2
- Policy updated using –max\_time\_out\_of\_sync option with nas\_replicate, from 1-1440 minutes [No more TO & HWM in V2]
- V1 supports 24TB while V2 supports 32TB
- If using FLR-E, both Source & Destination settings must be the same
- Replication does not support the Celerra Data Deduplication until NAS version 5.6.43

### **Internal Checkpoint Nomenclature:**

**V1:** root\_restart\_ckpt, root\_suspend\_ckpt, <pfs\_name>\_repl\_restart\_1, root\_new\_ckpt, root\_restore,ckpt

**V2:** <name>\_root\_rep\_ckpt

### **Benefits of V2 vs. V1:**

- Easier to setup—can do everything from Celerra Manager wizard & GUI
- Destination file system created automatically for VDMs and regular replication sessions

--One-to-many & Cascading

--Better performance

--No restart checkpoints needed to restart replication sessions

--Beginning with 5.6.42, support for VDM cascading and VDM one-to-many replication sessions

**Note:** So now can Replicate one-to-many & cascading for File Systems, iSCSI LUNs, & VDM objects

--Beginning with 5.6.43, support for FLR & Data Deduplicated file systems (F-RDE)

## **SAVVOL INFORMATION NAS 5.6:**

→Only a single SavVol is used by Snapsure RO checkpoints, Writeable checkpoints, and Replication sessions, per production file system, and for Replication, a SavVol is also used on the Destination side

→If file system size is >than 10GB, then the default SavVol size is 10GB

→If file system size is <than 10GB, but >than 64MB, then default SavVol size will equal the production file system size

→If file system size is <than 64MB, then SavVol size will be 64MB

→Range of SavVol size can be 64MB – 16TB

→By default, the SavVol is extended by 10GB when the default HWM of 90% capacity, is reached

→If not enough space to extend by 10GB, will use remaining 10% HWM space, and if more space is needed, will then begin inactivating & deleting oldest Checkpoints first, and finally, becomes inactive if it cannot auto-extend after using all reclaimed ckpts

→A best practice recommendation is to manually create the SavVol at 10% the size of the PFS, meaning a 1TB fs would have a minimum SavVol size of 100GB [if left to the system defaults, the size would only be created as 10GB]

→Use fs\_ckpt –modify to change auto-extension HWM or total SavVol size allowed

**%full=90** [HWM range is 10-99; if set to 0, disables the auto-extension of SavVol feature]

**maxsavsize=T | G | M** →sets limit on max size in MB, GB, TB, of a given SavVol

→For Replication Sessions, autofreeze=yes | no [default] if SavVol becomes full and cannot auto-extend

→For Replication Sessions, autoro=yes | no [default] if SavVol becomes full and cannot auto-extend

## **SNAPSURE CONFIGURATION LINE:**

### **/nas/sys/nas\_param**

#### **ckpt:10:20:20**

10 = Control Station polling interval in seconds

200 = Maximum rate in MB/sec at which a file system is written to

20 = % of total system space allowed for the creation and extension of all SavVols—this value can be changed between 20-99

**Note:** The only value on this checkpoint line that should ever be changed is the % Total system space

→If the 20 percent value is reached, the following message appears in the command output, and also in the /nas/log/cmd\_log.err and /nas/log/nas\_log.al.err files:

Error 2238: Disk space: quota exceeded

Feb 1 15:20:07 2008 NASDB:6:101 CKPT volume allocation quota exceeded

### **Replication Destination SavVol Issue:**

AR113894 finds that Destination SavVol creation size will be the size of whatever data is in the Source PFS. Fixed in 5.6.37.

### **USING NAS FS QUERIES TO DISPLAY SAVVOLS IN USE ON THE CELERRA:**

# nas\_fs -query>Type=uxfs -format:"%q" -fields:Checkpoints -query:\* -format:"ID=%s, Name=%s, VolumeID=%s,

VolumeName=%s\n" -fields:ID,name,VolumeID,VolumeName

ID=26, Name=root\_rep\_ckpt\_23\_1282\_1, VolumeID=106, VolumeName=vp106

ID=27, Name=root\_rep\_ckpt\_23\_1282\_2, VolumeID=106, VolumeName=vp106

ID=31, Name=root\_rep\_ckpt\_28\_1289\_1, VolumeID=113, VolumeName=vp113

ID=32, Name=root\_rep\_ckpt\_28\_1289\_2, VolumeID=113, VolumeName=vp113

ID=33, Name=root\_rep\_ckpt\_28\_1345\_1, VolumeID=113, VolumeName=vp113

ID=34, Name=root\_rep\_ckpt\_28\_1345\_2, VolumeID=113, VolumeName=vp113

ID=37, Name=root\_rep\_ckpt\_35\_1397\_1, VolumeID=124, VolumeName=vp124

ID=38, Name=root\_rep\_ckpt\_35\_1397\_2, VolumeID=124, VolumeName=vp124

### **ENABLING V1 REPLICATION ON 5.6 FOR LEGACY ENVIRONMENTS:**

1. Check current status

# **/nas/sbin/deploy\_replication -status**

1 = V1 ready

2 = V2 ready

3 = V1 ready & licensed

4 = V2 ready & licensed

or

# **/nas/sbin/deploy\_replication -status -verbose**

'replicatorV2' license is enabled.

2. Enable V1

# **/nas/sbin/deploy\_replication -v1**

3. Check Replicator License in Celerra Manager and apply to enable

### **ENABLING V2 REPLICATION:**

1. If status 3 is returned, delete v1 license first, then enable v2

```
# nas_license --delete replicator
```

```
# /nas/sbin/deploy_replication -v2
```

### **CASCADE REPLICATION SESSIONS:**

--Defined as replicating from A-to-B, and B-to-C [C-to-D not allowed]

--Only file system and iSCSI LUN objects can use cascading sessions

--Time\_out\_of\_Sync value should be set to the same for each leg

### **ONE-to-MANY REPLICATION:**

--Defined as up to four Replication sessions originating from a Source file system to four different target file systems

--Only file system and iSCSI LUN objects can use one-to-many

### **DOWNGRADING FROM REPV2 to REPV1:**

```
1. # mv /nas/site/nas_license /nas/site/nas_license.old
```

```
2. # nas_license -init
```

```
3. # /nas/sbin/deploy_replication -v1
```

```
4. # nas_license --create replicator
```

### **Switching to V2 or V1 Replication:**

```
/nas/sbin/deploy_replication -v2 | -v1
```

```
nas_license --delete replicator | replicatorV2
```

```
nas_license --create replicatorV2 | replicator
```

→Supports (2) types of file system replication:

a.) one-time copy

b.) ongoing replication

→Supports File System, VDM, and iSCSI objects

**Note:** Replicate VDM first, then production file systems. Must have iSCSI target & LUN available at destination before replicating

→Supports Local, Loopback, and Remote directional Replication sessions

→IPREPV2 Supports (1)-to-many replications; rearchitecture; file system cascading (a-to-b-to-c, etc.); up to 1024 Sessions per DM

**Note:** This feature will be upgraded after landing on NAS 5.6, using a /nas/sbin/upgrade\_repv1\_to\_repv2 script, which logs to

```
/nas/log/RepUpgradeV1V2.log. V1 is not going to be supported past NAS 5.6
```

### **TROUBLESHOOTING IPREPV2:**

```
# .server_config server_2 -v "replicate2"
```

```
replicate2 {
```

```
    checkdb
```

```
    displayrepsnapsignatures
```

```
        { fsid=<fsId> | vdmid=<vdmId> | lunid=<lunId> targetname=<targetName> }
```

```
    | displayrepsnaps
```

```
        { name=<aliasName> | id=<repSessionId> }
```

```
    | displaynamedb
```

```
        { [id=<repSessionId>] }
```

```
    | displayrepconfigsignature
```

```
        [ name=<aliasName> | id=<repSessionId> | vsid=<vsidStr> ]
```

```
    | checknamedb
```

```
        { mode=<critical | stale> }
```

```
    | skiprecovery
```

```
        { id=<repSessionId> }
```

```
    | recoverrepession
```

```
        { id=<repSessionId> }
```

```
    | resetrepconfigsignature
```

```
        { vsid=<vsId> vstype={filevolume} [alias=<repAliasName>]
```

```
# .server_config server_2 -v "vvs list"
```

```
# .server_config server_2 -v "replicate2 displayrepconfigsignature name=REPL_VMFS03"
```

```
1273048843: DPSVC: 7: VersionSetContext::findReplicaContextInt() failedrepName: alias:REPL_VMFS03
```

```
status:DpRequest_Context_NotFound
```

```
1273048843: DPSVC: 7: last message repeated 3 times
```

```
1273048843: DPSVC: 7: ReplicaContext::repReqInProgressRef() c:1 aliasName:REPL_VMFS03
```

```
1273048843: DPSVC: 7: findReplicaContext():alias=REPL_VMFS03 repTypeIN:0 repType:2
```

```
1273048843: DPSVC: 7: VersionSetContext::verReqInProgressRef() c:1 vsidStr:42:fs42_T1_LUN103_APM00082800574_0000
```

```
1273048843: DPSVC: 7: displayRepConfigSignature:: repType 2
```

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!

1273048843: UFS: 7: inc ino blk cache count: nInoAllocs 1: inoBlk c76e4004

1273048843: DPSVC: 7: DpVersion::getRepConfigSignature() expected crc:3003544296 and actual crc: 3003544296 of SourceFS: 0 and DestinationFS: 0

1273048843: DPSVC: 7: ReplicaConfigSignature crc intact.

-----REP\_CONFIG\_SIGNATURE-----

repAliasName : REPL\_VMF03

repSessionName : fs26\_T1\_LUN3\_APM00094702040\_0000\_fs42\_T1\_LUN103\_APM00082800574\_0000

repType : 2 (destination)

appLabel :

repSessionState : 4 (configured)

repStoppedReason : 0 (DpRequest\_CmdNotDefined)

For FS Replication

srcFsId : 0 dstFsId : 0

For iSCSI Replication

srcLunNum : 3 srcTargetName : iqn.1992-05.com.emc:apm000947020400000-1

dstLunNum : 103 dstTargetName : iqn.1992-05.com.emc:apm000828005740000-1

For VDM Replication

srcVdmId : 0 srcVdmName :

dstVdmId : 0 dstVdmName :

srcVsidStr : 26:fs26\_T1\_LUN3\_APM00094702040\_0000

srcVsType : 1 (VST\_FILE\_VERSION\_SET)

srcVersionName1 : fs26\_T1\_LUN3\_APM00094702040\_0000.ckpt000\_356569888734220

srcVersionName2 : fs26\_T1\_LUN3\_APM00094702040\_0000.ckpt000\_356569888811010

dstVsidStr : 42:fs42\_T1\_LUN103\_APM00082800574\_0000

dstVsType : 1 (VST\_FILE\_VERSION\_SET)

dstVersionName1 : fs42\_T1\_LUN103\_APM00082800574\_0000.ckpt000\_5777226191182803

dstVersionName2 : fs42\_T1\_LUN103\_APM00082800574\_0000.ckpt000\_5777226191255876

srcSnapFsId1 : 0 srcSnapFsId2 : 0

dstSnapFsId1 : 0 dstSnapFsId2 : 0

interConnectId : 20003 dstInterConnectId : 20003

srcIp : 10.1.31.250 dstIp : 10.2.14.151

slaValue : 30

crcVal : 3003544296

Old Replication Configuration Signature

repAliasName : REPL\_VMF03

appLabel :

repType : 2 (destination)

repSessionState : 5 (stopped)

interConnectId : 20003 dstInterConnectId : 20003

srcIp : 10.1.31.250 dstIp : 10.2.14.151

slaValue : 30

1273048843: DPSVC: 7: ReplicaContext::repReqInProgressDeref() c:0 aliasName:REPL\_VMF03

1273048843: DPSVC:10: DpContext::removeVersionSetContext() failed vsidStr:42:fs42\_T1\_LUN103\_APM00082800574\_0000

ref:1 status:DpRequest\_Context\_Busy

1273048843: DPSVC: 7: VersionSetContext::verReqInProgressDeref() c:0 vsidStr:42:fs42\_T1\_LUN103\_APM00082800574\_0000

1273048843: ADMIN: 6: Command succeeded: replicate2 displayrepconfigsignature name=REPL\_VMF03

## **DELETING ORPHANED REPLICATION CHECKPOINT:**

**\$ /nas/sbin/rootnas\_fs -d root\_rep\_ckpt\_25\_34130\_2 -o umount=yes -ALLOW REP\_INT\_CKPT\_OP**

## **5.6 DART INTERCONNECTS (DIC):**

→NAS 5.5.19-4 rolled out Phase I of the DART Interconnect (DIC) system for local or loopback iSCSI Replication. The project goal was to allow the ability of DART Servers to intercommunicate using HTTP in a Client/Server relationship. Part of the new DIC mechanism creates system-generated passwd entries with names based on storage array back ends, and also consists of a constant UID/GID value of 9000. These new passwd entries (alias records) are used by DIC to authenticate requests between DART Servers using XML packets over HTTP. With NAS 5.6 and RepV2, the DIC interconnect is replaced with an interconnect framework, used both for authentication and throttling information between DART servers. Do not delete, or otherwise modify, these entries.

## **EXAMPLE OF SYSTEM-GENERATED PASSWD ENTRY ON DATA MOVER:**

# cat /etc/passwd

APM000716005140000\_APM000716005140000:ef0wkKxT/c0z.:9000:9000:aSGsZJrseqsu0lsa2coly5tHpX::ndmp\_md5

**VERIFYING & VALIDATING INTERCONNECTS:**

```
# nas_cel -interconnect -list
id      name      source_server destination_system destination_server
20001  loopback   server_2     sludge4       server_2
30001  loopback   server_3     sludge4       server_3
```

```
# nas_cel -interconnect -validate id=20001
```

validating...ok

**QUERYING HTTP SERVER ON DATA MOVER FOR DIC INFO:**

```
# server_http server_2 -info DIC
```

server\_2 : done

DIC FACILITY CONFIGURATION

```
Service name      : Dart Interconnect server
Comment          : DIC facility for Data Movers exchange
Active           : True
Port             : 5081
Threads          : 10
Max requests     : 300
Timeout          : 60 seconds
```

ACCESS CONTROL

```
Allowed IPs       : any
Authentication    : digest ,Realm : DIC_Authorization
Allowed user      : APM000716005140000_APM000716005140000
```

SSL CONFIGURATION

```
Mode              : OFF
Persona          : default
Protocol         : default
Cipher            : default
```

```
# server_http server_2 -info DIC_L
```

server\_2 : done

DIC\_L FACILITY CONFIGURATION

```
Service name      : Celerra Interconnect server (L)
Comment          : CIC facility for long commands support
Active           : True
Port             : 5084
Threads          : 10
Max requests     : 300
Timeout          : 60 seconds
```

ACCESS CONTROL

```
Allowed IPs       : any
Authentication    : digest ,Realm : DIC_Authorization
Allowed user      : APM000716005140000_APM000716005140000
```

SSL CONFIGURATION

```
Mode              : OFF
Persona          : default
Protocol         : default
Cipher            : default
```

```
# server_http server_2 -info dic_s
```

server\_2 : done

DIC\_S FACILITY CONFIGURATION

```
Service name      : Celerra Interconnect server (S)
Comment          : CIC facility for short commands support
Active           : True
Port             : 5083
Threads          : 10
Max requests     : 300
Timeout          : 60 seconds
```

ACCESS CONTROL

```
Allowed IPs       : any
Authentication    : digest ,Realm : DIC_Authorization
Allowed user      : APM000716005140000_APM000716005140000
```

SSL CONFIGURATION

Mode : OFF  
Persona : default  
Protocol : default  
Cipher : default

# server\_http server\_2 -info dic\_d

server\_2 : done

DIC\_D FACILITY CONFIGURATION

Service name : Celerra Interconnect server (D)  
Comment : CIC facility for dart-dart commands support  
Active : True  
Port : 5085  
Threads : 20  
Max requests : 300  
Timeout : 60 seconds

ACCESS CONTROL

Allowed IPs : any  
Authentication : digest ,Realm : DIC\_Authorization  
Allowed user : APM000716005140000\_APM000716005140000

SSL CONFIGURATION

Mode : REQUIRED  
Persona : anonymous  
Protocol : default  
Cipher : default

**REBUILDING THE DIC PASSWD ON DATA MOVERS:**

# nas\_cel -update id=0

operation in progress (not interruptible)...

id = 0  
name = sludge4  
owner = 0  
device =  
channel =  
net\_path = 192.1.4.218  
celerra\_id = APM000716005140000

**Note:** This command also recreated and populated server\_3 with a default passwd entry to support the DIC feature.

**REPV1 to REPV2 CONVERSION ISSUE RELATED TO DIC ENTRIES:**

If a standby server was converted to a primary while at NAS 5.5, and replication sessions were in place, and the system was subsequently upgraded to NAS 5.6, during the REPV1 to REPV2 conversion a failure occurs because the standby (now primary) server never had all the DART interconnect passwd entries required for all the remote systems. See ARs 120114 & 108256. Fix will be to run nas\_cel -update after server\_setup to ensure that DIC passwd entries are run for all systems seen by nas\_cel.

**DART INTERCONNECT TROUBLESHOOTING:**

# .server\_config server\_2 -v "dic service"

1247670598: DIC: 6: 5081 DIC text/dic  
1247670598: DIC: 6: DicXmlAsyncMsgService [ on 1 cpost] URL=/dic/DicXmlAsyncMsgService  
1247670598: DIC: 6:  
1247670598: DIC: 6: 5083 DIC\_S text/dic  
1247670598: DIC: 6: DicXmlAsyncMsgService [ on 1 cpost] URL=/dic\_s/DicXmlAsyncMsgService  
1247670598: DIC: 6:  
1247670598: DIC: 6: 5084 DIC\_L text/dic  
1247670598: DIC: 6: DicXmlAsyncMsgService [ on 1 cpost] URL=/dic\_l/DicXmlAsyncMsgService  
1247670598: DIC: 6:  
1247670598: DIC: 6: 5085 DIC\_D text/dic  
1247670598: DIC: 6: SSL DIC\_D[1 0xe72fe904]  
1247670598: DIC: 6: repDicSvcV2 [ on 1 cpost] URL=/dic\_d/repDicSvcV2  
1247670598: DIC: 6: DpTunnelDicServiceV2 [ on 1 post] URL=/dic\_d/DpTunnelDicServiceV2  
1247670598: DIC: 6: DpTunnelDicMsgServiceV2 [ on 1 post] URL=/dic\_d/DpTunnelDicMsgServiceV2  
1247670598: DIC: 6: DpTaskDicServiceV2 [ on 1 thread] URL=/dic\_d/DpTaskDicServiceV2  
1247670598: DIC: 6: DicXmlSyncMsgService [ on 1 post] URL=/dic\_d/DicXmlSyncMsgService  
1247670598: DIC: 6: DicXmlAsyncMsgService [ on 1 cpost] URL=/dic\_d/DicXmlAsyncMsgService  
1247670598: DIC: 6:

### **REPLICATION INTERCONNECTS—DART & CS:**

- DART Interconnects are required for Local & Remote Replication sessions between Data Mover pairs
- Only single interconnect built for each Data Mover pair
- Celerra Mgr wizard will create bi-directional interconnects, otherwise, the User will need to configure from Replication tab on each side of the Replication environment
- For Replication, Data Mover interconnects maintain the structure for source and destination interfaces, bandwidth information, and interface usage counts

### **UPDATING CELERRA INTERCONNECTS WITH NEW CELERRA ID (CMU):**

- 1) Edit server “start” files with correct Celerra ID or CMU (nas\_cel –list):

```
# vi /nas/server/slot_2/start
```

```
setid celerra=00019010488028CF
```

- 2) Comment out the dpinit service for each server in the “eof” file, then reboot

```
# vi /nas/server/slot_2/eof
```

```
# dpinit
```

```
# server_cpu server_2 –reboot now
```

- 3) Delete the old interconnect database

```
# server_dbms server_2 –db –delete icon_db
```

- 4) Remove comment from “eof” dpinit line and boot.cfg files, then restart dpinit service

```
# .server_config server_2 “dpinit”
```

- 5) Update the interconnects using the following:

```
# nas_cel –update id=0
```

### **5.6 NAS CEL INTERCONNECT COMMAND:**

**# nas\_cel -interconnect -create** [creates a single interconnect between two Data Movers in same or different cabinets, authentication, bandwidth throttling scheduler, & IP addresses on Source & Destination]

**-destination\_interfaces** [name of interfaces]

**-bandwidth** [can specify periods by time and bandwidth: “MoTuWeThFr07:00-18:00/2000,/8000” translates into 2000KB/sec limit from 7-6:00p.m. Mon-Friday, and 8000KB/sec at all other times]

**-modify** [can modify source & destination addresses, interconnect names, throttling, crc]

**-pause** [sets throttle value to 0, suspending transfer]

**-resume** [set throttle back to value prior to suspending]

**-delete** [deletes the interconnect configuration, but only if not being used by a Replication session]

**-info** [displays interconnect information]

**-list** [displays interconnects on entire Celerra]

**-validate** [verifies that Data Mover pairs are authenticated and validates all source and destination interface pairs]

#### **Types of Interconnects:**

- a.) Control Station interconnects, created automatically by nas\_cel
- b.) Loopback Interconnect, created automatically for each DM
- c.) Local Replication Interconnect, created between pair of DMs for source and peer sides
- d.) Remote Replication Interconnect, created between local and remote DMs

**Note:** An interconnect provides all available bandwidth for replication sessions by default

#### **Creating Interconnect Example:**

```
# /nas/sbin/nas_cel –interconnect –create ns0_dm2 –source_server server_2 –destination_system ns1_dm2 –destination_server server_2 –source_interfaces ip=192.168.10.50, 192.168.10.55 –destination_interfaces ip=192.168.10.25, 192.168.10.30 –bandwidth MoTuWe 07:00-18:00/2000, /8000 –crc yes
```

### **5.6 NAS CEL COMMAND:**

→Used to register Source and Destination Control Stations with each other, run as root, requires passphrase, run from each Control Station to register the peer CS

```
# nas_cel –create ns1 –ip 10.250.60.10 –passphrase nasadmin
```

```
# nas_cel –list [verify registrations]
```

### **5.6 NAS REPLICATE COMMAND:**

→Used to create & manage replication sessions for file systems, VDMs, and iSCSI LUNs

→DP Manager on Data Mover is used to keep source and destination objects synchronized, using HTTP over TCP/IP between Data Movers. DP Mgr also controls the replica engine, which uses RCP over TCP/IP to transmit data between Source & Target Data Movers.

→Must specify the iSCSI LUN destination for iSCSI

```
# nas_replicate –list [list of replication sessions and status]
```

```
-info [detailed information about specific replication sessions]
```

-**create** [used to create a replication session]  
-**source -fs** [source fs to replicate]  
-**destination -fs** [replication target]  
-**source\_interface** [IP address or interface name for source]  
-**destination\_interface** [IP address or interface name for destination]  
-**max\_time\_out\_of\_sync** [max time before ckpt refreshed & new delta set sent to destination; range 1-1440 minutes; default 20 min.]  
-**overwrite\_destination** [overwrite destination with changes from source]  
-**background** [run as background process and return to prompt]  
-**source -vdm** [source vdm] -**destination -vdm** [destination vdm]  
-**source -lun -target** [source iSCSI LUN & Target IQN for replication] -**destination -lun -target** [destination iSCSI LUN & Target IQN to overwrite]  
-**start** [Used to restart a replication session using differential or full copy, from source side, usually after session was stopped. Can also change interconnect and interfaces, as well as policies, during the restart]  
-**start -reverse** [restarting replication session and reversing replication direction—issued from destination side, usually after a failover or switchover has been run. VDMs need to be mounted RW & Loaded, file systems need to be RW or RO, and iSCSI LUN needs to be exported to Host as iSCSI Target. If commonbase exists, a differential copy is performed. If no commonbase exists, a full copy must be done.]  
-**reverse** [run from Source side and reverses direction of Replication without data loss]  
-**modify** [used to modify source & destination Data Mover addresses, Time Out of Sync, Name of Session, but can be run only from Source side when replication is in configured state]  
-**stop** [stops replication session] -**stop -delete** [stops replication session and deletes checkpoints & configuration]  
-**refresh** [forces update from source to target, run from source side]  
-**failover** [Used for DR failover, issued on Destination side, stops Replication session, makes destination file system RW, terminates any deltaset, makes Source RO if reachable, involves some data loss as delta sets are not transferred]  
-**switchover** [Stops replication gracefully after syncing source & destination so that no data is lost, used for testing, issued from Source side only, makes destination RW, source RO]

**Note:** V2 replication automatically creates destination file system for file system and VDM replication, and creates two checkpoints for source and destination. After first full copy, the destination ckpt1 is refreshed (ckpt2) to become the common-base for source ckpt1. Then, the source ckpt2 is refreshed to create delta between original source ckpt1 and latest copy (differential copy). Once this transfer is completed, the destination ckpt2 is used as the common base to source ckpt2

## **5.6 NAS COPY COMMAND:**

# **nas\_copy -source -fs | -source -chkt | -from\_base | -refresh | -overwrite\_destination | -full\_copy**

## **5.6 NAS TASK COMMAND:**

# **nas\_task -list | -remote\_system | -info | -abort <id> -mover -remote\_system | -delete** [used to track progress of replication commands locally and remotely, with ability to abort or delete tasks]

# **nas\_task -info xxxx -remote\_system <cs\_remote>**

**Note:** nas\_task -delete only removes task from list but does not kill the actual task itself?

## **DART TASK LIST:**

\$ .server\_config server\_2 -v “cmdSvcDisplay”

## **iSCSI LUN REPLICATION REQUIREMENTS:**

- Destination LUN needs to exist and be RO
- iSCSI LUNs are supported for Replication using one-to-many, and cascading, etc.
- Target IQNs are required for -create
- During normal iSCSI Replication, destination iSCSI LUNs are not accessible, and only becomes accessible after failover, switchover, or stop events
- During normal iSCSI Replication operation, destination LUN can be snapped with CLI, promoted, then mounted and exported for Hosts
- iSCSI snaps are crash consistent, not application consistent
- iSCSI snaps db is not stored on Control Station, only on Data Mover

## **CREATING iSCSI REPLICATION SESSION:**

# **nas\_replicate -create -iscsi1 -source -lun 0 -target iqn.1992-05.com.emc:apm0074003149000-2 -destination -lun 0 -target iqn.1992-05.com.emc:amp00071500149000-1 -interconnect ns0\_dm2**

## **SETTING UP V2 REPLICATION ON CELERRA:**

### **USING CELERRA MANAGER:**

1. Enable Replicator license: Celerra Home>Licenses>Replicator V2
2. Select Wizards>Replication>New Replication>Wizard Steps  
Select a Replication Type > File System | VDM | iSCSI LUN  
Select Destination Celerra  
Select Data Mover Interconnect  
Select Replication Sessions's Interface

Select Source

Select Destination

Update Policy

Overview/Results

3. Or use Replications tab

**FROM CLI:**

1. Setup Source &amp; Destination communication using nas\_cel -create &amp; verify using nas\_cel -list

# **nas\_cel -create cs\_source -ip 192.168.10.10 -passphrase nasadmin**# **nas\_cel -create cs\_dest -ip 192.168.10.5 -passphrase nasadmin**# **nas\_cel -list**

2. Setup DM-to-DM Source &amp; Peer Interconnects &amp; verify using nas\_cel -interconnect -validate id=&lt;interconnect\_ID&gt;

# **/nas/sbin/nas\_cel -interconnect -create cs\_source -source\_server server\_2 -destination\_system cs\_dest -destination\_server server\_3 -source\_interfaces ip=192.168.10.25, 192.168.10.30 -destination\_interfaces ip=192.168.10.50, 192.168.10.55 -bandwidth MoTuWe07:00-14:00/0800 -crc yes**# **/nas/sbin/nas\_cel -interconnect -create cs\_dest -source\_server server\_3 -destination\_system cs\_source -destination\_server server\_2 -source\_interfaces ip=192.168.10.50, 192.168.10.55 -destination\_interfaces ip=192.168.10.25, 192.168.10.30 -bandwidth MoTuWe07:00-14:00/0800 -crc yes**# **nas\_cel -interconnect -list | -modify | -pause | -resume**# **nas\_cel -interconnect -list**

| id    | name     | source_server | destination_system | destination_server |
|-------|----------|---------------|--------------------|--------------------|
| 20001 | loopback | server_2      | hammer2            | server_2           |
| 30001 | loopback | server_3      | hammer2            | server_3           |

# **nas\_cel -interconnect -validate id=20001**

validating...ok

3. Create Replication session and verify using nas\_replicate -list, -info id=x

# **nas\_replicate -create rep1 -source -fs src\_fs1 -destination -pool admin\_pool -interconnect id=200003 -max\_time\_out\_of\_sync 15****Note:** A replication session creates (4) internal checkpoints, 2 for Source, 2 for Destination, baseline and differential images

--Basically, create the Replication session from Source to Destination

--A full copy of Ckpt1 on the source fs is copied to the Destination

--After Ckpt1 is copied to the Destination, the Destination Ckpt1 is refreshed and then becomes the common base between Destination and Source file systems

--On the Source, once the common base has been established, a differential checkpoint is taken on the source fs as Ckpt2 and includes the delta of any changes--this is called the differential copy which is then transferred to the Destination.

--Once the differential has been transferred to the Destination, the Destination Ckpt2 is refreshed to become the new common base with Source Ckpt2.

--Then, in accordance with whatever update policies are in place, the cycle continues with differentials copied over and new common bases being established each time

**REFRESHING DESTINATION SESSION:**# **nas\_replicate -refresh fs1\_x -background**# **nas\_task -info <id>****DELETE REPLICATION SESSION:**# **nas\_replicate -delete <name or id=x> -mode source | destination | both****STOP REPLICATION SESSION:**# **nas\_replicate -stop fs1 -mode source | destination | both****START REPLICATION SESSION:**# **nas\_replicate -start <fs\_name or session\_id> | -interconnect <name or interconnect\_id> | -max\_time\_out\_of\_sync 10****Note:** Session can be restarted using full or differential copy, and interconnect, interfaces, and policies can be modified**STARTING A FAILED OR SWITCHED OVER REPLICATION SESSION IN REVERSE DIRECTION:**# **nas\_replicate -start <fs\_name or session\_id> -reverse -interconnect id=30000 -overwrite\_destination****REVERSING A REPLICATION SESSION:**# **nas\_replicate -reverse fs1****Note:** Run on Source side, reverses direction of replication, places source RO**MODIFY REPLICATION SESSION:**# **nas\_replicate -modify****Note:** Use modify to change name of session, source & destination interfaces, max time out of sync, run from Source**3 TYPES OF REPLICATION FAILOVER:**# **nas\_replicate -failover**

**Note:** Used in disaster scenario, issued from destination Control Station. Stops any data transfers, changes destination to RW, and performs failover. Considered a DL event since this action will incur data loss. Direction of replication not changed.

#### # nas\_replicate –switchover

**Note:** Graceful failover run from Source side, no data loss, useful when testing failover. Synchronizes Source & Destination, stops replication session, mounts Source RO, mounts Destination RW, and does not start replication. Direction of replication not changed.

#### # nas\_replicate –reverse

**Note:** Used to change direction of replication, issued from Source side, synchronizes Source & Destination, stops replication, mounts source RO, mounts destination RW, starts replication in reverse direction from differential copy.

#### TROUBLESHOOTING/LOGS:

##### SET FOLLOWING DEBUG PRIOR TO TESTING IP REPLICATION:

```
$ .server_config server_x "logsys set severity CMD=LOG_DEBUG"
$ .server_config server_2 -v "cmdSvcDisplay" [task list on Data Mover]
$ .server_config server_2 -v "cmdQueueDbTest display" → looks for any queued tasks
$ .server_config server_2 -v "testDB display" → displays dbms records related to running tasks
1249327570: CMD: 6: Displaying DBMS Records
```

```
-----  
KEY=APM000832011840000_0_2612:4294969908_Initiator  
-----abridged-----
```

```
-----  
KEY=APM000832011840000_0_2612:8589937204_Initiator
```

**Note:** Here are two DBMS records for the Task\_Id referenced above

```
$ .server_config server_2 -v "testDB delete key=APM000752005020000_0_29619:4294996915_Server"
$ .server_config server_2 -v "cmdQueueDbTest delete key=524_APM00075200502_0000_462_APM00075200501_0000"
/nas/log/nas_log.al.mgmtd [Logs communications between DART & CS]
/nas/log/sys_log & server_log
/nas/log/webui/cli.log [GUI & CLI commands recorded here]
/nas/log/webui/apl_tm.log [legacy commands if V1 is used]
```

#### CLEANING UP HUNG REPLICATION TASKS:

**Note:** Generally, hung or defunct tasks are seen during Upgrade checks. See emc204889 & emc225820 for the proper cleanup methods

1. Identify running tasks

#### # nas\_task -list

| ID   | Task State | Originator    | Start Time                   | Description             | Schedule | Remote System |
|------|------------|---------------|------------------------------|-------------------------|----------|---------------|
| 2612 | Running    | root@168.159+ | Wed Jun 10 13:00:26 EDT 2009 | Create Replication fs1. |          |               |

2. Try to abort the tasks (Source & Target Control Stations)

#### # nas\_task –abort 2612 –mover server\_2 –remote\_system <remote\_cs\_name>

3. Continue with the following actions if the tasks cannot be aborted:

- a) Delete the following DBMS records:

```
$ .server_config server_2 -v "testDB display" # Lists the testDB record.
$ .server_config server_2 -v "testDB delete key={key from display command}" # Deletes the testDB record.
$ .server_config server_2 -v "cmdQueueDbTest display" # Lists the cmdQueueDbTest record.
$ .server_config server_2 -v "cmdQueueDbTest delete key={key from display command}" # Deletes the cmdQueueDbTest record.
```

- b) Su to root user and create console tool link:

```
# cd /nas/tools
```

```
# ln -s console _console
```

- c) Run the console tool, identify and destroy tasks:

```
./_console
```

At the prompt Command {Enter "exit" to quit } **>>taskmgr.dump** # Lists the tasks.  
2250036

Command {Enter "exit" to quit} **>> taskmgr.destroy 2250036** # Deletes the task.

- d) Reboot the Data Mover(s).

- e) Continue with upgrade.

#### Some Output Examples:

#### # .server\_config server\_2 -v "cmdSvcDisplay"

1249327540: CMD: 6: CMD Context: List of all Cmd Contexts :

1249327540: CMD: 6: Task\_Id RefCount

1249327540: CMD: 6: APM000832011840000\_0\_2612:4294969908

**Note:** Here it outputs a task list, which in this case is an orphaned task on the Control Station for Replication

## # .server\_config server\_2 -v “testDB display”

1249327570: CMD: 6: Displaying DBMS Records

-----  
KEY=APM000832011840000\_0\_2612:4294969908\_Initiator

-----abridged-----

KEY=APM000832011840000\_0\_2612:8589937204\_Initiator

**Note:** Here are two DBMS records for the Task\_Id referenced above

# .server\_config server\_2 -v "testDB delete key=APM000832011840000\_0\_2612:8589937204\_Initiator"

# .server\_config server\_2 -v "testDB delete key=APM000832011840000\_0\_2612:4294969908\_Initiator"

**Note:** Above two commands deleted the dbms records seen with testDB display

## # nas\_task -delete 2612

OK

## **5.6 NDMP VBB SUPPORT:**

→NDMP 3-Way VBB support (correction to handle unexpected EOTape during restores)

## **5.6 CS SECURITY ENHANCEMENTS:**

→Various security issues; PKI (X.509 certificates from CS CLI or GUI to support Data Movers—Certificates for DM can be signed by CS or other External Host), no dynamic ports, CS Security auditing; Linux kernel upgrade from RH 7.2 to RH Enterprise Linux 4—RHEL 4; External directory support for CS authentication following EMC CAMS CSO (Common Sign-On) initiative; Support for SSL or IPSEC for Control Station commands to allow for encryption over IP; passwd management on CS; CS Security auditing

**Note:** CAMS CSO provides authentication, Role & User management, and cryptography for Celerra Manager in communications with external directory servers. The RSA Common Security Toolkit (CST) is replacing CSO in a Cognac maintenance version.

## **5.6 CELERRA PUBLIC KEY INFRASTRUCTURE (PKI—X.509 Certificates):**

→Framework for distributing & maintaining private & public keys, and digital certificates, for protecting data, authenticating users, encrypting or using digital signatures

→PKI provides authentication services for SSL [DHSM, RepV2, OpenLDAP uses SSL]

→PKI is not supported with CIFS or NFS

→Managed from Celerra Manager>Security tab>Public Key Certificates or via CLI using server\_certificate or nas\_ca\_certificate

→Public Key Certificates folder is used for certificate and public key management

→CA certificates used to verify signatures on public key certificates received

\$ server\_certificate server\_2 –ca\_certificate –import –filename /home/nasadmin/cacert.pem [Import CA]

\$ server\_certificate server\_2 –ca\_certificate –list | -info <id> | -delete <id>

→Data Movers use a ‘Persona’ as a container to have keys and certificates managed for them by the CS. Persona is used to provide private key and certificate when acting as a Server to Client connections or as a Client to another server

→Each ‘Persona’ has up to 2 sets of keys and certificates (current & next), with key sizes 2048 or 4096 bits in length

→Only a single ‘Persona’ allowed per Data Mover

\$ server\_certificate server\_2 –persona –list | -info default

\$ server\_certificate server\_2 –persona –generate default –key\_size 2048 –cn “celerra1.pvt.dns” –filename

/home/nasadmin/newcertrequest.pem [Generate keys and certificate request external CA]

\$ server\_certificate server\_2 –persona –generate default –cn “celerra1.pvt.dns” –cs\_sign\_duration 24 [CS generated key & cert]

\$ server\_certificate server\_2 –persona –import default –filename /home/nasadmin/newcert.pem [Import public key cert]

\$ server\_certificate server\_2 –persona –delete default –both [delete a persona key set]

## **DISPLAYING CONTROL STATION CERTIFICATE:**

# /nas/sbin/nas\_ca\_certificate -display

Certificate:

Data:

Version: 3 (0x2)

Serial Number: 1 (0x1)

Signature Algorithm: sha1WithRSAEncryption

Issuer: O=Celerra Certificate Authority, CN=ballpeen2

Validity

Not Before: Jan 9 13:59:09 2008 GMT

Not After : Jan 7 13:59:09 2013 GMT

Subject: O=Celerra Certificate Authority, CN=ballpeen2 -----abridged-----

## **GENERATING NEW CA FOR CONTROL STATION:**

# /nas/sbin/nas\_ca\_certificate -generate

## **USING DATA MOVER PERSONA AND CONTROL STATION AS CA FOR CERTIFICATES:**

→When using the Control Station as the CA (Certificate Authority), the Certificate Request is automatically received from the Data Mover, generated, signed, and then returned to the DM

→A Control Station cannot be used to “sign” Certificates for external Hosts

→Must specify the duration in months that the certificate is valid for a CS CA

1. Create private key and certificate for use with the Control Station as CA [Not an External CA]

# **server\_certificate server\_2 -persona -generate default -key\_size 4096 -cs\_sign\_duration 8 -cn "w2k.pvt.dns"**

server\_2 :

Starting key generation. This could take a long time...

Done

**Note:** Certificate request sent by Data Mover to Control Station, generated, signed, and returned

# **server\_certificate server\_2 -persona -list**

server\_2 :

id=1

name=default

next state=Not Available

CURRENT CERTIFICATE:

id=1

subject=CN=w2k.pvt.dns

expire=SAT JUN 5 11:23:34 PDT 2010

issuer=O=Celerra Certificate Authority;CN=emcnas\_i0

# **server\_certificate server\_2 -persona -info -all | -info default**

server\_2 :

id = 1

name = default

next state = Not Available

Current Certificate:

id = 1

subject = CN=name;CN=1.2.3.4

issuer = O=Celerra Certificate Authority;CN=ballpeen2

start date = TUE JAN 29 22:52:37 EST 2008

end date = THU SEP 25 23:52:37 EDT 2008

serial number = 03

signature alg. = sha1WithRSAEncryption

public key alg. = rsaEncryption

public key size = 4096

version = 3

### **CHECKING FOR EXPIRED KeySets:**

→Run \$ server\_certificate –persona –list and check for expired certificate date

### **CLEARING EXPIRED KeySets:**

\$ **server\_certificate server\_2 -persona -clear default -both**

### **DISPLAYING CA CERTIFICATES:**

\$ **server\_certificate server\_2 -ca\_certificate -info 2**

### **CHECKING FOR EXPIRED CERTIFICATES:**

\$ **server\_certificate server\_2 -ca\_certificate -list** (check expire= dates)

### **DELETING CA CERTIFICATES:**

\$ **server\_certificate server\_2 -ca\_certificate -delete 1**

### **USING EXTERNAL CA CERTIFICATES:**

1. Generate the Data Mover’s persona public/private key set and request that a CA sign the certificate

\$ **server\_certificate server\_2 -persona -generate default -key\_size 4096 -cn 'name;1.2.3.4'**

server\_2 :

Starting key generation. This could take a long time...

Done

\$ **server\_certificate server\_2 -persona -generate default -key\_size 4096 -cn 'name;1.2.3.4' -ou 'my.org;my dept' -organization EMC -location Hopkinton -state MA -country US -filename /tmp/server\_2.certrequest.pem**

**Note:** Example of key generation using specific DM information and writing to a file

2. Write the certificate request to a file

# **server\_certificate server\_2 -persona -info default >server\_2.certreq.pem**

server\_2 :

id = 1

name = default

next state = Request Pending → Certificate request shows pending until CA provides signed certificate

3. Send the .pem file to the External CA

4. Import the signed certificate to the Data Mover:

**\$ server\_certificate server\_2 -persona -import default**

server\_2 : Please paste certificate data. Enter a carriage return and on the new line type 'end of file' or 'eof' followed by another carriage return.

**Note:** Follow the screen instructions after pasting the certificate data onto the screen

5. Verify:

**# server\_certificate server\_2 -persona -info default**

server\_2 :

id = 1

name = default

next state=Not Available → this is the key indicator to say that the certificate was successfully imported to the Data Mover

**# server\_certificate server\_2 -ca\_certificate -list**

## **5.6 CONTROL STATION PKI MANAGEMENT:**

**# /nas/sbin/nas\_ca\_certificate -display | -generate**

Certificate:

Data:

Version: 3 (0x2)

Serial Number: 1 (0x1)

Signature Algorithm: sha1WithRSAEncryption

Issuer: O=Celerra Certificate Authority, CN=ballpeen2

Validity

Not Before: Jan 9 13:59:09 2008 GMT

Not After : Jan 7 13:59:09 2013 GMT

Subject: O=Celerra Certificate Authority, CN=ballpeen2

Subject Public Key Info:

Public Key Algorithm: rsaEncryption

RSA Public Key: (2048 bit)

Modulus (2048 bit):

00:fb:d7:19:3b:e0:5b:9f:c3:a2:1b:c8:d8:d5:5f:

-----output abridged-----

X509v3 Basic Constraints:

CA:TRUE

X509v3 Subject Alternative Name:

DNS:ballpeen2, DNS:ballpeen2.w2k.pvt.dns, DNS:192.1.4.250

Signature Algorithm: sha1WithRSAEncryption

45:67:eb:31:94:f8:b5:34:e0:61:5d:8c:c5:11:a6:1d:86:4b: ---abridged----

## **USE CASE:**

### **CONTROL STATION ADMIN ROLES USING LDAP OVER SSL:**

→ Normal process should be to import a Public X.509 Certificate from a CA to the Control Station using the Security>Administrators>Domain Settings tab in Celerra Manager, “Upload New SSL Certificate”

→ This should provide the Public Key which CS can use to decrypt certificates signed by the CA. With Directory Servers, the CS can then communicate to the Directory Server and receives a certificate signed by the CA, which also contains its public key, which CS can decrypt and then use to communicate with the Directory Server

→ Control Station would initiate the LDAP connection to the Directory Server, which responds with LDAP Certificate, which has been signed by the site CA Certificate, and validates the SSL handshake with LDAP. So it's the site CA certification that must be installed on the CS.

**Issues:** ARs 150154 & 155672

1) Import file must be named with ".cer" suffix and be in PEM/Base64 format

2) If a site uses multiple certs in its CA chain, than all the certificates must be aggregated in a single file, which is then used to import to the Control Station

→ This may require that AD certificates, which are in DER binary format, be converted to PEM format and concatenated into a single text file

**Converting From DER to PEM using openssl:**

**# openssl x509 -in <DER\_cert.der> -inform DER -out <PEM\_cert.pem> -outform PEM**

**Troubleshooting Control Station LDAP SSL Certificate Issues:**

**# /nas/sbin/cso\_config**

```
usage: cso_config
-create { -domain <domain name> -pServer <primary_server>
-bServer <backup_server> -port <port_num>
-sslEnabled <true/false>
-cert <certificate_path>
-type <defaultAD | customAD | other>
-bindDn <binding_address>
-name <admin_name> -password <admin_password>
-userPath <user_search_path>
-groupPath <group_search_path>
-userAttr <user_name_attribute>
-groupAttr <group_name_attribute>
-groupClass <group_class>
-groupMember <group_member>
-autoUserCreate <true/false>
| -F <config file name> -password <admin_password> }
| -test
# /nas/sbin/cso_auth -test
# /nas/sbin/cso_config --debug on | --debug off
/nas/site/CA
/nas/site/CA/ca_certificate.pem | key.pem, etc.
/nas/site/cso_cert
/nas/site/cso_auth.cfg
/nas/site/cso_auth_config.xml
```

### **OPENSSL VERSION:**

```
# rpm -qa openssl
openssl-0.9.8g-5.4.01.EMC
```

**Note:** Updated in a 5.6 maintenance release

### **5.6 SECURITY CS AUDITING:**

Audit Logs for User & Kernel space components to be written to /nasmcd/var/auditing

```
/sbin/auditctl | /sbin/ausearch (used to find audit events of interest) | /sbin/aureport | /sbin/audited | /etc/audited.conf (uses one keyword per line for configuration information) | /etc/audit.rules (rules applied on startup)
```

/sbin/audited status | -s will report on status of auditing on the Control Station

/sbin/service audited start | stop | status | restart | reload | rotate | condrestart

#### **What actions are Logged?**

Admin access of rootfs of Data Movers; Access to sensitive CS files; Changes to auditing; Users authenticating to system

→Upgrade Hardening

→FileMover improvements; Config API, HTTPS support using SSL & PKI, Retention times set by Policy Engine (creates WORM stub file)

### **5.6 CELERRA MANAGER ENHANCEMENTS:**

- Rename file system
- Rename Data Mover
- Extend Checkpoint SavVol
- Modify NetBIOS name for NT CIFS servers or Standalone CIFS servers
- Modify Alias names for W2K, W2K3, NT, or Standalone CIFS servers
- Can change CIFS server interfaces, except for the default CIFS server
- More Provision Monitoring
- Added Security & Administrative Roles

### **PROVISIONING MONITORING:**

- Celerra Manager Basic Edition only allows Event Notifications for file system usage stats
- Celerra Manager Advanced Edition allows event notifications and monitoring of file system stats, file system usage projection, storage pool management, storage projection, data mover load
- File System notifications are polled every 10 minutes, and when condition is met, notification occurs, with only one notification per 24 hour period
- Notifications are based on Storage Usage & Projections that includes file systems and storage pools, via emails, alerts, traps
- 26 weeks of historical data kept, with default 2 weeks displayed

Celerras>Storage>Pools>Usage Statistics—new tab for storage stats

Celerras>Data Movers>Server\_2>File System>renamed Statistics to “I/O Statistics” and added “Usage Statistics” tab

### **SECURITY ENHANCEMENTS CELERRA MGR:**

--Encryption, Public Key Certificates, Role-based access, Password requirements

--Group & User Management facilities

--Remote Celerra authentication & auto-user creation upon login to Celerra Mgr

**Note:** Privileges and Roles are enforced in four CLI commands: nas\_ckpt\_schedule; nas\_copy; nas\_replicate; & nas\_task

## **5.6 ADMINISTRATIVE ROLES IN CELERRA MANAGER (GUI only):**

→Roles are based on Control Station management only, not DART, and are not object specific

→Admin roles based on privileges based on job function & area of responsibility

→An admin user account is always associated with a Primary Group [Pre-defined and Custom groups, local or domain-based]

→All groups are assigned a Role, which defines the privileges of the Users assigned to the Group

→Users can belong to multiple Groups

→All admin roles have Read privileges to view objects

→Modify privileges can make changes to existing objects

→Full Control privileges can create, delete, change objects

### **(3) Levels of Access defines Privileges:**

Read, Modify, Full Control

**/nas/site/role** →Role definitions are based on ID:name:type:description:privileges

### **CELERRA MANAGER SECURITY TAB:**

#### **→Administrative Roles for Celerra Manager: Celerras>Security: Users; Groups; Roles; Domain Settings**

##### **Security>Administrators >Users:** →two default users (nasadmin & root)

nasadmin: Primary Group “nasadmin” with Operator Role, and Groups “fullnas” with Nasadmin Role

root: Primary Group root with Root Role, with Groups “bin, daemon, sys, adm, disk, wheel” with Nasadmin Role

→Users will be automatically created during login if members of defined groups

→Membership in a Primary group and up to 16 other Groups

→All custom users belong to Nasadmin group with role of ‘Operator’

→UID Mappings can be auto-selected or specified when creating, and are added to users\_db on Control Station

### **/nas/site/users\_db**

500:0:1:: →New user ‘Samuel’ added & auto-assigned first UID of 500

→User name mappings added to pst\_db after logging into Celerra Mgr for the first time

### **/nas/site/pst\_db**

samuel:Student@10.127.114.45:4

**Note:** Manually defined membership in network group, while system auto adds to nasadmin group

### **Unique Celerra Manager Logins for NASADMIN & ROOT Accounts:**

# cat /nas/site/pst\_db/cat usrenv2Id.tbl

nasadmin:unknown\_login@192.1.4.212:1

root:unknown\_login@192.1.4.212:2

nasadmin:administrator@gloria.w2k.pvt.dns:3

root:administrator@frodo.w2k.pvt.dns:5

**Security>Administrators >Groups (role) GID** →predefined groups shown—custom groups can be local or domain-based

backup (backup\_operator) 506 →access to filesystems, checkpoints, VTLUs

filemover (filemover\_application) 507 →full control filemover & tasks

fullnas (nasadmin) 503 →RW access everything except root-only functions

nasadmin (operator) 201 →RO access to everything

network (network\_admin) 501 →RO access DMs, CS, CIFS Servers; RW network devices & services

opadmin (operator) 502 →RO access to everything

root (root) 0 →RW access to everything

security (security\_operator) 504 →RW access to user & role mgmt only

service (nasadmin) 500 →RW access to everything other than root-only function

storage (storage\_admin) 505 →RO access DMs; Full access storage, volumes, pools, file systems

### **Security>Administrators >Roles:** →default system-defined roles ([/nas/site/role](#))

1:root:system-defined:System Root Role:adminDomain,adminGroup,adminRole,adminUser, etc.

2:nasadmin:system-defined:Standard EMC Administrator Role:adminDomain,adminGroup,adminRole, etc.

3:operator:system-defined:Standard EMC Operator Role:task=\*!\*=collect,download,list,ping,validate

4:security\_operator:system-defined:Security Operator...Role Management and User/Group Management, etc.

5:backup\_operator:system-defined:EMC Backup Operator Role with full access to ckpt and VTLU, etc.

6:network\_admin:system-defined:EMC Network Admin Role:device, etc.

7:storage\_admin:system-defined:EMC Storage Admin Role:alert,clarionSystem, etc.

8:filemover\_application:system-defined:FileMover Application Role:dhsmConnection, etc.

### **Security>Administrators>Domain Settings:**

→Various domain properties displayed on this page

→Allows use of Active Directory or other Directory services for remote authentication of Users logging in through Celerra Mgr

### **REMOTE CELERRA MGR AUTHENTICATION/AUTO-USER CREATION:**

--Administrative User accounts can be local or mapped to domain-based accounts

--Domain-based User accounts require access to LDAP directory server for authentication before login can be successful, and must belong to at least one of the default or custom-created Groups on Celerra Mgr

--Domain-based users that are members of custom groups created in Celerra Mgr will be auto-created when the User account logs in, provided the LDAP server is available for authentication, and User uses domain password & name

### **REQUIREMENTS FOR USING DIRECTORY SERVICES:**

--FQDN

--IP Address of AD Server

--SSL Certificates if enabled [Port 636 when enabled, otherwise port 389 used if disabled]

--Use a Directory Services server type [Default Active Directory; Custom Active Directory; Other Directory Servers]

--Username & Password [Distinguished names & search paths for OpenLDAP]

### **OTHER DIRECTORY SERVICES [OpenLDAP]:** Additional fields required when defining OpenLDAP Directory Services:

Distinguished Name: cn=administrator,cn=Users, dc=nas10, dc=emc,dc=com

User Search Path: cn=Users, dc=nas10,dc=emc,dc=com

User Name Attribute: cn

Group Search Path: cn=Users, dc=nas10,dc=emc,dc=com

Group Name Attribute: cn

Group Class: <select from dropdown>

Group Member: uniqueMember

### **CELERRA PASSWORD MANAGEMENT:**

#### **Default Password Policies:**

→Min. password length 8 characters [6-15 possible]; Max. 3 attempts to define a passwd before failing; Min. 3 characters for new password that were not in the old passwd; Min. of 1 numeral in the password; Force passwd change after expiration period

→Password settings defined in /etc/pam.d/system\_auto, using pam\_cracklib.so, but not turned on, by default. Running the nas\_config –password command enables the use of policies

**Note:** Use /nas/sbin/nas\_config –password to set & modify password policies

→Password policies turned on or modified using /nas/sbin/nas\_config –password

→Use nas\_config –password –default to reset policies to defaults

→Policies do not apply to Nasadmin & Root User accounts

→Policies do not apply to user accounts authenticated by LDAP or AD services

#### **Password Expiration Policies:**

--Not managed or enabled by nas\_config utility

--Edit /etc/login.defs to modify or update available settings

--/etc/login.defs defines auto UID/GID min & max values used by the system, as well as password aging [default=99999 days]

### **PASSWORD CHANGE ISSUE FOR NASADMIN & ROOT:**

AR117242 & emc184679 documents that if CLI # passwd nasadmin command is used to change passwords, and then CS is rebooted, the password reverts back to the original. It appears that the “shadow-“ file is being copied back over the “shadow” file that does get changed during the password change command.

### **Changing nasadmin or root passwords using Celerra Manager with NAS 5.6:**

1. Log into Celerra Manager as root user

2. Navigate to Celerras>Security>Administrators>Users>rightclick & select properties for either nasadmin or root account, then change the password in the "Password" and "Confirm Password" boxes, then click apply.

**Note:** Using Celerra Manager properly updates the password and immediately synchronizes the change to /etc/shadow and /etc/shadow- files

### **References:**

→Celerra Manager Online Help seems to be the main documentation on Administrative Roles

### **5.6 FILEMOVER ENHANCEMENTS:**

--DHSM service can be stopped & restarted without impact to CS [server\_http service –stop | -start]

--FileMover commands issued by DART to port 5080, by CS to port 5081, Policy Engine can change port

--Threads can be updated on-the-fly

--fs\_dhsm offers ability to use HTTPS for connections

### **DART provides HTTP Server:**

# server\_http server\_x –service dhsm –start

### **5.6 VERIFYING DART HTTP SERVICE:**

# server\_http server\_2 -info

server\_2 : done

DHSM FACILITY CONFIGURATION

Service name: EMC File Mover service

Comment: Service facility for getting DHSM attributes

Active: True → This output indicates DART is running the HTTP Server

## **VERIFYING HTTP CLIENTS:**

# .server\_config server\_2 -v "httpclient audit"

1211900252: HTTP\_CLIENT: 6: dump of all http client:

1211900252: HTTP\_CLIENT: 6: -----

1211900252: HTTP\_CLIENT: 6:

dump of all http peers:

1211900252: HTTP\_CLIENT: 6: -----

1211900252: HTTP\_CLIENT: 6: HTTPCLIENT audit HTTP\_STATUS:200 OK

## **Troubleshooting HTTPS:**

--Setup Null encoding to troubleshoot problems (normally encrypted)

\$server\_http server\_x -modify DHSM -sslcipher NULL [when DM is an http server]

\$server\_config server\_x -v "param ssl cipher=NULL" [when DM is an http client]

--verify that certificate is good using server\_certificate

--verify that policy engine can connect to CS port 443

--verify xml\_api service running on CS

## **5.6 WINDOWS 2008 (Longhorn) SUPPORT:**

→ Windows Longhorn compatibility (DART can co-exist as a Windows 2003 member server)

→ Compatible with Vista clients

### **Windows 2008 Features:**

Self-Healing File System—system can perform online repairs, no more chkdsk offline, etc.

Multiple Sessions with Terminal Server—allows for faster User logons

Clean Shutdown—O/S will wait to shutdown applications gracefully vs. traditional 20 second timer

Kernel Transaction Manager (KTM)—better management of threads seeking access to same resources

SMB2—better network performance and scalability

Address Space Load Randomization (ASLR)—better protection against malware by using different memory locations to run kernel programs

Windows Hardware Error Architecture (WHEA)—standardization of errors reported on HW systems

Windows Server Virtualization—some hot pluggability added in SP release for hardware components

PowerShell—scripting language

Server Core—ability to install Servers for more specialized roles, without all the other baggage that usually comes

### **SMB2 Protocol:**

Support for more concurrent open files and number of shares

Better transaction support

Client Side Encryption

Support for symbolic links

Support for compounding operations, resulting in less overhead traffic

Support for larger buffer sizes

### **Limitations:**

→ W2008 Clusters not supported [new HA Shares and cluster technology]

→ Windows 2008 domain needs to run in Windows 2003 compatibility mode if NTLM is used for authentication

→ Or a DC change can be made: Computer Configuration>Administrative Templates>System>Netlogon>Allow cryptography

algorithms compatible with Windows NT 4.0

**Note:** Registry location is HKLM>Software>Policies>Microsoft>Netlogon>Parameters>Allow NT4Crypto

→ Not compatible with Vista clients using UAC (User account Control)

→ W2008 Offline folders not supported

→ W2008 FSR (File Replication Service) not supported

→ W2008 EFS (Encrypted File Systems) not supported

→ SMB2 protocol not supported (support added with 5.6.42)

## **5.6 OPENLDAP SUPPORT:**

→ iPlanet & OpenLDAP Directory Services support using SSL (RFC 2307)

→ Extension of iPlanet LDAP support for Network Information Service per RFC 2307, expected to be a replacement for traditional NIS architecture

→ Can support only a single LDAP domain, with DM as LDAP client

→ Uses file-based configuration /etc/ldap.conf file vs. iPlanet client configuration profile [DUAConfigProfile]

### **AUTHENTICATION RULES:**

--If binddn not used & SSL disabled = Anonymous [no SSL]

--If binddn & password used, but SSL disabled = Simple [no SSL]

--If binddn not used, sselpersona not used, but SSL enabled = anonymous over SSL

--If binddn & password used, and SSL enabled = Simple over SSL

--If binddn not used, and SSL enabled, and sslpersona configured = SSL-based CA

#### **SPECIFYING SIMPLE AUTHENTICATION:**

```
$ server_ldap server_2 -set -p -domain Celerra.pvt.dns -servers 10.241.168.21 -binddn
```

```
"uid=admin,ou=Administrators,ou=TopologyManagement,o=NetscapeRoot"
```

#### **SETTING LDAP DOMAIN:**

```
$ server_ldap server_2 -set -domain celerra1.pvt.dns -servers 10.241.168.21 [Using FQDN method]
```

```
$ server_ldap server_2 -set -basedn ou=celerra1, o=pvt, c=dns -servers 10.241.168.21 [Distinguished Name Format Base DN]
```

#### **ENABLING SSL:**

Configure on LDAP server first

```
$ server_ldap server_2 -set -sslenabled y [Uses port 636 by default]
```

#### **SETTING SSL PERSONA:**

```
$ server_ldap server_2 -set -sslpersona default
```

#### **SPECIFYING CIPHER SUITE:**

```
$ server_ldap server_2 -set -sslcipher default
```

#### **NAS 5.6 IPV6 PHASE I:**

→Available via RPQ only, supporting NFSv2, v3, FTP, TFTP & NDP protocols

→CIFS & NFSv4 are not supported with the GA release of 5.6

→CLI configuration

→DNS, NIS, local file, & LDAP lookup Name Service support

→Supported with DM failover & HA networking

→Hidden support for Stateless Autoconfiguration of IPv6 addresses per Autoconfig RFC

#### **STATELESS AUTOCONFIG IPv6:**

→Mandatory part of IPv6 protocol but does not work well with CIFS applications

1. Set param and reboot Server

```
param IPv6 autoconf=1
```

2. Enable AutoConfig manually using server\_ifconfig for Linklocal Broadcast Domain or automatically after using .server\_config EUI-64 address

```
$ server_ifconfig server_2 -create -Device cge0 -name cge0 -protocol IP6 3ffe:0000:3c4d:0015:0435:0200:0300:ED20
```

```
$ server_ifconfig server_2 -create -Device cge0 -name cge0 -protocol IP6 3ffe:0000:3c4d:0015:0435:0200:0300:ED20 /48
```

**Note:** Creating interface with a different network prefix length than the default 64

```
$ server_ifconfig server_2 -all -ip6
```

```
$ .server_config server_2 -v "ifconfig cge0ll device=cge0 protocol=IP6 local=link vlan=0"
```

3. DART listens to Router Advertisements and configures Global IPv6 addresses for each Network prefix advertised

4. Names for addresses will be auto-created and assigned by DART for each device

cge0\_0\_bd= broadcast device cge0 with Vlan 0

#### **CREATING LINK LOCAL ADDRESS FOR A DEVICE:**

```
$ server_ifconfig server_2 -create -Device cge1 -name cge1 -protocol IP6 fe80::260:16ff:fe0b:1234
```

#### **NAME SERVICE INFORMATION:**

```
$ server_dns server_2 -ip6
```

```
$ server_nis server_2 -ip6 | -status
```

#### **NFS EXPORT SYNTAX:**

```
$ server_export server_2 -Protocol nfs -option ro=172.30.55.0/255.255.255.0:[ef20:4567:8900::9abc:1000] /fs1
```

**Note:** Example of access list with both IPv4 and IPv6 addresses—note the [ ] for IP6 and colon separator between addresses

#### **DAD (Duplicate Address Detection):**

a) Two Neighbor Solicitations are sent for every IP6 address when an interface is first brought online, or configured, at 200ms intervals

b) If a Neighbor Advertisement is received within 600ms, the address is marked as a Duplicate and requires User intervention

#### **DART PARAMS FOR DAD:**

param IPv6 dadtimer [time between each Neighbor Solicitation]

param IPv6 dadCount [number of NS's sent]

#### **CLEARING AUTO CONFIGURED INTERFACES:**

```
$ .server_config server_2 -v "ifconfig clearauto"
```

#### **CHECKING CONNECTIVITY:**

```
$ server_ping6 server_2 -interface cge0 3ffe:0000:3c4d:0015:0435:0200:0300:00aa
```

```
$ server_ping6 server_2 -send 3ffe:0000:3c4d:0015:0435:0200:0300:00aa
```

**Note:** Examples of pinging from cge0 to a network host and continuous pinging to a network host

#### **ENABLING VLAN ON A DEVICE:**

```
$ server_ifconfig server_2 -create -Device cge0 -name cge0 -protocol IP6 3ffe:0000:3c4d:0015:0435:0200:0300:ED20 vlan=40
```

```
$ server_ifconfig server_2 cge0 vlan=40
```

**REMOVING THE VLAN TAG FROM A DEVICE:**

```
$ server_ifconfig server_2 cge0 vlan=0
```

**NAS 5.6 GIDMAP ENHANCEMENTS:**

→gid\_map enhancements for DUDL—automatic copy of gid\_map kept, new default size 262152 bytes. Corruption of original and copy of gid\_map would still require restore from backup

**MANUALLY RECOVERING GID MAP FROM LOST+FOUND:**

- server\_mount server\_x -o gid32Ignore fs1 /tmp
- look for gid\_map in lost\_found [identify by size 262152]
- verify file \$.server\_config server\_2 -v "gidmap chkfile /tmp/lost+found/fs1"
- copy file to file system .etc directory as gid\_map
- remount file system without the gid32Ignore option

**RESOLVING GID MAP CHECKSUM ISSUES:**

- Look for checksum issue in the Server Log
- Add to /nas/server/slot\_x/param  
param ufs fixGidCorruption=1
- Reboot Data Mover and checksums should be corrected on remount
- Remove param and reboot Data Mover

**GID Map Commands:**

```
$ server_mount server_x -o gid32Ignore fs1 /tmp [mount without the gid32 function]
$ .server_config server_x -v "gidmap chkfile /tmp/<path>" [verifying a gid_map file]
$ server_file server_x -put <file> /fs1/\Vetc/gid_map [copying gid_map to a location]
$ .server_config server_x -v "gidmap resetchecksum /fs1/.etc/gid_map" [resetting checksum of gid_map]
```

→Conversion of IUsermapper, Shares, Secmap to Berkeley database (NameDB)—dbms database format

**GID MAP ISSUE DURING UPGRADES TO NAS 5.6 (emc197578):**

File Systems may not be mounted after upgrading to NAS 5.6 because the gid\_map file size has been corrupted. PreUpgrade check has been built to flag this so that corrective action can be taken prior to the upgrade, rather than recovering file systems after the upgrade. Correct size of gid\_map for NAS 5.6 is 262152.

**PreUpgrade Check:**

E000124: File system (fs01 ufs /fs01 server\_2) mounted on server\_2 gidmap file is the wrong size (524288) it should be (262144). Check Primus Solution EMC197340 for additional information.

**Recovery Steps if File Systems are unmounted:**

- Permanently unmount the affected file system.  

```
# server_umount server_X -p <fs_name>
```
- Mount the affected file system to an ALTERNATE mountpoint with rw and gid32ignore options. IMPORTANT: Make sure you an ALTERNATE mountpoint so that the file system cannot be accessed via its original exports and shares during this procedure.  

```
# server_mount server_X -o rw,id32ignore <fs_name> /<TEMP_mountpoint>
```
- Get a copy of the wrong-sized gid\_map file from the affected file system.  

```
# server_file server_X -get /<TEMP_mountpoint>\Vetc/gid_map /<workdir>/gid_map_<fs_name>
```
- Verify the wrong size, truncate the file to correct size, and verify new size.  

```
# ls -l /<workdir>/gid_map_<fs_name>
-rw-r--r-- 1 root root 524288 Sep 17 17:20 gid_map_<fs_name>
# dd if=/<workdir>/gid_map_<fs_name> of=/<workdir>/gid_map_<fs_name>_fixed bs=256k count=1
1+0 records in
1+0 records out
# ls -l /<workdir>/gid_map_<fs_name>_fixed
-rw-r--r-- 1 root root 262144 Sep 17 17:25 gid_map_<fs_name>_fixed
```
- Put the fixed gid\_map file on the affected file system.  

```
# server_file server_X -put /<workdir>/gid_map_<fs_name>_fixed /<TEMP_mountpoint>\Vetc/gid_map
```
- Permanently unmount the affected file system.  

```
# server_umount server_X -p <fs_name>
```
- Remount file system to original mountpoint making sure to specify "rw" option and any other original mount option recorded in the prestep.  

```
# server_mount server_X -o rw,<other original options> <fs_name> /<original_mountpoint>
```
- Clean-up any temporary mountpoints created during this procedure, and clean-up or save working directory files as desired.

**NAS 5.6 PERFORMANCE STATMON TOOL (server\_stats):**

→Creation of performance statistics tool for DART similar to sysstat output and Unix SAR (nas\_stat & server\_stats)  
→Serves as single interface for statistic collection for DART components CPU, CIFS, NFS, CACHES, NET, DVOL, FS VOL  
→Two types of collection: Summary and Table collections

**Note:** Data Mover will use a “statmonService” listening on port 7777, and can be disabled in /nas/server/slot\_x/eof file by commenting out:

**/nas/server[slot\_\*/eof]**

# statmonService start port=7777 allow=128.221.252.100:128.221.252.101:128.221.253.100:128.221.253.101

--Statistics will be high-level system summaries (by kibibytes = 1024 bytes), and in-depth table collections by logical unit

**\$ server\_stats**

USAGE:

server\_stats &lt;movername&gt;

[ { -summary {basic|cifs|nfs|caches} [...] ]

| -table {net|dvol|fs|voll|cifs|nfs} [...] →Table format to be used

| -table {net|dvol|fs|voll|cifs|nfs}

| -sort &lt;field\_name&gt;

| -order {asc|desc}

| -lines &lt;lines\_of\_output&gt; ] [...] ]

} [...] ]

[-count &lt;count&gt;] →Number of reports by interval period, default is 10,000

[-port &lt;port&gt;] →Default port to DMs is 7777

[-interval &lt;seconds&gt;] →Default is 15 seconds

[-terminationsummary {no|yes|only}] →Gives user choice of seeing ctrl+c termination summary message or not. ‘Only’ option suppresses time series and only reports the summary. Default is ‘yes’.

[-format {text|csv}] →Default is text

[-titles {never|once|repeat}] →Controls generation of titles for reports

**# server\_stats server\_2**

server\_2 CPU Network Network dVol dVol

Timestamp Util In Out Read Write

% KiB/s KiB/s KiB/s KiB/s

09:39:42 0 0 0 128 11

09:39:57 0 0 0 128 477

09:40:12 0 0 0 128 4

09:40:27 0 1 1 129 116

server\_2 CPU Network Network dVol dVol

Summary Util In Out Read Write

% KiB/s KiB/s KiB/s KiB/s

Minimum 0 0 0 128 4

Average 0 0 0 128 125

Maximum 0 1 1 129 477

**Note:** Default interval is 15 seconds. Stop server\_stats polling to console using ctrl + c**\$ server\_stats server\_2 -summary basic | cifs | nfs | caches****Note:** Basic provides system overview, CIFS provides cifs summary, NFS provides nfs summary, & Caches provides memory summary for DNLC, Open Files, & Buffers.**\$ server\_stats server\_2 -i 3 -c 5****Note:** Example of basic summary with 3 second interval for total of 5 counts**\$ server\_stats server\_2 -summary basic,nfs,cifs,caches -table net,dvol,fsvol,nfs,cifs -interval 10 -format csv -count - > /tmp/stats.txt****Note:** Example of a command used to investigate a performance issue with SMB2, results written to the file in 10 second increments**5.6 CCMD MESSAGES:**

→CCMD will change logging mechanisms on DART &amp; CS. LogEvent interface for DART will be enhanced to handle CCMD messages, on-disk format will be changed, and logmsg interface will be enhanced for Server Log. CS sys\_log format will be changed to CCMD, and nas\_eventcollector, nas\_eventlog, PostEvent CLI, ReportEvent API, nas\_event, etc., all will be enhanced to support.

**5.6 FILE SYSTEM FAULT CONTAINMENT/FAILOVER:**

→File System fault containment and failover to be allowed [user must initiate manual fsck]

→Auto-FSCK will not be started automatically on detection of file system corruption. Rather, the fs will fail to mount. This will allow Data Mover to come back up with all other file systems active.

→Failover from production slot to failover slot will succeed if mount fails on any file system by continuing to mount all others and completes the failover process

**# server\_param server\_2 -facility ufs -info skipFsck -v**

server\_2 :

name = skipFsck

facility\_name = ufs

default\_value = 1

current\_value = 1  
configured\_value =  
user\_action = none  
change\_effective = immediate  
range = (0,1)  
description = If enabled, fsck will not be run on a corrupted FS while booting up and FS will be left unmounted.

#### **UNCORRECTABLE SECTOR ISSUE:**

See AR114835. GA release will not have either Rvector or cse\_recover utilities for repairing Uncorrectables until Maintenance release 2, so the fallback will be to use Volcopy in the interim.

#### **5.6 DUMP ENHANCEMENTS:**

→Dump enhancement adds two new Panic Handler codes

RC\_DART\_PANIC\_HANDLER\_COMPLETE(16) →CS initiates DM failover after initial panic

RC\_DART\_PANIC\_DUMP\_COMPLETE(13)

#### **5.6 iSCSI SNAPS IMPROVEMENTS:**

→iSCSI is a network service that runs on top of TCP/IP on the Data Mover

→iSCSI snaps & replication management via CLI, iSCSI Replication integration (CLI & GUI), iSCSI AIX NACA Support (SCSI-3 error handling over full duplex and non-interlocked parallel buses; if CHECK condition is encountered, all other commands are put on hold until the condition is cleared)

→Writes are optimized since indirect block copies occur asynchronously

##### **PLU:**

Production Logical Unit—primary iSCSI object that is a file that has space persistently reserved in a file system. Space is always reported as consumed from Celerra file system perspective, but Client may only be using x\_number of actual blocks.

##### **SLU: “Snapshots”**

Snap Logical Unit—contains pointers to blocks allocated by PLU.

--Size is number of used blocks written to PLU at time of snap creation, or number of new blocks allocated since a previous snap

--Created by Host Application, Replication Manager, or now, via Celerra CLI

--Up to 2000 snaps per Data Mover

--Snapshots are crash consistent only

##### **TWS:**

Temporary Writable Snap—Becomes a Writeable version of the SLU when the SLU is promoted to an iSCSI Initiator; used during Snap Promotion and Restore. All contents deleted when a snap is demoted.

##### **Commands:**

**\$ nas\_license –create iscsi**

##### **CREATING SNAPSHOT:**

**\$ server\_iscsi server\_2 –snap –create –target ns22 –lun 0 –data “snap1” | -info** (Application label = CelerraCS for CS-created snaps) | -list | -modify | -delete | -promote <snap\_name> -initiator <initiator\_name> | -Demote <snap\_name> -Force [cmd extended to manage snaps from Dart]

##### **DELETING SNAPSHOT:**

--Must demote first

--Use -Force switch if a snapshot created by some application other than CS CLI

**# server\_iscsi server\_2 –snap –delete snap1 | -target ns22 –lun 0**

##### **RESTORING PLU FROM SNAPSHOT:**

**# server\_iscsi server\_2 –snap –restore snap1**

##### **SETTING UP ISCSI ON CELERRA:**

1. Create Target

**# server\_iscsi server\_2 –target –alias ns22 –create 1000:np=10.127.62.31**

2. Create LUN

**# server\_iscsi server\_2 –lun –number 0 –create ns22 –size 20G –fs fs\_iscsi –vp yes**

3. Set LUN Mask

**# server\_iscsi server\_2 –mask –set ns22 –initiator iqn.2006-01.com.emc.iscsi:rh-199 –grant 0**

4. Start & Verify iSCSI Service

**# server\_iscsi server\_2 –service -start**

**# server\_iscsi server\_2 –service -status**

#### **5.6 WRITEABLE SNAPS (WCKPT):**

→Writeable Snap creation—ability to turn baseline RO Snap/Checkpoint into R/W file system using Snapsure v2.5 technology

→Changed blocks are written to the SavVol [which is shared with RO checkpoints]

→Limit of (16) writeable snaps per PFS, but only one per RO baseline checkpoint snap

→Can have a total of 96 RO checkpoints, and 16 Writeable snaps, per file system

→Typically useful for backups, testing, fsck's, etc.

→Writeable snaps cannot be “scheduled” as can regular Snapsure Checkpoints, nor RO snaps if using Writeables

- Can manage Writeable snaps from GUI & CLI
- Can create, delete [must delete Writeable before baseline can be deleted], or restore Writeable snaps
- Baseline and Writeable snaps cannot be refreshed
- Deleted Writeable snaps return space to SavVol [though SavVol size never contracts and is reclaimed until all snaps deleted]
- WCKPT not visible from .ckpt directory or from Explorer
- Cannot take a Writeable Snap from Internal checkpoints

#### **BITMAP:**

Identifies the changed data blocks

#### **BLOCKMAP:**

Records addresses of saved data blocks

#### **Allowed Operations with Writeable Snaps:**

Can perform fs\_copy, nas\_copy, MPFS, nas\_fsck, NDMP, NMFS, mount & umount, export of WCKPT

#### **SavVol Full Scenario:**

- Code will automatically delete the oldest RO snaps to create enough space for expanding SavVol [inactivates Writeable if required]

#### **Creating Writeable Snap:**

**\$ fs\_ckpt fs1\_ckpt1 –Create –readonly n**

**Note:** Mounts automatically, but is not exported

#### **Default naming convention:**

**fs1\_ckpt1\_writeable1**

#### **Writeable Snap Restore—mount RO if doing manual CLI restore:**

**\$ /nas/sbin/rootfs\_ckpt fs1\_ckpt1\_writeable1 –n fs1 –Restore**

**Note:** Writeable restore from GUI remounts as RO before the Restore

#### **5.6 iSCSI LUN RECOVERY/HARDENING:**

→iSCSI LUN Recovery for data recovery and hardening (iSCSI LUNs are special files within the UxFs file system)—purpose of feature is to try and maintain LUN availability for recovery, etc.

1. First step would be to run FSCK to correct file system metadata

**Note:** Version Files are no longer truncated by nas\_fsck. Affected blocks are marked by a special -2 address. Each Version File represents a separate iSCSI LUN. File systems can contain a number of iSCSI LUNs, therefore Version Files. Version Files are presented to iSCSI Host as SCSI disk [LUN]

2. Second step would be to invoke iSCSI Host recovery application checks

**\$ server\_iscsi server\_2 –lun –info 1**

**Note:** Command summary will show lun as “Healthy” or “Corrupt”

#### **5.6 iSCSI DR ENHANCEMENTS:**

→Creation of a DR framework using Client/Server architecture

→Client agent used on iSCSI host

→Automated recovery steps and centralized recovery operations from RM GUI [Replication Manager 5.1]

#### **5.6 STORAGE HARDENING:**

→Storage Hardening: Basic Volume creation changes, maintenance states, and I/O failure handling

#### **5.6 DBCHK ENHANCEMENTS:**

→Checks DART in-memory tables against Control Station db

→Adds –local option to check db files from customer site

→DBCheck Enhancements

Checks for volume consistency between CS & DMs; checks AVM groups & profiles/storage profiles; checks Replication sessions

→Checks local and loop-back

**# /nas/tools/dbchk -**

usage: dbchk -fhlpqwxvVs

- f fast, skip long running checks
- h help, display this message
- l local, use files in current directory
- p probe, use server\_devconfig
- q quite, suppress all output and only set exit code
- v verbose, print extra information
- w warning, do additional checks for warnings
- x extra, additional long running error checks
- V Volume consistency check mode [can not be used with -l option]
- s Suggestion mode [used with (-V) option only]
- R Replication V2 consistency check mode[can not be used with -l option]

#### **5.6 NEW LOCATION FOR NASDB BACKUPS:**

**# ls -la /celerra/backup**

```
-rw-r--r-- 1 nasadmin nasadmin 326360 Jan 10 09:01 _dbms_backup.01.tar.gz
-rw-r--r-- 1 nasadmin nasadmin 241856 Jan 10 08:01 _dbms_backup.02.tar.gz
-rw-r--r-- 1 nasadmin nasadmin 326360 Jan 10 09:01 _dbms_backup.OK.tar.gz
drwx----- 2 root root 16384 Jan 9 09:13 lost+found
-rw-r--r-- 1 nasadmin nasadmin 2315660 Jan 10 01:01 _nasbkup.01.tar.gz
-rw-r--r-- 1 nasadmin nasadmin 2314237 Jan 10 02:01 _nasbkup.02.tar.gz
-rw-r--r-- 1 nasadmin nasadmin 2314361 Jan 10 03:01 _nasbkup.03.tar.gz
-rw-r--r-- 1 nasadmin nasadmin 2313857 Jan 10 04:01 _nasbkup.04.tar.gz
-rw-r--r-- 1 nasadmin nasadmin 2313852 Jan 10 05:01 _nasbkup.05.tar.gz
-rw-r--r-- 1 nasadmin nasadmin 2313973 Jan 10 06:01 _nasbkup.06.tar.gz
-rw-r--r-- 1 nasadmin nasadmin 2314591 Jan 10 07:01 _nasbkup.07.tar.gz
-rw-r--r-- 1 nasadmin nasadmin 2314596 Jan 10 08:01 _nasbkup.08.tar.gz
-rw-r--r-- 1 nasadmin nasadmin 2314757 Jan 10 09:01 _nasbkup.09.tar.gz
-rw-r--r-- 1 nasadmin nasadmin 2315529 Jan 9 22:01 _nasbkup.10.tar.gz
-rw-r--r-- 1 nasadmin nasadmin 2315532 Jan 9 23:01 _nasbkup.11.tar.gz
-rw-r--r-- 1 nasadmin nasadmin 2315598 Jan 10 00:01 _nasbkup.12.tar.gz
-rw-r--r-- 1 nasadmin nasadmin 3639417 Jan 10 09:01 nasdb_backup.1.tar.gz
-rw-r--r-- 1 nasadmin nasadmin 3425910 Jan 9 09:17 nasdb_backup.b.tar.gz
```

**Note:** Other location remains in /nas/var/backup—no more backups in /home/nasadmin

## NAS 5.6 DM CAPACITY INCREASES:

### DATA MOVER CAPACITY CHART FOR NAS 5.6:

|       |         |            |                                            |
|-------|---------|------------|--------------------------------------------|
| NSX   | 32TB/FC | 64TB/ATA*  | Clarion and/or Symmetrix, up to 4 backends |
| NS80  | 32TB/FC | 64TB/ATA*  | " "                                        |
| NS40  | 24TB/FC | 64TB/ATA*  | " "                                        |
| NS20  | 16TB/FC | 48TB/ATA** | ICA-only                                   |
| NS350 | 10TB/FC | 16TB/ATA   |                                            |
| NS700 | 16TB/FC | 32TB/ATA   |                                            |
| NS500 | 10TB/FC | 32TB/ATA   |                                            |
| 514   | 8TB/FC  | 32TB/ATA   |                                            |

\*64TB supported for ATA using B2D or MDL solutions [using RepV2]

\*\*48TB supported for ATA using B2D or MDL solutions [using RepV2]

## 5.6 PSEUDO IP DEVICES:

→Special Psuedo IP devices for Morgan Stanley, which will allow “pseudo” device to be created using server\_sysconfig, and management of RIP using server\_rip. Limit of one pseudo device per physical Data Mover, and no creation on VDMs. Purpose of feature is to allow for Disaster Recovery for Clients connecting to common destination IP address to access a service (NFS, CIFS, etc). RIPv2 will be supported and tunable via the server\_rip command.

## 5.6 CYCLINDER GROUP CACHE IMPROVEMENT:

→Cylinder group cache retention changes

Cylinder group cache will be periodically flushed to storage, but in-memory copies of cache are retained and apply LRU algorithms

## 5.6 AIX NACA SUPPORT:

→NACA support for AIX 5.2/5.3 O/S [Normal Auto Contingent Allegiance bit]

Error recovery feature in Unix for SCSI commands

If SCSI control byte is set for NACA=1, then a clear ACA request is needed to clear the error on the device

Celerra iSCSI targets will support Abort Task Set, Clear Task Set, and Clear ACA (new)

## SETTING UP AIX HOST FOR CELERRA iSCSI:

I. Install EMC ODM package on AIX host by copying package to /tmp directory, uncompressing, the using SMIT tool to install  
# /tmp/smit installp [Install “EMC CELERRA AIX Support Software”]

II. Configure iSCSI on AIX

- Run SMIT tool>Devices>iSCSI>iSCSI Protocol Devices>Change
- Select iscsi0 and use lsattr to get attributes and AIX Initiator Name
- Create Celerra file system(s), create target and bind to portal, then grant access rights to AIX Initiator Name
- Enter Celerra target info in /etc/iscsi/targets file on AIX host

10.169.1.10 3260 iqn.1992-05.com.emc:s2t1

e. Discover Celerra iSCSI Targets and Luns

cfgmgr -l iscsi0 | lsdev

## 5.6 CELERRA EVENT ENABLER Version 4.0.2(CELERRA CEE) [aka, CELERRA CEPA; CEPP]:

→Latest version 4.5.1

→64-bit version introduced with 4.5.0.4 July 2009

→Overall feature name is CEPA, Celerra Event Publishing Agent

→CEE stands for “Celerra Event Enabler”, which is a new framework that supports both CAVA and CEPA event publishing and processing [aka CEPP or CEPA]

→The CEPA feature is a system whereby DART CIFS servers will provide certain Event Notifications [file/directory create, open, close, delete, etc.; metadata file/directory & acl changes] via triggers and a delivery mechanism, to the CEPA Servers, which in turn directs the information to the CQM-based application for policy decisions, with CIFS Clients receiving notification from DART on whether the action was successful or has errored

→With the use of the CEPA API, 3rd Party CQM vendors can integrate with Celerra

→CEPA uses an API to support the processing and publishing of events from DART to registered applications

→CAVA is no longer referred to separately

→CQM stands for the Content/Quota Management Class of Application, which ties into the CEPA API to register, receive, and act on events taking place on Celerra CIFS Servers using IDL & XML DTD files [later rollouts will support Auditing, Indexing, & Monitoring Classes of Applications for Windows & NFS]

→The CEPA Client runs as the EMC CAVA service on a Windows Server

→CAVA has been in place for years and is the EMC agent running on a Windows system that communicates with AV engines for scanning of CIFS files

→Currently qualified vendors for CQM applications are NTP Software and Northern Parklife

→CEE is a framework to be used by CAVA and CEPA (Celerra Event Publishing Agent) facilities for CIFS protocol only, during the initial 5.6 release. Later CEPA code enhancements will provide Auditing and Event triggers for HTTP and NFS, etc.

**Note:** The CEE framework is installed using the EMC\_CEE\_Pack.exe. The Windows service that CEE runs under is the “EMC CAVA” service—‘Provides Anti-Virus and Event Publishing services for EMC NAS’. The CAVA service needs to be changed to run by a CAVA/CEE Windows Domain User account, not the default System account. Use the “Celerra Management” MMC snapin to grant the “EMC Virus Checking” & “EMC Event Notification Bypass” rights to the Domain Account used to run the “EMC CAVA” service.

→Phase I Cognac will cover the CQM Classes of Applications, which are based on Pre-Event blocking

→Phase II will add Auditing of CIFS events as a Class of Applications, using DatAdvantage (Varonis) as the 3<sup>rd</sup> party application

→CEPA is a facility within the CEE framework, responsible for processing & publishing events from DART to applications registered via the CEPA API

→CEPA API is the interface between the Windows CEPA Client and the CQM applications, using XML over RPC

→CEPA is an EMC agent running on a Windows system for providing external applications with notification and control functions regarding changes made by CIFS clients on file systems, in real time. Applications can register to receive event notifications and context (i.e., metadata) from Celerra CEPA.

### **CEPA FRAMEWORK:**

→DART contains the CAVA & CEPA functionality, called mechanisms, which consist of event triggers and a delivery method

→“CEPA functionality” is how we should refer to the Data Mover piece

→The Windows Server running the CEPA Facility is the Client API, and also runs the CAVA Facility, which runs as the EMC CAVA service

→“CEPA Client” or Server is how we should refer to the Windows API

→The CQM class of applications registers with the CEPA Client API, though all events are driven to the CQM application, not pulled

→Endpoints refers to the IP addresses of the Host that runs the CQM application, which communicates with CEPA Service via RPC

### **CEPA LIMITATIONS:**

→First release only applies to CIFS environment

→CLI interface only, no Celerra Manager functionality

→Only a single CEPA pool name is defined in the cepp.conf file (can use multiple CEPA Clients)

→CEPA Client API & CQM applications communicate using XML over RPC

→First Cognac release offers only CQM (Content/Quotas Management)

→Data Mover CEPA facility acts as the originator of the CAVA or CEPA triggers, using Pre-Event Blocking of events flagged until CQM application delivers policy decision for each event

→Data Mover as CEPA facility sends out a regular heartbeat (every 10 secs.) to registered applications, and sends events of interest to the pool via round-robin service if more than one CEPA client is defined

→By default, with CEPA facility running on Data Mover, all file systems except the rootfs are subject to the Event Notifications setup in the cepp.conf file

### **Use following command to exclude specific file systems from CEPA:**

**# server\_mount server\_2 -option nocepp fs5 /fs5**

### **CLI INTERFACE:**

**# server\_cepp server\_2 -**

**server\_2 :**

Error 2100: usage: server\_cepp { <movername> | ALL }

**-service { -start | -stop | -status | -info } | -pool { -info | -stats }**

**\$ server\_cepp server\_x -service -stop | -start | -status | -info | -pool (-info -stats)**

**Note:** Service does not run by default and requires a valid /.etc/cepp.conf file on the Data Mover

## **DART CEPA EVENT POOL & STATES:**

DART's CEPA facility will place events into a CEPA pool to maintain "state", where a default heartbeat every 60 seconds checks each event pool member. A 10-second heartbeat checks each registered application.

### **EVENT POOL STATES:**

OFFLINE—CEPA Pool member is unavailable [do not send events]

ONLINE—CEPA Pool member is available [send events]

ERROR\_CQM\_OTHER—CEPA Pool member is ready, but cannot communicate with registered consumers [do not send events]

ERROR\_CQM\_NOT\_FOUND—CEPA Pool member ready, but cannot communicate with registered applications [do not send events]

UNKNOWN ERROR

### **CEPA APPLICATION STATE:**

ONLINE—CEPA communication to at least one registered application and is "ready to go"

OFFLINE—CEPA is up but cannot communicate to any applications

UN-REGISTERED—applications are present but not yet registered with CEPA

### **SUPPORTED CEPA EVENT TRIGGERS:**

OpenFileNoAccess, OpenFileRead, OpenFileWrite

CreateFile

CreateDir

DeleteFile

DeleteDir

CloseModified, CloseUnmodified

RenameFile

RenameDir

SetAclFile

SetAclDir

## **HOW DOES CEPA WORK?**

a) Celerra CEPA facility creates an Event trigger, which trap the Event & blocks the Client's request stream

b) Event Tuple & Context is formed on DART

**Note:** An event Tuple is: Event Tuple—Event type + associated metadata

c) Event Tuple is sent to the CEPP Pool as a CEPP request (using ONC RPC interface between CEPA Server & DART) for processing by the Server

d) CEPA Client passes request to CQM application for business policy decision using the CEPA API

e) Based on CQM policy rules, a response is returned from CQM application to the CEPA Client and on to the DART CEPA facility.

**Note:** If the response is "success", client request is allowed to proceed via DART CIFS Share. If response is "failure", client request is aborted and DART provides failure or error message back to client.

### **CEPA INSTALLATION OVERVIEW:**

→Install the framework on the Windows client from the [EMC\\_CEE\\_Pack\\_4.0.2.exe](#), which installs the CEPA framework, and runs under the control of the EMC CAVA service

**Note:** If installing on a system where CAVA was previously installed,, uninstall the previous CAVA version, then install the [EMC\\_CEE\\_Pack\\_xxx.exe](#)

→CQM applications must Register with the CEPP framework

→CEPA framework user context should be Domain Users who have EMC Virus Checking & EMC Event Notification Bypass rights, with the System account in the logon properties of the SCM [Service Control Manager] changed to run by the select User account.

→CEPP is disabled by default for root file systems on DART and enabled by default on all others. To disable Event Tuple formation [CEPA event triggers] from a specific file system, use the following mount option:

**"file mount ufs rw /home 2=2 nocepp"**

## **INSTALLING THE CEE FRAMEWORK/CONFIGURING CEPA ON CELERRA & WINDOWS:**

### **OVERVIEW STEPS:**

1. Ensure that a CIFS server, with Shares, is up & running in a typical AD environment [W2k, W2k3, etc]
2. Create cepp.conf file on the Control Station and transfer the file to the Data Mover .etc directory
3. Install the CEPA Client API application on a Windows system [Celerra Event Publishing Agent]
4. Go to the "EMC CAVA" service on the Windows CEPA Client, double-click the EMC CAVA service, then assign the selected Domain User account & password to run the service
5. Using the Celerra Management MMC snapin for the CIFS Server, set the CEPA user rights for EMC Event Notification Bypass and EMC Virus Checking
6. Perform Registry Edit on CEPA Client to define CQM application(s)
7. Reboot CEPA Windows Client
8. Start the CEPA facility on the Data Mover
9. Verify CEPA environment

**1. As a prerequisite, ensure that a CIFS server, with Shares, is up & running for a typical AD environment [W2k, W2k3, etc]**

**2. Create cepp.conf file using one of the following formats as a guide, then transfer the file to the Data Mover .etc directory:****cepp.conf EXAMPLE 1**

```
cifsserver=dbms          →Name of CIFS Server running CEPA client
surveytime=90             →Heartbeat Interval between checks for CEPA Client [default 60 secs, range 5-120]
ft level=1 location=/fs1 size=5   →CMR8 fault tolerance added (0-3); location=buffer file; size=size of buffer file
msrpcuser=domain.ceeuser    →Username for CAVA service
pool name=sepapool \        →pool name on Data Mover for CEPA [only a single pool name supported]
servers=192.1.4.236 \       →IP Address of CEPA Windows Server
preevents=* \              →Max. of 13 defined events, use asterisk to include all
postevents=* \              →Max. of 13 defined events, use asterisk to include all
posterevents=* \            →Max. of 13 defined events, use asterisk to include all [min. 1 event defined on one line]
option=ignore \ or option=denied →If CEPA not available, ignore errors, or return ‘access denied’ if using ‘denied’
reqtimeout=500 \            →Timeout ms when checking/waiting for access [default 1000, range 500-5000ms]
retrytimeout=50             →Timeout ms for retrying access request to CEPA Client [default 250, range 50-5000ms,
value must be equal to or greater than Request Timeout value]
```

**Caution:** Using “option=denied” in the cepp.conf file will result in loss of ReadWrite access to CIFS shares if the CEPA Client is unavailable for any reason! See emc178796 for more information.

**SYNTAX RULES:**

First two lines of the cepp.conf file [cifsserver & surveytime] are Global Variables, and must be defined on separate lines. The remaining entries can be defined on a single text line, or separated by individual lines using a backslash “\” with a space between the \ and the last character on the line [last line of the cepp.conf would not use a slash “\”]. Events can be defined individually [each event separated by the Pipe symbol “|”] or in total by using an asterisk \* to include all thirteen events. Multiple Server IP addresses for CEPA Servers would also be separated by the | symbol.

**cepp.conf EXAMPLE 2:**

```
cifsserver=dbms
surveytime=90
pool name=sepapool servers=192.1.4.236 preevents=* postevents=* posterevents=*
option=ignore reqtimeout=500 retrytimeout=50
```

**# server\_file server\_2 -put cepp.conf cepp.conf**

**3. Install the CEPA Server application on a Windows system [Celerra Event Publishing Agent]****EMC\_CEE\_Pack\_4.0.2.exe**

**Note:** This is a 36MB file shipped with all new systems on a separate kit called “Celerra Event Enabler Software Kit”, version 4.0.2

**4. Go to the “EMC CAVA” service on the Windows CEPA Server [Start>Run: services.msc], double-click the EMC CAVA service, then assign the selected Domain User account & password to run the service [Example: cepa\_user]****5. Using the Celerra Management MMC snapin for the CIFS Server, set the CEPA user rights for EMC Event Notification Bypass and EMC Virus Checking:**

a.) Celerra Management>Data Mover Management>Data Mover Security Settings>User Rights Assignment>EMC Event Notification Bypass

b.) Celerra Management>Data Mover Management>Data Mover Security Settings>User Rights Assignment>EMC Virus Checking

**Note:** This is the same Domain User account used to run the EMC CAVA service

**6. Perform Registry Edit on CEPA Server to define CQM applications**

Start>Run: regedit>HKLM>Software>EMC>Celerra Event Enabler>CEPP>CQM>Configuration>Endpoint

a. For test environment without NTP or Northern CQM software, doubleclick “EndPoint” and type “CeppEndPoint”

b. For NTP or Northern, doubleclick “EndPoint”:

Enter “ntp” or “northern” if CQM application installed locally

Enter “ntp@xxx.xxx;xxx.xxx” or “northern@192.1.4.236;192.1.4.237” to indicate location of Remote CQM application servers

**Note:** Use semicolon to separate multiple Remote CQM application Server IP addresses

c. Reboot CEPA Server

**7. Start the CEPA facility on the Data Mover:**

**# server\_cepp server\_2 -service -start**

**# server\_cepp server\_2 -service -status**

server\_2 : CEPP Started

**8. Verify:**

**# server\_cepp server\_2 -service -info**

server\_2 :

CIFS share name = \\DBMS\CHECK\$

cifs\_server = DBMS

heartbeat\_interval = 90 seconds

|           |                 |                       |             |               |
|-----------|-----------------|-----------------------|-------------|---------------|
| pool_name | server_required | access_checks_ignored | req_timeout | retry_timeout |
| sepapool  | No              | 0                     | 500         | 50            |

# server\_cepp server\_2 -pool -info

server\_2 :

```
pool_name = sepapool
server_required = No
access_checks_ignored = 0
req_timeout = 500ms
retry_timeout = 50ms
pre_events = OpenFileNoAccess,OpenFileRead,OpenFileWrite,CreateFile,CreateDir,Dr
post_events = OpenFileNoAccess,OpenFileRead,OpenFileWrite,CreateFile,CreateDir,r
post_err_events = OpenFileNoAccess,OpenFileRead,OpenFileWrite,CreateFile,Creater
CEPP Servers:
IP = 192.1.4.236, state = ONLINE, vendor = Unknown
```

# server\_cepp server\_2 -pool -stats

server\_2 :

```
pool_name = sepapool
Event Name      Requests    Min(us)     Max(us)     Average(us)
OpenFileNoAccess 587        934        35296       2115
OpenFileRead    264        932        34790       2689
CreateFile      80         1013       9194        1901
CreateDir       24         1001       3482        1699
DeleteFile      128        1359       4231        1682
DeleteDir       35         1029       2847        1505
CloseModified   76         997        3486        1563
CloseUnmodified 856        893        34596       2370
```

Total Requests = 2050

Total Min(us) = 893

Total Max(us) = 35296

Total Average(us) = 2224

#### **TROUBLESHOOTING CEPA:**

1. Use server\_cepp command to check statistics, Cepa Client status, Cepa facility status on Data Mover, etc.
2. Set debug logging and check server\_log messages

\$ .server\_config server\_2 -v "logsys set severity CEPP=LOG\_DEBUG"

\$ .server\_config server\_2 -v "param cepp"

| Name        | Location   | Current    | Default    |
|-------------|------------|------------|------------|
| cepp.Traces | 0x026f63f8 | 0x00000000 | 0x00000000 |

\$ .server\_config server\_2 -v "param cepp Traces=1"

**Note:** With 5.6, you may need to use \$ server\_log server\_x -i option to see trace output in the server log

3. Verify CEPA Service is running and using correct User account with correctly assigned EMC rights
4. Check Windows Event Logs on CEPA Clients
5. Have customer verify CQM applications [NTP Software or Northern Parklife, etc.]
6. Check connectivity to CQM and CEPA Clients
7. Verify that CEPA facility is running on the Data Mover
8. Verify whether “option=ignore” or “option=denied” is set in the cepp.conf file, which if set to the latter may result in CIFS ReadWrite denied if the CEPA Client is not available

#### **PARAMETERS IN 5.6.46:**

# .server\_config server\_2 -v "param cepp"

| Name              | Location   | Current    | Default    |
|-------------------|------------|------------|------------|
| cepp.RPCtype      | 0x03545098 | 0x00000003 | 0x00000003 |
| cepp.Traces       | 0x03544678 | 0x00000000 | 0x00000000 |
| cepp.clientTraces | 0x03545058 | 0x00000000 | 0x00000000 |
| cepp.lowWMFTSize  | 0x035446b8 | 0x00000050 | 0x00000050 |
| cepp.maxReqTO     | 0x02867bdc | 0x00001388 | 0x00001388 |

#### **COMMON ERROR MESSAGES AND CAUSES:**

# server\_cepp server\_2 -service -info →Data Mover CEPA service down

Error 13162905619: server\_2 : The CEPP facility is stopped.

# server\_cepp server\_2 -service -status →Data Mover CEPA service down

server\_2 : CEPP Stopped

# server\_cepp server\_2 -pool -info →Data Mover CEPA service down

Error 13162905619: server\_2 : The CEPP facility is stopped.

```
# server_cepp server_2 -service -start →Syntax problem with cepp.conf file
    Error 13162905607: server_2 : No CEPP server is defined.
# server_cepp server_2 -service -start →Missing cepp.conf file in /etc
    Error 13162905622: server_2 : Cannot open '/etc/cepp.conf'.
# server_cepp server_2 -service -start →Syntax problem in cepp.conf file
    Error 13162905604: server_2 : Invalid server IP: 192.1.4.236preevents=*postevent
s=*posterevents=*option=ignorereqtimeout=500retrytimeout=50ms.
# server_cepp server_2 -pool -info →EMC CAVA service not running
CEPP Servers:
IP = 192.1.4.236, state = OFFLINE, vendor = Unknown
# server_cepp server_2 -service -start
    Error 13162905605: server_2 : Invalid pre-event list: postevents=*.  

# cat cepp.conf
preevents=\ →Missing space before "\", invalid syntax in cepp.conf file
# server_cepp server_2 -service -start
    Error 13162905613: server_2 : No CEPP pre-events / post-events defined.  

# cat cepp.conf
preevents= \
postevents= \
posterevents= \ →need at least one event listed on one of these lines
```

#### Option=ignore vs. Option=denied in cepp.conf file

→When option=ignore is set, server\_cepp –service –info shows that “server\_required” is set to “No”, and in a practical sense means that if the EMC CAVA service or CEPA Client disappears, CIFS Users will still have access to the Shares  
→When option =denied is set, the server\_cepp –service –info shows that “server\_required” is set to “Yes”, and means that if the CEPA Client or CAVA Service hung, CIFS RW access to the Shares would result in Error Deleting (or Copying) File or Folder popup message—see emc178796

#### CEPA REFERENCES:

Using Celerra Event Enabler, 300-006-002 A01  
Using Celerra Event Publishing Agent (CEPA), 300-006-003 A01  
Celerra Event Publishing and Processing for NAS 5.6--Engineering TTT  
Man Pages & Command Reference Guide

#### CEPA COM VENDORS:

NTP Software QFS, Northern Parklife NSS  
**FUTURES:**  
New class of application for Auditing of CIFS events, using Post\_Events, hence is not blocking—Cognac maint release  
CEE Monitor will provide centralized configuration & management of CEPA, along with a GUI display  
CEE Indexing class of application, e.g. Google Search Appliance

#### NAS 5.6 CELERRA DBMS [Database Management System] aka BDB:

With the introduction of NAS 5.6, the Celerra implements a collection of Berkeley databases that are managed under the DBMS environment (sleepycat.com). The various “applications” that use DBMS are: Usermapper, Secmap, VDMs, Celerra Replicator, and other obscure functions. One of the main purposes of this feature was to offload the nameDB databases from the rootfs, which consumed large numbers of inodes, often leading to DU events and the need to manually extend the rootfs to create more inodes.

The two features that are most affected by this new database format are Usermapper & Secmap, which are converted from the nameDB format to DBMS format during the upgrade. The conversion from nameDB to DBMS is automatic and seamless, but it is important to know that the Usermapper and Secmap databases are still managed & repaired strictly from the “application” and not through DBMS [e.g., server\_usermapper and server\_cifssupport -secmap, respectively].

#### DBMS:

→DBMS consists of a series of independent Berkeley databases to support each “application” that uses it  
→Used by Secmap, Usermapper, VDMs, Replication, NAS Events, NFSv4, & other applications  
→Comprised of a set of Base Names & Tables (databases) that are made up of rows with Primary Key values and raw data  
→Consists of Tables, Transaction Manager, Lock Manager, Record Cache Manager, and Secondary Key Tables, which are created in memory only, having Primary & Secondary Key values

#### LIMITATIONS:

--As of 5.6 GA, the ShareDB database for CIFS exports remains and has not yet been converted to DBMS format  
--Databases are not Celerra global, but managed for each Data Mover or VDM separately  
--DBMS does not automatically compensate for disk size constraints or change its default memory cache values [i.e., manual tuning]  
--DBMS maintenance would require the CIFS service or other Application [e.g., Usermapper] to be stopped to close the database, though the server\_dbms -db -check function itself is not enabled for Usermapper or Secmap in the initial Cognac release. Instead, any Usermapper or Secmap database maintenance would occur at the application level, using server\_usermapper or server\_cifssupport -secmap, as examples.

# server\_dbms server\_2 -db -check Usermapper

server\_2 :

Error 13159432299: server\_2 : Invalid checkTable action. See AR111218 for more detail.

# server\_dbms server\_2 -db -check Secmap

server\_2 :

Error 13160874015: server\_2 : Invalid checkTable action.

#### **Pre-NAS 5.6 Roots Inode Consumption & Database Types:**

Server Shares [export.shares] consume 1 inode per Share in ./etc/shares/@COMPNAME or @GLOBAL [ShareDB]

Usermapper entries consume 4 inodes for every User or Group entry [nameDB]

SecMap consumes 2 inodes for every User or Group entry added [nameDB]

#### **NAS 5.6 Roots Inode Consumption & Database Types:**

Server Shares are unchanged and still consume 1 inode for every Share created [ShareDB--convert to DBMS format in 5.6 maint.?]

Usermapper entries have been converted to dbms and inodes reclaimed by rootfs [DBMS]

SecMap entries have been converted to dbms and inodes reclaimed by rootfs {DBMS}

#### **MANAGING DBMS ON CELERRA:**

# server\_dbms server\_2

server\_2 :

Error 2100: usage: server\_dbms { <movername> | ALL }

```
-db
  -list [<db_name>]
  |-delete <db_name>
  |-compact [<db_name>]
  |-check <db_name>
  |-repair <db_name>
  |-stats [<db_name> [-table <name>]] [-reset]
  |-fullbackup -target <pathname>
  |-incrbackup -previous <pathname> -target <pathname>
  |-restore [<db_name>] -source <pathname>
  |-service -stats [transaction | memory | log | lock | mutex] [-reset]
```

# server\_dbms server\_2 -db -list|grep BASE

BASE NAME : Usermapper →Usermapper db [Tables: aliases; usrmapping; idxname; usrmapping; usrgroupmapping; usrmapping; groupmapping; usrmapping]  
BASE NAME : db\_Cmd →Stores in-progress cmds for Repl. or NAS Events [Tables: CmdTable & ConfigurationTable]  
BASE NAME : V4NameSpace →NFSv4 namespace database [Table: pseudofs]  
BASE NAME : Secmap →Secmap mapping database [Table: Mapping]  
BASE NAME : icon\_db →Interconnect records used by DART for Replication [Tables: icon\_tables & rsys\_table]  
BASE NAME : queue\_db →Stores session, task ids, operations, callers [Table: cmdQueue\_table]

**Note:** These are the various dbms databases for Cognac 5.6 release. Each database may have multiple tables.

#### **HOW CAN I DETERMINE ACTUAL SIZE OF TOTAL DBMS DATABASES:**

# server\_dbms server\_2 -db -list|grep Size [server\_dbms only gives sizes for each “BASE NAME” database—add them to see totals or run /nas/sbin/dbms\_backup –totalsize]

**Size : 1826816**

|      |   |         |
|------|---|---------|
| Size | : | 8192    |
| Size | : | 8192    |
| Size | : | 327680  |
| Size | : | 1032192 |
| Size | : | 434176  |
| Size | : | 8192    |
| Size | : | 8192    |

**Size : 24576**

|      |   |       |
|------|---|-------|
| Size | : | 16384 |
| Size | : | 8192  |

**Size : 8192**

|      |   |      |
|------|---|------|
| Size | : | 8192 |
|------|---|------|

**Size : 970752**

|      |   |        |
|------|---|--------|
| Size | : | 970752 |
|------|---|--------|

**Size : 16384**

|      |   |      |
|------|---|------|
| Size | : | 8192 |
| Size | : | 8192 |

**Size : 8192**

Size : 8192

# /nas/sbin/dbms\_backup -totalsize

2854912

# server\_dbms server\_2 -db -stats Usermapper

**Note:** Outputs stats on the various databases, but in rather obscure format that is not user-friendly. The output is segmented by various Modules: LOG; LOCK; TXN; MPOOL; MUTEX

# server\_dbms server\_2 -db -stats &lt;db\_name&gt; -table &lt;table\_name&gt;

# server\_dbms server\_2 -service -stats [appears to output similar info as -db -stats, but is likely a global output]

# server\_dbms server\_2 -db -check Usermapper

server\_2 :

Error 13159432299: server\_2 : Invalid checkTable action.

**Note:** CCMD error string is logged to Server Log, minus the actual CCMD error message number

2008-01-09 14:44:45: USRMAP: 3: Invalid checkTable action.

# nas\_message -i 13159432299

MessageID = 13159432299

BaseID = 107

Severity = ERROR

Component = DART

Facility = USRMAP

Type = STATUS

Brief\_Description = Invalid \${action,8,%s} action.

Full\_Description = Checking and Repairing Usermapper database could not be performed.

Recommended\_Action = Use server\_cifssupport -usrmap to check and repair Usermap

**Note:** Apparently, the server\_dbms db check and repair functionality does not work, and the CCMD Error message that is generated is useless since server\_cifssupport cannot be used to repair a Berkeley database.

# server\_cifssupport server\_2 -usrmap

server\_2 :

Error 2100: usage: server\_cifssupport { &lt;movername&gt; | ALL }

```

    -usrmap
    -list
        {[ -name <name> [-domain <domain_name>]
          | -domain <domain_name>
          | -sid <SID>
          | -uid <user_id>
          | -gid <group_id> }
        { [-user] [-group] }

    | -create
        { -user | -group }
        { -name <name> -domain <domain_name> -sid <SID> }

    | -delete
        { -name <name> [-domain <domain_name>]
          | -sid <SID> }

    | -report
    | -migration

```

# server\_cifssupport server\_2 -usrmap -list -name w2ku30006 -domain w2k

server\_2 : done

USRMAP USER MAPPING TABLE

| SID                                     | UID    | GID    | Name          |
|-----------------------------------------|--------|--------|---------------|
| S-1-5-15-242a3a09-6bc5c62-3f32a78a-4982 | 200001 | 200000 | w2k\w2ku30006 |

# server\_cifssupport server\_2 -usrmap -list -uid 200001

server\_2 : done

USRMAP USER MAPPING TABLE

| SID                                     | UID    | GID    | Name          |
|-----------------------------------------|--------|--------|---------------|
| S-1-5-15-242a3a09-6bc5c62-3f32a78a-4982 | 200001 | 200000 | w2k\w2ku30006 |

# server\_cifssupport server\_2 -usrmap -report [not useful]

server\_2 : done

USRMAP GENERAL INFORMATION

State : Enabled

# server\_cifssupport server\_2 -secmap

server\_2 :

Error 2100: usage: server\_cifssupport { <username> | ALL }

```

-secmap
-list
[ -name <name> -domain <domain_name>
| -domain <domain_name>
| -sid <SID>
| -uid <user_id>
| -gid <group_id> ]
|-create
{ -name <name> -domain <domain_name>
| -sid <SID> }
|-verify
{ -name <name> -domain <domain_name>
| -sid <SID> }
|-update
{ -name <name> -domain <domain_name>
| -sid <SID> }
|-delete
{ -name <name> -domain <domain_name>
| -sid <SID> }
|-export [-file <filename>] |-import -file <filename> |-report |-migration

```

**# server\_cifssupport server\_2 -secmap -report**

server\_2 : done

#### SECMAP GENERAL INFORMATION

Name : server\_2

State : Enabled

Fs : /

Used nodes : 4019

Used blocks : 704512

#### SECMAP MAPPED DOMAIN

Name SID

W2K S-1-5-15-242a3a09-6bc5c62-3f32a78a-ffffffffff

**# server\_cifssupport server\_2 -secmap -list -name w2ku30006 -domain w2k**

server\_2 : done

#### SECMAP USER MAPPING TABLE

| UID    | Origin     | Date of creation        | Name          | SID                                     |
|--------|------------|-------------------------|---------------|-----------------------------------------|
| 200001 | usermapper | Tue Jan 8 16:48:25 2008 | W2K\W2kU30006 | S-1-5-15-242a3a09-6bc5c62-3f32a78a-4982 |

#### DBMS REFERENCES:

##### --Man Pages for server\_dbms and server\_cifssupport:

server\_dbms allows for backup and restore of the databases, as well as statistics about the db environment. Specifically, server\_dbms can perform recovery of db's from media or application corruption, as well as perform checks & fixes of the databases.

#### --Celerra Command Reference Manual

#### NAS EVENTS FOR DBMS:

**# nas\_event -l -c DART -f DBMS** → Facility ID = 122

DART(1)

|--> DBMS(122)

BaseID Severity Brief\_Description

1 ALERT(1) The environment database of the VDM might be corrupted and a recovery procedure must take place

2 CRITICAL(2) Only \${freeblocks,3,%llu} free blocks in the root file system (fsid \${fsid,2,%u}) of the VDM \${vdm,8,%s}.

3 ALERT(1) The root file system (fsid \${fsid,2,%u}) of the VDM \${vdm,8,%s} is full. There are only \${freeblocks,3,%llu} free blocks.

#### HOW TO DETERMINE IF A MIGRATION HAS OCCURRED?

**Note:** A migration occurs when upgrading from any version prior to NAS 5.6 that contains SecMap or Internal Database entries

**# server\_cifssupport server\_2 -usrmap -migration**

server\_2 : done

SECMAP MIGRATION INFORMATION → Supposed to show Usermapper migration status, but title is mislabeled

Start : Wed Jan 9 14:13:32 2008 → Time is in GMT, not local

End : Wed Jan 9 14:13:32 2008

Status : Database has been successfully migrated.

#### Server Files on Rootfs After NAS Upgrade:

```
# pwd  
/nas/quota/slot_2/.etc  
# cat secmap.migration  
1199888015  
1199888022  
# cd usrmapper;ls  
usrmapper.migration usrmap.settings  
# cat usrmap.settings  
Identity=128.221.252.2,128.221.253.2 →This is Unix epoch time in seconds since 1/1/1970, which translates to Wed, 9 Jan 2008 14:13:32 UTC  
1199888023  
# cat usrmap.settings
```

Identity=128.221.252.2,128.221.253.2 →These equate to Usermapper running on Primary & Secondary Internal networks Server\_2  
Started=on

**Server Log:** DBMS migration begins after CIFS Service is started on Server reboot during NAS Upgrade

**Usermapper Conversion Entries:**

```
2008-01-09 09:13:32: USRMAP: 6: Starting usermapper service  
2008-01-09 09:13:32: USRMAP: 6: 8: The migration of the Usermapper database to the Celerra Network Server version 5.6 format has started.  
2008-01-09 09:13:32: USRMAP: 6: The migration of the Usermapper table "aliases" has started.  
2008-01-09 09:13:32: USRMAP: 6: The migration of the Usermapper table "aliases" was successful. "0" records have been migrated.  
2008-01-09 09:13:32: USRMAP: 6: The migration of the Usermapper table "groupmapnamesid" has started.  
2008-01-09 09:13:32: USRMAP: 6: The migration of the Usermapper table "groupmapnamesid" was successful. "24" records have been migrated.  
2008-01-09 09:13:32: USRMAP: 6: The migration of the Usermapper table "idxname" has started.  
2008-01-09 09:13:35: USRMAP: 6: The migration of the Usermapper table "idxname" was successful. "4019" records have been migrated.  
2008-01-09 09:13:35: USRMAP: 6: The migration of the Usermapper table "usrgrpmapnamesid" has started.  
2008-01-09 09:13:38: USRMAP: 6: The migration of the Usermapper table "usrgrpmapnamesid" was successful. "3995" records have been migrated.  
2008-01-09 09:13:38: USRMAP: 6: The migration of the Usermapper table "usrmapc" has started.  
2008-01-09 09:13:38: USRMAP: 6: The migration of the Usermapper table "usrmapc" was successful. "3" records have been migrated.  
2008-01-09 09:13:38: USRMAP: 6: The migration of the Usermapper table "usrmapgrpc" has started.  
2008-01-09 09:13:38: USRMAP: 6: The migration of the Usermapper table "usrmapgrpc" was successful. "24" records have been migrated.  
2008-01-09 09:13:38: USRMAP: 6: The migration of the Usermapper table "usrmapusrc" has started.  
2008-01-09 09:13:43: USRMAP: 6: 1: The migration of the Usermapper database to the Celerra Network Server version 5.6 format has started.  
2008-01-09 09:13:43: USRMAP: 6: 9: The Usermapper database has been successfully migrated.  
2008-01-09 09:13:43: USRMAP: 6: 2: Usermapper service has been enabled.
```

**Secmap Conversion Entries:**

```
2008-01-09 09:13:35: SECMAP: 6: 1: The migration of the secmap database to the Celerra Network Server version 5.6 format has started.  
2008-01-09 09:13:35: SECMAP: 6: The migration of the secmap table "S-1-5-15-242a3a09-6bc5c62-3f32a78a-ffffffffff" has started.  
2008-01-09 09:13:39: SECMAP: 6: The migration of the secmap table "S-1-5-15-242a3a09-6bc5c62-3f32a78a-ffffffffff" has been successful. "4019" records have been migrated.  
2008-01-09 09:13:42: SECMAP: 6: 2: The secmap database has been successfully migrated.
```

2008-01-09 09:13:42: SECMAP: 6: server\_2 database started

**TROUBLESHOOTING DBMS:**

Server Log

CCMD Error messages

```
# .server_config server_2 -v "bdb ls" →Lists the various dbms DB Names and Tables  
# .server_config server_2 -v "bdb hide" →Rehide the dbms directory  
# .server_config server_2 -v "bdb show" →unhides the dbms database directory as /.etc/.dbms  
# ls -la  
d----- 5 root bin 1024 Mar 3 2008 .dbms  
# ls -la *  
-rw-rw---- 2 root bin 8192 Jan 9 09:14 db.db_Cmd  
-rw-rw---- 2 root bin 8192 Jan 9 09:14 db.icon_db  
-rw-rw---- 2 root bin 8192 Jan 9 09:14 db.queue_db
```

```
-rw-rw--- 2 root bin 8192 Jan 9 09:14 db.Secmap
-rw-rw--- 2 root bin 8192 Jan 30 16:36 db.Usermapper
-rw-rw--- 2 root bin 8192 Feb 26 13:32 db.V4NameSpace
-rw-rw--- 2 root bin 8192 Jan 9 09:13 sk.db_Cmd.CmdTable.parentTask
-rw-rw--- 2 root bin 364544 Feb 7 22:57 sk.Secmap.Mapping.xid
-rw-rw--- 2 root bin 8192 Jan 9 09:13 tb.db_Cmd.CmdTable
-rw-rw--- 2 root bin 8192 Jan 9 09:14 tb.db_Cmd.ConfigurationTable
-rw-rw--- 2 root bin 8192 Feb 26 13:32 tb.icon_db.icon_table
-rw-rw--- 2 root bin 8192 Jan 9 09:15 tb.icon_db.rsys_table
-rw-rw--- 2 root bin 8192 Jan 9 09:13 tb.queue_db.cmdQueue_table
-rw-rw--- 2 root bin 614400 Feb 8 07:51 tb.Secmap.Mapping
----- 2 root bin 8192 Jan 30 16:36 tb.Usermapper.aliases
----- 2 root bin 8192 Feb 7 23:00 tb.Usermapper.groupmapnamesid
----- 2 root bin 385024 Feb 7 22:57 tb.Usermapper.idxname
----- 2 root bin 466944 Feb 7 22:57 tb.Usermapper.usrgrpmapnamesid
----- 2 root bin 8192 Feb 7 22:57 tb.Usermapper.usrmapc
----- 2 root bin 8192 Feb 7 23:00 tb.Usermapper.usrmapgrpc
----- 2 root bin 765952 Feb 7 22:57 tb.Usermapper.usrmapusrc
-rw-rw--- 2 root bin 8192 Feb 26 13:32 tb.V4NameSpace.pseudofs
```

# tar -zcpf /home/nasadmin/dbms.tar.gz .dbms → Tar & zip the .dbms directory for forwarding to Eng.

### DBMS PARAMS:

# server\_dbms server\_2 -service -info

server\_2 : done

| PARAM          | ACTUAL  | DEFAULT IS |
|----------------|---------|------------|
| tx_max         | 20      | 20         |
| cachesize      | 8388608 | 8388608    |
| lg_bsize       | 32768   | 32768      |
| lg_max         | 5242880 | 5242880    |
| lg_regionmax   | 65536   | 65536      |
| lk_max_lockers | 1000    | 1000       |
| lk_max_locks   | 1000    | 1000       |
| lk_max_objects | 1000    | 1000       |

# .server\_config server\_2 -v "param dbms"

| Name                        | Location   | Current    | Default                                                |
|-----------------------------|------------|------------|--------------------------------------------------------|
| ---                         | -----      | -----      | -----                                                  |
| dbms.checksum               | 0x016444e0 | 0x00000001 | 0x00000001                                             |
| dbmsckptInterval            | 0x01644b68 | 0x0000003c | 0x0000003c                                             |
| dbms.deadlockDetectInterval | 0x01644b6c | 0x0000000a | 0x0000000a                                             |
| dbms.deadlockStrategy       | 0x01644b78 | 0x00000004 | 0x00000004                                             |
| dbms.highWaterMark          | 0x01644b70 | 0x0000000a | 0x0000000a                                             |
| dbms.lowWaterMark           | 0x01644b74 | 0x00000014 | 0x00000014                                             |
| dbms.maxUnusedLogs          | 0x01644b7c | 0x0000000a | 0x0000000a                                             |
| dbms.param.dm.cache         | 0x02677cc4 | 0x00800000 | 0x00800000                                             |
| dbms.param.dm.lgbsize       | 0x02677cc8 | 0x00008000 | 0x00008000 →Max trans. log buffer size BDB, DM (32kb)  |
| dbms.param.dm.lgmax         | 0x02677ccc | 0x00500000 | 0x00500000                                             |
| dbms.param.dm.lgregionmax   | 0x02677cd0 | 0x00010000 | 0x00010000 →Max log region size BDB, DM (65kb)         |
| dbms.param.dm.maxlockers    | 0x02677cd4 | 0x000003e8 | 0x000003e8 →Max number locking entities BDB, DM (1000) |
| dbms.param.dm.maxlocks      | 0x02677cd8 | 0x000003e8 | 0x000003e8 →Max number locks BDB, DM (1000)            |
| dbms.param.dm.maxobjects    | 0x02677cdc | 0x000003e8 | 0x000003e8 →Max number locked objects BDB, DM (1000)   |
| dbms.param.dm.txmax         | 0x02677cc0 | 0x00000014 | 0x00000014 →Max concurrent transactions, DM (20)       |
| dbms.param.vdm.cache        | 0x02677ce4 | 0x00100000 | 0x00100000                                             |
| dbms.param.vdm.lgbsize      | 0x02677ce8 | 0x00001fa4 | 0x00001fa4 →Max trans. log buffer size BDB, VDM (8kb)  |
| dbms.param.vdm.lgmax        | 0x02677cec | 0x00500000 | 0x00500000                                             |
| dbms.param.vdm.lgregionmax  | 0x02677cf0 | 0x00010000 | 0x00010000 →Max log region size BDB, VDM (65kb)        |
| dbms.param.vdm.maxlockers   | 0x02677cf4 | 0x00000064 | 0x00000064 →Max number locking entities BDB, VDM (100) |
| dbms.param.vdm.maxlocks     | 0x02677cf8 | 0x00000064 | 0x00000064 →Max number locks BDB, VDM (100)            |
| dbms.param.vdm.maxobjects   | 0x02677cfc | 0x00000064 | 0x00000064 →Max number locked objects BDB, VDM (100)   |
| dbms.param.vdm.txmax        | 0x02677ce0 | 0x00000014 | 0x00000014 →Max concurrent transactions, VDM (20)      |

### PARAMETER TUNING:

Maximum number of concurrent transactions for CIFS or VDM Server are 20 decimal (14 hex):

```
# .server_config server_2 -v "param dbms param.vdm.txmax" "param dbms param.dm.txmax"
```

dbms.param.vdm.txmax INT 0x02677ce0 20 20 (0,4294967295) FALSE REBOOT 'NA'

## **BACKUP & RESTORE OF DBMS DATABASES:**

**Note:** See AR114590 for issues with manual restores using server\_dbms

### **I. AUTOMATIC BACKUP OF DATABASES:**

→The nasdb\_backup keeps up to three dbms backups in **/nas/var/backup/**

```
-rw-r--r-- 1 nasadmin nasadmin 241856 Jan 10 07:01 _dbms_backup.01.tar.gz
```

```
-rw-r--r-- 1 nasadmin nasadmin 241856 Jan 10 08:01 _dbms_backup.02.tar.gz
```

```
-rw-r--r-- 1 nasadmin nasadmin 241856 Jan 10 08:01 _dbms_backup.OK.tar.gz
```

→Backups are automatically initiated by /nas/sbin/nasdb\_backup script, which is run as Cron job, but calls **/nas/sbin/dbms\_backup**

```
-rwxr-x--x 1 nasadmin nasadmin 13515 Dec 20 12:07 dbms_backup
```

```
# /nas/sbin/dbms_backup
```

Usage: dbms\_backup

```
-totalsize
```

```
| -backup -file <filename> [ -dir <download_directory> ] | -help
```

### **CONTENTS OF DBMS BACKUP:**

```
# tar -zvxf celerra/backup/dbms_bak.1.tar.gz →Unzips and untars to /DBMS_BACKUP.2008-01-10_09-28/server_2/
```

```
# tar -zvxf celerra/backup/_dbms_backup.02.tar.gz →Unzips & untars contents to /DBMS_BACKUP.2008-01-30_15-32
```

DBMS\_BACKUP.2008-01-30\_15-32/

DBMS\_BACKUP.2008-01-30\_15-32/dbms\_download.2008-01-30\_15-32.log

### **DBMS\_BACKUP.2008-01-30\_15-32/server\_2/**

DBMS\_BACKUP.2008-01-30\_15-32/server\_2/db.queue\_db

DBMS\_BACKUP.2008-01-30\_15-32/server\_2/tb.queue\_db.cmdQueue\_table

DBMS\_BACKUP.2008-01-30\_15-32/server\_2/db.Usermapper

DBMS\_BACKUP.2008-01-30\_15-32/server\_2/tb.Usermapper.groupmapnamesid

DBMS\_BACKUP.2008-01-30\_15-32/server\_2/tb.Usermapper.usrmapgrpc

DBMS\_BACKUP.2008-01-30\_15-32/server\_2/tb.Usermapper.usrgrpmapnamesid

DBMS\_BACKUP.2008-01-30\_15-32/server\_2/tb.Usermapper.usrmapusrc

DBMS\_BACKUP.2008-01-30\_15-32/server\_2/tb.Usermapper.idxname

DBMS\_BACKUP.2008-01-30\_15-32/server\_2/tb.Usermapper.usrmapc

DBMS\_BACKUP.2008-01-30\_15-32/server\_2/tb.Usermapper.aliases

DBMS\_BACKUP.2008-01-30\_15-32/server\_2/db.icon\_db

DBMS\_BACKUP.2008-01-30\_15-32/server\_2/tb.icon\_db.rsys\_table

DBMS\_BACKUP.2008-01-30\_15-32/server\_2/tb.icon\_db.icon\_table

DBMS\_BACKUP.2008-01-30\_15-32/server\_2/db.Secmap

DBMS\_BACKUP.2008-01-30\_15-32/server\_2/sk.Secmap.Mapping.xid

DBMS\_BACKUP.2008-01-30\_15-32/server\_2/tb.Secmap.Mapping

DBMS\_BACKUP.2008-01-30\_15-32/server\_2/db.db\_Cmd

DBMS\_BACKUP.2008-01-30\_15-32/server\_2/tb.db\_Cmd.ConfigurationTable

DBMS\_BACKUP.2008-01-30\_15-32/server\_2/sk.db\_Cmd.CmdTable.parentTask

DBMS\_BACKUP.2008-01-30\_15-32/server\_2/tb.db\_Cmd.CmdTable

### **II. CREATING MANUAL DBMS DATABASE BACKUPS:**

→Creates tar.gz backup and writes to default location /celerra/backup unless a specific path is requested

#### **1. USING dbms\_backup to create dbms backup:**

```
# mkdir /celerra/backup/restore
```

```
# cd /celerra/backup/restore
```

```
# /nas/sbin/dbms_backup -backup -file /celerra/backup/restore/dbms_backup.tar.gz
```

Backup server\_2 ... Done.

Archive and compress ... Done.

Remove the temporary directory ... Done.

```
# ls -la
```

```
-rw-r--r-- 1 root root 340714 Jan 30 15:57 dbms_backup.tar.gz
```

**Note:** The backup is written to the path indicated. If no path specified, the backup is written to /celerra/backup as a tar.gz file.

#### **2. Using server\_dbms command to create manual backup:**

```
# server_dbms server_2 -db -fullbackup -target /celerra/backup/restore/db
```

server\_2 : done

```
# server_dbms server_2 -db -fullbackup -target /celerra/backup/restore
```

server\_2 :

Error 2237: server\_2 : Execution failed: system error Directory not empty (errn)

**Note:** This method of creating a backup does not use tar.gz and requires an empty directory in order to write the files. Command will create a directory if not created ahead of time.

# cd /celerra/backup/restore/db

# ls -la

```
-rw-r--r-- 1 root root 8192 Jan 30 16:08 db.db_Cmd
-rw-r--r-- 1 root root 8192 Jan 30 16:08 db.icon_db
-rw-r--r-- 1 root root 8192 Jan 30 16:08 db.queue_db
-rw-r--r-- 1 root root 8192 Jan 30 16:08 db.Secmap
-rw-r--r-- 1 root root 8192 Jan 30 16:08 db.Usermapper
-rw-r--r-- 1 root root 8192 Jan 30 16:08 sk.db_Cmd.CmdTable.parentTask
-rw-r--r-- 1 root root 364544 Jan 30 16:08 sk.Secmap.Mapping.xid
-rw-r--r-- 1 root root 8192 Jan 30 16:08 tb.db_Cmd.CmdTable
-rw-r--r-- 1 root root 8192 Jan 30 16:08 tb.db_Cmd.ConfigurationTable
-rw-r--r-- 1 root root 8192 Jan 30 16:08 tb.icon_db.icon_table
-rw-r--r-- 1 root root 8192 Jan 30 16:08 tb.icon_db.rsys_table
-rw-r--r-- 1 root root 8192 Jan 30 16:08 tb.queue_db.cmdQueue_table
-rw-r--r-- 1 root root 614400 Jan 30 16:08 tb.Secmap.Mapping
-rw-r--r-- 1 root root 8192 Jan 30 16:08 tb.Usermapper.aliases
-rw-r--r-- 1 root root 8192 Jan 30 16:08 tb.Usermapper.groupmapnamesid
-rw-r--r-- 1 root root 385024 Jan 30 16:08 tb.Usermapper.idxname
-rw-r--r-- 1 root root 466944 Jan 30 16:08 tb.Usermapper.usrgrpmapnamesid
-rw-r--r-- 1 root root 8192 Jan 30 16:08 tb.Usermapper.usrmapc
-rw-r--r-- 1 root root 8192 Jan 30 16:08 tb.Usermapper.usrmapgrpc
-rw-r--r-- 1 root root 757760 Jan 30 16:08 tb.Usermapper.usrmapsrc
```

### INCREMENTAL BACKUPS?

# server\_dbms server\_2 -db -incrbackup -previous /celerra/backup/restore/db –target /celerra/backup/restore/db/incr  
server\_2 : done

**Note:** Incremental backup creates directory, but seems to act the same as a full backup

### RESTORING DBMS DATABASES:

#### I. Catastrophic Recovery

- Entire environment needs to be restored
- Quiesce all applications first
- DM rootfs must be consistent
- Database operations for backups and restores do not appear to be logged in sys\_log or server\_log—can set “logsys set severity DBMS=LOG\_DBG3” but the server\_log entries do not explain what is taking place

# server\_dbms server\_2 -db -restore –source /celerra/backup/dbms\_bak.1.tar.gz

#### II. Application Restore

- Application must first be stopped
- Delete existing dbms Usermapper database, or overwrite the existing db via the restore process

# server\_usermapper server\_2 –disable

# server\_dbms server\_2 -db –delete Usermapper

# server\_dbms server\_2 -db -restore Usermapper –source /celerra/backup/restore/dbms\_backup.tar.gz

server\_2 :

Error 2237: server\_2 : Execution failed: system error Not a directory (errno 20)

**Note:** Cannot do individual DBMS database restore from a tar.gz backup file!

→Either unzip and untar the dbms\_backup.tar.gz file, or find a directory that contains the necessary files from a server\_dbms -db –fullbackup in order to perform the Usermapper database restore

#### Example:

# server\_dbms server\_2 -db -fullbackup -target /celerra/backup/restore/db

#### Untar & Restore Example:

# tar -zxf dbms\_backup.tar.gz

# cd /DBMS\_BACKUP.2008-01-30\_15-57/server\_2;ls -la

```
-rw-r--r-- 1 root root 8192 Jan 30 15:57 db.db_Cmd
-rw-r--r-- 1 root root 8192 Jan 30 15:57 db.icon_db
-rw-r--r-- 1 root root 8192 Jan 30 15:57 db.queue_db
-rw-r--r-- 1 root root 8192 Jan 30 15:57 db.Secmap
-rw-r--r-- 1 root root 8192 Jan 30 15:57 db.Usermapper
-rw-r--r-- 1 root root 8192 Jan 30 15:57 sk.db_Cmd.CmdTable.parentTask
-rw-r--r-- 1 root root 364544 Jan 30 15:57 sk.Secmap.Mapping.xid
-rw-r--r-- 1 root root 8192 Jan 30 15:57 tb.db_Cmd.CmdTable
```

```
-rw-r--r-- 1 root root 8192 Jan 30 15:57 tb.db_Cmd.ConfigurationTable  
-rw-r--r-- 1 root root 8192 Jan 30 15:57 tb.icon_db.icon_table  
-rw-r--r-- 1 root root 8192 Jan 30 15:57 tb.icon_db.rsys_table  
-rw-r--r-- 1 root root 8192 Jan 30 15:57 tb.queue_db.cmdQueue_table  
-rw-r--r-- 1 root root 614400 Jan 30 15:57 tb.Secmap.Mapping  
-rw-r--r-- 1 root root 8192 Jan 30 15:57 tb.Usermapper.aliases  
-rw-r--r-- 1 root root 8192 Jan 30 15:57 tb.Usermapper.groupmapnamesid  
-rw-r--r-- 1 root root 385024 Jan 30 15:57 tb.Usermapper.idxname  
-rw-r--r-- 1 root root 466944 Jan 30 15:57 tb.Usermapper.usrgrpmapnamesid  
-rw-r--r-- 1 root root 8192 Jan 30 15:57 tb.Usermapper.usrmapc  
-rw-r--r-- 1 root root 8192 Jan 30 15:57 tb.Usermapper.usrmapgrpc  
-rw-r--r-- 1 root root 757760 Jan 30 15:57 tb.Usermapper.usrmapusrc
```

# pwd

/celerra/backup/restore/DBMS\_BACKUP.2008-01-30\_15-57/server\_2

# server\_dbms server\_2 -db -restore Usermapper -source /celerra/backup/restore/DBMS\_BACKUP.2008-01-30\_15-57/server\_2

server\_2 : done

#### **Restart Usermapper Application & Verify Records:**

# server\_usermapper server\_2 -enable

server\_2 : done

# server\_usermapper server\_2 -Export -user users

server\_2 : done

# server\_usermapper server\_2 -Export -group groups

server\_2 : done

# wc -l users groups

5196 users

30 groups

**Note:** Same number of Users and Groups verified before and after “Restore” testing

#### **III. Application Table Restore**

→Application must first be stopped

# server\_usermapper server\_2 -disable

# server\_dbms server\_2 -db -restore Usermapper -table usrmapusrc -source /home/nasadmin/usrmap/restore

### **RESTORING USERMAPPER & SECMAP DBMS DATABASES WITHOUT DELETING DBMS DATABASE FIRST:**

#### **1. Disable the Secmap or Usermapper "application" so as to "close" the dbms database in question**

# server\_usermapper server\_2 -disable (Closes the dbms Usermapper database)

# server\_setup server\_x -P cifs -o stop (Closes the Secmap database)

#### **2. Locate an acceptable dbms backup in /celerra/backup:**

# ls -la /celerra/backup/\_dbms\*

-rw-r--r-- 1 nasadmin nasadmin 1248366 Jan 30 21:01 /celerra/backup/\_dbms\_backup.01.tar.gz

-rw-r--r-- 1 nasadmin nasadmin 1248371 Jan 30 20:01 /celerra/backup/\_dbms\_backup.02.tar.gz

-rw-r--r-- 1 nasadmin nasadmin 340742 Jan 30 16:01 /celerra/backup/\_dbms\_backup.OK.tar.gz

#### **3. Unzip and untar the selected dbms backup file:**

# tar -zvxf /celerra/backup/\_dbms\_backup.02.tar.gz

#### **4. Navigate to the directory where the dbms database files were extracted:**

# cd /celerra/backup/DBMS\_BACKUP.2008-01-30\_15-32/server\_2

#### **5. Perform the Restore of the Usermapper or Secmap database:**

**Note:** Can only perform separate database restores from extracted files, not directly from tar.gz backup archives

# server\_dbms server\_2 -db -restore Usermapper -source /celerra/backup/DBMS\_BACKUP.2008-01-30\_15-32/server\_2

# server\_dbms server\_2 -db -restore Secmap -source /celerra/backup/DBMS\_BACKUP.2008-01-30\_15-32/server\_2

#### **6. Restart the Usermapper or Secmap databases:**

# server\_usermapper server\_2 -enable (Opens the dbms Usermapper database)

# server\_setup server\_x -P cifs -o start (Opens the Secmap database)

#### **7. Verify:**

# server\_usermapper server\_2 -Export -user <filename> | -group <filename>

# server\_cifssupport server\_2 -secmap -export -file <filename>

# server\_dbms server\_2 -db -list -->Check "BASE NAME" database information for Usermapper and Secmap

### **EXAMPLE OF DELETING & RESTORING USERMAPPER DBMS DATABASE:**

# server\_dbms server\_2 -db -delete Usermapper

server\_2 :

Error 13161332737: server\_2 : The Berkeley DB function deleteTable has return with the error Permission denied .

**Note:** Above error really means that the Usermapper application must be stopped before the db can be deleted

### **1. Perform manual backup of DBMS or verify existence of good backup in /celerra/backup**

# ls -la /celerra/backup/\_dbms\*

```
-rw-r--r-- 1 nasadmin nasadmin 333185 Jan 10 13:01 /celerra/backup/_dbms_backup.01.tar.gz  
-rw-r--r-- 1 nasadmin nasadmin 166805 Jan 10 14:01 /celerra/backup/_dbms_backup.02.tar.gz  
-rw-r--r-- 1 nasadmin nasadmin 166805 Jan 10 14:01 /celerra/backup/_dbms_backup.OK.tar.gz
```

**Note:** If restoring individual databases from tar.gz backup, you must first unzip and untar the backup file and then restore from the untarred location of the database and table files

# server\_dbms server\_2 -db -fullbackup -target /home/nasadmin/usrmap/restore

### **2. Stop Usermapper Service**

# server\_usermapper server\_2 -d

server\_2 : done

### **3. Verify current Usermapper dbms size and number of User and Group entries**

# server\_dbms server\_2 -db -list

```
BASE NAME      : Usermapper  
Size          : 1642496
```

# server\_usermapper server\_2 -Export -user users | -group groups

# wc -l users groups

5194 users

24 groups

### **4. Delete the Usermapper database**

# server\_dbms server\_2 -db -delete Usermapper

server\_2 : done

# server\_dbms server\_2 -db -list |grep -i usermapper

# server\_usermapper server\_2 -Export -user users [output will be 0 byte file, db is empty]

### **5. Restore Usermapper database from good backup location**

# server\_dbms server\_2 -db -restore Usermapper -source /home/nasadmin/usrmap/restore

server\_2 : done

### **6. Restart Usermapper Application & Verify**

# server\_usermapper server\_2 -enable

server\_2 : done

# server\_dbms server\_2 -db -list |more

```
BASE NAME      : Usermapper
```

Size : 1642496

# server\_usermapper server\_2 -Export -user users.bak

server\_2 : done

# wc -l users.bak

5194 users.bak

**Note:** Cannot perform the restore to individual dbms databases with the Application started, as the following shows when Usermapper is Enabled. Corrective action would be to stop Usermapper and then perform the restore.

# server\_usermapper server\_2

server\_2 : Usrmapper service: Enabled

Service Class: Primary

# server\_dbms server\_2 -db -restore Usermapper -source /home/nasadmin/usrmap/restore

server\_2 :

Error 2237: server\_2 : Execution failed: HTTP error 409 Conflict

### **CLOSING DBMS DATABASES:**

# server\_setup server\_2 -P cifs -o stop

# server\_dbms server\_2 -db -list

```
BASE NAME      : Secmap
```

Version : 1

State : Closed

Backup enabled : YES

Comment : CIFS Secure mapping database.

**Note:** Stopping CIFS closes only the Secmap database. Use server\_usermapper server\_2 -disable to close the Usermapper database. It is not yet known how the other dbms databases are closed.

### **DBMS MAINTENANCE:**

# server\_dbms server\_2 -db -compact Usermapper

server\_2 : done

**Note:** Compact command can reduce overall size of database, but ironically, seems to use more rootfs space & inodes

**Before Compact:**

# server\_df server\_2 /

server\_2 :

| Filesystem | kbytes  | used  | avail   | capacity | Mounted on |
|------------|---------|-------|---------|----------|------------|
| root_fs_2  | 3842416 | 11312 | 3831104 | 0%       | /          |

# server\_df server\_2 -i /

server\_2 :

| Filesystem | inodes  | used | avail   | capacity | Mounted on |
|------------|---------|------|---------|----------|------------|
| root_fs_2  | 3841022 | 1522 | 3839500 | 0%       | /          |

# server\_dbms server\_2 -db -list |head -7

Size : 1900544

**After Compact:**

# server\_dbms server\_2 -db -compact Usermapper

server\_2 : done

# server\_df server\_2 /

server\_2 :

| Filesystem | kbytes  | used  | avail   | capacity | Mounted on |
|------------|---------|-------|---------|----------|------------|
| root_fs_2  | 3842416 | 16184 | 3826232 | 0%       | /          |

# server\_df server\_2 -i /

server\_2 :

| Filesystem | inodes  | used | avail   | capacity | Mounted on |
|------------|---------|------|---------|----------|------------|
| root_fs_2  | 3841022 | 1523 | 3839499 | 0%       | /          |

# server\_dbms server\_2 -db -list |head -7

Size : 1642496

**5.6 UPGRADE CONSIDERATIONS:**

Peform nameDB backup of Secmap and Usermapper

**NAS 5.6 NTXMAP FEATURE:**

→Feature allows Users that may have different names in Unix and Windows, through the use of explicit unix-to-windows or windows-to-unix name mappings (using an ntxmap.conf file), to access file systems from either protocol while maintaining proper access

→This feature is not designed for environments where explicit mappings are required for more than 1000 users, as the ntxmap.conf file must be sequentially parsed for every mapping request, which could lead to authentication latency issues, potentially impacting CIFS user login and work sessions

→A populated ntxmap.conf file, with the explicit cross-protocol name mappings, is placed in the /.etc directory of the Data Mover

→The NTXMap feature provides mappings for Windows credential mappings and Unix to Windows mappings

**Windows credential mapping mechanism:**

CIFS User connects to the Data Mover, SIDs and Names are looked up via queries to the DCs, the ntxmap.conf file is then parsed for a matching Windows user name, and Domain if defined. If a match occurs, the Windows name is then mapped to the corresponding Unix name, and a UID is assigned via the normal mapping mechanisms

**Unix to Windows mapping mechanism:**

Unix user connects to Data Mover, Unix name is parsed in ntxmap.conf file, and if matched, then maps the corresponding windows user and domain names, and SIDs are obtained through the normal mapping mechanism for a Windows user

**NTXMAP.CONF FILE:**

```
# cat ntxmap.conf
w2k:w2ku40000=:unix40000
:W2KU40006=:unix40006
*:w2ku40007=:UNIX40007
w2k:cifs=20user=:unix40008
w2k:==00fcser=:unix40009
```

**Defining the syntax for the ntxmap.conf file:**

**domain:cifs\_user=:unix\_user**

w2k = Windows domain short name, or “netbios” name, not case-sensitive, and this field can be blank or can use a wildcard \* to indicate match the name in any Domain

w2ku40000 = Windows user name, not case-sensitive, use hex ascii [=20 →space] to represent a space in the user name, use hex unicode values to represent foreign characters [==00fc ü →umlaut]

==: Colons are required on each side of the = sign, and the = sign represents bidirectional mappings, which mean that a user can be mapped from either CIFS to Unix or from Unix to CIFS

Unix40000 = Unix name, which is case sensitive

**CHECKING NTXMAP CONFIGURATION FILE:**

```
# server_checkup server_2 -test CIFS -subtest ntxmap
```

Ntxmap : Checking the ntxmap configuration file..... Pass

```
# server_checkup server_2 -test CIFS -subtest ntxmap
```

-----Checks-----

Component CIFS :

Ntxmap : Checking the ntxmap configuration file..... Fail

-----CIFS : Ntxmap Warnings-----

Warning 17457807362: server\_2 : There is a syntax error in the file /etc/ntxmap.conf at line 4. Some users of the Data Mover cannot access some of the resources.

--> Fix the syntax error of the configuration file of the Data Mover. Use the following steps: 1) Get the config. file using server\_file -get command. 2) Edit the config. file using any text editor. 3) Restore the configuration file on the Celerra using server\_file -put

```
# cat ntxmap.conf
```

```
:w2ku40003=:unix40003
w2k:w2ku40004=:unix40004
W2K:w2ku40005=:unix40005
w2k::=:unix40006
```

→Line 4 missing username

**VERIFYING SECMAP ENTRIES FOR NTXMAP:**

```
# server_cifssupport server_2 -seemap -list |grep ntxmap
UID Origin Date of creation Name SID
36970 ntxmap Sat Feb 2 05:17:33 2008 W2K\W2ku40000 S-1-5-15-242a3a09-6bc5c62-3f32a78a-2b65
36973 ntxmap Fri Feb 8 12:51:08 2008 W2K\W2ku40003 S-1-5-15-242a3a09-6bc5c62-3f32a78a-2b68
36971 ntxmap Fri Feb 8 03:39:00 2008 W2K\w2ku40005 S-1-5-15-242a3a09-6bc5c62-3f32a78a-2b6a
```

**USE CASE 1:****CIFS User accessing Share for first time [No ntxmap.conf file]:**

Usermapper=W2kU40004

Secmap=W2k\W2kU40004 from “Origin” usermapper

Ntcred=(mapped=W2kU40004.W2K) →No ntxmap file

```
# .server_config server_2 -v "usrmap dump"
```

```
1205772648: USRMAP: 6: S-1-5-15-242a3a09-6bc5c62-3f32a78a-2b69:*:32768:32768:user W2kU40004 from domain
W2K:/usr/S-1-5-15-242a3a09-6bc5c62-3f32a78a-2b69:/bin/sh
```

```
# server_cifssupport server_2 -seemap -list -name w2ku40004 -domain w2k
```

| SECMAP USER MAPPING TABLE |            |                          |               |                                         |
|---------------------------|------------|--------------------------|---------------|-----------------------------------------|
| UID                       | Origin     | Date of creation         | Name          | SID                                     |
| 32768                     | usermapper | Mon Mar 17 16:50:32 2008 | W2K\W2ku40004 | S-1-5-15-242a3a09-6bc5c62-3f32a78a-2b69 |

```
# .server_config server_2 -v "ntcred if=cge0 user=w2ku40004"
VDM=server_2 CIFS SERVER=DBMS[W2K] dump cred at 0xe06c9a04
1205772858: SMB: 6: USER 'W2K\w2ku40004 (mapped=W2kU40004.W2K)'
Auth=KERBEROS CredCapa=0x2
SID = S-1-5-15-242a3a09-6bc5c62-3f32a78a-2b69
```

**USE CASE 2:****CIFS User already mapped [then, adding ntxmap.conf file]:**

Usermapper=unix40004 →Usermapper name updated from W2KU40004

Secmap=W2k\W2kU40004 from “Origin” ntxmap →Updated mapping origin

Ntcred=(mapped=unix40004.W2K) →NTCred shows ntxmap mapping

```
# .server_config server_2 -v "usrmap dump"
```

```
S-1-5-15-242a3a09-6bc5c62-3f32a78a-2b69:*:32768:32768:user unix40004 from domain
W2K:/usr/S-1-5-15-242a3a09-6bc5c62-3f32a78a-2b69:/bin/sh
```

```
# server_cifssupport server_2 -seemap -list -name w2ku40004 -domain w2k
```

| SECMAP USER MAPPING TABLE |        |                          |               |                                         |
|---------------------------|--------|--------------------------|---------------|-----------------------------------------|
| UID                       | Origin | Date of creation         | Name          | SID                                     |
| 32768                     | ntxmap | Mon Mar 17 15:22:51 2008 | W2K\W2ku40004 | S-1-5-15-242a3a09-6bc5c62-3f32a78a-2b69 |

```
# .server_config server_2 -v "ntcred if=cge0 user=w2ku40004"
VDM=server_2 CIFS SERVER=DBMS[W2K] dump cred at 0xcfcb6b004
1205767657: SMB: 6: USER 'W2K\w2ku40004 (mapped=unix40004.W2K)'
Auth=KERBEROS CredCapa=0x2
SID = S-1-5-15-242a3a09-6bc5c62-3f32a78a-2b69
```

**USE CASE 3:****CIFS User accessing Share for first time [using ntxmap.conf file]:**

Usermapper=unix40004

**Secmap=W2k\W2kU40004 from “Origin” ntxmap**

**Ntcred=(mapped=unix40004.W2K)**

**# .server\_config server\_2 -v "usrmap dump" |grep unix40004**

1205772281: USRMAP: 6: S-1-5-15-242a3a09-6bc5c62-3f32a78a-2b69.\*:32768:32768:user unix40004 from domain W2K:/usr/S-1-5-15-242a3a09-6bc5c62-3f32a78a-2b69:/bin/sh

**# server\_cifssupport server\_2 -seemap -list -name w2ku40004 -domain w2k**

SECMAP USER MAPPING TABLE

| UID | Origin | Date of creation | Name | SID |
|-----|--------|------------------|------|-----|
|-----|--------|------------------|------|-----|

|       |        |                          |               |                                         |
|-------|--------|--------------------------|---------------|-----------------------------------------|
| 32768 | ntxmap | Mon Mar 17 13:58:56 2008 | W2K\W2kU40004 | S-1-5-15-242a3a09-6bc5c62-3f32a78a-2b69 |
|-------|--------|--------------------------|---------------|-----------------------------------------|

**# .server\_config server\_2 -v "ntcred if=cge0 user=w2ku40004"**

VDM=server\_2 CIFS SERVER=DBMS[W2K] dump cred at 0xe06e0604

1205762346: SMB: 6: USER 'W2K\w2ku40004 (mapped=unix40004.W2K)'

Auth=KERBEROS CredCapa=0x2

SID = S-1-5-15-242a3a09-6bc5c62-3f32a78a-2b69

#### USE CASE 4:

**CIFS User already mapped from ntxmap [removing ntxmap.conf file]:**

**Usermapper=unix40004**

**Secmap=W2k\W2kU40004 from “Origin” ntxmap**

**Ntcred=(mapped=W2KU40004.W2K)**

**# .server\_config server\_2 -v "usrmap dump" |grep unix40004**

1205772281: USRMAP: 6: S-1-5-15-242a3a09-6bc5c62-3f32a78a-2b69.\*:32768:32768:user unix40004 from domain W2K:/usr/S-1-5-15-242a3a09-6bc5c62-3f32a78a-2b69:/bin/sh

**# server\_cifssupport server\_2 -seemap -list -name w2ku40004 -domain w2k**

SECMAP USER MAPPING TABLE

| UID | Origin | Date of creation | Name | SID |
|-----|--------|------------------|------|-----|
|-----|--------|------------------|------|-----|

|       |        |                          |               |                                         |
|-------|--------|--------------------------|---------------|-----------------------------------------|
| 32768 | ntxmap | Mon Mar 17 13:58:56 2008 | W2K\W2kU40004 | S-1-5-15-242a3a09-6bc5c62-3f32a78a-2b69 |
|-------|--------|--------------------------|---------------|-----------------------------------------|

**# .server\_config server\_2 -v "ntcred if=cge0 user=w2ku40004"**

VDM=server\_2 CIFS SERVER=DBMS[W2K] dump cred at 0xe06e0604

1205762346: SMB: 6: USER 'W2K\w2ku40004 (mapped=W2KU40004.W2K)'

Auth=KERBEROS CredCapa=0x2

SID = S-1-5-15-242a3a09-6bc5c62-3f32a78a-2b69

## ***COGNAC MAINTENANCE RELEASE PROGRAM***

**COGNAC 5.6 MAINTENANCE RELEASE 1 CMR1:** May 2008, 5.6.37.x

→CSA enhancements, new Checklist for install issues, Popup to advise using crossover cable for discovery issues

→New NS20 (NS20 Single Blade; NS20 Dual Blade; NS20FC Single Blade; NS20FC Dual Blade) and NS40 (NS40 Single Blade; NS40 Dual Blade) Installation Guides [replacing old Placemat docs]

→Support for Symm DMX4 Enginuity 5773, with support for SSD drives (73 & 146GB) and Solutions Enabler 6.5

**Note:** Not Thin Provisioning support

→Support for MPFS Write Back (new call, FMP\_WRITE)

**SYMMETRIX EFD DRIVES (Enterprise Flash Drives using flash-based solid state drives):**

--Symm Enginuity 5773

--SSD referred to as Tier 0 storage, 10x faster response times, 30x greater IOPs, less power required

--Uses 73 & 146GB drives with Raid5 3+1 or 7+1 only, min. of 1 Hot Spare per 32 drives

--new AVM pool symm\_ssd, not greedy, allows striping and uses 32K as default stripe size

--cannot mix with other drive types, nor use for SRDF or TimeFinder

--Symm Powervault drives must be traditional HDD

**ENTERPRISE FLASH DRIVE PERFORMANCE:**

--1ms latency vs. 6-12ms latency for 15K RPM HDDs

--For random Reads, will be 10x faster and provide up to 30x greater IOPs

--20x faster destaging of data to EFDs than HDD

--5 to 8x faster for sequential Read operations

**COGNAC 5.6 MAINTENANCE RELEASE 2 (CMR2):** June 2008, 5.6.38.x

→RepV1 to V2 upgrade script [upgrade method to abort current V1 sessions and use internal restart checkpoints to upgrade to V2], and Celerra Manager GUI enhancements [rightclick fs, iSCSI lun, or VDM to create new replication session; Bandwidth Schedule function]

→Symmetrix 5773 Thin Provisioning support (ability to present more storage to application than is really allocated, until needed)

→Backend Change Monitor tool officially released with NAS code (see emc195237 for more info)

→DUDL alerts will show on Home page Celerra Mgr

### **SYMMETRIX THIN PROVISIONING:**

--Manage via CLI or Symmetrix Management Console

### **BACKEND CHANGE MONITOR:**

→Collects backend configuration information, then tracks changes via SCCS

→Feature is configured during installation or upgrade to 5.6.38.2 or higher, as seen in NAS Upgrade log output  
upgrade.5.6.38-2.Jun-05-14\16.log

```
58      Set up backend monitor tool      0
=====
```

```
15:51 [ 58/67 ] Set up backend monitor tool      1 second
```

Start at: Thu Jun 5 15:51:00 2008

Complete at: Thu Jun 5 15:51:00 2008

→Feature is automatically kicked off by the NASDB backup script once every 24 hours, relying on timestamp of /celerra/backendmonitor/tmp directory to know when the last time it ran

```
# view /nas/sbin/nasdb_backup
```

```
# run backend configuration collector
```

```
/celerra/backendmonitor/bin/backend_change_monitor -d -m queryall
```

→Feature does not log when it runs in the sys\_log or elsewhere, but does update files in /celerra/backendmonitor directory if changes are captured

→Feature may indirectly create CallHome Warnings for backend or zoning changes that are flagged during the bi-weekly PAHC checkup

→Limited to monitoring of Cisco, McData, & Brocade switches, with SNMP queries for zone info only on Cisco & McData

→SCCS is used to track configuration changes, up to 1 year only

→Feature allowed to consume only 7MB in its home directory, /celerra/backendmonitor, which is an LVM directory

→By default, feature does not monitor any Switches—must set this up via separate –addentry commands to zonemonitor.cfg file

→Live configuration file is /nas/site/zonemonitor.cfg

### **CallHome Example:**

```
# cat /nas/log/checkup-run.080415-143003.log (PAHC Log)
```

Storage System: Check if backend zone correct session

Symptom:

\* The Data Mover in slot 2 HBA 5006016039a008f5 and Symmetrix/CLARiiON

HBA 5006016039a02c85 are in different zones on switch 10.32.22.64.

Action : The zones are not configured correctly.

### **sys log entry for CallHome:**

Apr 15 14:32:00 2008:87519264790::nas\_checkup has posted a warning. See /nas/log/checkup-run.080415-143003.log for more details.

**FEATURE DOES TWO THINGS:** Data Collection and Switch zone checks

### **DATA COLLECTION:**

#### **/celerra/backendmonitor/bin/backend\_change\_monitor**

**Note:** This is run on-demand by User, or by NASDB backup script daily, tracking system configuration changes

→Script runs multiple queries

run\_query\_clariion [runs hostip port –list, storagegroup –list, and checks if Integrated or not]

run\_query\_symm [list logins symmask]

run\_query\_dm [runs fcp show and fcp bind show on Data Movers]

run\_query\_switch [does snmpwalk for Cisco and McData switches]

### **ZONE CHECKS:**

#### **/celerra/backendmonitor/bin/check\_zone**

**Note:** This is run by PUHC and/or PAHC (situational/bi-weekly), checking zoning between Blades & Backend

→Check zone parses the data collection by the collection script, looking at the following rules:

--DM WWN zone vs. Clariion/Symm WWN zone, if not in same zone, will report bad zoning configuration

--DM WWN zone vs. Clariion/Symm WWN zone, if one is in a zone but the other is not, reports bad zoning configuration

--If neither DM or Backend WWNs are in a zone, will not report anything

### **Working Directory for Backendmonitor:**

#### **/celerra/backendmonitor**

```
# ls -la
```

```
drwxr-xr-x 2 nasadmin nasadmin 1024 May  7 13:51 bin
```

```
drwxr-xr-x 3 nasadmin nasadmin 1024 May  8 14:01 data
```

```
drwxr-xr-x 2 nasadmin nasadmin 1024 May  8 14:01 etc
```

```
drwx----- 2 nasadmin nasadmin 12288 Feb 25 16:06 lost+found
```

```
drwxr-xr-x 2 nasadmin nasadmin 1024 Jun 9 21:01 tmp
# ls -la * -->Bin directory is where the two scripts reside
bin:
-rwxr-xr-x 1 root  root  18066 Jun  2 13:07 backend_change_monitor
-rwxr-xr-x 1 root  root   5707 Jun  2 13:07 check_zone
data:
-rw-rw-r-- 1 nasadmin nasadmin  3 May  8 14:01 isintegrated
-rw-rw-r-- 1 nasadmin nasadmin 3144 May  8 14:01 naviportlist.192.168.100.200
-rw-rw-r-- 1 nasadmin nasadmin 3144 May  8 14:01 naviportlist.192.168.200.201
-rw-rw-r-- 1 nasadmin nasadmin 861 May  8 14:01 navistoragegroupelist.192.168.100.200
-rw-rw-r-- 1 nasadmin nasadmin 861 May  8 14:01 navistoragegroupelist.192.168.200.201
drwxr-xr-x 2 nasadmin nasadmin 1024 Jun  1 00:01 SCCS
-rw-rw-r-- 1 nasadmin nasadmin 510 May  8 14:01 slot_2
-rw-rw-r-- 1 nasadmin nasadmin 226 May  8 14:01 slot_2_persistent
-rw-rw-r-- 1 nasadmin nasadmin 510 May  8 14:01 slot_3
-rw-rw-r-- 1 nasadmin nasadmin 226 May  8 14:01 slot_3_persistent
etc:
-rwxr-xr-x 1 nasadmin nasadmin 185 Jun  2 13:07 weekly-backendmonitor.conf
-rw-rw---- 1 nasadmin nasadmin 871 Jan  1 1970 zonemonitor.cfg
tmp:
```

#### **Staging Directory for ChangeMonitor Install Files:**

#### **/nas/tools/backendmonitor\_tools**

```
# ls -la
-rwxr-xr-x 1 root root 18066 Jun  2 13:07 backend_change_monitor
-rwxr-xr-x 1 root root  5707 Jun  2 13:07 check_zone
-rwxr-xr-x 1 root root  188 Jun  2 13:07 monthly-backendmonitor.conf
-rwxr-xr-x 1 root root  185 Jun  2 13:07 weekly-backendmonitor.conf
```

#### **Default Configuration Files:**

**/celerra/backendmonitor/etc/weekly-backendmonitor.conf**

**/celerra/backendmonitor/etc/zonemonitor.cfg**

**/nas/site/zonemonitor.cfg** [this is the real configuration file location]

```
# cat weekly-backendmonitor.conf
# rotate log files weekly
# keep 12 weeks worth of backlogs
# and compress
/celerra/backendmonitor/data/SCCS/s.*.1 {
    missingok
    nocreate
    weekly
    compress
    rotate 12
}
# cat zonemonitor.cfg
# This is the configuration file for automatic configuration change
# monitor.
# Anything after the # character is ignored. Blank lines are also ignored.
# To get back to the default configuration file, simply remove or
# rename this file and a new one will be created
# next time automatic configuration change monitor runs.
# You may change this file if you wish,
# and if you do so, it will be recreated, and
# you will not receive any updates to this file should you
# choose to upgrade your control station software later.
# The following configurations are used by SNMP query.
# Please do not modify these unless instructed by
# authorized service personnel.
# zone-snmpquery ipadd1 switch-model1 community1
# For example:
# zone-snmpquery 10.6.27.85 cisco public
```

```
# zone-snmpquery 10.6.89.81 medata public  
# zone-snmpquery 172.24.97.203 brocade public
```

#### **BACKENDMONITOR CLI COMMANDS:**

**# /celerra/backendmonitor/bin/backend\_change\_monitor -config**

The list of switches being monitored:

No switch is being monitored.

#### **MODIFYING ZONEMONITOR CONFIG FILE:**

```
# /celerra/backendmonitor/bin/backend_change_monitor -addentry 10.6.27.85 cisco public  
# /celerra/backendmonitor/bin/backend_change_monitor -addentry 10.6.89.81 medata public  
# /celerra/backendmonitor/bin/backend_change_monitor -addentry 172.24.97.203 brocade public
```

**# /celerra/backendmonitor/bin/backend\_change\_monitor -config**

The list of switches being monitored:

10.6.27.85(cisco)

10.6.89.81(medata)

172.24.97.203(brocade)

#### **REMOVING SWITCH FROM MONITORING:**

**# /celerra/backendmonitor/bin/backend\_change\_monitor –removeentry 10.6.27.85**

#### **BACKEND CHANGE MONITOR COMMAND & HELP MENU:**

**# /celerra/backendmonitor/bin/backend\_change\_monitor -help**

usage: backend\_change\_monitor

| -help

| -config

| -addentry <hostname/IP> <switch model> <SNMP community>

| -removeentry <hostname/IP>

| -mode <option>

Options:

-help - this help screen

-config - Display the current configuration

-addentry - Add a switch to be monitored. <hostname/IP>  
is the hostname or IP address of the switch,  
<switch model> is either cisco, brocade, or  
medata, and <SNMP community> is the community  
string used for SNMP query against the switch

-removeentry - Remove a switch from being monitored.  
<hostname/IP> is the hostname or IP address of  
the switch

-mode - run checks specified by <option>. The option  
is comma seperated, and it can be:  
queryall: do all query.  
queryzone: query fibre switch config.  
querydm: query DM related issue.  
queryclarion: query Clariion.  
querysymm: query Symm.

## **COGNAC 5.6.39 MAINTENANCE RELEASE 3 (CMR3):** August 22, 2008, 5.6.39.5

#### **NS20 UPDATE:**

→Increase drive capacity from 60 to 90 FC or SATA drives, based on Flare 26 Patch 012 minimum

→Ability to perform “data in place” upgrades from NS20 to NS40 (Swapout of SPs, support for 240 drives)

**Notes:** Upgrades from CX3-10 or CX3-10F to CX3-40F. No upgrade allowed if using an NS20 with SATA vault drives.

#### **Upgrade Options:**

NS20 Integrated to NS40 Integrated (NS40 with Headhunter IO card on SPs, but without the FC-Enabled option)

NS20 Integrated to NS40FC

NS20FC to NS40FC

#### **REPLICATION ENHANCEMENTS:**

→IPRepV2 internal replication checkpoints will be hidden from GUI and CLI output by default (AR116012), but there will be an option to view the internal checkpoints. \$ nas\_fs –list –all and \$ fs\_ckpt –list –all will be added to output internal checkpoints, and an option called ‘All Checkpoints Including Those Used Internally’ will be added to the GUI.

→Replication Wizard enhancements in Celerra Mgr

#### **NX4 RELEASE:**

- Release of new NSX platform and support for Maynard Control Station
- Support for new Email User notification feature (aka, Local email support)
- CSA changes to support setting up ConnectHome email and Email User notification feature
- Support for new NX4 Backend Event Reporting mechanism (not based on Naviagent; based on StorageAPI)

### **EMAIL USER NOTIFICATION FEATURE:**

- Introduced with 5.6.39 in conjunction in NX4, but will also support NS20/NS40/NS80/NSX for new installations or upgrades
- Email User feature is independent of the ConnectHome Email feature—In otherwords, you could have a system where ConnectHome was not Enabled at all and yet the Email User notification was, and email would be delivered to the Recipient defined in the Email User feature
- Email User feature absolutely requires that the Control Station be properly configured with entries that would allow proper DNS lookup and resolution of the “Recipient Email Addresses” by FQDN  
e.g., bigdog@fido.pvt.dns

#### **Example of /etc/resolv.conf:**

```
# cat /etc/resolv.conf
domain fido.pvt.dns
search fido.pvt.dns
nameserver 192.168.4.10
nameserver 192.168.5.10
```

- Email User notification feature will allow customers to receive CCMD descriptive Email alerts for all Celerra CallHome events
- CallHome Immediate or Materials triggers Email Notification
- Basic configuration during CSA setup
- More extensive configuration possible from CLI or GUI
- Based on Sendmail application and using /var/log/maillog
- Driven by same nas events engine as are ConnectHome feature
- Feature requires Recipient Email address to be enabled
- Becomes part of the DBMS BDB system
- Email Server field is truly optional, and can be blank as long as DNS is properly configured to resolve the FQDN of the Recipient email address

#### **GUI Configuration:**

- Can configure as nasadmin or root user, unlike ConnectHome, which requires root login

#### **Celerras>Support>Email User**

Service Enabled checkbox

Recipient Email Addresses:

CC Email Addresses: (Optional)

Email Server: (Optional) →Would be required if sending email external to company's DNS domain, along with proper DNS search entries in /etc/resolv.conf. Email Server entries are written to line 101 in the /etc/mail/sendmail.cf configuration file. Use /sbin/service sendmail restart to restart Sendmail program.

Subject Prefix: Celerra Notification

Sender Email Address: (Optional)

**Note:** GUI changes for most fields are put into effect immediately. If changes are made to the “Email Server” section, the entries are written to the /nas/sys/email\_user.config immediately, but are not written to /etc/mail/sendmail.cf until an event trigger for Email User is generated, at which point the sendmail.cf file will be updated with new Email Server. Email User test button in GUI should work. Alternatively, use /sbin/service sendmail restart to update file immediately. AR126368.

#### **TESTING EMAIL USER FEATURE FROM CELERRA MANAGER>SUPPORT>Email User>Test**

- Test button should display “Test Email Notification Service. OK” and send a test message to the Recipient:

#### **TEST EMAIL USER MESSAGE:**

Date: Thu, 18 Sep 2008 14:32:12 -0400

From: smiley@acme.com

To: tmatta@sun1.hosts.pvt.dns

Subject: Celerra Notification - nx4.w2k.pvt.dns, INFO, EmailUser

Event Time: Sep 18 14:32:12 2008

Brief Description: Test event for Email User.

Full Description:

Test event for Email User.

Recommended Action:

No action is required.

Celerra Name: nx4.w2k.pvt.dns (192.1.4.230)

Celerra Model: NX4FC

Celerra Serial #: APM00081800232

NAS Version: 5.6.40-2

**CLI:****# nas\_emailuser**

Error 2100: Usage:

```
nas_emailuser
  -info
  | -test
  | -modify [ -enabled { yes | no } ]
    [ -to <email_addr>[,...] ]
    [ -cc <email_addr>[,...] ]
    [ -email_server <email_server> ]
    [ -subject_prefix <email_subject> ]
    [ -from <email_addr> ]
```

| -init →Creates a configuration file if one does not exist, /nas/sys/.email\_user.config

**# nas\_emailuser -info**

Service Enabled = Yes

Recipient Address(es) = user@domain.company.xxx.com

Carbon copy Address(es) =

Email Server = 10.68.150.240

Subject Prefix = Celerra Notification

Sender Address =

**Enabling/Disabling Email User Feature:****# nas\_emailuser -modify -enabled no | yes**

OK

**Note:** The recipient email address must be configured prior to enabling email notification**# nas\_emailuser -modify -enabled yes -to smiley@acme.com**

OK

**Example When Trying to Enable Feature without Valid Recipient:****# nas\_emailuser -modify -enabled yes**

Error 13693812739: The value for attribute 'emailTo' is invalid for the requested action.

**Example of Adding Multiple Recipient Addresses:****# nas\_emailuser -modify -cc user1@acme.com,user2@acme.com**

OK

**# nas\_emailuser -info**

Carbon copy Address(es) = user1@acme.com,user2@acme.com

**Recreating Default Email Configuration:****/nas/bin/nas\_emailuser –init** [Recreates default configuration file if none exists--/nas/sys/.email\_user.config]**Email User & ConnectHome Configuration File/Ownership/Permissions:**

# ls -la /nas/sys/.email\*

```
-rw-r--r-- 1 root  root  711 Jan 16 16:30 connecthome.config
-rw-rw-r-- 1 root  fullnas 130 Jan 14 15:08 .email_user.config
-rw-r--r-- 1 nasadmin nasadmin 10 Jan 14 15:08 ..email_user.config.sig
```

**Default Emailuser configuration (/nas/sys/.email\_user.config):****# nas\_emailuser -info**

Service Enabled = No

Recipient Address(es) =

Carbon copy Address(es) =

Email Server =

Subject Prefix = Celerra Notification

Sender Address =

**Backup Copy of Email User Configuration File:**

- If the file /nas/sys/.email\_user.config gets deleted or corrupted, may see following message

**# nas\_emailuser -info**

Error 13693812737: The email user service configuration file does not exist.

**# nas\_message -info 13693812737**

MessageID = 13693812737

BaseID = 1

Severity = ERROR

Component = APL

Facility = EmailUser

Type = STATUS

Brief\_Description = The email user service configuration file does not exist.

Full\_Description = The email user service configuration file, /nas/sys/.email\_user.config, does not exist.

Recommended\_Action = Copy the backup file /nasmcd/CHomeFiles/.email\_user.config to /nas/sys/ or use -init option to create the configuration file.

2. If available, copy the /nasmcd/CHomeFiles/.email\_user.config file back to /nas/sys/

3. Or, to recreate the config file: /nas/sbin/nas\_emailuser -init

#### **Testing EmailUser Feature:**

# **nas\_emailuser -test**

OK

# **nas\_logviewer -t /nas/log/sys\_log |tail**

Aug 28 10:32:09 2008:96109330532::Test event for Email User.

#### **Rules for Recipient Email Addresses:**

The mailbox portion of the fully qualified domain name (FQDN) uses the following ASCII characters: a-z, A-Z, 0-9, !, #, \$, %, &, ` , \*, +, -, /, =, ?, ^, \_, ` , {, |, }, ., and ~. Periods cannot be the first or last character in the mailbox. The maximum number of characters, including spaces and commas, is 255. Each email address can have a maximum of 63 characters.

#### **Comments on Configuring Email User feature:**

→ Set the sender email so that bounced emails go back to the Administrator

→ Sender email may be required for some email servers, depending on Customer's environment

→ If emails are sent external from Customer's network, then the Email Server field must be filled in with an Email Server that has access to the external network, along with configured DNS entries in /etc/resolv.conf for the external network

→ If the email send fails at the first email server, then sendmail will continue to retry for a few times, then gives up

→ CSA Wizard will be used to configure both ConnectHome email, and Email User notification

Issue with Email User Feature with NAS Upgrades:

→ During an upgrade to 5.6.39.5 or higher, the Email User feature must be initialized with its default /nas/sys/.email\_user.config configuration file. However, there are certain DNS misconfigurations on the Control Station that can cause the System Management service to take a long time to complete its initialization during the Upgrade process, and in turn may cause a Task 65 failure when trying to setup the email user feature:

**/install\_mgr -m upgrade**

=====Tasks=====

13:26 [ 65/68 ] Initial emailuser 15 seconds

Waiting for mgmtd to complete the initialization ...

nas\_emailuser failed to initialize: Error 13958709250: No response from Celerra

Manager for task Init Email Notification Service..

Error : Failed to initialize the emailuser feature.

Result: Failure

=====Task [65/68] "Initial emailuser" has failed.

#### **Workaround:**

1. Stop the System Management daemon on the Control Station:

# killall mgmtd

2. Tail log and wait for the Management Daemon to complete its restart:

# tail -f /nas/log/nas\_log.al.mgmtd

2008-11-17 12:50:32.409: INIT: 1: MGMTD initialization complete:

3. Initialize the default Email User configuration:

# /nas/bin/nas\_emailuser -init

OK

4. Start the NAS Upgrade

#### **Default contents of the hidden .email\_user.config file:**

# **/nbsnas/bin/nas\_emailuser -info**

Service Enabled = No

Recipient Address(es) =

Carbon copy Address(es) =

Email Server =

Subject Prefix = Celerra Notification

Sender Address =

#### **SCSI-3 PERSISTENT RESERVATION SUPPORT:**

→ Celerra now supports SCSI-3 Persistent LUN Reservations for Celerra iSCSI Targets (restrict or share access to SCSI LUNs)

→ SPC-3, SBC-2, & SAM-3 are used for this implementation

→Main purpose is for iSCSI target & cluster support for W2K8 & MS 2008 iSCSI Failover Clusters

**Note:** Client nodes will be able to register and reserve disk access rights as an I/O fencing mechanism for iSCSI luns where disks may be shared, as in Clusters. Persistent reservations are managed by the SCSI (SPC-3) subsystem

→Reservations are persistent through I\_T nexus loss, target resets, LUN resets, & power loss

→Cannot use both SCSI-2 Reserve/Release and SCSI-3 PR at the same time

#### **Configuring PR with MS Windows 2008 Clusters:**

1. Connect all hosts in the cluster for the Celerra iSCSI target

2. From MS GUI Failover Cluster Manager, create cluster

3. Optionally configure MPIO and add Celerra as “EMC Celerra” with five spaces between EMC and Celerra

#### **Server Log Example for code prior to Support with Windows 2008 iSCSI Clusters:**

2008-08-04 12:11:55: VLU: 3: SCSI op code 5e not supported

2008-08-04 12:11:58: VLU: 3: SCSI op code 5f not supported

#### **Persistent Reservation OpCode Commands:**

0x5E →Persistent Reservation In OpCode (Host cluster queries access rights db for SCSI LUNs from iSCSI targets)

0x5F →Persistent Reservation Out OpCode (Host cluster registers or reserve access rights to SCSI LUNs)

#### **PERSISTENT RESERVATION OUT SERVICE ACTIONS:**

REGISTER | RESERVE | RELEASE | CLEAR | PREEMPT | PREEMPT AND ABORT | REGISTER AND IGNORE EXISTING KEY | REGISTER AND MOVE

#### **PERSISTENT RESERVATION IN SERVICE ACTIONS:**

READ KEYS | READ RESERVATION | REPORT CAPABILITIES | READ FULL STATUS

#### **Configuring APTPL:**

→Purpose of APTPL is to be able to recover all PR information and state after power loss

→Disabled by default on Celerra

→Celerra can store PR information on disk if Hosts are configured and have requested this (APTPL—Activate Persist Through Power Loss—this can be enabled via Celerra parameter)

→MS Windows 2008 does not use the APTPL feature, though other Host Cluster implementations do

#### **# server\_param server\_2 -facility iscsi -info EnableAptpl**

server\_2 :

```
name      = EnableAptpl
facility_name = iscsi
default_value = 0
current_value = 0
configured_value =
user_action = reboot DataMover
change_effective = reboot DataMover
range     = (0,1)
description = Support PR Persist Through Power Loss
```

#### **# server\_param server\_2 -facility iscsi -modify EnableAptpl -value 1**

server\_2 : done

Warning 17716815750: server\_2 : You must reboot server\_2 for EnableAptpl changes to take effect.

#### **SCSI-3 Persistent Reservation Issue Resulting in Long Data Mover Failover Times--AR124930:**

→NAS 5.6.39, 5.6.40, & 5.6.41 exhibit long Server failover times if iSCSI LUNs are in use

#### **Workaround:**

Disable following parameter and reboot the Data Mover

#### **# server\_param server\_x -f iscsi -modify EnableAptpl -value 1**

#### **CLARIION LUN RENAME FEATURE (emc192005):**

→Celerra LUN rename feature for Clariion systems [Default CLARIION behavior is to name a LUN as “LUN 1”, etc., with Navisphere, and “Virtual Disk 201”, etc., with NaviExpress

#### **DEFAULT CLARIION LUN NAMES:**

→If Navisphere is used as the management tool, LUNs that belong to Storage Groups are labeled “LUN 6”, etc.

→If NavisphereExpress is used as the management tool, LUNs are known as “Virtual Disks” and default to names such as “Virtual Disk 201”, etc.

#### **FEATURE RULES:**

→Celerra LUNs will automatically be renamed during the diskmarking process or Celerra Manager Rescan

→Celerra LUNs, Storagegroup name, and Initiator Records, are renamed if Control Station Hostname is renamed via Celerra Mgr or CSA process

→Default Storage Group name

#### **Celerra\_<CS\_hostname>**

```
# /nas/sbin/navicli -h 10.241.168.188 storagegroup -list |grep "Group Name"
```

**Storage Group Name:** Celerra\_sludge4

→Default Server Initiator Record name will be the same as the Storage Group name, per the following format

**Celerra\_<CS\_hostname>**

**Example:**

```
# /nas/sbin/navicli -h 10.241.168.188 port -list |grep Server
```

**Server Name:** Celerra\_sludge4

Server IP Address: 10.241.168.187

**LUNs are renamed using the following convention, with Control LUNs appended with (NAS/OS):**

**Celerra\_<CS\_Hostname>\_<LUN\_ID>\_<dVolume>**

```
# /nas/sbin/navicli -h 10.241.168.188 getlun 16 -name
```

Name Celerra\_sludge4\_16\_d7

**For example, on a factory system where Control Station uses default name, “emcnas\_i0”:**

Celerra\_emcnas\_i0\_0\_root\_disk (NAS/OS)

Celerra\_emcnas\_i0\_8\_d9

→If the Control Station hostname is changed via the CSA process or when using Celerra Manager, a special

/nas/http/webui/bin/hostname.pl script is run that Updates Hostname; Runs nas\_cel –update; Restarts JServer; starts background job to run setup\_clariion –check\_and\_update –all, which updates Celerra Storage Group & Initiator Records names with the updated Hostname info, and finally, calls nas\_diskmark, which updates the LUNs with the new Hostname info

→If a LUN name is manually renamed or customized using Navisphere or NaviExpress, then the LUN will not be renamed during next nas\_diskmark/server\_devconfig -create operation

→If nas\_disk –rename is used to rename a dvolume ID number, the corresponding information is updated in the LUN name

→If nas\_disk –delete is used to remove a dvolume from the Celerra database, the LUN is renamed to the default CLARiiON LUN naming convention (e.g., LUN 6)

→Feature can be disabled by editing /nas/sys/nas\_param file

→Celerra Control LUNs are easily identified by the “(NAS/OS)” label after the name

→LUNs deleted from NAS DB will always revert back to the “LUN x” naming convention

**CONTROL STATION HOSTNAME RENAME WEBUI SCRIPT:**

**# /nas/http/webui/bin/hostname –s <new\_hostname>**

**Note:** This should be equivalent to changing Control Station Hostname from Celerra Mgr, from the CLI

**COMMANDS THAT UPDATE LUN NAMES:**

**nas\_disk –delete** →In this case, restores to the default CLARiiON “LUN x” name

**nas\_disk –rename** →Updates dvolume ID (e.g., d10 to d20), and updates dvolume ID info in LUN name

**nas\_diskmark –mark –all** →Diskmarks any new volumes and renames with new Celerra LUN name

**server\_devconfig –create** →Diskmarks any new volumes and renames with new Celerra LUN name

**WHAT THE VARIOUS SCRIPTS DO FOR THIS FEATURE:**

**/nas/sbin/setup\_clariion –check\_and\_update –all**

→Updates Storage Group and Celerra Initiator Names, if required

→Calls nas\_diskmark script, which then updates LUN names if required

**/nas/bin/nas\_diskmark**

→If called separately, will rename any Celerra-owned lun from its default CLARiiON name [e.g., LUN 20, Virtual Disk 201, etc] to the new Celerra\_cshostname\_lunid\_dvolume naming convention

→Otherwise, nas\_diskmark is called by the setup\_clariion –check\_and\_update –all script to update LUN names

**LUNs REVERT TO DEFAULT CLARIION NAMES IF DVOLUMES DELETED FROM CELERRA DB:**

**# /nas/sbin/navicli -h 192.1.4.231 getlun -name**

LOGICAL UNIT NUMBER 6

Name Celerra\_nx4\_6\_d11

**# nas\_disk -d d11 -perm**

**# /nas/sbin/navicli -h 192.1.4.231 getlun -name**

LOGICAL UNIT NUMBER 6

Name LUN 6 →This was d11 as shown above

**Note:** Does not get renamed back to default NaviExpress names, which are ‘Virtual Disk xxx’ etc.

**DISABLING LUN RENAME FEATURE:**

→By default, the feature is enabled, but can be easily disabled by editing the /nas/sys/nas\_param file to change the clar\_rename\_luns from true to false:

**# cat /nas/sys/nas\_param**

#clar\_rename\_luns:<true|false>

clar\_rename\_luns:true

**PERMANENTLY DISABLING LUN RENAME FEATURE:**

**Note:** Ordinarily, any changes made to the /nas/site/nas\_param file are reset to code defaults, say after a NAS Upgrade. In order to disable the LUN Rename feature persistently, make the following edit to the /nas/site/nas\_param file:

```
# vi /nas/site/nas_param
```

```
#clar_rename_luns:<true/false>
```

#### **LUN RENAME LOGGING:**

→See /nas/log/symapi.log for entries related to LUN renaming

|                         |      |                  |                |                                  |
|-------------------------|------|------------------|----------------|----------------------------------|
| 2008-07-01 06:43:07.590 | 7788 | 1 EMC:Celerra CS | User Generated | Renamed LUN SL7E1080700121 00006 |
|-------------------------|------|------------------|----------------|----------------------------------|

#### **Celerra\_emcnas\_i0\_6\_d12**

|                         |       |                  |                |                                  |
|-------------------------|-------|------------------|----------------|----------------------------------|
| 2008-07-01 06:45:30.883 | 10369 | 1 EMC:Celerra CS | User Generated | Renamed LUN SL7E1080700121 00006 |
|-------------------------|-------|------------------|----------------|----------------------------------|

#### **LUN 6**

### **TCP/IP LARGE RECEIVE OFFLOAD (LRO) FEATURE:**

→TCP/IP Large Receive Offload (LRO) [aggregation of multiple incoming packets into fewer, larger, packets to reduce CPU load]

→DART implements in software at Data-Link network layer

→Feature will be disabled by default

#### **Enabling/Disabling LRO:**

```
# .server_config server_2 -v "param tcp lro=1" and add param to /nas/server/slot_x/param file
```

```
# .server_config server_2 -v "param tcp lro=0"
```

#### **Displaying and Clearing LRO Stats:**

```
# .server_config server_2 -v "bcm cge0 displayLroStat" "bcm cge0 clearLroStat"
```

### **WINDOWS IDENTITY MANAGEMENT FOR UNIX FEATURE (IMU or IdMU):**

→Windows IdMU allows DART to use AD as the single naming service source for both NFS & CIFS protocols

→IdMU is an enhancement to retrieve Users and Groups from AD using LDAP and MS-IdMU schema (RFC2307)

→Only a single domain is support

→DART uses Bind + Password or Bind SSL to connect to AD server when using LDAP protocol

→IdMU provides a NIS Server to NFS clients and supports Unix Netgroups [i.e., a NIS replacement in multi-protocol environment]

→IdMU supported with Windows Server 2003 R2 and newer

→No Kerberos support for IdMU yet (NAS 6.0)

→Requires Celerra configuration for “ldap” in /etc/nsswitch.conf file for Users/Groups/Netgroups (and Hosts if not using DNS), and AD Servers supporting IMU Schema

→Group accounts can be Local or Domain-mapped from entries in /etc/group file on the Control Station (GID for Domain-mapped groups is needed only for use on the Control Station)

→User accounts are also Local or Domain-mapped from entries in /etc/passwd file on the Control Station (UID for Domain users needed on Control Station and controls client access privileges)

→Domain User accounts must also have membership in at least one Domain-mapped Group in order to login to the CS

#### **Default Windows Template Configuration Files:**

```
# ls -la /etc/ldap*
```

```
-rw-r--r-- 1 root bin 1814 Aug 13 10:24 /etc/ldap.conf.idmu_template_v1 (SFU3.5 Schema Windows 2003 R1)
```

```
-rw-r--r-- 1 root bin 2007 Aug 13 10:24 /etc/ldap.conf.sfu35_template_v1 (IMU Schema Windows 2003 R2)
```

#### **Configuring IdMU for Celerra:**

1. Set LDAP as a resolver in /etc/nsswitch.conf file for Users, Groups, Netgroups

2. Set param cifs resolver=1 on Data Mover

**Note:** IdmU feature only supports single Windows Domain

3. Disable Secmap on Data Mover (only done if UID/GIDs will be changing)

**Note:** When using LDAP cache, all Users must be authenticated by AD—Secmap is not queried for the LDAP function

4. Disable Usermapper on Data Mover

5. Copy modified Template file as /etc/ldap.conf file

6. Edit /etc/ldap.conf file ‘Containers’ section to reflect AD domain information

```
nss_base_passwd cn=Users,dc=eng-domfsu,dc=lcsc?one
```

```
nss_base_group cn=Users,dc=eng-domfsu,dc=lcsc?one
```

```
nss_base_hosts cn=Computers,dc=eng-domfsu,dc=lcsc?one
```

```
nss_base_netgroup cn=netgroup,cn=eng-domfsu,cn=DefaultMigrationContainer30,dc=eng-domfsu,dc=lcsc?one
```

7. Set the LDAP Client for AD on DART using one of the two possible authentication methods

```
# server_ldap server_2 -set -p -basedn dc=emc,dc=com -binddn cn=administrator,cn=Users,dc=emc,dc=com -servers
```

```
10.240.80.200
```

```
# server_ldap server_2 -set -p -basedn dc=emc,dc=com -binddn cn=administrator,cn=Users,dc=emc,dc=com =servers
```

```
10.240.80.200 -sslEnabled y
```

#### **Migrating NIS Database to AD using “Microsoft Identity Management for UNIX” Snapin:**

→Use “NIS Data Migration Wizard” to migrate NIS maps to AD (maps are passwd, group, hosts, netgroup)

→Create NIS maps on UNIX NIS Server using # ypcat passwd >passwd, # ypcat group > group, etc.

→Create UNIX users manually using ADUC>Users>username>Properties>UNIX Attributes tab

#### **How IdMU Would Work for CIFS User:**

1. CIFS Server would get Windows credentials for User [User SID + Group SIDs]
2. CIFS Server would build credentials
  - a. Query secmap for UID/GID for User SID
  - b. If not found, queries AD for username of SID
  - c. If nsswitch.conf configured with LDAP, uses username to obtain UID/GIDs
  - d. If Usermapper configured and above fails, will create UID/GID

#### **Troubleshooting/Configuring LDAP:**

```
# server_ldap server_x -info -v | -clear | -set | -lookup -user <username> or -group, -hostbyname, -netgroup
```

```
# .server_config server_x -v "logsys set severity LDAP=LOG_DBG2"
```

```
# server_param server_2 -facility ldap -list
```

server\_2 :

|                   |          |         |         |                                                                       |
|-------------------|----------|---------|---------|-----------------------------------------------------------------------|
| param_name        | facility | default | current | configured                                                            |
| cacheEnable       | ldap     | 2       | 2       |                                                                       |
| cacheTTL          | ldap     | 600     | 600     | →Default TTL for LDAP cache is 10 minutes, credentials & UID/GID info |
| cacheMaxUsers     | ldap     | 10000   | 10000   |                                                                       |
| SecurityLayer     | ldap     | 2       | 2       |                                                                       |
| cacheMaxHosts     | ldap     | 10000   | 10000   |                                                                       |
| cacheMaxGroups    | ldap     | 10000   | 10000   |                                                                       |
| cacheMaxNetgroups | ldap     | 10000   | 10000   |                                                                       |

```
# server_param server_2 -facility ldap -info cacheEnable -v
```

server\_2 :

|                      |                                                             |
|----------------------|-------------------------------------------------------------|
| name                 | = cacheEnable                                               |
| facility_name        | = ldap                                                      |
| default_value        | = 2                                                         |
| current_value        | = 2                                                         |
| configured_value     | =                                                           |
| user_action          | = reboot DataMover                                          |
| change_effective     | = reboot DataMover                                          |
| range                | = (0,2)                                                     |
| description          | = Switches off/on the ldap cache and optional offline mode. |
| detailed_description |                                                             |

Switches off/on the ldap cache and optional offline mode. It may be useful to disable the cache for testing. The offline mode means that in the case of a cache hit with an expired entry, if we cannot update the entry because none of the ldap servers are available, then the ldap service will report the expired entry. This feature may be disabled to enhance the security, by setting the cache-only mode 1.

#### **OTHER FEATURES 5.6.39:**

→Support for Celerra Servers as a Host in Navisphere, NST, or NaviExpress wizards, using CS IP address and name as default

→CSA support for MPFS iSCSI systems

**Note:** NS40 and NS40FC for MPFS will use CSA to do normal factory-install configuration of Celerra. CSA can configure back-end iSCSI ports, CHAP, and start MPFS service on Server. Powerlink Landing Page and Installation Guide will be provided.

→Celerra Gateway support for Fleet Arrays (Proteus Flare 28)

→Java JRE 1.5.12 Upgrade (Clients that connect to Celerra Manager will need to upgrade to 1.5 JRE before using)

→Updated Celerra Simulator (Blackbird)

#### **COGNAC 5.6 MAINTENANCE RELEASE 4 (CMR4):** Oct 15, 2008, 5.6.40.3

→Support for 128TB capacities on Blades (qualification effort, no code changes required)

#### **STIG HARDENING:**

→STIG hardening script for CS security for DISA, DLA Federal agencies (once released, can be applied to any 5.5/5.6 version)

**Note:** Script used on CS is “nas\_stig –on | -off” and can be applied to harden security. Backported to 5.5.36, RPQ only.]

```
# /nas/tools/nas_stig
```

nas\_stig

```
-on  
| -off  
| -status  
| -verify  
| -version
```

**Note:** Once implemented, a /etc/nas\_stig.state file is created, and actions logged in /nas/log/nas\_stig.log

**# /nas/tools/nas\_stig -status**

STIG hardening has never been performed on this system (state file is not present)

**ORACLE DNFS SUPPORT:**

→Oracle 11g dNFS support (Oracle NFSv3 client) in order to mitigate an NFSv3 weakness related to data integrity in XID generation by the client. Feature implements a checksum of first 200 bytes of each NFS operation to guarantee data integrity. Feature disabled by default since it incurs an NFS performance hit:

**# server\_param server\_2 -facility nfs -info transChecksum -v**

server\_2 :

```
name      = transChecksum
facility_name = nfs
default_value = 0
current_value = 0
configured_value =
user_action = none
change_effective = immediate
range     = (0,1)
description = NFS Request Checksum
detailed_description
```

Dart-wide param indicating whether NFS request CRC should be calculated.

**# .server\_config server\_2 -v "xidcachestats dump"**

Xid Cache Statistics: Total lookups = 11959

inProg Hits 0, resend reply 0, drop 0, redos 0, drops 0, misses 11959, reforward 0 inProg counts 0 forward Counts 0, CRC

**mismatches 0**

1220455722: ADMIN: 6: Command succeeded: xidcachestats dump

→Maynard CS support for NS20/40 systems

→Multi-byte support for share names, share comments, tree quota path & comments in GUI only [allows native language inputs in support of internationalization efforts]

→FileMover reading of archived data block retrieval enhancement for pass-thru from 8k to 32k (see AR75700)

**Note:** This is a Read Buffer Size enhancement from 8k to 32k

→Replication V2 enhancements (hide option for internal checkpoints; GUI enhancements)

**Note:** Internal checkpoints are exposed in CLI and GUI with 5.6. This enhancement will allow an option to hide the internal checkpoints from view in GUI

**COGNAC 5.6 MAINTENANCE RELEASE 4+ (CMR4+ 5.6.41.2) Dec 2008:**

→NS-120/NS-480 Sleet release support, nasStorageAPI-7.0-12

**Note:** This code will only ship for the Sleet platforms and is not to be used in CCA or for any other hardware platform installs or upgrades

→MPFS support for Sleet hardware and integration into CSA

**Note:** CSA should be able to start the MPFS service on the Data Mover, create the Storage Group, configure iSCSI IP addresses on backend, and set CHAP authentication. Both NS-120 & NS-480 platforms will use FC Enabled option when configuring MPFS over Fibre Channel, and iSCSI option when configuring MPFS over iSCSI.

**COGNAC 5.6 MAINTENANCE RELEASE 5 (CMR5 5.6.42) Dec 2008:**

→Secure FTP using SSL

→Support for SMB2

→System Management performance improvements

→Statistics enhancements

→Cascade replication support on VDMs—up to 4 sessions from Source to Targets; Adds support for two-hop Cascading for RepV2

→Multi-byte support for share names, share comments, tree quota path, tree quota comments

→CSA/CPW qualification for NX4, NS20, NS40, Sleet2, Sleet4

→Dual Maynard CS support NSX/NS80

→64-bit Support for MS VSS Provider for iSCSI for Windows 2003/2008—runs as a Windows service interfacing between Volume Shadow Copy Service and Celerra iSCSI Snapshots. Allows VSS-Enabled backup applications to make shadow copies of iSCSI luns.

**Feature Specifics:**

**DART FTP SECURE (FTPS):**

→Sends FTP traffic over SSL for encryption and authentication (Keys & Certificates managed by PKI)

**Note:** Secure FTP would use SSH to tunnel the FTP session over the connection

→ftp daemon to run during boot.cfg time (ftpd)

→Configuration stored in XML file (/etc/ftpd/conf/ftpd.xml)

- Administered via server\_ftp commands
- Supports new commands (AUTH—Authentication/Security Mechanism; PROT—Data Channel Protection Level; PBSZ—Protection Buffer Size; and CCC—Clear Command Channel)
- Follows RFC959, RFC2228, and RFC4217
- Clients using SSL to Port 990 can use this feature for secure authentication and data transfer on Celerra (control (Port 990) & data channel (Port 9890 SSL encoded))

**\$ server\_ftp –service | -start | -stop | -stats | -server –status**

**\$ server\_ftp server\_x –info | -modify**

**\$ server\_certificate server\_x**

**# server\_ftp server\_2**

```
-info  
| -service {-status | -start | -stop }  
| -service -stats [-full | -reset]  
| -modify  
  [-controlport <port>]  
  [-dataport <port>]  
  [-defaultdir <path>]  
  [-homedir {enable|disable}]  
  [-keepalive <period>]  
  [-highwatermark <threshold>]  
  [-lowwatermark <threshold>]  
  [-deniedusers <path>]  
  [-welcome <path>]  
  [-motd <path>]  
  [-timeout <length>]  
  [-maxtimeout <length>]  
  [-readsize <size>]  
  [-writesize <size>]  
  [-umask <mask>]  
  [-maxcnx <number>]  
  [-sslcontrol {disable|allow|require|requireforauth}]  
  [-ssldata {disable|allow|require}]  
  [-sslpersona {anonymous|default|<name>}]  
  [-sslprotocol {default|ssl3|tls1|all}]  
  [-sslcipher {default|<list>}]  
  [-sslcontrolport {port}]  
  [-ssldataport {port}]
```

**Note:** Must create certificate on DART and import to the client for FTP Secure SSL feature

## **SMB2 SUPPORT:**

### New features:

#### **Command Crediting**

→Client requests simultaneous operations; Server grants based on resources--protects against replay attacks & number of outstanding requests; Server can set Client credit to 0 to preserve resources; credits allow parallel operations such as copy chunk on Server side

#### **Asynchronous Command Processing**

→Server can send asynchronous response if an operation will take a long time, client can opt out. Async operations are SMB2\_NOTIFY, SMB2\_CREATE, SMB2\_LOCK

#### **Command Compounding of related operations**

→bunching of client requests in single transport message; performs more extended processing per roundtrip; server also bundles responses in transport message

→Examples of compounding are Create + Query Info; Create + Query Dir + Query Dir + Close; Create + Change Notify

#### **Durable Handle**

→allows for clients to be able to sustain temporary disconnections, only granted for exclusive access to a file; cached for 16 minutes after client disconnect; tunable via DART parameter

#### **Symbolic Links**

→clients can create symbolic links similar to UNIX; target can be a file or directory, relative or absolute UNC path; but feature is disabled by default on Vista & 2008 clients, and requires Administrative privileges; stored as uxfs symbolic link on Celerra, hence can be traversed by NFS clients; deletion of a symbolic link deletes the link, not the target

#### **Server Side Copy Chunk**

→client does data copy on same server without pulling data back to client; copy data from one file system to another if on same Data Mover; Copy, Xcopy, and Robocopy will use without the backup privilege set; EMCopy will use copy chunk with backup privilege, which overrides access rights checking on sources and resources Security Descriptors

#### **Other Information:**

--Improved performance and Unicode 3.0 command set support

--Used on Vista and Windows 2008 systems, Ports 445 & 139

--Only NTSTATUS codes returned, no SMB errors

--Simplies CIFS command set from 113 to 19 commands

--Will be disabled by default, used on Netbios, Standalone, or Compnames

**\$ server\_cifs server\_x –add security=NT,dialect=SMB2 | SMB2only** (type NT1 to disable once enabled)

--One of the main features is that the use of UNIX-like symbolic links will be supported in CIFS

--Server\_cifs and server\_cifs –o audit output the best way to ID a system running SMB2

Max protocol = SMB2

--\$server\_stats server\_x –table cifs –i 3 –terminationsummary no

--SMB2 supports larger buffer sizes and can send multiple commands within the same packet

--SMB2 uses strong hash using HMAC-SHA256 (vs. MD5 used in SMB) and signs session keys of User sending packet

#### **Typical SMB2 Dialog:**

SMB2\_SESSION\_SETUP (NTLMSPP negotiate msg, using SMB1)

SMB2\_SESSION\_SETUP (NTLMSPP response)

SMB2\_SESSION\_SETUP (NTLMSPP authenticate)

SMB2\_SESSION\_SETUP (NTLMSPP response)

SMB2\_TREE\_CONNECT

SMB2\_TREE\_CONNECT (response)

SMB2\_CREATE

SMB2\_CREATE (response)

SMB2\_READ

SMB2\_READ (response)

SMB2\_CLOSE

SMB2\_CLOSE (response)

SMB2\_TREE\_DISCONNECT

SMB2\_TREE\_DISCONNECT (response)

### **SYSTEM MANAGEMENT PERFORMANCE IMPROVEMENTS:**

New objects added to cache (interface, CIFS server, device)

Prefetching to occur for cached object data, but only for User nasadmin (Mover, VDM, File System, Migration File System, Checkpoint, Volume, Share, Export, Mount)

#### **STATISTICS ENHANCEMENTS:**

Celerra Statistics Framework (perfStats) provides common API for developers, built into standard kernel, and are accessible via server\_stats command; most common type of stats are counters and facts

Four new statistics types [MinMaxAvg; Stopwatch; MinMaxAvgSw; SStopwatch]

Counters are reported in terms of rate change per second

Facts are reported as point-in-time value

#### **Default setting is for server stats to show rate of change since previous sample:**

**\$ server\_stats server\_x –su basic –i 2 –c 5 –type rate**

#### **To see just the change in value since last sample:**

**\$ server\_stats server\_x –su basic –i 2 –c 5 –type diff**

#### **To see overall change since initial sample:**

**\$ server\_stats server\_x –su basic –i 2 –c 5 –type accu**

#### **REPv2 ENHANCEMENTS:**

One-to-many replications now supported for VDM Source to (4) VDM Destinations

VDM Cascading for two hops now supported, with each session independent of the other (limit is not enforced)

Source VDM can be in loaded or mounted State (but only one specific VDM in loaded state for Replication environment)

Option to disable CRC checking for Replication sessions has been removed

#### **MULTI-BYTE SUPPORT:**

Celerra Manager supports input of native languages for Shares & Comments, Tree quota paths & Comments, supporting Unicode 3.0

#### **Supported Multi-byte interfaces:**

MMC, APL, Dart, Celerra Mgr, Command Service, CBM DLL, SMI-S/CIM-NAS, CIC, SNMP, MAC XML, MAC RPC, NMI Handling, DART AV, CAVA API, CAVA Engine, CIFS API, CBM API, NDMP, DHSM API, DVT, XML-API

Supports multi-byte data in CLI, XML-API, CCMD, and Logs

## **COGNAC 5.6 MAINTENANCE RELEASE 6 (CMR6 5.6.43.8) Feb 23 2009:**

### **Foxglove release: NS-960 Integrated, NS-960FC, NS-960iS, and NS-G8**

XBlade based on Wildcat-S architecture. First Celerra platform to use CX4 hardware for Blades, as well as the SLIC Ultraflex IO modules. Blades are dual socket, 4-core Intel 2.33GHz processors with 1.33GHz dual-buses and quad channel memory, shipping with 8GB memory, though only 4GB is usable prior to Barossa release. A DME will consist of 1-2 Blades, 2 Earthquake management switches, and 4 Power Supplies. Foxglove will support 2-4 DMEs with 1-2 Blades each, and 1-2 Control Stations.

### **NS-960 product can have 2-8 Blades, but depends on model:**

NS-960 2-4 Blades and maximum of 480 drives

NS-960FC 2-6 Blades and maximum of 960 drives for use with FC Hosts

NS-960iS supports 480 drives

NS-G8 supports 2-8 Blades

### **NS480-4 Support (4-Blade Support)**

#### **FLR Compliance (FLR-C)**

#### **FILE LEVEL RETENTION: (aka FLR-E, FLR-C, CWORM)**

**Purpose:** FLR-E currently protects against file changes from NFS, CIFS, or FTP clients, but does not prevent administrative meddling of the file system, deletion of the file system, or tampering of the system clock, hence FLR-C. Intent is to meet Storage compliance regulations, as put forth by the SEC Rule 17a-4(f). Storage characteristics should have Non-rewriteable, non-erasable format; automatic storage media recording verification for quality and accuracy; serialized original and duplicate units of storage media, with time-date stamp; capacity to download indexes and records; testimony from 3<sup>rd</sup> party vendor that storage meets the FLR compliance standards. A FLR-C compliant file system will support NFS, CIFS, and FTP access. FLR Enterprise (i.e., CWORM feature) currently supports three file states—CLEAN, WORM, EXPIRED, and FLR-C adds a fourth, APPEND\_ONLY. All files start in a CLEAN state, and ordinarily progress to WORM, then EXPIRED, though files can go from EXPIRED back to WORM when clients extend the retention life of a file. APPEND\_ONLY allows a file to transition between CLEAN and WORM states, depending on whether it is Read-Only or Writeable. WORM state can be determined by the DHSM API.

#### **FLR-C LIMITATIONS:**

→NDMP backups do not preserve the FLR attributes. Restore to a file system with file-level retention enabled. If restoring expired files, will have infinite retention date. If necessary to retain file expiration status, restore to non-FLR file system, then copy to FLR file system.

#### **FLR FILE STATES:**

CLEAN—normal file state

APPEND\_ONLY—files cannot be deleted or renamed—existing data cannot be modified, but new data can be added to a file

WORM—files cannot be deleted, renamed, modified, or appended to

EXPIRED—files cannot be renamed, modified, or appended to, but can be deleted

#### **File Level Retention (FLR) COMPLIANCE:**

--FLR will now require license activation in Celerra Manager

--FLR has four possible states: CLEAN, WORM, EXPIRED, APPEND\_ONLY

--FLR will use a new FLR software clock to prevent administrators from tampering with system clock to change FLR status

**Note:** Software clock ticks only when file system is mounted RW. Will catch-up with system time through slow adjustments.

--FLR protection to be enhanced so as to prevent administrators from destroying and restoring FLR file systems

--FLR data integrity verifies that stored data can be read back (all writes verified by reading back data)

**Note:** This can be disabled via param as it is an expensive operation

--Soft infinite retention period (means retention period can be increased) is now a hard infinite period that cannot be changed

#### **FLR-C Features:**

Tamper Proof Clock—use of a software clock vs. hardware system clock to prevent administrator tampering

**Note:** If FLR clock is behind system clock, will advance up to 138 seconds per hour

File System Protection—intended to prevent administrators from destroying protected files in a FLR-C file system

**Note:** File System contents cannot be modified or deleted by Clients, Users, Administrators when set to FLR-C

Data Verification—to provide data integrity for data written to storage by doing read back from storage—DM will panic when a failure of this integrity check occurs after two retries

Default Infinite Retention period—if specific retention time is not set, a file will get the default retention period, which is not changeable for FLR-C

Activity Logging—logging of file actions taken to a FLR\_logs directory

Append\_only Files—ability to append to a file without changing existing data. File can be WORM or APPEND\_ONLY state.

#### **Creating FLR-C or FLR-E File System from CLI (Can also do from GUI):**

\$ nas\_fs –name flr\_1 –type uxfs –create size=10G pool=id=3 storage=SINGLE worm=compliance (enterprise) –option slice=y, mover=server\_2

#### **FLR Parameters:**

# server\_param server\_2 -facility FLRCCompliance -list

server\_2 :

param\_name facility default current configured

ActivityLogRP FLRCompliance 7 7

writeverify FLRCompliance 0 0

**\$ server\_param server\_2 –facility FLRCompliance –modify –value writeverify 1** (1 Enables Write Verify, default is off, 0)

#### Verifying File System FLR-C:

# nas\_fs -i file5

id = 28

name = file5

acl = 0

in\_use = True

type = udfs

**worm = compliance with no protected files**

worm\_clock= Fri Nov 13 10:45:33 EST 2009

worm Max Retention Date= No protected files created

**# /nas/tools/\_fs\_db server\_2 readsb 97 psb**

About to read superblock from volume 97 and sector 16

ncyl=263

magic=**0xdcbaabc** →This magic value means that the file system is set to FLR-C mode

#### Restoring Checkpoint to FLR-E File System:

**\$ /nas/sbin/rootfs\_ckpt flr1\_ckpt1 -name flr1\_ckpt2 -Restore -Force**

#### FLR Toolkit Version 3.0:

--A windows application that can be used to monitor, query, report, and verify status of FLR file systems

Dashboard—snapshot of CIFS share with report, charts, and showing when files will expire

Report Generator—detailed reports on CIFS Shares

Query Builder—searching for files on CIFS Shares

Monitor Service—monitor pathnames and auto set retention dates in directories

--can use CLI to apply retention dates or query files from CIFS shares

#### FLR Toolkit Version 3.4:

→Available from Powerlink Celerra Tools page as an installable: EMC\_FLR\_Toolkit\_v3.4.0.zip

→Windows utility designed to manage and report on Celerra File Retention at file system, directory, & file levels

→Explorer-based monitoring and reporting capability, and some CLI querying and management capability as well

→Toolkit does not support Append-only files, or FLR-C file systems

→Program Files\EMC\FLRToolkit\FLRService\FLRMonitorService.log

#### Flare 28 Mira Support:

EFD solid state drive support will be on Trident & Dreadnought arrays only, for 73GB 520BPS 4GB FC interfaces; support for 400GB EFD drive. New AVM pool, clarssd\_r5 with new disk type “CLSSD” configured in RAID 5 4+1 or 8+1

**Note:** AVM rules do not allow mixing disk types in a pool or extending file systems from one pool to a different pool type

#### General SSD Rules:

--Used only in RAID 5 4+1 or 8+1 configurations

--Not used as Vault drives

--Not recommended for use as Celerra Control LUNs in case of Gateway systems

--Drive type cannot be mixed within same DAE or within a RAID group

--Hot Spare can only be SSD for SSD

--SSD drives are valid for CX4-960 and CX4-480 arrays

--one Hot Spare recommended for every 30 drives

**Note:** No mention at all about best practice cache settings Celerra (write enabled, read disabled) vs. Clariion (both write & read cache disabled)

#### Celerra Data Deduplication/File-level Dedup and Compression (Celerra F-RDE; Celerra Dedupe, etc.):

##### “File Level Deduplication (single instancing) and File Compression”

→Key attributes of Celerra Deduplication are ‘single file instancing’ and ‘file compression’

##### Purpose:

File Level Redundant Data Elimination (F-RDE) seeks to eliminate duplicate file instances and performs compression (50%), providing for storage efficiency. For example, 70 file copies will have 70 inodes with metadata, but the actual file data will be a single file. Since this feature will have a performance impact, the targeted files will be those that are relatively inactive, with an access age threshold.

##### Operation:

Based on last access time and modification times, DART will scan file systems for eligible files and then perform the R-RDE function by putting file data in hidden F-RDE store (EDRS store) as a stub file using FileMover technology. The “name” of a deduped file is not hidden and looks online to the User (hence de-duplicated). Hidden deduped files are placed in the slash etc store of the file system and are not visible to Clients. The data is “compressed” on the way into the ‘store’ and “uncompressed” on the way out of the ‘store’.

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
Reads are done in pass-through mode without recalling the data, while writes and other administrative actions will perform the recall of data.

#### **DEDUPLICATION FEATURE:**

- File Level Deduplication does not require a lot of CPU or Memory, and provides about 10% space savings, while the compression component requires lots of CPU, but improves space savings by up to 50%
  - Note:** This is because deduplication uses a background, asynchronous operation that operations on inactive files
  - Files are selected based on Size and Age (accessTime & modificationTime)
  - Deduplication uses an internal policy engine, a Compression engine based on EMC RecoverPoint, and a Deduplication engine based on EMC Avamar hashing
  - Deduplication is based on file data, NOT metadata
  - On average, compression offers a 50% space savings
  - Policy Engine on Data Mover scans file system for de-dupe candidates (up to a million files in 3 minutes using 5% CPU) as a background asynchronous operation
  - Policy Engine uses 1 thread for deduplication and 1 thread for reduplication (recall)
  - Policy Engine logs exceptions and deduplication completion in Server Log under DEDUPE facility
  - Dedupe candidates are sent to the EDRS store for deduplication, compressed, then given a unique handle for each file stored
  - Policy Engine puts the file offline using FileMover and the unique file handle—original file replaced with stub file that references the data in the EDRS store
- Note:** RDE stub files will not display in CIFS as offline (until NAS 5.6.47)
- Dedupe works with Replication V2 file systems
  - Single file instancing (deduplication) can be disabled by changing the param singleInstancingEnabled to 0, which allows the compression feature to continue to run
  - From Celerra Manager, turn Deduplication On, Off, or Suspended, see Status, and see Statistics for deduplication (all other tasks require CLI)
  - Reduplication is the process of recalling deduplicated files from the EDRS store, such as when turning “Off” dedupe, but command will not execute unless the file system has enough space to allow. Or, if a client modifies or writes to a file, that file will be reduplicated and uncompressed
  - With the original Phase I version (5.6.43) of Deduplication, write operations to a deduplicated file leads to a full recall of the file, and write latency

#### **How Does Single File Instancing Work?**

→ When files are compressed, a hash is derived and applied to the file data, then compared to the existing hidden Data Store. For those files where the contents are identical, the hash will be the same. So, if the hash is found to already exist in the Data Store, a Stub file is placed in the file system to replace the data, hence we get “single file instancing” [aka, ‘file level deduplication’] of files with the same exact data content.

#### **Celerra Manager Deduplication Properties:**

Files scanned: 70  
Files deduped: 70 (100% of total files)  
Original data size: 503 MB (102% of current file system capacity)  
Space saved: 428 MB (85% of original data size)

#### **F-RDE LIMITATIONS:**

- Replication V1 not supported with deduplicated file systems
  - File systems running 5.6.47 deduplication cannot be mounted on pre-5.6.47 code due to a magic number change
  - Deduplication will not occur on paths >than 1024 bytes in length, creating errors [See emc231579—to be fixed 5.6.49]
  - 2010-01-08 15:23:43: DEDUPE: 3: 2: Deduplication of file system id 846 was aborted after 10 failure(s), 121571720 KB available on FS. Last reported error was: NameTooLong.
  - Files 24K and below are not deduplicated or compressed by default, though lower values can be specified via param
  - By default, system will abort dedupe rather than extend SavVol for Snapsure
  - By default, files >than 200MB will not be deduplicated for CIFS timeout reasons
  - iSCSI LUNs cannot be deduplicated
  - The F-RDE deduplication feature can be run in conjunction with Celerra FileMover
  - Writes to de-duped files initiates recall
  - NDMP (NVB) volume-based backups are supported, but only with full restores
- Note:** Downside here is that single file restores are not supported from NVB. Instead, for single file backup and restores, use Checkpoints.
- PAX NDMP & Network backups will reduplicate & uncompress the files for the backup application, though the files themselves are left deduplicated & compressed on the Celerra file system. Because of this, backup performance will be impacted. Also, restoring files from regular NDMP or Network backups will be restored in a non-deduplicated state (overwrites what might be deduplicated files?).
  - Over 64k instances of the same file will not be deduped, but will be compressed
  - FSCK cannot detect/repair stub files for RDE store

→For MPFS file systems, IO to any files in the RDE store will revert to NFS or CIFS

→Does not support Writeable Checkpoint feature

→Deduplicated blocks are not counted in Quotas (recommendation to use ‘File Size’ Quotas)

→Does not support CDMS mgfs file systems

→Deduplication scanning will abort after (10) errors are encountered. This is default behavior and the error thresholding can be increased to a max of 1000 errors:

### **param dedupe peMaxErrors=1000**

#### **THREE STATES FOR DEDUPLICATION:**

on (deduplication on for a file system)

off (deduplication turned off and file system undo’s any deduplicated files—called “reduplication”)

suspended (deduplication suspended)

#### **Commands:**

**# fs\_dedupe -list | -info | -all | <fs\_name> or <fs\_id>| -modify | -state {off | suspended | on}**

**Note:** Setting state to on implements Policy Engine. Setting state off recalls all files in RDE store. Setting state to suspended means Policy Engine will not look for dedupe candidates on the file system

**# fs\_dedupe -modify fs5 -state suspended**

Done

**Note:** Use the following command to force dedupe suspension, as in a situation where the dedupe task may not be suspended and the file system is unmounted:

**# .server\_config server\_2 -v "pe action=disable fsid=<file system ID>"**

**# fs\_dedupe -info fs5**

Id = 29

Name = fs5

Deduplication = Suspended

As of the last file system scan (Mon Feb 16 13:19:15 EST 2009):

Files scanned = 70

Files deduped = 0 (0%)

File system capacity = 1008 MB

Original data size = 646 MB (64% of current file system capacity)

Space saved = 0 MB (0% of original data size)

**# fs\_dedupe -info fs5**

Id = 29

Name = fs5

Deduplication = Off

**# fs\_dedupe -modify fs5 -state on**

Done

**# fs\_dedupe -list**

| id | name | state | status | time_of_last_scan            | original_data_size | usage | space_saved |
|----|------|-------|--------|------------------------------|--------------------|-------|-------------|
| 29 | fs5  | On    | Idle   | Mon Feb 16 13:19:15 EST 2009 | 646 MB             | 64%   | 0 MB (0%)   |

**# fs\_dedupe -info fs5**

Id = 29

Name = fs5

Deduplication = On

Status = Idle

As of the last file system scan (Mon Feb 16 13:19:15 EST 2009):

Files scanned = 70

Files deduped = 0 (0%)

File system capacity = 1008 MB

Original data size = 646 MB (64% of current file system capacity)

Space saved = 0 MB (0% of original data size)

#### **TURNING DEDUPLICATION OFF RE-DUPLICATES ALL FILES:**

**Note:** But only if there is room to do so, as seen in the following—had to extend the file system to turn it off

**# fs\_dedupe -modify cifs1 -state off**

Error 12048: Additional space 57 MB needed on this file system to perform this operation.

**# fs\_dedupe -modify cifs1 -state off**

Done

**# fs\_dedupe -info cifs1**

Id = 32

Name = cifs1  
 Deduplication = Off  
 Status = Re-duplicating(28% complete)

### **DEDUPLICATION STATUS:**

→Using fs\_dedupe -info, the various states for dedupe are Status = Scanning; Status = Idle; State = Reduplicating

#### **# server\_param server\_2 -facility dedupe -list**

server\_2 :

| param_name               | facility | default | current | configured                                                                                                |
|--------------------------|----------|---------|---------|-----------------------------------------------------------------------------------------------------------|
| fileExtensionExcludeList | dedupe   | "       | "       |                                                                                                           |
| singleInstancingEnabled  | dedupe   | 1       | 1       |                                                                                                           |
| minimumSize              | dedupe   | 24      | 24      | →Min size that file will be deduped & compressed, default is 24k                                          |
| savVolThreshold          | dedupe   | 90      | 90      |                                                                                                           |
| accessTime               | dedupe   | 30      | 30      | →Files accessed within this number of days will not be deduplicated                                       |
| maximumSize              | dedupe   | 200     | 200     | →largest file size in MB that will be deduplicated, default is 200MB to avoid CIFS write timeout failures |
| caseSensitive            | dedupe   | 0       | 0       |                                                                                                           |
| throttle.cpuLowTrigger   | dedupe   | 40      | 40      |                                                                                                           |
| throttle.cpuHighTrigger  | dedupe   | 75      | 75      |                                                                                                           |
| minimumScanInterval      | dedupe   | 7       | 7       | →number days between scans, default 7 days                                                                |
| modificationTime         | dedupe   | 60      | 60      | →Number of days before file will be deduplicated without mod. time change                                 |

### **TESTING DEDUPLICATION MANUALLY (Single file instancing and compression):**

**Note:** With the default parameter settings, the first time any files would be deduped on a file system would be during the first scan that occurs after the 30 day accessTime setting, and then only if the file had not been accessed within the 30 day accessTime, or modified within the 60 day modificationTime window. In order to test this feature, we will temporarily disable the accessTime and modificationTime parameters.

1. Create a CIFS Server, and a file system with Shares that are accessible from a Windows client on the domain
2. From the Windows client, copy a set of identical files to multiple directories on the Celerra file system
3. For NAS Versions 5.6.43 to 5.6.46, use the server\_param facility to disable the AccessTime & ModificationTime on a Data Mover basis by setting each param to 0:

| param_name       | facility | default | current | configured |
|------------------|----------|---------|---------|------------|
| accessTime       | dedupe   | 30      | 30      | 0          |
| modificationTime | dedupe   | 60      | 60      | 0          |

**Example:**

#### **# server\_param server\_2 -facility dedupe -modify modificationTime -value 0**

server\_2 : done

#### **# server\_param server\_2 -facility dedupe -modify accessTime -value 0**

server\_2 : done

**Verify changes:**

#### **# server\_param server\_2 -facility dedupe -list|grep Time**

|                  |        |    |   |
|------------------|--------|----|---|
| accessTime       | dedupe | 30 | 0 |
| modificationTime | dedupe | 60 | 0 |

4. With NAS 5.6.47 and later, change the parameter values on a per file system basis (i.e., only on the file system used for the testing):

#### **# fs\_dedupe -modify file2 -access\_time 0**

#### **# fs\_dedupe -modify file2 -modification\_time 0**

**Verify changes:**

#### **# fs\_dedupe -info file2**

5. Verify file system usage prior to turning Deduplication scanning on:

#### **# server\_df server\_2 file2**

server\_2 :

|            |         |         |        |          |            |
|------------|---------|---------|--------|----------|------------|
| Filesystem | kbytes  | used    | avail  | capacity | Mounted on |
| file2      | 2519984 | 2199880 | 320104 | 87%      | /file2     |

6. To turn deduplication on, or to initiate a new scan immediately, use the "# fs\_dedupe -modify file2 -state on" CLI command. From Celerra Manager, you would first "Suspend" and then turn back "On" deduplication in order to initiate a new scan, or just "On" if deduplication was never turned on:

**Note:** Scanning will take place and deduplication and compression of qualified files will occur

#### **# fs\_dedupe -modify file2 -state on**

Done

7. Use fs\_dedupe to monitor the progress of the file system deduplication scan operation:

**# fs\_dedupe -info file2**

```
Id          = 25
Name        = file2
Deduplication = On
Status      = Scanning (0% complete)
```

As of the last file system scan (N/A):

```
Files scanned    = 113
Files deduped   = 83 (0% of total files)
File system capacity = 0 MB
Original data size = 476 MB (0% of current file system capacity)
Space saved     = 0 MB (0% of original data size)
```

-----output abbreviated-----

8. When the fs\_dedupe -info reports the deduplication scan status as "Idle", deduplication scanning has completed. Review the results using server\_df and fs\_dedupe:

**# server\_df server\_2 file2**

server\_2 :

| Filesystem | kbytes  | used           | avail          | capacity   | Mounted on |
|------------|---------|----------------|----------------|------------|------------|
| file2      | 2519984 | <b>1135984</b> | <b>1384000</b> | <b>45%</b> | /file2     |

**# fs\_dedupe -info file2**

```
Id          = 25
Name        = file2
Deduplication = On
Status      = Idle
```

As of the last file system scan (Wed Dec 2 16:22:59 EST 2009):

```
Files scanned    = 265
Files deduped   = 200 (75% of total files)
File system capacity = 2460 MB
Original data size = 2143 MB (87% of current file system capacity)
Space saved     = 1034 MB (48% of original data size)
```

-----output abbreviated-----

9. After testing has been completed, you can restore the accessTime and modificationTime parameters to the default values, 30 days &amp; 60 days, respectively, or to whatever other value is desired.

**DISABLING SINGLE FILE INSTANCING (Deduplication):****Note:** This disables “deduplication” but still allows “compression” to work**# server\_param server\_2 -facility dedupe -info singleInstancingEnabled**

server\_2 :

```
name        = singleInstancingEnabled
facility_name = dedupe
default_value = 1
current_value = 1
configured_value =
user_action  = none
change_effective = immediate
range       = (0,1)
```

description = This parameter is used to enable or disable single instancing capabilities of the Celerra deduplication feature.

→Set singleInstancingEnabled to 0 to disable this feature, while retaining the compression feature

**REDUPLICATING FILES:**

→Files are reduplicated if Clients modify or write to them

→Also, turning the dedupe feature “Off” will reduplicate all files to the file system, but requires enough space to do so

**# fs\_dedupe -modify cifs1 -state off**

Error 12048: Additional space 78 MB needed on this file system to perform this operation.

**WHAT HAPPENS TO DEDUPLICATION IF FILE SYSTEM FULL?**

→Cannot turn off deduplication feature

→Cannot duplicate anymore files

→Deduplication feature remains set to “on” for the file system, but is “Idle”. Further scanning will not occur until file system space is provided.

### Server Log:

2009-07-07 12:56:25: DEDUPE: 3: PE: ProcessFile failed on file /cifs1/deduped/dir3/DX9-10-06/Apr2006\_d3dx9\_30\_x86.cab with error: NoSpace  
2009-07-07 12:56:25: DEDUPE: 3: 2: Deduplication of file system id 32 was aborted after 10 failure(s), 88 KB available on FS. Last reported error was: NoSpace.

2009-07-07 12:56:25: DEDUPE: 6: PE: Task 1 aborted scan on fsid 32, duration: 0 seconds

2009-07-07 12:56:25: DEDUPE: 6: 5: The deduplication scan on file system id 32 has just completed.

**Note:** Ironically, the Log says the scan completed, but in reality, it has been aborted.

### Celerra Manager records Error:

X Slot 2: Deduplication of file system id 32 was aborted after 10 failure(s), 88 KB available on FS. Last reported error was: NoSpace.

### Resolution:

1. Clicked on “Auto Extend Enabled” in Celerra Manager for the file system (or manually extend the file system)
2. From Celerra Manager file system properties page, under the ‘Deduplication’ section, click on “Suspended”, “apply”, then click on “On”, then ‘apply’. From the CLI, issue the following command:

# **fs\_dedupe -modify cifs1 -state on**

Done

**Note:** This will initiate a new scan of the file system

# **fs\_dedupe -info cifs1**

|               |                           |
|---------------|---------------------------|
| Id            | = 32                      |
| Name          | = cifs1                   |
| Deduplication | = On                      |
| Status        | = Scanning (35% complete) |

As of the last file system scan (Tue Jul 7 10:04:36 EDT 2009):

|                      |                                                |
|----------------------|------------------------------------------------|
| Files scanned        | = 53                                           |
| Files deduped        | = 70 (46% of total files)                      |
| File system capacity | = 984 MB                                       |
| Original data size   | = 503 MB (51% of current file system capacity) |
| Space saved          | = 11 MB (2% of original data size)             |

# **fs\_dedupe -info cifs1**

|               |         |
|---------------|---------|
| Id            | = 32    |
| Name          | = cifs1 |
| Deduplication | = On    |
| Status        | = Idle  |

As of the last file system scan (Tue Jul 7 13:09:12 EDT 2009):

|                      |                                                |
|----------------------|------------------------------------------------|
| Files scanned        | = 188                                          |
| Files deduped        | = 188 (78% of total files)                     |
| File system capacity | = 984 MB                                       |
| Original data size   | = 920 MB (93% of current file system capacity) |
| Space saved          | = 741 MB (80% of original data size)           |

## **FILTERING USING EXCLUSIONS:**

### Excluding File Extension Types:

\$ **server\_param server\_x -facility dedupe -modify fileExtensionExcludeList -value .mp3**

### Excluding Directories:

\$ **server\_param server\_x -facility dedupe -modify dirExcludeList -value temp\_dir**

### Disabling Single File Instancing:

\$ **server\_param server\_x -facility dedupe -modify singleInstancingEnabled -value 0** (default is 1)

### Hidden Parameters:

dedupe rdeLoggingEnabled →Controls logging of RDE stub file events to /.etc/dhsm.log

dedupe fileExtensionIncludeList →Can dedupe specific file extensions via colon delimited list

### For Client IO Errors when accessing deduped files?

--Users will need to restore from backup

--Recovery Tool can generate list of files that need to be restored from backup and also to verify status of deduped files in file system

**Note:** Need to obtain RecoveryTool from engineering (CHECK\_DEDUP)

# **.server\_config server\_2 -v "sspeache"**

## **CONTROLLING WHEN DEDUPLICATION OCCURS:**

1. Deduplication will occur when first enabling the system for deduplication (if timestamps and rules allow), and thereafter, as per the Rules of the Policy Engine
2. If deduplication scanning is occurring during a bad timeframe (middle of working day), simply turn on the dedupe session again for the file system at the desired timeframe (-modify -state on)

## **TROUBLESHOOTING DEDUPLICATION:**

**# .server\_config server\_2 -v "pe action=**

```
action=enable fsid=n
action=disable fsid=n
action=off fsid=n
action=start fsid=n
action=stop fsid=n
action=pause fsid=n
action=resume fsid=n
action=status fsid=n
action=free fsid=n
action=trav fsid=n
action=rmdb fsid=n
action=listfs
```

**# .server\_config server\_2 -v "pe action=listfs"**

```
1246911461: UFS: 7: inc ino blk cache count: nInoAllocs 1: inoBlk df674a84
1246911461: DEDUPE: 6: -----
1246911461: DEDUPE: 6: FSid = 32
1246911461: DEDUPE: 6: DB Version = 1
1246911461: DEDUPE: 6: RDE State = Enabled
1246911461: DEDUPE: 6: Last Enabled = Mon Jul 6 16:14:24 2009
1246911461: DEDUPE: 6: MntPnt = /cifs1
1246911461: DEDUPE: 6: fsSize = 503984
1246911461: DEDUPE: 6: fsFree = 428152
1246911461: DEDUPE: 6: fsFiles = 65
1246911461: DEDUPE: 6: Fv = 50
1246911461: DEDUPE: 6: KBv = 369056
1246911461: DEDUPE: 6: Fm = 50
1246911461: DEDUPE: 6: KBm = 369016
1246911461: DEDUPE: 6: Last Scanned = Mon Jul 6 16:14:24 2009
1246911461: DEDUPE: 6: Last Scan Duration = 15
1246911461: DEDUPE: 6: Valid = Valid
1246911461: DEDUPE: 7: There are 1 tasks on the queue: head: 0xdf5cd004, tail: 0xdf5cd004
1246911461: DEDUPE: 7: -----
1246911461: DEDUPE: 7: Task 32, state: IDLE, sequence: 1246910423, address: 0xdf5cd004
1246911461: DEDUPE: 7: -----
1246911461: ADMIN: 6: Command succeeded: pe action=listfs
```

#### Example of Aborted Dedupe Scan Operation:

**\$ nas\_logviewer /nas/log/sys\_log|grep aborted**

```
Jul 9 20:29:06 2010:DART:DEDUPE:ERROR:2:Slot 2:::1278725346:Deduplication of file system id 29 was aborted after 10
failure(s), 374225912 KB available on FS. Last reported error was: SavVolAtCapacity.
```

**\$ .server\_config server\_2 -v "pe action=status fsid=29"**

```
1279029088: DEDUPE: 6: -----
1279029088: DEDUPE: 6: FSid = 29
1279029088: DEDUPE: 6: Task State = IDLE
1279029088: DEDUPE: 6: RDE State = Enabled
1279029088: DEDUPE: 6: Prev: fsSize = 2500902152
1279029088: DEDUPE: 6: Prev: fsFree = 374225912
1279029088: DEDUPE: 6: Prev: fsFiles = 1686413
1279029088: DEDUPE: 6: Prev: Fv = 191793
1279029088: DEDUPE: 6: Prev: KBv = 436308864
1279029088: DEDUPE: 6: Prev: Fm = 114908
1279029088: DEDUPE: 6: Prev: KBm = 425508776
1279029088: DEDUPE: 6: Prev: Scan Time = Fri Jul 9 11:20:29 2010
1279029088: DEDUPE: 6: Prev: Scan Duration = 32917
1279029088: DEDUPE: 6: Prev: Valid = Aborted
1279029088: DEDUPE: 6: Curr: fsSize = 0
1279029088: DEDUPE: 6: Curr: fsFree = 0
1279029088: DEDUPE: 6: Curr: fsFiles = 0
1279029088: DEDUPE: 6: Curr: Fv = 0
```

1279029088: DEDUPE: 6: Curr: KBv = 0  
1279029088: DEDUPE: 6: Curr: Fm = 0  
1279029088: DEDUPE: 6: Curr: KBm = 0  
1279029088: DEDUPE: 6: Curr: Scan Time =  
1279029088: DEDUPE: 6: Curr: Scan Duration = 0  
1279029088: DEDUPE: 6: Curr: Valid = Stopped  
1279029088: ADMIN: 6: Command succeeded: pe action=status fsid=29

**EXPOSE SLASHETC TO SEE DEDUPE DATA FILE:**

# .server\_config server\_2 -v "shadow showetc <filesystem\_name>"

# cd slashetc  
[root@fox1 slashetc]# ls  
ACLdata ACLrecord gid\_map pedb quotas.config ufsasyncDB

# cat pedb  
1:1:1246911264:503984:427936:70:70:515936:70:515880:1246975476:15:1

**HIDE SLASHETC:**

# rm -rf slashetc

**VALIDATING STUB FILE WITH EDRS REPOSITORY:**

\$ .server\_config server\_3 -v "dhsm action=query\_offlineattrs path=/sparse/dense/cifs1/96\_803\_48\_1"

1241620022: KERNEL: 9: >>abce1 getting first hold on c62df4a8, nActive = 100139

1241620022: KERNEL: 9: last message repeated 1 times

1241620022: DHSM: 6: Offline attributes:

1241620022: DHSM: 6:

OFFLINE\_PATH=dart://rde/<**1969,4a08b297**>AQAAAAA/190/aa/AadBdxj3CPb8p6rsZX0wvjchdpt2AAAAAAAAAAAAAA-8AAAAAAA1MAA (inumber and generation count of file in hex—convert inumber ‘1969’ to decimal and match to stub file inode)

1241620022: DHSM: 6: OFFLINE\_MTIME=<invalid>

1241620022: DHSM: 6: INFO=

1241620022: DHSM: 6: READ\_METHOD=<NONE>

1241620022: DHSM: 6: PE\_ID=PE V1.0

1241620022: DHSM: 6: ORIGINAL\_BYTES\_USED=876544

1241620022: DHSM: 6: DEDUPE\_LINKCOUNT=2

1241620022: DHSM: 6: DEDUPE\_FSIZE=151612

**DEDUPE SAVVOL FULL ISSUE:** emc210954

→By default, deduplication is set to stop when 90% of the SavVol has been consumed, but in practice, deduplication may stop just before the 90% threshold is reached.

→The SavVol parameter “savVolThreshold” may need to be revised to 100% or turned off (set to 0) if the intent is to keep deduplication active without concern for the SavVol capacity.

# server\_param server\_2 -facility dedupe -modify savVolThreshold -value 100

# fs\_ckpt id=29 -l  
id ckpt\_name creation\_time inuse fullmark total\_savvol\_used ckpt\_usage\_on\_savvol  
32 checkpoints1 04/01/2009-22:47:46-EDT y 90% 89% 0%

**Server Log:**

DEDUPE: 3: 2: Deduplication of file system id 29 was aborted after 10 failure(s), 2653147840 KB available on FS.

Last reported error was:SavVolAtCapacity.

DEDUPE: 6: PE: Task 1 aborted scan on fsid 29, duration: 98 seconds

**ROUTINE SERVER LOG ENTRIES:**

2009-07-06 16:08:00: DEDUPE: 6: PE: Task 1 starting scan on fsid 32

2009-07-06 16:08:00: DEDUPE: 6: StoreHost: Created repository root on file system 32

2009-07-06 16:08:01: DEDUPE: 6: Domain not found for fsid 32, creating.

2009-07-06 16:08:01: DEDUPE: 6: Domain initialized for fsid 32

2009-07-06 16:08:07: DEDUPE: 6: PE: Task 1 finished scan on fsid 32, duration: 7 seconds

2009-07-06 16:08:17: DEDUPE: 6: 5: The deduplication scan on file system id 32 has just completed.

**Using Celerra Data Deduplication Tech Module:**

‘By default the system is configured to abort deduplication operations on a file system before it causes the SavVol to extend. This avoids the SavVol expanding due to deduplication activity.’

**Hiding NFS Exports:**

--Ability to hide exports from clients via export process (see param mount.forceFullShowmount)

--Client that tries to access a hidden export that they do not have export access for will see “does not exist”

**DHSM Enhancement:**

--Scripts enhanced to handle file names with ampersand, >, <, and apostrophe characters

## **Increased UNIX Group Support:**

--Group support increased to 128 UNIX groups for O/S types, such as AIX and Redhat, that can support more than 16 UNIX groups in the RPC credential (NFS uses AUTH\_UNIX authentication mechanism, traditionally set at 16 groups maximum)

--Celerra now can store up to 128 UNIX groups, but requires setting the maxgroups param

# **mount -o maxgroups=64 server\_2:/mntpoint /fs1**

**Note:** Mount example from AIX client

# .server\_config server\_2 -v “param security maxgroups” (maxgroups=128 to change to 128 group support)

security.maxgroups INT 0x02dbd738 16 16 (16,128) TRUE REBOOT ‘Define the max number of extra group in Unix credential’

# **server\_param server\_2 -facility security -info maxgroups -v**

server\_2 :

```
name      = maxgroups
facility_name = security
default_value = 16
current_value = 16
configured_value =
user_action = reboot DataMover
change_effective = reboot DataMover
range     = (16,128)
description = Define the max number of extra group in Unix credential
detailed_description
```

In Unix a user can belongs to many groups, when a user is authenticated a credential is created, where this information is stored. By default Celerra is supporting only 16 groups stored in the credential for authentication. This parameter allows to extend the number of groups we are storing to 128. This is also impacting the credential passed to Celerra in NFS requests in the RPC. Bumping this parameter is allowing to support clients able to set more than 16 groups in the RPC credential, like AIX or RedHat.

## **SYR Enhancements:**

--A new script to Data Mover system usage patterns on customer boxes has been added to /nas/tools/.sys\_stats

--Purpose is to collect CPU & Memory statistics and deliver via the Monthly SYR callhome when log\_config is run. Log\_config converts the histogram data into XML format

**Note:** .syr\_stats is not intended to be run manually, just by Cron Jobs and when log\_config runs monthly

# **/nas/tools/.syr\_stats -report**

usage: /nas/tools/.syr\_stats [ -update | -init | -report [cpulmemorycacheduncachederrors] | -help ]

This script is used by SYR to collect statistical data and should not be run manually

--There is also a new cron job to collect CPU & Memory utilization on Data Movers every 5 minutes. The information collected by the .syr\_stats -update Cron is placed in a histogram file stored in the /nas/var/log directory:

### **Histogram File:**

**/nas/var/log/syr\_stats**

### **Cron Job:**

/nas/site/cron.d/nas\_sys

**0-59/5 \* \* \* \* root /nas/tools/.syr\_stats -update > /dev/null 2>&1**

--Feature is enabled on new installations when log\_config -start is called, which creates the Cron Job

--For NAS Upgrades, log\_config -status checks to see if CallHome & the Cron Job for log\_config is configured. If configured, then the SYR collection is considered “Enabled”, and log\_config will restart to enable the new stats collection tool

**41 3 2 \* \* \* root /nas/sbin/log\_config -d -c >/dev/null 2>&1**

### **How the svr\_stats script is setup during Install or Upgrades:**

# **/nas/sbin/log\_config -start -all**

Adding entry to /etc/cron.d/nas\_sys

41 3 2 \* \* \* root /nas/sbin/log\_config -d -c >/dev/null 2>&1

Adding stats entry to /etc/cron.d/nas\_sys

0-59/5 \* \* \* \* root /nas/tools/.syr\_stats -update > /dev/null 2>&1

Creating Data Mover stats log file

### **Symptom when log\_config & sys\_stats CRON Jobs are not present:**

# **/nas/sbin/log\_config -status**

disabled

### **PUHC Warning—emc231857:**

Data Movers : Checking for excessive memory utilization..... Warn

Data Movers: Check for excessive memory utilization

Symptom: The script /nas/tools/.syr\_stats cannot be found on  
the system.

### **PAHC or Post-Install Warning—emc232647:**

Data Movers: Check for excessive memory utilization

Symptom: Cannot determine the memory usage status.

Action : Run "/nas/tools/.syr\_stats -report cached" manually to investigate further.(abridged)

## **COGNAC 5.6 MAINTENANCE RELEASE 7 (CMR7 5.6.44.4) Mar 2009**

### **DYNAMIC SRDF:**

→Allows DR site to mirror with Source after failover, with the idea to prevent long down-time during restore operations (due to synchronization between RDF volumes) and to provide better disaster protection in case DR site experiences failure while in an activated state

→Bin file changes: Enable RDF dynamic mode; Set Dynamic RDF flag on Celerra Volumes

**# /nas/sbin/nas\_rdf –activate -reverse**

**Note:** So, after failover, data is being mirrored from the R2 side to the R1 side. As data blocks are written to the R2, they are also synchronized back to the original source, or R1 side

### **> 2TB LUN SUPPORT:**

→Support for LUNs greater than 2TB on NS-960 and Gateways, which translates into 16TB luns

→Celerra Integrated NS-960 platforms create only a single LUN per Raid Group if using the setup\_clariion script

**Note:** With NAS 5.6.45, the Celerra Manager Provisioning Wizard is also capable of creating >than 2TB LUNS on ATA drives

→All other Integrateds will fallback to creation of dual luns per RG and <than 2TB luns when using setup\_clariion. You must specify a storage profile when using setup\_clariion that will be capable of creating large luns, such as the following example:

**# /nas/sbin/setup\_clariion -init**

The following 7 template(s) available:

1. ATA\_RAID5\_HS\_6+1\_6+1

2. ATA\_RAID5\_HS\_4+1\_4+1\_HS\_HS\_HS\_HS

3. ATA\_RAID5\_HS\_4+1\_8+1

4. ATA\_RAID10\_HS\_R10\_R10\_R10\_R10\_R10\_R10\_R10\_R10

5. ATA\_RAID6\_HS\_4+2\_6+2

6. ATA\_RAID6\_HS\_12+2 →This template yields a 10.7TB LUN from a shelf of 1TB SATA drives

Enclosure 1\_0.

Created disk group 8, luns 17

**Note:** I created a 5TB file system on the 10.7TB LUN, and it took 50 minutes until the file system was converted from rawfs and became mounted on the Server

→Gateways will support >than 2TB luns

### **POST & BIOS UPDATE:**

→NS-960 Integrated & Gateway will flash new firmware during NAS Upgrade or Install: POST 7.30 & BIOS 4.60

**Note:** emc208053 documents an issue with firmware in which the Blade may fail to boot (AR137603—fixed with BIOS 4.60 POST 7.30). If this happens when running the CSA, the error seen would be 1004021000 for Blade 2 (Blade 2 is booting up (rc=0), 1004031000 for Blade 3, etc.

Workaround: Reseat the Blade, but first make sure that the white “hand” LED is not lit, which would indicate an unsafe condition, with the blade doing a firmware update, etc.

### **SAVVol:**

SavVol default extension & creation size increased to 20GB

## **COGNAC 5.6 MAINTENANCE RELEASE 8 (CMR8 5.6.45)—July 2009**

→NS-960 CSA Provisioning Support will be available with a later NAS release

→Dell NX4 OEM Phase II support

→Celerra Gateway qualification for Tigon Symmetrix backend (V-Max hardware), but not for any new features it may present, with support for microcode 5874.1xx

→Support for Libra FLARE (04.28.000.5.704) for Gateway and Integrateds

### **CSA PROVISIONING WIZARD (CSA CPW):**

→The CPW functionality has been around for some time now, but has been restricted to EMC and Partner use only, and made available via a separate Powerlink download

→With NAS 5.6.45.5, the CSA with the Provisioning Wizard functionality becomes part of the mainstream NAS code release, and will be made generally available for customers

→The platforms where the CSA with CPW functionality has been qualified, are the NX4, NS20, NS40, NS-120, NS-480, NS-960

**Note:** You can also use the Celerra Manager Provisioning Wizard functionality to accomplish the same storage provisioning as that found in the CSA utility. Also note that >than 2TB LUNs is capable for NS-960 platforms when using Celerra Mgr provisioning.

### **MODEL ORDERING CHANGES:**

**Interim Model Ordering System--July:**

→In July, the Order Entry system for NS-480/NS-960 platforms will present one base array model, dropping the original AUXF & AUXI names: **NS480-AUX or NS960-AUX**

→Customers would add FC or iSCSI slices, to get the desired configuration

#### **New Model Ordering System:**

→In August, the Order Entry system for the NS-960 will have two base models:

**NS-960AUX or NS-960AUX8** (AUX8 indicates the use of the new 8Gbps Glacier FC IO Module)

#### **NS-960 Rules:**

→One Base AUX array will have (2) 4-port 4Gbps Tomahawk FC IO Module

→Another Base AUX array will have (2) 4-port 8Gbps Glacier FC IO Module

→Mfg or Upgrade can add (1) or (2) 4Gbps FC IO Modules to the array

→Mfg or Upgrade can add (1) or (2) 8Gbps FC IO Modules to the array, in lowest available slot position before 4Gbps cards

→Mfg or Upgrade can add (1) or (2) 10Gbps iSCSI Poseidon IO Module, in lowest available slot position before 1Gbps iSCSI cards

→Mfg or Upgrade can add 1-4 1Gbps iSCSI IO Modules

→960 Drive models can add an additional 4Gbps FC IO Module for BE connectivity

SP slot 0 & 1 must be all 4Gbps FC (AUX) or all 8Gbps FC (AUX8), and are always populated

SP slot 2 can be empty, be 4Gbps FC (when requiring >than 480 drives), or be 1/10Gbps iSCSI

SP slot 3 can be empty, be 4/8Gbps FC (for FC or MPFS over FC), or be 1/10Gbps iSCSI

SP slot 4 can be empty, be 4/8Gbps FC (5th FC for 8 Blades & more FC hosts), or be 1/10Gbps iSCSI (Mview or MPFS over iSCSI)

SP slot 5 can be empty, or have 1/10Gbps iSCSI (Mview or MPFS over iSCSI)

If MPFS over iSCSI selected, slots 2, 3, 4, & 5 would be 1/10Gbps iSCSI

#### **NS-960 SP Configurations (up to 4-Blade & 480 Drives):**

(2) FC SLICs, slots 0 & 1

(2) FC SLICs, slots 0 & 1, and (4) iSCSI SLICs, slots 2, 3, 4, 5

#### **NS-960 SP Configurations (>4-Blades or >than 480 Drives):**

(4) FC SLICs, slots 0, 1, 2, 3

(4) FC SLICs, slots 0, 1, 2, 3, and (1) iSCSI SLIC, slot 4

(5) FC SLICs, slots 0, 1, 2, 3, 4

(5) FC SLICs, slots 0, 1, 2, 3, 4, and (1) iSCSI SLIC, slot 5

**Note:** Customers select AUX or AUX8, then add additional slices per the above rules, to create the desired configuration.

→Fibre Channel Tomahawk SLIC is 4Gbps 4-port card

→Fibre Channel Glacier SLIC is 8Gbps 4-port card

→iSCSI Harpoon SLIC is 1Gbps 2-port card

→iSCSI Poseidon SLIC is 10Gbps 2-port card (Requires Taurus Flare 29)

#### **NS-480 Rules: NS480-AUX; NS480-AUX8**

→Base AUX array will have (2) 4-port 4Gbps Tomahawk FC IO Module (no 8Gbps Glacier)

→Mfg or Upgrade can add (1) 4Gbps FC IO Module to the array

→Mfg or Upgrade can add (1) 8Gbps FC IO Module to the array, in lowest available slot position before any 4Gbps FC cards

→Mfg or Upgrade can add (1) or (2) 10Gbps iSCSI Poseidon IO Module, in lowest available slot position before 1Gbps iSCSI cards

→Mfg or Upgrade can add (1), (2), or (3) 1Gbps iSCSI IO Modules

SP slot 0 & 1 can only be 4Gbps FC (AUX), and are always populated

SP slot 2 is either empty or populated with 1 or 10Gbps iSCSI

SP slot 3 can be empty, or be 4/8Gbps FC, or be 1/10Gbps iSCSI

SP slot 4 can be empty, or be 1/10Gbps iSCSI

If FC or MPFS over FC selected, slot\_2 is empty, and slot\_3 is 4 or 8Gbps FC

If FC or MPFS over FC selected, & Mirrorview over iSCSI is selected, slot\_2 will have an iSCSI slice, & slot\_3 is a 4 or 8Gbps FC slice

If MPFS over iSCSI selected, slots 2, 3, & 4 are iSCSI

#### **NS-480 SP Configurations:**

(2) FC SLICs, slots 0 & 1

(2) FC SLICs, slots 0 & 1, and (3) iSCSI, slots 2, 3, 4

(3) FC SLICs, slots 0, 1, 3

(3) FC SLICs, slots 0, 1, 3, and (1) iSCSI, slot 2

#### **NS-120 Rules: NS120-AUX, NS120-AUXF; NS120-AUXF8**

→Base AUX array will have (1) 4-port 4Gbps Tomahawk FC IO Module (no 8Gbps Glacier)

→Mfg or Upgrade can add (1) 4Gbps FC IO Module to the array

→Mfg or Upgrade can add (1) 8Gbps FC IO Module to the array, in lowest available slot position before any 4Gbps FC cards

→Mfg or Upgrade can add (1) 10Gbps iSCSI Poseidon IO Module, in lowest available slot position before any 1Gbps iSCSI cards

→Mfg or Upgrade can add (1) or (2) 1Gbps iSCSI IO Module

→NS-120 is handled differently & will use AUX, AUXF8, or AUXF in the order entry system, dropping the AUXI designation

SP slot 0 will have 4Gbps FC (AUX), and is always populated

SP slot 1 can be empty, or be 1/10Gbps iSCSI

SP slot 2 can be empty, or be 4/8Gbps FC, or be 1/10Gbps iSCSI

SP slots 3 & 4 are not used

#### **NS-120 SP Configurations:**

(1) FC SLIC, slot 0 (Port 1 empty, not used)

(2) FC SLICs, slot 0 & 2

(1) FC SLIC, slot 0, (2) iSCSI SLICs, slots 1 & 2 for MPFS over iSCSI

(2) FC SLICs, slots 0 & 2, (1) iSCSI SLIC, slot 1 for Mview over iSCSI

#### **GLACIER 8Gbps FC IO MODULE:**

→Libra Flare 28.7 introduces support for 8Gbps FC IO module

→Support for new 4-port Glacier 8Gbps FC SLIC on SPs for Gateways and the NS-120, NS-480 & NS-960 platforms (Aug 2009)

**Note:** NS-120 & NS-480 arrays supporting FC will add only a single Glacier card per SP, which will be used to support 4 or 8Gbps Host access. All FC IO cards on the NS-960 arrays, however, will support the Glacier IO module.

→Arrays using Glacier card will use a new CLARiiON model name:

#### **CX4-120C8, CX4-240C8, CX4-480C8, CX4-960C8**

→Glacier is a 8Gbps replacement to the 4Gbps Tomahawk FC IO Module, and is RoHS 6 compliant

→For new installs and array upgrades—can only add 8Gbps FC modules to existing array, not replace 4Gbps with 8 Gbps

→Currently, neither Blades nor SPs themselves support 8Gbps internally on the array—would only be used by external Hosts

→5.6.45 qualifies support for the 8Gbps FC IO Cards (Glacier) in the NS-960/NS-G8 Blades, though they will not begin shipping until November 2009

→8Gbps FC support will be added for NS-960/G8 Blade IO Modules in the future

**Note:** NS-120 & NS-480 Blades are not capable of running 8Gbps

→8Gbps FC uses optical SFP+ transceivers and LC connectors, with OM2 or OM3 cables (10/150 meter distance, respectively)

#### **MOJITO 10GbE/1GbE ETHERNET MODULE:**

→Support for the Mojito FrontEnd Ethernet IO card for NX4, NS-120, NS-480, & NS-G2 platforms, Q1 2010

**Note:** Mojito consists of 2-ports 10GbE with optical SFPs (will not support 1GbE speeds) and 2-ports 1GbE copper with RJ45 connectors. The 10GbE interface will support FSN failover from 10G-to-1G; 10G-to-10G; & 10G-to-LACP 1G ports

→Blade part number 100-520-066

#### **# server\_sysconfig server\_2 -pci |tail**

Broadcom 10 Gigabit Ethernet Controller

0: fxg0 IRQ: 24

txflowctl=disable rxflowctl=disable

0: fxg1 IRQ: 25

txflowctl=disable rxflowctl=disable

#### **# cat data\_mover\_resume.server\_2.xml**

<?xml version="1.0"?>

<DataMoverResume Host="server\_2">

  <RESUME\_INFORMATION\_BLADE

**EMC\_PART\_NUMBER="100-520-066"**

    EMC\_ARTWORK\_REVISION=" "

    EMC\_ASSEMBLY\_REVISION="A03"

    EMC\_SERIAL\_NUMBER="CF21B091200008 "

    VENDOR\_NAME="CELESTICA "

    LOCATION\_OF\_MANUFACTURE="THAILAND "

    YEAR\_OF\_MANUFACTURE="2009"

    MONTH\_OF\_MANUFACTURE="4 "

    DAY\_OF\_MONTH\_OF\_MANUFACTURE="10"

    ASSEMBLY\_NAME="**Mojito NS3-40 Dual 10G NAS** "

#### **Troubleshooting Mojito 10GbE:**

**\$ .server\_config server\_x -v "xena fxg0 stat"** [Viewing port statistics on the 10GbE interface]

**\$ .server\_config server\_x -v "xena fxg0 macstats"**

#### **DUAL CONTROL STATION DLm 2.0:**

→Dual Control Station for DLm product (includes NS-120 & NS-960)

→Achieved through use of existing DLm external switch for connecting the 2<sup>nd</sup> Control Station

→Control Station used as management station for the DLm, and is used to pass SNMP traps, deliver ConnectEMC callhomes, and to provide for ESRS connectivity to DLm

#### **MAC OS 10.5 SUPPORT:**

→Support for MAC 10.5 as NFSv3 client

→Support for SMB1 CIFS protocol

#### **PROVISIONING WIZARD for CELERRA MANAGER (aka, UMPW--Unified Manager Provisioning Wizard):**

→Celerra Manager Provisioning Wizard, new naming convention

→This wizard allows users to provision storage for Integrated and FC enabled systems from Celerra Manager, which is a big change from past versions, where only the pure Integrated system was allowed to use Celerra Manager

**Note:** With the advent of this capability, the Celerras>Storage>Systems>array\_name>Configure option is no longer presented

→For NX4, NS20, NS40, NS-120, NS-480, & NS-960 systems

→Celerra Manager “Provision Storage” wizard is available by rightclicking array (Celerras>Storage>Systems>array) or from Wizards>Provisioning>Provision Storage

**Note:** Provisioning capabilities presented are the same as found in the CPW utility. Only a single RAID type can be specified with the provisioning tool at a time, meaning that setup\_clariion may configure a shelf of disks differently because it uses multiple RAID type templates.

→For De-provisioning, Celerra Manager allows the deletion of Hot Spares, and unused Disk Groups (Raid Groups), which is similar to the CLI nas\_disk –delete –perm –unbind, which removes from the Celerra DB and from the array backend (cannot remove individual dvolumes from Celerra Manager as can be done with nas\_disk –delete)

**nas\_storage -delete id=x -spare <spindle\_ID> | -group <diskgroup\_ID>** →Equivalent functionality as with Celerra Mgr

**Note:** As a feature, Deprovisioning was actually included with 5.6.42, which also allowed delete of dvolumes (since removed)

#### **CLARiiON VIRTUAL PROVISIONING SUPPORT:**

→Mira FLARE (.504) introduced Virtual Provisioning Phase I, and Libra FLARE (.704) expands with Phase II

→CLARiiON Virtual Provisioning is a concept whereby more storage can be presented to an application than is physically available. The idea is to be able to increase capacity utilization on the array, simplify storage management, and reduce application downtime.

→Virtual Provisioning groups a collection of physical disks into “Thin Pools”, which are nested under the “Storage Pools” icon in Navisphere, along with traditional LUN RAID Groups. “Thin LUNs” are created within the “Thin Pools” and then added to a Storage Group.

→Main difference between Thin and Thick LUNs is that the former uses less actual physical space—space is only allocated from the Thin Pool when an application uses it

→Celerra supports this feature on Gateway and FC Enabled systems only (since Navisphere must be used and licensed)

→Celerra supports the use of CLARiiON Virtual Provisioning in the creation of file systems or iSCSI LUNs—Thin LUNs are added to the Celerra Storage Group, and diskmarked for use on the Celerra

→Celerra creates a “Storage Pool” for AVM based on the name of the Thin Pool, as created by Navisphere, and creates a separate “Storage Pool Profile” for each corresponding “Thin Pool” where “Thin Luns” are added to the Celerra Storage Group, as shown below:

#### **# nas\_pool -list**

```
id    inuse  acl   name
40    n      0     Thin Pool 0_FNM00083800203
41    y      0     Thin Pool 1_FNM00083800203
```

**Note:** Each time new Thin Luns are added to the Celerra that originate from different CLARiiON Thin Pools, a new Storage Profile Name is assigned and added to the ‘pools’ file

#### **# cat /nas/volume/pools|tail**

```
40:Thin Pool 0_FNM00083800203:0:Thin Pool Thin Pool 0 on FNM00083800203:1:1:y:n:n:80:y:y:
41:Thin Pool 1_FNM00083800203:0:Thin Pool Thin Pool 1 on FNM00083800203:1:1:y:n:n:82:y:y:
```

#### **# nas\_disk -info d9**

```
id      = 9
name    = d9
acl     = 0
in_use  = True
pool    = Thin Pool 0_FNM00083800203
size (MB) = 51199
type    = CLSTD
protection= RAID5(2+1)
stor_id  = FNM00083800203
stor_dev = 0006
volume_name = d9
storage_profiles = Thin Pool 0_FNM00083800203
virtually_provisioned = True
```

→Celerra does not rename Thin LUNs, but does adjust the HLU value to a legal Celerra data lun value

#### **USING NAVICLI:**

**# /nas/sbin/navicli -h 10.241.168.150 storagepool -list**

**# /nas/sbin/navicli -h 10.241.168.150 thinlun -list**

**# /nas/sbin/naviseccli -h 10.241.168.179 -user nasadmin -password nasadmin -scope 0 thinlun -create -type Thin -capacity 5 -sq gb -poolId 2 -sp a**

#### **CLARiiON VIRTUAL PROVISIONING RULES:**

→Not supported with Vault drives or Celerra Control LUNs

- Not supported with CPW or UMPW
  - Not supported with MirrorView or SAN Copy
  - Not supported with Reserved LUN pools, clone Private LUNs, or as a component of a metaLUN
  - Cannot use Celerra VP with CLARiiON VP for the same LUNs
  - Best Practice to not mix virtual and dense LUNs, mix Disk Technology, or Disk sizes, within a pool
  - Best not to use Thin LUNs for high performance applications
  - Thin LUN size limited from 1GB to 14TB
  - For VMWare configurations, setup VMWare virtual disks as “zeroedthick”—storage is reserved in the Datastore, but the VMWare kernel does not initialize all the blocks until the guest O/S begins writing to uninitialized blocks
  - Do not create sparse iSCSI luns from file systems that are built on CLARiiON Thin Luns
  - Feature requires installation of the VirtualProvisioningEnabler on the array (CLARiiON Support Site)
- Note:** Download from CLARiiON Support Site: [VirtualProvisioningEnabler-01.01.5.001-fleet64\\_free.ena](#)
- Feature requires configuration of Thin Pools and Thin LUNs via CLI, or Navisphere, either via the Storage Provisioning Wizard, by Thin Pools icon under Storage Pools>Create Thin Pool, or by rightclicking Array\_name or Storage Pool icon > Storage Pool Operations >Create Storage Pool>Storage Pool Type:
    - o RAID Group    o Thin Pool
      - Thin Pools can be expanded via Navisphere/CLI by adding disk drives to the pool
      - When Thin LUNs are deleted, space is reclaimed in the Thin Pool by ‘Consumed Capacity’
      - CLARiiON Thin Pools are supported with FC, SATA, or EFD drives in R5 & R6 configurations

**Note:** RAID 5 is the default selection when creating Thin Pools and is recommended--when expanding a R5 Thin Pool, add disks in multiples of 5, minimum of 3 disks. RAID 6 can be used, and is recommended for SATA drives--add disks to R6 Thin Pools in multiples of 8 drives, minimum of 4 disks.

- Supported with IP Replication (configure virtual provisioning on Source only, Destination will synchronize automatically)
- Not supported on NX4 since it's based on Flare 23

### **THIN POOL TERMINOLOGY:**

- Thin Pool ‘User Capacity’ is the total raw capacity of the drives, less the overhead associated with the RAID protection, that can be used by all Thin LUNs in the pool
- ‘Consumed Capacity’ is what is actually allocated from the total ‘User Capacity’, and includes Array overhead in building a Thin LUN (requires minimum of 2GB per Thin Lun), and Host file systems created
- ‘Total Subscribed Capacity’ is the total capacity used by the Host(s) (file systems, iSCSI LUNs, etc.)
- ‘Percent Subscribed’ is derived from Total Subscribed Capacity divided by User Capacity \* 100
- % Full Threshold (Consumed Capacity / User Capacity) is an alert monitoring mechanism on the array that issues a Warning when the threshold is crossed, and a Critical alert when pool usage crosses 85%. By default, Thin Pools are created with a 70% threshold value.
- ‘Oversubscribed by’ is when ‘Total Subscribed Capacity’ exceeds the ‘User Capacity’ of the Thin Pool

### **NS-G2 REBRANDING of NS40G:**

- Rebranding the NS40G as the NS-G2 product. /nas/sbin/model will show the system as an NS-G2 based on new TLA part number from manufacturing. Upgraded systems will not change and will remain NS-40G.
- Differences between NS-G2 and NS40G is that NS-G2 will support the Mojito Ethernet card, Maynard CS only, provides 3-year warranty, offers Enhanced support, and recommends additional 1U space by Control Station for access

### **/nas/sbin/t2vpd**

**Note:** Parses for known NS-G2 TLA numbers from resume, & use nas\_xml command to set model to NS-G2 vs. NS40G

### **NS-G2 Enclosure TLA's/Models (Single & Dual-blade systems):**

Factory model NSG2-1A/2A

100-520-679, 100-520-680 for Single & Dual Blade enclosure, 4-port Cu 1GbE Ethernet card

Factory model NSG2-1B/2B

100-520-681, 100-520-685 for Single & Dual Blade enclosure, 2-port Optical 1GbE, 2-port Cu 1GbE Ethernet card

Factory model NSG2-1C/2C

100-520-686, 100-520-687 for Single & Dual Blade enclosure, 2-port 10GbE Optical, 2-port 1GbE Cu Mojito Ethernet card

**Note:** FRUs & CRUs will be the same as presented for front-end on NS-120/NS-480 Celerra

### **Hardware Commands:**

# **/nas/sbin/t2vpd -s 2** →Blade & SFP info from RESUME PROM

# **/nas/sbin/t2vpd -e 0** →Enclosure TLA information

### **64-BIT CEE SUPPORT:** CEE Version 4.5.0.4

→64-bit Windows application for 64-bit Operating Systems

→64-bit CEE offers IPv6 runtime support and replaces ONC/XDR [ONC/RPC] protocol communications with MS-RPC using XML payloads (eliminates need to continue licensing the 32-bit ONC product from Distinct)

→Both 32-bit and 64-bit CEE kits will be maintained

EMC\_CEE\_Pack\_Win32\_4.5.0.4 (32-bit version)

EMC\_CEE\_Pack\_x64\_4.5.0.4 (64-bit version)

→New CEE version backwards compatible with ONC protocol if NAS is down-rev

→Adds new fault tolerance feature

→Adds Auditing class support based on DatAdvantage from Varonis

#### **Upgrading from 32-bit to 64-bit CEE:**

1) Would require uninstall of existing CEE or CAVA package, followed by installation of the 64-bit version

2) viruschecker.conf & cepp.conf files would require editing to add IPv6 addresses

addr=192.16.20.15:192.16.20.16:[001:0db8:85a3:0000:0000:8a2e:0370:7334]

3) Edit cepp.conf file to add new fault tolerance level line & msrpcuser=<cee\_username> line

#### **CEPA AUDITING CLASS SUPPORT:**

→CEPA Phase 2 Auditing class application support [DatAdvantage—from Varonis]

→Adds two new events for auditing: OpenDir and CloseDir

**Note:** Auditing clients will register with CEPA API and use “post-event” reporting from Celerra to gather auditing events without blocking CIFS calls. CEPA API uses an IDL & XML DTD file to interact with the application host (XML over RPC). Application could be co-resident within the CEE framework using local RPC calls, or remote using MS-RPC calls. Windows clients that perform file or directory create, open, close, close after modification, rename, delete, set acl functions can be audited. CEPA will gather events and context to create the tuple that is delivered with the CEPA API to registered applications. “Context” consists of event type, UNC, NetBIOS name of server, File owner SID, File user SID, file size, dwDesiredAccess, dsCreationDisposition.

#### **New Lines for Cepp.conf file:**

**ft level=0** →Fault Tolerance feature

**msrpcuser=domain.ceeuser** (domain name used to run CAVA service on CEE system)

#### **New Fault Tolerance Capability:**

ft level=0 location=/fs1 size=5 →with loss of CEPA heartbeat, continue and tolerate lost events (default setting)

ft level=1 →with loss of CEPA heartbeat, continue and use a persistent buffer file as a circular event buffer for lost events (once buffer is full, events will be overwritten)

ft level=2 →with loss of CEPA heartbeat, continue and use a circular event buffer file until buffer is full, then stop CIFS

ft level=3 →with loss of CEPA heartbeat, stop CIFS

**Note:** Without location specified, circular event buffer file is written to root; without size specified, default is 1MB for persistent buffer file.

**# server\_cepp server\_x –pool –stats –all** (to observe auditing statistics)

#### **MSRPC VC/CEPP ISSUE:**

→If CEPP cannot perform a reverse DNS lookup of the IP Address to a Name, the Server Logs will become flooded with the following entries, because it tries to use MSRPC, which will fail: See AR149912.

2009-07-21 09:16:54: 26041909248: SMB: 6: netbios session request failed return type 83 error 82 (see rfc 1002) on server 198 with client FILESRV01AVBK with port 139

#### **Workaround:**

→Set “param cepp RPCtype=1” to disable MSRPC protocol for VirusChecker and CEPA and to use ONC-RPC

1= ONCRPC

2= MSRPC

3= both (default)

#### **CST 1.1 TOOLKIT:**

→Replacing Control Station CAMS Common Sign-On (CSO) module with the RSA Common Security Toolkit (CST) module—this is used for communications & authentication with external directory servers

→Allows sharing of encryption keys, such as for dual CS environments

#### **SYMMETRIX V-MAX TOLERANCE:**

→Support for V-Max backends, using SymAPI version V7.0.0.0-914 or higher

→Minimum microcode level that supports Solutions Enabler 7.0 (i.e., SymAPI 7.0) is 5671

→NAS upgrades to 5.6.45 or higher will be blocked if any backends are detected as Symm 5 or older, or if the Symm Microcode level is below 5671

#### **SYMMETRIX SUPPORT:**

→Symm 5 & 5.5 are not be supported with 5.6.45 because the Solutions Enabler SymAPI version is V7.0

Symm Family: 5.0 & 5.5; Model 81xx, 84xx, 87xx (Microcode 5566, 5567, 5568), & 82xx, 85xx, 88xx (5567, 5568)

→Also, DMX & DMX2 are not supported unless running microcode 5671 or higher

#### **8K TREE QUOTAS:**

→Support for 8K tree quotas per file system (from 2048 to 8096 tree quotas per file system)

### **COGNAC 5.6 MAINTENANCE RELEASE 9 (CMR9 5.6.46)—RFE Aug 2009**

→CPW integration into DART release cycle was scheduled, but is now postponed. Instead, a separate CPW version will be posted on Powerlink.

### **COGNAC 5.6 MAINTENANCE RELEASE 10 (CMR10 5.6.47) GA Dec 2009:**

**CELERRA PANIC CALLHOME/LOGGING ENHANCEMENTS—See emc227235:**

- Serviceability feature for providing system ‘snapshot’ & panic header info in the CDATA section of the CallHome payload for all Blade Panics (<than 100kb)
- When triggered, the /nas/tools/.status\_check script runs a series of commands to gather a Snapshot of the system state, along with the panic header file, and inserts the information into the CDATA CalHome XML file as plain text. The contents of another new feature, the /nas/log/inventory\_status file are also delivered, but for all CallHomes (and not just panics)
- Useful for initial triage as it does not require Support to dial back into the system to determine panic header and whether system failed over properly or not
- Script basically collects panic header, failover status, NAS information, Array information, and snippets from sys\_log, server\_log, SP logs, etc.

**New BoxMonitor Panic Events:****BoxMonitor:CRITICAL:802 to 815** (slots 2-14)

- Original BoxMonitor Panic Event IDs (BoxMonitor:CRITICAL:302 to 315) are kept to record actual panic time, but by themselves, no longer call home

**How does the new feature work?**

- A Panic triggers Event 302-315, logs to sys\_log, calls recover\_slot, which calls .status\_check, then Event 802-815 fires & sends XML CallHome with payload
- Goal is to run .status\_check within 30 seconds, with a hard cutoff at 2 minutes, at which point the script will send whatever it has collected, in the CallHome
- Panic CallHomes are limited by System to once every 24 hours per Blade

**/nas/sys/nas\_eventlog.cfg file**

```
disposition range=802-802 threshold=1 rearm=1000 resetafter=86400, callhome plaintext
```

**BoxMonitor sys log events for slot\_3 panic with old & new event ID:**

- ```
Nov 5 12:25:18 2009:CS_PLATFORM:BoxMonitor:CRITICAL:303:::1257441918:Slot 3 has panicked.
Nov 5 12:25:25 2009:CS_PLATFORM:BoxMonitor:DEBUG:905:::1257441925:Slot 3 recover_slot /nas 3 has started.
Nov 5 12:29:20 2009:CS_PLATFORM:BoxMonitor:CRITICAL:803:::1257442160:Slot 3 has panicked - post panic logging complete: >>PANIC in file: ../../dskdump.cxx at line: 1836 : for new panic event test. Dump information can be found in /nas/log/post_panic_info_2009_11_05_12:29:03_EST_slot_3.log.
```

**Script & Input file:****/nas/tools/.status\_check** → Main script for this new feature

- Gathers panic header; failover recovery status; NAS version; model info; log collection status; getreason codes; sys\_log, cmd\_log, cmd\_log.err, /var/log/messages info; FCP info; server version and info; NAS Cel list; exports; NFS & CIFS Stats; storage list; Navicli getagent, getcrus, getcache, getdisk, getlun –trespass, getlog -50

**/nas/tools/.status\_checks.csv** → Editable input file to .status\_check for commands to run with .status\_check**Log Files:****Post Panic Info Log:****/nas/log/post\_panic\_info\_<date>\_slot\_x.log**

```
-rw-r--r-- 1 root root 70515 Nov 10 11:27 post_panic_info_2009_11_10_11:26:59_EST_slot_2.log
-rw-r--r-- 1 root root 2454 Nov 10 11:27 panic_headers.log
```

- The post\_panic\_info\_<timestamp>\_<slot\_x>.log holds system summary info from .status\_check, with up to (14) unique post\_panic\_info log files kept

**Panic Headers Log:****/nas/log/panic\_headers.log**

- This file holds up to 2000 lines of panic header info, trimmed back to 2000 lines daily, as needed, via log trimmer in /nas/sys/nas\_mcd.cfg

**Log Trimmer:****/nas/sys/nas\_mcd.cfg**

- This is where the Panic Headers log Trimmer lives

```
daemon "Paniclog Trimmer"
  executable  "/nas/sbin/log_trimmer"
  optional    no
  autorestart yes
  ioaccess    no
  cmdline     "-n /nas/log/panic_headers.log 2000 1 24 h t 2 n"
```

**Can run the .status check script manually:****# /nas/tools/.status\_check -showAll** (hidden switch)

```
.status_check: -slot <slot_num> [ -f <csv_input_file> ] [ -quiet ] [ -verbose ] [ -timeout ]
               | -csv_op [ check | clean ] [ -f <csv_input_file> ]
```

-slot: Mandatory argument. Information will be collected with regards to the server in this slot.

-quiet: Operation will display no output. | -verbose: Operation will display detailed process information.  
-f: Absolute path to custom CSV file  
-timeout: Overwrite the default timeout given for this operation to complete (default 120 secs)  
-csv\_op: Mandatory argument. Followed up by either check or clean  
check: Check if the file is sane and not malformed.  
clean: Make CSV file user readable and properly justified

## # /nas/tools/.status\_check -s 2

Spawning worker threads to collect information...done  
Waiting for jobs to finish.....done  
Collection log saved to: /tmp/.system\_status\_slot\_2.txt → Information saved here when .status\_check run manually

### NAS Event Information:

#### # nas\_event -list -action callhome|grep -i 802

802 CRITICAL(2) Slot \${SLOT},24,%u} has panicked - post panic logging complete: \${PANIC\_INFO},8,%s}. Dump information can be found in \${src\_file\_path},8,%s}.

#### /nas/sys/nas\_eventlog.cfg file contents

```
disposition range=802-802 threshold=1 rearm=1000 resetafter=86400, mail user
    disposition range=803-803 threshold=1 rearm=1000 resetafter=86400, mail user
    disposition range=804-804 threshold=1 rearm=1000 resetafter=86400, mail user
disposition range=802-802 threshold=1 rearm=1000 resetafter=86400, callhome plaintext
    disposition range=803-803 threshold=1 rearm=1000 resetafter=86400, callhome plaintext
    disposition range=804-804 threshold=1 rearm=1000 resetafter=86400, callhome plaintext
```

**Note:** The 86400 = seconds, which translates into once every 24 hours

### Contents of Post Panic Info File:

```
# cat post_panic_info_2009_11_10_11\:\:26\EST_slot_2.log|grep "##"
##--Panic header--##
##--Failover / recovery status--##
##--NAS Version--##
##--Model--##
##--Automatic Collection Status--##
##--Reason Codes--##
##==sys_log Info==##
##==cmd_log Info==##
##==cmd_log,err Info==##
##==System Messages==##
##--FCP Topology--##
##--Server Version--##
##--Server Information--##
##--NAS Cel--##
##--Exports--##
##--server_log Info--##
##--NFS Stats--##
##--CIFS Stats--##
##--Storage List--##
##--Navicli getagent--##
##--Navicli getcrus--##
##--Navicli getcache--##
##--Navicli getdisk -state--##
##--Navicli getlun -trespass--##
##--Navicli SPA getlog -50--##
##--Navicli SPB getlog -50--##
```

### Abridged Example of CallHome XML File Info:

```
<Event>
    <SymptomCode>802</SymptomCode>
    <Category>Status</Category>
    <Severity>Critical</Severity>
    <Status>Failed</Status>
    <Component></Component>
    <ComponentID></ComponentID>
    <SubComponent></SubComponent>
    <SubComponentID></SubComponentID>
```

```

<CallHome>Yes</CallHome>
<FirstTime>2009-11-10T11:27:16</FirstTime>
<LastTime>2009-11-10T11:27:16</LastTime>
<Count>1</Count>
<EventData>
```

<![CDATA[CCMD ID: 78928610082

Brief Description: Slot 2 has panicked - post panic logging complete: >>PANIC in file: ./dskdump.cxx at line: 1836 : testing new panic handling feature. Dump information can be found in /nas/log/post\_panic\_info\_2009\_11\_10\_11:26:59\_EST\_slot\_2.log.

Full Description: The DART OS running on the Data Mover slot has panicked. The specified dump information that describes this error can be found in the dump log.

Recommended Actions: To resolve this problem, perform the following: To resolve this problem, perform the following:

1. Using the CLI, log in as nasadmin (or as a user with NAS administrative privileges) and "su" to root (requires the root password).
2. To check for the existence of the logs, run the "cd /nas/var/emcsupport" command and then the "ls" command to see if the log collection file exists.
  - \* An example file is support\_materials\_APM00073701085.080206\_1401.zip.
3. To check for the existence of the dump and header files, run the "cd /nas/var/dump" command and then the "ls" command to see if the dump files exist.
  - \* Example dump files are dump\_APM00073701085.080206\_1411.dump.gz and header\_APM00073701085.080206\_1411.txt.
4. If the Data Mover continues to be unresponsive or for assistance, search the Knowledgebase on Powerlink as follows:
  - a. Log in to <http://powerlink.emc.com> and go to Support > Knowledgebase Search > Support Solutions Search.
  - b. Use the message ID or text from the error message's brief description to search.

A list of all dump headers can be found by performing the following actions:

1. To view the post\_panic\_info file, run the "cd /nas/log/" command and then run "cat post\_panic\_info | less" and look at the specified dump information.

System Configuration Snapshot: Tue Nov 10 11:15:02 EST 2009

Control Station Slot No: 0

/nas/sbin/model -option -list:

OPTION ENABLED: NOT AVAILABLE

/nas/bin/nas\_inventory -list:

| Component                         | Type            | Status | System ID                         |
|-----------------------------------|-----------------|--------|-----------------------------------|
| Battery A                         | Battery         | OK     | CLARiiON CX4-960 FNM00083800203   |
| Battery B                         | Battery         | OK     | CLARiiON CX4-960 FNM00083800203   |
| Celerra NS-960 FNM000838002030000 | Celerra         | OK     | Celerra NS-960 FNM000838002030000 |
| CLARiiON CX4-960 FNM00083800203   | CLARiiON        | OK     | CLARiiON CX4-960 FNM00083800203   |
| Control Station 0                 | Control Station | OK     | Celerra NS-960 FNM000838002030000 |
| DME 0 Blower A                    | Blower          | OK     | Celerra NS-960 FNM0008380020300   |

-----output abridged-----

#--Panic header--##

DART time of dump: Tue Nov 10 08:13:48 2009 UTC

Product: EMC Celerra File Server

Uptime: 010 days, 22:24:16

Version: T5.6.47.6

IP Address: 128.221.252.2

Host Name: server\_2

COMPRESSED = YES

DART panic/fault message:

>>PANIC in file: ./dskdump.cxx at line: 1836 :

testing new panic handling feature

-----output abridged-----

#--Failover / recovery status--##

Status\_check collection script version: 1.0

Time on Control Station: Tue Nov 10 11:27:00 EST 2009

Information being collected for slot: 2

Standby slot for this slot is: 3

Failover status: SUCCESS

Recovery status: SUCCESS

#--NAS Version--##

5.6.47-6

##--Model--##

NS-960

##--Automatic Collection Status--##

Current configuration:

The automatic collection feature is enabled and the automatic transfer feature is disabled

Maximum number of dumps allowed in /nas/var/dump: 2

##--Reason Codes--##

10 - slot\_0 primary control station

4 - slot\_2 configured(Post Code = 0x61 ; Mid Code = 0x80 ; Blade Code = 0x2c ; Slic Code = 33686017)

5 - slot\_3 contacted(Post Code = 0x61 ; Mid Code = 0x80 ; Blade Code = 0x2c ; Slic Code = 33686017)

##==sys\_log Info==##

Nov 9 10:01:13 2009:CS\_PLATFORM:NASDB:INFO:306:::1257778873:nasdb\_backup: Celerra database backup done.

Nov 9 11:01:08 2009:CS\_PLATFORM:NASDB:INFO:300:::1257782468:nasdb\_backup: NAS\_DB checkpoint is in progress.

-----output abridged-----

##==Status Report==#

Collection status: SUCCESS

The overall execution time for this collection was: 16 seconds

**Panic still triggers Auto Log Collect Script Feature, which is independent of the Panic Enhancement feature:**

→Auto Log Collection & Dump collection still takes place independently

**/nas/var/emcsupport**

-rw-rw-r-- 1 nasadmin nasadmin 2527422 Nov 5 12:32 support\_materials\_FNM00083800203.091105\_1229.zip

**/nas/var/dump**

-rw-rw-r-- 1 nasadmin nasadmin 28059525 Nov 5 12:29 dump\_FNM00083800203.091105\_1229.dump.gz

-rw-rw-r-- 1 nasadmin nasadmin 1181 Nov 5 12:29 header\_FNM00083800203.091105\_1229.txt

**UNIFIED BLOCK STORAGE EXPERIENCE (UBSE):**

→New Marketing characterization for Celerra Integrated offerings

→Unified Block Storage the new Integrated terminology

→A major goal of this initiative is to support all CLARiiON CX4 SP I/O Modules configurations on Celerra Unified models (e.g., 8Gbps FC; 10Gbps iSCSI)

→Allows for Native Block CLARiiON FC & iSCSI support--LUN can be presented via FC or iSCSI to Hosts, and can support all the CLARiiON Layered Applications

→A supporting goal is to shift to “Base Models” in the Order Entry process—customers order a ‘Base NS-480’, ‘Base NS-960’, etc., then add whatever legal combinations of SLICs are allowed for the SPs, along with a choice of MPFS over iSCSI/FC, or native CLARiiON iSCSI/FC licenses

→Marks the transition from Celerra iSCSI as a product, to CLARiiON iSCSI, phasing out the former

→However, for the time-being, the /nas/sbin/model command will still display the three variants that we know today (e.g., NS-120, NS-120iS, NS-120FC), but will transition next year to only a single /nas/sbin/model name per Celerra platform (e.g., NS-120 only)

→Be advised that the Install Guides, and other formal documentation, have already removed the use of the “FC” and “iS” terms

→A related nas\_inventory feature was introduced to assist providers in Remote System identification (nas\_inventory will be discussed in detail in next Topic)

→/nas/sbin/model can no longer be used alone to define the actual system ‘type’

→/nas/sbin/model expanded with “–option –list” to show license option for a particular system (read from /etc/be\_sg\_info file)

→/nas/sbin/nas\_hw\_upgrade expanded to allow Mfg or Field to enable or disable license flag options depending on Customer Sales Order for the Unified models, used during Installs and HW Upgrades

→During factory installation, either the “FC Enabled” or “Integrated” option is selected, then after installation, the has\_hw\_upgrade script is run, with appropriate switches selected based on Licenses and Physical SLICs ordered

→/nas/log/nas\_xml.xml contains model name used for display purposes

**Protocol Licenses available when ordering:**

NS120/480/960-NBOPTUL MPFS over FC →MPFS FC (mpfsfc flag)

NS120/480/960-NBOPTUL MPFS over iSCSI →MPFS iSCSI (mpfssi flag)

NS120/480/960-NBOPTUL CLARiiON FC →CLARiiON FC (clarionfc flag)

NS120/480/960-NBOPTUL CLARiiON iSCSI →CLARiiON iSCSI (clarionis flag)

**Note:** CLARiiON FC or ISCSI is referred to as Native Block storage (NB)

**Protocol License Flags/Options:**

→License “flags” are added or deleted by using nas\_hw\_upgrade script

→License “flags” are written to /etc/be\_sg\_info file

→License “flags” are displayed via /nas/sbin/model –option –list

**License Flag Translation Table:**

**/nas/sbin/model –option –list**

**tail /etc/be\_sg\_info**

**/nas/sbin/nas\_hw\_upgrade**

|                                |                       |                             |
|--------------------------------|-----------------------|-----------------------------|
| OPTION ENABLED: MPFS FC        | MPFSF_ENABLED=YES     | -option -enable -mpfsfc     |
| OPTION ENABLED: MPFS iSCSI     | MPFSI_ENABLED=YES     | -option -enable -mpfsis     |
| OPTION ENABLED: CLARIION FC    | CLARIIONF_ENABLED=YES | -option -enable -clariionfc |
| OPTION ENABLED: CLARIION iSCSI | CLARIIONI_ENABLED=YES | -option -enable -clariionis |
| OPTION ENABLED: SAN            | FC_ENABLED=YES        | -fc_option -enable          |

**Note:** Each line represents a protocol license option and the different way it's displayed, depending on what's being looked at

**Manufacturing Rules:**

→If MPFS over FC and/or MPFS iSCSI, system installed as an “Integrated”

→If CLARiiON FC and/or CLARiiON iSCSI, system installed as an “FC Enabled”

→Specific license “flags” are written to /etc/be\_sg\_info file during installation when /nas/sbin/nas\_hw\_upgrade script is run

**Upgrade/Reinstall Rules:**

→Field will be expected to update license flags per the upgrade procedures or during system reinstalls—for Upgrades, the Field may need to refer to the Sales Order for licensing & SLIC HW info

**Special Installation Note:**

→During fresh NAS installs, the system gives an option for “FC Enabled” or “Integrated”. You would select “FC Enabled” if the “CLARiiON FC” or “CLARiiON iSCSI” options were ordered, otherwise “Integrated”.

→Display model name is a function of Install choice, the nas\_hw\_upgrade flags enabled, and actual HW I/O modules used

**NAS HW UPGRADE FLAG OPTIONS:**

# **/nas/sbin/nas\_hw\_upgrade** →Used to add, remove, or list protocol Flag options, writes to be\_sg\_info file

```
nas_hw_upgrade -fc_option {-enable|-disable}  
nas_hw_upgrade -option {-enable|-disable} {-mpfsfc -mpfsis -clariionfc -clariionis}  
nas_hw_upgrade -option -list {-mpfsfc -mpfsis -mpfs -clariionfc -clariionis -clariion -san}
```

**/nas/log/nas\_hw\_upgrade.log**

→Logs upgrade and license flag option activities

# **tail /etc/be\_sg\_info**

# **/nas/sbin/model -option -list** →Output of this command reads from /etc/be\_sg\_info file

```
OPTION ENABLED: MPFS FC  
OPTION ENABLED: MPFS iSCSI  
OPTION ENABLED: CLARIION FC  
OPTION ENABLED: CLARIION iSCSI  
OPTION ENABLED: SAN
```

**Translating upgrade licenses into flag options (i.e., MPFS over FC, MPFS over iSCSI, CLARiiON FC, CLARiiON iSCSI):**

→Based on CPQO rules, customer orders SLIC upgrades per rules established by CLARiiON, for the CX4 system that they have, and order appropriate protocol licenses, resulting in the Sales Order

→Field/Support may need to establish if any of the applicable licenses are on the Sales Order

→Use the SLIC I/O Upgrade procedures to help translate what licenses apply to which nas\_hw\_upgrade flag option (Celerra Procedure Generator)

**UBSE System Model Display Rules:**

→If the MPFSIS flag is set during install, system reports as “iS” model

→If MPFSIS flag is set, and the MPFSFC flag, system still reports as “iS” model

→If MPFSFC flag is set, or no flags are set, system reports as plain Integrated model (NS-960)

→If CLARiiON iSCSI, CLARiiON FC, or FC Enabled option during install are set (alone or in combination with each other), then system reports as an “FC” model, with SAN flag option and FC Enabled=Yes flag set

→NAS Upgrades to 5.6.47 will not change display model output

→During factory build-to-ship, system model display is reflected by FC Enabled vs. Integrated choice, and the flag options selected when using the nas\_hw\_upgrade script

→In the field, system model display and options will be changed or updated based on the use of the nas\_hw\_upgrade script, presumably in conjunction with hardware and license upgrade activities

→/nas/log/nas\_xml.xml file contains the displayed model name based on install and flag options used during install/upgrade

**CSA SETUP MPFS RULE:**

→CSA will only offer “Setup MPFS” configuration option if MPFS FC and/or MPFS iSCSI flags are enabled

**EXAMPLES OF SYSTEM MODELS:**

**Base NS-120 Example for CLARiiON iSCSI:**

→A base NS-120 is ordered with a Clariion iSCSI license and physical iSCSI I/O modules for the array

→Clariion iSCSI or Clariion FC automatically means the system is installed with the “FC Enabled” option

→It shows up at CSA configuration time as an “NS-120FC” with the “SAN” option enabled, and the factory used the nas\_hw\_upgrade command to set the -clariionis Flag

# **/nas/sbin/model**

NS-120FC

# **/nas/sbin/model -option -list**

OPTION ENABLED: CLARIION iSCSI

OPTION ENABLED: SAN

# /nas/sbin/nas\_hw\_upgrade -option -enable -clarionis (Flag applies if iSCSI modules ordered with the NS-120 system)

**Base Model without any License Flags Ordered:**

# /nas/sbin/model -option -list

OPTION ENABLED: NOT AVAILABLE  No Flag options are set

# /nas/sbin/model

NS-960

**Example output when only FC Flag set:**

# /nas/sbin/nas\_hw\_upgrade -fc\_option -enable

# /nas/sbin/model

NS-960FC

# cat /etc/be\_sg\_info | grep FC

FC\_ENABLED=YES

# /nas/sbin/model -option -list

OPTION ENABLED: SAN

**Example output when only MPFSIS Flag is set:**

# /nas/sbin/nas\_hw\_upgrade -option -enable -mpfsis

# /nas/sbin/model

NS-960iS

# /nas/sbin/model -option -list

OPTION ENABLED: MPFS iSCSI

# tail /etc/be\_sg\_info

FC\_ENABLED=NO

MPFSI\_ENABLED=YES

**Note:** iS designation used only if the MPFSI Flag is enabled (and the MPFSF Flag can be enabled), but no FC Enabled, no Clariion iSCSI, or no Clariion FC enabled flags are set

**Example output when MPFSIS flag is set, and either Clariion iSCSI or Clariion FC flags are set:**

→ Always results in FC Model + SAN flag, or both, results in FC Model & SAN Flag also set

# /nas/sbin/nas\_hw\_upgrade -option -enable -clarionfc -clarionis

# /nas/sbin/model

NS-960FC

# /nas/sbin/model -option -list

OPTION ENABLED: MPFS iSCSI

OPTION ENABLED: CLARIION FC

OPTION ENABLED: CLARIION iSCSI

OPTION ENABLED: SAN

# tail -4 /etc/be\_sg\_info

MPFSI\_ENABLED=YES

CLARIIONF\_ENABLED=YES

CLARIIONI\_ENABLED=YES

FC\_ENABLED=YES

**UNIFIED STORAGE MODELS:**

**NS-120 Model:**

Base NS-120 offers 2 blades, 120 drives, NAS only (i.e., no licenses)

Array models are AUX, AUXF, & AUXF8

**NS-480 Model:**

Base NS-480 offers 4 blades, 480 drives, NAS only (i.e., no licenses)

Array models are AUX

**NS-960 Model:**

Base NS-960 offers 4 blades, 480 drives, NAS only (no licenses), and 4 or 8Gbps FC Modules on the SPs slots 0 & 1

Array models are AUX and AUX8

**NS-120 UNIFIED MODELS:**

**NS-120 Base NS120-AUX (1) 4-port 4Gbps FC card (slot 0)**

slot\_1 → Can have (1) 2-port 1Gbps iSCSI card or (1) 2-port 10Gbps iSCSI card

slot\_2 → Can have (1) 4-port 4Gbps FC card, or (1) 4-port 8Gbps FC card, or (1) 2-port 1Gbps iSCSI card

**Note:** Single BE port slot\_0 port 0 for DAE's

**NS-120FC Base NS120-AUXF (1) 4-port 4Gbps FC card (slot 0)**

slot\_1 → Can have (1) 2-port 1Gbps iSCSI card or (1) 2-port 10Gbps iSCSI card

slot\_2 → Can have (1) 4-port 4Gbps FC card

**Note1:** Single BE port slot\_0 port 0 for DAE's

**Note2:** The AUXF variant is phased out in Q3 2010

**NS-120FC Base NS120-AUXF8 (1) 4-port 4Gbps FC card (slot\_0)**

slot\_1 → Can have (1) 2-port 1Gbps iSCSI card or (1) 2-port 10Gbps iSCSI card

slot\_2 → Can have (1) 4-port 8Gbps FC card

**Note1:** Single BE port slot\_0 port 0 for DAEs

**Note2:** The AUXF8 variant will go away Q3 2010

**NS-120 SLIC SLOT RULES:**

→ If 10Gbps iSCSI card selected, must go in lowest eligible slot, which is slot\_1

→ Only a single 10Gbps iSCSI card allowed per SP on the NS-120

→ If 8Gbps FC card selected, must go in lowest eligible slot, which is slot\_2

**NS-480 UNIFIED MODELS:**

**NS-480 Base NS480-AUX (2) 4-port 4Gbps FC cards (slot\_0 & 1)**

slot\_2 → Can have (1) 2-port 1Gbps iSCSI card or (1) 2-port 10Gbps iSCSI card

slot\_3 → Can have (1) 4-port 4Gbps FC card, or (1) 2-port 1Gbps iSCSI card, or (1) 4-port 8Gbps FC, or (1) 2-port 10Gbps iSCSI

slot\_4 → Can have (1) 2-port 1Gbps iSCSI card

**Note:** Two BE ports slot\_0 & slot\_1 ports 0 & 1 for DAEs

**NS-480 SLIC SLOT RULES:**

→ If 10Gbps iSCSI card selected, must go in lowest eligible slots, starting with slot\_2 and/or slot\_3

→ Add any 1Gbps iSCSI to the right of 10Gbps iSCSI slots

→ If 8Gbps FC card selected, must go in slot\_3

→ Only (2) 10Gbps iSCSI cards allowed per SP on NS-480

**NS-960 UNIFIED MODELS:**

**NS-960 Base1 NS960-AUX (2) 4-port 4Gbps FC cards (slot\_0 & 1)**

**NS-960 Base2 NS960-AUX (2) 4-port 8Gbps FC cards (slot\_0 & 1)**

slot\_2 → Can have (1) 4-port 4Gbps, or (1) 2-port 1Gbps iSCSI, or (1) 4-port 8Gbps FC, or (1) 2-port 10Gbps iSCSI (FC can be used for Host, Blade, or Backend connectivity)

slot\_3 → Can have (1) 4-port 4Gbps, or (1) 2-port 1Gbps iSCSI, or (1) 4-port 8Gbps FC, or (1) 2-port 10Gbps iSCSI (FC for Host or Blade connectivity)

slot\_4 → Can have (1) 4-port 4Gbps, or (1) 2-port 1Gbps iSCSI, or (1) 4-port 8Gbps FC, or (1) 2-port 10Gbps iSCSI (FC Host connectivity)

slot\_5 → Can have (1) 2-port 1Gbps iSCSI card, or (1) 2-port 10Gbps iSCSI card, or (1) 4-port 4Gbps FC, or (1) 4-port 8Gbps FC (FC Host connectivity)

**Note:** Two BE ports slot\_0 & slot\_1 ports 0 & 1 for DAEs, plus optional (4) BE ports if FC card added to slot\_2

**NS-960 SLIC SLOT RULES:**

→ If >than 480 drives needed, add a 4Gbps FC card in slot\_2 for BE DAEs

→ If 10Gbps iSCSI card selected, must go in lowest eligible slots, starting with slot\_2, slot\_3, slot\_4, or slot\_5, but only a max of (20 of these to a configuration

→ Max. of (2) 10Gbps iSCSI cards per SP

→ Add any 1Gbps iSCSI to the right of 10Gbps iSCSI slots

→ If 8Gbps FC card selected, goes in lowest eligible slot, starting with slot\_2, 3, or 4

→ Add any 4Gbps FC cards to the right of 8Gbps FC cards, for max. of (2)/configuration

**AUX ARRAY MODELS:**

NS120-AUX → NS-120 with 4Gb FC slot\_0 only

NS120-AUXF → NS-120 with 4Gb FC slot\_0 & slot\_2

NS120-AUXF8 → NS-120 with 4Gb FC slot\_0 & 8Gb FC slot\_2

NS480-AUX → NS-480 with 4Gb FC slot\_0 & slot\_1

NS960-AUX → NS-960 with 4Gb FC slot\_0 & slot\_1

NS960-AUX8 → NS-960 with 8Gb FC slot\_0 & slot\_1

**Additional SLIC Models:**

AUX-M4GF-FE-A = 4Gb FC

AUX-M8GF-FE-A = 8 Gb FC

AUX-M1GI-A = 1Gb iSCSI

AUX-MXGI-A = 10Gb iSCSI

AUX-M4GF-BE-A = 4Gb FC for BE (Backend)

**LICENSE MODELS WITH TRANSLATIONS:**

NS120-NBOPTL + NAV-NS120 → Native FC/iSCSI license + Navisphere lic.

NS480-NBOPTL + NAV-NS480 → Native FC/iSCSI license + Navisphere lic.

NS96-NBNV480L → Native FC/iSCSI license for < than 480 drives, with Navisphere

NS96-NBNV960L → Native FC/iSCSI license for > than 480 drives, with Navisphere

**Note:** Existing rules apply if ordering MPFS over FC or MPFS over iSCSI

**Example MPFS over FC (Linux, 1 Host, 2 CPUs):**

MPFS-CPU-LIC + MPFS-CPU-T1 (qty 2) + MPFS-LNX-DCD

**EXAMPLE OF HOW TO CHANGE MODEL DISPLAY:**

**Scenario:**

# /nas/sbin/model

NS-960

# cat /etc/be\_sg\_info

--edited for brevity----

MPFSF\_ENABLED=YES

# /nas/sbin/model -option -list

OPTION ENABLED: MPFS FC

**Changing from NS-960 to NS-960iS Model:**

# /nas/sbin/nas\_hw\_upgrade -option -enable -mpfsis

Checking if running on primary CS...yes

Checking if running as root...yes

Checking for integrated system...yes

Checking if model is supported...yes

Checking if backend is supported...yes

Options enabled:

MPFSI\_ENABLED=YES

Model: NS-960iS

Command succeeded

# /nas/sbin/model

NS-960iS

# tail /etc/be\_sg\_info

-----edited for brevity-----

MPFSF\_ENABLED=YES

MPFSI\_ENABLED=YES

# /nas/sbin/model -option -list

OPTION ENABLED: MPFS FC

OPTION ENABLED: MPFS iSCSI

**USE CASE: NS-960 SP IO MODULE UPGRADE 5.6.47**

**Note:** Started with simple NS-960 system, with only (2) FC IO Modules per SP, in slots\_0 & 1. Used the NST tool to add a single 2-port iSCSI IO Module to SPs in Slot\_2 (A2 & B2).

**Current System configuration before licensing update:**

# /nas/sbin/model

NS-960

# /nas/sbin/model -option -list

OPTION ENABLED: NOT AVAILABLE

**Updating Hardware Licensing per procedure:**

# /nas/sbin/nas\_hw\_upgrade -option -enable -mpfsis -clarionis

Checking if running on primary CS...yes

Checking if running as root...yes

Checking for integrated system...yes

Checking if model is supported...yes

Checking if backend is supported...yes

Options enabled:

MPFSI\_ENABLED=YES

CLARIIONI\_ENABLED=YES

FC\_ENABLED=YES → This equals the OPTION ENABLED: SAN seen in model -option -list output

Model: NS-960FC

Command succeeded

**Note:** Creates /etc/be\_sg\_info.orig backup file, then adds the new licensing options to the /etc/be\_sg\_info file. All nas\_hw\_upgrade activity is captured in the /nas/log/nas\_hw\_upgrade.log file.

/nas/log/nas\_hw\_upgrade.log

**Results:**

# /nas/sbin/model

NS-960FC

# /nas/sbin/model -option -list

OPTION ENABLED: MPFS iSCSI

OPTION ENABLED: CLARIION iSCSI

OPTION ENABLED: SAN

# tail /etc/be\_sg\_info → HW Upgrade added following lines to the be\_sg\_info file, read by the model –option –list command

MPFSI\_ENABLED=YES

CLARIIONI\_ENABLED=YES

FC\_ENABLED=YES

**Note:** Running the nas\_hw\_upgrade –option –disable command will reverse the upgrade

**USE CASE: MPFSFC OPTION**

# /nas/sbin/nas\_hw\_upgrade –option -enable -mpfsfc

Checking if running on primary CS...yes

Checking if running as root...yes

Checking for integrated system...yes

Checking if model is supported...yes

Checking if backend is supported...yes

Options enabled:

MPFSF\_ENABLED=YES

Model: NS-960

Command succeeded

# /nas/sbin/model

NS-960

# /nas/sbin/model –option -list

OPTION ENABLED: MPFS FC

# tail /etc/be\_sg\_info

-----edited-----

MPFSF\_ENABLED=YES

**IDENTIFYING CELERRA SYSTEM REMOTELY:**

→ New Serviceability feature ‘nas\_inventory’ allows for remote identification of hardware configurations on the Celerra & Array

→ Consists of Collection Script, Cron Job, nas\_inventory command, and Inventory Cache file for CallHome delivery

→ Overall, the Cron Job uses /nas/sbin/collect\_inventory to maintain a system inventory cache, written and updated to the /nas/log/inventory\_status file, and is delivered in the XML payload for all CallHome events & Email User notifications

→ nas\_inventory command takes what already existed in the Celerra Manager Inventory tab and converts to CLI output

**NAS Inventory CLI:**

# nas\_inventory –list –location –info –all | –tree | –report –fields –noheader

# nas\_inventory –report –fields location, status

-list switch outputs all components by Name, Type, Status, and System ID

-location switch shows component locations & how the location is defined, which can then be used with -info

-tree switch lists the components in a Tree hierarchy

-info displays details of all components; specify component by “ “ location field to output details of that component, similar to output seen in RESUME

-report switch is hidden, displays information on components, such as part #'s

**Collect Inventory Script:**

/nas/sbin/collect\_inventory

→ collect\_inventory Cron is run 15 minutes past the hour, whenever NAS is started, or manually by a User, and calls nas\_inventory and /nas/sbin/model –option –list commands to create the inventory\_status file

**Cron Job:**

/nbsnas/site/cron.d/nas\_sys

15 \* \* \* \* nasadmin /nas/sbin/collect\_inventory > /dev/null 2>&1 (15 minutes past every hour)

**Inventory Cache File:**

/nas/log/inventory\_status

→ Cron Job creates & then updates the status of system components in a /nas/log/inventory\_status cache file

→ The inventory\_status file maintains most recent output from the collect\_inventory script, which in turn calls /nas/bin/nas\_inventory & /nas/sbin/model –option –list. Only the most recent nas\_inventory & model –list outputs are maintained in the inventory-status file.

/nas/sbin/model –option -list

/nas/bin/nas\_inventory -list

# cat /nas/log/inventory\_status

System Configuration Snapshot: Thu Nov 12 09:15:01 EST 2009

Control Station Slot No: 0

/nas/sbin/model -option -list:

OPTION ENABLED: CLARIION iSCSI

OPTION ENABLED: SAN (output abbreviated)

**CallHomes/Email Notifications:**

→The “inventory\_status” cache file is delivered as a “System Configuration Snapshot” with every CallHome event and is parsed into the CDATA XML section of the CallHome payload

**Relevant Troubleshooting Logs:**</nas/log/webui/cli.log></nas/log/apl/tm.log>**Sample Outputs:**# [nas\\_inventory -list](#)

| Component                           | Type            | Status | System ID                           |
|-------------------------------------|-----------------|--------|-------------------------------------|
| Battery A                           | Battery         | OK     | CLARiiON CX4-960 FNM00083800203     |
| Battery B                           | Battery         | OK     | CLARiiON CX4-960 FNM00083800203     |
| Celerra NS-960FC FNM000838002030000 | Celerra         | OK     | Celerra NS-960FC FNM000838002030000 |
| CLARiiON CX4-960 FNM00083800203     | CLARiiON        | OK     | CLARiiON CX4-960 FNM00083800203     |
| Control Station 0                   | Control Station | OK     | Celerra NS-960FC FNM000838          |

-----abbreviated-----

# [nas\\_inventory -info -all](#)

Location = system:NS-960FC:FNM000838002030000  
 Component Name = Celerra NS-960FC FNM000838002030000  
 Type = Celerra  
 Status = OK  
 Variant = NS-960FC  
 Version = 5.6.47-8  
 Serial Number = FNM00083800203  
 -----abbreviated-----

# [nas\\_inventory -tree](#) (abbreviated example)

| Component                           | Type       | Status |
|-------------------------------------|------------|--------|
| Celerra NS-960FC FNM000838002030000 | Celerra    | OK     |
| Data Mover 2                        | Data Mover | OK     |
| Data Mover 2 IO Module 0            | IO Module  | OK     |
| Data Mover 2 IO Module 0 SFP 0      | SFP        | OK     |
| Data Mover 2 IO Module 0 SFP 1      | SFP        | OK     |
| Data Mover 2 IO Module 0 SFP 2      | SFP        | OK     |
| Data Mover 2 IO Module 0 SFP 3      | SFP        | OK     |

# [nas\\_inventory -report -fields location,status](#)

"location","status"  
 "system:NS-960FC:FNM000838002030000","OK"  
 "system:NS-960FC:FNM000838002030000\clariionSystem:CX4-960:FNM00083800203","OK"  
 "system:NS-960FC:FNM000838002030000\clariionSystem:CX4-960:FNM00083800203\cache:write:","OK"

# [nas\\_inventory -info "system:NS-960FC:FNM000838002030000\clariionSystem:CX4-960:FNM00083800203\iomodule::A0"](#)

Location = system:NS-960FC:FNM000838002030000\clariionSystem:CX4-960:FNM00083800203\iomodule::A0  
 Component Name = IO Module A0  
 Type = IO Module  
 Status = OK  
 Variant = 4 PORT FIBRE IO MODULE  
 Storage System = CLARiiON CX4-960 FNM00083800203  
 Serial Number = CF2YW082602082  
 Part Number = 103-054-100C  
 History = EMC\_PART\_NUMBER:103-054-100C  
 EMC\_ARTWORK\_REVISION:C01  
 EMC\_ASSEMBLY\_REVISION:C03  
 EMC\_SERIAL\_NUMBER:CF2YW082602082  
 ---lines deleted for display----  
 VENDER\_SERIAL\_NUMBER:N/A  
 VENDOR\_NAME:N/A  
 LOCATION\_OF\_MANUFACTURE:Hopk, MA USA  
 YEAR\_OF\_MANUFACTURE:2008  
 MONTH\_OF\_MANUFACTURE:10  
 DAY\_OF\_MONTH\_OF\_MANUFACTURE:02

## **CELERRA DEDUPLICATION PHASE 2 (Celerra F-RDE):**

- Celerra Deduplication feature was introduced with NAS 5.6.43
- Deduplication is file compression and single-file instancing of identical data files
- Celerra “compresses” files for file system space savings (estimated average 50% savings), & “single instances” identical files by keeping a hashed copy of the data in the hidden data store, while stub files are kept for all inodes that reference the same content in the file system
- Celerra uses EMC RecoverPoint technology for file compression
- Celerra uses EMC Avamar hashing technology for single file instancing
- Celerra Deduplicatiion differs from Data Domain deduplication in that the former is an inline dedupe for backup to disk, while Celerra supports production & archival file dedupe

### **5.6.47 introduces a number of major enhancements:**

- Sparse write capability, the ability to write to existing deduplicated files without having to “recall” or “rehydrate” the file while writes take place (use of migration inode in the stub file with updated block addresses)
- Sparse write also allows deduplication of file sizes up to 8TB now without CIFS client timeouts (writes directed to migration inode blocks in stub files)
- If deduplicated files are modified so that size of changed blocks plus total number of blocks is larger than logical size by 10%, then deduplication of the file will be dynamically done, as opposed to waiting for next deduplication scan
- Support for VMWare in NFS environments for duplication of large VMDK files
- Support for large files without rehydration
- Ability to deduplicate “active” files, not just “inactive” ones
- NDMP PAX Backups/Restores preserves “compressed” file sizes, but not “single instanced” files since NDMP requires all instances of file in stream
- Ability to apply deduplication parameters to specific file systems [i.e., on a per file system basis vs. parameters that applied to the whole Data Mover only]
- Ability for CIFS clients to see & change compression attributes on files/folders
- Prior to 5.6.47, dedupe parameters were applied using server\_param –facility dedupe command. With 5.6.47 and above, we now use the fs\_dedupe command to change param values at a file system or Data Mover level.
- Feature does not support block-level deduplication, such as for iSCSI Luns

### **What can CIFS Clients do?**

#### **CIFS CLIENT COMPRESSION BEHAVIOR:**

- >CIFS clients can view compression attributes on files & folders, as well as manipulate them, but only on a file system with Deduplication enabled, and with the cifs\_compression\_enabled attribute turned on (which it is by default)
  - Use Explorer>Properties to see the true file ‘Size’ vs. “Size on disk” (the latter of which is the compressed size)
  - Use Explorer>select file or folder >Properties>Advanced> \_\_Compress contents to save disk space
  - When selecting a directory to compress, the CIFS client can elect to compress just that folder, or all subfolders and files
- Note:** The technical materials refers to the Client action as bypassing the dedupe policies, since the client can compress any file or folder that it chooses, and the Celerra dedupe engine does not compress folders
- By default, once deduplication has been turned on for a file system, the cifs\_compression\_enabled attribute is enabled by default
  - Only files meeting the minimum size requirements are compressed by Celerra deduplication during scanning
  - Folders are not automatically compressed by Dedupe, only by Client action using Explorer
  - For a compressed folder, any uncompressed files that are copied to it, will be compressed
  - For an uncompressed folder, any compressed files that are copied to it, will be uncompressed
  - CIFS client can compress entire file system by selecting toplevel folder and selecting Advanced, then box to “Compress contents to save disk space” and then “\*Apply changes to this folder, subfolders and files”

**Note:** One of the effects of this is that all files, regardless of size, show as compressed.

→Turning off the cifs\_compression\_enabled attribute for a particular file system will uncompress all files and folders that may have been previously compressed, either by dedupe or CIFS client

→Explorer popup error seen if trying to compress files or folders without deduplication enabled on the Celerra file system

#### **“Error Applying Attributes:**

#### **An error occurred applying attributes to the file.”**

#### **Behavior when turning on or off the cifs compression enabled parameter on a file system:**

- If cifs compression is set to no, all files that were compressed per the deduplication engine on Celerra, will show as uncompressed on the clientside. However, the system “remembers” what the previously compressed files were, because once you restore the cifs compression param to yes, all originally compressed files will again show as compressed.
- Similar behavior is seen when Client forces compression of all folders and files. If compression is disabled, then all files & folders show up as uncompressed. However, once cifs compression is re-enabled, the system remembers and all files and folders that were specifically compressed before, will show that attribute again.

#### **DEDUPLICATION UPGRADE STORY WITH 5.6.47:**

- See emc223821 for some background

→If Deduplication is not enabled on any file systems prior to 5.6.47, then the NAS 5.6.47 upgrade has zero impact and is a non-issue  
→If Deduplication is enabled prior to 5.6.47, the Upgrade to 5.6.47 will not impact anything, but any deduplicated files that are subsequently written to “after” the upgrade, will cause the file system to be updated with a new ‘magic’ number 0xdbadface (pre 5.6.47 magic number is 0xbadface)

→Engineering’s recommendation is to backup any file systems that have deduplication enabled prior to upgrading to 5.6.47  
→Once a file system running 5.6.47 has had the Magic number updated, then that file system becomes unmountable on any earlier NAS versions

→The 5.6.47 Upgrade code itself requires a mandatory reboot of the Standby Server, the idea being that if a problem occurs on a Primary server and it fails over, then the file systems will be able to remount, whereas if the Standby was still running a previous NAS version, it might not be able to mount the file systems

**Note:** Lab testing would indicate in practice that this situation would not really happen, since you need to be able to mount the file system first at 5.6.47, then have a User write to a “deduplicated” file, before the magic number would be updated, and since the server would thus already be running 5.6.47, the purported problem would not exist

#### **What happens to pre-5.6.47 file systems running Deduplication, once upgraded to 5.6.47?**

→After the upgrade to 5.6.47, the file system Magic number would be updated either after new files were deduplicated, or whenever the first pre-existing deduplicated file was written to

#### **Server log:**

2010-01-20 09:41:55: DEDUPE: 4: last message repeated 1 times

2010-01-20 09:41:55: DEDUPE: 6: Domain initialized for fsid 28

2010-01-20 09:53:18: DEDUPE: 6: Replacing magic number 0xbadface(195951310) with RDE magic number 0xdbadface

2010-01-20 09:53:18: DEDUPE: 6: Updated on-disk super blocks with RDE magic number change

# /nas/tools/\_fs\_db server\_2 readsb 102 psb |grep magic

About to read superblock from volume 102 and sector 16

magic=0xdbadface

#### **Trying to mount 5.6.47 deduplication file system with NAS Versions prior to 5.6.47:**

→As long as the magic number was not updated on the file system running Deduplication, it could still be mounted on pre-5.6.47 NAS code

→For 5.6.47 systems where the file system magic number has been updated to 0xdbadface, the file systems would be unmountable by any pre-5.6.47 code

#### **Trying to mount 5.6.47 file system with new magic number on 5.6.46 Server:**

# server\_mount server\_2

file2 on /file2 ufs,perm,rw,<unmounted>,<corrupt>

#### **Server log:**

2010-01-20 10:17:18: UFS: 6: Volume name:102

2010-01-20 10:17:18: UFS: 3: invalid superblock 0xd45a4000:

magic 0xdbadface, bsize 0x2000, ssize 0x2000

should be 0x11954 or 0xbadface, 0x2000, 0x2000, <= 0x2000

2010-01-20 10:17:18: UFS: 4: 7: File system 28 is corrupted and will remain unmounted.

2010-01-20 10:17:18: UFS: 3: initFS failed, initRootNode returned status DirtyFileSystem

2010-01-20 10:17:18: ADMIN: 3: Command failed: file mount ufs rw /file2 102=28 rw

2010-01-20 10:17:18: ADMIN: 3: Command failed: export "/file2"

**Note:** Despite the log entries above, the file system isn’t really corrupted at all and can be mounted on 5.6.47 DART

→Support for Thin file cloning, which are writeable snapshots of files (up to 2000 clones). Useful in VMWare environments where VMDK can be distributed to multiple users

#### **UPGRADE ISSUE FOR IP REPLICATION:**

**Note:** If you upgrade the Source side to 5.6.47, but Destination is running 5.6.43, etc., the destination file systems will show up as unmounted and corrupt. From this actual case, the recommendation was to stop the replication sessions until the destination could be upgraded to 5.6.47, at which point Replication should be able to be restarted. See emc230290.

fs01 on /fs01 ufs,perm,ro,<unmounted>,<corrupt>

fs02 on /fs02 ufs,perm,ro,<unmounted>,<corrupt>

#### **MODIFYING DEDUPLICATION PARAMS:**

##### **Setting Deduplication Values from Celerra Manager:**

→Can turn on deduplication when creating new file system

→File Systems tab ‘Deduplication Settings’ section, used to modify params globally or per file system

→No separate License required for deduplication

→Can also fine-tune duplication from CLI using fs\_dedupe

##### **Modifying Deduplication Parameters via CLI:**

→Use the server\_param facility prior to 5.6.47. After 5.6.47, use the fs\_dedupe facility to change params on a per file system basis, or for changing global Data Mover parameter values.

# server\_param server\_2 -facility dedupe -list

**Note:** Only outputs global parameter values that have been changed from the defaults

**Use fs\_dedupe to tune deduplication parameters on a file system basis:**

**# fs\_dedupe -modify fs1 -access\_time 0**

**# fs\_dedupe -modify fs1 -modification\_time 0**

**Use fs\_dedupe to list, modify, or clear (back to default) deduplication parameter settings on a Server basis:**

**# fs\_dedupe -default -set server\_2 -access\_time 0** →Changes global default values on the Server

→Changes global default values on a Server basis. All file systems created and enabled for deduplication will inherit the global Server values.

**RESTORING DEFAULTS GLOBALLY & PER FILE SYSTEM:**

**Setting Deduplication Values back to default values on a Server Basis:**

**# fs\_dedupe -default -clear server\_2**

→Resets the Server back to its original code default values

**Setting Deduplication Values back to default values on a File System:**

**# fs\_dedupe -clear file6**

→You can also reset or change default values for individual params on a file system

**VERIFYING PARAM VALUES GLOBALLY AND PER FILE SYSTEM:**

**Verifying Current Global Server Dedupe Values:**

**# fs\_dedupe -default -info server\_2**

Server parameters:

server\_2

|                             |   |
|-----------------------------|---|
| Case Sensitive              | = no                                    |
| Duplicate Detection Method  | = sha1                                  |
| Access Time                 | = 15 →Example of a global default value |
| Modification Time           | = 15                                    |
| Minimum Size                | = 24 KB                                 |
| Maximum Size                | = 8388608 MB                            |
| File Extension Exclude List | =                                       |
| Minimum Scan Interval       | = 7                                     |
| Savevol Threshold           | = 90                                    |
| Backup Data Threshold       | = 90                                    |
| CPU % Usage Low Water Mark  | = 40                                    |
| CPU % Usage High Water Mark | = 75                                    |
| Cifs Compression Enabled    | = yes                                   |

**Verifying Original Code Default Global Server Values:**

**# .server\_config server\_2 -v "param dedupe" |egrep "access|modification"**

dedupe.accessTime 0x0361f3f8 0x00000000 0x0000000f

dedupe.modificationTime 0x0361f438 0x00000000 0x0000000f

**Verifying File System Values:**

**# fs\_dedupe -info file6**

|               |         |
|---------------|---------|
| Id            | = 54    |
| Name          | = file6 |
| Deduplication | = On    |
| Status        | = Idle  |

As of the last file system scan (Thu Dec 17 11:02:14 EST 2009):

|                      |  |
|----------------------|--|
| Files scanned        | = 40   |
| Files deduped        | = 29 (73% of total files)                      |
| File system capacity | = 984 MB                                       |
| Original data size   | = 220 MB (22% of current file system capacity) |
| Space saved          | = 200 MB (91% of original data size)           |

File system parameters:

|                             |              |
|-----------------------------|--------------|
| Case Sensitive              | = no         |
| Duplicate Detection Method  | = sha1       |
| Access Time                 | = 0          |
| Modification Time           | = 0          |
| Minimum Size                | = 24 KB      |
| Maximum Size                | = 8388608 MB |
| File Extension Exclude List | =            |
| Minimum Scan Interval       | = 7          |
| Savevol Threshold           | = 90         |
| Backup Data Threshold       | = 90         |

Cifs Compression Enabled = yes

Pathname Exclude List =

### When are Files Reduplicated?

--when deduplication is turned off

--when CIFS client decompresses a file

--when DHSM API reduplicates a file

--when sum of written blocks + size in EDRS store is greater than logical file system by 10%

### Exposing the EDRS File Store on a given file system (Do not expose the store without Eng. involvement):

# .server\_config server\_2 -v "storeHost exposeRepository /file3"

1258129678: DEDUPE: 6: Throttle running in Full mode (1% busy)

1258129678: UFS: 7: inc ino blk cache count: nInoAllocs 2: inoBlk 202df904

1258129678: DEDUPE: 6: StoreHost: ExposeRepository succeeded

/nasmed/quota/slot\_2/file3/etc

# ls -la

total 288

dr-xr-xr-x 3 root bin 80 Nov 13 11:27 .

drwxr-xr-x 7 root root 1024 Nov 13 11:08 ..

**drwx----- 5 root bin 1024 Nov 13 11:27 EDRS\_v1**

-r--r--r-- 1 root bin 262152 Nov 12 13:22 gid\_map

# cd EDRS\_v1;ls -la

drwx----- 12 root bin 1024 Nov 13 11:21 AQAAAAA

drwx----- 2 root bin 1024 Nov 12 14:57 Q09ORg

drwx----- 2 root bin 1024 Nov 13 11:21 VEVNUAABAAAAA

### Hiding the EDRS Store:

# .server\_config server\_2 -v "storeHost hideRepository /file3"

1258130068: DEDUPE: 6: Throttle running in Full mode (1% busy)

1258130068: UFS: 7: inc ino blk cache count: nInoAllocs 2: inoBlk 202df904

1258130068: DEDUPE: 6: StoreHost: HideRepository succeeded

### SavVolAtCapacity Behavior (emc229138):

→ Deduplication scanning will abort if Snapsure checkpoints are in use, and the SavVol usage is at 81% capacity or higher (though param suggests that it should not abort until 90% of SavVol is used)

**Note:** Eng. explanation is that the dedupe value represents 90% of the SavVol HWM, which by default is 90%. So, 90% of 90 is 81.

→ See Server Log messages and Celerra Manager alerts

→ This is the design behavior for Deduplication—don't want to have deduplication extending SavVol by default

### Workarounds:

--See emc229138 for all the workarounds and details

--Extend SavVol on the file system using Celerra Manager > File Systems > rightclick file system name > Checkpoints > Extend Checkpoint Storage

--Or, disable the “Savevol Threshold” deduplication param by setting to 0

# fs\_dedupe -modify file1 –savvol\_threshold 0

### FEATURE INTEROPERABILITY WITH DEDUPE:

--NDMP PAX backup/restore supports compressed files, but not ‘single instance’ files

--NDMP VBB backup/restore supported, but not single file or file-by-file restore

--Block Quota policy (size on disk) is affected by deduplication, but not File Size Quota policy, which relates to logical file sizes

--Snapsure can increase SavVol space used when using deduplication (changes)

--MPFS clients will fallback to NFS/CIFS when accessing deduplicated files

--ReplicationV2 supports deduplication

--Replication V1 is not supported—See emc242296:

1) In the case of an existing file system with deduplication already enabled, if you attempt to setup Replication V1, all the steps will succeed without error, yet the Replication Session itself does not get established.

### Example:

# cat /nas/log/cmd\_loggrep dupe1

2010-05-27 16:35:46.307 db:0:12778:S: fs\_copy -start dupe1\_ckpt2 dupe1\_dest:cel=nx4-2 -fromfs dupe1\_ckpt1 -option monitor=off

2010-05-27 16:35:56.864 db:0:12778:E: fs\_copy -start dupe1\_ckpt2 dupe1\_dest:cel=nx4-2 -fromfs dupe1\_ckpt1 -option monitor=off

**Note:** So, even though all commands succeeded, including the above incremental for the 2<sup>nd</sup> checkpoint, no RepV1 session actually started

2) In the case where Replication V1 is already established, the commandline will provide an Error 12050 when trying to turn on deduplication, preventing deduplication from being started.

### Example:

# fs\_dedupe -modify dupe2 -state on

Error 12050: The file system contains Replicator V1 sessions.

## **EMC CELERRA FAST SOLUTION (Fully Automated Storage Tiering):**

→Search for “EMC CELERRA FAST” for a complete description of the solution

### **TOP TALKERS FEATURE (5.6.47) [-top replaced in 6.0]:**

→Feature to help customers identify NFS/CIFS clients consuming the most resources, and to assist with identifying storage problems

→Server\_stats and counters have been enhanced to identify NFS/CIFS Clients

→CLI version only with this release, no GUI support

→Default list of TopTalkers is for 20 systems, but is customizable

# server\_param server\_2 -facility statmon -list|grep top

```
top.updateInterval      statmon      0      0
top.maxClients         statmon      20     20
```

\$ server\_param server\_2 -facility statmon -modify toptalker.maxClients (default client list is 20, editable from 1-256 max)

\$ server\_param server\_2 -facility statmon -modify toptalker.updateInterval -value xx (off by default, 0-86400 seconds)

→Update interval off by default (configure from 0-86400 secs.), recommendation to use 600 secs to minimize DART resource impact

→Update interval defines frequency that DART polls network connections to build top talkers list—runs internally

#### **Turning Feature On:**

# server\_stats server\_2 -top cifs

ERROR (13421969432): The top collection options are currently disabled on the server\_2 Data Mover.

# server\_param server\_2 -facility statmon -modify top.updateInterval -value 5

server\_2 : done

**Note:** Must set an updateInterval value to begin gathering Top Talker statistics

#### **Example Commands:**

# server\_stats server\_2 -top

```
[{ -summary {basiclcifslnfslcaches}{...}
  | -table {netldvollfs vollcifslnfs}{...} | -table {netldvollfs vollcifslnfs}
    | -sort <field_name> [-order {asc|desc}] [-lines <lines_of_output>]
    | -count <count> [-port <port>] [-interval <seconds>]
    | -terminationsummary {no|yes|only} [-format {text|csv}]
    | -type {ratelendifflaccu} [-titles {never|once|repeat}]}
```

# server\_stats server\_2 -top cifs -te no -i 5

# server\_stats server\_2 -top cifs | nfs

# .server\_config server\_2 -v "topTalker" "topTalker cifs" "topTalker nfs"

## **FAST UNMOUNT AND FAST MOUNT ON REBOOT OR FAILOVER:**

→Feature designed to improve reboot and failover/fallback times

→Focus is on improving file system unmounting and mounting

→Mount begin & end commands are executed in parallel (async mounts)

→Unmount time reduced by multi-threading vnode & data flush operations

**Note:** From a global freelist traversal of each file system to a multi-threaded single traversal of the global freelist

→For Failover, CS sends bulk request to Data Mover to mount file systems, file systems mounted one by one by DART, but mount end is sent back to CS after all are mounted

→For Reboots, unmount and remount operations are done using boot.cfg order

→Mount requests are queued and handled by a default group of 64 threads

# .server\_config server\_2 -v "param cfs" | grep Parallel

```
cfs.maxParallelMount 0x01fb1edc 0x00000040 0x00000040
```

**Note:** 40 hex = 64 decimal

## **BOOT.CFG & SERVER LOG ENTRIES:**

### **Pre-5.6.47 mount entries:**

mount begin

file mount udfs rw / 12=2 rw

----edited----

file mount udfs rw /CSAShare 184=54 rw

mount end

### **New Fastmount entries in 5.6.47:**

mount fastmount

mount begin

file mount udfs rw / 12=2 rw

----edited----

file mount udfs rw /file5 97=28 rw

mount end

# server\_log server\_2 -s |grep mount

2009-11-12 22:58:14: ADMIN: 6: Command succeeded: mount fastmount

2009-11-12 22:58:14: ADMIN: 6: Command succeeded: mount begin

-----edited----

2009-11-12 22:58:14: ADMIN: 6: Command succeeded: mount end

#### **Fast Mount Issue with NMFS:**

→See emc233009 to disable the ‘Fast Mount’ feature if using NMFS

#### **CELERRA SRDF HEALTHCHECK HARDENING:**

→Purpose of this feature is to verify SRDF state and configuration

→Healthchecks are run at end of nas\_rdf –init and –localinit

→Healthchecks are run at beginning of nas\_rdf –activate & -restore

→These healthchecks are not built into the PUHC

→When problems noted, an appropriate CCMD error message with recommended action is printed to screen and to the /nas/log/nas\_rdf.log

#### **(6) Warning Messages Available:**

##### **MessageID = 17726767122**

--Destination Standby slots do not matchup with Source slots

--Run nas\_rdf –init on the destination system to fix

--Message can be generated at nas\_rdf –activate or after –init

##### **MessageID = 13431799813**

--Checks if Source R1’s are RW mounted and Destination R2’s

--Log into Source to verify if file systems mounted and Data Movers restored

--Issue nas\_rdf –restore on Source if necessary

--Message can be generated before nas\_rdf –restore operation

##### **MessageID = 17726767105/13431799811**

--Checks if all devices synchronized between Source and Destination

--Check devices on each side and determine action required

--Message can be generated before –init, before –activate, before –restore

##### **MessageID = 17726767117**

--Device Group configuration check

--Compare symrdf list of devices to symrdf –g device group devices, and add missing devices to appropriate device groups

--Message can be generated after –init, after –localinit, before –active, and before –restore

##### **MessageID = 17726767118**

--Match devices via probe to those in device groups

--Add missing devices

--Message can be generated after –localinit or –init, before –activate or –restore

##### **MessageID = 17726767119**

--Display failed or degraded backend devices (symdev list –service\_state notnormal)

--Verify devices and escalate to Service Provider if required

--Message can be generated after –localinit & -init, and before –activate & -restore

#### **FLARE 29 TAURUS SUPPORT (04.29.000.5.003):**

→5.6.47 introduces Integrated & Gateway support for Flare 29

→With Flare 29, SPs no longer require rebooting for changing IP addresses (Feature added by CLARiiON per NAS request due to Proxy Arp requirements). For Celerra Integrateds, make sure to use the clarion\_mgmt –modify –spa\_ip –spb\_ip when changing IP addresses on the SPs.

→Flare 29 introduces CLARiiON LUN shrink, but only for Windows 2008 Server Host LUNs with this initial Flare release (Basic & Dynamic disk volumes)

→Total LUNs per Storagegroup increases to 1024 for CX4-480, CX4-960 (Gateways & NS-480/960)

→Total LUNs per Storagegroup increases to 512 for CX4-120 backends (Gateways & NS-120)

**Note:** setup\_clariion & CPW updated

→Support for 2TB SATA drives (low power 5400rpm drives)

→Flare 29 provides support for 2-port Ultraflex 10Gbps iSCSI ‘Poseidon’ modules with optical SFPs on the array [10Gbps iSCSI can be used as Native iSCSI solution, MirrorView solution, or MPFS over iSCSI solution]

→Flare 29 NDU’s have a “Celerra Compatibility Check” screen where the user must select a box certifying that the NAS Version is compatible with the FLARE version

#### **CLARiiON DISK POWER SAVINGS FEATURE:**

Celerra supports file systems hosted on SATA II disks in Raid Groups (Storage pools) using the Flare 29 power management spindown capability after an established idle I/O period (CLARiiON default is 30 minutes)—Celerra will tolerate spinups after an I/O request is generated

**Note:** Initial testing did not show that 5.6.47 worked, but did at 5.6.48.

#### **Enabling Power Savings on Array and on Raid Groups for SATA II Drives:**

- 1) Rightclick Array serial > Power Savings > Select checkbox to “Enable Storage System Power Savings”, and checkboxes for desired RAID Groups to “Allow Power Savings”  
> Click on “Allow Power Savings” for qualified Raid Groups under “Storage System Pool Settings”
- 2) Rightclick Raid Group > Power Savings > Click on “click here to enable”

#### Using CLI to set Power Savings Globally on the Array:

# /nas/sbin/navicli -h 10.241.168.150 powersaving -globalsettings on

Turning Global Power Saving settings on will spin down all eligible unused disks and all Power Saving settings on Storage Pools will take effect. Do you want to proceed?(y/n) y

#### Verifying Power Savings Mode from GUI:

Rightclick Raid Group > Power Savings > Click on Disks tab and look at “State” column

Or, expand the Physical >Enclosure > Disks display in Navisphere, and observe the drive power state

**Note:** When actually in Power Savings mode, the GUI will show Disk Drives or Raid Group disk drives in the “Low Power” state

#### Verifying Power Savings Mode Enabled on Array from CLI:

# /nas/sbin/navicli -h 10.241.168.150 powersaving -info

System Idle Time: 30 minutes

Global Power Saving Settings: ON

#### Verifying Power Savings Mode Enabled on specific Raid Group:

# /nas/sbin/navicli -h 10.241.168.150 getrg 8 -all |grep Power

Power Savings Setting: ON

RAID GROUP Power Savings Eligible: YES

Is RAID GROUP in Power Savings Mode: NO

**Note:** The ‘Is RAID GROUP in Power Savings Mode: NO’ line indicates that the disks are not in “Low Power” savings mode

#### Example when Disks are in Power Savings Mode YES:

# /nas/sbin/navicli -h 10.241.168.150 getrg 8 -all |grep Power

Power Savings Setting: ON

RAID GROUP Power Savings Eligible: YES

Is RAID GROUP in Power Savings Mode: YES

#### ADDITIONAL POWER SAVINGS MODE INFO:

→From Navisphere, Disk State shows up as “Low Power” in Raid Group > Power Savings > Disks properties (Low Power)

→Also, all unbound SATA drives will also automatically go into Power Saving mode (Low Power)

→On the DAE itself with SATA drives, each Drive that is in Power Saving mode will flash the Green Power LED once every second

→Testing access to CIFS files on a file system hosted on SATA II drives in “Low Power” mode took less than 30 seconds to power-up drives and display files

#### CLARiiON NATIVE iSCSI/10Gbps SUPPORT:

→Flare 29 Arrays support 2-port 1Gbps iSCSI and/or 2-port 10Gbps iSCSI Ultraflex IO modules in various configurations

→Celerra has qualified the use of 10Gbps “Poseidon” iSCSI for NS-120, 480, & 960 arrays, as well as all Gateway platforms

→Poseidon replaces the 1Gbps iSCSI Harpoon module, and requires FLARE 29

→With the introduction of the 10Gbps iSCSI I/O module, Celerra supports CLARiiON Native iSCSI (not NX4), and also marks the beginning of the transition from file-based iSCSI Celerra support to CLARiiON native iSCSI lun support

**Note:** Integrated Arrays now support Host—attached iSCSI luns or FC luns

→Celerra Integrated supports ‘native’ CLARiiON iSCSI, and traditional Celerra MPFS over iSCSI

→Native Block License (NBOPT-L), for either iSCSI/FC, replaces FCOPT license

→Part of the benefit of ‘native’ CLARiiON iSCSI is that LUNs can be presented to either FC or iSCSI without data migration

NS-120 arrays can have up to (4) iSCSI modules, (2) of which can be 10Gbps

NS-480 arrays can have up to (6) iSCSI modules, (4) of which can be 10Gbps

NS-960 arrays can have up to (8) iSCSI modules, (4) of which can be 10Gbps

#### CLARiiON NATIVE vs. CELERRA iSCSI:

→CLARiiON native iSCSI support offers unlimited LUN size (>2TB)

→CLARiiON iSCSI can use CLARiiON & Navi features

→CLARiiON iSCSI luns support drive spin down feature

→CLARiiON iSCSI luns support PowerPath encryption

→Celerra file-based iSCSI luns support Netware & IP NAS mgmt

→Celerra file-based iSCSI luns are still limited to 1MB under 2TB

→Celerra iSCSI does not support Navi software features

→Celerra iSCSI does not support drive spin down feature of Flare29

→Native CLARiiON iSCSI for Celerra platforms require purchase of Native Block Option (right to use FC or iSCSI on the array) and Navisphere [NSxxx-NBOPTL]

→NS-960 platform, however, has Navisphere bundled with the NativeBlock License

**Note:** Celerra FCOPT license will no longer be offered

#### 8Gbps FC SLIC SUPPORT FOR NS-960/NS-G8 BLADES:

→NS-960/G8 Blades now support the ‘Glacier’ 8Gbps Fiber Channel I/O modules for connectivity between Blades and SPs, though backend buses & DAE’s still only operate at 4Gbps

**Note:** Celerra AUX arrays [NS-120/480/960] have supported 8Gbps FC I/O modules since NAS 5.6.42

→Sold only with new Installations at this time (no upgrade path yet)

→Auto-negotiates 2/4/8Gbps

→8Gbps SFP is different than 4Gbps

→Allowed in any SLIC slot where a FC I/O Module is allowed

→Backend array models offered for NS-960 will be NS960-AUX (4Gbps FC) and NS960-AUX8 (8Gbps FC)

### **EFD ENHANCEMENTS:**

→Flash Drives have been supported on CX4-480/960 since NAS 5.6.43

→In broad terms, about 10-30x faster than high-speed FC drives

→Best when used for random I/O workloads

→Adds Raid 1/0 support [min. 3 drives for HS]

**Note:** Clariion EFD luns are configured with Write & Read cache disabled, while Celerra best practices dictates that EFD Celerra LUNs be configured with Write cache enabled and Read cache disabled.

### **Restrictions:**

Celerra does not currently support having mixed FC & EFD drives in the same enclosure

### **Supported EFD Raid Types:**

#### **Raid1/0 1+1, Raid5 4+1, Raid5 8+1**

**Note:** Pools are clarefd\_r10 & clarefd\_r5, respectively

→Raid 1/0 1+1 & Raid5 4+1 will create four SP-balanced & equal sized LUNs per RG (setup\_clariion)

→Raid5 8+1 will create eight SP-balanced & equal sized LUNs per RG (setup\_clariion)

**Note:** Multiple Luns & threads are required to leverage EFD high performance

→Naming convention changed from “SSD” to “EFD”, hence AVM pool name clarssd\_r5 becomes clarefd\_r5, clarefd\_r10, etc., default 256K stripe size

→Cannot add EFD drives to empty slots in a DAE with FC drives

**Note:** In practice, setup\_clariion (and Provisioning wizard from Celerra Manager) fails when there are FC & EFD drives in the same enclosure. The NAS Support Matrix has a note that says though Clariion supports mixed FC & EFD drives in the same Enclosure, Celerra does not.

→Celerra FAST would be one example of a NAS solution using EFD, CLARiiON FAST another example, and next year, CLARiiON will introduce an extension to CLARiiON Write Cache using EFD drives

→Striping used whenever 2 or more luns available, otherwise concatenated

→EFD uses 16 internal channels, services 16 simultaneous I/Os vs. FC 1 I/O

→CLARiiON LUNs have queue depth of 12, therefore we need multiple LUNs per RG to have enough queued I/O to satisfy the EFD drives 16 channels

→At the same time, using multiple luns per RG yields multiple striped dVolumes on Celerra, which increases I/O queue depth

**Note:** Celerra has a normal queue depth of 8 I/O’s per dvol

### **Setup clarion EFD Templates:**

4+1 R5, HS, 1+1 R1/0, 1+1 R1/0, 4+1 R5

4+1 R5, HS, 1+1 R1/0, 1+1 R1/0, 1+1 R1/0, 1+1 R1/0, HS

HS, 1+1 R1/0, 1+1 R1/0, 1+1 R1/0, 1+1 R1/0, 1+1 R1/0, 1+1 R1/0, 1+1 R1/0

### **Example Output of EFD Drive:**

# /nas/sbin/navicli -h 128.221.252.200 getdisk 0\_0\_11

Bus 0 Enclosure 0 Disk 11

Vendor Id: STEC

Product Id: ZIV2A210 CLAR200

Product Revision: 100H

Lun: Unbound

Type: N/A

State: Unbound

Hot Spare: NO

Prct Rebuilt: Unbound

Prct Bound: Unbound

Serial Number: STM000100300

Sectors: 0 (0)

Capacity: 187847

Private: Unbound

Bind Signature: 0x0, 0, 11

Hard Read Errors: 0

Hard Write Errors: 0

Soft Read Errors: 0  
Soft Write Errors: 0  
Read Retries: N/A  
Write Retries: N/A  
Remapped Sectors: N/A  
Number of Reads: 446  
Number of Writes: 6  
Number of Luns: 0  
Raid Group ID: This disk does not belong to a RAIDGroup  
Clariion Part Number: DG118032673  
Request Service Time: N/A  
Read Requests: 446  
Write Requests: 6  
Kbytes Read: 349  
Kbytes Written: 137  
Stripe Boundary Crossing: None  
Drive Type: FC SSD  
Clariion TLA Part Number: 005048998  
User Capacity: 0  
Idle Ticks: 5288  
Busy Ticks: 188  
Current Speed: 4Gbps  
Maximum Speed: 4Gbps

### **2TB SATA DRIVE SUPPORT:**

→Support for 2TB 5400rpm drives for NS-960, NS-G8, NS-480, NS-120, NS-G2, & NX4  
→2TB drives can be used within the same enclosure with other SATA drives (not vault)  
→NS-960 would support 896TB usable capacity with 2TB drives

### **LARGE LUN SUPPORT:**

→Large LUN support for up to 16TB luns is extended to NS-120/480  
**Note:** Previously limited to Gateways & NS-960 platform (5.6.44)  
→Large LUN support affects setup\_clarion and CPW/UMPW utilities, which will now allow the creation of up to 16TB luns for NS-960, NS-480, & NS-120  
**Note:** Only the NS-960, using the R6 12+2 template would reach the 16TB LUN limit because it also enforces the single LUN per RG rule (useOneLUN)

### **MirrorView/S GA:**

→MView/S for Celerra has been supported via RPQ since 5.5.23  
→MView/S is now a “GA” offering  
→Number of luns in Consistency Group increases from 16 with Flare 28 to 32 with Flare 29 for NS-120  
→Number of luns in Consistency Group increases from 32 with Flare 28 to 64 with Flare 29 for NS-480 & NS-960  
→Overall, MView/S seen as a more viable Disaster Recovery solution than before

### **SMB 2.1 SUPPORT:**

→SMB 2.02 protocol supported for Vista and Windows 2008 R1 since NAS 5.6.42  
→Support now extended to SMB 2.1, used by Windows 7, Vista SP2, and Windows 2008 R2, SMB2 leases and Unbuffered writes  
→SMB2.1 enhancements are all related to increasing CIFS performance by reducing network traffic overhead  
→SMB2 leverages Client “leases” for better data cache and synching with Server vs. SMB1 Opportunistic Locks  
→With SMB 2.1 leases, Client lease can span multiple opens and multiple connections  
→New param cifs smb2.capabilities, set to 0x00000003 by default, supports SMB2\_CAP\_GLOBAL\_DFS & SMB2\_CAP\_GLOBAL\_LEASING, but not Global Large MTU feature (later release)

### **Three Types of SMB2 Leases:**

Read-Caching Lease—multiple clients (or same) can cache read the same file(s)  
Write-Caching Lease—Only single client can write cache, using byte-range lock  
Handle-Caching Lease—multiple clients (or same) can cache open file handles

### **Unbuffered writes—FILE FLAG WRITE THROUGH:**

→Clients can perform synchronous Write-through operations to Server (no server-side buffering), regardless of how file was opened  
→Clients can do synchronous & asynchronous writes for same File Handle, and keep data cache synchronized with server for longer periods

### **Large MTU:**

→SMB2 large MTU capability is not supported with this release

### **Enable SMB2 on Celerra via server cifs cmd:**

# server\_cifs server\_2 -add security=NT,dialect=SMB2 | SMB2only

**Before:**

# server\_cifs server\_2

Max protocol = NT1

**After:**

# server\_cifs server\_2

Max protocol = SMB2

**Use server\_cifs audit to determine SMB protocol version/Leases:**

# server\_cifs server\_2 -o audit fullgrep "ProtoFid"

```
| Proto=SMB2.10, Arch=Win2K8, RemBufsz=0xffff, LocBufsz=0xffff, popupMsg=1  
| Fid=87, FNN=0x1bafe850(FREE,0x00000000,0), FOF=0x00000000 Lease=RHW FILE=/foo.txt
```

**Note:** RHW=Read Handle Write caching lease

**MPFS SMB 2.02 SUPPORT:**

→ 5.6.47 qualifies SMB2.02 on Vista SP1 and Windows 2008 MPFS Clients

→ Clients will send 0x293 Server Capability during FMP Session create to indicate ability to do SMB2

**WINDOWS RESTRICTED GROUP GPO SUPPORT:**

→ Restricted Groups allow system administrators to define & control membership for security-sensitive groups

→ Restricted Groups support Local Users, Domain Users, and Domain Groups

→ Restricted Groups GPO updated every 90 minutes, whenever CIFS is started, or manually when using server\_security –update –policy gpo

→ Usually applied to the Local Groups on Workstations or member Servers (e.g., CIFS Server)

→ A Restricted Groups domain policy can define which Users or Groups are members of a Restricted Group, and also which groups the Restricted Groups can belong to

→ For DART, this becomes a new way to define and enforce Local Groups content

→ Restricted Groups 'Members' property page defines who is a direct member of the group

→ Restricted Groups 'Memberof' property page defines Restricted Group membership in other Groups

# server\_log server\_2 -s → Example of GPO update every 90 minutes

2009-11-20 04:09:46: LGDB: 6: Restricted Groups enforced for deduper

2009-11-20 05:39:50: LGDB: 6: Restricted Groups enforced for deduper

**MULTI-BYTE CHARACTER SUPPORT FOR CELERRA INTERFACES:**

→ Feature allows ability to input, store, and display multi-byte characters using Celerra management interfaces (e.g., CLI server\_cifs, fs\_dhsm, server\_cdms; MMC snapins; XML-API; & Celerra Manager), such as for local languages

→ While not true Internationalization or Localization, this is a step towards it

→ 5.6.47 extends multi-byte support from CIFS Sharename & Tree Quota fields introduced with 5.6.40, to many other fields [CIFS Compname, Netbios, or Alias names; CIFS Comments; Domain or Workgroup names; OU name; User name for User Quotas; Mountpoint & File System names; Checkpoint name; CDMS source share and export names & paths; DFS Root name]

→ Requires Unicode setting for the Data Mover

→ Follows Unicode 3.0 standard, default encoding HTML, XML (ISO/IEC 10646)

→ Browsers and Host OS's must support UTF-8 (variable length character encoding for Unicode, ranging from 1-4bytes, U+0000 to U+10FFFF)

→ UTF-8 can encode any valid Unicode character using multiple bytes (1-4 bytes)

**Issues—See emc229599:**

Known problem with ECC, which does not support multi-byte characters. Issue is currently understood to have the following potential impact to ECC when inputting Celerra Object names using non-ASCII characters, for any of several available interfaces [e.g., Celerra Manager, MMC Snap-in, CLI, etc]:

--With NAS 5.6.40, Celerra supports multi-byte characters in CIFS Share names. As such, if Share names use >than 63 non-ASCII characters, then ECC may fail to properly discover the Celerra.

--With NAS 5.6.47, Celerra extended multi-byte characters support to include File System, Mountpoint, NFS Export, and Quota User names (among several other Object name fields not used by ECC). As such, for any of the Object name fields listed in the next bullet, ECC may fail to discover the Celerra if the stated Fields use >than 63 non-ASCII characters. Additionally, if the stated Fields contain any non-ASCII characters, the non-ASCII characters themselves will be garbled in the XHMP summary collected by the ECC engine, but for any ASCII characters used within the same object name, these will be displayed correctly. The underlying data and names themselves on the Celerra system are not impacted. This is solely a Discovery/Display issue with ECC.

**List of Celerra Object Field Names that are used by ECC:**

[Fields which Celerra supports for non-ASCII characters, and are therefore susceptible to the Discovery/Display issue]

File System Names → Supported with 5.6.47

Mountpoint Names → Supported with 5.6.47

CIFS Share Names → Supported with 5.6.40

NFS Export Names → Supported with 5.6.47

Quota User Names → Supported with 5.6.47

→These names can cause ECC Celerra Discovery failures if using >than 63 Unicode characters in the object names

→In addition, Object names collected by ECC will be garbled in the ECC report for any names that use Unicode characters

#### **Suggested Workarounds:**

1) Use only ASCII characters in the above listed Object Field Names when using ECC for Celerra discovery.

2) If Unicode characters must be used, in order to successfully "discover" the Celerra, keep all applicable Object Field names under 64 characters, but also be aware that if any on-ASCII values are used in the above named fields, the non-ASCII characters will appear garbled in the ECC summary.

#### **OUT OF FAMILY REPLICATION V1 SUPPORT (5.5 to 5.6):**

→Goal of this feature is to provide customers a more manageable method for updating replication infrastructure by supporting the upgrade of target systems to 5.6.47 while source systems are running a minimum revision of 5.5.39

**Note:** check\_nas\_upgrade checks were put into place that requires any 5.5 V1 environment to have the Source side at 5.5.39 before upgrading the environment to NAS 5.6. See AR165855.

→Upgrade methodology would be to leave the 5.5 source site replication sessions running while upgrading target sites to 5.6 (Do not suspend sessions)

→Important to know that “management” of the replication environment is limited while running dissimilar code families on source vs. targets [i.e., cannot start new sessions; cannot restart suspended or inactive sessions; cannot reverse or suspend sessions; cannot do fs\_copy; cannot resync failed over sessions]

→Customers can failover from the 5.5 source to 5.6 target, but cannot fallback until source is upgraded to 5.6

→File System copies, restarts, or creating new sessions will fail when the sites are at dissimilar code versions

→Attempted upgrades to 5.6.47 on source with destinations running 5.5 will PUHC fail; PUHC fail if fs\_copy is running; PUHC fail if source is less than 5.5.39

#### **Upgrading Bi-Directional Replication-Example:**

→Situation where each side of replication is both a Source and Target

Source1 (Target2) ↔ Source2 (Target1)

→Suspend replication sessions from Source2 to Source1 to break the bi-directional dependency

→Upgrade Source2 to 5.6.47, then Source1 to 5.6.47, then restart replication sessions from Source2 to Source1

#### **OUT OF FAMILY REPLICATION V2 SUPPORT:**

→Goal of this feature is to be able to support out-of-family V2 replication for any version 5.6.47 or higher, with NAS 6.0

→ReplicationV2 will be compatible with full management capability across the CLI & GUI

→The only caveat here is that both Source & Target sides must be running at least 5.6.47 or later, meaning that the Source could be running 6.0 while the Target runs 5.6.47, or vice versa—the order would not matter

→All the commands and functionality would be available

→One caveat is that there is no guarantee that some new feature or functionality might not affect this strategy, and the exceptions will be handled as they occur

→One example of interoperability issues with GUI's is that a system running 5.6 would have to use the Celerra Manager interface, while a system running 6.0.xx would use the new Unisphere GUI framework (GUI that will be used to manage both CLARiiON and Celerra)

#### **CSA ENHANCEMENTS:**

→Prior to 5.6.47, CSA could only configure Server\_2—now, can configure all Primary Servers [i.e., DNS, NTP, Unicode, Timezones]

→Prior to 5.6.47, CSA(CPW) could not configure LACP or Ethernet Channel on some newer network device names [e.g., fxg0, cge-3-0, etc.]—it only recognized devices named cge0 – cge3. Now, CSA wizards will be able to display & use all available network devices on the Blades

→From “Attach an iSCSI LUN” GUI screen, the “Local Host IP Address” will now display all bound IP addresses

→Warning message will be displayed if trying to mount an iSCSI LUN on the local Host when it has already been assigned to another initiator

→For Dell NX4 systems only, a unique IQN will be created for each new target based on IQN string plus unique target name assigned

#### **Multiple Connections per iSCSI session:**

--Ability to create up to (10) internet IP address connections per iSCSI session [will use Local Host IP addresses instead of 0.0.0.0 for target login]—user will be able to select which IP address to use for a particular session connection login

--If no existing connection between Host and iSCSI Target, will create iSCSI session and the connection between source portal and Target (iscsicli logintarget)

--If connection between Host & Target already exists, will allow adding new connection (iscsicli addconnection)

#### **CPW/CSA SUPPORT FOR MPFS:**

→With 5.6.47, the CSA/CPW will present a wizard to configure MPFS for the Celerra and for Linux clients using Best Practice guidelines for FC & iSCSI, the goal being ready for MPFS I/O at the end of configuration

→MPFS Setup wizard can also be used to ‘reconfigure’ an existing MPFS setup

→Builds FC environment or iSCSI environment as needed on Celerra & Hosts

→MCM tool being removed from Powerlink (CSA capability replaces)

→'Setup MPFS' wizard discovery will generate an error if at least 4-8 equal sized luns are not available for Raid5

(clar\_r5\_performance), Raid6 (clar\_r6), or Raid 1/0 (clar\_r10)

→Users should go back to the "Select a Wizard" page and select "Provision unused disks", select Custom mode, and as an example for Raid 1/0, select 4 disks for Raid 1/0 to create a Raid Group with 4 luns, and then repeat again to create another RG with 4 luns

**The 'Welcome to the CSA' screen will only present the 'Setup MPFS' button under the following rules:**

→If model is NS40-AUX-C, 'Setup MPFS' will display for MPFS over iSCSI

→If model is NS20/40FC, 'Setup MPFS' will display for MPFS over FC

→If model is NS-120, 480, or 960, the /nas/sbin/model -option -list command is run to check whether any MPFS flags are enabled

--If only MPFS iSCSI flag is set, 'Setup MPFS' will display for MPFS over iSCSI

--If only MPFS FC flag is set, 'Setup MPFS' will display for MPFS over FC

--If both MPFS iSCSI & FC flags are set, 'Setup MPFS' will display for mixed environment for MPFS over iSCSI and MPFS over FC

**MPFS Setup Chronology:**

CSA > 'Select a wizard' > 'Setup MPFS' (Presented only if Wizard detects an MPFS-enabled Celerra /etc/be\_sg\_info)

--MPFS discovery occurs for MPFS over iSCSI, MPFS over FC, or both, checks to make sure a Storagepool is available with LUNs, and also offers the option of setting up just the Celerra, or both Celerra & MPFS Hosts

--If MPFS already configured, User presented option to reconfigure

--If applicable, User is prompted to setup/verify IP's on iSCSI ports on SPs

--Setup CHAP

--Create>Select MPFS StoragePool and number of members to stripe at 256KB

--Wizard configures backend [Setup iSCSI ports, Create MPFS StorageGroup, Create/Extend MPFS StoragePool, Add LUNs to MPFS Storage Group, Start MPFS Service on Data Mover]

--Add Hosts screen [enter Username, Password, & IP address], with option to use an Import Host List

--Install MPFS Client Software [Linux version 5.0.31.7]

--Set iSCSI parameters using Best Practices or Custom options for Host

--Set sysctl parameters

--Create/Mount MPFS enabled NFS Exports [If creating export, select the mpfs\_r10\_pool, name the export, and give a size to file system]

--Add export to Host

--Host Configuration Summary displays

--Host Configuration Status

--Verify MPFS exports and luns on the Host using # mpfsinq

**OTHER CELERRA MPFS ENHANCEMENTS:**

→SMB 2.0.2 support for Vista SP1 & Windows 2008 MPFS clients

→Support for CentOS 5.3, AIX 6.1, and Windows 2008 (but not R2)

→Windows 2008 (x64 &x86) support for MPFS over FC/iSCSI

→RHEL 4.8 & 5.3 (32 & 64-bit) support for MPFS over FC/iSCSI

→Support for 10Gbps iSCSI I/O modules on CLARiiON SPs

→Increased Host support on FC or iSCSI Ports for Arrays running Flare 29

| CX4 Models                        | 120 | 240  | 480  | 960  |
|-----------------------------------|-----|------|------|------|
| Max Initiators FC ports           | 256 | 256  | 256  | 512  |
| Max Initiators 1Gbps iSCSI ports  | 256 | 256  | 256  | 256  |
| Max Initiators 10Gbps iSCSI ports | 256 | 512  | 1024 | 1024 |
| Max Initiators/SP                 | 256 | 512  | 1024 | 4096 |
| Max Initiators/Array              | 512 | 1024 | 2048 | 8096 |

**WILDCAT-S VE HARPERTOWN CPU MODULE:**

→Introduction of a new motherboard for NS-960/G8, Harpertown Wildcat-S VE (original motherboard Clovertown Wildcat-S)

→Actual shipments expected January 2010 using Coldfire management firmware version v2.99. Earlier mgmt firmware versions do not support Harpertown.

**IOMEGA CIFS COPY SUPPORT TO CELERRA SHARES:**

→CIFS copy support from Iomega ix4-200d StorCenter device to Celerra

**AVM CHANGE:**

→AVM auto-extend default change from 10GB to 20GB, meaning that file system <than 20GB will auto-extend by the original size of the file system, and one >than 20GB will auto-extend by 5% of its size, or by 20GB, whichever is larger

**COGNAC 5.6 MAINTENANCE RELEASE 11 (CMR11 5.6.48.7 Patch 705):**

→Tolerance for Symmetrix Madeira 5874.207 release (Symm released in Jan 2010) for all Gateways

→Tolerance for Symmetrix FAST v1 solution for non-Celerra hosts

→Dell OEM NS-120/480/960 support

- Support for NS-480 and NS-960 Dense Storage products
- Porting the CCC (Celerra Configuration Collector) to the Control Station—enhances Collect Support Materials
- Celerra Event Enabler (CEE) version 4.5.2.2
- Support for RHEL 5.0 update 4 for MPFS Linux clients
- DART support VMWare plugin (see below)

**Note:** In order to “Add a Celerra” using the plug-in, it will check to make sure that the correct NAS version is on the Celerra

## **CELERRA CONFIGURATION COLLECTOR:**

→Basically, the CCC is a java collection tool to gather configuration details from the Celerra system, and is outputted to a raw XML file that is only used by the NSD [Network Solution Designer] application (at this time)

### **When does the CCC information get collected?**

- 1) Automatically whenever the automatic log collection script runs

**Note:** Default behavior, though you could comment out the “GetCCC” option in the /nas/site/automaticcollection.cfg file

- 2) Manually if invoked using /nas/tools/automaticcollection –getlogs or /nas/tools/collect\_support\_materials -CCC

- 3) Or manually if invoked from Celerra Manager > Support > Log Collection > Collect

### **Scripts called by the Log Collection utilities:**

/nas/tools/automaticcollection –getlogs →Automatically or manually started, calls collect\_support\_materials

/nas/tools/collect\_support\_materials -CCC or -CCC-quota →Main collection script with CCC options

/nas/tools/collect\_config →Called by log collection script, in turn calls java –jar /nas/tools/ccc.jar \$1\$2

**Note:** -symmdisks is automatically added as a variable whenever symm backend is seen. –quota option is added when specifying –CCC-quota flag manually

/nas/tools/monitor\_collect\_script →script started by collect\_support\_materials to monitor processes once every 60 seconds, and kills any processes that last 15 minutes

### **Configuration Files:**

/nas/site/automaticcollection.cfg →GetCCC option enabled by default

/nas/tools/CCC.cfg →Change “Timeout 300” to a different value

### **Logs:**

# ls -la /nas/log/ccc.log\*

```
-rwxrwxrwx 1 nasadmin nasadmin 621251 Mar 2 13:49 ccc.log
-rwxrwxrwx 1 nasadmin nasadmin 143979 Mar 1 17:38 ccc.log.1.gz
-rwxrwxr-x 1 nasadmin nasadmin 188727 Jan 11 13:05 ccc.log.2.gz
```

### **Log Collection Files:**

# ls -la /celerra/ccc

```
-rw-r--r-- 1 root root 1033802 Mar 1 16:03 CCC_FNM00083800203_10_03_01_16-03-08.xml
-rw-r--r-- 1 root root 1033802 Mar 2 13:05 CCC_FNM00083800203_10_03_02_13-05-21.xml
```

# ls -la /nas/var/emcsupport

```
-rw-rw-r-- 1 nasadmin nasadmin 6996440 Mar 1 16:08 support_materials_FNM00083800203.100301_1603.zip
-rw-rw-r-- 1 nasadmin nasadmin 6998172 Mar 2 13:09 support_materials_FNM00083800203.100302_1305.zip
```

**Note:** The CCC\_<serial>\_<date&time>.xml file is located in the “BG\_Output” directory in the zip collection

# zipinfo -l support\_materials\_FNM00083800203.100302\_1305.zip |grep BG

```
-rw-r--r-- 2.3 unx 1033802 tx 59194 defX 2-Mar-10 13:09
```

fox1.100302\_1305/BG\_Output/CCC\_FNM00083800203\_10\_03\_02\_13-05-21.xml

### **Example of XML content:**

# more CCC\_FNM00083800203\_10\_01\_22\_14-07-29.xml

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<CCC_Document xmlns:CELERRA="http://www.emc.com/celerra">
  <CELERRA:Version>1.4.0</CELERRA:Version>
  <CELERRA:Celerra>
    <CELERRA:Name>fox1</CELERRA:Name>
    <CELERRA:Serial>FNM00083800203</CELERRA:Serial>
    <CELERRA:Model>NS-962FC</CELERRA:Model>
    <CELERRA:Version>5.6.48-5</CELERRA:Version>
    <CELERRA:DMSlots>2</CELERRA:DMSlots>
    <CELERRA:Software_Licenses>
      <CELERRA:Advanced>online</CELERRA:Advanced>
```

-----abridged-----

## **EMC Celerra Plug-in for VMware Version 1.0.7 ('Nimitz' release): Requires 5.6.48.701 patch!**

### **FEATURES:**

→NFS Datastore Provisioning

→VM Full Clone

→VM Fast Clone

→VM Compression

**Overall Plug-in Purpose:**

NFS Datastore Provisioning; Array-based compression & cloning of VMs

**Note:** Celerra creates ufs file system from array and presents to VMware ESX servers via NFS Export, and vSphere plug-in creates the NFS Datastore presentation, from which the VMDK Snaps (Fast Clone) or VMDK Copies (Full Clone) are made for VMs

**VMware PLUG-IN CAPABILITIES:**

--View Celerra properties for VMs, NFS Datastores, etc via the Plugin

**1) Provision or Extend Storage for NFS Datastores:**

--Create & mount NFS Datastores on ESX Servers via Celerra file systems (or use existing)

--Filesystems created with Virtual Provisioning and Auto-extend enabled

--NFS Exports use best practice mount options

--Extends existing Celerra NFS Datastores storage (aka file systems, max. size of 16TB Celerra limitation)

**2) Optimize Storage using Compression (aka Celerra Dedupe):**

--Ability to compress or uncompress VMs that reside on Celerra NFS Datastores

**3) Create VM Clones as Full Copies or Fast Snap Copies:**

--Ability to “clone” VMs that reside in Celerra NFS Datastores

--Fast clone support – VM .vmdk files are “snap” copies, limited to the same Datastore file system where the VM is stored

--Full clone support – VM .vmdk files are full copies, limited to Datastore file systems on the same physical Data Mover

**Plug-in Specifics:**

→The plug-in is a Windows executable that installs on the Virtual Center vCenter 4.0 or VMWare vSphere 4.0 Client system

→Properties are available by rightclicking objects in vSphere [VMs; Hosts; Clusters; Datacenters; Folders; Datastores]

→ESX data path plugin, vStorage API integration for NFS—better integration between storage platforms and ESX server

**Note:** vCenter plug-in will allow an admin to make a copy of a VM as a full clone or fast (thin) clone

→Qualified with ESX 3.5 and 4.0 Hosts

→Plug-in is designed to simplify management of virtual infrastructure on NFS Data Stores

→Plug-in provides storage optimization by offering compression of VM’s & associated VMDK files (via NAS Deduplication code)

a) Reads only uncompress the portion of the file requested

b) Writes go to a set-aside file for background compression

→Replication Manager snap management & recovery [Celerra Snapshot checkpoints of NFS datastore; SRM integration]

a) Snap management initiated via RM v5.2

b) VMSnap executes on all VMs within the datastore (application consistent copies)

c) Celerra Snapshot creates copies of VMs in the NFS Datastore

d) Restores of individual VMs or entire Datastore

e) Can perform offline backups

→VMware Site Recovery Manager (SRM 4.0) Integration

a) Celerra Replicator management by SRM with NFS Storage Replication Adapter (Replication between Celerras)

b) Celerra supports SRM failover and fallback (with NFS Failback plug-in)

**MANAGING CELERRAS FROM PLUG-IN:**

[View > Home > Solutions and Applications: EMC Celerra](#)

**PROVISION OR EXTEND STORAGE:**

→Provision Storage wizard allows creation of the NFS Datastore on a particular Control Station and Data Mover, using either a new NFS Export (new file system with Virtual Provisioning enabled by default) or an existing NFS export

→Default Celerra best practice settings for the NFS Export is to mount with “uncached” & “noscan” options

→Add the Celerra to the VM Infrastructure using the “Provision Storage” dialog, or if a VM operation for cloning or compressing is done on a VM to which the Celerra has not yet been added

→Can extend the file system using the “Extend Storage” wizard, which extends the presented NFS Datastore

**Note:** Max size is 16TB per Celerra file system limit

**VM FULL CLONE:**

→Rightclick VM > EMC Celerra > Full Clone or Fast Clone

→Creates full & separate copy of a VM (vmx & vmdk files) in the NFS Datastore using Celerra DHSM full clone api

→Limited to the NFS Datastore on the same Data Mover

→Maximum clone count is based on number of CPU cores in Datacenter Cluster or ESX Host

**VM FAST CLONE:**

→Creates file level snap of a VM’s vmdk file and a full clone of a VM’s vmx file

→Limited to the NFS Datastore of the parent VM

→Maximum clone count based on the number of CPU cores in the Cluster or ESX Host

→Initial fast clone snap size is only 32kb, but expands after the VM’s are powered on, and depends on guest O/S size

→Fast clone properties will list the parent

→Fast Clones can only be made on VMs on the same file system

→VM’s with a Fast Clone cannot be deleted (leaves .vmdk file)

→Fast Clones cannot be compressed

#### **STORAGE OPTIMIZATION: COMPRESS/DECOMPRESS**

→Compression (Celerra Deduplicatiion) is not enabled by default when creating the NFS Datastore, but can be configured via Advanced Options during creation, or via Properties dialog box after creation of the Datastore

**Note:** Setting compression automatically enables deduplication on the underlying Celerra file system, but shows up as “suspended”. The Deduplication API is called whenever plug-in actions are taken.

→Compress or decompress VMs, approximately 30-50% (rightclick VM > EMC Celerra > Compress/Decompress)

→Compression properties will show compression characteristics

→Compression not allowed on Branch File (Fast clone) or Version File [a base file (aka VM) with fast clones]

→Plug-in does not support NMFS

→Plug-in does not support Compression or Fast Clones when using VMware Fault Tolerance feature

#### **PREREQUISITES FOR USING CELERRA VMWARE PLUGIN:**

--NAS 5.6.48.7 + 701 patch

--ESX Server 3.5 or 4.0

--vCenter Server 4.0

--vSphere Client 4.0

--NFS license enabled on Celerra

--Network connectivity between Data Mover, Control Station, ESX server, vSphere client

--DHSM enabled and configured on Celerra for CS & vSphere Client access (cloning & compression)

--EMC Celerra Plug-in for VMware installed on vSphere client

--Available Storage Pools on Celerra for automatic creation of NFS file system and export for NFS Datastore

#### **Current Limitations:**

→Deduplication does not work for version files, which are fast clones

→No way to monitor progress of deduplication for full clones, or failure message from client

#### **USING THE CELERRA VMWARE PLUGIN:**

##### **I. Install the EMC Celerra Plug-in for VMware on the vSphere Client Server**

##### **II. Setup the DHSM User and Service on the Data Mover:**

**Note:** The DHSM setup will be required if using the Plug-in to compress or uncompress VMs in the NFS Datastore, or cloning or viewing properties.

###### **a) Create DHSM User on the Data Mover(s)**

```
# /nas/sbin/server_user ALL -add -md5 -passwd dhsm_user
```

Creating new user dhsm\_user

User ID: 525

Group ID: 525

Home directory:

Changing password for user dhsm\_user

New passwd:

Retype new passwd:

###### **b) Enable digest Authentication**

```
# /nas/bin/server_http ALL -modify dhsm -authentication digest -users dhsm_user
```

server\_2 :

#### DHSM FACILITY CONFIGURATION

Service name : EMC File Mover service

Comment : Service facility for getting DHSM attributes

Active : False

Port : 5080

Threads : 16

Max requests : 300

Timeout : 60 seconds

#### ACCESS CONTROL

Allowed IPs : any

Authentication : digest ,Realm : DHSM\_Authorization

Allowed user : dhsm\_user

#### SSL CONFIGURATION

Mode : OFF

Persona : default

Protocol : default

Cipher : default

###### **c) Start DHSM Service on Data Movers**

```
# /nas/bin/server_http ALL -service dhsm -start
```

server\_2 : done

#### DHSM FACILITY CONFIGURATION

Service name : EMC File Mover service  
Comment : Service facility for getting DHSM attributes  
Active : True

#### d) Add vSphere Client IP & Control Station IP to DHSM access list

# /nas/bin/server\_http ALL -append dhsm -hosts 10.250.148.134

-----abridged-----

#### ACCESS CONTROL

Allowed IPs : 10.250.148.134,10.250.148.184

**Note:** Add the IP address of the Control Station in order to view get\_attributes, or to manually test compressing, decompressing, cloning, etc., but using the tools found in the /nas/tools/dhsm directory

#### III. Add the Celerra Control Station to the vSphere Plug-in:

[View > Home > Solutions and Applications: EMC Celerra > Add](#)

Add Celerra – EMC Celerra Plug-in

Celerra Control Station Hostname / IP: 10.250.148.184

Celerra Control Station Username: nasadmin

Celerra Control Station Password: passwd

DHSM Username: dhsm\_user

DHSM Password: passwd

**Note:** Observe the Recent Tasks section at the bottom of the vSphere screen to monitor success or failure of the Add operation

#### IV. Create a new NFS Datastore:

[Datacenter > rightclick, EMC Celerra > Provision Storage](#)

**Note:** Enter Datastore name, Control Station, Data Mover, Data Mover interface, Storage Pool, then enter details for new or existing NFS export, as well as any advanced properties (e.g., Virtual Provisioning; Enable Compression (dedupe), etc)

→ NFS Datastore created at Datacenter level is visible to all ESX Server Hosts, while NFS Datastore created at Host level is visible to that Host only

→ When creating from existing NFS Export, best practice popups recommends Virtual Provisioning, and mount options nocache, and uncached

→ Plugin should provide popup error if there was a problem during the NAS Storage Provisioning operation, though a provisioning error will not stop the Celerra from creating and exporting the NFS file system (so you may need to cleanup exports and file systems once the datastore issue has been figured out)

#### EXAMPLE OF SUCCESSFUL NAS NFS DATASTORE CREATION USING PLUGIN:

# server\_mount server\_2 |grep uncached

fs1 on /fs1 udfs,perm,rw,noscan,uncached

Dstore1\_184 on /Dstore1\_184 udfs,perm,rw,noscan,uncached

datastore2\_184 on /datastore2\_184 udfs,perm,rw,**noscan,uncached**

# server\_export server\_2

server\_2 :

export "/dstore\_225" root=192.1.8.229 access=192.1.8.229 → VMKernel IP address on the ESX Host Server

#### EMC NFS Plugin.log:

Issuing command to Celerra:

https://10.241.168.30/servlets/wizards?object=filesystem&command=new&properties=name=Dstore1\$storageMB=10240\$type=uxfs\$rwMover=server\_2\$storagePool=36\$thinProvisioningEnabled=True\$flrState=off\$autoExtendEnabled=True\$autoExtendMaxMB=20480\$rdeState=off\$autoExtendHWM=90

DEBUG 2010-03-03 11:22:16 [com.emc.celerra.tools.engine.apl.HttpsConnection]: Issuing command to Celerra:

https://10.241.168.30/servlets/wizards?object=export&command=new&properties=path=/Dstore1\$mover=server\_2\$rootHosts=192.1.8.227\$accessHosts=192.1.8.227

DEBUG 2010-03-03 11:22:17 [com.emc.celerra.vmware.nfs.engine.CloneController]: mount NFS Datastore Dstore1 on Host 192.1.8.225 by using NFS Export /Dstore1 and Datamover interface 192.1.8.241

#### vSphere Client Plugin:

View > Inventory > Datastores > Dstore1: rightclick and “Browse Datastore...” or go to “EMC Celerra > Properties” to see properties of the Datastore [Shows VMs in the Datastore; Compression is not enabled by default]

#### vSphere Client plug-in GUI:

#### Plug-in Properties for Datacenter, Cluster, and ESX Host objects:

EMC Celerra > Compress; Decompress; Provision Storage; Properties; Help

#### Plug-in Properties for VMs:

EMC Celerra > Compress; Decompress; Full Clone; Fast Clone; Properties; Help

→ For “Provision Storage”, you create an NFS Data Store based on a new or existing NFS Export, after logging into the Celerra with username and password, and providing a dhsm username and password

→Rightclick VM > EMC Celerra > Properties | Full Clone | Fast Clone | Compress | Decompress

→Once created, you can rightclick the NFS Datastore, select EMC Celerra>then go to “Compress; Decompress; Extend Storage; or Properties”

**Note:** You must setup a DHSM user account and service on the Data Movers that are hosting the VMware NFS Datastores, and then add the vSphere Host to the DHSM access list. This is necessary because VM properties, cloning, and compression cannot be done without the Celerra’s DHSM api.

#### Troubleshooting Plug-in:

→Plug-in activities are logged on the Windows Client

**Program Files>VMWare>Infrastructure>Virtual Infrastructure Client>Plugins>EMC Celerra>log>EMC\_NFS\_Plugin.log**

#### USING DHSM TOOLS FROM CONTROL STATION TO TEST PLUGIN:

##### Viewing attributes of objects in the NFS Datastore:

```
# /nas/tools/dhsm/get_attributes -u dhsm_user -p nasadmin -d 192.1.8.241 nfs:/dstore4/2k3_cloned/2k3_cloned.vmdk
<DHSM_GET_ATTRS PATH="nfs:/dstore4/2k3_cloned/2k3_cloned.vmdk" DEDUPE="True"/>
```

##### Compressing Single File in Datastore:

```
# /nas/tools/dhsm/dedupe_file -u dhsm_user -p nasadmin 192.1.8.241 nfs:/dstore4/Win2k3157/Win2k3157.vmdk
<DHSM_DEDUPE_FILE PATH="nfs:/dstore4/Win2k3157/Win2k3157.vmdk"
```

```
# /nas/tools/dhsm/get_attributes -u dhsm_user -p nasadmin -d 192.1.8.241 nfs:/dstore4/Win2k3157/Win2k3157.vmdk
<OFFLINE_ATTRS
```

OFFLINE\_PATH="dart://rde/AQAAAAA/1e1/c2/AQm8fBIIINUL69uSy4DMQHkYwfhQUAAAAAAAAAAAAAAwAAAAAA  
AAACAw;338b91b1" →This line indicates that the file is compressed

##### Uncompressing Single File in Datastore:

```
# /nas/tools/dhsm/redupe_file -u dhsm_user -p nasadmin 192.1.8.241 nfs:/dstore4/Win2k3157/Win2k3157.vmdk
<DHSM_REDUPFILE PATH="nfs:/dstore4/Win2k3157/Win2k3157.vmdk"
```

##### Creating a Full Clone of VMDK File:

```
# /nas/tools/dhsm/clone_file -f -u dhsm_user -p nasadmin 192.1.8.241 nfs:/dstore4/Win2k3157/Win2k3157.vmdk
nfs:/dstore4/Win2k3157/w2k3_cl.vmdk
```

<DHSM\_CLONE\_FILE PATH="nfs:/dstore4/Win2k3157/Win2k3157.vmdk"

CLONE\_PATH="nfs:/dstore4/Win2k3157/w2k3\_cl.vmdk" CLONE\_TYPE="Full"

##### Creating a Fast Clone of VMDK File:

→Difference from previous command is to drop the “-f” from the commandline syntax

#### EMC CELERRA PLUGIN FOR VMWARE Version 1.1.9 (Eisenhower release): June 2010

→Enhancement to the initial Version 1.0.7 Celerra VMware plugin release

#### PRACTICAL INFORMATION ON THE PLUG-IN:

--The initial plug-in release for Celerra was version 1.0.7, March 2010

--A new and enhanced plug-in was GA’ed in June 2010, version 1.1.9

--The plug-in setup.exe is downloadable from Powerlink

--The plug-in installs on Windows 2003 running VMware vSphere 4.0

--The vSphere communicates to the Celerra using HTTPS

--Celerra Screenshot is one of several DART features that are leveraged in the creation of fast clones for VMs

--When you create the DHSM connections, make sure that the ESX Host VMkernel IPs, and Control Station IP, are added to the DHSM connection (dhsmsetup.pl script does not necessarily advise you to do this otherwise)

--The Celerra DHSM API is used for Cloning, Compression, & Viewing of properties within the vSphere 4.0 Client

--VMs are seen as a ‘Regular file’ from within the vSphere GUI properties

#### Datastores:

--Supports only NFS Datastores, not FC or iSCSI block Datastores

--Datastores are built on Celerra file systems and are intended to be dedicated for ESX Server use only, using the vSphere plug-in functionality, and not for general use by other NFS or CIFS Hosts

--Datastores are created by default with Celerra Virtual Provisioning and file system auto-extend enabled

--Datastores are used as a container for Virtual Machines, including Full or Fast clones of Virtual Machines

#### Compression:

--Optimizes storage through the use of Celerra deduplication for compression of VMs

--Overall performance of VMs is said to be within 10% of uncompressed VMs

--However, compression is never automatically applied and must be manually initiated by the vSphere Client (using the plug-in capability), and applies only to VMs within datastores

--You can initiate ‘compression’ from various object levels (Datacenter, Datastore, ESX Host, individual VM, etc.)

--Be aware that having ‘Compression Enabled’ on a datastore does not automatically compress VMs or Clones that get added to (or created in) the datastore—the User must manually initiate an operation to compress any new VMs or Clones added to a datastore

--Compression attributes and space savings on VMs can be seen via ‘Properties’ from the plug-in

--Compression information is not visible from the Celerra CLI

#### Compression Error Examples—Product Functioning as designed:

**Action:**

Try to compress a Virtual Machine that already has Fast Clones created underneath (aka, Version file, or Base VM with Fast Clones)

**Result:** The action taken is not allowed, no popup, just results in a Task Error

“Compress vm184-2 Error Unable to compress a version file(Base VM with Fast Clones). This operation is not allowed.”

**Action:**

Try to compress a Fast Clone

**Result:** The action taken is not allowed, no popup, just results in a Task Error

“Compress vm184-2clone1 Error Unable to compress a branch file(Fast Clone). This operation is not allowed.”

**VM Properties using Plug-in:**

--A Virtual Machine is seen as a ‘Regular file’ from within the context of the vSphere plug-in properties

**VM path is:**

**/dstore184-3/w2k3-184**

**Full Clones:**

--Full clones are complete copies of a VM, and can reside on the original datastore file system, or other datastore file systems, provided it is on the same Celerra Data Mover

--Full clones are seen as Regular files in vSphere properties

--Full clones hang off the root of the datastore file system to which they are copied, and under a full\_clones subdirectory

**Full Clone of ‘w2k3-184’ in datastore ‘dstore184-3’:**

**# ls -la dstore184-3/full\_clones**

drwxr-xr-x 2 root root 1024 Jun 30 13:27 clone00001

**Note:** In this example, clone00001 is ‘w2k3-184-clone1’ in vSphere

**Properties of Full Clone using Plug-in:**

--Full clone of VM ‘w2k3-184’ is a complete & independent copy

--Like VMs, Full clones are also ‘Regular files’

--You can create multiple clones using ‘Clone Count’ field, and assign your own naming convention with ‘Clone Name’ field, Power On automatically, Integrate with VMware View

**Full clone path:**

**/dstore184-3/full\_clones/clone0001**

**Fast Clones (aka Branch file):**

--Fast clones are snapped copies of a VM, and can only reside on the original file system datastore where the parent VM resides. In otherwords, the Fast Clone is a child of the parent VM and lives in a sub-directory on the parent file system.

--A VM that has a child Fast Clone is called a ‘Base VM’, or ‘Version file’

**Fast Clone from parent VM ‘vm184-2’ in datastore ‘dstore184-3’:**

**# ls dstore184-3/vm184-2 |grep clone\***

clone00001

clone00002

**Properties of Fast Clone using Plug-in:**

--Fast clone of ‘vm184-2’

--Fast clones are called Branch files, or thin clones of a Base file (the parent VM)

--You can create multiple clones using ‘Clone Count’ field, and assign your own naming convention with ‘Clone Name’ field, Power On automatically, Integrate with VMware View

**Fast Clone Path:**

**/dstore184-3/vm184-2/clone0001**

**CREATING A VIRTUAL MACHINE USING vSPHERE:**

--The following steps provide an overview of how you might create a VM on a newly created Celerra NFS datastore, and is strictly a function of the vSphere Client and not the Celerra plug-in

**Steps for Creating a Virtual Machine on a Celerra datastore:**

1. Create an NFS datastore using the plug-in
2. Create the VM by rightclicking the Datacenter object in vSphere
  - a.) Highlight Datacenter name > New Virtual Machine > Create New Virtual Machine
  - b.) Choose a name for the VM, Datacenter/ESX Host, and datastore to host the VM
  - c.) Select a desired Operating System from the list (e.g., Windows 2003)
  - d.) Proceed through the ‘Create a Disk’ and ‘Ready to Complete’ screens and the VM gets created
3. After creating the VM, you must then actually ‘install’ the O/S
  - a.) Rightclick the new VM > Edit Settings
  - b.) Upload a Windows 2003 ISO image file to a datastore location using the “Browse Datastore” and ‘Upload’ functionality within vSphere
    - c.) Select “CD/DVD Drive 1” from hardware list, then select “Datastore ISO File” and browse to the previously uploaded ISO file, and finally, check the box at the top of the screen for “Connect at power on”
    - d.) Highlight the VM name, and click on the “Console” tab located on the righthand side of the vSphere screen

e.) Click the green triangle button at the top of the screen to power on the VM, and install the Windows O/S within the context of the ‘Console’ screen display

**PLUG-IN PRE-REQUISITES:**

- Minimum NAS 5.6.48.7 + 701 patch
- ESX Server 3.5 or 4.0
- vCenter Server 4.0
- vSphere Client 4.0
- VMware View Manager 4.0
- NFS license enabled on Celerra
- Network connectivity between Data Mover, Control Station, ESX server, vSphere client, VMware View client
- DHSM enabled and configured on Celerra for Control Station & vSphere Client access
- EMC Celerra Plug-in for VMware installed on vSphere Client
- Available Storage Pools on Celerra for the automatic creation of NFS file systems and exports (for NFS Datastores)

**PLUG-IN FEATURES FOR 1.1.9:**

**VMware View 4.0 integration:**

- VMware View is integrated to VSphere using an ‘Add’ dialogue option in the plug-in
- ‘Refresh Desktops’ applies to Fast Clones only—refreshing a Fast Clone Desktop deletes the existing Fast Clone, creates a new Fast Clone with the same name from the ‘base’ VM, then places new Clone back in the VMware View Pool as a ‘refreshed’ Desktop
- Full and Fast clone integration into VMware View
- Allows for automatic registration of Fast clones in VMware View

**No Scan Mount Option for disabling CAVA scanning:**

- By default, Celerra file systems will mount with the “noscan” option so as to disable CAVA virus scanning on the NFS datastores

**Script Setup of DHSM on Celerra:**

- DHSM script sets up the DHSM user, service, & connection to the Celerra and the vSphere VMkernel [previously, all manual steps]
  - a) Upload dhsmsetup.pl perl script to any location on the Control Station, using a program such as WinSCP3
  - b) As Root user, execute the script & answer the prompts for UID, GID, etc:  
**# perl dhsmsetup.pl <Data Mover name> <dhsm\_user> <vSphere\_host IP>**
  - c) Append the CS IP address to DHSM connection if you want to use the Control Station DHSM troubleshooting utilities

**Users can view Space Savings after running compression:**

- Before and after space savings are now displayed in GUI properties for a compressed object

**CONFIGURING THE CELERRA PLUG-IN ENVIRONMENT:**

**I. Download the EMC Celerra Plug-in for VMware from Powerlink and install on the vSphere Client**

**II. Setup the DHSM User, Service, and Connection(s) on the Data Mover:**

**Note:** The DHSM setup is required for using Compression, Cloning, or to View properties of objects within the plug-in

**a) Create DHSM User on the Data Mover(s)**

**# /nas/sbin/server\_user ALL -add -md5 -passwd dhsm\_user**

Creating new user dhsm\_user

User ID: 525

Group ID: 525

Home directory:

Changing password for user dhsm\_user

New passwd:

Retype new passwd:

**b) Enable digest Authentication**

**# /nas/bin/server\_http ALL -modify dhsm -authentication digest -users dhsm\_user**

server\_2 :

**DHSM FACILITY CONFIGURATION**

Service name : EMC File Mover service

Comment : Service facility for getting DHSM attributes

Active : False

Port : 5080

Threads : 16

Max requests : 300

Timeout : 60 seconds

**ACCESS CONTROL**

Allowed IPs : any

Authentication : digest ,Realm : DHSM\_Authorization

Allowed user : dhsm\_user

**SSL CONFIGURATION**

Mode : OFF

Persona : default  
Protocol : default  
Cipher : default

**c) Start DHSM Service on Data Movers**

```
# /nas/bin/server_http ALL -service dhsm -start
DHSM FACILITY CONFIGURATION
Service name : EMC File Mover service
Comment : Service facility for getting DHSM attributes
Active : True
-----abridged-----
Allowed IPs : 10.250.148.134,10.250.148.184
```

**Note:** Add the IP address of the Control Station in order to view get\_attributes, or to manually test compressing, decompressing, cloning, etc., using the tools found in the /nas/tools/dhsm directory

**e) Verify DHSM Service:**

```
# server_http server_2 -info
DHSM FACILITY CONFIGURATION
Active : True
ACCESS CONTROL
Allowed IPs : 10.241.168.134,10.241.168.184,192.1.4.195
Allowed user : dhsm_user,dhsm_user1
```

**III. Add the Celerra Control Station to the vSphere Plug-in:**

**View > Home > Solutions and Applications: EMC Celerra > Add > EMC Celerra**

|  |   |
|--|---|
| Celerra Control Station Hostname / IP: | 10.250.148.184  |
| Celerra Control Station Username:      | nasadmin  |
| Celerra Control Station Password:      | *****<br>DHSM Username: dhsm_user<br>DHSM Password: ***** |

**Note:** See ‘Recent Tasks’ section in vSphere screen to monitor success or failure of actions

**IV. Integrate VMware View into the vSphere environment:**

**a. Install and configure View Manager on a Windows system**

- (1) Install VMware View Agent 4.0.1
- (2) Install VMware View Connection Server 4.0.1

**Installed Services:**

VMware View Connection Server—VMware View service

VMwareVDMDS—VMware LDAP directory services

- (3) All Programs > VMware > View Administrator Console
- (4) Click Configuration > Add license details
- (5) Click ‘Servers’ and add the vSphere/vCenter Server IP, username, and password to register

**b. Add View Manager to vSphere**

- (1) View > Home > Solutions and Applications: EMC Celerra > Add > VMware View Manager
- Vmware View Manager Server:

Username:

Password:

**PERFORMING PLUG-IN TASKS:**

**PROVISIONING NFS DATASTORES:** Provision Storage by creating new file system or using existing NFS export

**Datacenter > rightclick, EMC Celerra > Provision Storage**

--Virtual Provisioning (with file system auto-extend) is selected by default

--Manually type the ‘Initial Capacity’ & ‘Max Capacity’ values, or else the “OK” button will not enable itself! AR371466

**Advanced Options:**

- High Water Mark=file system extend threshold, default 90% full (50-99% avail.)
- ‘Uncached’ mount option is a default (NFS best practice)—NFS writes are not cached on the Server, are written straight through to storage disks
- No Prefetch unchecked means Read ahead Prefetch on large files is used
- Enable Compression not a default—when checked Dedupe is in ‘suspended’ state on file system, giving ability= to compress VMs within the Datastore
- Celerra AntiVirus Scanning--unchecked = “noscan” mount option NFS BP
- Export to Subnet--allows datastore export to a subnet

**ESX Settings:**

- Set Timeout Settings--based on best practice NFS heartbeat settings for ESX Server exports
- HostHeartbeat=12
- HostHeartbeatMaxFailures=10
- HostHeartbeatDelta=5
- HostHeartbeatTimeout=5
- See Program Files\Vmware\Virtual Infrastructure Client\Plugins\EMC Celerra\cfg\vmware\_nfs.ini file for many more ESX Export settings
- See the EMC Celerra Plug-in for VMware Troubleshooting Checklist for definitions of all the ESX Server parameters

#### **DATASTORE EXAMPLE--Celerra File System Properties for ‘dstore184-3’:**

# nas\_fs -info dstore184-3

```
name    = dstore184-3
auto_ext = hwm=90%,max_size=12288M,virtual_provision=yes
deduplication = Suspended (When Compression selected during creation)
```

**Note:** Compression was selected during “dstore184-1” creation, represented by deduplication = Suspended

# server\_mount server\_2

```
dstore184-3 on /dstore184-3 ufs,perm,rw,noscan,uncached
```

# server\_export server\_2

```
export "/dstore184-3" root=192.1.8.229:192.1.8.216 access=192.1.8.229:192.1.8.216 (ESX Server VMkernel IP addresses)
```

#### **What Happens during NFS Datastore creation?**

- a) Celerra file system created from selected Storage Pool (1fs per NFS datastore)
- b) File system mounted based on selected/default options
- c) File system exported for NFS root access to ESX Server VMkernel IP addresses
- d) NFS Datastore created in vSphere for selected ESX Hosts
- e) NFS export options are updated on ESX Hosts

#### **OPTIMIZING STORAGE:**

- Compresses or decompresses VMs (at Cluster, Datacenter, Datastore, ESX Host, Folder, or VM object level)
- Compression not enabled by default
- Compression sets file system deduplication into “suspended” state, and plug-in actions will be able to automatically compress or decompress
- Compression not allowed on Fast Clones (aka Branch Files), which are snapshots of the parent VM, and referenced back to the parent VM
- Compression not allowed on a VM (Version file, base file) that has had Fast Clones created
- Compression typically yields 30-50% space savings

#### **CLONING VIRTUAL MACHINES (VMs):**

- Can create Full independent copies of VMs or Fast clone snaps of VMs
- Total number of clones allowed varies with the number of CPU cores in the ESX Cluster or destination ESX Host
- Full clones are full and independent copies of a VM
- Full clones can be made from one datastore to another as long as the datastores reside on the same physical Data Mover
- Fast clones are file-based snapshots of VMs, with relationship to parent
- Fast clone destinations limited to the same NFS Datastore as parent VM
- Multiple clones can be created by incrementing “Clone Count” field
- Fast clones cannot be compressed
- VMs with associated Fast clones cannot be deleted
- Fast clone snap size is only 32kb, but expands to size of O/S after Fast clone VM is powered on

#### **VMware VIEW INTEGRATION:**

- Full/Fast clone integration with VMware View
- One difference between Full and Fast clone is that you do not choose a Destination NFS Datastore (can only snap to the same Dst ore as the parent VM)
- Add View Manager to vSphere and vSphere to View Manager
- View allows Fast clones to be automatically registered
- View allows refreshing of Desktop VMs on Fast clones only, within View Pools

#### **EXAMPLE--Refreshing “view1\_fast1” Fast clone:**

- a) Select a new or existing View Pool to add the clones
- b) Create full clone pool ID and name
- c) Full clone Pool desktop settings
- d) View Pools and Desktop view of clones integrated with VMware View
- e) Refresh Desktop using plug-in & View integration

#### **Refresh performs the following steps, as seen below:**

- 1) Fast clone ‘view1\_fast1’ is powered off
- 2) Fast clone is then deleted

- 3) A clone directory is created in the Datastore
- 4) A new Fast clone of the base VM ‘View1’ is created
- 5) The Fast clone is registered as a VM in the Datacenter
- 6) Refreshed Fast clone ‘view1\_fast1’ is reconfigured

#### **EXTENDING STORAGE:**

--Extending the NFS Datastore means extending the underlying Celerra file system (hence the default file system auto-extend value)  
--From example below, the (Max: 100.05GB) is the total available space in the Storage Pool on the Celerra

#### **USING CELERRA DHSM TOOLS:**

--Provided the dhsm\_user, service, and connection are setup properly, the Celerra Control Station can use CLI to perform many similar functions as done directly by the plug-in  
--Refer to the Troubleshooting Checklist document on the Partners Web site, for dhsm tool examples

#### **# ls /nas/tools/dhsm**

clone\_file → Use to create Full or Fast clones  
get\_attributes or set\_attributes → Use to view or set attributes on objects in the datastore  
dedupe\_file → Compressing files  
redupe\_file → Uncompressing files

#### **CELLERRA PLUG-IN LIMITATIONS:**

--Plug-in does not support Celerra NMFS  
--Plug-in does not support compression or fast cloning when using VMware Fault Tolerance  
--Compression not allowed on Fast clone or base VM with Fast clones  
--Deduplication does not work for ‘version files’, aka ‘fast clones’  
--The plug-in requires that Data Movers in the same Celerra cabinet have the same DHSM username and password

#### **TROUBLESHOOTING/PRIMUS SOLUTIONS:**

##### **Key Log on vSphere Client (EMC NFS Plugin.log):**

Program Files > VMWare > Infrastructure > Virtual Infrastructure Client > Plugins > EMC Celerra > **EMC\_NFS\_Plugin.log**

##### **Plug-in Configuration File (vmware\_nfs.ini):**

→Based on key parameters & best practices for ESX  
Program Files > VMWare > Infrastructure > Virtual Infrastructure Client > Plugins > EMC Celerra > cfg > **vmware\_nfs.ini** [See Online Help for details; To change logging levels from LogLevel=Info (default), set LogLevel=Error, or LogLevel=Debug]  
→See Program Files\VMware\Virtual Infrastructure Client\Plugins\EMC Celerra\cfg\vmware\_nfs.ini file for many more ESX Export settings  
→See the ‘Troubleshooting Checklist’ reference for complete definitions of all the parameters in the vmware\_nfs.ini file

##### **vSphere Task & Events--Recent Tasks:**

Use the Task & Events section to troubleshoot and observe plugin actions, at Datacenter, Datastore, ESX Host, VM level, etc.

##### **DHSM Verification on Data Mover:**

##### **# server\_http server\_2 -info | grep Active**

Active : True

##### **Knowledgebase Solutions:**

emc235511 Invalid NAS version detected by plugin  
emc235510 Celerra...not found—need to register CS with plugin  
emc234921 Error...creating datastore—needs IP configured on VMkernel  
emc235684 Cannot view properties of VM...None of Data Movers interfaces are reachable from vCenter—DHSM not setup & enabled

##### **REFERENCES:**

##### **GA release typically contains the following reference materials:**

EMC Celerra Plug-in for VMware Read Me First\*  
EMC Celerra Plug-in for VMware Release Notes\*  
EMC Celerra Plug-in for VMware Solution Guide\*+  
EMC Celerra Plug-in for VMware.zip setup.exe file\* (plug-in installable)  
EMC Celerra Plug-in for VMware Troubleshooting Checklist\*\*  
\*Located on Powerlink.com (Use Advanced Search: ‘EMC Celerra Plug-in’)  
\*\*Located at EMC Services Partner Web > Celerra > Celerra Plug-in for VMware Troubleshooting Checklist

+An updated Solution Guide to be posted soon

##### **Powerlink:**

Powerlink > Support > SW Downloads and Licensing > Downloads A-B > Adapters for Third Party Applications  
Celerra Plug-in for VMware  
Celerra Plug-in for VMware 1.1.9 Read-Me  
Celerra Plug-in for VMware 1.1.9 (.zip file with Read-Me; Release Notes; DHSM Script Setup instructions; dhsmsetup.pl Perl script; setup.exe)

#### **EMC CELERRA PLUGIN FOR VMWARE Version 1.1.20 (Aug 2010):**

→No new features with this release, just bug fixes

#### **FUTURE CELERRA VMWARE PLUGIN SUPPORT:**

- Support for FC/iSCSI LUN block provisioning of VMFS Datastores on CLARiiON luns from VI client
- Data restore from Celerra Checkpoints
  - a) Granular file restores to VM from available checkpoints  
Rightclick VM > EMC Celerra > Recover Files
  - b) Virtual Machine restores from Celerra Checkpoints  
Rightclick VM > EMC Celerra > Recover VM [Replace Active VM with recovery VM or Convert recovery VM to new VM]
- VDI Enhanced View Support

#### **COGNAC 5.6 MAINTENANCE RELEASE 12 (CMR12 5.6.49.3)**

- Mfg will begin using 500GB HDD for Maynard
- Mfg will begin using Flare 29 Patch 006 for NS-120, NS-480, & NS-960 platforms
- No new features in this release

#### **COGNAC CMR13 5.6.50 MAINTENANCE RELEASE:**

- Tolerance for Flare 30 Zeus, similar to NAS 6.0

#### **COGNAC CMR14 5.6.51 MAINTENANCE RELEASE:**

- Will tolerate arrays that use FCoE I/O modules for other Hosts, and supports Flare 30.5 Jupiter release, which introduces the FCoE I/O modules on the array
- Tolerance for Symmetrix Danube microcode on Gateways

#### **CELLERRA DENSE RACK SOLUTIONS (Yields 30 drives per 3U space vs. 15 drives):**

- GA June 2010, NOT customer installable
- NS-480 & NS-960 will introduce an Integrated factory-built option for a 42U Dense rack solution (No field installed to cust. Racks)
- Design allows for the installation of two 3U D15 DAE's in each rack row through the use of tandem rails, which will double the number of drives stored from 15 to 30 per tier (CTA Cable Track Assembly & CMA Cable Management Arm)

**Note:** Uses 3U & 6U rail adapters

- New CTA FC coupler with use of smaller 28gauge HSSDC2 cables
- SATA & EFD drives only for both solutions (except Vault will remain FC R5 4+1 + HS)

**Note:** Thermal design prevents use of FC drives in anything other than the Vault location

- NS-960 will ship with Dense racks to support up to 960 drives

- NS-960 will use a unique NS-42URK16DAE (NS system + 16 DAEs) rack for system, then an NS-42 12-tray rack for add-on DAEs & DMEs [NS-42URK24DAE—24 DAEs]

#### **NS-480 Dense Rack Product [NSD-480]:**

- NS-480 is offered with only a single 42U dense rack, max. 24 DAE's in 12 tandem drawers, totaling 360 drives
- NS-480 will use the 12-tray NS-42URK24DAE rack
- Single Control Station, 2 or 4 Blades, 1-12 tandem drawers for DAE's

→SP's can ship with 2-3 FC IO Modules, and, 2-3 iSCSI IO Modules, for a total of (5) IO modules max

#### **NS-960 Dense Rack Product [NSD-960]:**

- Min. config with single 42U rack, max. config with (3) 42U dense racks (8 DAE's in 1<sup>st</sup> rack, 12 DAE's in next two racks)

#### **Dense Rack Restrictions:**

- 2 & 5 meter cables are 100 Ohm
- Only vault drives can be FC, all other drives would be either SATA and/or EFD
- No data-in-place upgrades for Clariion
- Not sold as field installable, only factory-racked and pre-configured
- Clariion 480 and 960 models only [Flare 28.003, or 28.704, or 29.001]

#### **BAROSSA (NAS 6.0.36.4)** GA September 2, 2010

→Customer installable [Only for Integrateds, using CSA. Gateways require EMC/Partner installation.]

→Customer upgradeable [Using USM—Unisphere Service Manager and Celerra Upgrade Tool wizards, but Integrateds only)

#### **Location of NAS 6.0 Code on Powerlink:**

Software > Software Downloads and Licensing > Downloads C > Celerra Software

**Note:** Includes Apps & Tools CD, Customer Documentation set, Release Notes, Install CD's and DVD images, Upgrade CD's and DVD images, tar.gz package and CD1.iso file required to run the CUT tool manually.

**Location of Customer Service and Customer Documentation Sets on Powerlink:**

Support > Technical Documentation and Advisories > Hardware/Platforms Documentation > Celerra Network Server > General Reference

- Celerra Network Server Documentation (User Edition)
- Celerra Network Server Documentation (Customer Service Edition)

**UNSUPPORTED HARDWARE:**

CNS-14, CFS-14, 507, 514, 517, 518 Data Movers; NS350, 500, 700 [these platforms cannot support 64-bit O/S]

**UNSUPPORTED FEATURES:**

**Replication V1:**

→ A series of PUHC checks will check for Replicator V1 license, sessions, checkpoints, or fs\_copy sessions, and fail the upgrade  
Error HC\_CS\_1620181221: Celerra Replicator (V1) sessions or checkpoints

Error HC\_CS\_1620181222: fs\_copy sessions

Error HC\_CS\_1620181223: Celerra Replicator (V1) license on

**Note:** Must either delete and remove all vestiges of RepV1 & license from system, or convert from V1 to V2 first (min. of 5.6.42 to convert), then upgrade to 6.0

**VG2/VG8 GALILEO GATEWAY MODELS:**

→ Galileo models are based on the Argonaut hardware, Westmere Intel CPU's for Blades, and Blackwidow Blade enclosure

→ No Prequalifier required for Installation, but is not customer installable—requires EMC or Partner Service installation

→ Installation is defined as Racking, Cabling, NAS installation, and System Registration

→ CSA & CPW are not supported on the Gateways, but the Registration Wizard will be supported

→ Ships in a basic EMC 40U rack from factory, or is available for field installation in a customer rack

→ HW warranty 3 years with Enhanced or Premium maintenance

→ SW warranty 90 days, beyond requires SW maintenance contract

→ Systems are upgradeable using USM Celerra Upgrade Tool (NAS 6.0) [Gateways are not yet officially customer upgradeable]

→ Common front-end hardware based on Argonaut (Westmere processor) and Blackwidow enclosure (Chassis, Power Supplies, Management modules)

→ FSB is replaced with a QPI (Quick Path Interconnect) which runs at 4.8Gbps

→ Maynard Control Station

→ Replaces NS-G2 & NS-G8

→ Based on Barossa 6.0 NAS release & Flare 30 Zeus

→ Factory or Customer-racked models

**EMC Celerra VG2 (Galileo A)—mid-range system replacement for NS-G2:**

→ 4-core 2.4GHz Westmere CPU, 3@2GB DIMMs totaling 6GB, 800Mhz DDR3 SDRAM memory, 1-2 blades, 1-2 Control Stations, and using only (4) IO slots/blade (5<sup>th</sup> slot is not used)

→ The VG2 can be direct-connected to a single storage backend, or attached to a max. of (4) backends via Fabric Switch

→ VG2 supports 64TB/blade

**EMC Celerra VG8 (Galileo B)—high-end system replacement for NS-G8:**

→ 6-core 2.8GHz Westmere CPU, 24GB memory (6@4GB DIMMs) 1066 MHz DDR3 memory, 2-8 blades, 1-2 Control Stations, and using (5) IO slots/blade

→ The VG8 can be direct-connected to a single storage backend, or attached to a max. of (4) backends via Fabric Switch

→ Supports 128TB/Blade, with Marketing leeway to bump this up to 256TB if necessary

**I/O MODULE TYPES:**

**FIBRE CHANNEL:**

→ Both systems will ship with the 4-port 8Gb FC I/O by default, though customers can select the 4-port 4Gb FC I/O module

**Note:** 4Gb card supports 1/2/4Gbps speeds, while 8Gb card supports 2/4/8Gbps speeds

4-port 8Gbps FC I/O Module (303-092-102B)--Glacier

4-port 4Gbps FC I/O Module (103-054-100C)--Tomahawk

→ Both systems can support up to (3) different Ethernet I/O modules

**ETHERNET:**

4-port 1GbE copper I/O module (303-121-100A)

2-port 1GbE copper + 2-port 1GbE optical Quad module (303-122-100A)

2-port 10GbE optical--max. of two per blade (303-081-103B)

**POST GA Q4 2010 HEATWAVE FCoE IO MODULE:**

→ Celerra VG2/VG8 blades will support FCoE IO for backend connectivity, either one or two FCoE SLICs, or mixed with FC

→ FCoE SLIC is Qlogic EP8112 10GbE Enhanced Ethernet

**Note:** Initial use on blades only for FCoE backend connectivity, not 10GbE frontend ethernet connectivity

→ Blades can support both FCoE & FC SLICs

→FCoE must be in slot\_0 if expected to boot from Backend via these connections [BE0, BE1]

→Otherwise, backend connections by FCoE in slot\_1 is possible [BE2/AUX2, BE3/AUX3]

→Example switch support would be Cisco Nexus 5010/5020 or Brocade 8000

#### **I/O MODULE RULES:**

→Single FC I/O module, default order entry is for 8GB FC 2/4/8Gbps, in slot\_0 only [4GB FC 1/2/4Gbps card optional]

→8Gb FC card uses SFP+ optical connectors, while the 4Gb FC card uses SFP optical connectors

→VG8 FC card comes with all (4) SFP's in the ports (2 ports to SPs; 2 ports for Tape drives)

→VG2 FC card comes with only (3) SFP's, but a 4<sup>th</sup> can be ordered (2 ports to SPs; 2 ports for Tape drives)

→Systems must use same hardware I/O module configuration, slot-for-slot, for all Blades

→Maximum of (2) 10GbE optical I/O modules per Blade

→Both Blades have (5) I/O slots, but VG2 only uses (4), while VG8 can use all (5)

→When adding I/O modules, must not skip slots, from left-to-right

→General rule-of-thumb, if adding multiple Ethernet card types, add 10GbE to leftmost open slot, then the 2portCU 1GbE/2portOpt 1GbE card, and then the 4-port CU 1GbE card

#### **SFP PART NUMBERS & DESCRIPTIONS:**

##### **019-078-032:**

→Standard multimode optical Fibre Channel SFP for 4Gbps FC IO Modules, supporting 1/2/4/GB (CX4, NS-120, NS40, etc)

##### **019-078-042:**

→Standard multimode optical Fibre Channel SFP+ for 8Gbps FC IO Modules, supporting 2/4/8Gbps CX4, Galileo Blades, etc

##### **019-078-041:**

→Standard multimode optical 10GbE SFP+ iSCSI or Ethernet, supporting Clariion 10GbE iSCSI or Celerra 10GbE Ethernet

#### **HARDWARE FOOTPRINT:**

VG2 Gateway footprint is 4U (2U blade enclosure, & 1U for each CS—1U space if CS1 is not ordered)

VG8 Gateway will have a footprint from 4U-10U maximum (2U @ four enclosures, 1U for each CS)

#### **I/O MODULE SUPPORT:**

##### **(2) different Fibre Channel I/O SLIC modules offered:**

--Default FC I/O SLIC module is the 8GB Glacier (2/4/8Gbps) with SFP+ connectors, though customers can still select the 4GB Tomahawk (1/2/4Gbps) FC I/O module when ordering (SFP connectors)

--Only (3) of the FC ports will be populated with SFPs (a 4th SFP can be ordered) on the VG2 [All ports have SFPs for VG8]

##### **(3) different Ethernet I/O SLIC modules offered:**

4-port 1GbE copper (303-121-100A) Thunderchild Broadcom 5715 controller

2-port 1GbE copper & 2-port 1GbE optical card (303-122-100A) Thunderbolt Broadcom 5715 controller

2-port 10GbE optical (303-081-103B) Poseidon Broadcom 5771 controller

→IO module slots are based on PCI-Express Gen 2 bus

→IO Modules used will be Tomahawk (4Gbps FC), Thunderchild, Thunderbolt, Poseidon I Broadcom 57710 controller (2-port 10Gbps), Glacier (8Gbps FC), and FCoE Heatware module, with PCI-Express Gen 2, post-GA release

**Note:** FC IO module on blades can be either 4Gbps or 8Gbps, but can only have (1) FC card, and in slot\_0

#### **IO Module Port Numbering:**

##### **FC Card:**

BE-0-0, BE-0-1, AUX-0-0, AUX-0-1

##### **Slot 1 Card:**

cge-1-0, cge-1-1, fge-1-0, fge-1-1 (Example)

##### **Slot 4 Card:**

cge-4-0, cge-4-1, fge-4-0, fge-4-1 (Example)

#### **DETERMINING PLATFORM/MODEL INFORMATION—EXAMPLE VG8:**

# /nas/sbin/model

VG8

# /nas/sbin/serial

FNM00101600060

# /nas/sbin/t2cab

3 - CELERRA IP CABINET

The Cabinet child is : 14 - a CELERRA ARGONAUT

# /nas/sbin/t2vpd -m

Argonaut\_B

# server\_sysconfig server\_2 -P

server\_2 :

Processor = Intel Six Core Westmere

Processor speed (MHz) = 2800

Total main memory (MB) = 24576

Model = VG8

QPI speed (MHz) = 6400  
 Bios Version = 04.73  
 Post Version = Rev. 24.20  
 Firmware Version = Rev. 10.50  
 Family ID = 0x0012  
 FRU ID = 0x000c

#### # nas\_inventory -list head

| Component                      | Type    | Status | System ID                      |
|--------------------------------|---------|--------|--------------------------------|
| Celerra VG8 APM000832011840000 | Celerra | OK     | Celerra VG8 APM000832011840000 |

→GUI also displays model name

#### **DART CONSOLE/T2CAP FACILITY:**

→New tier2 utility using a t2console/t2cap listener that will continually record management switch information (from Coldfire firmware) into separate files for each Data Mover in the /var/log/t2cap\_slot\_x.log

→Only available with Celerra Argonaut systems (VG2/VG8, etc.)

→Each mgmt switch is responsible for reporting console information for its blades to the Control Station

#### # /nasmcd/sbin/t2cap

T2cap - Version 1.0 - 11/24/09

usage: t2cap <slot> -f

<slot>: Sends console output to the screen.

-f : Writes console output to a file of the form /var/log/t2cap\_slot\_\*

Source revision 5.2.0.1 - 06/12/03

#### /var/log/t2cap\_slot\_2.log

-rw-r--r-- 1 root root 20325 Jun 14 21:28 t2cap\_slot\_2.log

**Note:** The t2cap logs record POST and bootup messages, as well as fibre channel login and normal server log boot.cfg tasks

#### Verifying that T2cap is running?

#### # ls -la /var/run |grep t2cap

```

-rw-r--r-- 1 root root 4 Aug 6 10:36 t2cap_slot_2.pid
-rw-r--r-- 1 root root 4 Aug 6 10:36 t2cap_slot_3.pid
-rw-r--r-- 1 root root 4 Aug 6 10:36 t2cap_slot_4.pid
-rw-r--r-- 1 root root 4 Aug 6 10:36 t2cap_slot_5.pid
-rw-r--r-- 1 root root 4 Aug 6 10:36 t2cap_slot_6.pid
-rw-r--r-- 1 root root 4 Aug 6 10:36 t2cap_slot_7.pid
-rw-r--r-- 1 root root 4 Aug 6 10:36 t2cap_slot_8.pid
-rw-r--r-- 1 root root 4 Aug 6 10:36 t2cap_slot_9.pid

```

#### # ps -ef |grep t2cap

```

root 4066 3489 0 Aug06 ? 00:00:01 /nasmcd/sbin/t2cap 2 -f
root 4067 3489 0 Aug06 ? 00:00:01 /nasmcd/sbin/t2cap 3 -f
root 4068 3489 0 Aug06 ? 00:00:00 /nasmcd/sbin/t2cap 4 -f
root 4069 3489 0 Aug06 ? 00:00:01 /nasmcd/sbin/t2cap 5 -f
root 4070 3489 0 Aug06 ? 00:00:00 /nasmcd/sbin/t2cap 6 -f
root 4071 3489 0 Aug06 ? 00:00:00 /nasmcd/sbin/t2cap 7 -f
root 4072 3489 0 Aug06 ? 00:00:00 /nasmcd/sbin/t2cap 8 -f
root 4073 3489 0 Aug06 ? 00:00:00 /nasmcd/sbin/t2cap 9 -f

```

#### **VG2/VG8 PRODUCT SERIAL NUMBER & PART NUMBERS:**

##### Product Part Numbers (PPN):

**900-567-004** (VG2 Galileo A)

**900-567-005** (VG8 Galileo B)

→The intentions are that the PSN & PPN will be persistent for enclosure replacements

#### # cat /etc/product\_numbers.db

#Product Numbers Database file

#This file contains the Product Serial Number(PSN) and Product Part Number(PPN).

Version=1.0

PSN=FNM00101600060

PPN=900-567-005

#### # tail /var/log/serial.log

---

Completed -- serial -set -psn FNM00101600060 -ppn 900-567-005 -force

Mon Jun 14 11:04:27 EDT 2010

---

#### Changing PPN Number:

→Use command: # /nas/sbin/serial -set -psn FNM00101600060 -ppn 900-567-005 –force

→Connect to Blade in enclosure, access POST Resume information, set PPN, reboot

#### **DME0 ENCLOSURE RESUME INFORMATION VG8:**

# cat /nas/log/enclosure\_resume.enclosure\_0.xml

```
<?xml version="1.0"?>
<EnclosureResume EnclosureID="0">
  <RESUME_INFORMATION_MIDPLANE>
    EMC_BARE_ASSEMBLY_PART_NUMBER="100-562-178      "
    EMC_BARE_ASSEMBLY_REVISION="A06"
    EMC_BARE_ASSEMBLY_SERIAL_NUMBER="CF2A5094200021  "
    EMC_PART_NUMBER="100-520-113      "
    EMC_ARTWORK_REVISION="      "
    EMC_ASSEMBLY_REVISION="A01"
    EMC_SERIAL_NUMBER="FNM00101600060      "
    EMC_PRODUCT_PART_NUMBER="900-567-005      "
    EMC_PRODUCT_SERIAL_NUMBER="FNM00101600060      "
    VENDOR_NAME="      "
    LOCATION_OF_MANUFACTURE="      "
    YEAR_OF_MANUFACTURE="      "
    MONTH_OF_MANUFACTURE="      "
    DAY_OF_MONTH_OF_MANUFACTURE="      "
    ASSEMBLY_NAME="GALILEO B DUAL DATA MOVER - INT "
```

#### **VG2/VG8 FRU/CRU INFORMATION:**

→Nimitz mgmt switch reports on most system FRUs except for SFPs & DIMMs, which DART reports on

→Coldfire mgmt switch firmware is flashed during upgrades, and setup\_enclosure determines the correct enclosure switch type in order to select the correct mgmtfirmware\_nm.s19 firmware & bootblock\_nm.s19 package for the system

#### **CRUs:**

Blade Encl Pwr Supply  
Blade I/O modules  
Blade SFPs  
Blade Management modules

#### **FRUs:**

Control Station  
Blade DIMMs  
Blade Assembly (Includes motherboard, not DIMMs—called a “CPU Module”)  
Blade Enclosure

#### **GALILEO HARDWARE UPGRADES AT GA:**

##### **Customer/EMC/Partners:**

Add SFP for 2<sup>nd</sup> tape connection VG2 only  
Add Blades VG2/VG8  
Add Blade I/O modules

##### **EMC/Partners only:**

Add 2<sup>nd</sup> Control Station VG2/VG8  
Add Blade Enclosure VG8 only  
Change I/O module for upgrade of same type module

#### **GALILEO BACKEND STORAGE SUPPORT:**

→VG2 supports direct connect (single CLARiiON only) or fabric switch connect to CLARiiON or Symmetrix (up to four arrays)  
→VG8 supports direct connect (single CLARiiON only) or fabric switch connect to CLARiiON or Symmetrix (up to four arrays)

#### **CLARiiON:**

CX300, 500, 700 [Flare 19, 24, 26]  
CX3-10, -20, -40, -80 [Flare 22, 24, 26]  
CX4-120, 240, 480, 960 [Flare 28, 29]

#### **SYMMETRIX:**

DMX1, 2, 3, 4  
VMAX, VMAX-SE

**Note:** Galileo also supports VMAX using 8GB FC with Enginuity 5874.207.168

#### **INCREASED STORAGE CAPACITY:**

→256TB/Blade on NS-960/NS-G8 architecture (file system still limited to 16TB)  
→iSCSI LUNs increased to 16TB

#### **DART 64-BIT OPERATING SYSTEM:**

- DART uses a 64-bit O/S, which allows for use of larger memory structures & addresses, increasing performance
- 64-bit DART does NOT mean 64-bit file systems [still 32-bit only]
- 4GB Memory systems will see a 10% performance penalty when upgraded to 64-bit
- Upgraded 6.0 systems will automatically use all physical memory (e.g., NS-960/NS-G8 will use all 8GB, whereas only 4GB used with 5.6)
- Control Station Linux remains 32-bit
- DART file systems remain 32-bit
- Ability to address >than 4GB memory
- Obsolete Hardware that does not support 64-bit instructions:  
510 DM, NS350, NS500, NS600, NS700, NS704; 507, 514, 517, 518 DMs, CNS14, CFS14

# **server\_sysconfig server\_2 -P**

server\_2 :

Processor = Intel Six Core Westmere  
Processor speed (MHz) = 2800  
Total main memory (MB) = 24576  
Model = VG8

## **6.0 ZEUS FLARE 30 (CX4-only) TOLERANCE:**

- Celerra will tolerate CX4 arrays running Flare 30, but many specific R30 features are not supported for use by Celerra
- Celerra does not support Auto-Tiering (CLARiiON FAST) policy
- Celerra does not support LUN Compression
- Celerra does not support DLU Fully Provisioned pool LUNs
- Celerra does not support mixed drive types in enclosure or Storage Pools

**See emc241591—nas diskmark -m -a:**

Warning 17717198855: FAST device 0033 on APM00083201184 was not marked because FAST is not supported in this release.

Warning 17717198861: Fully provisioned CLARiiON pool device 0065 on APM00083201184 was not marked because it is not supported in this release.

Warning 17717198863: CLARiiON pool device 000B on APM00083201184 was not marked because compression is not supported in this release.

→Zeus supports SATA EFD drives. SATA EFD & FC EFD drives are treated as the same drive types by Celerra (e.g., clarefd/clefd pools)

**Note:** SATA EFD drives are expected to be 100-200GB Samsung drives using Chariot paddle card

→Unlike CLARiiON, Celerra does not support the use of mixed drive technologies in the same Enclosure or Pool (Regular SATA, FC, EFD), will produce diskmark errors

**emc242009 examples:**

- Mixed drives in enclosure will show CPW GUI Error: ‘No provisioning can be done’
- nas\_diskmark -m -a: ‘Warning...Device...not marked...cannot contain mixed disk drive types.’
- setup\_clariion -init: ‘...disk devices can’t be configured...’

**nas diskmark:**

Warning 17717198856: Device 000B on APM00083201184 was not marked because a single CLARiiON device cannot contain mixed disk drive types.

**Setup clariion -init** script complains that it has no matching template for a Shelf with mixed drives:

*16 disk devices can't be configured using User Defined.*

*Configuration template doesn't completely match hardware.*

*Some disks may need to be relocated to fully configure the selected template.*

**EMC FAST Cache (EFD Drives):**

- FAST caching extends the arrays caching capability by mapping frequently accessed data to Flash drives, dramatically increasing the scale of caching [73GB – 2TB], as well as increasing system performance—DRAM cache cannot be scaled the same on a cost effective basis
- FAST Cache is created in RAID1 mirrored Read/Write mode, using a Policy engine and 64KB memory map strategy
- Requires FAST Cache enabler on SPs, and SP reboot
- Configurable on RAID Group and Pool-based luns via Unisphere
- Once installed FAST Cache is enabled by default on newly created Pool or RAID luns
- FAST Cache is non-volatile, meaning that data is preserved on power loss
- FAST Cache is not used on Vault drives
- Celerra CSA/CPW allows user to skip configuration of EFD drives to allow for FAST Cache configuration first
- FastCache can be enabled on individual RAID Group luns
- FastCache is enabled at the Storage Pool level only, meaning all luns within a pool will have the capability
- Luns and Pools created after installing the FASTCacheEnabler will be enabled by default
- Once enabled, applications can experience performance improvements as frequently accessed chunks of data are copied to FastCache, and demoted when usage falls

→FASTCache works with extents of 64KB (DRAM works at 2-16KB), operating at millisecond to microsecond speeds

→FASTCache uses a Policy engine (manages IO flow into and out of cache) and Memory map of 64KB chunks of storage, with a memory map copy stored in array DRAM

#### Creating FAST Cache from CLI:

# /nas/sbin/navicli -h 10.250.40.25 cache -type efd -create disks <disknames> -mode ro | rw -raidtype <rtype>

# /nas/sbin/navicli -h 10.250.40.25 cache -type efd -info | -stats | -status | -disks

#### Creating FAST Cache from Unisphere:

→Unisphere > System > Manage Cache

#### STORAGE POOL ENHANCEMENTS FLARE 30:

→Support for RAID 5, 6, & 1/0 for Pools, though any given pool can only be one RAID type

→Support for mixed drive types within Pools, but not for RAID Groups (this is how FAST works, by using different drive types for tiering within a single pool of luns)

→Support for up to three different drive types (FC, SATA, EFD) per pool (though Celerra does not support different drive types)

→CLARiiON Best Practice with Flare 30 is to have heterogeneous pools of mixed disk types, to support FAST Tiering

**Note:** Hetero pools support EFD, SATA, FC three tiers, and a single RAID protection type, and should be the same size

→Max number of drives 955 per Storage Pool on CX4-960 (RAID Groups limited to 16)

→Added support for DLU fully provisioned luns in Storage pools (previously just Thin luns in pools)

**Note:** DLU Thick LUNs cannot be used as WIL luns or Clone Private luns (CPL)

→CLARiiON now supports 14TB lun size

→Thin LUNs can now be expanded (both thick and thin luns can be expanded)

→Pool luns can now be shrunk (but only with Windows 2008 hosts)

#### LUN shrink:

a) Windows 2008 Server only, use Disk Mgmt to shrink a file system

b) Use Diskraaid utility from VDS provider in EMC Solutions Enabler to shrink pool LUN

c) Array background process migrates the Flare, Thick, or Thin LUN to a new Thin Lun, returning 8k chunks to the Storage Pool

#### CLARiiON LUN TYPES:

FLU Flare luns are traditional RAID Group luns, and are not TLU or DLU luns, and are supported by Celerra 6.0

TLU Thin luns are not fully provisioned pool luns, and are supported by Celerra 6.0

DLU Thick luns are fully provisioned pool luns, and not supported by Celerra 6.0

#### FLARE LUNs vs. THICK LUNs vs. THIN LUNs:

→Flare & Thick lun performance is best, Thin lun performance not as good

→Flare lun Raid Group disks limited to 16, while Thick & Thin Storage Pool disk limit is 955 disks

→Flare luns support all RAID protection types, Thick & Thin support only R5, R6, and R1/0

→Flare luns are expanded using Metaluns, Thick & Thin luns do not use Metaluns for expansion

→Compression converts Flare and Thick luns into Thin luns

→Flare luns do not support Auto-Tiering (FAST), while Thick & Thin luns do

→Flare luns support disk spin-down, Thick & Thin luns in Storage Pools do not

→Flare & Thick luns support RLP (Reserve Lun Pool), while Thin luns do not

→Flare luns support WIL (Write Intent Log) & CPL (Clone Private Lun), while Thick & Thin luns do not

#### CLARiiON LUN CONCEPTS & CELERRA SUPPORT:

→FLU ‘Flare’ Luns are traditional CLARiiON LUNs created in RAID Groups and are not TLU or DLU luns, and 6.0 supports

→TLU ‘Thin’ Luns are created in “Pools”, and are supported by 6.0

→DLU ‘Direct’ Luns are also created in “Pools”, but are NOT supported by 6.0

→Auto-tiering, or CLARiiON FAST, can only be set on Luns created from “Pools”, and is NOT supported by 6.0

→CLARiiON LUN compression is NOT supported by 6.0

→Enclosures with mixed disk technologies is supported by CLARiiON, but is NOT supported by 6.0 and diskmarking will fail

#### CLARiiON VIRTUAL PROVISIONING:

→Much has changed since VP was introduced with Flare 28. The concept of ‘Thin Pools’ no longer exists. You must create a ‘Pool’ under the Unisphere > Pools/RAID Groups section, then create LUNs and check the box for “Thin” if you want thin LUNs. The default would be to create Thick Direct LUNs (aka, fully provisioned)

→VirtualProvisioning enabler requires SP reboot

→Thin Luns in the Pools are also called TLU’s, whereas fully provisioned luns in the Pools are called DLU’s

→Celerra supports Clariion Thin Luns, but only if Auto-tiering policy is not set

Unisphere > Storage > LUNS > LUN 22 >properties > Tiering tab > Set the “Tiering Policy” to “\*No Data Movement”

#### Celerra will diskmark TLU Thin Luns if Auto-Tiering is turned off:

# nas\_diskmark -m -a

Discovering storage (may take several minutes)

done

Info 26306752254: APM00083201184 reassigned LUN 0022 in storage group 'Celerra\_sleet-120' from host id 0008 to 0024

Info 26306752254: APM00083201184 reassigned LUN 0023 in storage group 'Celerra\_sleet-120' from host id 0006 to 0025

**Celerra will not diskmark DLU pool luns, as they are not supported:**

# **nas\_diskmark -m -a**

Discovering storage (may take several minutes)

done

Warning 17717198861: Fully provisioned CLARiiON pool device 0065 on APM00083201184 was not marked because it is not supported in this release.

Warning 17717198861: Fully provisioned CLARiiON pool device 0064 on APM00083201184 was not marked because it is not supported in this release.

**Celerra will not diskmark LUNs that have a CLARiiON FAST Tiering Policy Set:**

# **nas\_diskmark -m -a**

Discovering storage (may take several minutes)

done

Warning 17717198855: FAST device 0033 on APM00083201184 was not marked because FAST is not supported in this release.

Warning 17717198855: FAST device 0032 on APM00083201184 was not marked because FAST is not supported in this release.

**FAST (Auto-Tiering):**

- Via policy on a per-lun basis, dynamically moves ‘hot’ data to higher performing drives, and ‘cold’ data to lower performing drives
- Requires Virtual Provisioning and AutoTiering enablers
- Once installed, all Pool luns will have AutoTiering enabled by default
- With Flare 30, CLARiiON FAST is now a two-way data relocation solution, with a flexible scheduler
- Tiering is now based on drive speed and capacity, within a Storage Pool
- Lun properties, “Tiering” tab, set each LUN to “No Data Movement” to disable auto-tiering

Name of the software package: -AutoTiering

**Note:** Installing AutoTiering enabler does not require SP reboots

→ Tiering requires that the User set the Tiering Policy and the Data Relocation Scheduler

**Celerra will not diskmark LUNs that have Compression set:**

# **nas\_diskmark -m -a**

Discovering storage (may take several minutes)

done

Warning 17717198863: CLARiiON pool device 000B on APM00083201184 was not marked because compression is not supported in this release.

→ You can uncompress the LUN under properties “Compression” tab

**Note:** You cannot compress DLU luns until after the LUN has been written to. You cannot compress RAID Group FLU luns without migrating the existing data to a Pool first. You can compress TLU thin luns without having any data in them.

**CLARiiON LUN COMPRESSION:**

- Ability to take a source fully provisioned Pool lun (DLU), or regular FLARE Raid Group lun (FLU), and migrate the lun to a new destination lun (pool) that then becomes “thin” (TLU), and after migration the original Flare or Thick lun is deleted, and the Thin lun is renamed with the original LUN number
- As mentioned earlier, Celerra does not support, though if you compress a Celerra Lun, the data is still available, but Celerra storage health checks and associated backend operations will subsequently fail
- Turning off LUN Compression changes the lun from Thin to Thick (aka, Fully provisioned DLU)
- Compression Enabler requires SP reboot
- So, compression can be applied to any Lun, and the results will be a new Thin Lun
- For Celerra, if you turn on compression, even though the data is migrated to a Thin lun and is still available to Clients, we will not be able to make any further changes to storage because health checks will fail diskmarking operations because of the compressed lun
- You cannot turn off LUN compression and migrate back to a RAID Group, it must remain in a Pool
- Stopping a Flare or Thick lun compression operation will make the Lun Thick (i.e., fully provisioned)

**Viewing Compression Attributes from CLI:**

# **/nas/sbin/navicli -h 10.241.168.179 compression -list**

Compression Feature State: On

LOGICAL UNIT NUMBER: 25

Name: Celerra\_sleet-120\_25\_d24

UID: 60:06:01:60:67:80:21:00:EC:8F:5D:49:0F:65:DF:11

Destination Pool ID: N/A

Current State: Compressed

Status: OK(0x0)

Percent Complete: 100

Rate: Medium

Total Capacity (Blocks): 281406912

Total Capacity (GBs): 134.185

Consumed Capacity (Blocks): 8389120

**Turning on LUN Compression via CLI for Flare LUN:**

# /nas/sbin/navicli -h 10.241.168.179 compression -on -l 23 -destPoolId 2 -rate high

**Turning on LUN Compression via CLI for Thick or Thin LUN:**

# /nas/sbin/navicli -h 10.241.168.179 compression -on -l 24

**CLARiiON SUPPORT FOR MIXED DRIVES WITHIN SAME “POOL”:**

→CLARiiON still does not support the used of mixed drive types within a RAID Group

→However, with Flare 30, CLARiiON does support the use of mixed drive types within a “Pool”

→Celerra does not support this, as demonstrated by the following

# nas\_diskmark -m -a →Message output below is inaccurate for CLARiiON, though the point is Celerra still does not support Discovering storage (may take several minutes)

done

Warning 17717198856: Device 000B on APM00083201184 was not marked because a single CLARiiON device cannot contain mixed disk drive types.

**CELERRA CLARiiON PUSH REGISTRATION:**

→Celerra will be able to provide CS Hostname, IP Address, Vendor string, and Unique ID for Blade registrations with the Array, thereby allowing for host-based protections of LUNs used by Celerra (to be implemented with Flare 31)

→Celerra upgrades where the Celerra Host initiator naming convention is not in use will not use this feature

**Sample Output navicli port -list:**

# /nas/sbin/navicli -h 10.241.168.179 port -list

Information about each HBA:

HBA UID: 50:06:01:60:C1:E0:D9:54:50:06:01:68:41:E0:D9:54

Server Name: sleet-120

Server IP Address: 10.241.168.178

HBA Model Description:

**HBA Vendor Description: EMC Celerra File Server 6.0** →New with 6.0

HBA Device Driver Name:

Information about each port of this HBA:

SP Name: SP A

SP Port ID: 1

**HBA Devicename: server\_3:scsi-0** →New with 6.0

Trusted: NO

Logged In: YES

Source ID: 1

Defined: YES

Initiator Type: 3

StorageGroup Name: Celerra\_sleet-120

**6.0 NEW LINUX KERNEL RHEL 5.2:**

→Based on Red Hat Enterprise Linux 5 (RHEL5)

→Kernel upgraded from 5.6 2.6.9 to 6.0 2.6.18

→Required for IPv6 support

→Required for additional security improvements

→Will not support CNS, CFS, NS350, NS500, NS600, or NS700 platforms

→Adds ipv6calc tool

→Will change ProxyARP to ProxyND service for the IPv6 support

# uname -r

2.6.18-128.1.1.6005.EMC

**STORAGE API VERSION 6.0.36-4:**

# rpm -qa |grep -i nasStorage

nasStorageAPI-8.2-12

**NEW CONTROL STATION PARTITIONS WITH 6.0:**

/celerra/audit →system partition for Security Auditing logs

/celerra/ccc →Celerra Configuration Collector, also later 5.6 releases

/celerra/commoncache →Unisphere C3 Celerra Common Cache mechanism to improve GUI performance

/celerra/wbem →ECOM & SMI-S binaries & logs

**CELERRA 6.0 EXPRESS INSTALL [EI] STORY:**

→”Express Install” replaces the traditional Classic Install, and are designed for factory reset of Integrated systems, Control Station recovery, normal Gateway system install, and uses either the boot:install dialogue, or “kickstart” answer file. A Control Station “recover” is issued for a recovery or CS replacement.

→Reinstalls of systems require backend cleanup, similar to current procedures, using a new cleanup script on the Installation media [/media/cdrom/tools/nas\_systemcleanup] or the traditional nas\_raid –s cleanup script from /tftpboot

→Installs are done via CLI only, requiring serial connection to the Control Station, and initiated with either the “install” or “kickstart” options.

→Gateway products require EMC/Partner installation, and will use the EI methodology for either “install” or “kickstart”.

→Integrated products ship out as pre-installed systems from Manufacturing, and make use of an Install Guide and CSA.

→An integrated Celerra would only use the EI methodology if a system needed to be reset to factory (mfg) condition.

→Express Installs are not much different than a classic NAS install of a system.

→Systems requiring EI install (Gateways) are not customer installable.

→Express Installs are slightly different from from Classic Installs (pre-6.0), though the look and feel are pretty much the same. EI is a little faster overall—uses a ghost imaging process.

→Install media consists of either a 2-disc bootable CD-ROM set, or a single bootable DVD disc, approximately 960MB total

**Note:** Stage 1 would be to scribe Linux image on IDE drive & bring CS online. Stage 2 would be to scribe NAS software on backend and bring up the system.

#### **From the Express Install CD-ROM or DVD, these are the EI options:**

**boot: install** [Interactive install, Question and Answers throughout]

**boot: kickstart** [Interactive, but most Q's upfront—recommended]

**boot: recover** [Control Station recovery option only, not for fresh installs]

#### **DUAL COMPACT DISC INSTALL MEDIA:**

##### **Explorer view of contents on CD1:**

images

isolinux

tools

.cd1

→CD1 is bootable and leads to the “boot:” prompt

##### **Explorer view of contents on CD2:**

images

isolinux

tools

.cd2

#### **DVD INSTALL MEDIA:**

##### **Explorer view of contents on DVD:**

images

isolinux

tools

.dvd

→DVD is bootable and leads to the “boot:” prompt

#### **Benefits of Express Install:**

→Provides an ability to restart from interrupted installs

→Provides a method for setting a system back to factory configuration

→Provides for a simplified & quicker Control Station replacement process

→Overall reduces the time for installations

→Better media integrity. Media files are verified by checksums after they are copied to the installation partition on the Control Station from the CD or DVD media.

#### **REINSTALLS OF EXISTING SYSTEMS:**

→Gateways would require manual cleanup of the Celerra configuration for reinstall

→Integrateds use the /media/cdrom/tools/nas\_systemcleanup script on EI CD1

→Optionally, Integrateds can also use the traditional /tftpboot/bin/nas\_raid –s cleanup script

→If the Celerra was previously “joined” to the CLARiiON Storage domain, make sure it is “unjoined” as part of the cleanup steps  
Ultimately, you may need to manually deconfigure a Celerra from the backend in order to perform a fresh install ( remove Luns; RGs; SG; Initiators, etc)

→Remember to set the SPs back to the Celerra internal network IPs (128.221.252.200 & 128.221.253.201, SPA SPB, respectively)

#### **Generic Integrated cleanup Steps:**

1) Stop Proxy ARP service (resets the SPs back to Celerra internal IPs) using “clariion\_mgmt –stop”

2) Mount 1<sup>st</sup> CD-ROM media and execute cleanup script

# mount /media/cdrom

# /media/cdrom/tools/nas\_systemcleanup -all

NAS Control/Data LUN Removal Utility - Version 0.3

Fri Jul 16 16:54:42 2010

Enter Service Password [or type "ABORT" to exit]: [**TotalDestruction**]

You are about to prepare the system for cleanup operation. NAS service will be stopped and all File Systems will be umounted from Data Movers.

Do you want to proceed with the cleanup? [YES/NO] yes

**Note:** The traditional /tftpboot/bin/nas\_raid –s cleanup still works and is another option without requiring 6.0 Media

3) Do the reinstall without assigning the CS Hostname & external IP address

4) Use the CSA utility to complete the configuration of the system

### **Example of Successful System Cleanup:**

# **/media/cdrom/tools/nas\_systemcleanup -all**

NAS Control/Data LUN Removal Utility - Version 0.3

Wed Sep 1 07:35:47 2010

Enter Service Password [or type "ABORT" to exit]:

You are about to prepare the system for cleanup operation. NAS service will be stopped and all File Systems will be umounted from Data Movers.

Do you want to proceed with the cleanup? [YES/NO] yes

Checking for integrated Celerra.....

Stopping NAS service....

Initiating NAS service stop ... OK

Umount all FileSystems from Data Movers to prepare data deletion...

Confirming that file systems are unmounted on each Data Mover ...

Slot\_2 ... OK

Slot\_3 ... OK

Mount NAS partition...

Mounting system mountpoints ...

/nas ... OK

WARNING!! OPERATION NOT REVERSIBLE

You are about to delete all user and system luns from the storage system.

You will not be able to recover your data once this is done.

Do you want to proceed with the cleanup? [YES/NO] yes

WARNING!! OPERATION NOT REVERSIBLE

You are about to delete the content of this IDE. The Control Station will not be able to boot unless you install new system image on the system.

This operation could take some time to complete, do not power off or reset the Control Station at this time

Do you want to proceed with the cleanup? [YES/NO] yes

Operation Succeeded.

### **SYSTEM CLEANUP NOT SUPPORTED ON GATEWAY SYSTEMS (VG8):**

# **/media/cdrom/tools/nas\_systemcleanup -all**

NAS Control/Data LUN Removal Utility - Version 0.3

Fri Jul 16 16:54:42 2010

Enter Service Password [or type "ABORT" to exit]:

You are about to prepare the system for cleanup operation. NAS service will be stopped and all File Systems will be umounted from Data Movers.

Do you want to proceed with the cleanup? [YES/NO] yes

Checking for integrated Celerra.....

Error!! Cannot determine current system configuration.

This tool can only run against integrated systems.

Please investigate and correct the issue shown, and retry this task later.

Operation Failed.

→Run CSA to complete “Installation” of system

### **TYPICAL INSTALL CONSOLE OUTPUT MESSAGES:**

boot: install

Is this a Secondary Control Station(y/n/a)?

please swap in the second CD, and press ENTER when ready

System Reboot ...

Restarting system.

Starting EMC NAS Factory Installation

Setting up the internal network...

Setting up the Enclosure (this may take several minutes) ...:Done

Is this the expected number of movers in the cabinet? [yes or no]: yes

Checking for existing System RG/LUN:

Setting up SPA-SPB communication

Backend Security Setup Successfully.

Rebooting Data Movers...

EMC NAS PXE-BOOT Setup

## **CELERRA 6.0 UPGRADE STORY:**

→Upgrades are supported from NAS 5.6 to 6.0, or within the 6.0 family

→CFS, CNS, & NSxxx (CX series) Celerra platforms no longer supported

### **CUT TOOL:**

→Celerra is customer upgradeable using USM (Unisphere Service Mgr—CUT tool—Celerra Upgrade Tool)

→USM replaces the CLARiiON NST tool

→CUT does not support Dual CS systems (PUHC check will prevent)

→CUT does not support patch/nas.exe upgrades

→CUT will upgrade Gateway systems, but we prefer only EMC does this right now, not customers

### **Several Upgrade Methods:**

1) GUI CUT tool is the preferred method for upgrades for Customers/EMC/Partners

2) CLI CD-ROM or DVD media using install\_mgr –mode upgrade

3) CLI ISO image loop mounted on local hda10 partition for 6.0 to 6.0 upgrades

**Note:** Upgrade media is different than Install media

→Out-of-Family Recovery during upgrades

boot: serialrescue

/mnt/sysimage (locates existing Linux install and mounts under /mnt/sysimage)

sh-3.2# /recover

### **Message Seen During Upgrades:**

# cat /etc/motd.standby

EMC Celerra Control Station Linux Tue Jul 13 17:33:25 EDT 2010

Warning!!upgrade is in progress from 6.0.36-2 to 6.0.36-4

Warning!!Please log off IMMEDIATELY if you are not performing the operation on the Celerra

## **INSTALL MANAGER CHANGES & REFERENCE TO GUI:**

**./install\_mgr -m upgrade -list**

**Note:** Lists out the Tasks and estimated time for the upgrade

**./install\_mgr -m upgrade**

-----abridged-----

You can do this upgrade as an online upgrade. In online upgrades:

\* You MUST reboot the blades AFTER the upgrade to run the new software.

\* You will not be prompted during the upgrade about blade reboots.

\* File service disruption times shown above do not apply.

Do you wish to do an online upgrade? [ yes or no ]? > no

You can now perform upgrades with an easy-to-use GUI. The GUI can be run from

any PC that has access to this Celerra. To use the GUI:

1. Install Unisphere Service Manager (USM) on the PC from the Celerra

Apps and Tools CD.

2. Open USM and login into the system you wish to upgrade.

Do you wish to continue to use install\_mgr instead of USM [ yes or no ]?

> yes

## **SEVERAL UPGRADE METHODS:**

### **1) GUI--Unisphere Service Manager—Celerra Upgrade Tool**

Preferred upgrade method for qualified Integrateds is to use the Unisphere Service Manager CUT Wizards (Celerra Upgrade Tool).

But, we prefer that only EMC Providers perform Gateway upgrades using CUT, and not customers.

### **2) CD-ROM or DVD Media**

An alternate method for upgrade of Gateways, dual Control Station systems, etc., is to use either the 2-disk DCD (Dual Compact Disc) set, or a single DVD ISO image, and use the traditional install\_mgr –mode upgrade script

### **3) ISO Loop Upgrade within 6.0**

Another method for 6.0 to 6.0 upgrades is to load the DVD ISO image directly to the local partition HDA10 partition on the IDE drive, copy the ISO image to the partition, loop mount the image, and run the install\_mgr upgrade script from /celerra/upgrade/EMC/nas. See emc245885 for the procedure.

**CD-ROM UPGRADE MEDIA:**

**Explorer View Contents for CD1:**

CLARiiON

EMC

images

isolinux

repodata

/disc1

.discinfo

.treeinfo

.version1

autorun.inf

EMC.ico

**Note:** Disc 1 (CD1) has bootable linux code, and RPM packages, approximately 340MB

**Explorer View Contents for CD2:**

CLARiiON

EMC

repodata

/disc2

.discinfo

.treeinfo

.version2

autorun.inf

EMC.ico

**Note:** Disc 2 (CD2) has NAS code, approximately 470MB. Files are staged to CS harddrive and verified with checksums (install\_mgr -m upgrade)

**5.6 to 6.0 Upgrade Partition:**

/dev/hda9 14G 2.2G 11G 17% /partitionU

**6.0 DOS PARTITION RECOVERY:**

# /nas/sbin/install\_init -r

**NEW 6.0 LOCAL IDE PARTITIONS:**

# /sbin/fdisk -l

-----edited----- →Three new IDE partitions carved out of the local disk

/dev/hda8 26224 28312 16779861 da Non-FS data

/dev/hda9 28313 30139 14675346 da Non-FS data

/dev/hda10 30140 30401 2104483+ 83 Linux

**DETERMINING IF A CONTROL STATION HAS A DVD CAPABLE DRIVE:**

# cat /proc/sys/dev/cdrom/info

Example of System that does not have DVD reader:

# cat /proc/sys/dev/cdrom/info |grep "read DVD"

Can read DVD: 0

Example of System that can read DVD:

# cat /proc/sys/dev/cdrom/info |grep "read DVD"

Can read DVD: 1

Only the following Control Stations are DVD capable:

100-520-581

100-560-974

100-520-665

**6.0 UPGRADES VIA DUAL COMPACT DISCS (DCD) or single DVD:**

→DCD uses different software images than for Express Installs

→Need to provide two CD's for 6.0 upgrade media

Disc 1 has bootable linux code, and RPM packages, approximately 340MB

Disc 2 has NAS code, approximately 470MB

→Or, with systems that are DVD capable, a single DVD can be used

→Files are staged to CS harddrive and verified with checksums (install\_mgr -m upgrade)

**DUAL COMPACT DISC Upgrade:**

1) Insert CD #1 and # mount /media/cdrom

2) Change to CDROM directory and start install\_mgr script:

# cd /media/cdrom/EMC/nas

# ./install\_mgr –mode upgrade

*You can reboot blades during or after the upgrade.*

\* **DURING the upgrade :**

*Users will lose access to files served by each primary blade for 7 minutes during its reboot. The blades will not failover. The upgrade will pause and allow you to select the blades to be rebooted 46 minutes after you start the upgrade.*

\* **AFTER the upgrade :**

*Users will not lose file access during the upgrade. The blades will run the old software until you reboot them using Celerra Manager or the CLI server\_cpu command*

*When do you wish to reboot blades [ during or after ]? > during*

#### **Upgrade Annovance:**

Even after answering the above prompt to reboot blades during the NAS upgrade, the system will pause at the interactive message requesting the user to press the Enter key to reboot the Standby Server, then again for the Primary Servers.

*12:16 [ 45/77 ] Upgrade Standby Blades 7 minutes*

*Ready to reboot the Standby Data Mover (3). Press the ENTER key to begin...*

**3) After the Linux upgrade and reboot, run install\_mgr -m upgrade again, it runs a health check, and then creates the upgrade partitions, then instructs user to use the CD #2:**

*At this point you must swap discs to continue the NAS upgrade.*

*To swap discs, do the following:*

**1. Unmount and eject the current disc with the commands:**

```
cd  
umount /media/cdrom  
eject
```

**2. Remove the "Celerra Disc 1" media from the CD or DVD drive, and insert the media labeled "Celerra Disc 2" in its place.**

**3. Mount the CD-ROM with the command:**

```
mount -r /media/cdrom
```

**4. Resume the upgrade with the command:**

```
/media/cdrom/EMC/nas/install_mgr -m upgrade
```

*Result: Stopped*

**Note:** You must cd /, exit root, cd / again, then login as Root user, and issue # eject cdrom in order to get the CD ejected

**4) Mount CD #2 and start install\_mgr script:**

```
# mount /media/cdrom # cd /media/cdrom/EMC/nas # ./install_mgr -mode upgrade
```

**5) Shortly after, instructed to reboot again, and run the install\_mgr again before NAS began upgrading**

#### **NAS 6.0 to 6.0 DVD ISO LOOP MOUNT UPGRADE PROCEDURE:**

**Note:** For situations where the USM is not eligible or practical, or where onsite assistance is limited, the DVD ISO Loop mount method for upgrading within 6.0 is easy to perform

1. Create mountpoints and mount the local IDE partition /dev/hda10:

```
# mkdir /celerra/dvdiso  
# mkdir /celerra/upgrade  
# mount /dev/hda10 /celerra/dvdiso  
# cd /celerra/dvdiso
```

2. Upload the DVD.iso image to /celerra/dvdiso, loop mount the image to /celerra/upgrade, then proceed with the upgrade:

```
# mount -t iso9660 -r -o loop /celerra/dvdiso/6.0.36.2_image_DVD.iso /celerra/upgrade
```

```
# cd /celerra/upgrade/EMC/nas  
# ./install_mgr -mode upgrade
```

3. Reboot the Control Station at the following message:

*The Control Station must be rebooted now. This will take a few minutes.*

*To continue with the NAS upgrade, do the following:*

**1. When the Control Station finishes booting, login as "root".**

**2. Mount the ISO and make it your working directory:**

```
mount -r -o loop /celerra/dvdiso/6.0.36.2_image_DVD.iso /celerra/upgrade  
cd /celerra/upgrade/EMC/nas
```

**3. Continue the upgrade with the command: ./install\_mgr -m upgrade**

*Press Enter key to reboot:*

4. After approximately five minutes, log back into the Control Station. You may need to remount the /dev/hda10 partition to /celerra/dvdiso before loop remounting the iso image to /celerra/upgrade. Resume the the upgrade.

```
# mount /dev/hda10 /celerra/dvdiso  
# mount -t iso9660 -r -o loop /celerra/dvdiso/6.0.36.2_image_DVD.iso /celerra/upgrade
```

```
# cd /celerra/upgrade/EMC/nas
```

```
# ./install_mgr -m upgrade
```

## **UPGRADE TROUBLESHOOTING LOGS:**

**Note:** There are a variety of upgrade errors or issues that can occur during either the DVD/CD upgrade method, or USM method.  
→For CLI or GUI upgrades that are in progress, review the /var/tmp/upgrade.log to see what the last message details are.

=====SUMMARY=====

*Upgrade is still in progress. The Control Station is going to reboot now.*

*When the Control Station comes back up, rerun install\_mgr to resume the operation.*

*Status: Requested Reboot*

*Actual Time Spent: 9 minutes*

*Total Number of attempts: 2*

**Log File:** </var/tmp/upgrade.log>

→For PUHC health check logs, click on the GUI “Pre-Upgrade Log” link after the PUHC has run in order to see the detailed contents, or from CLI, go to /nas/log and view the contents of the check\_nas\_upgrade.xx.log file

## **UPGRADING SYSTEMS WITH DUAL CONTROL STATIONS:**

→Run the install\_mgr –mode upgrade script on CS0, as you would normally do

→At the end of the upgrade, the console screen directs you to use install\_mgr to upgrade CS1 (But see CPG for more verbose details)

=====SUMMARY=====

*Congratulations!! Upgrade for NAS software to release 6.0.36-4 succeeded.*

*To complete the upgrade, you must upgrade the standby Control Station.*

*Take the CD out of the primary control station, put it in the standby Control*

*Station, and start the upgrade on the standby.*

→CD2 contains the emcnassby code [/media/cdrom/EMC/nas/package/emcnassby\_6.0.36-4.i386.rpm]

→Use install\_mgr –mode upgrade to launch upgrade on CS1 (takes only about 10-15 minutes)

## **6.0 UPGRADES USING UNISPHERE SERVICE MANAGER (USM, formerly NST):**

→The CLARIION NST was renamed Unisphere Service Manager, and incorporates the Celerra Upgrade Tool (CUT) [Java-based]

→The Celerra Upgrade Tool (CUT) is a series of three wizards for upgrading qualified Celerra systems by Customers/Field/Partners

→The Download Software wizard determines the appropriate upgrade package and files from the dynamic Powerlink catalog (for 5.6 to 6.0, or 6.0 to 6.0 upgrades), and downloads to the Windows client’s EMC\Repository\Celerra\downloads\6.0.36.upg directory.

The .tar package is used for in and out-of-family upgrades, while ISO is also required for out-of-family upgrades: Release Notes; .tar package; .iso file

→“CUT” is now the preferred method for upgrading Celerra

→However, we are not officially endorsing that customers use CUT for Gateways, just Integrated systems, and, Systems with Dual Control Stations do not qualify for using CUT at all

→If you try to use the USM over an EMC network, the firewall will prevent the “Download Software” wizard from working (See emc250893). Basically, you need to download the necessary files (.iso & .tar) from Powerlink and prestage them on your laptop in a directory named “<nas\_version>.upg” in order for the USM CUT tool to run the ‘Pre-Upgrade Health Check’ and ‘Perform Software Update’ wizards

→Failed USM upgrades can be restarted from GUI, or via CLI install\_mgr—strongly recommended that Field folks keep a copy of the target upgrade media on CDs whenever performing upgrades

→USM can be launched as a stand-alone application, or from the Unisphere GUI (yet still requires that the client have the USM application installed)

→Must login as root user to use the CUT to upgrade from 5.6 to 6.0 (root is the default account and is greyed out and cannot be changed)

→Within 6.0, you can use root, nasadmin, or an imported array admin account, provided the Celerra is Joined and participating in the Storage Domain, and the imported accounts have been given the correct privileges [Group “fullnas(nasadmin)”, Client Access “Control Station shell allowed”]

→Celerra functionality is nested underneath the “Software” icon, with three Wizards that consist of the “Celerra Upgrade Tool”

### **Download Software wizard**

### **Pre-upgrade Health Check wizard**

### **Perform Software Update wizard**

→CUT content for Celerra is encapsulated in JAR files

→Supports NS20/40/80, NSX, NX4, NS-120/480/960, NS80G, NS-G2, NS-G8, VG2, VG8

→Localization for (8) languages (English; Spanish; Italian; German; Brazilian Portuguese; Chinese; Korean; Japanese)

→CUT cannot be used on dual Control Station upgrades

→CUT can only be used to upgrade from 5.6 to 6.0, or within the 6.0 family

→CUT cannot upgrade a system with Replicator V1 to 6.0

## **USM APPLICATION DOWNLOAD LOCATION ON POWERLINK:**

Support > Software Downloads and Licensing > Downloads T-Z > Unisphere Service Manager (USM) :

## Unisphere Service Manager (USM) Windows 1.0.0.1.0490

### USM USER ACCOUNTS, ROLES, and allowed ACTIONS:

| Local Celerra users  | Role                   | USM Actions  |
|--|------------------------|--|
| nasadmin*  | nasadmin               | Can login with Scope local and run all three CUT wizards           |
| root**   | root                   | Can login with Scope local and run all three CUT wizards           |
| Migrated Array users+  | imported_administrator | Can login with Scope global, but cannot run any of the CUT wizards |
| + Requires granting the user the Group Role of “fullnas(nasadmin)” and Client Access right of “Control Station shell allowed” in order to perform CUT upgrades, but only within the 6.0 family |                        |  |
| *The nasadmin account can only be used for upgrades within the 6.0 family, not from 5.6 to 6.0   |                        |  |
| **Root account is mandatory for CUT upgrades from 5.6 to 6.0   |                        |  |

\*The nasadmin account can only be used for upgrades within the 6.0 family, not from 5.6 to 6.0  
\*\*Root account is mandatory for CUT upgrades from 5.6 to 6.0

### USER ACCOUNT BEHAVIOR FOR USM & CLARIION DOMAIN MEMBERSHIP:

- For NAS 5.6 to 6.0 Upgrades, you must use the “root” user account and password (it is hard-coded for root user)
- For in-family 6.0 Upgrades, you can use either root or nasadmin “local” accounts
- Or, if you wish to use the migrated domain administrative accounts (provided the Celerra has properly ‘joined’ the CLARiiON Storage Domain), you must first grant the migrated user accounts the “fullnas(nasadmin)” Group role and “Control Station shell allowed” Client access rights in order to perform CUT upgrades, but only within the 6.0 family

**Note:** See emc245870

### Use LDAP Option:

→Option to use LDAP credential when logging into a system via USM

→Celerra needs to be configured with an LDAP directory server connection, as well as group mappings to the LDAP server

### USM UPGRADE METHODOLOGY FOR CELERRA SYSTEMS:

1. As a general rule, upgrade the Celerra NAS software first, after registering the Celerra system serial number on Powerlink
2. Login to the Celerra IP Address from USM, and run the wizards in the order that they appear on screen, to upgrade NAS
3. Login to the CLARiiON IP Address from USM, using an array admin account Scope Global, and run the wizards to upgrade FLARE or other software

### USM PUHC & UPGRADE LOGS:

→After the PUHC has run, there is a GUI link to “Pre-Upgrade Log” that displays the details of the PUHC checks

→The PUHC files are retained on the USM client and on the Control Station

### Control Station Logs during upgrades:

# ls -la /var/tmp/upgrade.log →Most useful log

/var/tmp/ →Many other temporary log files are kept here and may be of some value, depending on the issue

### Control Station Logs after upgrades:

/nas/var/log/check\_nas\_upgrade.100806-142946.log

/nas/log/upgrade.xxx.log

/var/tmp/.CUT/xxx

/var/tmp/cus.log

### USM Client Log location:

C:\EMC\repository\celerra\logs\HEALTH\_CHECK\_WIZARD\_10.241.168.30\_1281032963243\_log.txt

### USM CUT UPGRADE ISSUES:

#### Flare vs. NAS Compatibility:

→The USM does not prevent anyone from upgrading to Flare 30 first, before NAS is upgraded to a version that supports Flare 30  
→emc251289 is a solution for reinstalling or upgrading the nasStorageAPI version on a system, especially to address the issue where Flare 30 was applied prior to upgrading the NAS 5.6 system to a level that supports Flare 30. In any event, the action to take should be to escalate ALL such cases immediately to Support for resolution—the fix often requires EE assistance as well.

--NAS 5.6.50-2 provides nasStorageAPI-7.4-7, which supports Flare 30

--NAS 6.0.36-4 provides nasStorageAPI-8.2-12, which supports Flare 30

#### EMC FIREWALL ISSUE:

→EMC employees will encounter a firewall issue when using CUT and trying to automatically download the upgrade files from the Powerlink catalog

#### See emc250893 for more details, but the workaround is as follows:

1. Download the necessary .iso & .tar files to the c:\EMC\repository\celerra\downloads\6.0.36.upg directory
  - pkg\_6.0.36.4\_emcnas.tar
  - 6.0.36.4\_emcnas\_CD1.iso
2. Run Pre-upgrade Health Check and Perform Software Update wizards

#### Alternative Powerlink Access:

From external internet connection, connect to www.emc.com

Click link on upper righthand side of screen for “Customer/Partner Login”

Enter your NT username and FOB credentials

#### OTHER ISSUES:

Emc251622—Customer download issues—need to register Celerra system on Powerlink in order for Software Download wizard to work: [Support > Product and Diagnostic Tools > Celerra Tools > Register Your Celerra](#)

Emc251355—Patchwork primus with link to an Engineering paper on using the USM to upgrade Celerra systems, including Flare

Emc246820—New install FC system, CSA tries to Join Celerra to domain and fails

Emc252069—FC system upgrade to 6.0, upgrade screen says it joined Celerra to domain on completion page, but upgrade log shows that it failed

**Comment:** FC Systems are not designed to Join the Celerra to the Storage Domain for either Installs or Upgrades

Emc240345—Celerra will no longer be able to communicate or manage the array if the array admin password is changed—need to do the manual nas\_storage –modify –security step to sync new password to Celerra security files—currently no automatic mechanism in place to keep these sync’ed

Emc245870—CUT wizard errors related to User account privileges

## **6.0 UNISPHERE [Replaces Navisphere, NaviExpress, & Celerra Manager]:**

- The Unisphere management interface replaces Celerra Manager with NAS version 6.0, and Navisphere for Flare 30+ systems
- The recommended interface of choice for Unified Storage (Celerra and CLARiiON) is the Unisphere Client, a stand-alone application that runs on a Windows client (previously known as the Offarray UI), and offers the same interface as Unisphere
- Unlike the native Unisphere that launches via web browser, the Unisphere Client can be used to manage CLARiiON arrays running Flare 19-30, except AX4-5 systems where Navisphere Express is the licensed management software

**Note:** Upgrade NaviExpress to full Navisphere, then use the Unisphere Client to manage the AX4-5 system

→ The Unisphere Client can connect to pre-NAS 6.0 systems, but will only display system Alerts (Must still use Celerra Manager for 5.6 and below)

→ The advantages to using the Unisphere Client would be to provide a consistent version and interface for all systems that are members of the Storage Domain, as well as providing better performance than the Browser-based Unisphere interface

→ A major goal for Unisphere is to be able to provide a single common interface to manage NAS and Storage, as well as to lay the groundwork for unifying CLARiiON block and Celerra file platforms together as a single product

### **UNISPHERE Layout:**

→ Unisphere GUI provides a common framework. Both Navisphere and Celerra Manager components plug into the framework with essentially the same underlying capabilities as they had before.

→ The new GUI emphasis is for task-based navigation, with tables and task lists, vs. traditional object-oriented representations.

→ For Tables, you can easily export almost anything into .csv Excel files.

→ Users can customize the Dashboard display, sort & filter values in Tables, hide and unhide columns in the various screens, etc.

Unisphere Dashboard is a quadrant of screens that displays Systems by Severity, Alerts by Severity, CLARiiON capacity, and Celerra capacity.

→ Through the ‘single sign-on’ concept for Celerra, and the Domain Master role with Flare 30, a User can manage multiple arrays and Celerras via a single log-on session (once setup to do so).

→ Unisphere can scale to 100,000 objects and manage 50 systems

→ Navi plug-ins are (CIM/XML) & Celerra plug-ins (XML/HTML)

→ ECUE Compliant (EMC Common User Experience standard for management applications)

### **Top level Navigation Bar:**

→ Top level Navigation defaults to an “All Systems” view with Dashboard, System List, Domains, Alerts, Support

→ Celerra System presents a System, Storage, Sharing, Replicas, Monitoring, Settings, and Support view

→ Clariion System presents a Storage, Replicas, Monitoring, Hosts, Settings, Support view

### **Automatic Refresh:**

→ Occurs when an alert is added or removed

→ Occurs when a storage system is added or removed

→ Occurs when a user completes an operation in a dialogue box or wizard

### **UNISPHERE LICENSING & POWERLINK DOWNLOAD LOCATION:**

→ Older NS systems that are sold with NAS 6.0 will sell with either a file-only “Unisphere for File” license (NS-480), or the “Unisphere for Unified” license, which replaces the Native Block Option and Navisphere Manager licenses (e.g., NS-480FC)

→ An NX4 with Fibre Channel option can be managed with Unisphere on the NAS side and NaviExpress on the array side, or upgraded to a “Unisphere for Unified” license, which allows the customer to use Unisphere for both the Celerra and CLARiiON

### **UNISPHERE CLIENT:**

→ In the CLARiiON realm, formerly known as the Navisphere OffArray UI

→ This tool is the preferred management application for mixed Celerra/Clariion domains

→ Benefits of this application are; better performance; more current/consistent interface that can manage Flare 19 – 30 systems, as well as Celerra 6.0 systems. For example, you could not use a Browser to launch Unisphere for a Flare 29 system, but if using the Client, the management application is based on the Unisphere framework.

→ For FLARE 30 arrays, Unisphere will enforce two levels of certificate validation

Low: All certificates pass, essentially bypassing certificate validation

Medium: All certificates must be validated

### **Unisphere Client Powerlink Download:**

## **EMC Unisphere Client (Windows) 1.0.0.1.0492**

### **UNISPHHERE STORAGE DOMAIN CONCEPTS:**

- The new terminology today, and with future releases, will be Unified Management for Block (CLARiiON) and File (Celerra) systems
- Outside of the consolidation of the CLARiiON and Celerra GUIs into Unisphere, one of the more significant changes is that the Celerra will now be able to join and participate as a member system in the CLARiiON Storage domain, or ‘Unified Storage Domain’
- When discussing “domains” in the context of Celerra and CLARiiON, we are really talking about CLARiiON Storage Domains, where array global admin user accounts are shared among Celerra and CLARiiON systems, and used for the “single sign-on” concept in Unisphere whereby you can log into a single system, using a global array admin user, and then manage all the ‘joined’ systems from the same account as a “multi-domain” environment
- The concept of “joining” a CLARiiON domain is new with NAS 6.0
- After “joining” the CLARiiON domain, the domain admin accounts are copied to the Celerra with Group role membership in the “nasadmin(operator)” group, which is equivalent to Read-Only privileges on the Celerra, and Client Access privileges for “Unisphere allowed” and “API access allowed”
- In order for the migrated array admin accounts to be able to manage the Celerra, you must login to the Celerra as root scope local, and assign the Group membership for “fullnas(nasadmin)” and the Client Access privilege for “Control Station shell allowed”

### **Unisphere > Settings >User Management > Users**

**Note:** Shows the list of User accounts on the Celerra

#### **Single Sign-on Capability:**

- When logging in using default Global scope, Unisphere will attempt to log-in to all systems in the System and Domain list (works because Clarion domain users are copied to the Control Station on installs and upgrades)
- Celerra Only model would require the same username and password on all Celerras
- Applies to the Unisphere Service Manager (USM) as well, but in order to be able to use the USM CUT tool wizards with the migrated global array admin accounts, you would need to set the proper permissions for the account on the Celerra **fullnas(nasadmin)**

#### **Control Station shell allowed**

#### **Setting Privileges on Migrated Array accounts:**

- a) Log into the Celerra system as Root user, Scope local
- b) Go to Settings > User Management > Users, assign the desired privileges for the migrated domain accounts to allow for “single sign-on” management of both the Celerra and CLARiiON system(s).
- c) Assign the “fullnas(nasadmin)” privilege, and check the box for “Control Station shell allowed”

**Note:** Migrated accounts are granted the Operator privilege, and the access rights for “Unisphere allowed”, “API access allowed”  
→Domain membership can be seen in the “System List” or Domains > Local > Systems list

#### **What's a Unified Storage Domain?**

- A collection of systems (CLARiiON and Celerra) that are grouped together for ease of monitoring and management
- When properly configured, a single administrative global array user account can be used for “single sign-on” into the storage domain, for the purpose of administering all participating member systems
- You can have a single local domain with multiple systems, or multiple domains with multiple systems

#### **TYPES OF UNISPHERE DOMAINS:**

##### **Celerra Only List Unisphere Domain**

- Unisphere login Scope is Local (e.g., NX4 integrated where array is managed using NaviExpress
- Multiple Celerras can be added to the “list”. If the user account and password exists on all Celerra systems, when using Scope Global, then the credentials will be authenticated for every Celerra on the list, and all the Celerras can be managed from one session.
- For any Celerras on the list that cannot authenticate the user credentials, the system will appear in the System List with a “Not Logged In” status, and the user can rightclick and login at that time
- No CLARiiONS participate in the ‘Celerra Only List’

##### **CLARiiON Only Unisphere Domain**

- In order to use single sign-on for management of all joined systems, the domain master needs to be Flare 30
- Domain with only CLARiiON arrays as member systems
- Unisphere Client can be used to manage Flare 19-30 CX4 systems
- No Celerras
- Unisphere login Scope is Global

##### **Mixed Domains Unisphere Domain**

- Both Celerra 6.0 systems and CLARiiON R19-30 can be participants in the storage domain, as long as one of the arrays is running Flare 30 as Domain Master
- This is the optimum structure, where ‘single sign-on’ comes into play
- Unisphere login Scope is Global

#### **Local Domains vs. Multiple Domains**

--CLARiiONs can be initialized with a default ‘local’ domain, where other Arrays and Celerras can be added as participants of the local domain, with a single Domain Master

--You can also have a group of domains in a Multi-Domain environment, each with their own CLARiiON and Celerra systems, either with or without single sign-on user accounts (each domain can have its own separate admin accounts, or can share across the domains)

--Initially, you would Add a new domain to an existing Local Domain

## **CELERRA JOIN BEHAVIOR DURING INSTALLS/UPGRADES FOR 6.0:**

--Gateway systems will not auto Join the Storage domain during installs

--Gateway systems will not auto Join the Storage domain during upgrades

--Integrateds (i.e., non-shared array) will auto Join the Storage domain during installs (Using CSA)

--Integrateds (i.e., non-shared array) will auto Join the Storage domain during upgrades

--FC Integrateds (i.e., with shared array) will not auto Join the Storage domain during installs

--FC Integrateds (i.e., with shared array) will not auto Join the Storage domain during upgrades

**Note:** A future maintenance release may change the behavior of FC Integrateds to that during Installs and Upgrades they ‘Join’ the storage domain similarly to the Integrateds

## **CELERRA MEMBERSHIP IN STORAGE DOMAINS:**

→With NAS 6.0, Celerra can participate as a member system in the Storage Domain

**emc252609**

--NAS 6.0 Celerras can be joined to the host array storage domain as a member system if the array is running the latest Flare 26 patch, or a higher version of flare (at least testing seems to support this).

--NAS 6.0 Celerras cannot join the host storage domain as a member system for an array running less than Flare 26/031

--All NAS 6.0 Celerras can, however, be joined to any Flare 30 storage domain as an alternative, and ideally, this is preferred in order to run the single sign-on Unisphere management.

→This allows for enterprises to use single sign-on to administer multiple systems from a single logon session

→There are certain rules involved with the way the Celerra joins the domain

→When the Celerra Joins a domain, the array admin accounts are copied over to the Celerra with certain Group role membership and Access rights, which varies depending on the Celerra system

→Both Unisphere and Unisphere Service Manager use the array administrative accounts

## **ATTRIBUTES OF MIGRATED ACCOUNTS ON CELERRA:**

→By default, Gateway and FC systems will migrate the array accounts with the Group roles of “imported\_administrator” and “nasadmin(operator), and Client Access rights for API & Unisphere allowed, which equals Read Only access on the Celerra

→Integrated systems without a shared backend, and a “nasadmin” array account, will be copied to the Celerra with equivalent privileges and membership as for the native Celerra nasadmin user [Fullnas & Operator], except for the Client Access right of “Control Station access allowed”, which must be manually added

→To change migrated User privileges, login to the Celerra as Root, Scope Local, and assign the migrated account the Group privilege for “fullnas(nasadmin)”, and the Client Access privilege “Control Station shell allowed”

**Note:** Without the CS shell allowed privilege, the User cannot launch CLI fromUnisphere, or run any of the USM CUT wizards

## **MANUALLY JOINING CELERRA TO THE STORAGE DOMAIN:**

**CLI Join Method:** Not supported for NAS 6.0

# /nas/sbin/navicli -h 10.241.168.179 domain -add 10.241.168.80

WARNING: You are about to add following node(s) to the domain.

10.241.168.80

Proceed? (y/n) y

Domain add operation failed.

Error 10.241.168.80 is not reachable.

**GUI Join Method:**

1) Log into the Array using either Unisphere or the Unisphere Client (if the array is pre-Flare 30) and global admin account

2) From the top of the Unisphere screen: Domains > Add/Remove Systems > Add: Enter IP Address for Celerra Control Station, then root Username and Password, Scope Global

3) Verify:

- a) Click on the ‘Local’ object in the Domains section—the Celerra should appear as a registered system
- b) Use Navicli to verify:

# /nas/sbin/navicli -h 10.241.168.179 domain -list

Node: vg8

IP Address: 10.241.168.80 →Celerra Control Station

Name: vg8

Port: 80

Secure Port: 443

-----abridged-----

Node: APM00083201184

IP Address: 10.241.168.180  
Name: SPB  
Port: 80  
Secure Port: 443  
IP Address: 10.241.168.179 (Master)  
Name: SPA  
Port: 80  
Secure Port: 443

f) Close Unisphere and log into the Celerra as ‘root’ with ‘Scope’ ‘Local’

**Settings > User Management > Users:**

**Note:** Should be able to see the migrated CLARiiON domain accounts

g) Assign the desired privileges to the migrated accounts

→By default, all migrated accounts are assigned membership in the Group Role ‘Operator(nasadmin)’, with Client Access privileges for ‘Unisphere allowed’ and ‘API access allowed’ enabled, but without “Control Station shell allowed” enabled

**Auto Joining:**

→Part of the upgrade tasks are to check and join the CLARiiON Storage domain for qualified models

```
=====  
21:03 [ 79/81 ] Add celerra to domain           30 seconds  
Celerra was added to the domain.  
=====
```

**Assigning Rights equivalent to the native nasadmin local account on the Celerra:**

To be able to manage the Celerra, check the box for the “fullnas(nasadmin)” privilege and the access privilege “Control Station shell allowed” to grant privileges equivalent to the local Celerra nasadmin user.

**FILES CREATED WHEN CELERRA JOINS THE STORAGE DOMAIN:**

# ls -la /nas/http/domain

```
-rw-r--r-- 1 apache apache 69 Apr 22 15:13 domain_list  
-rw-r--r-- 1 apache apache 72 Apr 22 15:13 domain_master  
-rw-r--r-- 1 apache apache 65 Apr 22 15:13 domain_users
```

# cat domain\_list

```
DOMAIN_ADDRESS_LIST=<10.241.168.184>,<10.241.168.185!10.241.168.186>,
```

# cat domain\_master

```
MASTERIP=10.241.168.185
```

```
DOMAINID={87B90648-1BEE-41F5-8897-164BDDE71AC1}
```

```
PORt=443
```

# cat domain\_users

```
nasadmin:imported_administrator:580ABE52065C603C7BCBB4E3F98F8BCA
```

**REMOVING CELERRA FROM A STORAGE DOMAIN (Unjoin)—emc247813:**

→If you attempt to remove the Celerra from the Storage Domain while logged in as the domain admin user, but the user has not been granted the “root(root)” Group Membership role on the Celerra to be removed, the operation will fail with the following popup message:

**“User not privileged**

**X The current user does not have the privilege to perform domain operations on the Celerra.”**

→If the above situation occurs, you must perform the following steps to remove the Celerra from the domain:

**GUI Option:**

1. Using Unisphere, log into the Celerra with user “root” and Scope local
2. Settings > User Management > Users, select the domain user, properties, and assign the Global(Role)Membership “root(root)”
3. Exit Unisphere, reopen Unisphere, log into the SP with the modified domain user account, Scope Global
4. Remove the Celerra from the domain using Domains > Add/Remove Systems > vg8 > Remove, and accept the popup Warning message that says, “Warning: Removing a system from the domain will erase/invalidate all domain user accounts on the system...”

→Alternatively, you can use the CLI to remove the Celerra from the domain

**CLI Option:**

# /nas/sbin/navicli -h 10.241.168.179 domain -list

```
Node:          vg8  
IP Address:   10.241.168.80  
Name:          vg8  
Port:          80  
Secure Port:   443
```

# /nas/sbin/navicli -h 10.241.168.179 domain -remove 10.241.168.80

WARNING: You are about to remove following node from the domain: 10.241.168.80

Proceed? (y/n) y

#### Verify from CLI or GUI:

- 1) **# /nas/sbin/navicli -h 10.241.168.179 domain -list** → Celerra no longer listed as a member
- 2) All domain files have been removed from the /nas/http/domain directory  
→ From Unisphere (logged into Celerra scope local), all the formerly migrated domain array accounts show a ‘State’ of ‘Defunct’, with an ‘Account Type’ called ‘Defunct Storage Domain’

#### JOIN ISSUES:

- If the Celerra is already listed, the Join function will not work from the GUI
  - # navicli domain -list** → If the CS is already listed, make sure it is unjoined before attempting to rejoin to the Storage Domain
  - For CSA join failures, check startup\_wiz.log
- Note:** FC Enabled Celerras do not Join automatically, by design, though CSA and Installs do claim to try and fail
- Use the correct User account and Scope: Make sure you have logged into the Celerra using Root with Scope “Global” to perform the Join [Caution--Scope Local Join appears to work, but gives no error]

#### REJOINING CELERRA TO SAME STORAGE DOMAIN:

- Follow same Join steps as outlined above
- The domain user accounts are migrated to the Celerra, but do not override the previous accounts with the same name. Rather, the migrated accounts are given a ‘user1’, ‘user2’ local name, but maps to the original migrated name

#### Example:

secured1 → Mapped User Name “secured”

#### KNOWN ISSUES/LIMITATIONS:

- If the array user passwords are changed, the changes are not automatically propagated to the Celerra. Users will not be able to login and manage the Celerra from Unisphere, and the Celerra communication with the array will become broken. For example, new disks or storageAPI updates cannot be completed. The solution is to make sure that the following CLI command is run on the Celerra anytime that the Array user account passwords are changed:

#### **# nas\_storage -modify -security**

**Note:** Enter username and new password to update the Celerra storageAPI database

#### LOGGING INTO CONTROL STATION CLI WITH MIGRATED ACCOUNTS:

- If you want to log into the Celerra using CLI (using SSH, etc.), there is some trickiness involved, depending on whether the NAS system was ‘joined’ to the Storage Domain as an Integrated or as a Gateway.
- Basically, the first rule to know is that migrated accounts are never assigned the ‘Client Access’ privilege for “Control Station shell allowed”, so at a minimum, the Celerra admin must first log into the Celerra Unisphere as Root Scope Local, and grant the “Control Station shell allowed” privilege.
- Once the above step is completed, you would generally use the following syntax when logging in via Putty/CLI/SSH:

**login as: migrated\_username@storageDomain**

**migrated\_username@storageDomain@10.241.168.80's password:**

- The exception is that for any migrated username that already exists in the local Celerra database (e.g., nasadmin), the migrated name is given a new “Local Name”, such as “nasadmin1” in the case of the array nasadmin account migrated for a native Celerra Integrated system, and so you would need to use the “Mapped User Name” when logging in via CLI, not the “Local Name”

#### CLI LOGIN EXAMPLES:

##### For Celerra Integrateds where array admin name is “nasadmin”:

1. Log into the Celerra via Unisphere, using Root user, Scope Local
2. Click on the Celerra System, navigate to “Settings > User Management”, highlight the array account “nasadmin1”, select “Properties”, check the box under “Client Access” for “Control Station shell allowed”, then apply
3. Open an SSH session to the Control Station, logon with the following user account syntax and appropriate password:  
**login as: nasadmin@storageDomain**

**nasadmin@storageDomain@10.241.168.184's password:**

[nasadmin1@sleet-480 ~]\$

##### For all other Celerras (e.g., Gateway), where admin names are different:

1. Log into the Celerra via Unisphere, using the Root user , Scope Local
2. Click on the Celerra System, navigate to “Settings > User Management”, highlight the array account “secured”, select “Properties”, check the box under “Client Access” for “Control Station shell allowed”, then apply
3. Open an SSH session to the Control Station, logon with the following user account syntax and appropriate password:  
**login as: secured@storageDomain**

**secured@storageDomain@10.241.168.184's password:**

[secured@vg8~]\$

#### UNIFIED STORAGE DOMAIN ISSUES?

- Though we claim a Unified Storage convergence story, with Unisphere as the single management interface for Block (CLARiiON) and File (Celerra), there are many nuances involved with setting up Single Sign-on capabilities, Joining/Unjoining Celerras, etc.

→ If logging into Celerra via Unisphere with the local Root account, you login as Root Scope Local, but in order to “Join” a Celerra to the storage domain, you need to login as Root Scope Global during the join process

- Single sign-on Unisphere management of both CLARiiON and Celerra systems first requires that the Array user accounts are properly configured on the Celerra (after making sure that the Celerra is first ‘joined’ to the domain)
- Similarly, User accounts must also be properly privileged in order to run any of the Celerra Upgrade Tool wizards in the USM for NAS Upgrade activities

→Array password changes are not integrated with the Celerra. i.e., password changes to the domain admin accounts are not automatically propagated to the Celerra. You must actively log into the Celerra CLI and run # nas\_storage –modify –security –username –password, and use the new password to synchronize the passwords in the Celerra databases (/etc/.clar\_security)

## **6.0 CS & DART IPv6 SUPPORT:**

→NAS 6.0 provides support to all Celerra features that require IPv6 addressing that were not previously supported

→XMLAPI & MMC snap-ins support IPv6

→IPv6 allows for 128 bit addressing, more efficient routing, elimination of broadcasts, more security, better autoconfiguration

→Minimum MTU of 1280 bytes, with header size 40 bytes

→DART maintains default routes based on router advertisements

→Adds support for CAVA, CIFS, Usermapper, iSCSI, Kerberos, ReplicatorV2, NFSv4, NTP, XLT, LDAP-based Directory client configuration, & pNFS support

**Note:** GUI support not included in this release

→Adds support for SNMPv2 & SNMPv3

→Supports IPv6 name resolution only, for LDAP (not IPv6 addresses)

→Supports Local File name resolution using IPv6 addresses

→Supports DNS AAA address to name and name to address resolution & operation of IPv6 addresses

### **Forward Lookup Zones**

-pgace110.emc.com

**NS10-blade    IPv6 Host (AAAA)    2620:0000:0000:0000:0000:abcd:0059**

→Supports NIS name resolution using names or IP addresses using IPv6

→Supports Etherchannel, LACP, IP Reflect, FSN, and Jumbo frames

→Supports DNS/DDNS; NIS; LDAP; NFSv2, v3, v4; FTP; TFTP; iSCSI; SNMPv1, v2, v3; Ping6; CIFS; CDMS/DHSM; CAVA; Replication V2

### **IPv6 Limitations:**

→Cannot use Unisphere to configure or modify IPv6 addressing, only the CLI interface [Unisphere can read/display IPv6 values]

→Cannot use the CSA to Install or Configure using IPv6 networks

→IPRepV2 works with IPv6, but should be either all IPv6 or all IPv4

→Cannot use Host names in lieu of IPv6 address

→Cannot use WINS on IPv6 networks

→Cannot use External Usermapper on IPv6 networks

→Cannot use MPFS on IPv6 networks

→Does not support CLARiiON iSCSI or MirrorView

→Cannot use Link Local addresses on anything other than core IPv6 protocols (e.g., ping6, ftp)

→Control Station can operate in a dual-stacked IPv4 & IPv6 environment, but not a pure IPv6 environment

→DART can use IPv6 exclusively (except that CS to DM communications will still be IPv4 internally)

→Configure IPv6 addresses on Control Station using nas\_cs

→Configure IPv6 aliases using nas\_config

### **Configuring IPv6 on Control Station:**

**# /nas/bin/nas\_cs**

```
nas_cs
  -info [ -timezones ]
  | -set [ -hostname <host_name> ]
    [ -ip4address <ipv4_address> ]
    [ -ip4netmask <ipv4_netmask> ]
    [ -ip4gateway <ipv4_gateway> ]
    [ -ip6address <ipv6_address[/prefix_length]> ]
    [ -ip6gateway <ipv6_gateway> ]
    [ -dns_domain <dns_domain_name> ]
    [ -search_domains <domain_name>[,...] ]
    [ -dns_servers <ip_addr>[,...] ]
    [ -session_monitor_timeout <days> ]
    [ -session_idle_timeout <minutes> ]
    [ -time <yyyymmddhhmm[ss]> ]
    [ -timezone <time_zone_str> ]
    [ -ntp_servers <ip_addr>[,...] ]
  | -clear [ -ip4gateway ]
```

```
[ -ip6address ]
[ -ip6gateway ]
[ -dns ]
[ -search_domains ]
[ -session_monitor_timeout ]
[ -session_idle_timeout ]
[ -ntp_servers ]
```

| -reboot

**CREATING, CHECKING, DELETING IPv6 ADDRESS ON CONTROL STATION:**

```
# nas_cs -set -ip6address 2620::abcd:109
```

```
# nas_cs -info
```

IPv6 Address = 2620::abcd:109/64

```
# nas_cs -clear -ip6address
```

**CONFIGURING IPv6 NTP ADDRESS ON CONTROL STATION:**

```
# nas_cs -set -ntp_servers 2620::abcd:23
```

OK

**OUTPUTTING CS CONFIGURATION INFO:**

```
# nas_cs -info
```

```
Host Name = sleet-480
Version = 6.0.34-0
Location = system:NS-480:APM000738018380000|controlStation::0
Status = OK
Standby Location =
Standby Status =
IPv4 Address = 10.241.168.184
IPv4 Netmask = 255.255.255.0
IPv4 Gateway = 10.241.168.128
IPv6 Address = 2620::abcd:109/64
IPv6 Gateway =
DNS Domain =
DNS Domain search order = localdomain
DNS Servers =
Session Monitor Timeout = 0 Days
Session Idle Timeout = 0 Minutes
NTP Servers = 2620::abcd:23.
System Time = Wed Jul 07 14:41:15 EDT 2010
```

```
# nas_cs -reboot
```

Info 26575634438: Reboot request was successfully sent to Control Station.

**CONFIGURING IPv6 ON DATA MOVERS:**

```
# server_ifconfig
```

```
-all [ -ip4 | -ip6 ]
| -create -Device <device_name> -name <if_name>
-protocol { IP <ipv4_addr> <ipmask> <ipbroadcast>
| IP6 <ipv6_addr>[/PrefixLength] }
```

```
# server_ifconfig server_2 -create -name cge0_ip6 -Device cge0 -p IP6 2620::abcd:60 [Global Unicast address]
```

server\_2 : done

```
# server_ifconfig server_2 -a -ip6
```

server\_2 :

```
cge0_ip6 protocol=IP6 device=cge0      →Global Unicast address
inet=2620::abcd:60 prefix=64
```

UP, Ethernet, mtu=1500, vlan=0, macaddr=0:60:16:1d:22:c4

```
cge0_0000_ll protocol=IP6 device=cge0   →Link Local auto. created when Unicast addr. created [_ll and fe80]
inet=fe80::260:16ff:fe1d:22c4 prefix=64
```

UP, Ethernet, mtu=1500, vlan=0, macaddr=0:60:16:1d:22:c4

```
loop6 protocol=IP6 device=loop
inet:::1 prefix=128
```

UP, Loopback, mtu=32768, vlan=0, macaddr=0:0:0:0:0 netname=localhost

```
# server_ping6 server_2 2620::abcd:60
```

server\_2 : 2620::abcd:60 is alive, time= 0 ms

**NFS Export with both IPv4 & IPv6 addresses, netgroups, and hostnames:**

**\$ server\_export server\_2 -Protocol nfs -o access=128.221.15.35:[2080:0:0:0:8:800:200C:417A]h1:netgrp1 /fs1**

**Note:** Use : colon between successive IPv6 address entries when exporting. Link local addresses are not supported for Exports.

### **NFS EXPORT & MOUNTING USING IPv6:**

**# server\_export server\_2 -P nfs -o access=[2620::abcd:138],root=[2620::abcd:138] /fs1**

**# mount \[2620::abcd:59\]:/fs1 /fs1** →Remote mounting from Sun Solaris

### **Managing IPv6 Neighbor Discovery and Routing Tables:**

**# server\_ip server\_2 -route -list**

server\_2 :

| Destination | Gateway | Interface | Expires (secs) |
|-------------|---------|-----------|----------------|
| 2620::/64   |         | cge-1-1   | 0              |

→Use the –route option to List, Create, or Delete entries in the IPv6 route table, which are updated automatically by the Neighbor Discovery protocol

**# server\_ip server\_2 -neighbor -list**

server\_2 :

| Address       | Link Layer Address | Interface | Type | State |
|---------------|--------------------|-----------|------|-------|
| 2620::abcd:10 | 0:50:56:a4:60:6c   | cge-1-1   | HOST | STALE |

→Use the –neighbor option to List, Create, or Delete neighbor cache entries for troubleshooting

### **PORTS LISTENING & ESTABLISHED ON SERVER:**

**# server\_netstat server\_2 -A inet6 -a | -A inet -a (IPv4)**

### **6.0 CELERRA ADMIN ROLES FOR CLI USERS:**

→Administrative Roles (Role-based privileges) were implemented with NAS 5.6, for GUI Users only

→Roles define what a user can do with a Celerra object, and users are always associated with a primary group, and each group has a defined role

→Role privileges are defined as Read, Modify, or Full Control

→By default, all Roles are allowed to view objects, or to have Read privileges

→NAS 6.0 adds Admin Roles capability for CLI Users, using predefined or custom roles

→An Authentication Library is used for the CLI that will allow certain Role-based actions to occur, depending on a User's Group membership role, and the action being taken [Action + Object + Role=result] using nas\_cmd commands

→This is a passive feature; admin roles cannot be managed from the CLI. To test, you would need to setup Admin Role accounts from the GUI, then test via CLI to see what actions are allowed or denied. Use the /nas/tools/authorize\_client utility to check and test user privileges.

→Root user account UID=0 is not subject to the Auth Library check and is skipped when using the CLI, however, though root has all access and control, it does go through a authorization check if using Unisphere.

→An important concept for Roles is that a single Group can only be mapped to one role, but many groups can map to the same role

### **Database Files:**

**/nas/site/role\_user** →Admin Role users and roles

**/nas/site/group\_db** →Admin Role groups and roles

**/nas/site/role** →System defined Users (root & nasadmin) & Groups for GUI

### **Deciphering a User's Roles from the database file:**

**# cat /nas/site/role\_user**

```
:nasadmin:201:3,2,
:nasadmin1:500:3,2,9,
```

→Where field 1 is the username, field 2 = UID, & field 3 = role(s). So, “nasadmin1” is the name of the account, has a UID of 500, and group roles for 3 (operator), 2 (nasadmin), & 9 (imported\_administrator), respectively, as defined in the /nas/site/role file

**# cat /nas/site/role** (output abbreviated and edited)

```
2:nasadmin:system-defined:
```

```
3:operator:system-defined:
```

```
9:imported_administrator:storage-defined:
```

### **Creating and configuring Admin Roles:**

1. Create User with Network Admin role using GUI

Unisphere > Settings > User Management, with Account Type local, a password & expiration, and a Primary Group role, making sure to check the Client Access boxes for ‘Unisphere Manager allowed’ and ‘Control Station shell allowed’

2. Login to CLI as the new network admin User & test privileges

**\$ nas\_fs -name net2 -create size=1000M pool=clar\_r5\_performance -o slice=y**

Error 13422428165: The user is not authorized to perform the specified operation.

3. Or use authorize\_client utility to test access privileges for a given User

**\$ ./authorize\_client networker filesystem list**

**\$ echo \$? 1** --> Return code of 1 indicates that the action is allowed for the user

**\$ ./authorize\_client networker filesystem modify**

**\$ echo \$? 2** --> Return code of 2 indicates that the action is not allowed for the user

**Verifying existence of new User:**

# cat /etc/passwd | grep networker

networker:x:500:501::/home/networker:/bin/bash

**Example of user ‘networker’ with Network Admin Role:**

→User logs into Control Station using SSH

→User tries to create a file system, but because of Roles, is not authorized to do so and the operation fails

\$ nas\_fs -name net2 -create size=1000M pool=clar\_r5\_performance -o slice=

Error 13422428165: The user is not authorized to perform the specified operation.

**Using the authorize\_client script to Test & Troubleshoot CLI Roles:**

/nas/tools/authorize\_client

Usage: authorize\_client <user\_name> <object\_name> <action\_name>

**Note:** See emc236631 for a link to the complete authorization library table that shows the CLI command, Object name, Action allowed, and Privilege required

**Using Return Codes to test <user\_name> against <object\_name> based on <action\_name>:**

\$ ./authorize\_client networker filesystem list

\$ echo \$?

1 --> Return code of 1 indicates that the action is allowed for the user indicated

\$ ./authorize\_client networker filesystem modify

\$ echo \$?

2 --> Return code of 2 indicates that the action is not allowed for the user indicated

→Above examples show that user ‘networker’ can list file systems but not modify them

**Possible Return Codes using authorize\_client and the \$ echo \$? query:**

AD\_ALLOW = 1,

AD\_DISALLOW = 2,

AD\_USER\_NOT\_FOUND = 3,

AD\_ERROR\_CANNOT\_OPEN\_FILE = 4,

AD\_ERROR\_CANNOT\_ALLOCATE\_MEMORY = 5,

AD\_ACQUIRELOCK\_FAILED=6,

AD\_RELEASE\_FAILED=7,

**References:**

→Appendix A of the Celerra Security Configuration Guide lists the Commands, Object Actions, Category, and predefined roles for all the Objects where Admin CLI roles apply. There are many commands that do not fall under the role-based access umbrella (e.g., nas\_halt, server\_log, nas\_cel, server\_mpfs, etc.), and would require the traditional nasadmin or root user accounts. Similarly, there are a number of commands that all Roles have the right to execute (server\_ping, nas\_inventory, server\_uptime, etc)

→emc236631 contains a list of Commands, Objects, Actions, & privilege for testing User privileges using /nas/tools/authorize\_client

**6.0 CONTROL STATION SECURITY AUDITING:**

→New auditing feature added to Control Station with RHEL 5.0 kernel

→Based on Linux auditd daemon, records, events, and logs

→Basic purpose is to audit security-relevant events

→Auditing is enabled by default on new installations, keeping track of commands run, changes to system files, changes made to auditing, changes to user accounts/roles, and user authentication to the system

→Systems upgraded to 6.0, however, are not enabled by default

→NAS services are not required to run the audit daemon, hence new 6.0 /celerra/audit partition on the local Control Station IDE drive

→Auditing records are sync’ed to /nbsnas/var/auditing/cs0 or cs1 every 180 seconds based on nas\_mcd.cfg configuration

→For CS0 replacement, code will restore the audit files. CS1 replacement would require manual restore of audit files

/sbin/ausearch →Reading the audit trail

/sbin/aureport →Summary reports of audit logs

→One of the associated goals is to audit access to Data Mover root file systems over the internal NFS network using a Mount Monitor process

# ps -ef | grep mt\_mon

root 9656 9654 0 Aug06 ? 00:04:08 /nasmed/sbin/mt\_mon

**Note:** Audit Daemon monitors rules and passes records to Mount Monitor

→Can configure CS to shutdown if auditing is disabled

**New Local Partition:**

/celerra/audit

**Note:** Purposes of creating an internal IDE partition on the CS is to store the logs locally, not depend on NAS Services, and have ability to log on either CS0 or CS1

**Control Station Files:**

/etc/audit/auditd.conf →Auditd behavior, logs, log rotation rules, up to 7 logs

/etc/audit/audit.rules →Defines what auditd will audit

**/celerra/audit/audit.log, audit.log.1, audit.log.2, etc.** →Where audit logs stored

**/sbin/auditd** →Audit Daemon executable, startup based on auditd.conf

**/sbin/auditctl** →Audit subsystem control executable

#### **Audit Record Types:**

SYSCALL—system calls

PATH—file being accessed

CWD—current working directory of process

USER\_XXX—events associated with a user

FS\_WATCH—file system object access when explicit watch is enabled

#### **Configuring Auditing on Upgraded Systems:**

1. Edit auditd.conf or audit.rules if custom auditing desired
2. Configure the auditd daemon service to run on system startup

**# /sbin/chkconfig --levels 2345 auditd on | off**

**# /sbin/chkconfig --list auditd**

```
auditd      0:off 1:off 2:on 3:on 4:on 5:on 6:off
```

3. Start the audit daemon service

**# /sbin/service auditd start**

```
Starting auditd: [ OK ]
```

#### **Difsync Backups from nas\_mcd.cfg:**

→backs up /etc/audit/auditd.conf & audit.rules, and /celerra/audit/audit.log to /nas/var/auditing/.backup, which is a link to /nas/var/auditing/cs0, every 180 seconds

#### **Backup location for audit.log and auditd.conf:**

#### **/nas/var/auditing**

```
lrwxrwxrwx 1 root root 21 Aug 3 20:34 .backup -> /nas/var/auditing/cs0  
drwxr-xr-x 2 root root 1024 Sep 9 00:20 cs0 →auditd.conf, audit.log, & audit.rules  
drwxr-xr-x 2 root root 1024 Aug 3 15:31 cs1
```

→During CS restore, auditing logs are restored to /nas/var/auditing/cs0

#### **Location of Audit Logs:**

**# ls -la /celerra/audit**

```
-rw----- 1 root root 7802979 Sep 8 20:00 audit.log  
-r----- 1 root root 15728749 Sep 8 18:30 audit.log.1 →Logs rotate after every 15MB of information recorded  
-r----- 1 root root 15728677 Sep 8 15:40 audit.log.2  
-r----- 1 root root 15728733 Sep 8 12:40 audit.log.3  
-r----- 1 root root 15728781 Sep 8 09:50 audit.log.4  
-r----- 1 root root 15728657 Sep 8 06:50 audit.log.5
```

#### **New Audit Directory on Control Station:**

#### **/etc/audit**

**/etc/audit/auditd.conf** →controls auditd behavior, log locations, log rotation, etc.

**Note:** This file defines auto log rotation rules and trimming—keeps up to 7 log files at 15MB each

**/etc/audit/audit.rules** →file telling auditd what to audit

**Note:** By default, logs commands run by users, list of sensitive files, changes to auditing, user accounts, roles, users authenticating to system

#### **ENABLING/DISABLING AUDITD:**

**/sbin/chkconfig --levels 2345 auditd on | off**

**# /sbin/chkconfig --list auditd**

```
auditd      0:off 1:off 2:on 3:on 4:on 5:on 6:off
```

**/sbin/ausearch** →reads audit trail, searches on many different keywords

**/sbin/aureport** →provides summary reports on logs

**# /sbin/aureport**

Summary Report

=====

Range of time in logs: 09/08/2010 09:50:03.666 - 09/09/2010 00:47:37.877

Selected time for report: 09/08/2010 09:50:03 - 09/09/2010 00:47:37.877

Number of changes in configuration: 0

Number of changes to accounts, groups, or roles: 0

Number of logins: 4

Number of failed logins: 9

Number of authentications: 8

Number of failed authentications: 4

Number of users: 2

Number of terminals: 10

Number of host names: 5

Number of executables: 28

Number of files: 27

Number of AVC's: 0

Number of MAC events: 0

Number of failed syscalls: 9

Number of anomaly events: 0

Number of responses to anomaly events: 0

Number of crypto events: 0

Number of process IDs: 2116

Number of events: 7911

/sbin/auditctl → kernel audit subsystem

/sbin/service auditd start | stop | status, etc.

### # /sbin/aureport --auth

Authentication Report

=====

# date time acct host term exe success event

=====

```
1. 09/08/2010 14:43:26 acct="nasadmin 765nj11.corp.emc.com ssh /usr/sbin/sshd yes 5755215
2. 09/08/2010 14:43:29 acct="root ? pts/0 /bin/su yes 5755262
3. 09/08/2010 17:04:43 acct="nasadmin 765nj11.corp.emc.com ssh /usr/sbin/sshd yes 5772543
4. 09/08/2010 17:04:48 acct="root ? pts/0 /bin/su yes 5772590
5. 09/08/2010 19:52:56 acct="nasadmin 765nj11.corp.emc.com ssh /usr/sbin/sshd yes 5792318
6. 09/08/2010 19:52:59 acct="root ? pts/0 /bin/su yes 5792365
7. 09/09/2010 00:11:04 acct="nasadmin uscscsuser1l29c.corp.emc.com ssh /usr/sbin/sshd no 5824591
8. 09/09/2010 00:11:08 acct="nasadmin uscscsuser1l29c.corp.emc.com ssh /usr/sbin/sshd no 5824593
9. 09/09/2010 00:11:15 acct="nasadmin uscscsuser1l29c.corp.emc.com ssh /usr/sbin/sshd no 5824595
10. 09/09/2010 00:11:32 acct="nasadmin uscscsuser1l29c.corp.emc.com ssh /usr/sbin/sshd no 5824597
11. 09/09/2010 00:11:57 acct="nasadmin uscscsuser1l29c.corp.emc.com ssh /usr/sbin/sshd yes 5824601
12. 09/09/2010 00:12:01 acct="root ? pts/0 /bin/su yes 5824648
```

### # /sbin/service auditd start

Starting auditd: [ OK ]

### # /sbin/service auditd

Usage: /etc/init.d/auditd {start|stop|status|restart|condrestart|reload|rotate|resume}

/var/log/audit

### # ps -ef |grep audit

```
root 20750 1 0 Mar24 ? 00:02:12 auditd
root 21465 13074 0 Mar24 ? 00:00:02 /bin/sh /nas/sbin/dirsync /etc/audit /nas/var/auditing/.backup 180 auditd.conf audit.rules
root 21466 13074 0 Mar24 ? 00:00:02 /bin/sh /nas/sbin/dirsync /celerra/audit /nas/var/auditing/.backup 180 audit.log
```

### Mount Monitor Process Roots:

→Mount monitor process watches for automounting of the Server rootfs, detects and issues the file system watch for rootfs

### # ps -ef |grep mt\_mon

```
root 20752 20750 0 Mar24 ? 00:00:57 /nasmcd/sbin/mt_mon
```

**Note:** FS Watch capability and monitoring tool for rootfs

→Audit Daemon writes to log and passes record to Mount Monitor for action, as required

## **6.0 C3 CELERRA COMMON CACHE IMPROVEMENTS:**

→New Celerra Common Cache service (C3) for use by WEB-UI, JServer, XML-API, SMI-S, & CLI clients

### Logs:

### /celerra/commoncache/logs/c30.log.0

**Note:** Logs will show if client request was not met via the cache, indicated by “doPassThru”. Log will also show when an Indication is received that updates the cache: “Deleted Mount thru indication”. Indications are also logged in the /nas/log/cel\_api.log.

### Configuration Files:

/celerra/commoncache/bin/config/c3Configuration.xml (cache behavior, properties) | c3Logging.properties (logging)

→C3 uses Port 9824

→Cache is kept up-to-date via Indications mechanism

→Cache can also be updated manually via Refresh button and Invalidate Cache and Refresh section from Tools in Unisphere

## **NAS 6.0 HC MESSAGE FORMAT & IDs FOR PUHC/PAHC:**

HC\_BE\_14505017348

→New series of messages and descriptions that start with an “HC\_xxx” ID

- nas\_message –info cannot output the details of HC messages
- Primus solutions written for all Error and Warning PUHC messages, but many have incomplete content
- The PUHC healthcheck descriptions and IDs can be viewed in the following files:

#### /nas/tools/upgrade\_check\_tools/catalog/en

**cs\_msg.xml** →Control Station messages

**dm\_msg.xml** →Data Mover messages

**be\_msg.xml** →Backend messages

**sys\_msg.xml** →System messages

#### **PAHC/PUHC EXAMPLES:**

Error HC\_SYS\_14505213997: NAS services are not started on the Control Station.

Action : Run "service nas start" to restart the service...

Control Station: Check auto transfer status

Warning HC\_CS\_18800050417: The automatic transfer feature is disabled.

Action : EMC recommends the automatic transfer feature to be enabled...

Storage System : Check if Fibre Channel zone checker is set up

Information HC\_BE\_27389919354: You have not set up...this check...zoning cannot be checked.

#### **Some checks are used for multiple variables, as in the following examples:**

Blades : Check DNS connectivity and configuration

Warning HC\_DM\_18800115743:

\* server\_2: The DNS service is currently stopped and does not contact any DNS server.

Warning HC\_DM\_18800115743:

\* server\_2: The Network Time Protocol subsystem (NTP) has been stopped or is not connected to its server.

Warning HC\_DM\_18800115743:

\* server\_2: The NIS domain 'hosts.pvt.dns' is defined with only one NIS server.

Warning HC\_DM\_18800115743:

\* server\_2: The NIS server '192.1.4.210' is not accessible on the network (no answer from 192.1.4.210).

## **6.0 SERVER STATS & NAS STATS STATISTICS ENHANCEMENTS:**

→Additional Blade statistics reported, -summary switch replaced with –monitor, -table switch removed

→Statistics are reported in real-time and in time-series fashion

→Stats are housed in Tree format, and a statpath is the fully qualified path to a particular statistic, separated by periods (e.g., rep.v2.Session)

→statmonService enabled by default on Blades, using port 7777, and collects raw values for all the statistics

→Additional options added to server\_stats interface, all Blade stats now reported

→New Control Station nas\_stats interface to allow creation & management of statgroups, based on existing system-defined statgroups, as well as the use of thousands of possible statpaths

**Note:** A User would define a particular group of stats to gather, create a custom-defined statgroup, then invoke the statgroup name during a server\_stats –monitor <custom\_group> operation

→Default stat update sampling period “interval” is every 15 seconds, but can be modified between 1-300 seconds

→Stat Groups were added to help users run groups of stats without typing each individual statpath name, using system-defined groups, as well as allowing creation of custom user-defined groups

#### **Server\_stats still the main reporting interface:**

# **server\_stats server\_2 -service -status**

server\_2 : The statmonService has started.

interface=INTERNAL

port=7777

allow=128.221.252.100:128.221.252.101:128.221.253.100:128.221.253.101

The statmonService is listening for incoming network connections.

Max Connections: 32, Current: 0

→The –summary and –table switches for server\_stats have been deprecated

→Use \$ server\_stats –list and –info to obtain more information about statpath names for a Blade

# **server\_stats server\_2 –service –start –port 7777**

# **server\_stats server\_2 -service -status**

server\_2 : The statmonService has started.

interface=INTERNAL

port=7777

allow=128.221.252.100:128.221.252.101:128.221.253.100:128.221.253.101

The statmonService is listening for incoming network connections.

Max Connections: 32, Current: 0

# **server\_stats server\_2 -list |grep ntp**

Family ntp

```

Counter      ntp.firstHits
Counter      ntp.firstMisses
Counter      ntp.hits
Compound     ntp.lastDelta
Fact        ntp.lastDelta.usec
Fact        ntp.lastDelta.sec
Counter      ntp.misses
# server_stats server_2 -info ntp.hits
server_2 :
name       = ntp.hits
description = Number of successful contacts with the NTP server
type       = Counter
member_stats =
member_elements =
member_of   =
→New -monitor option replaces -summary for statpaths and statgroups
# server_stats server_2 -monitor basic-std (formerly -summary basic)
server_2 CPU Network Network dVol dVol
Timestamp Util In Out Read Write
% KiB/s KiB/s KiB/s KiB/s
13:06:59 0 14 1 0 12

```

**Note:** -monitor can be used to gather information on multiple stats by comma-separating each stat name, or, by using a single set of statpaths with sorting options

```

-monitor -action {status|enable|disable}
[ { -monitor {statpath_name|statgroup_name}[,...]
| -monitor {statpath_name|statgroup_name}
  [-sort <field_name>]
  [-order {asc|desc}]
  [-lines <lines_of_output>]
# server_stats server_2 -monitor net.device -sort device
server_2 device Network Network Network Network Network Network
Timestamp In In In Out Out Out
Pkts/s Errors/s KiB/s Pkts/s Errors/s KiB/s
13:15:24 mge0 23 0 29 14 0 2
mge1 0 0 0 0 0 0

```

#### Old Statistic Commands removed:

**server\_mpfs**  
**server\_cifssstat**  
**server\_nfssstat**

#### Example:

**# server\_nfssstat server\_2**

server\_2 :

Info 26306752351: server\_2 : This command has been deprecated and replaced with server\_stats command.

Error 2237: Execution failed: Segmentation fault: Operating system signal. [A\_ERROR\_SERVICE.last\_fatal\_error\_code]

#### Top Talkers -top switch from 5.6.47 replaced with the following:

##### CIFS Clients:

**# server\_stats server\_2 -i 1 -c 1 -te no -m cifs.client**

**Note:** Reports CIFS stats by Client IP addresses/Username for Reads/Writes, Suspicious operations, Averages for SMB1 & SMB2

##### NFS Clients:

**# server\_stats server\_2 -i 1 -c 3 -m nfs.client**

**Note:** Reports NFS stats by Client IP addresses for Read/Writes, Suspicious operations, Averages, for NFSv2, NFSv3, NFSv4

##### Quota Trees:

**# server\_stats server\_2 -i 1 -c 2 -te no -m fs.qtree**

**Note:** Reports file systems stats by quota tree Read/Write operations per second

#### **# server\_stats**

server\_stats <mountname>

-list

```

| -info [-all<statpath_name>[,...]]
| -service { -start [-port <port_number>]
  | -stop

```

```

| -delete
| -status }

|-monitor -action {status|enable|disable}
[{| -monitor {statpath_name|statgroup_name}|,...]
|-monitor {statpath_name|statgroup_name}
  [-sort <field_name>]
  [-order {asc|desc}]
  [-lines <lines_of_output>]

[-count <count>]
[-interval <seconds>]
[-terminationsummary {no|yes|only}]
[-format {text |titles {never|once|<repeat_frequency>}}|csv}]
[-type {rate|diff|accu}]
[-file <output_filepath> [-overwrite]]

```

**Note:** For printing stats, default format is text. If using .csv, use –file option to avoid seeing multiple title lines in output. For “–type”, it depends on whether the statistic is one that increases monotonically (network bytes), or one that displays a point-in-time value (CIFS connections). ‘Rate’ displays rate of change since previous sample. ‘Diff’ displays change in value since previous sample. ‘Accu’ displays change in value since initial sample. Default display ‘type’ is ‘rate’.

#### Server Stats Examples:

```
$ server_stats server_2 -monitor basic-std -interval 5 -count 5 -type rate | diff | accu
$ server_stats server_2 -monitor cifs-std -i 3 -c 5 (5 iterations of cifs-std statistics group at 3-sec. intervals)
$ server_stats server_2 -monitor cifsops-std -i 5 -c 3 (SMB dialect statistics outputs for CIFS)
$ server_stats server_2 -monitor netDevices-std -i 5 -c 3 (Network device summary every 5 secs, three times)
```

**Note:** Use “–terminationsummary no” option to not display the Min/Aver/Max summary for the devices (also, “te no”)

#### CONTROL STATION NAS STATS COMMAND:

→New command to help manage Statistic Groups

# nas\_stats

USAGE:

nas\_stats

```

-groups
{
  -list
  |-info [-all<statgroup_name>|,...]
  |-create <statgroup_name>
    [-description "<description_line>"]
    {<statpath_name>|<statgroup_name>}[,...]
  |-modify <statgroup_name>
    { [-rename <new_statgroup_name>]
      [-description "<description_line>"]
      [<statpath_name>|<statgroup_name>|,...] }
  |-add <statgroup_name>
    {<statpath_name>|<statgroup_name>}[,...]
  |-remove <statgroup_name>
    {<statpath_name>|<statgroup_name>}[,...]
  |-delete <statgroup_name> [-Force]
  |-database
    {-recover [-Force] →Tries to restore Statistic Groups db from latest NAS_DB backup; recreates default db if required
     |-verify }
```

#### Creating & using a custom StatGroup:

```
# nas_stats -groups -create nfsv4_group nfs-std,nfsOps-std,nfs.v4.op
'nfsv4_group' created successfully
```

# nas\_stats -groups -list

| Type   | Name            |
|--------|-----------------|
| System | basic-std       |
| System | caches-std      |
| System | cifs-std        |
| System | cifsOps-std     |
| System | diskVolumes-std |
| System | metaVolumes-std |
| System | netDevices-std  |
| System | nfs-std         |

System nfsOps-std  
User nfsv4\_group →Example of a custom-defined group name  
**Note:** Provides a list of system-defined Stat Groups, and user-defined Stat Groups, if created

```
# nas_stats -groups -info nfsv4_group
name      = nfsv4_group
description = nfsv4_group
type      = User-defined
member_stats = nfs-std,nfsOps-std,nfs.v4.op
member_elements =
member_of   =
```

```
# server_stats server_2 -monitor nfsv4_group
# nas_stats -groups -info basic-std
name      = basic-std
description = The basic system-defined group.
type      = System-defined
member_stats = kernel.cpu.utilization.cpuUtil,net.basic.bytesIn,net.basic.bytesOut,store.readBytes,store.writeBytes
member_elements =
member_of   =
# nas_stats -groups -create <statgroup_name> -description "description_here" <statpath_name(s)> [comma separated]
# nas_stats -groups -modify <stat_group> -rename | -description "new_descript" | -delete <statgroup_name> | -add | -remove
# nas_stats -groups -database -verify
```

Database is healthy.

#### **# nas\_stats -groups -database -recover**

Latest healthy database modified last on Fri Jul 23 13:50:45 EDT 2010.  
Any updates performed after the latest backup will be lost. Continue? [Y/N] y

nas\_stats: recovery completed successfully.

#### **# server\_stats server\_2 -service -status**

server\_2 : The statmonService has started.

interface=INTERNAL

port=7777

allow=128.221.252.100:128.221.252.101:128.221.253.100:128.221.253.101

The statmonService is listening for incoming network connections.

Max Connections: 32, Current: 0

#### **# server\_stats server\_2 -monitor -action status**

server\_2 : Statistics are enabled.

#### **New options under server stats:**

**-list** →lists all statpaths and statgroups

**-info** →detailed info on statpaths/statgroups

**-monitor** →specify a statpath or statgroup to collect stats on

**-service** →managing statmonService on the Data Mover

#### **Statistic Types:**

Family—corresponds to a major feature or sub-system in DART

Compound—stats that are grouped together

Set—collection of compound instances

Correlated Set—correlations across sub-systems

Element—an instance in a set or correlated set

Counter—an increasing value statistic

Fact—point in time information

Computed—create a statistic from 2 or more other stats

Statistic Group—collection of stats defined at the Control Station, an evolution of the Summary & Table collections

#### **Summary and Table collections replaced by statistic family groups:**

| System-defined statistics group name | server_stats collection name |
|--------------------------------------|------------------------------|
| basic-std                            | -summary basic               |
| caches-std                           | -summary caches              |
| cifs-std                             | -summary cifs                |
| nfs-std                              | -summary nfs                 |
| cifsOps-std                          | -table cifs                  |
| diskVolumes-std                      | -table dvol                  |
| metaVolumes-std                      | -table fsvol                 |
| netDevices-std                       | -table net                   |

nfsOps-std -table nfs

**CREATE & USE NEW STATGROUP FROM SYSTEM-DEFINED GROUPS & STATPATHS:****1. Create new statgroup**

# nas\_stats -groups -create newgroup basic-std,caches-std,ftp.auth.cifsFailures -description "New group"

'newgroup' created successfully

# nas\_stats -groups -info newgroup

```
name      = newgroup
description = New group
type      = User-defined
member_stats = basic-std,caches-std,ftp.auth.cifsFailures
member_elements =
member_of  =
```

**Note:** User nas\_stats -groups -add or -remove to add or remove statpaths to existing User-defined statgroups**2. Invoke collection of stats for the new statgroup**

# server\_stats server\_2 -monitor newgroup

| server_2  | CPU   | Network | Network | dVol  | dVol  | DNLC    | OF Cache | Buffer | Failed           |
|-----------|-------|---------|---------|-------|-------|---------|----------|--------|------------------|
| Timestamp | Util  | In      | Out     | Read  | Write | Hit     | Hit      | Cache  | CIFS             |
| %         | KiB/s | KiB/s   | KiB/s   | KiB/s | KiB/s | Ratio % | Ratio %  | Hit %  | Authorizations/s |
| 13:47:05  | 0     | 20      | 1       | 0     | 14    | -       | -        | 100    | 0                |
| 13:47:20  | 0     | 11      | 1       | 0     | 12    | -       | -        | 100    | 0                |
| 13:47:35  | 0     | 31      | 2       | 0     | 34    | -       | -        | 100    | 0                |

| server_2 | CPU   | Network | Network | dVol  | dVol  | DNLC    | OF Cache | Buffer | Failed           |
|----------|-------|---------|---------|-------|-------|---------|----------|--------|------------------|
| Summary  | Util  | In      | Out     | Read  | Write | Hit     | Hit      | Cache  | CIFS             |
| %        | KiB/s | KiB/s   | KiB/s   | KiB/s | KiB/s | Ratio % | Ratio %  | Hit %  | Authorizations/s |
| Minimum  | 0     | 11      | 1       | 0     | 4     | -       | -        | 100    | 0                |
| Average  | 0     | 219     | 8       | 0     | 206   | -       | -        | 100    | 0                |
| Maximum  | 0     | 1219    | 38      | 0     | 1155  | -       | -        | 100    | 0                |

**TROUBLESHOOTING SERVER STATS:**

# .server\_config server\_2 -m nfs.client -TRACE

**6.0 SMI-S PROVIDER—Web Based Enterprise Management (WBEM):**

→Features represents the Storage Management Initiative-Specification to provide an SMI-S compliant API for storage products (SMI-S 1.3)

→SMI-S is an infrastructure designed to manage storage devices

→CIM client, CIM server, &amp; Provider

→SMI-S is based on CIM, with clients able to manage Celerra via CIM

→Basically Web Based Enterprise Management (WBEM) using Common Information Model (CIM XML)

→Embeds the WBEM infrastructure on the CS in the form of a CIM Server (aka, ECOM, using CIMOM—CIM Object Manager) and a Provider

→ECOM &amp; SMI-S are embedded on the Control Station

→ECOM is the EMC Common Object Model CIM Server

**Starting ECOM Service:**

→vi edit /nas/sys/nas\_mcd.cfg and uncomment “cim conf”, “cim server”, and “SMIS” logs, stop &amp; restart NAS services daemon “cim server”

```
executable  "/celerra/wbem/bin/start_cim_server"
optional    no
autorestart yes
cmdline    "/celerra/wbem"
```

---

```
# daemon "SMISPlugin Log Trimmer"
#   executable  "/nas/sbin/log_trimmer"
#   optional    no
#   autorestart yes
#   ioaccess   no
#   cmdline    "-n /nas/log/smisi/SMISPlugin.log 1000 1 2 h t 4 y "
```

---

```
# daemon "SMIS cimomlog.txt Log Trimmer"
#   executable  "/nas/sbin/log_trimmer"
#   optional    no
#   autorestart yes
```

```
#    ioaccess    no
#    cmdline     "-n /nas/log/smis/cimomlog.txt 1000 1 2 h t 4 y"
# ps -ef |grep -i cim
root  25995 19683 0 10:19 ?      00:00:00 /bin/sh /celerra/wbem/bin/start_cim_server /celerra/wbem
root  26007 19683 0 10:19 ?      00:00:00 /nas/sbin/log_trimmer -n /nas/log/smis/cimomlog.txt 1000 1 2 h t 4 y
# ps -ef |grep -i ecom
root  28071 25995 3 10:20 ?      00:00:01 ECOM
```

#### Troubleshooting:

```
# ls -la /nas/log/smis
```

```
-rw-r--r-- 1 root  root  16738 Jul  2 18:34 cimomlog.txt
-rw-r--r-- 1 root  root    0 Jul  2 18:32 HTTP_trace.log
-rw-r--r-- 1 root  root  2458 Jul  2 18:34 securitylog.txt
-rw-r--r-- 1 root  root  22057 Jul  2 18:34 SMISPlugin.log
```

#### WBEM PORT:

```
# netstat -a |grep 5988
tcp    0      0 *:5988          *.*        LISTEN
```

## **6.0 MS IDENTITY MANAGEMENT FOR UNIX PHASE II (IDMU):**

→IDMU is a Microsoft solution that offers a centrally managed environment, under Active Directory, for both Windows and UNIX users

→Microsoft Active Directory, Identity Management for UNIX is comprised of a series of tools to help integrate a Windows environment with a UNIX environment, where the AD Domain Controller becomes the Master NIS Server and uses Password Synchronization to automatically synchronize passwords between Windows and UNIX users, for centralizing management of the Windows and UNIX environments

→IDMU is installed on an AD server, along with “Server for NIS”, “Administration Components”, “Password Synchronization”, and “Other Network File and Print Services: Microsoft Services for NFS”

→Celerra has supported IDMU since 5.6.39

→Kerberos authentication has been added to IDMU support for 6.0

→IDMU is found in Windows 2003 R2 and later

#### MS IDMU Active Directory Configuration:

--Server for NIS; Administration Components; Password Synchronization; Other Network File and Print Services: Microsoft Services for NFS  
--Create or configure an existing Group with UNIX GID and NIS Domain  
--Configure AD Users with UNIX username, etc.  
--Configure User Name Mapping,NIS for the NFS service, Simple Mappings, etc.

#### Celerra Configuration:

1. Make a copy of the default /etc/ldap.conf.idmu\_template\_v1 file and edit to reflect LDAP Base domain and AD Domain name

a) Edit the following fields to reflect your LDAP Base domain and AD domain name

```
# Replace "dc=mydomain,dc=com" by your base DN.
```

```
# If you have a dedicated container for netgroups, replace
```

```
# "cn=netgroup,cn=mydomain,cn=DefaultMigrationContainer30" by the right DN.
```

```
nss_base_passwd  cn=Users,dc=mydomain,dc=com?one
```

```
nss_base_group   cn=Users,dc=mydomain,dc=com?one
```

```
nss_base_hosts   cn=Computers,dc=mydomain,dc=com?one
```

```
nss_base_netgroup cn=netgroup,cn=mydomain,cn=DefaultMigrationContainer30,dc=mydomain,dc=com?one
```

2. Push file out to Data Mover as /etc/ldap.conf

3. Copy /nas/sys/nsswitch.conf.tmpl file and add “ldap” to each search order line, then push to DM as /etc/nsswitch.conf  
passwd: files ldap nis  
group: files ldap nis  
hosts: dns ldap nis files  
netgroup: files ldap nis

4. Configure and start LDAP client on Data Mover

**# server\_ldap server\_2 -set -basedn dc=idmu,dc=emc,dc=com -servers 192.1.8.158 -kerberos -kaccount compname\$**

Note: Where basedn is LDAP basedn; Servers = LDAP server; Compname = CIFS Name\$ (\$ sign required!)

5. Start NIS Service and set CIFS Resolver for NIS lookups:

a) **# server\_nis server\_2 hosts.pvt.dns 192.1.4.210**

server\_2 : done

b) **# server\_param server\_2 -facility cifs -modify resolver -value 1**

server\_2 : done

#### Testing/Troubleshooting:

```
# server_ldap server_2 -lookup -user <username>
```

# server\_ldap server\_2 -info

# server\_ldap server\_2 -i -v (Verifying LDAP configuration and connection to LDAP domain)

**Note:** Check Base DN, connection state, config files, etc.

### **BLACKBIRD:**

→Blackbird to become 64-bit to support DART, and will have libraries to support 32-bit Linux Control Station

→Runs on VMWare Workstation or Player host-based virtualization

### **6.0 SNMPv3 SUPPORT:**

→A new SNMPD daemon, supporting SNMPv1, SNMPv2, and SNMPv3, runs on the Data Mover

→SNMP consists of managed network devices, agents, and network management system applications

→SNMP is used to establish a baseline state for devices, and then uses network watchdog ‘agents’ to monitor, sending alerts or trap notifications to the management station when something changes/errors

→SNMPv3 uses a User Security Model (USM) to provide SNMP security on the network

→\$ server\_snmp is replaced with \$ server\_snmpd

# server\_snmpd server\_2

```
-info  
|-service {  
    -status  
    | -stop  
    | -start  
|-modify [ -location <sys_location> ]  
    [ -contact <sys_contact> ]  
    [ -community { -clear | <community> } ]  
|-user {  
    -list  
    | -create <name> -authpw -privpw  
    | -delete <name>  
    | -modify <name> -authpw -privpw
```

→SNMP collects and processes network information at fixed or random intervals, based on watchdog agents, which sends alerts via Traps

→MIBs databases are used to manage & monitor devices

→Managed devices collect & store NMI (Network Mgmt Information) and send to NMS (Network Mgmt System) using SNMP

→Main benefit of SNMPv3 are security and remote configuration enhancements (authentication, encryption, message integrity)

→Support for SNMPv1, v2c, and v3 for DART-only

→Current MIBs and IPv6 MIBs supported

→Will use MD5 for authentication and DES for privacy

### **CONFIGURING SNMPv3:**

#### **1. Create User & Authentication/Privacy passwords:**

# server\_snmpd server\_2 -user -create snmp\_user -authpw -privpw

Enter authentication password:

Confirm authentication password:

Enter privacy password:

Confirm privacy password:

server\_2 :

OK

#### **2. Set SNMP Location & Contact:**

# server\_snmpd server\_2 -modify -location "svt" -contact "me"

server\_2 :

OK

# server\_snmpd server\_2 -info

server\_2 :

enabled = Yes

location = svt

contact = me

community =

# server\_snmpd server\_2 -user -list

server\_2 :

snmp\_user

9bRL1SD73nrc9BbF5FdHYA82KeP1P36

# server\_snmpd server\_2 -service -status

server\_2 :

SNMP Running

### **USING SNMPGET ON CONTROL STATION:**

# snmpget -v3 -u snmp\_user -l authPriv -a MD5 -A nasadmin -x DES -X nasadmin 128.221.252.2 sysContact.0

SNMPv2-MIB::sysContact.0 = STRING: me

### **DISABLING SNMPv1 & v2:**

# server\_snmpd server\_2 -modify -community -clear

server\_2 :

OK

### **ENABLING SNMPv1/v2:**

# server\_snmpd server\_2 -modify -community private

### **CPW CHANGES:**

→Provides user with the option to skip provisioning of EFD disks as disk volumes, and instead reserve them for CLARiiON EFD Cache functionality

→If user wishes to use the CLARiiON EFD cache feature, they should proceed to Unisphere and configure EFD Cache on the array

### **Excerpt from CPW screen:**

“EFD Information

This system has unused EFD disks...By default, EFD disk provisioning is bypassed. To provision EFD disks, check the following box.

— Allow EFD disks provisioning for Celerra storage.”

### **CSA CHANGES:**

→CSA screen messages and popups will now conform to an Event Code that can be queried using nas\_message (aligns with CCMD messaging)

### **6.0 REPLICATION V1 NOT SUPPORTED:**

→IP Replication V1 is not supported on 6.0 and systems will not upgrade if any of the following conditions are seen by the PUHC

→Only resolution would be to remove all vestiges of RepV1 before upgrading

→See emc242421 for more details

### **V1 Replication Sessions running:**

Error HC\_CS\_14505083114: There are Celerra Replicator (V1) sessions or checkpoints on this Celerra. Replicator (V1) has been deprecated and is not supported in the NAS version to which you are upgrading.

### **V1 FS Copy Sessions running:**

Error HC\_CS\_14505083115: There are fs\_copy sessions on this Celerra. Replicator (V1) has been deprecated and is not supported in the NAS version to which you are upgrading.

### **V1 License still active:**

Error HC\_CS\_14505083116: There is a Celerra Replicator (V1) license on this Celerra. Replicator (V1) has been deprecated and is not supported in the NAS version to which you are upgrading.

### **6.0 CELERRA FAST REBOOT (aka Warm Reboot--Fixed times vs Variable times):**

→Main purpose to reduce the ‘fixed time’ for reboots (much of which involved BIOS/POST hardware checks) from 3 minutes to 30 seconds

→‘Variable times’ will still apply to overall system reboot/failover times, but are mitigated by improvements introduced with NAS 5.6.47 & later

→Still, the normal data mover warm reboot is extremely fast & a big improvement

→Warm reboot skips hard HW reset, BIOS, POST, & DOS loading steps

→Warm reboot is the default action for server\_cpu server\_2 –reboot now

→Must issue server\_cpu server\_2 –reboot –cold now to do cold reboot

→Certain conditions will not qualify for warm reboot, such as HW malfunction, Data Mover panics, t2reset issued, or for manual or automatic failover

→Purpose is to allow for faster reboots of the Data Mover by resetting the Operating System only, not the motherboard

→Warm reboot will skip the BIOS & POST steps, which are Fixed reboot times

**Note:** Still see a reference to the BIOS during Warm reboot, however

2010-05-06 10:24:50: ADMIN: 6: Warm reboot begins. Shutting down the server ...

2010-05-06 10:27:15: KERNEL: 6: System BIOS Rev Sledgehammer

→Fast/Warm reboot actually has Data Mover loading NAS image and config files first before Control Station tells Data Mover to then proceed with reboot

**Note:** Server failover & fallback does “Cold” reboot on the slot that fails over, and the slot that fails back

# server\_log server\_2.faulted.server\_3 -s -a |grep -i cold

2010-05-10 15:26:53: ADMIN: 6: Cold reboot begins. Shutting down the server ...

# server\_log server\_3 -a -s |grep -i cold

2010-05-10 15:31:49: ADMIN: 6: Cold reboot begins. Shutting down the server ...

→Warm Reboots are used during Installs and Upgrades

→NS20/40/80/NSX support this feature since they can flash firmware, unlike older platforms which flash from DOS, which is not supported by Warm Reboot

→Firmware updates may require two reboots, one Warm Reboot which flashes firmware, which then detects BIOS/POST change and requires Cold Reboot

→Default behavior for server\_cpu is to attempt a Warm Reboot, but if warm reboot does not qualify, will do a Cold Reboot

**Note:** Users need to know that the default server\_cpu syntax will warm reboot. Must stipulate server\_cpu –reboot –cold now.

#### **DEFAULT SERVER CPU IS WARM REBOOT:**

2010-05-07 09:15:51: KERNEL: 3: Warm Reboot: Trying CS ip address 128.221.252.100

2010-05-07 09:15:51: KERNEL: 6: Warm Reboot: boot.bat: gload -or c:\bin\nas.exe c:\slot\_3\boot.cfg

2010-05-07 09:15:59: ADMIN: 6: Warm reboot begins. Shutting down the server ...

#### **Outline of Traditional Boot Process**

→Controlled DART shutdown ~ 5 secs

#### **New Warm Boot Process**

DART shutdown ~ 5 secs

→Hardware reset ~ 10 secs

→BIOS ~ 50 secs

→Extended POST ~ 50 secs

→DOS load ~ 5 secs

→gload ~ 5 secs

Bootloader ~ 5 secs

→DART init ~ 55 secs

DART init ~ 20 secs

#### **Server Log for panic reboot:**

2010-05-07 09:30:25: ADMIN: 6: Command succeeded: param kernel autoreboot=600

**Note:** Log does not specify whether this is a Warm or Cold reboot

#### **Server Log for Warm Reboot:**

# **server\_cpu server\_2 -reboot -m now**

# **server\_log server\_2 -a -s |grep -i warm**

2010-04-05 15:13:27: ADMIN: 6: Warm reboot begins. Shutting down the server ...

#### **Cold Reboot:**

# **server\_cpu server\_2 -reboot cold -m now**

server\_2 : reboot in progress 0.0.0.0.0.0.0.0.1.1.1.3.3.4.done

#### **Server Log for Cold Reboot:**

# **server\_log server\_2 -a -s |grep -i Cold**

2010-04-05 15:05:24: ADMIN: 6: Cold reboot begins. Shutting down the server ...

# **nas\_logviewer /nas/log/sys\_log |grep -i cold**

Apr 5 15:07:36 2010:DART:CHAMII:INFO:103:Slot 2:::1270494490:Info: POST\_INFO: 04/05/10 19:07:05 : INFORMATION:  
Power Up Status (0x86): Coldfire Power Down caused reset;

#### **FILE SYSTEM & NFS PARAMETER CHANGES (all require rebooting blade):**

→Open NFS files, directory name lookup cache, vnode cache, and NFS threads are being replaced by regular DART parameters, as follows:

# **server\_param server\_2 -facility file -list**

| param_name  | facility | default | current | configured |
|-------------|----------|---------|---------|------------|
| dnlcNents   | file     | 384000  | 384000  |            |
| cachedNodes | file     | 256000  | 256000  |            |

→DNLC replaced by dnlcNents, provides mappings between filenames and files in all file systems on the server. Default size depends on physical memory of Blade: 4GB=484k; 8GB=768k; 24GB=2.304m

→Nodes replaced by cachedNodes, the vnode inode & block cache for files & directories globally for all server file systems.  
4GB=256k; 8GB=512k; 24GB=1.536m

# **server\_param server\_2 -facility nfs -list**

| param_name  | facility | default | current | configured |
|-------------|----------|---------|---------|------------|
| nthreads    | nfs      | 256     | 256     |            |
| ofCachesize | nfs      | 240000  | 240000  |            |

→Nfsd replaced by nthreads, the # threads dedicated to NFS requests. 4GB=256; >than 4GB=512

→Openfiles replaced by ofCachesize, the size of the NFS open files cache. 4GB=240k; 8GB=360k; 24GB=3080m

→All of the above parameters require System reboot before going into effect

#### **6.0 NDMPCOPY:**

→Open source program for copying/migrating file system data between Data Mover file systems in the same or different Celerra, or within same Data Mover

→Download the utility from [www.ndmp.org](http://www.ndmp.org) (or SUSE and Linux 1.2 version from Apps & Tools CD)

→Utility supported Red Hat 5.2, 5.3, HP-UX 11.3, Solaris 10, SuSE Linux 10 & 11

→Supports multi-protocol file systems, preserves CIFS & NFS permissions

→Supports data copy from 5.6 to 6.0 systems, 6.0 to 5.6, 6.0 to 6.0

→Requires use of username & password for NDMPCopy

**Note:** Create user and passwd for each blade using server\_user command

**LIMITATIONS/ISSUES:**

→Not recommended to run on the Control Station (may run over internal network)

→NDMPCopy uses clear text passwords

→Supports only NDMPv2 protocol

→Copy operation may hang if network communication not successful

→Copy operation may hang if Incremental Level 0 backup has no changes

→Does not support VBB or Tape Silvering

**NDMPCopy Example:**

# ./ndmpcopy -h

Usage: ndmpcopy src\_filer:/src/dir dest\_filer:/dest/dir

```
[ -sa none | user:password ] [ -da none | user:password ]
[ -sport ndmp_src_port ] [ -dport ndmp_dest_port ]
[ -dhost ndmp_dest_ip_addr ] [ -level ndmp_dump_level ]
[ -v ] [ -q ] [ -dpass ] [ -h ]
```

Defaults:

```
src_auth_type      = text
src_auth_user      = root
src_auth_password  =
dest_auth_type     = text
dest_auth_user     = root
dest_auth_password =
ndmp_src_port      = 0 (0 means NDMP default, usually 10000)
ndmp_dest_port      = 0 (0 means NDMP default, usually 10000)
ndmp_dump_level    = 0 (valid range: 0 - 9)
ndmp_dest_ip_addr  = (no default: user needs to override dest_filer value)
verbosity          = noisy
different_passwords = no
```

# **ndmpcopy <blade\_ip>:/fs1 <blade\_ip>:/fs20 -sa user1:passwd -da user2:passwd2**

**Note:** Specifying Username and Passwords for each Data Mover

**NDMPCOPY FROM ONE FS TO ANOTHER- DIFFERENT CELERRAS (From RHEL Host):**

# ./ndmpcopy 10.127.62.126:/fs2/bin 10.127.62.127:/fs3 -sa ndmp\_copy:nasadmin -da ndmp\_copy:nasadmin -sport 10000 -dport 10000 -v

Connecting to 10.127.62.126.

Connecting to 10.127.62.127.

10.127.62.126: CONNECT: Connection established.

10.127.62.127: CONNECT: Connection established.

10.127.62.126: LOG: server\_archive: emtar vol 1, 2 files, 0 bytes read, 585460 bytes written

10.127.62.126: HALT: The operation was successful!

Waiting for 10.127.62.127 to halt too.

10.127.62.127: LOG: server\_archive: emtar vol 1, 2 files, 585460 bytes read, 0 bytes written

10.127.62.127: HALT: The operation was successful!

The transfer is complete.

Elapsed time: 0 hours, 0 minutes, 6 seconds.

**NDMPCOPY FROM ONE FS TO ANOTHER SAME BLADE (From RHEL Host):**

# ./ndmpcopy 10.127.62.126:/fs2 10.127.62.126:/it/copy -sa ndmp\_copy:nasadmin -da ndmp\_copy:nasadmin -sport 10000 -dport 10000 -v

Connecting to 10.127.62.126.

Connecting to 10.127.62.126.

10.127.62.126: CONNECT: Connection established.

10.127.62.126: CONNECT: Connection established.

10.127.62.126: LOG: server\_archive: emtar vol 1, 47 files, 0 bytes read, 15028938 bytes written

10.127.62.126: HALT: The operation was successful!

Waiting for 10.127.62.126 to halt too.

10.127.62.126: LOG: server\_archive: emtar vol 1, 47 files, 15028938 bytes read, 0 bytes written

10.127.62.126: HALT: The operation was successful!

The transfer is complete.

Elapsed time: 0 hours, 0 minutes, 6 seconds.

**SERVER LOG OF SUCCESSFUL NDMPCOPY OPERATION:**

2010-07-14 19:36:36: NDMP: 4: Session 006 (thread ndmp006) < Backup type: dump >  
2010-07-14 19:36:36: NDMP: 4: Session 006 (thread ndmp006) Name: UPDATE Value: y  
2010-07-14 19:36:36: NDMP: 4: Session 006 (thread ndmp006) Name: TYPE Value: dump  
2010-07-14 19:36:36: NDMP: 4: Session 006 (thread ndmp006) Name: USER Value: root  
2010-07-14 19:36:36: NDMP: 4: Session 006 (thread ndmp006) Name: HIST Value: n  
2010-07-14 19:36:36: NDMP: 4: Session 006 (thread ndmp006) Name: FILESYSTEM Value: /fs2  
2010-07-14 19:36:36: NDMP: 4: Session 006 (thread ndmp006) Name: LEVEL Value: 0  
2010-07-14 19:36:36: NDMP: 4: Session 006 (thread ndmp006) Name: EXTRACT Value: y  
2010-07-14 19:36:36: NDMP: 4: Session 006 (thread ndmp006) tape service is on a remote side, CALLER\_ADDRESS 0x00015ae3eb.  
2010-07-14 19:36:36: NDMP: 4: Thread ndmp006 uses a set of Pax threads with handle 0.  
2010-07-14 19:36:36: NDMP: 6: Thread ndmp006 filterDialect could not be set - prereqs not met [pax\_cmd\_thread::setFilterDialect]  
2010-07-14 19:36:36: NDMP: 4: Thread ndmp006 fsSize of /fs2 is 15630336 (./pax\_ndmp.cxx: 1196)  
2010-07-14 19:36:36: NDMP: 4: Session 007 (thread ndmp007) Name: UPDATE Value: y  
2010-07-14 19:36:36: NDMP: 4: Session 007 (thread ndmp007) Name: TYPE Value: dump  
2010-07-14 19:36:36: NDMP: 4: Session 007 (thread ndmp007) Name: USER Value: root  
2010-07-14 19:36:36: NDMP: 4: Session 007 (thread ndmp007) Name: HIST Value: n  
2010-07-14 19:36:36: NDMP: 4: Session 007 (thread ndmp007) Name: FILESYSTEM Value: /fs2  
2010-07-14 19:36:36: NDMP: 4: Session 007 (thread ndmp007) Name: LEVEL Value: 0  
2010-07-14 19:36:36: NDMP: 4: Session 007 (thread ndmp007) Name: EXTRACT Value: y  
2010-07-14 19:36:36: NDMP: 4: Session 007 (thread ndmp007) < Backup type: dump >  
2010-07-14 19:36:36: NDMP: 4: Session 007 (thread ndmp007) tape service is on a remote side, CALLER\_ADDRESS 0x00015abd37.  
2010-07-14 19:36:36: NDMP: 4: Thread ndmp007 uses a set of Pax threads with handle 1.  
2010-07-14 19:36:36: NDMP: 6: Thread ndmp007 filterDialect could not be set - prereqs not met [pax\_cmd\_thread::setFilterDialect]  
2010-07-14 19:36:36: NDMP: 4: Thread ndmp007 fsSize of /fs2 is 15630336 (./pax\_ndmp.cxx: 1196)  
2010-07-14 19:36:36: NDMP: 6: Session 006 (thread nasa00) backup dir /fs2 is not on read-only checkpoint.  
2010-07-14 19:36:36: PAX: 6: Thread nasa00 dedup backup ratio threshold of /fs2 is 90%  
2010-07-14 19:36:36: NDMP: 4: Thread nasa00 Backup root directory: /fs2  
2010-07-14 19:36:36: NDMP: 4: Thread nasa00 write count 15028938  
2010-07-14 19:36:36: NDMP: 4: Thread nasa00 time used 0 sec  
2010-07-14 19:36:36: NDMP: 4: Thread nasa00 write rate 85828 KB/Sec  
2010-07-14 19:36:36: NDMP: 4: Thread nasa00 read rate 0 KB/Sec  
2010-07-14 19:36:36: NDMP: 4: Thread nasa00 average file size: 312KB  
2010-07-14 19:36:36: NDMP: 4: Thread nasa00 dir or 0 size file processed: 4  
2010-07-14 19:36:36: NDMP: 4: Thread nasa00 1B -- 8KB size file processed: 6  
2010-07-14 19:36:36: NDMP: 4: Thread nasa00 8KB -- 16KB size file processed: 3  
2010-07-14 19:36:36: NDMP: 4: Thread nasa00 16KB -- 32KB size file processed: 4  
2010-07-14 19:36:36: NDMP: 4: Thread nasa00 32KB -- 64KB size file processed: 6  
2010-07-14 19:36:36: NDMP: 4: Thread nasa00 64KB -- 1MB size file processed: 21  
2010-07-14 19:36:36: NDMP: 4: Thread nasa00 1MB -- 32MB size file processed: 3  
2010-07-14 19:36:36: NDMP: 4: Thread nasa00 32MB -- 1GB size file processed: 0  
2010-07-14 19:36:36: NDMP: 4: Thread nasa00 1G more size file processed: 0  
2010-07-14 19:36:36: NDMP: 6: Thread nasa00 server\_archive: emctar vol 1, 47 files, 0 bytes read, 15028938 bytes written  
**# ndmpcopy <blade\_ip>:/fs1 <blade\_ip>:/fs20 -sa user1: -da user2: -dpass**

**Note:** Allows system to prompt to enter password for each Blade

#### **Troubleshooting NDMPCopy:**

```
$ server_log server_x -f  
$ server_pax server_2 -stats -reset  
$ server_pax server_2 -stats -verbose  
$ ndmpcopy -v -v -v -v →Debug information and port information
```

#### **6.0 INCREASED NDMP BACKUP STREAMS:**

→Support increased from 4 to 8 NDMP concurrent backup streams  
→But, can only have >than 4 streams if Blade memory 8GB or more  
→Default value left at 4 streams, any changes to param requires reboot

**# server\_param server\_2 -facility NDntDataStreams -value 8**

server\_2 :

Error 4418: server\_2 : 8 is not in range (1,4)

**Note:** Above error means that Data Mover does not have 8GB memory, so command fails

**# server\_param server\_2 -facility NDMP -info concurrentDataStreams**

```
server_2 :
name          = concurrentDataStreams
facility_name  = NDMP
default_value   = 4
current_value   = 4
configured_value =
user_action     = reboot DataMover
change_effective = reboot DataMover
range          = (1,8)
description     = Number of concurrent data streams
```

**Note:** The range 1,8 will only be reflected by systems running 8GB memory or greater. Systems under 8GB memory will only show a possible concurrent stream range of 1,4

#### **SETTING STREAMS TO 8 VIA CLI:**

**# server\_param server\_2 -facility NDMP -modify concurrentDataStreams -value 8**

server\_2 : done

Warning 17716815750: server\_2 : You must reboot server\_2 for concurrentDataStreams changes to take effect.

**# server\_param server\_2 -facility NDMP -info concurrentDataStreams**

server\_2 :

```
name          = concurrentDataStreams
facility_name  = NDMP
default_value   = 4
current_value   = 8
configured_value = 8
```

-----abridged-----

#### **SETTING STREAMS TO 8 VIA UNISPHERE:**

Click Celerra system > Settings > Data Mover Parameters > Select Server, NDMP Facility, ALL Parameters, highlight param to change, Properties, then set Value and follow instructions to reboot

#### **6.0 NDMP PORT RANGE FEATURE:**

→ Celerra has the ability to specify a range of ports for NDMP use

→ The default NDMP port ranges are 1024-65535, and are also min & max

→ The minimum number of ports allowed for a range is 32, though the best practice recommendation is to have at least a 100 ports in the range

→ Recommended ranges to use are 49152-65535

→ No server reboots required

**# server\_param server\_2 -facility NDMP -info portRange -v**

server\_2 :

```
name          = portRange
facility_name  = NDMP
default_value   = 1024-65535
current_value   = 1024-65535
configured_value =
user_action     = none
change_effective = immediate
range          = *
description     = Port range for NDMP data/mover listening
```

detailed\_description

Specifies the range of ports allowed for NDMP data/mover listening.

#### **Example:**

**# server\_param server\_2 -facility NDMP -modify portRange -value 62000-63000**

server\_2 : done

**# server\_param server\_2 -facility NDMP -info portRange**

server\_2 :

```
name          = portRange
facility_name  = NDMP
default_value   = 1024-65535
current_value   = 62000-63000
configured_value = 62000-63000
user_action     = none
change_effective = immediate
range          = *
```

description = Port range for NDMP data/mover listening

## **6.0 SAMBA 4 SUPPORT:**

- Samba4 runs on UNIX host as a 2003 AD server, complete with AD, LDAP, DNS, & Kerberos KDC server
- Purpose is to allow Unix and Linux systems ability to co-exist in a Windows AD environment
- Samba 4 supports AD logon & authentication (but does not support DFS)
- Requires Windows client running MS Client and Administration pack

### **Configuring Celerra Data Mover:**

**\$ server\_param server\_2 -facility ldap -modify SecurityLayer –value 0**

**Note:** Setting param to zero disables Signing & Encryption and sets the LDAP Bind security layer to 0

→Use regular CIFS create and Join commands

## **6.0 pNFS SUPPORT:** (Parallel Network File System, an extension of NFSv4.1)

- pNFS stands for Parallel Network File System, an extension of the NFSv4.1 standard, and an EMC implementation
- pNFS supports the NFSv4.1 standard included in Open Source Linux distributions, such as found in Fedora 11
- pNFS operates similarly to Celerra MPFS in that Hosts obtain direct block data access to the array (FC or iSCSI), via metadata mappings (pNFS) from Celerra file systems using NFS, for file to block-level addresses
- Support defined in RFC 5661 and NFSv4.1, but Celerra will only initially support Host block access to either CLARiiON or Symmetrix arrays using iSCSI or FC, not file or object access
- Storage Access protocols from Host Clients to array are FC or iSCSI for direct block access
- Control Protocol from Celerra to array is done using FC
- pNFS file mapping Protocol operates between Host Clients and Celerra
- Supported only for Celerra CLI
- Fedora 11 would represent a typical pNFS client
- Statistics for pNFS would be read from NFSv4 statistics
- No MPFS licensing required, only the regular NFS license
- Celerra pNFS is an EMC implementation of an Open Source standard using pNFS based on NFSv4.1, Block FC or iSCSI storage access protocols, and Fibre Channel control protocol, and has the FMP protocol built into the pNFS standard

### **Operational Overview:**

- LINUX clients using NFSv4.1, and pNFS RPMs, and either iSCSI or FC connectivity to the Array
- LINUX clients access data via NFS Exports from Celerra, or directly from array using FC & iSCSI
- Data Mover servers as metadata server to mapping data blocks for direct Host access to array
- iSCSI or FC network between the Hosts and the Array for direct data block access
- Celerra NFS and NFS protocols for file access permissions

**\$ server\_nfs server\_2 -pnfs -service -start | -stop**

### **pNFS CONFIGURATION:**

#### **I. Celerra Configuration Steps:**

1. Enable NFSv4.1 & pNFS on the Data Mover by adding “hivers=4” to “config” file line with nfs config:

**# vi /nas/server/slot\_2/config**

**nfs config hivers=4** →Add hivers=4 to the nfs config line

2. Reboot Data Mover and verify that NFSv4.1 and pNFS Services have started:

**# server\_nfs server\_2 -v4**

server\_2 :

----- nfsv4 server status -----

\* Service Started \*  
\* pNFS service Started \*

----- NFSv4 Clients -----

Confirmed Clients : 1  
UnConfirmed Clients : 0  
Domain Name : Not Defined

----- NFSv4 State -----

Opens : 0  
Locks : 0  
Delegations: 0  
Layouts : 0

3. Create Meta volume from a single dvol for each file system to be created:

Storage > Volumes > Create >Type: \*Meta

**Note:** Select an available dvvolume to create the ‘meta’ volume from for the pNFS file system

4. Create a file system from the Meta volume:

Storage > File Systems > Create > Create from: Meta Volume

5. Mount the file system for pNFS protocol:

**# server\_mount server\_2 -o pnfs pnfs /pnfs**

server\_2 : done

# **server\_mount server\_2**

pnfs on /pnfs uxf,perm,rw,pnfs

**6. Export the file system to Linux host for root access:**# **server\_export server\_2 -P nfs -o root=10.241.168.184 /pnfs**

server\_2 : done

**II. CLARiiON Configuration Steps:**

1. Configure SP iSCSI SLIC ports with IP addresses for Host access
2. Create Raid Group and two LUNs (for this example) for the Celerra file systems
3. Add LUNs to Celerra Storage Group
4. Rescan or run diskmark operation on Celerra to pull new storage volumes into the Celerra database
5. Register Linux Host iSCSI initiator records on the array  
--Obtain IQN from /etc/iscsi/initiatorname.iscsi file  
--Set 'Initiator Type' to CLARiiON Open with Failover Mode: failovermode 4  
--Add Host name & IP address in the "Host Agent Information" section
6. Create Fedora Linux Host Storage Group & add Celerra LUNs previously added to Celerra Storage Group  
--Assign the HLU values as 0 & 1 for the Fedora Host  
--From Hosts tab, move the Linux Host from 'Available Host' to 'Hosts to be Connected'

**III. Fedora Host Configuration Steps:****1. Load the iSCSI Initiator Utilities:**# **yum install iscsi-initiator-utils**

Loaded plugins: presto, refresh-packagekit

|                    |              |
|--------------------|--------------|
| updates/metalink   | 16 kB 00:00  |
| updates            | 4.5 kB 00:00 |
| updates/primary_db | 5.7 MB 00:04 |

Setting up Install Process

Resolving Dependencies

--&gt; Running transaction check

---&gt; Package iscsi-initiator-utils.x86\_64 0:6.2.0.870-10.fc12.1 set to be updated

--&gt; Finished Dependency Resolution

Dependencies Resolved

---

| Package | Arch | Version | Repository | Size |
|---------|------|---------|------------|------|
|---------|------|---------|------------|------|

---

Installing:

iscsi-initiator-utils x86\_64 6.2.0.870-10.fc12.1 fedora 489 k

Transaction Summary

---

| Install | 1 Package(s) |
|---------|--------------|
|---------|--------------|

---

Upgrade 0 Package(s)

Total download size: 489 k

Is this ok [y/N]: y

Downloading Packages:

Setting up and reading Presto delta metadata

fedora/prestodelta | 1.6 kB 00:00

Processing delta metadata

Package(s) data still to download: 489 k

iscsi-initiator-utils-6.2.0.870-10.fc12.1.x86\_64.rpm | 489 kB 00:00

warning: rpmts\_HdrFromFdno: Header V3 RSA/SHA1 signature: NOKEY, key ID 57bbccba

fedora/gpgkey | 3.2 kB 00:00 ...

Importing GPG key 0x57BBCCBA "Fedora (12) &lt;fedora@fedoraproject.org&gt;" from /etc/pki/rpm-gpg/RPM-GPG-KEY-fedora-x86\_64

Is this ok [y/N]: y

Running rpm\_check\_debug

Running Transaction Test

Finished Transaction Test

Transaction Test Succeeded

Running Transaction

Installing : iscsi-initiator-utils-6.2.0.870-10.fc12.1.x86\_64

1/1

Installed:

iscsi-initiator-utils.x86\_64 0:6.2.0.870-10.fc12.1

Complete!

## **2. Modify and record iSCSI Initiator IQN name:**

# more /etc/iscsi/initiatorname.iscsi

**InitiatorName=iqn.1994-05.com.fedora:6eaf3214e054**

a) vi edit initiatorname.iscsi and add Hostname to last field, as follows:

# more /etc/iscsi/initiatorname.iscsi

InitiatorName=iqn.1994-05.com.fedora:**fedora-202**

**Note:** Save the IQN name for registration on the CLARiiON

## **3. Load and verify the NFS and Block layout drivers on the Host:**

# modprobe nfslayoutdriver

WARNING: Deprecated config file /etc/modprobe.conf, all config files belong into /etc/modprobe.d/.

# modprobe blocklayoutdriver

WARNING: Deprecated config file /etc/modprobe.conf, all config files belong into /etc/modprobe.d/.

# lsmod

Module           Size Used by

blocklayoutdriver 36800 0

nfslayoutdriver 18423 0

nfs 353047 2 blocklayoutdriver,nfslayoutdriver

## **4. Start iSCSI Service and perform iSCSI Target discovery:**

# iscsidadm -m discovery -t st -p **10.127.62.234**

Starting iscsid: [ OK ]

10.127.62.235:3260,2 iqn.1992-04.com.emc:cx.fnm00100900045.b4

10.127.62.234:3260,1 iqn.1992-04.com.emc:cx.fnm00100900045.a4

## **5. Remote mount Celerra pNFS file system and verify:**

# mkdir /pnfs

# mount -t nfs4 -o minorversion=1 **10.127.62.126:/pnfs1 /pnfs**

# cd /pnfs

## **6. Verify iSCSI Disk mounts on the Fedora Host:**

# fdisk -l

# grep LAYOUT /proc/self/mountstats

PNFS\_LAYOUTGET: 0 0 0 0 0 0 0 0

PNFS\_LAYOUTCOMMIT: 0 0 0 0 0 0 0 0

PNFS\_LAYOUTRETURN: 0 0 0 0 0 0 0 0

# mount

10.127.62.126:/pnfs1 on /pnfs type nfs4 (rw,minorversion=1,addr=10.127.62.126,clientaddr=10.127.62.202)

## **7. Copy data to the remote mount and observe on Celerra using server\_stats to verify operation**

# dd if=/dev/zero of=/pnfs/file\_75GB bs=1024 count=**75000000**

# server\_stats server\_x -monitor nfsOps-std

# server\_nfs server\_2 -v4

----- NFSv4 State -----

Opens : 1

Locks : 0

Delegations: 0

Layouts : 1

**Note:** If the Opens and Layouts values were 0, it would mean that the Data was transferred via the Data Mover and not NFSv4

## **CST SUPPORT:**

→RSA's CST 2.0 Common Security Toolkit support for LDAP authentication

**Note:** Replaces existing CSO libraries

→Purpose of CST for Celerra is for authenticating Remote LDAP users

## **Troubleshooting Remote Logins and LDAP:**

/nas/sbin/cst\_config utility and debugging of ldap communications by setting “–debug on” | -debug off

**Note:** Check /var/log/secure and /nas/http/logs/error\_log

/nas/site/cst

/nas/http/logs/error\_log

/nas/log/webui

/var/log/secure

/var/log/messages

Jul 2 13:57:04 sleet-480 sshd[13718]: Accepted password for nasadmin from 10.4.146.212 port 1230 ssh2

Jul 2 15:24:02 sleet-480 Unisphere: Authentication succeeded for user nasadmin1 from 10.4.146.212

## **CEPA SUPPORT FOR NFS AUDITING 4.6.6.2 (Celerra Event Enabler):**

- Celerra Event Enabler version 4.6.6.2
- Supports IPv6
- Implements auditing of NFS postevents
- Third party tool for auditing from DatAdvantage (Varonis)
- CEPA NFS events initiated via Celerra mount options  
--ceppnfs or nocepp

### **Events audited for NFS are:**

FileRead; FileWrite; CreateFile; CreateDir; DeleteFile; DeleteDir; RenameFile; RenameDir; SetSecFile; SetSecDir

- Server\_mount options for using CEPA for CIFS & NFS

# server\_mount |grep cepp

```
[ ceppcifs | ceppnfs | ceppcifs,ceppnfs | nocepp ]
```

## **6.0 MISCELLANEOUS FEATURES:**

- VSS 2.0 support iSCSI
- VDS 1.1 iSCSI support
- MPFS Client support Windows 2008 R2 SP2 64-bit; Windows 7; RHEL5; CentOS 5.4 (over iSCSI only); SLES 10 SP3 & 11
- MAC O/S 10.6 support for CIFS SMB1 clients (already support NFSv3)
- Celerra iSCSI LUN to CLARiiON iSCSI LUN migration procedures
- Prepare Celerra plug-ins, CSA, Registration Wizard, Install Manager, PAHC & PUHC for full Internationalization of multibyte character strings and XML or Java properties [Barossa maintenance release will then provide Language packs for various interfaces]
- Tolerance for 8Gb FC cards on V-Max, Enginuity 5874.207.168 and later

## **BAROSSA REFERENCES:**

- Powerlink link to NAS 6.0 code & docs—Software Downloads > Celerra Software
- Unisphere Domain Functionality white paper (not yet posted)
- Unisphere Client download under Software Downloads > T-Z
- Partners Web site for CLARiiON
- SABA 6.0 Training materials
- 6.0 and VG2/VG8 PSBs
- 6.0 Release Notes
- Documentation—Customer doc pdfs located on same page as Celerra Software downloads on Powerlink:  
Celerra\_Documentation.zip

### **Following path leads to both User and Customer Service Editions of Documentation:**

Support > Technical Documentation and Advisories > Hardware/Platforms Documentation > Celerra Network Server > General Reference

- Primus articles
- Location of Celerra Brownbag presentations: emc233640, download from link
- EMC One Network—Celerra & CLARiiON links to various docs

<http://one.emc.com/clearspace/docs/DOC-10122>

### **USM Help:** (Available offline)

--Launch the USM application, and while at the login screen, click on the upper righthand questionmark icon to launch the Help pages

### **Unisphere Help:** (Available offline)

--Launch the Unisphere Client application, and while at the login screen, click the icon on the lower right side of the screen for 'Help', rightclick and answer the 'allow blocked content' popups, and on the Unisphere help screen, click on the upper left link that says, 'Show TOC'. You should now have access to all of the built-in Help content.

### **eLearning Links:**

NAS 6.0 - Security, Auditing, and OS Support : <https://learning.emc.com/Saba/Web/Main/goto/361843361>

Celerra VG2 and VG8 Architectural Overview : <https://learning.emc.com/Saba/Web/Main/goto/361719023>

Celerra VG2 and VG8 Technical Presales : <https://learning.emc.com/Saba/Web/Main/goto/363306161>

## **BAROSSA BMR1 6.0.40 MAINTENANCE RELEASE:** ETA Dec 2010

→ The main feature in this release is Celerra tolerance for Jupiter Flare 30.5, and FCoE I/O Modules on the SPs

→ Additionally, the release offers Japanese, Chinese, and Korean Unisphere Localization, and Unisphere Language packs

## **EMC CELERRA FAST (Fully Automated Storage Tiering) SOLUTION:**

### **What is it?**

→ Celerra FAST is a “solution” based on File Tiering (aka archiving) using the Rainfinity FMA (File Management Appliance) policy-based file utilization engine, Celerra FileMover API, and various Storage tiers as potential ‘resting places’ for inactive/unimportant data

→ Celerra FAST is based on file archiving (Marketing prefers the term ‘tiering’ vs. ‘archiving’), whereas CLARiiON FAST is block-based, using CLARiiON LUN migration technology

**Note:** There are potential DUDL scenarios here, since any LUN not using the Celerra/CLARiiON LUN naming convention (5.6.39+), risks being migrated under the CLARiiON FAST solution

→ Celerra FAST is not technically tied to the NAS 5.6.47 release, or even FMA 7.3, but for Marketing/Sales/Program Mgmt reasons, is being rolled out as a “solution” based on CMR10 & FMA 7.3 or FMA/VE ESXServer 3.5, via Direct/Channel Express

→ Celerra FAST is not customer-installable

→ Celerra FAST can be viewed as either ‘drive-based’ or ‘platform-based’ tiering

### **Why use it?**

→ Reduce costs--companies find it difficult to keep up with ever-expanding storage needs and costs, and see automatic storage tiering solutions as a way to ease the management costs and reduce capital expenditures

→ Create operational efficiencies--Tiering strives to optimize storage usage (hence costs) by keeping the most “active” data on the more expensive Primary Storage tier, while moving less active data to lower cost tiers. Tiering can reduce backup times and reduce administrative costs because it’s automated.

→ Improve service levels--Besides optimizing storage usage, Tiering can improve service levels & performance for important applications

### **How does it work?**

#### **Example:**

→ The Primary storage tier might be hosted by Celerra on EFD flash-drive luns. Using FMA File Archiving policies, along with file matching rules (access & modification times, size, filenames and/or extensions, etc.), filesystems are scanned, and files are moved first to a Secondary storage tier (FC drives), and then using Multi-tiering file & stub policies, with additional file matching and archive destination rules, files are moved again to a final Tertiary storage tier (SATA drives).

→ Consideration of various Celerra features needs to be taken into account, such as Deduplication (Secondary tiers only), File Retention for stub protection, Celerra PAX NDMP backup, Celerra Snapsure, & IP Replication.

→ The Celerra FAST solution is based on FMA 7.3, and involves the archiving (using FMA) of aging/inactive data from the high-cost Primary Tier to lower cost Secondary (FC Drives) and/or Tertiary (SATA drives) tiers, which could be Celerra, Centera, Windows, NetApps, or Atmos targets.

→ Celerra FAST is a linear solution going in one direction only (Highest to Lowest cost tiers). CLARiiON LUN-based FAST is also a one-way migration, but uses NaviAnalyzer tools to find “hot” lun candidates to migrate from FC to EFD drives, or “cold” luns from FC to SATA drives. Symmetrix FAST is bi-directional, and moves volumes back and forth from different drive tiers, as required.

### **File Retrieval Behaviors:**

→ With Celerra or Data Domain secondary or tertiary storage, file retrievals are ‘in-band’, meaning that the FileMover API handles the recall of files to Primary Celerra storage

→ With Centera or Atmos as secondary or tertiary storage, file retrievals are “out-of-band”, meaning that it requires the use of the FMA to retrieve the files

### **Celerra FAST Components:**

Celerra FileMover API (NAS 5.6.47)

Rainfinity FMA version 7.3 or FMA/VE version 7.3

Primary, Secondary, and optional Tertiary Storage tiers

**Note:** Tiers involve Primary Storage on Celerra, and archiving to other tiers [Celerra, Centera, Atmos, Windows, NetApps]

### **Planning Considerations for using FAST:**

Dataset should involve large files (stubs are 8kb, so >than 8kb for archiving)

Dataset should have a reasonable number of inactive files (if all files were active, solution would not work)

Be aware of FMA archive file limits (75 or 250 million)

Select a migration strategy & backup methodology

Determine single cabinet or multi-cabinet solution

NFS supports only single Tier archiving, CIFS supports multiple tiers

### **Migration Best Practices:**

--Single Tier Migration involves copying dataset to Primary Tier, then archiving in chunks to the Secondary Tier

--Multiple Tier Migration involves copying dataset to a temporary location, archiving from that location, then copying the remaining files and stubs to the Primary Tier using FVA, or some other stub-aware migration tool

### **Backup Best Practices:**

--Backup complete dataset using automatic passthrough reads so as not to “recall” files from Secondary tiers

--Backup Primary and Secondary Tiers separately using stub-aware backups

### **Replication Best Practices:**

--Replicate Primary and Secondary Tiers to separate destinations as independent sessions to better protect stubs and all the data

### **Overall Celerra FAST Best Practices:**

→Use NAS 5.6.47

→If using IP Replication in the environment, consider using ‘delayed stubbing’ policy with FMA in order to let Replication complete to the Destination before stubbing files on the Source

→If using IP Replication, consider replicating Primary Tier and Secondary Tier to separate locations to maintain stubs and content

→If stub protection is required, use File Retention and FMA Retention policy on file systems

→Consider using automatic file system extension in combination with Virtual Provisioning and Deduplication (the latter two on Secondary tiers only)

→Single Tier Migrations involve copying dataset to Primary tier, then archiving in chunks to Secondary tiers

→Multiple Tier Migrations copy dataset to a temporary site, archive from that site to a Secondary tier, then copies remaining files & stubs from temp site to the Primary tier using stub-aware migration tool

→Use Backups with auto passthrough reads so as not to “recall” files from Secondary tiers, or Backup Primary & Secondary tiers separately using stub-aware backup

#### **Other Information:**

→FMA accesses Celerra FileMover service using TCP port 5080

→Stub files reside on the Source Tier, and would be updated by the FMA engine as the file contents are moved from Tier to Tier

→When Retention Periods are used with files (WORM), stub protection is offered with NAS 5.6, preventing Stub files from being accidentally deleted or modified (FMA 7.3 only supports stub retention with NAS 5.6 Celerra to Centera archiving)

→Primary Storage Tier space requirements would need to support “active” files + 8kb Stub files used in the tiering process

→Celerra Deduplication is recommended for use with the FAST solution, but only for Secondary or Tertiary storage tiers, not on the Primary storage tier

→NFS files only support single Tiering, whereas CIFS supports Multi-tiered FAST

→With Rainfinity FMA version 7.3 (Q4 2009), Celerra FAST is a multi-tiered storage solution that consists of Celerra, Atmos, Centera, or Windows storage in a variety of Tier A, B, or C configurations

→While Centera or Atmos can be in Tiers B or C, further archiving cannot be done from them, only to them, so they are endpoints

**Note:** In otherwords, there is only a 3<sup>rd</sup> Tier if Celerra is used for Tier A and Tier B. If Centera or Atmos were in Tier B, there would be no Tier C. Celerra filermove is what pushes/pulls files from Tier to Tier.

→Solution leverages EFD, FC, and 2TB SATA drives, as well as NAS Deduplication feature

→Solution uses Rainfinity FMA and FMA/VE policy engines

#### **RAINFINITY FMA 7.3 & FMA/VE (ESXServer 3.5):**

##### **How does it work?**

→The FMA policy engine basically walks through the source file system to identify files to archive, and/or stub files to relocate. Files that match defined rules and policies will be archived to the desired location. Stub file offline locations are compared to a destination rule, and moved if the path differs from the rule. The offline path in the stub is then updated with new destination path, FMA database is updated, stub file is updated, and previous location of file data is deleted

→Rainfinity File Management uses a GUI to Configure storage platforms (Celerra, Netapps, Centera, Windows, Atmos), create Policies and file rules to perform archiving, and a Schedule function to setup tasks to execute policies at scheduled times, or on-demand, either as a Simulation or actual archiving. There is also an Archived Files function to view Simulated or archived files by list or by log activity

→Raininfinity also employs a CLI capability

→Stub files always reside on the Primary Tier. Stub files would be updated if the file contents were migrated from Secondary to Tertiary tiers, and then after a delay period, the previous file content’s location would be deleted

→FMA also allows the ability to migrate an entire archive set (Repository) from one tier to another

##### **Other Information:**

→FMA 7.3 supports a new policy called “multi-tier” archiving—file matching rules & archive destinations are used to relocate data across multiple tiers (e.g., Tier 1 Primary Storage, Tier 2 Secondary Storage, Tier 3 Tertiary Storage), scanning both files and stubs

→FMA 7.3 supports a new policy called “multi-tier stub policy”—this policy scans only stub files from already archived files (via the stub files located on the Source), and then migrates data from Secondary to Tertiary tiers

→FMA 7.3 uses a new “simulation” feature to evaluate the effects of a policy—essentially runs a real-time scan with reporting, without actually executing archiving [Previous feature was “preview” which only did this against snapshot metadata, not real-time]

→FMA 7.3 can migrate a NAS repository (Secondary archive tier) to another Tier NAS, Centera, or Atmos repository

→FMA 7.3 supports archiving using NAS 5.5 or 5.6

→FMA 7.3 now supports delayed stubbing when archiving from NAS to NAS—this would be used in situations where the archived tier was also being replicated to another Celerra location. Purpose of the delayed stubbing would be to not actually create the stub file on the Source system until Replication has completed. The idea here is that if the Archive Celerra were to become unavailable before Replication finished, the data itself would be orphaned.

→FMA 7.3 supports stub retention when archiving from Celerra running NAS 5.6, to Centera only

→FMA 7.3 supports Orphan and Stub delete policies that can be run Daily, Weekly, Monthly, or as Scheduled

**Note:** An orphan file is a file on Secondary storage without a stub file to reference it. Orphan files are identified via stub scanning. Orphans can be set to be automatically deleted, or in the case of Retention Files, deleted only when retention period has expired.

→FMA version 7.3 also allows the use of a single set of credentials for callback operations for all Celerras, rather than the previous requirement that each Celerra to have defined IP addresses (called IP Filtering)

→FMA archiving is based on file matching expression policies that are initiated by Request, Schedule, or Capacity triggers

→FMA 7.3 supports Atmos as an archive destination using Atmos API REST to write or recall data

**Note:** Atmos is EMC's cloud computing solution. Cloud computing is a model delivering convenient, on-demand network accessed to a shared pool of resources that are centrally managed but globally available.

#### **FMA 7.3 LIMITATIONS:**

→Windows 2003/2008 support is limited to only NTLM, no Kerberos, and no SMB Signing support

#### **How does a multi-tier archive policy work?**

1) FMA walks file system to identify files to archive, or stubs to relocate, based on defined policies

2) If a stub matches a rule, offline path is compared to archive destination in the rule. If it matches, nothing happens. If it does not match, then the archived file is copied to new destination. If a file matches a rule, it is archived to the destination specified, and the source is replaced with the stub file.

3) Offline path in the stub is updated with the new destination and in the FMA database

4) Location of archived file is deleted after specified delay period

#### **FMA 7.3 vs. FMA/VE 7.3:**

→FMA supports Celerra or NetApps at Primary Tier, FMA/VE only supports Celerra

→FMA supports Celerra, NetApps, Windows 2003/2008, Centera, or Atmos at the Secondary Tier, FMA/VE supports only Celerra/Centera

→FMA supports Celerra, NetApps, Windows 2003/2008, Centera, or Atmos at the Tertiary Tier, FMA/VE supports only Celerra/Centera

→FMA 7.3 increases file limit from 200 to 250 million, and FMA/VE 7.3 increases file limit from 50 to 75 million

#### **FMA SOURCES:**

Celerra and NetApp systems

#### **FMA TARGETS:**

Celerra, Centera, NetApp, Windows 2003/2008, Atmos

#### **FMA/VE SOURCES:**

Celerra

#### **FMA/VE TARGETS:**

Celerra and Centera

#### **FMA 7.25 vs. 7.3**

| <b>FMA Version 7.25</b>             | <b>FMA/VE Virtual Edition</b> | <b>FMA</b>                                |
|-------------------------------------|-------------------------------|---|
| Single Tier                         | VMWare ESX 3.5 + SW           | FMA 2U HW appliance + SW                  |
| Sources                             | Celerra                       | Celerra, Netapps                          |
| Targets                             | Celerra                       | Celerra, Netapps, Centera                 |
| File Limits                         | 50 million                    | 200 million                               |
| <b>FMA 7.3/VE (Virtual Edition)</b> |                               | <b>FMA 7.3</b>                            |
| Platform                            | VMWare ESX 3.5 +SW            | FMA 2U HW appliance + SW                  |
| File Limits                         | 75 million                    | 250 million                               |
| Sources                             | Celerra                       | Celerra, Netapps                          |
| Targets                             | Celerra, Centera              | Celerra, Netapps, Centera, Atmos, Windows |

**Notes:** Centera & Atmos are final tiers, data cannot be tiered from them. FMA 7.3 supports archiving from 5.5 or 5.6 Celerra. FMA 7.3 is a multi-tiering solution, whereas 7.25 is single tier only

#### **EMC File Management Appliance--FMA Version 7.3.1 'Denver':**

→One of the main purposes of this release is to bring FMA/VE into parity with capabilities of FMA

→The 'Rainfinity' name is officially dropped with this release

→Adds support for NetApp as an archive source FMA & FMA/VE

→Adds Data Domain, Windows, and NetApps as archive targets for FMA & FMA/VE

→Adds Data Domain as a NAS Repository

→Adds support for 250 million archived files for FMA/VE (from 75 million with 7.3)

→Adds support for (4) virtual CPUs (from 2 CPUs)

→Adds virtual support for 500GB HD (from 150GB)

→Support for NetApp OnTap 7.3 pass-through recall in FMA and FMA/VE editions (did full recall of files prior to 7.3)

→Data Domain is treated as a standard NFS archive, and NFS NAS Repository

→Data Domain supports stub management, orphan deletion, multi-tiering, and NAS migrations

→Adds support for NetApp as a (OnTap 7.3) source for Atmos destination archive, with full & pass-through recall support

→Provides upgrade path from FMA to FMA/VE

#### **TARGET SYSTEMS SUPPORTED BY 7.3.1:**

Celerra; Centera; Atmos on Premise; DataDomain; NetApp; Windows

#### **SOURCE SYSTEMS SUPPORTED BY 7.3.1:**

### **FMA Version 7.3.2 ‘Durango’:**

- Main enhancement will be the ability to convert DX/NAS stub files to FMA stub files without requiring recall
- DX NAS version 3.1 will use a Windows FileWalker executable to convert stubs one file system at a time
- DX NAS conversion does not support Deduplication or File Retention features
- Support for Atmos OnLine solution
- Supports Gen5, Gen6, & Gen7 hardware platforms

### **RAINFINITY FMA RULES:**

- Archiving based on Last Accessed Time, Last Modified Time, File Size, File Name/Extensions, High Watermarks
- Scheduling based on Now, Daily, Specific Day/Time of Week/Month, Once at future Day/Time, When capacity threshold reached
- Delete policy for orphaned files and when archived files retention period expires

### **RAINFINITY FMA/VE HIGH AVAILABILITY:**

- requires Virtual Center Server
- requires cluster of ESX servers
- requires NFS, iSCSI, or FC SAN Shared Datastore

**Note:** Virtual Center would monitor the ESX Servers and failover all VMs to another ESX Server if there was a problem, which requires that the datastore be shared among the ESX servers

### **Celerra FAST Best Practices:**

- Do not use deduplication on the primary tier
- Do not use the automatic file extension feature on the primary tier
- Disable Read recall policy on the primary tier and limit the age of snapshots
- Enable deduplication on Secondary tiers, with accessTime and modificationTime parameters disabled, set to 0
- Enable virtual provisioning on Secondary tiers, and allow space for disk growth
- Enable automatic file extension on Secondary tiers
- Delete orphan files from Secondary tier after 30 days
- Disable client access to Secondary tiers
- When possible, implement FAST solution first, then setup IP Replication between systems
- Celerra stub files can be read using /nas/tools/dhsm/get\_attributes

```
$ /nas/tools/dhsm/get_attributes -u rsadmin -p Stud3nt123 -w server_2 '/podE/LICENCE.TXT'|egrep
```

```
"WORM_STATE|RETENTION"  
    WORM_STATE="Worm"  
    RETENTION="1259176994"
```

### **IP Replication Recommendations for Celerra FAST:**

- If using IP Replication, the recommendation would be to have separate replication Sessions for Primary and Secondary tier file systems to protect the content on both Primary & Secondary tiers

### **Backup Recommendations for Celerra FAST:**

- The primary recommendation would be to backup the source file system using automatic passthrough Reads so that archived files from other Tiers are not recalled to the Primary tier (read\_pass\_through, which is the default anyways)

**Note:** Full restores, however, could be problematic as they would restore active and contents of the stub files to the Primary tier, which would likely exceed its capacity. Another disadvantage is that the backup window is larger since it must backup everything.

- If backup Windows are a big concern, could use a stub-aware backup solution to backup Primary & Secondary tiers separately, though this introduces special considerations when doing Restores to ensure that stub files retain proper links to the secondary tier.

### **Case of the orphaned file when using SnapSure:**

1. Create file\_1 on primary file system, then archive to secondary tier file system, which leaves source file\_1 as a stub file
2. Run a Snapshot of the primary file system, then delete the Stub file for file\_1
3. Create file\_2 on primary file system, then archive to secondary file system, which leaves stub file\_2 on source
4. Run FMA multi-stub policy to update stub files, and file\_1 will be listed as orphaned
5. Restore snapshot taken in step 2, run stub scanner update, and now file\_1 will show as archived, while file\_2 will be orphaned (since the snapshot restore overwrote the file\_2 stub)

**Note:** One of the points of this exercise is to show that file\_1's data does not get deleted just because the stub is deleted, hence it is accessible after the checkpoint restore

### **CLARiiON FAST LUN MIGRATOR (aka FASTLite, FAST I) Q4 2009:**

- FAST LUN Migrator is a feature which can identify and migrate LUNs that can benefit from Storage Tiering
- Migration scenarios are for moving Hot LUNs to EFD drives, or Cold LUNs to SATA drives
- CX4 platforms running FLARE 29 for LUN migration (though Flare 28 can do the lun analysis)
- Requires FAST LUN Migrator tool installation on a Windows platform, and a FAST LUN Migrator Enabler UPF install package on the array
- NaviAnalyzer and NAR archive files are used to generate the report of “Hot” and “Cold” luns for recommended migration. From the list of recommended LUNs, the CLARiiON FAST LUN Migrator tool will then automatically move luns to either a higher performance drive (EFD) for ‘Hot’ luns, or to archival drives (SATA) for ‘cold’ luns—outputs to XML and CSV

**Note:** Analysis Phase (Windows Host) & Migration Phase (CLI)

→LUN migration recommendation list must be edited to add destination RAID Group, and is used by the Migration Utility as an input file to actually perform the migration

**Note:** The User must kick off the actual migration

→CLI only for the migration piece

# fastlunmigrator lunanalyze <narfile> | fastlunmigrator –address <IP\_addr> lunassist –start –file <LUNID\_RGID\_list> -rate low | medium | high | asap | -cancel | -status

→CLARiiON FAST LUN Migrator tool will exclude LUNs with name “Celerra” from the report and .csv file. During migration, will also check and exclude any luns with name “Celerra” in the .csv input file

## **COMPARISON OF EMC FAST SOLUTIONS:**

### **CELERRA FAST:**

→Handles file-based migrations only, using Rainfinity GUI or CLI (create policy, defined rules, save policy)

→Demote across EFD, FC, or SATA drives on Celerra, Centera, or Atmos with Rainfinity FMA 7.3

→Archiving is one-way, from EFD to FC, FC to SATA, EFD to SATA

### **SYMMETRIX FAST:**

→2-way migrations allowed

→Handles Volumes for Mainframe CKD or FBA Open Host LUNs using SMC or CLI interface

→Can Promote or Demote luns (Bi-directional) across EFD, FC, or SATA drives

→Symm FAST supported with Celerra CMR11

### **CLARIION FAST:**

→1-way migration only

→Uni-directional LUN-based migration using CLI

→Not fully automatic, input file of lun candidates required to run the migration

→Migrates “Hot” luns from FC to EFD, and “Cold” luns from FC to SATA

→CLARiiON FAST will not recommend Celerra LUNs for migration if Celerra LUNs follow the 5.6.39 naming convention: e.g.,

[Celerra\\_sleet-120\\_17\\_d8](#)

### **CELERRA FAST:**

→1-way multi-directional tier-to-tier migration

→Uni-directional file-based migration, using Rainfinity FMA 7.3 GUI/CLI

→Demote files across Celerra, Centera, Windows, Netapps, or Atmos tiers

→Archiving is intended to be one-way

→A single Celerra cabinet solution might consist of EFD to FC, FC to SATA, or EFD to SATA archiving [marketing prefers the term ‘tiering’]

## **CELERRA and VMWARE**

→ESX Server is a physical platform that hosts the ESX Server and Virtual Machines (Guest Operating Systems)

→ESX Server can boot from a local drive or from iSCSI LUN or FC LUN on a storage array

**Note:** Requires installation of iSCSI or FC HBA on ESX

→VMWare Virtual Machines are instances of Operating Systems running in a virtualized environment on a single physical computer, usually one installed with an ESX Server. The physical machine is represented by software, with virtual RAM, CPU, NIC, Disks.

→The ESX Server abstracts processor, memory, storage, and networking resources for use with Virtual Machines (Virtualization)

→ESX Server installs a hypervisor layer as the virtualization layer which interacts directly with the underlying hardware on behalf of the subsequent Guest VMs or OSs that are installed later

→VMFS are high performance cluster file systems for ESX, defined by a set of encapsulated files, and stored on physical disks and partitions

→VMIInfrastructure Client used to connect to ESX Server, and also within the Virtual Center framework for multiple ESX Servers

→Virtualization allows one to install & run multiple O/S’s on a single hardware platform as VMs, and transfer the VM from one ESX Server platform to another

→An ESX Server sees each VM as a compilation of files (configuration file, virtual disk file, NVRAM Bios file, Log file)

→VMs do not directly access the underlying hardware

→Another method of using Virtualization is to use Host-based operating systems to install the VMWare Server or Workstation environment, which requires that the underlying ‘Console OS’ be installed first (Linux, Windows, etc)

→Datastores identify the specifics about the Storage that is being used by each VM machine

→Datastores can be created as NFS, iSCSI, or FC SAN Shared Datastores

## **OPTIMIZING VMWARE ESX SERVER WITH CELERRA**

→Celerra supports VMWare for CIFS, NFS, iSCSI, or FC protocols

→VMFS (Virtual Machine File Systems) are used to create the Datastore for the particular protocol to be supported

→Celerra recommends aligning VM storage partitions to 8kb tracks when using Block-based storage (iSCSI) for Windows VMs

**Note:** Partition Starting Offset would be set at 65,536 bytes—do not use 64kb alignment for NFS

→Use uncached option for Celerra file systems for NAS environment if IO is random and mixed workloads

**Note:** Server mount option “uncached”

→Consider disabling Celerra prefetch for NFS random and mixed workloads

**Note:** Server mount option “noprefetch”

→Use a hardware iSCSI HBA on the ESX Server for Celerra iSCSI if using sequential workloads (otherwise software iSCSI fine)

### **INSTALLING ESX SERVER ON CELERRA iSCSI LUN:**

1. Install the iSCSI HBA software (Qlogic)
2. Boot system and enter HBA BIOS to obtain HBA’s IQN name
3. Create Celerra file system; iSCSI Target; iSCSI LUN; iSCSI LUN Mask; Configure access to IQN name of HBA iSCSI Initiator and iSCSI LUN number; Configure CHAP; Enable iSCSI service on Data Mover
4. Configure iSCSI HBA in BIOS for IP address, mask, gateway; CHAP name and password; Target iSCSI Name, IP Address, and Boot LUN; set adapter boot mode to manual; reboot the system
5. Enter system BIOS and set iSCSI LUN to be the first Boot device
6. Install ESX Server onto the iSCSI LUN (iSCSI device driver is on the VMWare ESX install disc)

**Note:** MaxRequestHoldTime should be set to 600 seconds on Windows clients to tolerate Celerra reboots/failover

### **INSTALLING ESX SERVER ON CELERRA FC LUN:**

1. Install FC HBA software
2. Boot into QLogic BIOS and configure to boot from the FC LUN; set boot mode to manual; reboot system
3. Create RAID Group, Bind LUNs, assign to Storage Group; add ESX Hostname & Control Station IP to register
4. Reboot Host ESX system and enter HBA BIOS and select FC LUN as primary boot device, noting WWPN
5. Enter system BIOS and set iSCSI LUN to be the first Boot device
6. Install ESX Server onto the FC LUN

→The above ESX installation procedures are rough and untested, and are probably not completely in the correct order

### **VMWARE CELERRA PLUGINS:**

#### **ECM Celerra VMWare View Deployment Plug-in:**

→NAS 5.6.42, ESX 3.5, Virtual Desktop Manager 2.1 or VMWare View 3.0, plugin installed on Windows 2003 SP2 VDI Installation Server

#### **EMC Celerra VMWare vCenter SRM Failback Plug-in:**

→vCenter Server 2.5, VI Client, Site Recovery Manager Server with Site Recovery Manager 1.x & EMC Celerra Replicator Adapter 1.0

#### **EMC Celerra Plug-in for vCenter:**

→NAS 5.6.48 plug-in for provisioning NFS datastores, performing fast or full clones of VMDK files, compressing or uncompressing VMs in NFS datastores

→Install plug-in on a vSphere Client

**Note:** See 5.6.48.7 CMR11 section for more details

### **EMC ATMOS COS, Revision 1:**

→GA March 2010 using Dell R610 Server, Titan 40U or 44U (Dense) Racks, SAS ports, 1TB 7200rpm or 2TB 5400rpm drives

### **CSA CPW (CELERRA PROVISIONING WIZARD) (5.6.43.x CMR6):**

→Powerlink release is targeted for Feb 2009 as 5.6.43.84, Internal EMC and Partners only for first release

→CPW to be integrated into the NAS code branch for CMR9 release (July 2009)

→CSA\_CPW will support NX4, NX4FC, NS20, NS40, NS-120, NS-120FC, NS-120iS, NS-480, NS-480FC, & NS-480iS systems

→Designed as GUI wizard for Customer, Partner, EMC basic implementations

→Provides for CIFS Share & iSCSI Lun creation and attachment to Host, and NFS Export

→Adding storage provisioning capability to CSA tool in support of CIFS, NFS, iSCSI deployments

→Presents an “Install” and “Configure” button on CSA Welcome screen (Install button would be for new install configurations, while Configure button would represent the new CPW provisioning functionality)

→Includes ability to configure several different High Availability configurations for Network Interfaces (FSN, LACP, EtherTrunk)

→Provisioning tool is Host independent and can be initiated from any number of Hosts, multiple times

→Wizard will include context-sensitve Help and ToolTip hover messages

#### **Version Requisites:**

→Systems running NAS 5.6.40 or higher can use the provisioning functionality of the CSA

→Systems running NAS 5.6.39 can run the 5.6.43.84 CSA, but provisioning would not be available, only the Install function

#### **CLASSIC CSA 5.6.43 vs. CPW\_CSA 5.6.43.84**

##### **5.6.43.9 Classic CSA-Powerlink:**

Supports NX4, NS20, NS40, NS-120, NS-480, & NS-960 platforms

This version CSA will support NAS 5.6.40.3 -- 5.6.43.8 NAS Installations on the above platforms

##### **5.6.43.84 CPW CSA—Powerlink, EMC & Partners only:**

Supports NX4, NS20, NS40, NS-120, NS-480 platforms

Supports NAS 5.6.39.5 -- 5.6.43.8 NAS Installations, but only on the above platforms

Supports NAS 5.6.40.3 -- 5.6.43.8 for Provisioning, but only on the above platforms

#### **Other Questions:**

--Either version of CSA can be installed on Windows 2000, 2003, XP, & Vista clients

--Either version of CSA can be used on the branded Dell NX4

#### **CPW BACKGROUND:**

--The Celerra Startup Assistant has been in use now since the introduction of the NS20 in August 2007

--CPW came about for growing need to have tools available for customers & partners to implement Celerra outside of PS

--“Celerra Provisioning Wizard” (CPW) is the name of the enhanced functionality/wizards that have been incorporated into the CSA

--Not considered a complete implementation tool, but geared towards basic storage provisioning, NFS, CIFS, & iSCSI configurations

--The 5.6.43.84 version of CSA is posted on Powerlink, and will support NX4, NS-120, NS-480, NS20, & NS40 systems running NAS 5.6.40 to 5.6.43, limited initially to Internal EMC & Partners use

--The 5.6.45 release integrates the CSA with CPW capability into the regular NAS release cycle, and is qualified for use with NX4, NS-120, NS-480, NS20 & NS40. Original intent was to also support NS-960, but this has moved out to the 5.6.47 release.

--Unlike earlier CSA versions, provisioning capability can be run on any qualified Windows client, repeatedly, as the tool is re-entrant

--Provisioning capability to be formally integrated into the NAS 5.6.46 release [i.e., CMR9] for use by all on platforms that qualify

--Presents new “Install” and “Configure” buttons on CSA Welcome screen

--“Install” functionality the same as the original CSA, intended only for initial configuration & setup of factory pre-configured install

--“Configure” functionality is the new provisioning piece, 4-wizards available after logging into Celerra CS

#### **CSA CPW GUI Look and Feel:**

The application launches the same way as before, but now presents two options on the “Welcome to the CSA” screen: Install & Configure

**Install** →Select this to perform the traditional CSA discovery and basic configuration of a new Celerra system

#### **Basic flow for Install:**

Install flow is similar to current behavior, with Welcome and Initialization screens to discover and select the Control Station MAC address; to input Control Station hostname and IP; option to change system passwords; configuring various network settings and IP addresses for SPs; ConnectHome and Email User feature setup; license screen; Cable Check Healthcheck; Pre-Configuration apply for Proxy ARP and setting backend IPs, and committing configuration changes, including password changes; MPFS Initialization screens if the system is detected as an iSCSI model, where IPs are assigned to iSCSI ports on the SPs and the MPFS service is setup on Server\_2; Select a wizard screen and option to skip wizards and “Register the system”; followed by all the relevant steps to perform the Registration of the Celerra.

**Configure** →Select this option anytime after the initial “Install” and system “Registration” are completed to run Provisioning or implementation wizards. The purpose of this option is to perform system setup for production. The wizard can be run any number of times from any number of different Windows Hosts.

#### **Basic flow for Configure wizards:**

Select the Configure option on the Welcome to the CSA screen>Log in to the Celerra>Select a Configuration Wizard (Provision unused disks; NFS; iSCSI; CIFS; Celerra Manager).

#### **Provision unused disks:**

→Storage Provisioning screen basically presents information on the Celerra system and provides a choice between Express vs. Custom modes. The main difference between Express vs. Custom are the choices available for building R5, R6, or R1/0 system pools.

**Express Mode:** Defaults to provisioning all available disks in a “Capacity” clar\_r5\_economy system pool, though there is an option to select RAID 6 for additional Raid protection (clar\_r6).

**Custom Mode:** This mode allows a User to reserve disks for later configuration. It also allows a combination of two of the following three configuration types: Capacity (clar\_r5\_economy), Protection (clar\_r6), and Performance (clar\_r6). You can select either one or two of the configuration types at a time.

#### **Create NFS export:**

→Use this wizard to select a Blade, select a network device or create a new IP address or configure an HA interface configuration, designate a name for the export (becomes file system and mountpoint name), select the desired storage pool, file system export size, and whether the export is exported to RO, RW, or Root Hosts, etc.

#### **Create/Attach iSCSI LUN:**

→Use this wizard to select a network device, create a new IP or HA configuration, create and attach a new iSCSI LUN to an existing or remote Host, attache an existing iSCSI LUN to a Host, select Blade and target name (iSCSI LUN name), selected desired storage pool, lun size, etc.

**Note:** If getting iSCSI attach failure errors in the CSA, make sure the Client system is running the latest version of the MS iSCSI software initiator (version 2.08). See emc207977. Check startup\_wiz.log for more information on attach failure.

#### **Create/Attach CIFS share:**

→Use this wizard to select a network device (or create new IP/HA configuration), create CIFS share and attach to local Host, or attach an existing Share to a Host, enter Blade name, Storage pool, Share name (file system name) and size, local drive number to use, and whether to use an existing CIFS Server or to create a new CIFS server (which involves screen to enter Compname, Domain FQDN, DNS IP addresses, Domain Admin and password, whether to enable Local User support, and option to select a specific interface IP.

**Open Celerra Manager:**

→Strictly an applet to open up Celerra Manager

**Wizard Choices (Select a wizard screen):**

**Provision unused disks**

**Create NFS export (Unix/Linux)**

**Create/Attach iSCSI LUN**

**Create/Attach CIFS share (Windows)**

**Open Celerra Manager**

**GENERAL CSA/CPW LIMITATIONS:**

→Tool initially available only for EMC & Partners, not customers

→The tool only allows for a single CIFS Share or NFS Export per file system created

→The tool Exports or Shares the file system at the top level, which exposes the .etc and lost+found system directories, which is against best practices (emc208120/AR132379)

→No Summary Report page at the end of each logical operation, such as iSCSI Lun, CIFS Share, or NFS export (AR136153)

→No VDM CIFS support

→For CIFS Shares, you can only map drive letters and not UNC paths

→Cannot upgrade the CSA application (remove old, install new)

→Cannot resize the CSA window (difficult to use on laptops, etc.)

→Cancel button within wizards will close entire CSA application, not current operation. Only way to maneuver between wizards if not completing a wizard is to use ‘back’ button

→No Help menu available when using ‘Install’ CSA functionality

→Single iSCSI LUN per Target (cannot attach multiple iSCSI luns to one target—emc206712)

→NTP service configuration for server\_2 is provided as an option when running the ‘Install’ portion of CSA, but is not offered or enforced in the ‘Configure’ portion of CSA, and even if a CIFS Join were to be successful, without a proper NTP service configuration, the Data Mover time will drift and result in a future CIFS outage (emc206837)

→iSCSI wizard may fail due to incompatible MS Initiator version (latest version 2.08)

<http://www.microsoft.com/downloads/details.aspx?FamilyID=12cb3c1a-15d6-4585-befd1319f825&DisplayLang=en>

→Many iSCSI failure popup messages are not particularly helpful in leading the User to resolve the issue

**CPW Release Schedule for Powerlink Support:**

→Powerlink release scheduled for end of February 2009--UMPW will be later

**CSA/CPW FILES & LOGS OF INTEREST 5.6.43.84:**

**Logging on the Celerra Control Station:**

**/nas/log/cpw**

# ls -la

```
-rw-r--r-- 1 apache apache 778 Jan 29 11:05 CIFS_cifs_share_20090129111405.xml
-rw-r--r-- 1 apache apache 852 Jan 29 09:56 iSCSI_10.241.168.59_10.241.168.172_20090129100521.xml
-rw-r--r-- 1 apache apache 453 Jan 28 11:24 NFSEXPORT_dumble_20090128112059.xml
-rw-r--r-- 1 apache apache 837 Jan 28 15:02 PROVISIONING_123317265432404_168.159.16.45-20090128150142.xml
-rw-r--r-- 1 apache apache 2008145 Jan 29 10:43 startup_wiz.log_USCSMATTATL1C_168.159.16.45
```

**Note:** The above files are representative of logs created on the CSA Client when running the provisioning wizard. Specifically, whenever the CSA is opened, and the “Configure” option selected, all activity is logged in the startup\_wiz.log. Any of the CIFS, iSCSI, NFS, or PROVISIONING wizards will generate a separately named and dated file for the activity, which is then saved on the Client system and also transferred to the Control Station /nas/log/cpw directory using HTTP. Additionally, the startup\_wiz.log is renamed to include the CSA Client hostname and IP address.

**# tail /nas/log/provision.log**

```
=====
/nas/sbin/setup_backend/provision object=clariionProvDisk action=list sessionID=123195975130499
Start pid 31637 Wed Jan 14 14:02:52 2009
=====
```

DRY RUN started

```
=====
clariionProvDisk diskType:string="FC" diskSizeGB:integer="300" total:integer="15" available:integer="4"
=====
```

```
=====
End pid 31637 Wed Jan 14 14:02:52 2009
=====
```

```
=====
/nas/sbin/setup_backend/provision object=clariionProvTemplate action=list sessionID=123195975130499
Start pid 31870 Wed Jan 14 14:03:12 2009
=====
```

DRY RUN started

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
clarionProvTemplate id:integer="1" numDisks:integer="4" storagePool:string="clar\_r10" capacityMB:integer="549556"  
mode:integer="2" diskType:string="FC" diskSizeGB:integer="300" reqSpares:integer="0"  
clarionProvTemplate id:integer="2" numDisks:integer="2" storagePool:string="clar\_r10" capacityMB:integer="274778"  
mode:integer="2" diskType:string="FC" diskSizeGB:integer="300" reqSpares:integer="0"

### Logging on the CSA Client:

C:\Program Files\EMC\CSA\startup\startup\_wiz.log  
C:\Program Files\EMC\CSA\startup\startup\_wiz.log\_xp1\_156.242.10.43  
C:\Program Files\EMC\CSA\test\CSA\_00151777b84e.xml [classic CSA install xml file]  
C:\Program Files\EMC\CSA\test\NFSEXPORT\_nfs1\_20090119174059.xml  
C:\Program Files\EMC\CSA\test\PROVISIONING\_123248750030077\_10.4.134.29-200901191647016.xml  
C:\Program Files\EMC\CSA\test\CIFS\_cpw2\_20090119182817.xml  
C:\Program Files\EMC\CSA\test\CIFS\_join\_20090202102955.xml

**Note:** At the successful conclusion of each CPW wizard, when the ‘Finish’ button is clicked, the respective xml files and startup\_wiz.log files are transferred to the Control Station over HTTP.

### /nas/log/nas\_log.al → commands run by CSA during setup

2009-05-21 06:19:18.351 db:0:18272:S: /nas/bin/nas\_storage -modify FNM00083800203 -network -spa 10.241.168.150 -spb 10.241.168.151  
2009-05-21 06:19:52.866 db:0:18272:E: /nas/bin/nas\_storage -modify FNM00083800203 -network -spa 10.241.168.150 -spb 10.241.168.151  
2009-05-21 06:19:53.019 db:0:18920:S: /nas/bin/nas\_storage -failback FNM00083800203  
2009-05-21 06:19:53.069 db:0:18920:E: /nas/bin/nas\_storage -failback FNM00083800203  
2009-05-21 06:20:09.764 db:0:19766:S: /nas/bin/nas\_cel -update id=0  
2009-05-21 06:20:09.918 server\_2:0:19796:S: /nas/bin/server\_file server\_2 -get passwd /tmp/passwd19774  
2009-05-21 06:20:09.938 server\_2:0:19796:E: /nas/bin/server\_file server\_2 -get passwd /tmp/passwd19774  
2009-05-21 06:20:10.566 server\_2:0:19876:S: /nas/bin/server\_file server\_2 -get passwd /tmp/passwd19850  
2009-05-21 06:20:10.587 server\_2:0:19876:E: /nas/bin/server\_file server\_2 -get passwd /tmp/passwd19850  
2009-05-21 06:20:10.884 server\_2:0:19963:S: /nas/bin/server\_file server\_2 -put /nas/server/slot\_2/passwd passwd  
2009-05-21 06:20:10.918 server\_2:0:19963:E: /nas/bin/server\_file server\_2 -put /nas/server/slot\_2/passwd passwd  
2009-05-21 06:20:11.277 server\_2:0:20040:S: /nas/bin/server\_file server\_2 -get passwd /tmp/passwd20007  
2009-05-21 06:20:11.297 server\_2:0:20040:E: /nas/bin/server\_file server\_2 -get passwd /tmp/passwd20007  
2009-05-21 06:20:11.616 server\_2:0:20135:S: /nas/bin/server\_file server\_2 -put /nas/server/slot\_2/passwd passwd  
2009-05-21 06:20:11.643 server\_2:0:20135:E: /nas/bin/server\_file server\_2 -put /nas/server/slot\_2/passwd passwd  
2009-05-21 06:20:11.999 db:0:19766:E: /nas/bin/nas\_cel -update id=0  
2009-05-21 08:20:19.406 db:0:20839:S: nas\_license -create advancedmanager=19906C74  
2009-05-21 08:20:19.428 db:0:20839:E: nas\_license -create advancedmanager=19906C74  
2009-05-21 08:20:20.039 db:0:20887:S: nas\_license -create cifs=74DC5D75  
2009-05-21 08:20:20.075 db:0:20887:E: nas\_license -create cifs=74DC5D75  
2009-05-21 08:20:23.684 db:0:21128:S: nas\_license -create nfs=4A4A5775  
2009-05-21 08:20:23.710 db:0:21128:E: nas\_license -create nfs=4A4A5775  
2009-05-21 08:20:24.471 db:0:21197:S: nas\_license -create iscsi=397BFF6F  
2009-05-21 08:20:24.500 db:0:21197:E: nas\_license -create iscsi=397BFF6F  
2009-05-21 08:20:25.369 db:0:21272:S: nas\_license -create snapsure=2E625A63  
2009-05-21 08:20:25.397 db:0:21272:E: nas\_license -create snapsure=2E625A63  
2009-05-21 08:20:28.152 server\_2:0:21428:S: server\_date server\_2 timezone -name America/New\_York  
2009-05-21 08:20:28.317 server\_2:0:21428:E: server\_date server\_2 timezone -name America/New\_York  
2009-05-21 08:20:34.905 server\_2:0:21872:S: /nas/bin/server\_date server\_2 090521082032  
2009-05-21 08:20:36.152 server\_2:0:21872:E: /nas/bin/server\_date server\_2 090521082032  
2009-05-21 08:20:42.510 server\_3:0:22359:S: /nas/bin/server\_date server\_3 090521082039  
2009-05-21 08:20:43.181 server\_3:0:22359:E: /nas/bin/server\_date server\_3 090521082039  
2009-05-21 08:20:44.344 server\_2:0:22580:S: server\_dns server\_2 -protocol udp w2k.pvt.dns 192.1.4.248 w2k.pvt.dns 192.1.4.248  
2009-05-21 08:20:44.403 server\_2:0:22580:E: server\_dns server\_2 -protocol udp w2k.pvt.dns 192.1.4.248 w2k.pvt.dns 192.1.4.248  
2009-05-21 08:20:49.676 db:0:23199:S: /nas/bin/nas\_diskmark -mark -all  
2009-05-21 08:21:32.054 db:0:23199:E: /nas/bin/nas\_diskmark -mark -all  
2009-05-21 08:21:34.108 ALL:0:25322:S: /nas/bin/server\_ifconfig server\_2 -upgrade  
2009-05-21 08:21:34.189 ALL:0:25322:E: /nas/bin/server\_ifconfig server\_2 -upgrade  
2009-05-21 08:21:34.808 ALL:0:25322:S: /nas/bin/server\_ifconfig server\_3 -upgrade  
2009-05-21 08:21:34.883 ALL:0:25322:E: /nas/bin/server\_ifconfig server\_3 -upgrade  
2009-05-21 08:38:30.878 server\_2:201:9180:S: /nas/bin/server\_setup server\_2 -P mpfs -option start  
2009-05-21 08:38:30.933 server\_2:201:9180:E: /nas/bin/server\_setup server\_2 -P mpfs -option start  
2009-05-21 08:38:36.683 server\_2:201:9579:S: /nas/bin/server\_mpfs server\_2 -set threads=128

### **CELERRA MANAGER CPW INTEGRATION (UMPW—Unified Manager Provisioning Wizard):**

- Functionality to be released with 5.6.45
- A separate qualification effort called UMPW (Unified Manager Provisioning Wizard) will bring CPW support into Celerra Manager
- Ability to de-provision LUNs, RAID groups, unused volumes, & spares via Celerra Manager for all platforms
- Ability to provision via Celerra Manager for NX4, NS20, NS40, NS-120, NS-480

### **CELERRA CLARIION ARRAY PRODUCTS:**

#### **CLARIION CLI's:**

Clariion uses a Java CLI (navicli.jar—discontinued Flare 26 release), a Classic CLI (navicli.exe—no longer installed by default beginning Flare 26), and Secure CLI (navisecccli.exe—Navi commands built on the latter as of Flare 22, introduced with Flare 19). Secure CLI replaces Java and Classic, outputs in XML, provides N+1 support, legacy Java & Classic support, security model, etc.

### **EMC CLARIION SOFTWARE RELEASES:**

#### **FLARE CODE VERSIONS:**

**Flare 7 →2.01** [first release to support CX200/400/600]

**Flare 11 → 2.04**

**Flare 12 → 2.05**

**Flare 13 → 2.06**

**Flare 13 → 2.06** [first release to support CX300, CX500, CX700]

**Flare 14 → 2.07**

**Flare**

→Patch 707 release Dec 2009, ECO for NAS NX4 is ECO70011; this version supports 4-port BOA FC IO module for CLARiiON, support for 2.5" SAS drives in same enclosure as 3.5" drives [2.5" SAS 10k rpm at 300GB & 146GB sizes], and support for 600GB SAS 15k rpm drives.

**Flare 24 → 3.24.x.5.x & 02.24.x** [Mars & Janus, CX3, CX300/500]. Mars supports CX [Fish CX300/500/700] and CX3 [Hammer] platforms, not AX, new 750Gb SATA II drives, ATA Drive map recovery, called Rebuild logging (NAS 5.5.26 Napa 5 release & 5.4.28.x). Flare 24 GA December 7, 2006. LUNS are now assigned to Global LUN pools until assigned to specific SPs—separate LUN folder in Navisphere; Support for iSCSI Jumbo Frames; Starting with Flare 24, SP access security can be managed by Navisphere Manager and an optional LDAP implementation through use of new LDAP/LDAPS login Scope, meaning that AD or LDAP can authenticate users. Classic NaviCLI Manager uses Global or Local Scope. Introduces Fast Bind of LUNs—luns are immediately available for use, though background zeroing occurs and may degrade performance til completed. Many TCE initiatives built into FLARE, such as simpler storage management, more wizard utilities, better pre-defined reports, capturing array configuration to XML file:

**# /nas/sbin/navisecli -h 192.168.1.200 arrayconfig –capture –output >/home/nasadmin/ex80\_template**

**SCOPE VALUES FOR SP LOGIN:**

Scope 0 Global

Scope 1 Local

Scope 2 LDAP

**# /nas/sbin/navisecli migrate –start –source 6 –dest 7 –rate low –user myname –scope 2 –password mypass**

**NAVISPHERE ROLES:**

Administrator; Manager; Monitor

**Note:** All 3<sup>rd</sup> Generation CX(3) platforms supported at this flare version

**Flare 26 →** Leo, GA Aug 31 2007, CX300/500/700 & CX3 arrays supported. CX200/400/600 not supported. Introduces RAID 6 to protect against dual drive failures in the same RG and media errors during rebuilds, etc.; implemented with even widths 4, 6, 8, 10, 12, 14, 16 only, with only one Proactive Spare allowed per Raid 6 LUN; Only navisecli supports creation of RAID 6 type; New Asymmetric Logical Unit Access (ALUA) failover mode without application involvement can allow access to LUN from all ports simultaneously; New security enhancements, Classic CLI filtering possible via Navi, but Secure CLI will become the default; Proactive sparing; NTPv4, etc. Supports IE 7.0, Netscape 7 & 8.0, FireFox 2.0; Minimum JRE 1.5; Introduces ability to add a “Security Administrator” role for administration in pure Flare 26 environments only—a Security Administrator can do everything except to see or bind luns, create mirrors, or install/update software; iSCSI Replication over MirrorView or SANCopy  
→Pisces patch release supports RAID 6 on AX4 array

**Note: Last flare version to support the 2<sup>nd</sup> Generation CX platforms, the CX300/500/700**

**Adding Security Administrator Role for User Secadmin with Global Scope:**

**# /nas/sbin/navisecli -user nat -password nat -scope 0 -address 10.241.168.149 security -adduser -user secadmin -password secadmin -scope global -role securityadministrator -o**

**# /nas/sbin/navisecli -user nat -password nat -scope 0 -address 10.241.168.149 security -list** (verifying entries)

**Disabling/Enabling ClassicCLI with Flare 26:**

**Note:** Flare 26 leaves ClassicCLI enabled by default, but provides the option to disable it

**NAVISPHERE:**

Navisphere>Arrayname>Rightclick and select “Enable/Disable Classic CLI” to enable or disable functionality on the fly

**USING CLI:**

**# /nas/sbin/navisecli -h 192.1.4.214 classiccli -disable**

Do you really want to disable Classic cli(y/n) y

ClassicCLI: Disabled Succesfully

**# /nas/sbin/navisecli -h 192.1.4.214 classiccli -status**

ClassicCLI ENABLED: No

**# /nas/sbin/navisecli -h 192.1.4.214 classiccli -enable**

Do you really want to enable Classic cli?(y/n) y

ClassicCLI: Enabled Succesfully

**# /nas/sbin/navisecli -h 192.1.4.214 classiccli -status**

ClassicCLI ENABLED: Yes

**# /nas/sbin/navisecli -h 192.1.4.214 classiccli -disable -o** [without prompt]

ClassicCLI: Disabled Succesfully

**Flare 28 →** Proteus 64-bit (mainly to address memory addressing limitations) Flare running on W2k3 to support new Fleet Clariion CX4 arrays, Nautilus, Ironclad, Trident, Dreadnought (New CX4-10 & replacements for CX3-20, CX3-40, & CX3-80, respectively), May 2008. Enhancements are Single SP Write Caching (GWCA—Greater Write Cache Availability—i.e., if an SP goes away, the other can still support Write Cache Enabled), Host I/O throttling to first five drives with queue depth ceilings, WWN port field bit allocation change, support for IPV6 with management module, Flexible I/O Ports—UltraFlex SLICs (FE or BE FC ports can be interchanged as needed—only fixed port will be port 0 on IO module 0, which is logical backend 0). NST will add a Flex IO Port

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
Wizard. No Flare reversions possible, upgrades performed using Software Assistant. Management LAN ports now visible & configurable: networkadmin –get –all

**Celerra Support:**

Gateway models support Flare 28/Fleet beginning with NAS 5.6.39.5

NS-120 & NS-480 Integrated support introduced with NAS 5.6.41.2

NS-960 Integrated support introduced with NAS 5.6.43

**Flare 28 Mira Release—Flare 28.5:** GA December 8, 2008; 04.28.000.5.501 (factory)/.504 (download)

Patch release support for Virtual Provisioning (Thin Pools & Luns), Enterprise Flash Drives (EFD) 73GB for all Fleet arrays, and Data-In-Place Upgrades (DIP); 960-Drive Expansion Enabler for CX4-960; support for 400GB EFD (005048999)

Minimum NAS version 5.6.43

**Flare 28 Libra Release—Flare 28 ++ May 2009, 04.28.000.5.704 [ECO 67564 Flare 04.28.000.706]:**

Support for 4-port 8Gbps FC SLIC (Glacier card—303-092-100B) as a FRU

Minimum NAS version 5.6.45

**Note:** Arrays will now be able to support 1Gbps iSCSI, 2/4Gbps fibre channel on the backend DAE loops, and 2/4/8Gbps FC for Front-end hosts. SFPs between 4Gbps array ports & 8Gbps host front-end ports are different. Arrays still limited to 4Gbps internally.

**Flare 29 →Taurus release Aug 2009 (04.29.000.5.001); Update to CX4 series**

→New 2-port 10GB Poseidon iSCSI SLIC card support (requires optical SFPs, not RJ45 LAN ports as for 1Gbps iSCSI)

→Virtualization-aware Navisphere Manager

**Note:** Integrates with VMWare vCenter APIs to help manage virtual machines and storage resources. A free EMC Storage Viewer for CLARiiON download is available as a plug-in to vCenter to map logical and physical relationships between VMs and Storage

→Ability to change SP IP addresses without rebooting

→VLAN tagging for Management and iSCSI Clariion ports

→SPCollect improvements

→EMCRemote replaced with Remotely AnyWhere (RA)

**ESRS Issue:**

Per emc220481, ESRS DRM does not dynamically update its connection records, so Users must manually select the correct tool when connecting. In Oct 2009 the DRM will be able to dynamically update its connection records.

**RemotelyAnywhere Connection Options:**

- a) Connect to Navisphere /setup page and add IP address to RA filter list [Set RemotelyAnywhere Access Restrictions], then connect to SP IP address & Port 9519: [<sp\_ip:9519]. Use clariion1992 clariion1992 as username password. RA Dashboard page, select RA remote control feature to use. Use “clariion clariion!” as user and password to log into the SP’s Windows operating system.
- b) Use RA when connected to the SP’s LAN service ports [laptop set to 128.221.1.249/254 with mask 255.255.255.248]
- c) Use RA when connected to SP’s Serial port, IP address 192.168.1.2 [Start>Settings>Network Connections>New Connection Wizard>Set up an advanced connection>Connect directly to another computer>Guest>CIN1>Anyone’s use>Add a shortcut>Finish. Properties box, General tab>Configure>115200bps>close>Enter “clariion clariion!” for user and password>Connect. Use RA client.
- d) Connect to ESRS client, startup RCC client, enter storage serial number, search, highlight system, Connect, select SP, highlight RemotelyAnywhere, Request Session, Invoke Application or enter Target Connection IP of SP with Port 9519 specified. Enter clariion1992 clariion 1992 on RA authentication page, Login. RA Dashboard opens, select RA feature to use.

→NST/CUT Tool for Celerra

→SSD Phase 2 Support for SATA drives

→Virtual Server Integration for ESX support & VMWare ESX 4.0 support

→Virtual Provisioning Phase 2 (LUN shrink for W2k8 only) support for CLARiiON Replication applications using FC or iSCSI:

MirrorView/A; MirrorView/S; SANCopy (one time copies or ongoing replication between arrays)

→MView/A & S now support up to 64 members in a Consistency Group

→Thin LUNs supported Source or Target for MView, SAN Copy

→Policy-based Idle Raid Group SATA II drive spindown

→Adoption of ECOM Common Object Manager model for CIMOM & Web Services

→Integration of NST & Navisphere

→Support for Internet Explorer 8

→SLIC conversions

→Increased Initiator and LUN support

→Online drive upgrades

→Serviceability items

→RAID 6 rebuild logging

→New Navisphere roles for Local Replication Only, Replication, and Replication/Recovery

**Remotely AnyWhere (RA) Utility:**

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
RA replaces EMCRemote on CLARiiON systems, and is used to access the SPs. RA provides for IP Address Filtering, which means that all RA clients must be added to filtering table found on the /setup page in Navisphere. Default LAN Service IPs for RA-- 128.221.1.249; 128.221.1.254; and Serial RA IP address is 192.168.1.2.

## **LOGGING INTO THE SP USING REMOTELYANYWHERE (R29 +):**

- 1) Connect a LAN cable between a Windows client and SP A's Service LAN port.
- 2) Set the IP address and mask of the Windows client to 128.221.1.249 / 255.255.255.0.
- 3) Open a browser and access SP A's setup program: 128.221.1.250/setup.
- 4) Scroll down to the "Set RemotelyAnywhere Access Restrictions" button and add the Windows Client's IP address to the access list (e.g., 128.221.1.249), then "Apply Settings."
- 5) Close the browser, wait a minute, then reopen and try to connect to the RemotelyAnywhere service on SP B using "https": https://128.221.1.251:9519/
- 6) Enter the RemotelyAnywhere credentials:  
User Name: clariion1992  
Password: clariion1992
- 7) To the left of the screen, click on the "Remote Control" option and wait for the logon screen to appear.
- 8) Click on the "ctrl + alt + del" icon, then logon:  
Username: clariion  
Password: clariion!
- 9) Log off the SP using Start > Shutdown > Log off clariion
- 10) Log off the RA session by clicking the "End Remote Session" box, in red, at the top of the screen.

### **NST Changes:**

- Supports installation of 4 or 8GB FC SLICs, 1 or 10GB iSCSI SLICs
- DRU, Disk Firmware, and Disk Installation Utility now supports EFDs
- CLARiiON Software Assistant will automatically detect and notify users when Celerra system is attached when performing NDU
- New model designations with Poseidon 10Gb iSCSI modules: CX4-120C8X, CX4-240C8X, CX4-480C8X, CX4-960C8X

### **Flare 30 → Zeus will be a SW & HW release; last major SW release for CX4 arrays.**

- Introduction of low-end DPE (Zodiac) based on Magnum platform, supporting 6Gb SAS 2-port Coromandel SLIC card for communication & control with Viper DAE, supporting boot from SATA or SAS drives—hence BE support only; 3.5" Viper DAE with (15) 3G/6G SAS drives or SATA drives; 450GB 15K 3Gb SAS drives; 1.5TB/2TB SATA 3Gb drives; 300/450/600Gb 15k 6Gb SAS drives; 300/450/600 10k 6Gb SAS drives; 100/200GB SATA EFD drives with Chariot Paddle card to translate SATA to FC; ODFU—Online Disk Firmware Upgrade capability—basically, where drives have redundancy within a Raid Group, one drive will be upgraded at a time; SSPG support; Advanced SNAPS; Virtual Array; EFD Cache Phase I; TwinAx Passive cable support for 10Gb iSCSI as an alternate to Optical Fibre cables (SFF-8431 standard)

### **6Gb SAS IO Module Support:**

- Support for Coromandel 6Gb SAS 2-port SLIC for SP to DAE BE connectivity (SFP to mini-SAS connectors)

### **Viper DAE Support 15@3.5" Drives:**

- Supports SAS or SATA drives
- Requires (2) new 6Gb SAS LCC's per DAE

### **RAID 6 Parity Shredding:**

- When a RAID group is degraded, provides ability to write data to a parity position in the event of a dead drive. Data is copied back when drive restored from shed data position, and parity rebuilds

- Data integrity feature for RAID

### **LUN Auto-Tiering (FAST):**

- Allows for automatic migration of luns between EFD, FC, SATA drive types
- Based on hot vs. cold data on luns
- Tiers can be built on drive type, speed, capacity, and RAID type
- Requires installation of the auto-tiering enabler
- Auto-tiering works on lun analysis, review, and migration of luns  
navicli storagepool –autoTiering –list –tiers | -create –disks <disklist>  
navicli lun –list –l xx –tieringpolicy –initialtier –tiers  
autotiering –schedule –modify

### **Auto-Tiering LUN Framework:**

Zeus will provide an Enabler for Auto-tiering for LUNs within StoragePools, using a Default Schedule  
C:\>navisecccli –h <ip> autoTiering –modify | -info –state | -info –rate | on off (default) | -tierPreference bestFit

### **DLU Pool LUNs:**

- DLU pool Luns will replace typical Flare Lun in the RAID group as RAID 1/0 Pools
- DLU pool Luns can also be used as Clone Private LUNs, MView Write Intent Log Luns, & Reserved Luns
- Comprised of fully provisioned DLUs, presented as a pool of luns to customer
- Expand LUN capability for DLU & TLU's
- LUN Shrink capability for DLU & TLU's

**LUN Compression 1.6:**

- Uses Recoverpoint Algorithms
- Allows for compression of a LUN during migration to a Thin LUN target
- Use naviseccli compression -list | -on -l xx | migrate -list

**EFD Flash Cache:**

- New storage tier that will utilize EFD drives as write cache, resides between DRAM Memory & Disk Drives

**SATA EFD Support:**

- Based on solid state flash technology, 100 & 200GB sizes
- Formatted with native 520 byte block size
- Drive uses a Chariot paddle card to translate SATA II to FC protocol when used in FC DAEs

**Mixed DAE & RAID Group Drive Support:**

- Support for FC & SATA EFD drives in same DAE
- Support for FC & SATA EFD drives in same RG

**Navisphere Trusted Communications:**

- Current communication path uses SSL
- Revised communication path will use SSL to encrypt, but add Certificate Validation to ensure clients are Trusted

**Unisphere:**

- New GUI interface used for Clariion and Celerra system management, replaces Celerra Manager and Navisphere

**Unisphere Service Manager:**

- New name for NST
- Same look as Unisphere Service Manager
- Can launch USM from Unisphere
- Introduces Celerra CUT tool within the USM framework for upgrading NAS systems
- Intelligent Advisories (TAND)

**Online Drive Firmware Updates (ODFU)**

- Ability to upgrade drives one at a time within Raid Groups to minimize impact, hence “online”, though the upgraded drive itself is offline during the firmware update
- New wizard in USM

**VMware vStorage APIs for Array Integration (VAAI) Support:**

- Purpose is to offload work from the ESX Server Host to the Storage array by performing work directly on the array
- VMware Full Copy Acceleration enhancement for cloning of file blocks
- VMware ATS solution in clustered environments to alleviate multiple lock situations to single LU's
- Array Accelerated Zero; Array Accelerated Copy; Array Accelerated Locking
- EMC Storage Viewer application on vSphere Client has been renamed to Virtual Storage Integrator

**Backend Path Failure (BPF) Support:**

- Current behavior if drives are available from one SP, but not the other, is to have FLARE fault the drive
- New behavior is to allow drive paths to failover to other SP until the BPF situation is fixed

**EMC Secure Remote Support (ESRS 2.0) IP Client for CLARiiON:**

- Secure centralized 2-way connection from EMC to customer storage for remote management
- Consists of ESRS IP Client; ConnectEMC; Navicli; Unisphere Host Agent; EMCRemote
- Personnel connect via an EMC Service Link
- System with highest version of Flare serves as Portal system
- Other arrays in the domain will be auto-registered
- Auto-registration of devices with DRM (Device Relationship Manager)
- Automatic Capture and Upload (ACU) and Re-active SP Collects (RSC) will be enabled by default

**Misc. Enhancements:**

- Ability to support multiple logins between the iSCSI Initiator and Target (supports initiator's with multiple interfaces that use single IQN; booting from iSCSI Luns where BIOS retains connection to initiator; VMware MPIO support; Linux & Solaris login to all ports by default)

→Enhancement to SPCollect by grouping collection commands together for concurrency

**FLARE 30.5 →**Jupiter release. Provides SP support for 2-port FCoE module on CX4-120/240/480/960. Introduces FAST Cache Read/Write support for CX4-480.

**FLARE 31 →**SCORPION Release for new CX5 Fighters arrays

→Flare 04.31.000.5.xxx Hellcat lite, Hellcat, and Lightning arrays based on Sentry platform; 6Gb SAS backend; Viper 3.5" @ 15 drive SAS DAE; Derringer 2.5" @ 25 drive SAS DAE

**CX3 ARRAYS & FLARE SOFTWARE:**

Flare is installed on first (4) disks on backend loop 0

Disks 0\_0 & 0\_2 have mirrored copies of flare for SPA, Disks 0\_1 & 0\_3 have mirrored copies of flare for SPB

**CLARIION ARRAY PRIVATE SOFTWARE SPACE**

Clariion array uses the first (5) physical drives on the array for Private space, dedicated to a variety of databases, partitions, etc.  
 CX600 Fish family uses NT O/S for Flare, CX700 Fish family converts to XP O/S, Fleet will use W2k3 O/S 64-bit

| <b>CX FISH PLATFORMS</b>           | <b>CX3 HAMMER PLATFORM</b> | <b>CX4 FLEET PLATFORM</b>      |
|------------------------------------|----------------------------|--------------------------------|
| <b><u>Disks 0-4:</u></b>           |                            |                                |
| DDBS (Data Directory Boot Service) | DDBS                       | DDBS                           |
| ODBS (Data Directory)              | ODBS                       | ODBS                           |
| Flare Database                     | Flare Database (legacy)    | Flare Database (legacy)        |
| Vault                              | Vault                      | Vault                          |
| DH Diagnostic Area                 | DH Diagnostic Area         | DH Diagnostic Area             |
| User LUNs                          | User LUNs                  | User LUNs                      |
| NAS LUNs                           | NAS LUNs                   | NAS LUNs                       |
|                                    | Future Growth              | Future Growth                  |
|                                    |                            | Impacted Sector List 1, List 2 |
|                                    |                            | Disk Obituary                  |
| <b><u>Disks 0-2:</u></b>           |                            |                                |
| External Databases                 | External Databases         | External Databases             |
| PSM                                | PSM                        | PSM                            |
|                                    | Boot Directives            | Boot Directives                |
|                                    | Flare Database             | Flare Database                 |
| <b><u>Disks 0-3:</u></b>           |                            |                                |
| Flare Boot Partitions (Flare)      | Flare Boot Partitions      | Flare Boot Partitions          |
| Utility Partitions                 | Utility Partitions         | Utility Partitions             |
| <b><u>Disk 4:</u></b>              |                            |                                |
| IR (Image Repository)              | Image Repository           | Image Repository               |
| <b><u>Disks 0-1:</u></b>           |                            |                                |
|                                    | Image Repository           | Image Repository               |
| <b><u>Total Private Space:</u></b> |                            |                                |
| 30GB                               | 33GB                       | 66GB                           |

**DECODING FLARE SOFTWARE BUILDS:**

**Example:** v02.04.0.60.5.005 [Flare Release 11]

Translation: v02 = 02 for 2GB Family; 04 = Minor Release Version, Release 4; 0 = SW Options, 0 = No Options 1 = AccessLogix; 60 = Hardware Identifier, 60 = CX600 (20=CX200; 40=CX400; 01=Chameleon2; 03=Longbow); 5 = SW Distr. Type, 5 = General Release; 005 = Build Iteration, 005 = SW Distr. Type >= 4

**CELLERRA/CLARIION FLARE vs NAS UPGRADE RULES:** (check eLab tables & emc203413)**GENERAL FLARE vs. NAS UPGRADE GUIDANCE:**

- See emc203413 for a Table that shows Flare vs. NAS compatibility for NAS 5.4, 5.5 and 5.6
- See CCA Guide for Flare Interoperability comments and guidance
- NX4, NS-120, NS-480, NS-960 systems use NST software assistant to perform FLARE NDU
- Older NS Integrateds (NS500, NS350, NS700, NX600, NS80, NS40, NX40G, NS20, NSX) should follow Celerra Procedure Generator guidance when upgrading FLARE with Celerra attached
- NST uses built-in FLARE Provider API for Rules Checking with Flare 23 and higher

**FLARE vs. NAS GUIDANCE IN NAS CCA GUIDE (Oct 2008):**

[http://3cgs.corp.emc.com/nascca/guide/nas\\_cca\\_guide.htm](http://3cgs.corp.emc.com/nascca/guide/nas_cca_guide.htm)

“The **number one rule** of FLARE / NAS Code interoperability is that **the FLARE can never be upgraded to a version that the NAS Code version doesn't support**. There are currently **NO** circumstances where this rule can be broken.”

**Example where one would upgrade FLARE prior to upgrading NAS, which does not break the above rule:**

“emc186425 - Do not upgrade or install NAS code 5.6 on a Celerra attached to a CLARiiON running Flare code version 24 Due to an issue specific to FLARE R24, the CLARiiON must be upgraded to R26 prior to upgrading the NAS Code to 5.6.”

**LEGACY UPGRADE RULES:**

1. General rule, upgrade NAS code first, then FLARE

**Note:** However, with Napa 5.5.26.4 Release, NAS code for CX Fish arrays will only support Flare 19 and higher, and for CX3 Hammer arrays, NAS 5 will only support Flare 22 and higher.

2. Previous CCA rules required (emc108354) that Celerra be offline before upgrading FLARE Code

**Note:** March 3, 2006 Clariion Online Update procedure released to allow NDU Upgrades with attached Celerra for NS & NSX systems only, with minimum Flare version 16, and NAS versions 5.3.23 ro 5.4.20 and above.

3. Likewise, always shutdown Celerra if adding ATA DAE or upgrading Flare with ATA drives involved
4. Upgrading from Flare 7/8 requires Utility Partition first, upgrade to Flare 11, then latest Flare
5. Prior to upgrades, verify that LUNs are not trespassed or in transition
6. Verify that there are no faults on the array, obtain resume information, run SPCollects and analyze, etc.
7. Install CAP2 program on laptop prior to upgrading array to Release 16

**Caution:** In Clariion world, NDU means that Array itself has valid failover software for SP's, meaning that upgrade occurs on one SP at a time, so whole system does not need to be down. Utility Partition is added to first four Fibre Channel Boot Drives [Vault Drives] and is meant to be used as a Recovery Image to recover from a corrupted boot image. As a general note, Clariion Procedure Generators are NOT specific-enough and do not unequivocally state that Celerra Servers should be offline during Flare Upgrades, though it does state that ATA chassis require all I/O to be halted. See Primus emc90820.

### **ATA DRIVES AND FLARE UPGRADES:**

- Stop I/O to array during NDU when ATA drives are involved
- Manually disable and zero write cache until update is complete

--Clariion rules state that ATA Chassis should have all LUNs assigned to one SP or the other, and not both

**Note:** FLARE upgrades include new firmware [FRUMON code] for LCCs on attached DAE and ATA chassis. The firmware upgrade on LCCs can cause loss of access to Array for Celerra Servers and is the ultimate reason for never performing an OnLine Flare Upgrade with Celerra. Primus emc93310.

### **ATA DRIVE ISSUES:**

- Clariion Flare upgrades should have Hosts offline if it involves ATA drives
- Clariion uses ATA in RAID 3 configurations only (Not Celerra)
- Drives have single port to backend via LCC card, as opposed to fibre channel, which has two ports to backend
- Flare 19 fixes major ATA drive issue with unrecoverable or invalid sectors being logged for bitflip problem

### **TRY NDU UPGRADE USING SKIPRULES SWITCH:**

**Note:** If you still encounter a failure, you may need to run the NDU Upgrade from the Off-Array SIW

**# /nas/sbin/navicli -h 10.241.168.57 ndu -install /home/nasadmin/flare19/FLARE-Operating-Environment-02.19.600.5.027.lst -skiprules -delay 360**

Item number: 0  
Name of the software package: FLARE-Operating-Environment  
Revision of the software package: 02.19.600.5.027  
Already Installed Revision 02.16.600.5.019  
Installable YES  
Disruptive upgrade: NO  
Ndu Delay: 360 secs

The requested package(s) will be installed. Do you wish to proceed? : (y/n)? y

### **MONITOR THE FLARE UPGRADE:**

**# /nas/sbin/navicli -h 10.241.168.57 ndu -status**

Is Completed: NO  
Status: **Disabling the cache**  
Operation: Install  
#/nas/sbin/navicli -h 10.241.168.53 ndu -status  
Is Completed: NO  
Status: **Running check scripts**  
Operation: Install  
#/nas/sbin/navicli -h 10.241.168.57 ndu -status  
Is Completed: NO  
Status: **Installing software on secondary SP**  
Operation: Install  
#/nas/sbin/navicli -h 10.241.168.57 ndu -status  
Is Completed: NO  
Status: **Activating software on secondary SP**  
Operation: Install  
#/nas/sbin/navicli -h 10.241.168.57 ndu -status  
Is Completed: NO  
Status: **Rebooting secondary SP**  
Operation: Install  
#/nas/sbin/navicli -h 10.241.168.57 ndu -status  
Is Completed: NO  
Status: **Waiting due to DMP delay**  
Operation: Install  
#/nas/sbin/navicli -h 10.241.168.57 ndu -status  
Is Completed: NO

Status: **Installing software on primary SP**

Operation: Install

# /nas/sbin/navicli -h 10.241.168.58 ndu -status

Is Completed: NO

Status: **Rebooting primary SP**

Operation: Install

# /nas/sbin/navicli -h 10.241.168.58 ndu -status

Is Completed: YES

Status: **Operation completed successfully**

Operation: Install

## **UPGRADING FLARE 19 PATCH 027 to 034 USING OffArray SIW & LAYERED APPLICATIONS:**

### **OffArray SIW Upgrade from Laptop:**

1. Download FLARE bundle: <http://www.cs.isus.emc.com/csweb2/dgweb/software/index.asp>

**CX600-Bundle-02.19.600.5.034.pbu**

2. Download the "**Navisphere OffArray UI's (Windows) 6.19.2.6.5**" from Windows Management Station section of website

3. Doubleclick and install the SIW executable: **Windows\_UIs\_6.19.2.6.5.exe**

4. Connect to Array using Start.html file, located at:

Program Files>EMC>ManagementUI>6.19.2.6.5>WebContent>**start.html**

5. Enter IP address at the Navisphere Connection window & log into the array using valid User and Password:

10.241.168.52>connect

Navisphere Login →User \*\*\*\*\*

6. Enter Eng. Mode: ctrl + shift + f12 & enter messner for password

7. Rightclick array>Software Operations>Software Installation Wizard>next>NDU Upgrade Delay (360)

Browse to .pbu file, doubleclick—unpacks and transfers files to Clariion SP

A list of available software appears>Click Next to start software installation

8. After upgrade begins>click Finish to modify software progress, or use ndu -status from Celerra

Installs sw on Secondary SP first

Reboots Secondary SP

Repeats same steps for Primary SP

9. After verifying software upgrade success, go back into Navisphere Manager>rightclick array>properties>Software>highlight Flare-Operating-Environment for 02.19.600.5.034 Active (Commit required) and commit the software

**# /nas/sbin/navicli -h 10.241.168.57 ndu -commit FLARE-Operating-Environment**

10. Verify using ndu -list:

**# /nas/sbin/navicli -h 10.241.168.57 ndu -list**

Name of the software package: FLARE-Operating-Environment

Revision of the software package: 02.19.600.5.034

Commit Required: NO

Revert Possible: NO

Active State: YES

Is installation completed: YES

11. After commit is completed, failback any trespassed Celerra LUNs as a result of the SP reboots during upgrade

**# /nas/sbin/navicli -h 10.241.168.57 getlun -trespass**

LOGICAL UNIT NUMBER 91

Default Owner: SP A

Current owner: SP B

LOGICAL UNIT NUMBER 0

Default Owner: SP A

Current owner: SP B -----output abridged-----

**# /nas/sbin/navicli -h 10.241.168.57 trespass lun 0**

**Note:** For MirrorView luns, may need to trespass luns from other side if following message seen

**# /nas/sbin/navicli -h 10.241.168.58 trespass lun 113**

Error: trespass command failed

Error returned from Agent

You are not allowed to transfer the LUN to the other SP (0x4000805a)

**Note:** In the above case, the mirrored lun was lun 28, which was successfully trespassed back from the other side

## **CELLERRA FLARE NDU UPGRADE ISSUES:** Flare 24 & 26

DIMS 180689

ETA emc173153 for Flare 24 and emc1877241 for Flare 26

### **Causes:**

A problem exists in all versions of Flare 24 prior to patch 016. When the ndu -install command is initiated, the NDU process must

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
first run a series of rules that ensures that the system is not faulted and is safe to proceed with an upgrade. In this particular instance, the runrules process can produce events that are greater than 255 characters, and causes a corrupt event log, and a resulting timeout and failure of the ndu command. For Flare 26, exact cause is unknown, but likely related to ssl\_socket communication timeouts.

**SYMPTOM:**

# /nas/sbin/navicli -h 192.1.4.214 ndu -install /home/nasadmin/CX3-40-Bundle-03.26.040.5.005.pbu -delay 360

Error returned from the target: 192.1.4.214

Please specify valid IP address of the storage system and try again.

**WORKAROUND:**

1. Make sure that Naviseccli security is updated with a valid Global CLARIION Administrator account and password:

# /nas/sbin/navisecccli -h 192.168.1.200 -AddUserSecurity -scope 0 -user nasadmin

Enter password:

2. Use the skiprules switch to force the FLARE Upgrade

Caution: Do not use -skiprules without a thorough array healthcheck first! If in doubt, do not continue this procedure.

# /nas/sbin/navicli -h 192.1.4.214 ndu -messner -install /home/nasadmin/CX3-40-Bundle-

**03.26.040.5.005.pbu -skiprules -delay 360**

Item number: 0

Name of the software package: FLARE-Operating-Environment

Revision of the software package: 03.26.040.5.005

Already Installed Revision 03.24.040.5.006

Installable YES

Disruptive upgrade: NO

Ndu Delay: 360 secs

The requested package(s) will be installed. Do you wish to proceed? : (y/n)? y

3. Use the -status switch to monitor the FLARE upgrade progress:

# /nas/sbin/navicli -h 192.1.4.214 ndu -status

Is Completed: NO

Status: Waiting due to NDU delay

Operation: Install

# /nas/sbin/navicli -h 192.1.4.214 ndu -status

Is Completed: YES

Status: Operation completed successfully

Operation: Install

4. Do the following post-upgrade steps:

Re-enable statistics logging on each SP.

Verify that Write cache has been enabled and array is healthy.

Fail back any Celerra LUNs to Owner SP.

Commit the FLARE version.

Restart NAS Services if applicable (single blade systems only)

**Note:** The use of –skiprules is similarly required in some situations involving Clariion Layered Applications

**FLARE 24 NDU MAY REQUIRE USE OF NAVISECCLI (emc235884):**

# /nas/sbin/navicli -h 192.168.1.200 ndu -install CX3-40-Bundle-03.24.040.5.022.pbu -delay 360

Unable to access jarfile /opt/Navisphere/bin/SIWMain.jar

Pre installation rules have been run to ensure the success of this software upgrade. The above conditions have been detected that need to be corrected before running this command again. If this condition persists please contact your EMC Service representative.

# /nas/sbin/navisecccli -h 192.168.1.200 -user nasadmin -password nasadmin -scope 0 ndu -install CX3-40-Bundle-03.24.040.5.022.pbu -delay 360

Running install rules...(edited for brevity)

Pre installation rules have been run to ensure the success of this

software upgrade. One or more rules has a warning. The installation can continue with warnings.

The requested package(s) will be installed. Do you wish to proceed? : (y/n)?

**LISTING RULES FOR A FLARE PACKAGE:**

# /nas/sbin/navicli -h 192.168.100.200 ndu -runrules CX3-40-Bundle-03.26.040.5.005.pbu -listrules | -

**verbose** (verbose list of rules)

Your configuration will run the following rules

=====

Special Conditions

Redundant SPs  
All Packages Committed  
No Trespassed LUNs  
Statistics Logging Disabled  
No Transitions  
Acceptable Processor Utilization  
No Active Hot Spares  
Version Compatibility  
No Un-owned LUNs  
No Active Replication I/O  
No System Faults  
Host Connectivity  
SP Event Log Verification

### **RUNNING NDU RULES FROM NAVICLI:**

**# /nas/sbin/navicli -h 10.241.168.150 ndu -runrules \***

Version Compatibility : Rule passed.  
Redundant SPs : Rule passed.  
Acceptable Processor Utilization : Rule passed.  
No Trespassed LUNs : Rule passed.  
No Transitions : Rule passed.  
No System Faults : Rule passed.  
All Packages Committed : Rule passed.  
Special Conditions : Rule has warning.  
Statistics Logging Disabled : Rule failed.  
Host Connectivity : Rule passed.  
SP Event Log Verification : Rule passed.  
No Un-owned LUNs : Rule passed.  
No Active Replication I/O : Rule passed.  
No Virtual Provisioning Transitions : Rule passed.

1 rule(s) have warnings.

1 rule(s) failed.

Detailed rule results:

=====

RULE NAME: Special Conditions

RULE REVISION: 6.28.20.1.40.1

RULE STATUS: Rule has warning.

---output abridged-----

### **BACKEND ENABLER:**

Backend Enabler software is used for “captive” Celerra/Clarion configurations that use NAS personality modules to “enable” the AUX ports when connecting DM’s to Backend ports. Certain “Integrated” models (NS600/700) require the Backend Enabler, but the **NS350, NS500, NS704, NS40, NS80** Celerras do not use AUX ports, therefore have no ‘personality’ module or Backend Enabler software to install. Use the normal Flare bundle from Clariion website for all Integrateds except the original NS600/NS700’s.

**Caution: Do not use the Backend Enabler on SAN systems or the ports will be disabled!**

### **FLARE 24 to 26 UPGRADE:**

Flare NDU on NS Integrated models running Napa 8 or higher requires use of Control Station CLI. Beginning with NX4 product, the NST will be used to perform Flare NDU on Integrated systems. Basically, if you use NST and get the following popup, then use of GUI is not supported: “This feature is not supported”

**# /nas/sbin/navicli -h 192.168.100.200 ndu -install /home/nasadmin/flare26/CX3-40-Bundle-03.26.040.5.005.pbu -skiprules -delay 360**

Exception : A network error occurred while trying to connect: '192.168.100.200'.Message : Error occurred because connection refused. Management Server is not running.

**Note:** Found that the only way to do the upgrade is to use the –skiprules syntax—the above network error was benign and the upgrade proceeded. Root cause of the problem is with a Flare 24 bug, where RunRules can generate Events that are >than 255 characters, which corrupts the Event Log, and causes NDU command to timeout, and fail. See ETA emc173153 for more detail.

### **CELERRA NS40/NS80:**

--NS systems designed to support new CX3 arrays, as well as traditional CX and DMX platforms

--Uses Dewars Control Station

## **NS40 & NS40G:**

- GA NAPA II release based on new Clariion CX3 arrays, replacement for NS500 Platform
- Based on Sledgehammer hardware, DM blades are dual Intel 2.8Ghz processors with 4GB memory & 800MHz fsb
- SAN Gateway versions can connect to most array backends, while Integrated model connects only to Sledgehammer (CX3-40)
- Projected replacement for NS500 & NS700 based on Clariion CX3-40 platform
- 1U Enclosures, 1 or 2 blades only (Data Movers), single Chivas Control Station (2GB memory), (4) Power Supplies, Internal Management Switch, (2) serial ports for DART console—COM1 & COM2
- NS40 Blade with NAS personality will have dual 2.8GHz Nocona P4 processors, (2) serial console ports on lower left side, (2) RJ45 connectors to Internal Management switch, (4) Broadcom 5704 Gigabit Ethernet Optical ports [cge0-cge3] [or (2) Copper fge ports & (2) Optical cge ports], (2) Agilent 1/2/4Gb Aux Fibre Ports, and (2) Agilent 1/2/4Gb BE Fibre Ports
- SAN blades for SP's can have either (2) or (4) Fibre ports each, while NAS blades have only (2) Fibre ports
- Internal Management switch built into Motherboard, as opposed to NSX, where Management Switches are FRU's, with (2) RJ45 ports for access
- SPs run in SAN AUX mode using copper Fibre Channel HSDCC2 cables between SPs and DMs for Integrated, and SAN mode for Gateway using Fibre Optic cables
- NS40 can have 16TB per DM blade using FC drives, or 32TB using mixture FC/ATA drives, and supports up to 240 drives

## **NS80 & NS80G:** Supported NAS 5.5.24.2 Napa 4

Gateway & Integrated models. Replacement to NS702 & 704G. Offers 40u cabinet, non-UPS, between (2) and (4) Xblade 60/65 Servers, 1 CS upgradeable to two; Support for 10GbE Ethernet IO module on X-blade 65 and 4Gb FC IO module; NS80G GA NAPA 4 release Oct 9, 2006; NS80 Integrated GA Nov 17, 2006.

### **NS80FC [Model name NS80-AUX-FC]:**

AUX TLA NUMBER: 100-520-828

AUX Model Number: NS80-AUX-FC

**Note:** Also requires Navisphere & Port-Enabler license (NS80-FCOPT-L)

#### **Minimum Configuration:**

| Model          | Description                  | Qty             |
|----------------|------------------------------|-----------------|
| NS80-FCOPT-L * | NS80 FC enablement License   | 1               |
| RACK-40U-60    | 40U EMC Rack                 | 1               |
| NS80CDME6x     | 2 x NS80 Blades + enclosure  | 1 [Xblade60/65] |
| NS80-AUX-FC    | NS80-AUX B/E with 4 FC ports | 1               |
| NS-4PDAE-80    | First Disk Shelf             | 1               |
| NS80-CS        | First Control Station        | 1               |

→Will be released May 2008, no upgrading to or from an NS80FC (new installation of NS80FC only)

→RPQ-only product, ships with 5.6.37.x or higher

→Single or Dual blades only (X-Blade 60 or X-Blade 65), no 2<sup>nd</sup> enclosure, Single or Dual CS's possible [DME60 or DME65]

→Non-CSA setup by CS using Installation procedure distributed with RPQ approval

→See Celerra Procedure Generator. The NS80FC is actually installed as a direct-connect NS80G gateway, and will report itself as such.

→Two AUX ports are available per SP for Open Hosts as direct-connect Gateway style system (4 total FC ports)

→There are no simple commands to tell whether a system is an NS80FC

**Note:** System will report as NS80G and CallHome as an NSX

/nas/sbin/model

NS80G

### **TS PROCEDURE FOR INSTALLATION OF NS80FC:**

#### **The basic steps required are as follows:**

- Enable Access Logix and Navisphere Manager on the array and assign customer LAN IP addresses to the SPs.
- Remove the Celerra control LUNs already on the array and destroy RG0.
- Insert floppy and installation CD into the Control Station and reboot it.
- Begin a destructive "serialinstall" with the customer-provided IP information for the CS.
- Installs as Gateway model and prompts for entering SPA & SPB IP Addresses, and array security account & password
- Select Direct-Attached when prompted by installation script.
- Complete the NAS installation.
- Put Additional 4 FC ports on CLARiiON, unused by NAS, into SAN for SAN host attachment.

### **CELLERRA LUNS:**

LUNs 0-5 are created on Raid Group 0, located on the first 5 disks of the array

LUN 00 →contains the DOS partition (NAS image & DM Config files), /nbsnas/dos & /nas/dos, located on /dev/ndal on the backend—Server Logs, Dumpfiles, and DOS partition are located here. LUN 00 size now 11263

LUN 01 →ufslog partition, LUN size 11263

LUN 02 →Linux CS0 partition, all systems except NS Series, size 2047

LUN 03 → Linux CS1 partition, all systems except NS Series, size 2047

LUN 04 →contains /nbsnas partition (NASDB CS config files), located on /dev/nde1, size 2047

LUN 05 →contains /nbsnas/var partition (NASDB backups, dumps, log files), located on /dev/ndf1, size 2047

## **CELERRA AUTO-ASSIGN & AUTO-TRESPASS GUIDANCE:** emc95145

### **For all Celerra systems attached to Clariion backends, the following applies:**

→Beginning with NAS 5.2.16.x and NAS 5.3.10.4 GA release, full NAS Package upgrades will change the default behavior of both Auto-Assign and Clariion Trespass Option for all Celerra systems with Clariion backends. Please note that DART patch would only change the Clariion\_no.trespass value, not the Auto-Assign value.

→Be aware that with NAS 5.4 release, the Auto-Assign feature will be Disabled because the Celerra Host will be able to negotiate its own Trespass capability, as do other Clariion-attached hosts.

### **NAS 5.5.30.4 Defaults:**

Auto-assign

# /nas/sbin/navicli -h 192.168.1.200 getlun 0 -aa |grep assign

Auto-assign: DISABLED

# /nas/sbin/navicli -h 192.168.1.200 getlun 0 -at |grep trespass

Auto-trespass: DISABLED

# /nas/sbin/navicli -h 192.168.1.200 arraycommpath

Current arraycommpath setting is: 0 [0=Disabled]

# /nas/sbin/navicli -h 192.168.1.200 failovermode

Current failovermode setting is: 0

### **VALUES FOR ALL CELERRA SYSTEMS WITH CLARIION BACK ENDS:**

clariion no\_tresspass=0 (Disabled) →Param default is 0 for NAS 5.2.16.x and higher

# .server\_config server\_2 -v "param clariion no\_tresspass" [Syntax for NAS 5.5]

clariion.no\_tresspass INT 0x010c3f00 0 0 (0,4294967295) FALSE REBOOT 'NA'

**Auto-trespass:** Disabled (this is the Navi value that we are talking about with the no\_tresspass param above)

**Auto-assign: Enabled or Disabled** →Disabled by default NAS 5.4/5.5 +

**arraycommpath=0** →Disabled by default NAS 5.5 (Value seen as 0 for Disabled & 1 for Enabled in Navisphere)

**failovermode=0**

### **DEFINITION AND PURPOSE OF CELERRA/CLARIION PARAMETERS:**

clariion no\_tresspass=0

**Auto-trespass: Disabled:** Default Clariion setting is 0 to disable this feature, and Celerra LUNs should also be disabled. If the feature were enabled, if an IO were sent by a Host to an SP for a LUN that it does not currently own, then the trespass command would fail the lun over to the SP and the IO would succeed. Normally, if a Data Mover cannot drive I/O down a certain path, it will send a trespass command to the alternate SP to take over--note that Hosts must also initiate the trespass back to the original SP after a trespass occurs. Setting value to 0 enables the Host to control trespassing for IO requests.

**Note:** In certain situations, setting the value to 1 is done to help trespass Celerra LUNs to their owner SP, but only as a temp measure

**Auto-assign: Disabled:** Disabled by default on Clariion. For Clariion, if an IO request comes to an SP that does not OWN the LUN, the SP will honor the IO request and failover the LUN if the peer SP is not present. If the peer SP is present, the IO request will be denied to the non-owner SP the ownership of a LUN when an SP fails. The trespassed LUN does not get automatically transferred back to its Owner unless specifically told to do so, or unless the system reboots. This feature does not affect Ownership of the LUN. Auto-assign is a mechanism devised in those configurations where failover software was not employed.

**Note:** Auto-Assign must be enabled to ensure that any “un-owned” LUNs are trespassed, works in conjunction with Clariion no\_tresspass parameter.

**arraycommpath: Disabled, 0:** Arraycommpath is a Clariion feature used by PowerPath to set the communication path to the Array for LUNZ support—Celerra does not use or recognize this feature, therefore HLUs and Storage Groups that are created for Celerra Systems should not have it enabled. ATF applications require that the setting be 0, not 1.

**failovermode: Disabled, 0:** Failover Mode works in conjunction with PowerPath and ATF applications. Celerra requires that this setting be set to 0 so that if IO fails to one SP, and DART then tries to send to an alternate path and SP, the non-owning SP will reject the command with a Check Condition 04/04/00 illegal request error. DART will recognize this scsi sense code and realize that it must initiate a trespass command to the alternate SP, thus failing over the LUN to the alternate SP, and re-driving IO down this path.

### **DETERMINING DATA MOVER IO FAILOVER PATHS NS704G:**

\$ .server\_config server\_2 -v "fcip bind show"

\*\*\* Persistent Binding Table \*\*\*

Chain 0000: WWN 500601603060090b HBA 0 SP-a0 Bound →I/O failover would be Chain 48 HBA1 SPA port 1

Chain 0016: WWN 500601683060090b HBA 0 SP-b0 Bound →I/O failover would be Chain 32 HBA1 SPB port 1

Chain 0032: WWN 500601693060090b HBA 1 SP-b1 Bound

Chain 0048: WWN 500601613060090b HBA 1 SP-a1 Bound

Chain 0064: WWN 500104f0008a3496 HBA 2 N\_PORT Bound

Chain 0080: WWN 5006048000000000 HBA 3 N\_PORT Bind Pending

**Note:** I/O failure down a specific HBA chain will cause failover to alternate HBA and different SP port on same SP. I/O failure down both chains for SPA will cause LUN trespass to SPB

## **VERIFYING VALUES:**

### **clarion no\_tresspass**

**\$ .server\_config server\_2 -v "param clarion" or "param clarion no\_tresspass"**

clarion.no\_tresspass 0x011cf874 0x00000000 0x00000000 -->Example output shows that value is set to 0, which is correct

**# .server\_config server\_2 -v "param fulldescription clarion no\_tresspass"**

clarion.no\_tresspass 0x010488dc 0x00000000 0x00000000

**Note:** Required syntax change for NAS 5.5

### **auto-assign**

1.) **#/nas/sbin/navicli -h spa storagegroup -list** (SAN-attached Systems only—run to obtain list of HLU/ALU luns)

2.) **#/nas/sbin/navicli -h 10.14.32.90 getlun -aa** (Use for both Integrated & Gateway systems)

**Note:** The storagegroup -list is shown to illustrate that the Celerra LUNs for Gateway systems will have an HLU number associated with them, and the above output in "step 1" shows the HLU-to-ALU association, which would be needed if values required changing

## **CHANGING ARRAYCOMMPATH/FAILOVER USING NAVISPHERE or CLI:**

ArrayCommpath & Failover Modes are assigned to the SP's WWNs on an initiator basis, but associated to a particular Data Mover Hostname for the HBA port—the values can be changed individually on SPA or SPB Ports 0 & 1.

### **UPDATING WWN HOST INITIATOR RECORDS FROM NAVISPHERE:**

1. If values need to be changed on all data movers, make the change on the Standby Server's Host Records for the appropriate SP and ports using the following methodology:

a.) Navisphere>APM Array>log into Eng. Mode using ctrl + Shift + F12, and password ‘messner’

b.) APM Array>Connectivity Status>highlight the Initiator Name/Server Name>Click on “Info” to see what the current Array Commpath and Failover Mode settings are [Correct values are Disabled and 0, respectively]

c.) APM>Connectivity Status>Highlight Initiator Name record>Group Edit>Select the correct WWN Initiator record from the “Available” WWN list and move to the “Selected” WWN list [Update only one record at a time]

d.) Select “Array Commpath: Disabled” [Disabled=0], “Failover Mode: 0”, and “Existing Host” [Make sure the change is applied to the intended Hostname associated with the WWN record being updated]

e.) Click o.k. to apply, click Yes on popup warning box to acknowledge the activity>Success ‘o.k.’

f.) Repeat the process for each WWN Initiator record that needs updating

| SERVER NAME (DM HOST NAME) | INITIATOR NAME (SP WWN)                         | PARTITION ID |
|----------------------------|---|--------------|
| laip2_dm3_p0               | 50:06:01:60:90:60:25:c4:50:06:01:68:10:60:25:c4 | A-1          |
| laip2_dm3_p1               | 50:06:01:60:90:60:25:c4:50:06:01:69:10:60:25:c4 | B-0          |
| laip2_dm2_p0               | 50:06:01:60:90:60:25:c4:50:06:01:60:10:60:25:c4 | A-0          |
| laip2_dm2_p1               | 50:06:01:60:90:60:25:c4:50:06:01:61:10:60:25:c4 | B-1          |

### **CLI:**

**# /nas/sbin/navicli -h 10.241.168.57 storagegroup -list |grep Name**

Storage Group Name: Celerra\_hammer1

Storage Group Name: Celerra\_nyip1

**# /nas/sbin/navicli -h 10.241.168.57 port -list -gname Celerra\_nyip1**

2. Failover production server to Standby slot, then change arraycommpath & failovermode values on the ‘faulted’ server slot, i.e., the slot not in production!

3. Fallback data mover and repeat as needed

### **UPDATING WWN HOST INITIATOR RECORDS FROM CLI:**

1. Make sure you correctly identify the Data Mover Host Initiator names that are associated with each SP port and run command against Standby Server Host records first:

**#/nas/sbin/navicli -h spa storagegroup -sethost -host laip2\_dm3\_p0 -arraycommpath 0 -failovermode 0**

**#/nas/sbin/navicli -h spa storagegroup -sethost -host laip2\_dm3\_p1 -arraycommpath 0 -failovermode 0**

WARNING: Changing configuration options may cause the array to stop functioning

correctly. Do you wish to continue (y/n)? y

2. Failover production server to Standby slot, then change arraycommpath & failovermode values on the ‘faulted’ server slot, i.e., the slot not in production! In otherwords, if Server\_2 were failed over to Standby Server\_3, the Host Initiator Records for “laip2\_dm2\_p0” and “laip2\_dm2\_p1”, which are tied to physical slot\_2 and not in production, would be used to change the values.

```
#/nas/sbin/navicli -h spa storagegroup -sethost -host laip2_dm2_p0 -arraycommpath 0 -failovermode 0  
#/nas/sbin/navicli -h spa storagegroup -sethost -host laip2_dm2_p1 -arraycommpath 0 -failovermode 0
```

3. Fallback data mover and repeat as needed for multiple Data Mover systems

#### arraycommpath

```
#/nas/sbin/navicli -h spa arraycommpath (Integrated systems only)
```

```
#/nas/sbin/navicli -h spa port -list -all (SAN-attached Systems only)
```

#### failovermode

```
#/nas/sbin/navicli -h spa failovermode (Integrated systems only)
```

```
#/nas/sbin/navicli -h spa port -list -all (SAN-attached Systems only)
```

Server Name: ns600g\_dm30

ArrayCommPath: 1

Failover mode: 1

**Note:** Both the arraycommpath & failovermode output can be misleading, as the array itself could have the values set to 1, while the individual Celerra Host Name records may be set to 0. Use Navisphere GUI to verify arraycommpath/failovermode values for the specific Host Name/Initiator Records in question. Navisphere should correctly show Celerra Host Name records with Array CommPath ‘Disabled’ and Failover Mode ‘0’.

#### NAS 5.5.31.6 CHANGES:

Upgrades will check to see if Array CommPath & Failover Mode are disabled/set to 0, and apply the default settings to each Host Initiator record one a time so as not to disrupt HA paths.

#### DATA MOVER, NAVISPHERE, & REGISTRATION PROCESS:

Windows Hosts automatically register their WWNs and IP address to the Array, and run as a Navi Server with two-way communication between the array. The Celerra Data Movers, on the otherhand, do not function as a Navi Server and register their WWNs via Hostnames, with only a one-way communication for the registration process. Because of this limitation (See AR99780), when trying to use the Storage Wizards in Flare 24 or higher, to create or allocate additional backend storage for the Celerra, this will fail because the Data Movers are not registered as Servers that run the real Navi agent, and the Data Movers simply do not show up in the Wizard lists. The proposed fix to this issue would use the Control Station Hostname & IP Address as the basis to provide a Navi-Server functionality so Wizards would work, and Data Mover WWNs would become registered via IP address and not just Hostname.

To make matters more confusing, whenever trying to manually Register Data Mover WWNs, the IP address field requires that a unique IP address be entered [though it isn’t used and can be any bogus IP address], along with the Hostname for the Initiator record. However, the factory installation of the new NS20/40 Integrateds bypassing the IP address problem and uses only the Hostname, meaning that an IP address will not be seen in the properties of the Host Initiator record.

#### Agent tab under Host Initiator Record Properties:

<hostname> is not a Managed Host, unable to retrieve Remote Agent Configuration.

### SETTING CORRECT SYSTEM VALUES FOR CELERRA/CLARIION SYSTEMS:

#### SETTING NO\_TRESPASS=0:

- 1.) `param clariion no_trespass=0` (set in /nas/site/slot\_param)
- 2.) `$.server_config server_2 -v "param clariion no_trespass=0"`
- 3.) Reboot Server

**Note:** Default value 0 for NAS 5.2.16.x --disables the Clariion LUN trespass mechanism, allowing Celerra Hosts to control LUN failover

#### SETTING AUTO-ASSIGN:

```
#/nas/sbin/navicli -h spa chglun -l 00 -a 1 [ALU lun number—run command to change each Celerra LUN]
```

#### SETTING AUTO-ASSIGN FOR CAPTIVE & NON-CAPTIVE BACKENDS:

- 1.) Run following While True Loop Script for Captive Systems

```
#for i in <Celerra LUN numbers>
```

```
>do
```

```
>/nas/sbin/navicli -h spa chglun -l $i -a 1
```

```
>done
```

- 2.) Run Storagegroup List command for non-captive backend and use ALU column for LUN numbers and repeat Step 1 script

```
#/nas/sbin/navicli -h spa storagegroup -list
```

**Note:** Non-Captive systems use AccessLogix & Storage Groups to control Host access to luns.

#### SETTING ARRAYCOMMPATH & FAILOVERMODE ON CAPTIVE & NON-CAPTIVE BACKENDS:

```
#/nas/sbin/navicli -h spa arraycommpath=0 (Captive Systems)
```

#**/nas/sbin/navicli -h spa failovermode=0** (Captive Systems)

#**/nas/sbin/navicli -h spa storagegroup -sethost -host ns600g\_dm30\_p0 -arraycommpath 0 -failovermode 0**

**0** (See new procedure outlined above for conducting with Celerra online but Production Server failed over)

**Caution:** With care, ArrayCommPath settings can be changed while the Celerra is online. However, the correct method for changing Array Commpath values with Servers online would be to change the values first on the Standby Server, then failover each production server and change the values on the Host Name Records of the "faulted" Server slot. In otherwords, the records should only be changed if the physical Server slot is not in production! If the changes are made to a Server while in a production status, there is a potential that IO could be mishandled, resulting in file system corruption, Data Mover panics, etc. If there is any question about this cautionary note, please contact your local EMC Support person.

**Note:** Storagegroups with Celerra Hosts must have arraycommpath & failovermode set to disabled. Repeat for each Celerra Host listed from “port –list –all” on Non-Captive backends [NS600G/NS700G, etc]

## **CHANGE DEFAULT LUN OWNERSHIP:**

#**/nas/sbin/navicli -h 192.168.59.71 chglun -l 0 -d 0** [0=SPA, 1=SPB]

# **/nasmcd/sbin/navicli -h 10.241.168.57 chglun -l 100 -d 1**

Error: chglun command failed

Error returned from Agent

The attribute(s) cannot be set for private lun(s)

**Note:** Cannot change default ownership from navicli for Private Write Intent Log LUNs

## **CELERRA LUN TRESPASS MECHANISM**

### **HOW CELERRA PERFORMS LUN TRESPASS WITH CLARIION:**

→DART uses scsi mode sense commands to trespass LUNs on SPs, not navicli. If an SP is failing, DM should receive sense key information of the failure, and after a number of retries, should drive IO down an alternate path & trespass the LUN. If Data Mover gets SK/ASC 05/04/00 scsi sense command, it will know that the LUN needs to be trespassed and will issue the trespass command to SP and re-issue the IO. Essentially, with setting failovermode=0 for DART LUNs, we are allowing the non-owning SP to reject any media access commands with a Check Condition 05/04/00 sense code error, meaning that DART will then know that it has to initiate the trespass to the SP in order to be able to drive the IO down the alternate path, in this case, to the other SP.

→Issue can occur with hardware failures, such as LCC card, in which sense key information is returned from both SP's that the lun is no longer owned, and lun could be trespassed from one SP to the original owner SP.

## **CELERRA INTEGRATED(Captive) vs. GATEWAY(SAN/Switch) CLARIION ARRAYS:**

### **CELERRA INTEGRATED CLARIION ARRAYS:**

#### **CAPTIVE = INTEGRATED = DIRECT-ATTACHED = DEDICATED STORAGE = NS600/NS700**

**Note:** All of this terminology is being supplanted by the use of “Unified Storage” to indicate the Celerra Integrated platform that can support all protocols (FC, iSCSI, NAS IP, MPFS)

# **nas\_storage -query:\*** -fields:IsCaptive -format: %s\n

False [returns false if not an Integrated, though reports False even for FC models]

--An integrated Celerra/Clariion solution means that the Storage Array system is dedicated to Celerra Servers only—No other Hosts are connected to the Array

**Note:** This definition has changed now that FC-enabled models are allowed [NS20FC/NS40FC], where certain Integrateds are allowed to “share” the backend array with other Hosts

--Uses CX600 or CX700 Arrays, with (2) SP's, (2) SPS's, and (1) or more DAE2 Shelves

--An Integrated Array does not use Optical SAN ports

**Note:** NS80's, NS20's, & NS40's now all use Fibre Optic SFPs to connect directly to the SPs.

--An Integrated system delivers NaviEventMonitor backend events to the Celerra as configured by the /nas/sys/nas\_mcd.cfg file using navilog\_mon, but only a few events will generate a CallHome—Events are posted in the sys\_log

--Data Movers are direct-connected to the Array using copper FC Arbitrated loop cabling [except all new models use Fibre Optic SFPs]

--Never Uses Access Logix

**Note:** Again, this is no longer true. Storagegroups and AccessLogix are used for NS80, NS40, & NS20 systems.

--Uses AUX modules

--Connects from DM's BE0 ports to SPA, while BE1 ports connect to SPB, using FC AL Copper cabling

**DM2-BE0 – AUX 0 SPA-2**

**DM2-BE1 – AUX 0 SPB-2**

**DM3-BE0 – AUX 1 SPA-3**

**DM3-BE1 – AUX 1 SPB-3**

--Connects to AUX0 for Backups using MIA and optical cabling from DM  
 --Integrated Arrays can be upgraded to SAN Gateway Arrays via SW and HW changes

--Change Internal IP address scheme per procedure found in Integrated Array Setup Guide

**Note:** Integrated NS Series systems are connected only to Clariion Storage Systems [NS700/NS500]

**IDENTIFYING LEGACY INTEGRATED SYSTEMS:****#cat /etc/hosts |grep -i APM**

```
192.168.1.200 A_APM00040601637 SPA # CLARiiON SP [Usually a captive backend will retain default IP structure for SP's]
192.168.1.201 B_APM00040601637 SPB # CLARiiON SP
```

**# nas\_storage -i APM00040601637** [Serial number of array from /etc/hosts]

```
id          = 1
arrayname   = APM00040601637
name        = APM00040601637
model_type  = RACKMOUNT
model_num    = 600 [Array Model Number: CX600]
```

**num\_storage\_grps = 0** [Lack of Storagegroups indicates a captive backend—no shared SAN]

**captive\_storage = True** [Captive = True for dedicated array for single host]

**# /nas/sbin/navicli -h 192.168.1.200 storagegroup -list**

Error: storagegroup command failed

This version of Core Software does not support Access Logix

**Note:** Also a key indicator as only SAN-attached uses Storagegroups & Access Logix

**# cat /nas/sys/nas\_mcd.cfg** (Gateway models do not use navilog\_mon for NaviEventMonitor)

daemon "Navi Event Monitor"

```
executable  "/nas/sbin/navilog_mon"
optional    no
autorestart yes
```

**IDENTIFYING FILE SYSTEMS ASSOCIATED WITH SPECIFIC LUN ON CAPTIVE SYSTEMS:****CELERRA FILE SYSTEM LUNS/DETERMINING FILE SYSTEM LUNS, etc:****Determining Which File Systems are built on Lun 27(hex=001B)?****# nas\_disk -l**

```
id  inuse sizeMB storageID-devID type name      servers
1   y     4095  APM00040601637-0000 CLSTD root_disk  1,2
2   y     4095  APM00040601637-0001 CLSTD root_ldisk 1,2
3   y     2047  APM00040601637-0002 CLSTD d3       1,2
-----output abridge-----
17  y     912405 APM00040601637-001B CLATA d17      2,1
18  y     912405 APM00040601637-001D CLATA d18      2,1
```

**Note:** Since this is a captive system without Storage Groups, the concept of HLU/ALU does not apply and the nas\_disk -list will give you the true LUN number in HEX for every dvolume on the Celerra. Convert the HEX number to DECIMAL. Device ID 001B is HEX for decimal 27, or LUN 27 on the Clarion array, using dvolume “d17”. So, any filesystem with “d17” is built in part from LUN 27.

**# nas\_fs -i ata**

```
id      = 24
name    = ata
symm_devs = APM00040601637-001D,APM00040601637-001B
disks   = d18,d17
```

**# server\_devconfig server\_2 -p -s -a**

```
chain= 0, scsi-0
tid/lun= 1/11 type= disk sz= 0 val= -5 info= DGC RAID 5 02061B0000DBCACL
chain= 16, scsi-16
tid/lun= 1/11 type= disk sz= 912405 val= 17 info= DGC RAID 5 02061B0000DBCACL
```

**Note:** Devconfig output will show Controller, Target, and Lun that “d17”, LUN 27, is mapped to from Celerra perspective—in this example SPB owns the LUN on Chain16.

**CELERRA GATEWAYS:****SAN = GATEWAY = FABRIC-CONNECTED = NS600G/NS700G/NS700G 4-Way**

--Gateway configurations allow other Hosts, outside of Celerra, to share the Clariion Array via a SAN network



**Note:** Only a SAN-attached system will contain Storage Groups, with HLU and ALU pairs for LUNs. To find LUNs on Integrated, use nas\_disk -list and take the device number (hex) and convert to decimal value, which then equals ALU/HLU LUN number.

### # /nas/sbin/navicli -h 10.241.169.35 getagent

Revision: 2.04.1.60.5.007 → Flare code shows value of 1 in third field, indicating use of Access Logix to manage Storage Groups for SAN-attached arrays

### # /nas/sbin/navicli -h 10.241.169.35 storagegroup -status

Data Access control: ENABLED → Means that Access Logix is installed

**Note:** Max. of 256 luns can be assigned to a storagegroup

### # /nas/sbin/setup\_backend/setup\_clariion2 list config APM00034202048 [Shows overall status of Backend]

System 10.241.168.52 is up

System 10.241.168.53 is up

Clariion Array: APM00030600872 Model: CX600 Memory: 2048

Enclosure(s) 0\_0,1\_0,0\_1,1\_1,0\_2,1\_2 are installed in the system.

Enclosure info:

-----  
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14  
-----

1\_2: 72 72 72 72 72 72 72 72 72 72 72 72 72 72 72 72  
FC UB  
-----abridged-----

Disk group info:

-----  
Disk Group ID: 0 r5 Disks: 0\_0\_0,0\_0\_1,0\_0\_2,0\_0\_3,0\_0\_4  
Disk Group ID: 2 r5 Disks: 0\_0\_6,0\_0\_7,0\_0\_8,0\_0\_9,0\_0\_10  
Disk Group ID: 3 r5 Disks: 1\_0\_0,1\_0\_1,1\_0\_2,1\_0\_3,1\_0\_4  
Disk Group ID: 4 r5 Disks: 1\_0\_5,1\_0\_6,1\_0\_7,1\_0\_8,1\_0\_9  
Disk Group ID: 5 r5 Disks: 1\_0\_10,1\_0\_11,1\_0\_12,1\_0\_13,1\_0\_14  
Disk Group ID: 8 r5 Disks: 0\_1\_0,0\_1\_1,0\_1\_2,0\_1\_3,0\_1\_4  
-----

Lun info:

-----  
Lun ID: 0 RG ID: 0 State: Bound root\_disk  
Lun ID: 1 RG ID: 0 State: Bound root\_ldisk  
Lun ID: 2 RG ID: 0 State: Bound d3  
Lun ID: 3 RG ID: 0 State: Bound d4  
Lun ID: 4 RG ID: 0 State: Bound d5  
-----abridged-----

Spare info:

-----  
Spare ID: 1 Disk: 0\_0\_5  
Spare ID: 9 Disk: 0\_0\_11

## NAS 6.0 Fibre Channel Link Settings are all Autonegotiate:

# .server\_config server\_2 -v "param fcTach" |grep linx  
fcTach.linx\_speed\_aux0 0x0005f4a790 0x00008000 0x00008000  
fcTach.linx\_speed\_aux1 0x0005f4a794 0x00008000 0x00008000  
fcTach.linx\_speed\_be0 0x0005f4a788 0x00008000 0x00008000  
fcTach.linx\_speed\_be1 0x0005f4a78c 0x00008000 0x00008000

## NS700 (integrated) & NS700G (gateway) PRODUCTS:

GA 9 February 2004. Midrange storage solution running DART 5.2 on Clariion CX700 or Symmetrix backends.

--SP speeds increased to 3.0Ghz and 533 Bus speed & NS700-AUX memory is 8GB.

--NS700, NS701-FD, NS702, NS702-FD consists of a “captive NS700 NAS head” [Data Mover], an AUX module, a Control Station, and Disk Array Enclosures

--NS700 SPs have 4 FC ports, two for Backend loops and two for Data Movers

**NS704G 4-Way:** GA Aug 4, 2004 → Compatible with NAS 5.3, 4-Data Mover and Dual-CS model

--NS704 connects to backend where SPs can support 4 backend loops and 4 Blades (8 FC ports/SP)

**NS700G PHASE I:** NS700 that can share some ports on array with a SAN. Data Movers in this model are direct-connected to CX700, but CX700 contains extra fibre ports connected to a switch to allow for other external hosts to access storage. Uses ‘Access Logix’ and StorageGroups to restrict access to LUNS to specific Hosts.

**NS700G PHASE II:** All hosts, including data mover, connect to fibre switch to access the CX700

**NS700:** Contains new ‘Ajaguar’ (9) network port NAS personality module with (7) copper GB ports and (2) optical GB ports

**WWN:** Each data mover has unique WorldWideName seed stored in PROM. Prefixed numbers for WWN will be different if plugged into a different chassis.

### **STORAGE GROUPS:**

Storage Groups contain WWNs and LUNs and Host fibre port HBA Initiators. NS700G requires use of Storage Groups. Storage Groups provide for LUN masking for specific Hosts. The LUN numbers presented to one Host in one Storagegroup could be reused for a different Host in a different StorageGroup. Use storagegroup –list

### **RESERVED CELERRA LUNS:**

LUNs 0-15 are reserved on the Celerra. Data LUNs should always be created on LUNs 16 and higher. NAS 5.3 prevents the illegal configuration of data LUNs on Celerra Reserved Luns.

### **CLARIION RESERVED LUNS:**

**\$ /nas/sbin/navicli -h 10.241.168.52 reserved -lunpool-list -freeluns**

Name of the SP: SP A

Unallocated LUNs: None

### **CELERRA NBS SERVICES:**

**Note:** NBS (Network Block Storage) is an EMC proprietary iSCSI technology that enables NBS network clients (such as the Control Station) to access block storage devices connected to NBS network servers (Data Movers). NBS transactions are more efficient than other protocols because it uses block data transfers. NBS is implemented in a Client-Server model, where pseudo disk instances of each storage object are exported to the Client by the Block Server--presenting remote storage devices as local disk devices. Client block device driver receives requests from applications and pass to Block Server using TCP. Block Server processes Client requests and replies. NBS version 2.0 with NAS 5.3. NBS runs over TCP. CS NBS Client is used to connect to backend “nas” & “dos” partitions using Data Mover as NBS Server.

### **COMPONENTS:**

Client & Server NBS processes

Internal Celerra network

Storage volumes exported by Data Mover

### **CONTROL STATION BOOT OPERATION & ARRAY COMMS WITH CLARIION SYSTEMS:**

The Celerra Control Station boots directly from the Storage System drives with the CFS-14 and CNS14 families. With the NS-family, however, such as the NS600 & NS700, the Control Station boots from an internal IDE drive and not from the Storage Array drives [though it does store a copy of its configuration data on the array]. The NS-family Control Station does not directly copy data to the Backend Array and does not have a Fibre Channel host bus (HBA), but uses Network Block Service (NBS) to access LUNs on the Array. The implications here are that the CS is dependent on NBS (a network service that runs on the Data Mover) to read and write to its database configuration on the Array, and also to conduct a NAS software install.

The Control Station sends NBS data over the Private LAN to the Data Mover, the latter which then forwards the information to the Array over Fibre Channel. NBS runs as a Client on the CS and as a Service on the DM. NBS is used only for block data.

### **VERIFYING NBS SERVICES:**

**# ps -ef |grep nd-clnt**

```
root 1062 1 0 Oct13 ? 00:00:02 [nd-clnt 0 1]
root 1067 1 0 Oct13 ? 00:00:06 [nd-clnt 4 5]
root 1073 1 0 Oct13 ? 00:00:00 [nd-clnt 5 6]
```

### **VERIFYING CS CONNECTIVITY TO NBS DEVICES:**

**# /sbin/fdisk -l /dev/ndf1**

```
Disk /dev/ndf1: 255 heads, 63 sectors, 229 cylinders
Units = cylinders of 16065 * 512 bytes
```

### **NBS SERVICES ON NS702G:**

**# ps -ef |grep nd-**

```
root 1210 1 0 Oct18 ? 00:00:55 [nd-clnt 0 1] → Process that retrieves NBS data from DM/backend
root 1211 1 0 Oct18 ? 00:44:29 [nd-clnt 4 5]
root 1212 1 0 Oct18 ? 00:01:16 [nd-clnt 5 6]
root 1213 1 0 Oct18 ? 00:01:49 [nd-sndd 0 1] → Process that sends NBS data to DM/backend
root 1214 1 0 Oct18 ? 00:03:19 [nd-sndd 5 6]
root 1215 1 0 Oct18 ? 02:08:43 [nd-sndd 4 5]
```

**# /nas/sbin/t2tty -C 2 "nbs info" (also use .server\_config -v to output same thing)**

1163176250: NBS: 4: \*\*\*\* Block Server \*\*\*\*

1163176250: NBS: 4: T5.5.24.2-NBSv6

1163176250: NBS: 4: Listening on port 5033  
1163176250: NBS: 4: 20 threads started  
1163176250: NBS: 4: TCP hi watermark: 524288  
1163176250: NBS: 4: Using default TCP low watermark

**# /nas/sbin/t2tty -C 2 "volume info NBS1"**

\*\*\*\* Basic Volume 1 : 0x1100204 Information: \*\*\*\*  
Total References:.....0x000b  
Total Blocks:.....0x15fff80  
Bytes Per Block:.....0x0200

**# .server\_config server\_3 -v "nbsid list"**

1163176381: NBS: 4: nbs add name=1 vol=NBS1 rw=192.168.1.100:192.168.1.101:192.168.2.100:192.168.2.101 share exclusive raw  
1163176381: NBS: 4: nbs add name=5 vol=NBS5 rw=192.168.1.100:192.168.1.101:192.168.2.100:192.168.2.101 share exclusive raw  
1163176381: NBS: 4: nbs add name=6 vol=NBS6 rw=192.168.1.100:192.168.1.101:192.168.2.100:192.168.2.101 share exclusive raw

**# /nas/sbin/t2tty -C 2 "nbsstat list"**

1163182039: NBS: 4: \*\*\*\*\*  
1163182039: NBS: 4: BlkSrvStat: Avg Rd IO size:0 kBytes, Avg Wr IO size:9 kBytes  
1163182039: NBS: 4: BlkSrvStat: RDcnt:0, RD Success:0, RD Fail:0  
1163182039: NBS: 4: BlkSrvStat: WRcnt:317, WR Success:317, WR Fail:0  
1163182039: NBS: 4: BlkSrvStat: RD:0, WR:585, RW:0 Kb/sec (output abridged)

**RECONFIGURING NBS ON A DM:**

**/nasmcd/sbin/setup\_slot -nbs -add 2**

**# cat /etc/nbs.conf**

```
#  
# Simple configuration file for nbs service  
#  
# Format: devIndex:NbsId:Host1,host2,...:  
0:1:server_2,server_3b,server_2b,server_3:  
4:5:server_2,server_3b,server_2b,server_3:  
5:6:server_2,server_3b,server_2b,server_3:
```

**# cat /etc/fstab**

```
/dev/nde1      /nbsnas    ext3    noauto,rw,sync 0 0  
/dev/nda1      /nas/dos     msdos   umask=002,noauto,rw,sync,gid=201    0 0  
/dev/ndf1      /nas/var     ext3    noauto,rw    0 0
```

**# cat /nas/server/slot\_2/nbs.cs**

```
volume disk NBS1 c0t0l0  
volume disk NBS5 c0t0l4  
volume disk NBS6 c0t0l5  
volume disk RDF7 c0t0l6  
volume disk RDF10 c0t0l9  
volume disk NBS1 c16t0l0  
volume disk NBS5 c16t0l4  
volume disk NBS6 c16t0l5  
volume disk RDF7 c16t0l6  
volume disk RDF10 c16t0l9  
nbsid add nbsid=1 vol=NBS1 exclusive raw share  
exportnbs add ip=192.168.1.100 nbsid=1  
nbsid add nbsid=5 vol=NBS5 exclusive raw share  
exportnbs add ip=192.168.1.100 nbsid=5  
nbsid add nbsid=6 vol=NBS6 exclusive raw share  
exportnbs add ip=192.168.1.100 nbsid=6  
nbsid add nbsid=7 vol=RDF7 exclusive raw ro=192.168.1.100:192.168.1.101  
exportnbs add ip=192.168.1.100 nbsid=7  
nbsid add nbsid=10 vol=RDF10 exclusive raw ro=192.168.1.100:192.168.1.101  
exportnbs add ip=192.168.1.100 nbsid=10  
nbs start
```

**STOPPING & RESTARTING NBS SERVICES:**

**#/sbin/service nbs stop | start**

**Note:** Above is one way to stop and restart NBS.

Another method might be less intrusive and doesn't require PXE booting of servers to restore access to the backend:

### # server\_cpu ALL -q now

**Note:** This command seems to stop NBS access to the backend from the Servers [Data Movers], but does not actually stop the NBS client processes on the Control Station [ps -ef |grep nd- still shows client services running], meaning that to recover NBS fully does not require PXE booting, just a simple reboot of any of the Data Movers.

### VERIFYING NBS VOLUMES:

#### # /sbin/service nbs status

Configured NBS devices:

```
254: 0  NbsId: 1
Disk_Name: nda
Capacity: 11534272 KB
Crnt_Server: 0X202a8c0
Server_List: 0X302a8c0 0X202a8c0 0X301a8c0 0X201a8c0
num_io: 48095   time 89843750 ms num_sect 1481313 que 0
Avg_blk: 15 K  throughput 8 K/s resp_tm: 9 ms, ops/sec 0 (output abridged)
```

#### # /sbin/service --status-all

Configured NBS devices:

##### 254: 0 NbsId: 1

```
Disk_Name: nda
Capacity: 4194240 KB
Crnt_Server: 0X201a8c0
Server_List: 0X301a8c0 0X201a8c0
num_io: 27211   time 1201536430 ms num_sect 390684
Avg_blk: 7 K  throughput 0 M/s resp_tm: 1 ms, ops/sec 0
```

##### 254: 16 NbsId: 5

```
Disk_Name: nde
Capacity: 2097088 KB
Crnt_Server: 0X201a8c0
Server_List: 0X301a8c0 0X201a8c0
num_io: 178982   time 1201536440 ms num_sect 2848996
Avg_blk: 7 K  throughput 0 M/s resp_tm: 5 ms, ops/sec 0
```

##### 254: 20 NbsId: 6

```
Disk_Name: ndf
Capacity: 2097088 KB
Crnt_Server: 0X201a8c0
Server_List: 0X301a8c0 0X201a8c0
num_io: 96341   time 1201536450 ms num_sect 5731596
Avg_blk: 29 K  throughput 0 M/s resp_tm: 11 ms, ops/sec 0
```

#### # cat /etc/nas\_device.map

##### DOSDSK=/dev/nda

OS1DSK=/dev/hda

OS2DSK=/dev/hda

VERDSK=/dev/hda

##### VARDISK=/dev/ndf

##### NBSDSK=/dev/nde

ENET\_INT0=eth0

ENET\_INT1=eth0:0

ENET\_EXT=eth1

ENET\_IPMI=

ENET\_SP=

#### #/nas ls -la \*

```
lrwxrwxrwx 1 nasadmin nasadmin 11 Dec 7 13:58 dev -> /nbsnas/dev
lrwxrwxrwx 1 root  root 11 Sep 20 16:22 dos -> /nbsnas/dos
lrwxrwxrwx 1 root  root 13 Dec 7 13:58 dosfs -> /nbsnas/dosfs
lrwxrwxrwx 1 root  root 15 Dec 7 13:58 install -> /nbsnas/install
lrwxrwxrwx 1 root  root 12 Dec 7 13:58 lock -> /nbsnas/lock
```

**Note:** Point of above output is to show that most directories & files in /nas are just links to /nbsnas

#### # /sbin/fdisk -l

Disk **/dev/nda**: 255 heads, 63 sectors, 1435 cylinders →/nas/dos

Units = cylinders of 16065 \* 512 bytes

Device Boot Start End Blocks Id System

/dev/nda1 \* 1 17 136521 6 FAT16

Disk **/dev/nde**: 255 heads, 63 sectors, 261 cylinders →/nbsnas

Units = cylinders of 16065 \* 512 bytes

Device Boot Start End Blocks Id System

/dev/nde1 1 230 1847443+ 83 Linux

Disk **/dev/ndf**: 255 heads, 63 sectors, 261 cylinders →/nas/var

Units = cylinders of 16065 \* 512 bytes

Device Boot Start End Blocks Id System

/dev/ndf1 1 230 1847443+ 83 Linux

**#/sbin/fdisk -l /dev/nda | /dev/nde | /dev/ndf** [To verify access to the partitions]

#### **TYPICAL BIND TABLE:**

**\$ .server\_config server\_3 -v "fcp bind show"**

\*\*\* Persistent Binding Table \*\*\*

Chain 0000: WWN 500601610060086d HBA 0 SP-a1 Bound

Chain 0016: WWN 500601690060086d HBA 1 SP-b1 Bound

Chain 0032: WWN 5006048000000000 HBA 2 N\_PORT Bind Pending

Chain 0048: WWN 5006048000000000 HBA 3 N\_PORT Bind Pending

\*\*\* Dynamic Binding Table \*\*\*

Chain 0000: WWN 500601610060086d HBA 0 ID 0 Inx 00:00 Pid 0000 S\_ID 0000ef Sys

Chain 0016: WWN 500601690060086d HBA 1 ID 1 Inx 01:00 Pid 0016 S\_ID 0000ef Non

Chain 0032: WWN 0000000000000000 HBA 2 ID 2 Inx 02:81 Pid 0032 S\_ID 000000 Non

Chain 0048: WWN 0000000000000000 HBA 3 ID 3 Inx 03:81 Pid 0048 S\_ID 000000 Non

**Note:** NBS must be configured on Chains 0 and 16 on direct-connect captive systems or else NBS may not function correctly. With Fabric-connected systems, NBS can be on Chains 0, 16, 32 & 48.

#### **CONTROL SYSTEM LUNS FOR CELERRA (Celerra Control LUNs):**

**Note:** LUNS define logical disks on the array presented to Hosts, and usually spans a group of physical disks known as a “RAID GROUP”. Several LUNS can comprise a Raid Group.

→NS500/NS700 LUNS 0-5 are always RAID GROUP 0 Raid5 4+1 & are located on first (5) disks on first shelf in array, and should be only LUNs in this Raid Group.

#### **LUN 0: /nas/dos or /nbsnas/dos**

/nas/dos partition on array from which Data Movers boot—actually a symbolic link to /nbsnas/dos for /dev/nda1

Data Movers use this partition to boot and load DART and also to obtain individual configuration files

#### **NAS PARTITION:**

/nas is actually mounted on the Control Station’s IDE drive on /dev/hda5, but many of the configuration database files/folders are symbolic links to the NBS partition on the array known as the “NASDB”, or NAS configuration database. So, if the Control Station is not able to talk to an NBS Server to actually get to the backend partition via the symbolic links, then the locally mounted /nas directory will contain only some configuration directories and files. Note that during normal operation, a “dirsync” process runs on the Control Station which copies any changes from /nas to /nbsnas.

#### **LUN 1: Used by data movers**

#### **LUNS 2 & 3: Reserved**

#### **LUN 4: /nbsnas**

/nbsnas is the mountpoint for the NBS partition /dev/nde1 on the array and contain mountpoints for “dos” & “var” partitions

/nbsnas contains NAS configuration database and is known as the home for the “NASDB” [/nbsnas/server, etc.]

#### **LUN 5: /nas/var symbolic link to /nbsnas/var**

/nbsnas/var is the mountpoint for the NBS partition /dev/ndf1, which contains the NASDB backups, dumps, log files, & other files

**Note:** To find correct array LUN for the Control Luns, run storagegroup -list and MAP HLU LUNs 0-5 to the ALU number on array

#### **DIRSYNC PROCESS TO KEEP NAS & NBSNAS PARTITIONS SYNCHRONIZED:**

**\$ ps ax |grep dirsync**

1701 ? S 1:16 /bin/sh /nas/sbin/dirsync /nas /nbsnas 180

1702 ? S 1:23 /bin/sh /nas/sbin/dirsync -c /nas/sys /nasmcd/CHomeFi

**Note:** There can be many different reasons for /nas or /nbsnas, or both, to become full--dumpfiles extracted to /nas, logfiles that have grown disproportionately large, or JServer database files have grown too large. To complicate matters, the NS Series Celerra, which runs NBS Services for Control Station access to the backend, runs a Dirsync process on the Control Station to keep the /nbsnas

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
partition in sync with the /nas partition. One of the quirks of this file syncing process is that you cannot properly delete files from either partition to reduce the file system capacity without stopping NAS Services first to stop the dirsync process, as files will be rewritten to the partition(s) from cache, even if you have manually deleted them and df -h shows healthy partition space.  
→Running pstree against NAS MCD process shows hierarchy of processes. “Dirsync” is a process designed to sync any file changes on the local /nas partition with the backend /nbsnas partition, every 3 minutes. It also syncs any changes from /nbsnas/sys to the local /nasmcd/CHomeFiles directory, every 5 minutes.

**# pstree -pau 25137 | grep dirs**

```
dirs,30078 /nas/sbin/dirs /nas /nbsnas 180
dirs,30118 /nas/sbin/dirs -c /nas/sys /nasmcd/CHomeFiles 300
```

## **CLEANING UP SPACE ON /NAS & /NBSNAS FOR NS CELERRAS:**

1. Run #df -h before any changes

**# df -h**

| Filesystem | Size | Used | Avail | Use% | Mounted on |
|------------|------|------|-------|------|------------|
| /dev/nde1  | 1.7G | 1.8G | 0     | 100% | /nbsnas    |
| /dev/hda5  | 2.0G | 2.0G | 0     | 100% | /nas       |

2. Stop NAS Services as Root

**Caution:** For dual-control station NS systems, you must first stop NAS Services on the Standby Control Station (CS1), then stop NAS on CS0/.

**# /sbin/service nas stop**

/nbsnas is in use by the following processes(uid)

```
/nbsnas: 25444 26044 26052 26053 26056
```

NAS:/etc/init.d/nas: ERROR: Failed to stop NAS services

**Note:** Please be aware that if you or another User are sitting in the /nas or /nbsnas partition, NAS may give an error saying it could not be properly stopped. In all cases, you must verify that both Box Monitor and NAS Services are stopped before continuing. The most likely situation is that Box Monitor will be stopped, but NAS will remain running--if you find this to be the case, you will need to kill the NAS processes manually:

**# ps -ef |grep boxm**

```
root 27221 24916 0 11:01 ttyS0 00:00:00 grep boxm
```

**# ps -ef |grep -i nas\_mc**

```
root 1189 1 0 Nov13 ? 00:00:00 /nasmcd/nas_mcd -h /nasmcd /nas/
root 1363 1189 0 Nov13 ? 00:00:00 /nasmcd/nas_mcd -h /nasmcd /nas/
root 1364 1363 0 Nov13 ? 00:00:00 /nasmcd/nas_mcd -h /nasmcd /nas/
```

**# pkill nas\_mcd**

pkill: 1484 - No such process

# ps -ef |grep nas\_mc

```
root 1189 1 0 Nov13 ? 00:00:00 /nasmcd/nas_mcd -h /nasmcd /nas/
root 1363 1189 0 Nov13 ? 00:00:00 [nas_mcd <defunct>]
```

**# kill -9 1189**

**# ps -ef |grep nas\_mc**

3. Manually remount partitions on Control Station:

**#mount /nbsnas #mount /nas #mount /nas/dos #mount /nas/var**

4. Delete necessary files from /nas, then the same files from /nbsnas--one way to determine what directories in /nas or /nbsnas are contributing to the full condition is to run the following from within the /nas or /nbsnas directory:

**# du -sh /\***

**Note:** In some cases you will find that JServer logs are contributing to most of the space being used on /nas & /nbsnas. In this situation, you will need to obtain customer approval before running #js\_cleandb. Also be aware that in older NAS Versions, #js\_cleandb only deleted files from the /nas partition--you must still manually delete the odb and sdb directories from /nbsnas/jserver. Finally, in many situations, log files have grown disproportionately large and can be zeroed out by running the following, but please do not run this against log files if you do not know the impact:

**# ls -la**

total 39048

```
-rw-r--r-- 1 root root 22827157 Nov 9 09:25 access_log.5
```

**# cp -i /dev/null access\_log.5**

cp: overwrite `access\_log.5'? y

**# ls -la**

```
-rw-rw-rw- 1 root root 0 Apr 11 2002 access_log.5
```

5. When partition cleanup with completed, run df -h to verify:

**# df -h**

| Filesystem | Size | Used | Avail | Use% | Mounted on |
|------------|------|------|-------|------|------------|
| /dev/nde1  | 1.7G | 857M | 828M  | 51%  | /nbsnas    |
| /dev/hda5  | 2.0G | 778M | 1.1G  | 41%  | /nas       |

**Note:** Keep in mind that just because the dirsync process is designed to keep /nbsnas in sync with any changes on /nas, the two partitions are not necessarily identical in size

6. Restart NAS services & verify that Box Monitor, NAS, and dirsync are running:

**# /sbin/service nas start****# ps -ef |grep nas\_m**

```
root 372 1 0 11:40 ? 00:00:00 /nasmcd/nas_mcd -h /nasmcd /nas/
root 483 372 0 11:40 ? 00:00:00 /nasmcd/nas_mcd -h /nasmcd /nas/
-----output abridged-----
```

**# ps -ef |grep boxm**

```
root 938 372 0 11:41 ? 00:00:00 /nas/sbin/nas_boxmonitor /nas -i
root 998 938 0 11:41 ? 00:00:00 /nas/sbin/nas_boxmonitor /nas -i
root 999 998 0 11:41 ? 00:00:00 /nas/sbin/nas_boxmonitor /nas -i
-----output abridged-----
```

**# ps -ef |grep dirsync**

```
root 25439 25312 0 10:58 ? 00:00:00 /bin/sh /nas/sbin/dirsync -skipj
```

**CONTROL STATION NETWORK INTERFACES:****NS600 CONTROL STATIONS:**

→eth0 as lower network port connected to local network switch  
→eth1 as upper network port connected to public network

**NS700 FALCON CONTROL STATIONS:**

→eth0 2<sup>nd</sup> port from left, connecting to primary local network switch  
→eth1 1<sup>st</sup> port from left, connects to cross-over to eth1 on CS1  
→eth2 3<sup>rd</sup> port from left, connects to secondary local network switch  
→eth3 4<sup>th</sup> port from left, connects to public network

**CONTROL STATION SERIAL PORTS:**

**Note:** Serial port devices are accessed and read via /dev/ttySx

/dev/ttyS0 Located on back of CS for serial modem access, COM1

/dev/ttyS1 Located on front of CS for serial console access, COM2

**Note:** Change console redirection from COM1 to COM2 by accessing BIOS

/dev/ttyS4 Located on 4-port serial cable and connects to data mover in slot\_2

/dev/ttyS5 Located on 4-port serial cable and connects to data mover in slot\_3

/dev/ttyS6 Located on 4-port serial cable and connects to data mover in slot\_4

/dev/ttyS7 Located on 4-port serial cable and connects to data mover in slot\_5

**Note:** Use following command to check if processes are using Serial ports—only Box Monitor should be using the ports  
#fuser /dev/ttyS4

**ADDING SECOND CONTROL STATION TO NS700 SYSTEMS TO SETUP CS IPMI NETWORK:**

**Note:** Refer to CNS Hardware Upgrade Procedures Guide, ‘Adding a Control Station to an NS700G 4-Way’ system

1. Connect all cables from CS-1 to DMs and new secondary switch
2. Setup secondary CS network on CS, on DM’s, & reconfigures NBS using:

**#/nasmcd/sbin/upgrade\_to\_dual\_intnet**

**Note:** Above script upgrades from eth0 primary internal network and eth0:0 secondary internal network, to the new CS\_IPMI network (192.168.3.0) that allows for communications between Control Stations, eth0:0 is deleted. All data movers must be up and NAS Services must be stopped or script will fail.

**CS IPMI NETWORK CONFIG:** Control Stations are networked between eth1 on CS0 to eth1 on CS1 for IPMI

**# cat ifcfg-eth1**

```
DEVICE=eth1
IPADDR=10.0.3.100
NETMASK=255.255.255.0
NETWORK=10.0.3.0
BROADCAST=10.0.3.255
ONBOOT=yes
```

**Troubleshooting:**

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
Upgrade logs are put in /tmp directory; Files upgraded are /etc/hosts; /etc/nas\_device.map; /etc/modules.conf; /etc/sysconfig/network-scripts/ifcfg-eth1 & ifcfg-eth2; /nas/site/nas\_param; /nas/server/slot\_x/ifconfig; and /etc/sysconfig/network-scripts/ifcfg-eth0:0 file is deleted. There is a problem noted in ETA emc149610 that the upgrade script can remove all VLAN tag ID's and MAC addresses from the CS & DM interfaces.

3. Power-up CS1 and answer prompts, if no mismatch, setup will complete [Control Station comes with software installed]

## **CLARIION TROUBLESHOOTING/NS TROUBLESHOOTING COMMANDS & TOOLS:**

### **BASIC TROUBLESHOOTING:**

# /nas/sbin/getreason

# /nas/sbin/getboxmask -r

# /nas/sbin/t2reset reboot -s 2

→NASDB\_BACKUP and RESTORE options

→Stop and Restart NAS [#/sbin/service nas stop | start]

### **VERIFY CS & NBS PROCESSES:**

#ps -ef |grep nd- [Should be (3) processes, one for each of the (3) NBS devices that are mounted]

**nd-clnt 0 1 →device /dev/nda [/nas/dos]**

**nd-clnt 4 5 →device /dev/nde [/nbsnas]**

**nd-clnt 5 6 →device /dev/ndf [/nas/var]**

# ps -ef |grep -i nas\_box [Should be minimum of (6) Box Monitor processes]

# /sbin/fdisk -l /dev/nda | /dev/nde | /dev/ndf [To verify access to the partitions]

→Ping data mover

→Stop and start NBS services if necessary & manually mount /nas /nas/dos /nas/var

**# /sbin/service nbs stop | start**

**# /sbin/service nbs status**

Configured NBS devices:

254: 0 NbsId: 1

Disk\_Name: nda

Capacity: 11534272 KB

Crnt\_Server: 0X202a8c0

Server\_List: 0X302a8c0 0X202a8c0 0X301a8c0 0X201a8c0

num\_io: 48095 time 89843750 ms num\_sect 1481313 que 0

Avg\_blk: 15 K throughput 8 K/s resp\_tm: 9 ms, ops/sec 0 (output abridged)

### **USING T2TTY FACILITY TO CHECK NBS:**

**/nas/sbin/t2tty -C 2 “logsys add output console” “camshowconfig” “nbsid list” “nbs info” “fcp bind show” “export” “share” “volume info NBS1” “nbsstat list”**

BlkSrvConfigID: : nbsid:1, idtype: 1, name:S1 . . .

BlkSrvConfigID: : nbsid:5, idtype: 1, name:S5 . . .

BlkSrvConfigID: : nbsid:6, idtype: 1, name:S6 . . .

**# /nasmd/sbin/t2tty -C 2 “logsys set output disk=root\_log\_2”** [slot id of DM—turns off output to console]

**# /nasmd/sbin/t2tty -C 2 “logsys add output console”**

**# /nasmd/sbin/t2tty -C 2 “logsys delete output console”**

### **Data Mover NBS Configuration File:**

**/nas/server/server\_1/nbs.cs**

**/nas/server/server\_2/nbs.cs**

**nbs.cs -> /nas/server/slot\_2/nbs.cs.ro**

### **CS NBS Config File:**

**# cat /etc/nbs.conf**

#

# Simple configuration file for nbs service

#

# Format: devIndex:NbsId:Host1,host2,...:

0:1:server\_2,server\_3b,server\_2b,server\_3:

4:5:server\_2,server\_3b,server\_2b,server\_3:

5:6:server\_2,server\_3b,server\_2b,server\_3:

### **T2TTY COMMANDS:**

→Use t2tty commands to send commands to DART serial console, force PXE boot, etc. [Stop NAS Services first]

→Can also use t2tty facility to send commands to DART via RPC [ t2tty –C 2 “ifconfig”]

#/nasmcd/sbin/t2tty -c 2 “logsys add output console” “ifconfig” “fcp bind show” etc.

#/nasmcd/sbin/t2tty -p 2 [Forces DM to PXE boot from DART kernel stored on CS]

**Note:** Disable Data Mover output to console when completed using “logsys set output disk=root\_log\_x” [x=slot]

#### **REBOOTING DATAMOVER USING PEER POWERCONTROL:**

**# /nas/tools/.server\_peer\_powerctrl -reboot 3** [Use only if other DM is not hung]

#### **MINICOM:**

→Use CS Minicom terminal program to observe data mover boot [Stop NAS Services before using]

**Note:** Setup data mover profile using minicom –s & then run “minicom dm2” to access data mover. Use esc + O + Q keys to access BIOS on bootup. This program is useful with NS Series, since it has a serial connection between CS & DM.

#### **T2PXE BOOTING DATA MOVERS:**

**Note:** CS acts as PXE Boot Server for Data Movers

→Stop NAS Services, Start PXE services on CS [/nasmcd/sbin/t2pxe –s –R 4]

#### **Conduct automatic PXE boot using following:**

#/nasmcd/t2tty –p 2 [NAS 5.2.7 and higher—see t2pxe boot section for more recent PXE Boot procedure]

#### **MANUAL PXE BOOT PROCEDURE TO ENTER POST:**

1. Stop NAS Services and manually remount /nbsnas, /nas, /nas/dos, /nas/var

2. Verify dhcp daemon running on CS: #ps –ef |grep dhcp

3. Reset data mover using t2reset reboot –s 2

4. Startup minicom session to the Data Mover

5. At beginning of the Extended POST, the following key sequence is seen on the screen:

*ABCab << Stopping after POST >> DEabcdeFGHabcdIJK*

a) Anytime during the above sequence, press “ctrl + c”, which will allow POST to complete, and then offers the “Storage System Failure...” message, at which point a password can be entered.

**Note:** “ctrl + c” replaces the previous method of using the “esc” key to interrupt POST. Use the “esc” key if ctrl + c does not work.

6. Enter POST configuration password: SHIP\_it or DB\_key (later versions, DB\_key puts you into FCCBOOT menu)

7. Select 40) ICA Sub-Menu and then 2) PXE Boot

8. Observe bootup and perform commands as necessary to troubleshoot

9. Exit minicom when finished (ctrl + A, then type ‘exit’) and disable PXE booting: #/nasmcd/sbin/t2pxe –e

#### **USING SETUP SLOT:**

**Note:** Use setup\_slot only if access to backend through NBS is available. Can help resolve Reason Code 14 issues.

1. Stop NAS Services

2. Mount /nas /nbsnas /nas/dos /nas/var

3. #/nasmcd/sbin/setup\_slot –init 2 [Will regenerate boot.bat & boot.cfg files & bootup data mover with configuration]

#### **CLARIION CACHE DIRTY CONDITION (aka, UNOWNED LUNS):**

#### **ACCESSING ARRAY VIA XP REMOTE DESKTOP TO CLEAR DIRTY CACHE:**

1. Startup XP Remote Desktop

2. Login using clarion clarion!

3. c:>admin tool

4. Select option 2: Recovery Menu →Clear Cache Dirty LU

Password: 29814

Enter LUN Number

5. FSCK file systems as appropriate

**Note:** If getlun does not show an owner, LUN is unowned. All LUNs that have had cache dirty condition should be FSCK’ed before being mounted back on the Celerra Server.

#### **CLEARING CLARIION DIRTY CACHE WITH FLARE 19+:**

Dirty Cache occurs when an SP cannot properly flush its dirty pages or modified data to disk, due to some unusual event

Dirty Cache means data lost

Flare 19 and higher, Dirty Cache can be fixed by right-clicking Unowned LUN in Navisphere>Bring LUN Online

#### **CLARIION UNOWNED LUNS:**

→All Hot Spares are Unowned LUNs held in a special Raid Group

→Orphaned LUNs are Unowned

→CRU signature error results in Unowned LUN, say a disk is physically removed before unbound

→Dirty Cache can also result in Unowned LUN

#### **CLARIION HOT SPARES:**

→Only invoked for User LUNs

→Drive type of HS must be the same as failed disk, except for Blizzard & Northstar drives, which can Hot swap for each other

→Size of LUNs cannot exceed size of Hot Spare

→Recommendation is to use (1) HS per 20 or 30 drives

→A disk rebuild reconstructs private data in flare db, and reconstructs RAID data on logical unit partitions

### **CHECKING FOR DIRTY CACHE FROM NAVICLI:**

# /nas/sbin/navicli -h 10.241.168.52 luncache -list

LOGICAL UNIT NUMBER 0

LUN Offline (Cache Dirty Condition): NO

### **CLEARING DIRTY CACHE FROM NAVICLI:**

# /nas/sbin/navicli -h 10.241.168.52 luncache 0 -clear

### **NAS 5.4 CLARIION STORAGE API CHECKS & HEALTHCHECK TOOL:**

# nas\_storage -list

# nas\_storage -sync -all | id=1 (nas\_storage -sync -all did not work)

**Note:** Synchronizes the Control Station's view of the backend Storage array with any changes that may have happened on backend

# nas\_storage -check -all | id=1

Discovering storage (may take several minutes)

Error 5010: Storage API code=3593: SYMAPI\_C\_CLARIION\_LOAD\_ERROR:

**Note:** Used as a healthcheck command for backend to check if High Availability is configured and available.

### **DETERMING DRIVE FIRMWARE LEVEL:**

# /nas/sbin/navicli -h 128.221.252.200 getdisk -capacity -product -rev -type -vendor -drivetype

Bus 0 Enclosure 0 Disk 0

Capacity: 68238

Product Id: ST373453 CLAR72

**Product Revision: 8A0C**

Type: 0: RAID5 1: RAID5 2: RAID5 3: RAID5 4: RAID5 5: RAID5 16: RAID5 17: RAID5

Vendor Id: SEAGATE

Drive Type: Fibre Channel

### **PEFORMING DRIVE FIRMWARE UPGRADE ON CELERRA INTEGRATED SYSTEMS:**

→As of this entry, several procedures are being written to address the different Integrated platforms [e.g., CX class, CX3 class, etc.]

→Drive firmware upgrades requires that all Hosts be offline, so Celerra Data Movers are halted for all the procedures

→All procedures involve the use of the NST tool and built-in Disk Firmware Upgrade wizard to perform the drive upgrade

→Basic flow is to identify the IP addresses used on the SPs, halt the data movers but leave the Control Station online, connect to either the Internal Celerra network to access the SPs, or connect directly to the array LAN ports using a mini-switch, along with a service laptop on the same network as the SPs, connect to SPA using the NST tool, perform the DFU activity using the wizard after logging into Eng. Mode (ctrl + shift + f12 SIR), then restore Celerra to operation after the Clariion firmware work is done

### **PERFORMING FIRMWARE UPGRADE ON NS350/500/600/700:**

**Note:** The following steps are an outline only, and may differ from the actual published procedure that will be released

1. Connect to Serial port on Control Station from laptop and open hyper-terminal session

2. Connect to Control Station over IP using SSH

3. Do healthcheck & identify SP IP addresses

/nas/sbin/getreason

/nas/bin/nas\_checkup

tail /etc/hosts

/nas/sbin/model

nas\_version

navicli -h getcrus | getagent | getcache (record Read and Write cache values) | getdisk -product -rev -type -vendor

4. Halt each Data Mover using server\_cpu –halt now, disconnect from SSH session, then remove Data Movers physically from backplane (or unplug the power to the Data Movers)

5. Configure laptop with IP address on Celerra Internal network (usually 192.168.1 or 128.221.252)

6. Plug laptop LAN port into 8-Port Allied Telesyn switch located behind the Control Station

7. Ping IPs of SPA and SPB to ensure connectivity (SPA IP 128.221.252.200; SPB IP 128.221.252.201)

8. Launch NST tool, connect to SPA, log into system, then proceed to Eng. Mode using ctrl + shift + f12 and password: SIR

9. Click on Engineering>Upgrade Disk Firmware: Disk Firmware Upgrade wizard Welcome screen

10. Follow prompts and browse to .fdf firmware upgrade package and proceed with upgrade

11. If Vault drives are involved in the upgrade, the DFU wizard will have to be run a 2<sup>nd</sup> time, as the code will only allow 2 vault drives to be upgraded during each pass. With vault drives involved, the DFU wizard will show a Red Error for the “Verifying second

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
pass disks upgraded” message, at which point the NST should be completely closed. Relaunch the NST, login, then run the DFU wizard to complete the remaining drive firmware upgrades.

**Note:** The firmware itself is flashed in seconds, with the majority of the time taken to allow both SPs to reboot a couple of times during the DFU process

12. Once NST says the firmware upgrade is completed, confirm array status by opening Navisphere, check for any faults, rightclick arrayname>Disk Summary to visually check firmware version, then open the SP Event Log and look for 778 event, which should have a message: “LCC Firmware complete on all.”

13. Reboot Control Station

14. Reinsert Data Movers

15. Wait about 5 minutes, then log into Control Station using SSH, and healthcheck the Celerra system

/nas/sbin/getreason

/nas/bin/nas\_checkup

server\_mount ALL

navicli -h getcache | getcrus | getlun -trespass

## **CLARIION STORAGE ARRAY PRODUCTS:**

→Clariion AX100i, CX300i, and CX500i Native iSCSI Storage Arrays, Feb 2005 [iSCSI support--NAS, NetWin, Symm, Connectrix] CX300, CX500, & CX700 GA 9 February, 2004, running embedded Windows XP O/S on SPs

→Each data block on array contains 8 bytes error checking data (LRC, Shedstamp, Writestamp, Timestamp), and SNiiFER runs in backup continuously checking block data integrity

→All fibre disks are dual-ported

→LCC A is located on right of system and B on the left, when viewed from rear

→Backend loop for CX arrays consists of a single FC\_AL loop from Primary port on DAE2 and through all 15 disks of the enclosure, then out to the next DAE via the Expansion port to the next LCC card, and all disks, eventually returning via the alternate LCC path and drives to the SP

→Backend loop consists of SP Port, Cables, Cable Ports on LCC/BCC, and Disk Drives

## **SP AMBER LIGHTS CX/CX3/CX4:**

¼ Hz indicates system doing BIOS

1 Hz indicates system doing POST

4 Hz indicates O/S is booting, but FLARE/NDUMON are not yet started

Solid Amber indicates a fault condition

Lights off indicates normal operation

## **CLARiiON CX3 ULTRASCALE SERIES:**

### **CX3-10 ULTRASCALE TACKHAMMER PLATFORM:**

GA Feb 14, 2007; Flare 24 patch 008 minimum; 60TB capacity

→CX3-10C offers both iSCSI and FC connectivity for Hosts, 64 Hosts maximum, 60 drives max., 128 Host initiators, 512 LUNs

→No hardware upgrades to CX3-10c until a future timeframe

→Clariion is marketing only the CX3-10c model, which will support either all FC drives, FC vault drives and ATA, or all ATA vault drives & ATA drives; Single 1.8GHz processor per SP; 1GB memory per SP; (2) 1Gb/s FC/iSCSI ports per SP, (2)4Gb/s FC front-end ports per SP, & (1) 4Gb/s FC BE port for DAE per SP; Single SPS with 2<sup>nd</sup> SPS optional; (2) LAN management ports; (2) Serial ports

→Base system comes with SPE3 enclosure & single 1000watt DC SPS, 4-power supplies/cooling fans, two SPs, single DAE with minimum 6 drives

#### **CX3-10C Front-End Connectivity:**

(2) 4Gb FC optical ports/SP

(2) 1Gb copper Ethernet ports for iSCSI/SP (iSCSI targets via GbE LAN, direct-connect to NIC, or connect to Host Initiator HBA)

#### **CX3-10C Back-End Connectivity:**

(1) 4Gb FC arbitrated loop port/SP

#### **CX3-10DC-FD ULTRASCALE SERIES:**

Model GA July 2007, NEBS certified with 2U chassis for DC power, FC ports only, Aries Flare code, no SPS units, no PowerPath 2GB memory per array; 64 Hosts per array; (4) FC-only ports per array & (2) BE ports; Max. 60 drives, 30TB, 512 LUNs; FC Drives

#### **CELLERRA CX3-10 VARIANTS:**

**Note:** CX3-10 is stripped down version of CX3-20—no Combo card and SP's have only 1GB memory

#### **CX3-10 for NS2-AUX (Celerra NS20):**

(2) FE FC Ports

(1) BE FC Port

#### **CX3-10F for NS2-AUXF (Celerra NS20 where SPs have additional Headhunter Quad IO Module):**

(6) FE FC Ports  
(1) BE FC Port

### **CX3-20/CX3-40 PLATFORMS:**

- Replacement to CX300/500 family
- Both CX3-20 & CX3-40 family supports 2 or 4GB UltraPoint disk array speeds (DAE3P) & 4GB FC port speeds
- First offering for Tiered Services (NS20 only) and for CRU (Customer Replaceable Units)
- Flare 22 patch 034, support increased to allow for upgrades from CX2/3/4/5/6/7 models to the CX3 platform, and for new 500GB SATA II disk drives
- Some configurations will allow both iSCSI and FC clients, though access to both protocols from the same Host is not allowed

### **CX3-20 ULTRASCALE JACKHAMMER PLATFORM:**

Flare 24 minimum, 55-59TB capacity, Single 2.8GHz processor per SP, 4GB total memory per array, max. Write Cache 1053MB; 120 drives max; 128 max Hosts; 1024 LUNs; CX3-20 arrays will be Customer-installable, and will use Navisphere Service Taskbar (NST) tool for limited hardware maintenance and software upgrades

#### **CX3-20 Front-End FC Connectivity:**

(2) FE FC ports to FC Switches or Servers/SP

#### **CX3-20 Back-End FC Connectivity:**

(1) BE FC port to DAE's/SP arbitrated loop

#### **CX3-20c ULTRASCALE SERIES:**

Same hardware as CX3-20 except for addition of iSCSI IO module; Flare 22 patch 506; max 120 drives, 128 Hosts, 256 initiators per array, 1024 LUNs

#### **CX3-20C Front-End FC Connectivity:**

(1) FE FC ports to FC Switches or Servers/SP

(4) FE 1GbE iSCSI ports (targets) to GbE LAN, Host NIC, or Host HBA Initiator/SP

#### **CX3-20C Back-End FC Connectivity:**

(1) BE FC port to DAE's/SP arbitrated loop

#### **CX3-20F ULTRASCALE SERIES:**

Max 120 drives, 128 Hosts, 256 initiators; 1024 LUNs per array; No iSCSI; Requires Flare 24, patch 007 minimum

**Note:** Ports labeled 0 Fibre – 5 Fibre, & BE0

#### **CX3-20F Front-End FC Connectivity:**

(6) FE FC ports to FC Switches or Servers/SP

#### **CX3-20F Back-End FC Connectivity:**

(1) BE FC port to DAE's/SP arbitrated loop

#### **CX3-20FDC or FD ULTRASCALE SERIES:**

Same hardware as CX3-20F except the FDC supports DC power for NEBs certification: Flare 24 patch 008 minimum--Flare 03.24.040.5.008; 4GB memory, 64 Hosts, (12) FC ports, (2) BE ports, 120 drives, 59TB, 1024 Luns per array; (6) optical Front-end ports & (1) Back-end copper port per SP

### **CX3-40 ULTRASCALE SLEDGEHAMMER PLATFORM:**

Flare 24 minimum, 114-119TB capacity

(2) 2.8GHz processors per SP, 8GB memory per array with max 2500MB Write cache; 240 drives max; 128 Hosts and 2048 LUNs;

#### **Front Chassis View, I-to-r, SPA, SPB:**

Power Supply Fan & Supply module, I-to-r, PS A0, PS A1, PS B0, PS B1 [green LEDS, steady or flashing amber with faults]

#### **Back Chassis View, I-to-r, SPB, SPA:**

IO modules, Fibre optical, Fibre copper, Ethernet, and Serial connections

Power LED green for normal; Fault LED 1/4Hz for BIOS checking; 1Hz for POST; 4Hz for Booting OS; 2Hz for NMI Button pressed

#### **CX3-40 Front-End FC Connectivity:**

(2) FE FC ports to FC Switches or Servers/SP

#### **CX3-40 Back-End FC Connectivity:**

(2) BE FC ports to DAE's/SP arbitrated loop

#### **CX3-40C ULTRASCALE SERIES:**

Flare 22 patch 506; Same hardware as CX3-40 except for new Combo iSCSI IO module; max 240 drives, 128 Hosts, 256 initiators per array; 2048 LUNs

#### **CX3-40C Front-End FC Connectivity:**

(2) FE FC ports to FC Switches or Servers/SP

(4) FE 1GbE iSCSI ports (targets) to GbE LAN, Host NIC, or Host HBA Initiator/SP

#### **CX3-40C Back-End FC Connectivity:**

(2) BE FC ports to DAE's/SP arbitrated loop

#### **CX3-40F ULTRASCALE SERIES:**

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
Max 240 drives, 128 Hosts, 256 initiators per array; 2048 LUNs per Storage system; No iSCSI ports this model; Flare 24, patch 007 minimum

**Note:** Ports will be labeled 0-3 Fibre for Front-end, and BE0-BE3 for Back-end

#### **CX3-40F Front-End FC Connectivity:**

(4) FE FC ports to FC Switches or Servers/SP

#### **CX3-40F Back-End FC Connectivity:**

(4) BE FC ports to DAE's/SP arbitrated loop

**Note:** Celerra will use the CX3-40F for the NS40F Model

#### **CX3-40FDC MODEL:**

Same hardware & IO ports as CX3-40F except the FDC supports DC power for NEBs certification—Flare 24 patch 008 minimum

### **CX3-80 ULTRASCALE HAMMERHEAD PLATFORM:**

Flare 24 minimum, 234-239TB capacity, Highend SAN/NAS storage, replacement for CX700, 4U Enclosure, Dual 3.6GHz Processors, 16GB memory per array; max. 480 drives, 256 Hosts, 2048 LUNs, 512 max. initiators per array; No iSCSI ports for this model; SAN Management Module allows communication between SPs; LCCs provide point-to-point comms between SPs and DAE3Ps; FC Copper SFP to HSSDC2 cable from SPs to DAE3Ps; HSSDC2 to HSSDC2 FC Copper cables between DAE3P enclosures; LC to LC FC optical cables between SAN Hosts and SPs; (8) port GbE module for NAS support (6 cu and 2 optical);

#### **Front-End Connectivity:**

(4) port fibre optical 4GB FC Module for FE SAN/NAS per SP

#### **Back-End Connectivity:**

(4) port arbitrated loop 4Gb/s FC BE ports per SP

### **ULTRASCALE DISK DRIVES:**

73GB or 146GB 2Gb/s FC 10k RPM drives; 73GB or 146GB 4Gb/s FC 15k RPM drives

300GB 2Gb/s FC 10k RPM drives

500GB 2Gb/s FC 7200rpm drives

500GB 4Gb/s SATA II 7200rpm drives

### **RAID TYPES FOR ULTRASCALE ARRAYS:**

RAID 0 3-16 drives; no data protection

RAID 1 2 drives; mirrored pairs

RAID 1/0 2-16 drives; mirrored and striped [Not supported until Wormwood platform]

RAID 3 4+1 or 8+1 parity

RAID 5 3-16 drives; parity striped across drives

#### **BE PORT TYPES:**

HSSDC2 copper arbitrated loop FC connectors

### **CLARiiON CX4 FLEET SERIES:** GA August 4, 2008, Flare 28 04.28.000.5.008 bundle used on all CX4 platforms

**Models:** CX4-120C, etc. “C” denotes that arrays are sold with combination Fibre Channel & iSCSI IO ports on SPs

Nautilus (CX4-120), Ironclad (CX4-240), Trident (CX4-480), and Dreadnought (CX4-960—4U chassis) to replace the CX3 UltraScale Series CX3-20, CX3-40, & CX3-80, respectively.

64-bit Flare 28 “Proteus” Windows 2003 kernel for SPs (CX3 ran XP)

Mix & match support for FC, iSCSI, SAS, GbE, 10Gb IO modules

Based on Black Widow system design for next generation of Clarion and NAS products

Single SP Write Caching (GWCA—Greater Write Cache Availability)

IO Throttling on vault drives

5-6 different IO modules available, billed as “self-installable”, though requires array reboot—SLIC IO modules will be customer installable, but EMC maintained

Increase in drive count and performance

Support for SAS/SATA IO modules and enclosures

ICA Image to be on a Flash card (image of boot partition for an SP)

IPV6 support

In Place Upgrades from CX or CX3

Replacement SPs will not ship with DIMMs or CPU board—replace components from existing SP

Write cache is not disabled during NDU operations, single SP failure, single SP removed or rebooted, etc.

#### **Jupiter Program—Flare 30.50x GA Q4 2010:**

→Rollout of the 2-port 10Gb FCoE SLIC for CX4 class arrays, as a direct connection from the array to an FCoE switch

→Allowing use of Active Twinax cabling for 10Gb iSCSI modules

→Modification of the CLARiiON CLI Test Event for Heartbeat

**navicli -h <sp\_ip> eventmonitor -heartbeat -insert**

- First release will not support layered application functionality
- Automatic checking for SnapShot Rollback status prior to doing a Clone Sync or Reverse Sync operation
- Support limited to Windows 2003/2008, VMware, and Linux hosts (Post GA support for SUN, AIX, HP-UX)

### **BLACK WIDOW SPO CPU:**

- Black Widow 2U SPE CPU module, with either a single 2.2GHz or Dual-core 1.6GHz processor system, up to 8GB memory, 64-bit architecture, for both Trident and Ironclad platforms, respectively
- (5) IO modules max. for both Blades & SPs, with max. of 20 ports per Blade & SP
- Akula SAN Mgmt module [with integrated Broadcom BCM5397 switch used to connect the Service & Mgmt LAN ports to either CPU, with a crosslink connection between ports for peer Mgmt]
- Nimitz NAS Mgmt module, with (2) LAN mgmt ports & (1) Uplink port (1 GbE), (1) micro DB-9 serial port, (1) USB port for external connections, and an NMI button, to be used with Trident & Ironclad Clariions
- Tomahawk 4-port Fibre Channel I/O module [1-4Gbps capable]. Any of the I/O ports could be configured for direct FE Host connectivity, Switched fabric connectivity, or BE DAE connectivity, and can also be Optical or Copper SFPs.
- Harpoon 2-port iSCSI I/O module
- Enclosure ID will be stored in midplane resume PROM

**Note:** Resume PROM is stored in non-volatile memory on the backplane. Software is used over serial buses to determine unique enclosure addresses for system components and then store them.

### **DREADNOUGHT (CX4-960):**

- CLARiiON Fleet CX4-960 Flare 28 Proteus 64-bit O/S, replacement for CX3-80, max 960 drives & 32GB system memory
- 4U SPE enclosure, Wildcat-S Quad-core Clovertown 2.33GHz processor; 16GB memory per SP, 512 initiators per array (256 Hosts per SP), 2-port iSCSI Harpoon IO modules, 4-port FC Tomahawk IO modules,
- Uses the Wildcat-S 3.0GHz dual socket CPU running at 1.33GHz FSB
- Supports Clariion, Symmetrix (first HW platform to support Symm—‘Eagle’ low-end & ‘Tigon’ high-end), and Clariion families
- Typhoon or Earthquake NAS Management modules/switch for the Dreadnought & Tigon Arrays (Q3-2008)
- Earthquake NAS module will have ability to redirect serial console to Mgmt LAN network, and implements heartbeat discovery algorithm over mgmt network to assign IDs to each enclosure
- Will support (8) Data Mover blades in DMEs
- Supports (6) IO modules per SP, with max of (16) FC ports, (8) iSCSI ports, & 15 Flex Ports per SP
- 960 FC and/or SATA II drives, up to 64 DAEs, 120 drives per loop, and 8 loops per array
- Supports 512 Initiators per SP, 1024 per Array, with 512 HA hosts
- Supports 4096 LUNs, 256 LUNs per RG, 256 LUNs per SG, 1024 SGs, RAID 0, 1, 3, 5, 1/0, 6

### **DREADNOUGHT EFD DRIVES (Enterprise Flash Drives, flash-based solid state drives):**

- GA date Oct 2008, for Dreadnought platforms only
  - 73GB STEC Zeus drives @52,000 read IOPS/18,000 write IOPS, compared to conventional drives of 250 read/125 write IOPS
  - Flare 28 base, Proteus and Mira
  - Minimum bus speed 2GB
  - Not supported for Vault drives
- Note:** Initially not allowed on Bus 0 [only bus 1 enclosure 0, called the “Enzo” enclosure]
- FC FSSD allowed in multiple enclosures, with no mixing of other drives within same enclosure
  - RAID 5 4+1 or 8+1 configurations
  - Drive type reported as Fibre Channel Solid State Drive
  - Drive format is 512 byte, meaning that Clariion has to use an SFD [Sata Filter Drive] to convert to 520 byte
  - Mira Flare program plans to support a 2GB FC 512byte STEC drive type, and a 4GB FC 520byte STEC drive type

### **TRIDENT (CX4-480) PSN 900-566-002:**

- CLARiiON Fleet CX4-480 Flare 28 Proteus, replacement for CX3-20, max 480 drives
- 2U SPE enclosure, Blackwidow Dual-core Single CPU Xeon 2.2Ghz processor, up to 8GB memory per SP, 256 Hosts per array (128 per SP), 2-port iSCSI Harpoon IO modules, 4-port FC Tomahawk IO modules
- Default Read Cache 512MB
- 1200 Watt 1U SPS per SP, for supporting SPA or SPB & first DAE LCCA or LCCB (Enclosure 0, Bus 0)
- (2) 400 Watt 2U Power Supplies per SP
- Supports (5) IO modules per SP, with max of (12) FC ports, (8) iSCSI ports, and 11 Flex ports per SP
- 5-480 FC and/or SATA II drives, (32) DAEs, 120 drives per loop, 4 loops per array
- Supports 256 Initiators per SP, 512 per array, with 256 HA Hosts
- Supports 4096 LUNs, 256/RG, 512 SGs, RAID 0, 1, 3, 5, 1/0, 6

### **IRONCLAD (CX4-240) PSN 900-566-001:**

CLARiiON Fleet CX4-240 Flare 28 Proteus, replacement for CX3-20, max 240 drives

2U SPE enclosure, Blackwidow dual-core Woodcrest Intel Xeon 1.60GHz processor, 4GB memory per SP, 256 Hosts per array (128 per SP), 2-port iSCSI Harpoon IO modules, 4-port FC Tomahawk IO modules

→Supports (5) IO modules per SP, with max of (8) FC ports & (8) iSCSI ports per SP, & 7 Flex Ports per SP

→5-240 FC and/or SATA II drives using (16) total DAEs (FC, SATA, SAS later on)

→Supports 256 Initiators per SP, 512 per array, and 256 HA Hosts

→Supports 1024 LUNs, 256 LUNs per RG, 256 SGs, RAID 0, 1, 3, 5, 1/0, 6

### **NAUTILUS (CX4-120) PSN 900-566-004:**

→CX4-120 array Q3 2008, replacement for CX3-10 array, 5-120 FC or SATA II drives

→Diskless SPE; max 120 drives; min 5 disks; Dual-core Single 1.2GHz CPU for SPs; 3GB memory per SP; FC, SATA, or SAS; new 1.2KW SPS

**Note:** If you look at the SP system itself, the CPU reports itself as 1.60GHz, which is what the native CPU horsepower is. For Nautilus, FLARE throttles down the CPU to 1.2GHz

→(2) 400 Watt 2U Power Supplies per SP

→Support for up to (5) I/O modules per SP, 2 BE FC ports per array, total of 8 FC ports per SP, 4 iSCSI ports per SP (No FLEX ports)

→Supports 128 Initiators per SP, 256 per array, 128 HA Hosts per array

→Supports 1024 LUNs, 256 LUNs per RG, >2TB LUN size, 16 drives per RG, 128 Storage Groups, RAID 0, 1, 3, 5, 1/0, 6

→Support for RAID 0, 1, 3, 5, 6, 1/0

→Data-in-place conversions from CX or CX3 arrays

→Self installation, upgradeable, replaceable components, Fibre Channel and iSCSI support

→SPs are Hot-Swappable

### **FLEET LED/CX4 LED INFORMATION:**

#### **SPS LEDs:**

##### **Active LED**

Top LED solid green means SPS is On-Line, Fully Charged, and Ready—i.e., operating normally; Blinking green means SPS battery is recharging, On-Line, Charging, Not Ready. Green LED is out if any Amber LEDs are lit.

##### **On Battery LED**

2<sup>nd</sup> LED from top; Amber LED lit = On-Battery mode, AC Power interrupted & SPS supplying battery power to SP/DAE; also illuminated during battery testing

##### **Replace Battery X LED**

3<sup>rd</sup> LED from top; Amber LED lit battery lifetime exceeded or self-test failure occurs, SPS not charged & not able to perform cache destaging function, resulting in write cache disabled

##### **Internal Check/Triangle Fault LED**

SPS has an internal fault, Amber LED lit, write cache disabled

### **SP TRIANGLE LED, AMBER LIGHT:**

¼ Hz indicates system executing BIOS

1 Hz indicates system executing POST

4 Hz indicates O/S is booting, but FLARE/NDUMON are not yet started

→Solid Amber LED indicates a fault condition

→Blinking Amber 1-3-3-1 sequence on bootup indicates Memory problem

→LED off indicates normal operation (or power is off to SP)

→NMI button will reflect 2Hz flashing Amber LED while dump is in progress, remaining until SP reboots

### **SP TRIANGLE LED, BLUE LIGHT:**

¼ Hz indicates Windows OS has booted

1 Hz indicates Flare driver starting

4 Hz indicates Flare drive started

→Solid blue LED indicates SP running in Degraded Mode

→Blue LED off indicates FLARE ready for IO and NDUMon started

### **SP HAND ICON LED, WHITE LIGHT:**

→This is a special LED represented by a “hand” and if illuminated with White light, means that the SP cannot be safely removed

White LED off indicates SP can be serviced normally

White LED on indicates SP cannot be safely removed (Peer could be panicking or rebooting without its Cache configuration; SP could be doing BIOS/POST or Resume prom update; SP could be dumping cache to vault)

### **SPE ENCLOSURE LEDs CX4-120 to 480:**

→Located front middle section of SPE between the 2<sup>nd</sup> & 3<sup>rd</sup> Power Supplies

→Top round LED solid Amber light indicates enclosure is faulted (normal operation is LED off)

→Bottom round LED solid Blue light means SPE is powered up

### **SP MANAGEMENT MODULE:**

Upper left LED off, no power

Upper left LED Amber, faulted

Upper left LED Green, operating normally

RJ45 ports, LED Green off, no link; Amber LED off, no activity; Green LED on, Link; Amber LED flickering, activity

### **SP POWER SUPPLY MODULES—4 modules located on Front of array:**

Solid green LED indicates powered on and normal operation

Solid amber LED indicates faulted module

Blinking amber LED indicates fault external to module (missing AC cord, SP removed, etc)

### **TOMAHAWK FC IO MODULE (SLIC):**

Upper LED off, no power; Upper LED Amber on, faulted; Upper LED Green on, normal powered on

#### **Fibre Port LEDs Green or Blue:**

--LEDs off, no link or SFP;

--Blinking 1Hz green, 1G/2G link speed port marked

--Solid green normal, 1G/2G link speed

--Solid blue normal, 4G link speed

--Blinking 1Hz blue, 4G link speed port marked

--alternating 1Hz blinking blue and green, SFP faulted or not supported type

### **HARPOON iSCSI IO MODULE (SLIC):**

Upper LED off, no power; Upper LED Amber on, faulted; Upper LED Green on, normal powered on

#### **iSCSI Port LEDs Yellow or Green:**

--Power LED, Green indicates good link and should be on

--Link/Act off could be connected or unconnected

--Link/Act blinking Amber 1Hz indicates transmit or receive activity

--Both Link/Act and Power LEDs blinking indicates port marking

### **DISK DRIVE LEDs:**

LED off means not powered on

Solid green LED indicates normal operation

Blinking green LED indicates drive access activity

Solid amber LED indicates drive faulted

Blinking amber LED indicates diplex flash command from navi

### **FLEET LEDs UNDER HW CONTROL:**

→Power supply fault LED

→Air mover module fault LED

→I/O module power fault power good LED

→Tornado power good LED

→System enclosure power on LED

→Diplex LCC fault

### **FLEET LEDs UNDER SW CONTROL:**

→Enclosure fault

→Management module Fault

→IO port RX Loss fibre channel

→IO port connector link and activity LEDs

→IO module power fault LED amber

→CPU module removed

→SP fault [Blue/Amber LED; 1/4Hz amber=BIOS; 1Hz amber=POST; 4Hz amber=Starting OS; 1/4Hz blue=OS booted; 1Hz blue=Flare drive init; 4Hz blue=Flare drivers completed; Off=Flare ready for IO; Steady amber=SP faulted]

→Tornado fault

→Disk drives

### **FLEET DISK DRIVE SUPPORT:** FC, SATA II, SSD

146GB 15k RPM 2Gb or 4Gb FC interfaces

400GB 10k RPM 2GB or 4Gb FC interfaces

300GB 15k RPM “ “

1TB 7200RPM SATA

1TB 5400RPM SATA

73GB SSD (STEC FC Solid State Drives) 2Gb FC Controller, 512-520BPS Emulation, 3.5" form factor

## **MISCELLANEOUS FLEET CLARiiON PROCEDURES/FRU INFORMATION:**

### **Changing SP IP Addresses for Customers:**

- 1) CLARiiON recommends “Destroy Security and Domain Information” using /setup program, then Restart Management Server, and do on both SPs
- 2) Change IPs on SPs and Restart Management Server for each SP
- 3) Connect to SP using Navisphere and select Yes to initialize security
- 4) Once connected to Navisphere, go to File>Set Up Domain>Select Master and highlight SPA

## **DIRECT CONNECTING TO CX4 SERVICE LAN & USING NST or NAVISPHERE:**

### **I. NST**

- 1) Configure laptop with an IP address on CLARiiON default system network:

**128.221.1.249**

**255.255.255.248**

- 2) Open NST, select Options>Network Address Translation (NAT) Connection, and enter default system IP addresses for each SP

**128.221.1.250 SPA**

**128.221.1.251 SPB**

- 3) Log in and use the NST

**Note:** The NAT limitation does not exist for normal public connection to the SP over the Management LAN port

### **II. NAVISPHERE**

- 1) Open web browser and enter public or hard-coded IP address for SPs (e.g., 128.221.1.250)

### **CX4 SERVICE LAN PORTS:**

→Unlike the CX3 arrays, where you physically connected to the Service Port on SPB in order to connect to SPA, the CX4 array Service LAN ports are cross-connected on the midplane between the SPs, meaning that you can physically connect to either SPA or SPB and then connect via Navisphere to either SP. For NST use, use NAT when connecting to the array.

**Note:** See emc199379 for more details.

→Service LAN port is the lefthand port indicated by the “wrench” symbol

→Management LAN port is the righthand port and is typically used by Celerra for the Internal management network configuration

→VLANs are used to separate the Service LAN ports from the Management LAN ports

→Crossover or straight Ethernet cables will work because the built-in BCM Broadcom switch is auto-MDIX enabled, and uses auto-sensing for Tx/Rx transmissions

**Note:** Important point here is that Support folks should NOT use a Hub or Switch between Laptop and SPA and SPB ports, as typically done on CX3 arrays—this may cause packet storm problems that are only resolved by reseating the Management Switch

### **SP MANAGEMENT SWITCHES:**

→Switches are hot swappable, meaning they can be safely removed and re-inserted on a live SP

### **Connecting to Fleet SPs Using Serial Port:**

- 1) Connect mini-db 9 serial cable to SP’s serial maintenance port (wrench symbol)
- 2) Open HyperTerminal and configure as 9600/8/N/1/Hardware Flow Control
- 3) Reboot SP from CLI or Navisphere GUI
- 4) Press esc key during “ABabCabcdefg.....” Output
- 5) Enter password: DB\_key to access Diagnostic Menu

### **Diagnostic Menu**

- 1) Reset Controller
- 2) Display Warning/Errors
- 3) DMI Log Sub-Menu
- 4) DDBS Service Sub-Menu
- 5) FCC Boot Sub-Menu

### **To access the CLARiiON Utility Toolkit Menu:**

→Select 4) DDBS Service Sub-Menu

- 1) Drive Slot ID Check
- 2) Utility Partition Boot
- 3) Force Degraded Mode

→Select 2) Utility Partition Boot

### **CLARiiON Utility Toolkit Main Menu**

- 1) About the Utility Toolkit
- 2) About this Array
- 3) Reset Storage Processor
- 4) Image Repository Sub-Menu
- 5) Plugin Sub-Menu

- 6) NVRAM Sub-Menu
- 7) Enable LAN Service Port
- 8) Enable Engineering Mode (Special password required)
- 9) Install Images to Recover OS
- 10) Save Logs

→Select 9) Install Images to Recover OS if you need to perform a Flare reinstall of the CLARiiON System (contingent on the vault drives having intact utility partitions and the base flare to reload)

#### **Engineering Mode for Toolkit Menu**

- 10) Install Images to Restore Factory Configuration
- 11) Invalidate Data Directory

### **REINSTALLING FLARE OS ON CLARiiON (ICA):**

- a. Ensure both SPs are plugged into the midplane
- b. Connect to Serial Maintenance port on SPA and open HyperTerminal session
- c. Reboot SPA and access ‘Diagnostic Menu’ during POST “ABabCabcdefg...” using esc key, then password DB\_key
- d. Select DDBS Service Sub-Menu>Utility Partition Boot
- e. Enter Engineering mode: 8) Engineering Mode  
password: wombat
- f. If restoring OS from existing partitions, invalidate data directory first: 11) Invalidate Data Directory
- g. Then select 9) Install Images to Recover OS

#### **Select Images to Install →1)**

- 1) ICA\_FLARE\_PARTITION\_REGION 04.28.000.5.003 (Selected this Image for FLARE)
- 2) ICA.Utility\_PARTITION\_REGION 04.28.000.5.001

Are you sure you want to install these images? y

#### **Select Storage Processors to install images for... →3)**

- 1) This SP (SPA)
- 2) Peer SP (SPB)
- 3) Both SP’s

--Select 3) “Both SP’s”

WARNING!! Continue with image installation? y

h. At this point, the ICA takes place in about 30 minutes

i. Reset Storage Processor at the end

**Note:** SPA will reboot several times to complete its FLARE installation—this will take awhile

j. Reboot or reset SPB

**Note:** When reapplying FLARE to system, found I had to also run the above procedure while connected to SPB, with SPA removed from its slot, but did not do the Invalidate Data Directory step—then everything worked.

#### **Other Menus of interest**

#### **CLARiiON Utility Toolkit Image Repository Menu**

- 1) Back to the Main Menu
- 2) List Image Repository Contents
- 3) Delete Files from the Image Repository
- 4) Copy Files from the RAM Disk to the Image Repository
- 5) Copy Files from the Image Repository to the RAM Disk

#### **Image Repository Contents**

- ICA Flare Partitions 0-8
- ICA Utility Partitions 0-5
- Persistent Events.db
- Parameters.cfg
- NonParameters.cfg
- ktrace-20081024\_000638-A.zip
- NonParameters.cfg.old

### **CLARiiON CX5 FIGHTERS SERIES:**

→Next generation CLARiiON arrays

→Using 6Gb SAS backends only (DPE & SPE-based)

**Note:** Future release will support SAS front-end ports

→Flare 31 ‘Scorpion’

→Samsung Flash drive support

→Maximum of (10) DAE’s allowed for a single back-end bus

→Mezzanine card on the SP’s is considered “onboard” or built-in I/O module, therefore will not be a FRU. Also contains SAS & FC controllers.

### **VCX5-75 (CX5-75 Hellcat Lite array):**

- New low-end platform
- Codenamed “Hellcat-Lite”
- Flare 31
- Uses a 3U Sentry Lite DPE consisting of 15-drives, (2) SP’s, single SPS, Power Supplies, Nehalem dual core 1.86GHz CPU Modules, I/O Modules, 4GB memory DIMMs, and (3) internal Fan packs that are not FRUs (loss of 2 fans will shutdown SP)
- Note:** DPE type with “getagent –cabinet” will display DPE7
- SAS, SATA & EFD drives, max. of 75 drives allowed
- Single 6Gb SAS backend bus using only Port 0, of the two SAS ports

### **VCX5-125 (CX5-125 Hellcat array):**

- Replacement for CX4-120
- Codenamed “Hellcat”
- Flare 31
- SAS, SATA & EFD drives, max. of 125 drives
- Uses 3U Sentry DPE consisting of 15 drives, (2) SP’s, single SPS, Power Supplies, dual core Nehalem 2.0GHz CPU Modules, I/O Modules, 8 Gb memory DIMMs, and (3) internal Fan packs that are not FRUs (loss of 2 fans will shutdown SP)
- Note:** DPE type with “getagent –cabinet” will display DPE7
- Two 6Gb SAS backend buses using SAS ports

### **VCX5-250 (CX5-250 Lightning array):**

- Replacement for CX4-240
- Codenamed “Lightning”
- Flare 31
- SAS, SATA & EFD drives, max. of 250 drives
- Uses a 3U Sentry DPE with 15 drives, and (2) SP’s, (2) SPS’s, Power Supplies, quad core 2.13GHz Nehalem CPU Modules, I/O Modules, 12GB memory DIMMs, and (3) internal Fan packs that are not FRUs (loss of 2 fans will shutdown SP)
- Will have more memory, faster CPU, and will use both SAS Ports for (2) back-end buses
- Dual 6Gb SAS backend buses

### **VCX5-500 (CX5-500):**

- Codenamed “Spitfire”, Scorpion Flare, SAS/SATA drives, XPE
- 2U SPE, 2.4GHz Westmere CPU’s for SPs, with 18Gb memory
- Supports SAS, SATA & EFD drives, 500 drives max. using 6Gb SAS ports for (4) backend buses

### **VCX5-1000 (CX5-1000):**

- Codenamed “Mustang”, Scorpion Flare, SAS/SATA drives, XPE
- Uses 2U SPE, 2.80GHz Westmere CPU for SPs, with 24Gb memory
- Supports SAS, SATA & EFD drives, 1000 drives max. using 6Gb SAS ports for (4/8) backend buses

#### **Available DAE’s:**

- VIPER—3U 15@3.5” SAS/SATA drives
- DERRINGER—2U 25@2.5” SAS drives
- VOYAGER—4U 60@3.5” SAS/SATA drives

#### **Available I/O Modules:**

Supercell—1Gbps 4-port FE iSCSI module, for Host or switch connectivity via copper GbE ports. All CX5 arrays except for Hellcat Lite. FRU/CRU.

Hypernova-V—1.5, 3, or 6Gbps SAS 2.0 4-port BE to DAE connectivity to SAS or SATA drives (Viper, Derringer, Voyager)

Cold Front—1.5, 3, or 6Gbps SAS 4-port FE module

Eruption-V—10GbaseT 2-port copper iSCSI module for host-attached (CAT 6/6a/7 cabling), and as 10GbE module for Data Movers  
Cat 6: 55 meters max distance for 10Gbps

Cat 6a/7: 100 meters max distance 10Gbps

Cat 5e: 1Gbps only

Poseidon II—10Gbps 2-port iSCSI module with upgraded chipset

#### **Growler Mezzanine Card:**

- This is an onboard I/O module that contains (2) Backend SAS ports and (4) Frontend FC ports [NOT a FRU]

#### **Heatwave**

- 10GbE 2-port FCoE FE module, to be used with CX4 & CX5 arrays
- Direct-attached Hosts are not supported--FCoE Hosts will connect via an FCoE switch port to which the array will connect to
- Connection interfaces will use 10G optical SFP+ connectors or TwinAx active connectors (1, 3, or 5M lengths)
- CX4-120, 480, & 960 support will be added with Jupiter FLARE Q3 2010
- CX5 support will be added in the Culham release timeframe
- VG2/VG8 will support FCoE Initiator mode connections to FCoE switches with Heatwave IO Module in slot\_0

#### **FCoE:**

Fibre Channel over Ethernet uses lossless full duplex Ethernet, with a dual-port FE IO Module on the SPs for sending FC frames natively over 10GbE Ethernet. FCoE represents a merging of SAN FC over IP Networks. FC0 & FC1 layers in the FC stack are replaced with Ethernet, which runs above TCP & IP layers, therefore FCoE is not routable (limited to a single subnet). Normal Ethernet does not have a flow control mechanism. Enhancements to the Ethernet standard were made to introduce flow control, making it “lossless” Ethernet, or Converged Enhanced Ethernet. Sold with two connector types, an optical SFP+ with optical cables and Active Twinax and 1/3/5M cables (supporting fibre channel frames over 10GbE). Only Jumbo frames of 2240 bytes are supported. Requires use of an FCoE switch and FCoE SLIC on the SPs. Direct connections between Host and Array using FCoE is not supported.

**FCoE Standard:** FC-BB-5

Suitable Ethernet extensions for FC-BB\_E usage are the PAUSE mechanism defined in IEEE 802.3-2008 and the Priority-based Flow Control mechanism defined in IEEE 802.1Qbb (these are the two extensions used to obtain lossless full duplex Ethernet networks).

**Supported Hosts:**

--Linux, Windows 2k3/2k8, VMware

**Ethernet Enhancements for FCoE:**

--Encapsulation of FC frames into Ethernet frames

--FIP, FCoE Initialization Protocol

--Extension to Ethernet standard for lossless Ethernet with Priority Flow Control (PFC)

**NAVICLI Shutdown Behavior changes:**

Both the **shutdownsp** & **shutdownpeersp** commands will shutdown and hold the SP in reset, and also powers down half the drives in the enclosure. To recover, you would need to use rebootpeerSP command from active SP.

The **iportconfig –replace & –upgrade** commands will also shutdown the SP and hold it in reset.

**Backend Bus:**

# **/nas/sbin/navicli -h 10.241.168.188 backendbus -get -all**

Bus 0

Current Speed: 2Gbps.

Available Speeds:

3Gbps.

6Gbps.

**Backend Port connectors:**

**backendbus -get -connstate**

→Will show connector state when mini-SAS HD connectors are inserted into BE SAS ports

**backendbus -get -prominfo**

→Information about Front-End SFP+ ports and mini-SAS HD connectors

**IO Port Info:**

iportconfig –list –iomodule –sp a

**CLARIION AX SERIES—2008:**

**AX4-5/AX4-5F8—MAMBA:**

→Uses Boomslang CPU Module [Sossaman 1.66GHz processor], 32-bit architecture, 2U Disk Processor Enclosure, 4Gb FC backend DAEs, 128 initiator limit per SP (64 HA hosts), max 512 LUNs, 256 LUNs per initiator, 3U DAE with 15 drives; 2U DAE with 12 drives, Release 23 Flare for this platform

→Supports 4-60 drives

→AX4-5F8(100-562-104) array variant for NX4 uses 02.23.050.5.703 Flare

**TYPES OF CLARIION BACKEND LOOPS:**

→Katana Backend loop requires all traffic to pass through each disk’s receiver and transmitter—if one disk R/T is bad, all traffic could be corrupted.

→Klondike ATA Backend loop uses a BCC for the loop

→Stiletto DAE2P & DAE3P Backend loop uses disks that switch into and out of the loop, meaning that only traffic to a bad disk is corrupted—soft SCSI errors isolated to single disk. Still, single bad cable or LCC port can corrupt all traffic and give soft SCSI errors across all disks in the loop

→All CX, CX3, & CX4 arrays use UltraPoint Disk Array Enclosures (DAE2P & DAE3P, respectively), which allow for direct connection to each drive via a switch mechanism within the LCC

**CLARIION DISK ENCLOSURE TYPES:**

→DAE2 Katana 2Gb Disk Drive Enclosure

→DAE2-ATA Klondike SATA I 2GB Disk Drive Enclosure

→DAE2P Stiletto 2GB Disk Drive Enclosure

→DAE3P Stiletto SATA-II Northstar 4GB Disk Drive Enclosure

→DPE2 CX2/3/4/500

→SPE CX700

→SPE3 Jackhammer, Sledgehammer (CX3-20/CX3-40)

→SPE2 HammerHead (CX3-80)

### **TOTAL OF 5 SYSTEM DRIVES FOR CLARIION USING RAID3:**

**SPA Boot Drives 0 & 2**

**SPB Boot Drives 1 & 3**

**Note:** FLARE Configuration & PSM located on first 3 drives

### **CLARIION AX100 STORAGE ARRAY (Piranha):**

Supported with NetWin 200 Version 2.0 & NetWin 110

Entry-level array with (12) Serial ATA (SATA) drives in 2U 3.5" rack-mounted enclosure, max. of 2TB

Supports Windows, Linux, NetWare Servers

AX100SC = Single Controller model/AX100 = Dual processor Controller model with mirrored cache

### **SPS BATTERIES:**

Tested once/week, default is that both are tested at once, which disables write cache, but in actuality, Flare code will allow only one SPS to test itself at a time. Whole purpose of SPS's are to destage data from write cache to vault disks in event of power failure, 2 minutes of backup time. At least one SPS must be fully charged for Write-Cache to be enabled.

### **CLARIION CX700 BARRACUDA STORAGE ARRAY:**

58.4TB capacity with (4) 3GHz processors and (8) back-end fibre ports, support for up to 256 HA Hosts

SPE enclosure houses SPs, Pwr Supplies, Fan modules, SPS units

Vault drive private space uses about 6.4GB per first five drives [SPA Boot on Disks 0 & 2; SPB Boot on Disks 1 & 3; PSM on disks 0-2; Vault area 2.1GB on each disk for all disks 0-4; Flare DB on disks 0-2; NAS core on disk 4; Image repository on disk 4

Mix ATA and Fibre Drives (max. 240 drives, either 4-loop configuration, 60 drives per BE port, or 2-loop configuration with 120 disks per BE port), dual 3.2GHz SPs, 200K IOPS

(4) @2GB Optical FC LC Front-end ports per SP for Host connectivity

(4) @2GB HSSDC-style copper Back-end FC ports per SP—two HSSDC connectors for BE0, BE1, and two HSSDC connectors for BE2/AUX0 & BE3/AUX1 as additional backend loops

(2) 2GB FC CMI [Communications Messaging Interface] channels between each SPs on 64-bit 100MHz PCI-X bus

High-end midrange array with dual 3-GHz processors per SP; Each SPE contains 4GB Cache per SP (3072MB Write, 144MB Read) and 2GB DAE2 disk array enclosure; ranging from 5-240 disks, either Fibre and ATA drives

Max. of 256 Hosts per array at 8/port, 240 drives, 2048 LUNs

**Fibre Channel Drives (DAE2 Katana LCC Cards):** 36GB & 73GB 10k/15k; 146GB 10k;

**ATA Drives (Klondike DAE2 BCC Cards):** 250GB 5400k ATA

### **Fibre Channel Drives (DAE2P Stiletto Enclosure):**

Drives are on a FC loop but only one disk is switched onto the loop at a time—switching technology

**Note:** Enclosures are daisy-chained together, but all traffic must pass through entire backend loop

Requires Flare 13, and has four backend fibre channel ports per SP, uses Windows XP embedded O/S

Max. 256 Hosts per array, max 16 drives per RAID Group, max 128 LUNs per RAID group, 2048 LUNs per array, max. 512 Storage Groups per array; Max 256 LUNs per Storage Group (Initiator)

Min. config is a 4U SPE, a 1U SPS, and a 3U DAE2 = 8U

Max. config is a 4U SPE, a 1U SPS, and (16) 3U DAE2's = 53U

Main difference in ATA enclosures vs. FC are the disks and a modified LCC called a BCC (Bridge Controller Card) that converts FC signals to Serial ATA and vice versa

Upgrade from CX300/CX500

### **CX700 xPB Module:** L-to-R

10/100 LAN port | Grn & Yel LEDs, pwr & fault + Bios, Post, Boot, respectively | SPS async comm. port | IOIOI async console port | Aux0 & Aux1 Fibre Ports | BE0 & BE1 for BackEnd fibre channel ports

### **CX700 NAS Personality Module:** L-to-R

GbE Ports 11, 12, 13 on one PCI-X Bus | GbE Ports 14, 15, 16 on other PCI-X Bus with Green Link light & Yellow activity LED

### **CX700 SAN Personality Module:** L-to-R

2Ghz Fibre Ports 0, 1, 2, 3 FrontEnd Host connections with Green Link LED

### **LAYERED CLARIION APPLICATIONS:**

#### **Clones (BCVs):**

Consistent clones in Flare 19

#### **SnapView Snapshots V2.19:**

Application that can create instantaneous point-in-time virtual copies of LUNs with copy-on-first-write capability, or full Clone point-in-time copies of LUNs; 8 per source LUN, 300 total per array, 100 Snapshot sessions per storage system ((2) types of Snapshots--copy on first write, and BCV clone copies). A Snap rollback to source LUN can be done instantly from any session. BCV's can be

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
incrementally updated from source; Snapshots & BCV's can be mounted RW on Hosts; Fracture a BCV before Host can access;  
Consistent snaps in release 19. Snaps are useful for backing up LUNs, cloning LUNs, etc.

#### **MirrorView/S:**

Full copy synchronous remote mirroring of LUNs between (2) or more arrays (FC4700/CX400/500/600/700).  
LUN of equal-size on a separate storage system that receives host writes simultaneously  
WIL (Write Intent Log) keeps track of in-flight writes to primary side, allowing SP to rebuild fracture log if SP is rebooted  
Consistency Groups keep all LUNs in a set mirrored (16 luns per consistency group, max of 16 consistency groups per CX600/700)  
Requires LUN Masking

#### **MirrorView/A:**

Full copy asynchronous remote mirroring of LUNs between (2) or more arrays, using delta sets  
Snapshots of both Source and Target LUNs can be made, but BCVs of Source/Target not supported  
Rules are similar to MirrorView/S

#### **SANCopy:**

--LUN copy across SAN between arrays or within the same array, both Full and Incremental, of LUNs, metaLUNs, BCV's (eliminates need for Hosts to move the data)  
--Useful for one-time copies of data, or for ongoing Replication between sites  
--Requires installation of the SANCopyEnabler-01.01.5.002-xpfree.ena on the array

### **OTHER APPLICATIONS:**

#### **PowerPath 4.4:**

Host-based path load-balancing and failover, with (8) different policies available [Basic load balancing disabled; Clariion Optimized; Least Blocks Load Balancing; Least I/O Load Balancing; No redirect; Request—uses O/S default; Round Robin; Symmetrix Optimized].

Supports I/O prioritization, FC and iSCSI arrays, automatic path failover, up to 16 HBA's per Host (Hosts that use PowerPath can allow trespassed luns to be restored automatically, whereas Celerra does not)

PowerPath 5.2 supports Windows 2008 hosts

#### **CLARAlert 6.2:**

Remote support for Clariion arrays with Windows management station and modem; notifications sent via emails or dial-up

### **CLARIION CX500 TARPON STORAGE ARRAY:**

NAS 5.3.11-4 and higher

Mid-size array with support for up to 128 HA Hosts, 3U Assembly, DPE2 houses both SPs, 15 disks, two pwr supplies, & cooling fans  
Fibre & ATA Drives mixed in same array, up to 120 drives total with max. of 120,000 iops; Dual SPs; Up to (7) added DAE2's;  
2GB disk array ranging from 5 -120 disks, Fibre or ATA [36GB 10/15k; 73GB 10/15k; 146GB 10k; 250GB 5400rpm ATA]

Requires Flare 13 and supports up to (2) GB Cache per SP [4GB total cache]

DPE2 contains dual 1.6GHz Intel P4 Xeon processors (SP's) with 533MHz FSB, and first (15) drives

DPE2 contains both CX500 SPs and up to (15) drives in the enclosure, while DAE2 contains (2) LCC cards and up to (15) drives  
Two SPS units--from back, RH SPS power-out & sense cables connect to pwr supply A and SPA, LH SPS to pwr supply B and SPB  
Flare resides on first (5) drives, user data resides on 6.4GB and higher portion of each drive

Max. 512 MetaLuns per Storage Array, max. lun size 2TB, and max. number of volumes to build lun is (16), max. storage groups 256  
(4) front-end FC ports and (4) back-end FC ports

(2)@2GB Fibre Channel CMI between SP's (Communications Messaging Interface), Fibre channel running on 100MHz bus

(2)@2GB Optical SFF LC (Lucent Connector) FE ports per SP—Port 0 used only for Host connection; Port 1 can be used for Host connection, unused, or for Remote Mirror connection & (2)@2GB HSSDC Copper BE fibre loops ports for each SP (BE0 & BE1)

**Note:** ATA implementations use BCC (Bridge Controller Cards) that are used to convert FC signals to serial ATA signals on the drives, and vice versa. CX300/CX500 use DAE2 with SP integrated into the LCC

Max. of 128 hosts/array, 4/port, 120 drives, 1024 LUNs

### **CLARIION CX500i/CX300i iSCSI ARRAYS:**

120/60 disks, respectively; 1.6GHz/800MHz SPs, respectively; (2)@1Gb copper ethernet iSCSI FE ports per SP; (2)@2GB BE ports per SP for CX500, (1)@2GB BE port per SP for CX300; 4GB & 2GB Cache per SP, respectively; PowerPath V4.3.1 supports iSCSI arrays; MirrorView and SAN Copy are not supported; Gbe Jumbo frames not supported; Only (1) login per iSCSI name per SP port; MS iSCSI HBA initiator uses same name for all NICs and allows only (1) login per Server to each SP array port; QLogic iSCSI HBA initiator can allow (1) login per HBA name to each array SP port.

### **CLARIION CX300 SNAPPER STORAGE ARRAY:**

Small-size array with large capacity—13TB

Fibre & ATA Drives, up to 60 drives [DPE2 module with 15 drives, which also houses SPs, PwrSupplies, Cooling fans, & up to (3) DAE2 enclosures of 15 drives each]; Upgrades from CX200 not supported; 1GB memory per SP; SP's run 800MHz PIII Tualatin-LP processor with 512KB Level 2 cache; Single or two SPS units

(2)@2GB FE fiber optical ports per SP; (1)@2GB BE HSSDC port per SP; (2) serial ports; Single LAN port

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
Disk array, ranging from 5 – 60 disk drives [36GB 15k; 73GB 10/15k; 146GB 10k; 250GB ATA]; Requires Flare 13, supports (2) GB Cache per SP, uses embedded Windows XP O/S

**Note:** The CX500/CX300 arrays are essentially the DAE2 model with the SP integrated into the LCC  
Max. 64 Hosts/array, 4/port, 60 drives, 512 Luns

### **MAX CAPACITY FOR FIBRE CHANNEL DRIVES:**

| Disk  | CX700  | CX500  | CX300 |
|-------|--------|--------|-------|
| 36GB  | 8.6TB  | 4.2TB  | 2.1TB |
| 73GB  | 17.5TB | 8.6TB  | 4.4TB |
| 146GB | 35TB   | 17.4TB | 8.6TB |
| 300GB | 72TB   | 36TB   | 18TB  |

**Note:** Max number of FC drives per RAID Group is 16

### **MAX CAPACITY FOR ATA DRIVES/ATA STORAGE:**

| Disk  | CX700 | CX500  | CX300  |
|-------|-------|--------|--------|
| 250GB | 56TB  | 26.2TB | 11.2TB |
| 320GB | 72TB  | 33.6TB | 14.4TB |

**Note:** Max number ATA drivers per RAID Group is 16, first DAE is FC with min. of (5) drives. 250GB drives are 7200RPM SATA; 320GB drives are 5400RPM PATA. Outside of first DAE, others can be mixed ATA or FC.

**Note:** PATA stands for Parallel ATA drives, with parallel buses from controller to disk using ribbon-type cable, running on 133Mbps buss. Serial ATA using serial communication between controller and disk and runs at 150mbps bus speeds.

### **CLARIION ATA DRIVES/STORAGE:**

Large drives faster than tape but cheaper than fibre-channel. Initial Celerra release consisted of single AVM profile for archiving purposes. Current RAID 5 6+1 Groups with ATA-DAE2 consist of (15) 250GB drives, for a total storage of 2.7TB per Shelf (max. of 4TB allowed per DM for Fibre Channel drives, while 16TB total allowed for ATA drives per DM). Non-integrated storage is storage allocated to Celerra using Access Logix storage groups included in SAN/NAS configurations. Default clarata\_archive profile will attempt to find two LUNs from different RAID groups to create a file system, or from a single unallocated LUN if that fails.

SATA—Serial ATA Drives, first type used

PATA—new 250GB & 350GB PATA Drives used with CX600/CX700.

### **FLARE 19 PATCH 21 ATA DRIVE ISSUES:**

250GB PATA drives have potential issue with systems that are not running Flare 19 Patch 21:

→Disk drive may report successful completion of Write that later results in error on Read: Flare 19 Patch 21 fix is to verify each write

→Disk drive may intermittently timeout during error handling, resulting in faulted drive: Flare 19 Patch 21 fix is to adjust the timeout issues to prevent

→Regular ATA drive rebuild operations will be enhanced by a factor of (3) after drive failures

→Background sniffer operations are 4x faster than previously (Detects and corrects many ‘Soft Media Error’ event 820)

### **SOFT MEDIA ERROR FROM SPCOLLECT:**

```
02/13/2007 15:08:54 Bus 3 Enclosure 0 Disk 11(6a0) Disk soft media error [0x00] 0      22
02/13/2007 18:46:44 Bus 3 Enclosure 0 Disk 11(820) Soft Media Error    [0x00] 0      5
02/13/2007 18:46:46 Bus 3 Enclosure 0 Disk 11(689) Sector Reconstructed [0x00] 267a748 28003000
02/15/2007 05:08:32 Bus 3 Enclosure 0 Disk 11(684) Parity Sector Reconstructed [0x00] abc4c8 28003000
```

### **WINDOWS O/S FOR ARRAY MODELS:**

→CX700/500/300 arrays run embedded XP  
→CX600/400/200 & FC4700 run embedded NT

### **NS700 PLATFORMS:**

#### **BACKENDS:**

CX700→Barracuda      CX400→Tarpon      CX200→Snapper

#### **NS-BARRACUDA: Integrated & Gateway Models**

**Specs:** BARRACUDA SPE 4U Assembly--(2) @3GHz P4 Xeon processors per SP; 533 MHz FSB 4GB 266MHz DDR Memory

**Front-End:** NAS personality card with (6) copper Broadcom 10/100/1000 ports & (2) optical GB ports

**Back-End:** Support for Symm 5.0 & 6.0 DMX only; CX600, CX400, CX200; FC-AL for NAS-only and direct-attached SAN/NAS; FC-SW for fabric-attached SAN/NAS with Access Logix required; Support for FLARE 13; Supports baseline 8GB memory per SP. Supports 15k 36GB FC, 10k 73GB FC, 15k 72GB FC, 10k 146GB FC, 250GB ATA, 320GB ATA. O/S runs on XP. CX700 has four buses to backend, so AUX-0 and AUX-1 can be used for BackEnd or Tape Drives [BE-2 & BE-3, respectively]

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
Raid 5 (4+1, 8+1); Raid 5 (4+1); Raid 1; Raid 5 (6+1) ATA drives; (3) Fan Modules; 2-minute SPS for Write Cache destaging as backend for SPE and first DAE2; Max. 240 drives per array; same DAE2 style as the CX600

- (4) 2GB front-end fibre LC ports per SP
- (4) 2GB back-end copper HSSDC fibre channel loops (BE0-BE3)

#### **Requirements:**

NAS 5.2 & Flare 13 that supports both non-Access Logix & Access Logix

Supports Timefinder/FS & RDF

#### **NS4B BARRACUDA:**

- Gateway NAS version with (4) Data Movers [3 Primary and 1 Standby]  
Enclosure 0 = Server\_2 & 3 Enclosure 1 = Server\_4 & 5
- Dual Control Station support [Falcon 1U SPO Storage Platform Operations]
- CS reset using IPMI v1.5 Intelligent Platform Management Interface [CS will use IPMI to reboot and detect presence of standby]—CS communicates with hardware via BIOS. IPMI consists of a physical link between Control Stations and can be verified with the following command if logging “IPMI Connection Failure” events. Cannot reboot the other CS if this link is down, though CS failover would still occur.

#### **#/nasmcd/sbin/t2tty -m emcnasotherIPMICS\_i3**

- Com1 & Com2 will be used to communicate to CS0 and CS1
- Broadcom Ajaguar NAS personality card optional [8 Ethernet and 1 Management ports]
- CS boots from Linux OS from onboard IDE driver and accesses NASDB & DOS via NBS protocol
- DMs PXE boot from CS ide drive if they cannot find DOS boot partition on backend storage—if no Enclosure ID is found, or DM MAC addresses mismatched, DM boots to minimal config from slot\_99 with reason code 14
- New Linux boot loader replaces LILO [GRUB]

#### **NS INTEGRATED SYSTEMS:**

- setup\_clariion script is run on CS, via Celerra Manager or CLI, to configure backend
- Availability of System\_defined Templates, as well as use of User\_defined Templates

#### **CX Standard Raid 5 Template:**

- 4+1 RAID 5, HS, 8+1 RAID5 for 1<sup>st</sup> Shelf
- (3) 4+1 RAID 5 for next shelf, following by alternating 4+1 RAID5 8+1 RAID5, and (3) 4+1 RAID5

#### **CX All 4plus1 Raid 5 Template:**

- 4+1 R5, HS, 4+1 R5, HS, HS, HS, HS 1<sup>st</sup> shelf
- All other shelves use (3) 4+1 R5

#### **CX Standard Raid 1 Template:**

- 4+1 R5, HS, 8+1 R5 on 1<sup>st</sup> shelf
- All other shelves have (7) RAID1 groups

#### **User Defined Templates:**

- greater flexibility in mixing shelves of diff. RAID types, each shelf must be defined by User, RAID3 can be used and mixed with RAID 5 shelves
- RAID3 Introduced in NAS 5.4 for ATA B2D solutions
- Configure RAID1 for logs and RAID5 for tablespace for Oracle solutions

**Note:** Arrays with ATA storage must be setup using User\_defined Templates, using CLI

#### **CHECKING SP IP ADDRESS CONFIGURATION:**

#### **# /nasmcd/sbin/navicli -h 10.241.168.52 networkadmin -get**

Storage Processor: SP A  
Storage Processor Network Name: nyip2spa  
Storage Processor IP Address: 10.241.168.52  
Storage Processor Subnet Mask: 255.255.255.0  
Storage Processor Gateway Address: 10.241.168.128

#### **CHANGING IP ADDRESSES OF SP's:**

#### **C:>navicli -h 192.168.1.201 networkadmin -set -address 192.168.1.200 –subnetmask 255.255.255.0 –gateway 192.168.1.100**

**Note:** SP's reboot when changing IPs, until Flare 29

#### **CHANGING CLARiiON SP IP ADDRESSES WITH GATEWAY CELERRA:**

#### **Prerequisites:**

NAS 5.6.40 or higher

See emc200831 for the published procedure. Please note that the following procedure contradicts the official CLARiiON procedure in that SP's should be removed from the Security Domain and all security destroyed before changing SP IP Addresses. The following

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!  
procedure to update the symapi with the changed IP addresses will not work if the Security Domain is removed. It's possible, yet untested, that the Clariion steps could be performed to change IP addresses, re-establish Domain Security, and then the Celerra symapi updated.

1. Change the IP address for SPA using Navisphere or CLI
2. After waiting for SPA to reboot and stabilize, repeat Step 1 for SPB

**Note:** Wait until array settles back to normal operation before continuing

3. Use either of the following commands to update the symapidb with the new IP addresses:

**# nas\_storage -modify id=3 -network -spa 10.241.168.40 -spb 10.241.168.41**

Changing IP address for APM00030600863

**# nas\_storage -check -all**

Discovering storage (may take several minutes)

Done

4. Update CLARiiON Security cache files on the Control Station:

**# nas\_storage -modify id=3 -security -username nasadmin -password nasadmin**

Setting security information for APM00030600863

Done

5. Verify security file updates:

```
# ls -la
-rw-rw-r-- 1 nasadmin nasadmin 276 Oct 21 15:49 .clar_security
-rw-rw-r-- 1 nasadmin nasadmin 0 Oct 21 15:49 .clar_security.lck
# ls -la /etc/.clar*
-rw-rw-r-- 1 nasadmin nasadmin 276 Oct 21 15:49 /etc/.clar_security
-rw-rw-r-- 1 nasadmin nasadmin 0 Oct 21 15:49 /etc/.clar_security.lck
# pwd
/nas/symapi/config
# ls -la
-rw-rw-r-- 1 nasadmin nasadmin 229 Oct 21 15:49 emcpwddb.dat
```

6. Vi update the /etc/hosts and /nas/site/sp\_info files with new IP addresses

7. Reboot Control Station, then verify IP communication & Navicli authentication with the array

**# /nas/sbin/navicli -h 10.241.168.40 security -list**

Username: nasadmin

Role: administrator

Scope: global

8. Trespass LUNs to Owner SP if required, using Celerra Manger or CLI ‘trespass lun x’

9. Run setup\_clariion -check\_and\_update -all

**# /nas/sbin/setup\_clariion -check\_and\_update APM00030600863**

CLARIION(s) APM00030600863 will be setup.

Setup CLARiiON APM00030600863 storage device...

System 10.241.168.40 is up

System 10.241.168.41 is up

Clariion Array: APM00030600863 Model: CX600 Memory: 4022

Validating APM00030600863 storage group 'Hammer1'

Setup of CLARiiON APM00030600863 storage device complete.

Discovering storage (may take several minutes)

Done

## **CELLERRA NSX SERIES:** (aka Celerra Hammerhead)

--Titan 40U-60 cabinet consists (single or dual 24 AMP service) of Enclosures, Control Stations, ATS Boards, UPS's, and Modems

--First released with NAS 5.4.14.3, Champagne code—Gateway solution only

--NSX System totals up to (4) Data Mover enclosures, (2) Control Stations, (2) UPS, and (2) Modems

--Units ship with minimum 4-Blades and up to 8-Blades maximum, and minimum of 2-Control Stations

--High-end NAS Gateway using up to (8) data movers [aka Blades, XBlade 60 used in this first generation model] (4-enclosures daisy-chained with Ethernet) and 1-2 Chivas Control Stations

--Designed to replace the CNS cabinets [upgrade from CNS or NS-Barracuda-4]

--Supports Clariion and/or Symmetrix

--No serial communication between CS and DM, only RPC via internal network

--Shipped with Titan cabinet of (4) DM enclosures pre-wired

**Note:** Chivas CS will only be used with NSX series

## **NSX CABINET Bottom-to-Top:**

### **NSX HAMMERHEAD FEATURES:**

- NS & NSX Gateway Models will have ability via ksnas.cfg file to automatically configure Zoning, WWN's, LUNs, RAID GROUPS, and STORAGE Groups during installation for CONTROL LUNS only, as well as Ethernet for CS & DM's
- boot: serialkickstart ksnas.cfg** → Installs are invoked with this command after booting from install floppy
- faster NAS installs as there are multiple PXE boots of servers [User LUNs created after install]
- New utilities to add, remove FRU's & enclosures, as well as to verify cabling: setup\_enclosure; enclosure\_status
- Updated t2tty facility for NSX that uses internal network to pass commands from CS to DM (no serial from CS to DM with NSX)
- New getreason -e facility and codes to troubleshoot cabling and hardware issues

### **XBLADE 65 H2G2 2<sup>nd</sup> GENERATION NSX:**

- New Xblade65 data mover supported with NAPA II NAS 5.5.22.2, GA Aug 11, 2006
- 3.6GHz P4 processor 4GB memory 800MHz fsb with (3) I/O Modules & Wildfire-4 motherboard:  
One ‘twister’ blade module with four 4GB FC ports [BE0, BE1, FE0, FE1] for tape and storage connections, one ‘blizzard’ blade module with two Optical 1GB Ethernet ports [fge0, fge1] and six Copper 10/100/1000 BaseT Ethernet ports [cge0, cge2, cge3, cge5], and a third ‘tempest’ module with a single Neterion Xframe Optical 10GbE port [fxg0]
- NSX supports 2-4 blades, 2 Control Stations, 2 ATS, 2 UPS, 2 Modems, contains Broadcom 5325M 5-port Management switch—3 ports [Port 0 Uplink eth0, Port 3 Downlink to Port 0 next enclosure, Port 4 to CS1 or UPS management port in Enclosure 2] are accessed from front of the switch and the other two are wired into the backplane and accessed from RS232 Serial Ports—A port is for Blade0, B port is for Blade1, requiring special USB-to-Serial cable Part # 038-003-157

### **NSX BLADE ENCLOSURES:**

NSX can have up to (4) Blade Enclosures, called Blade Enclosure 0 – 3, each with (2) Data Mover “Blades” per Enclosure  
Each Enclosure consists of (2) DM's, (2) Management switches, & (2) Power Supplies

Data Mover module contains Fibre IO module on left side and Ethernet IO module on right side, memory in middle, CPUs in back

### **ENCLOSURE MONITORING:**

4 fans; 2 pwr supplies; 2 Management Switches; 2 DM blades; 2 I/O Modules per DM blade;

Cabinet UPS Monitoring & new FRU faults; (1) I/O annex per DM blade

**Note:** NSX/NS80 IO Annex sits underneath the Blades, with left-half of the Annex for Server\_3 and right-half of Annex for Server\_2. IO Annex supports 10GbE Ethernet IO card. See 100-561-787 for IO Annex Assembly.

### **REPLACING ENCLOSURES:**

→ Stop NAS, Replace Enclosure, run following commands:

```
# /nas/sbin/setup_enclosure -removeEnclosure <encl id>
# /nas/sbin/setup_enclosure -checkSystem
# /nas/sbin/setup_enclosure -addEnclosure <encl id>
# /nas/sbin/setup_enclosure -checkCable
# /nas/sbin/setup_enclosure -changeSubnet 0 | 1 <new subnet>
# /nas/sbin/setup_enclosure -checkSystem autoRepair
```

→ Configure new WWNs and Storage Group configurations, restart NAS

```
# /nas/sbin/enclosure_status -e 0 -v
```

**Note:** Output below is for Enclosure 0—adjust syntax when troubleshooting other Enclosures. Checks for Alarms, Status, and FRU failures

|                           |                               |
|---------------------------|-------------------------------|
| DEVICE A                  | DEVICE B                      |
| -----PRESENCE-----        |                               |
| 00 60 16 05 79 4E         | MAC 00 60 16 05 78 A4         |
| Typhoon                   | Hardware Platform Typhoon     |
| Absent                    | I/O Annex Absent              |
| Present                   | Peer Typhoon Present          |
| Present                   | Compute Blade Present         |
| Powered On                | Compute Blade Powered On      |
| Present                   | Power Supply Present          |
| Inactive                  | Manufacturing Mode Inactive   |
| Inactive                  | Margin High Mode Inactive     |
| Inactive                  | Margin Low Mode Inactive      |
| Inactive                  | PSA I2C Bus in Reset Inactive |
| Inactive                  | PSB I2C Bus in Reset Inactive |
| -----BROADCOM STATUS----- |                               |
| Inactive                  | B /P4 FullDuplex Inactive     |
| Inactive                  | B /P4 100 MBPS Inactive       |
| Inactive                  | B /P4 Up Inactive             |

Active Dnlink/P3 FullDuplex Active  
Active Dnlink/P3 100 MBPS Active  
Active Dnlink/P3 Up Active  
Active BladeB/P2 FullDuplex Active  
Active BladeB/P2 100 MBPS Active  
Active BladeB/P2 Up Active  
Active BladeA/P1 FullDuplex Active  
Active BladeA/P1 100 MBPS Active  
Active BladeA/P1 Up Active  
Active Uplink/P0 FullDuplex Active  
Active Uplink/P0 100 MBPS Active  
Active Uplink/P0 Up Active

-----ENCLOSURE STATUS-----

Inactive PS 12V\_2 Disabled Inactive  
Inactive PS 12V\_1 Disabled Inactive  
Inactive SP\_V2 Not Present Inactive  
Inactive SP\_V1 Not Present Inactive  
Inactive Blower Speed Up Inactive  
Active 12V\_SBY Voltage Active  
Active 12V\_2 Voltage Active  
Active 12V\_1 Voltage Active

-----ENCLOSURE ALARMS-----

Pass 12V\_2 OverCurrent Pass  
Pass 12V\_1 OverCurrent Pass  
Pass 12V\_SBY High Pass  
Pass 12V\_SBY LO Pass  
Pass 12V\_2 HIGH Pass  
Pass 12V\_2 LO Pass  
Pass 12V\_1 HIGH Pass  
Pass 12V\_1 LO Pass  
Pass Blower Power D Fault Pass  
Pass Blower Power C Fault Pass  
Pass Blower Power B Fault Pass  
Pass Blower Power A Fault Pass  
Pass Blower Tach D Error Pass  
Pass Blower Tach C Error Pass  
Pass Blower Tach B Error Pass  
Pass Blower Tach A Error Pass  
Pass PS Fault Pass  
Pass PS Fan 2 Fault Pass  
Pass PS Fan 1 Fault Pass  
Pass PS Overtemp Pass  
Pass PS Seated Fault Pass  
Pass Multi Fan Fault Pass  
Pass Peer PS Fault Pass  
Pass DC Present Pass  
Pass AC OK Pass  
Pass SP\_B Size Fault Pass  
Pass SP\_A Size Fault Pass  
Pass 12V\_SBY OverCurrent Pass  
Pass PS Link1 Fault Pass

-----BROADCOM ALARMS-----

Pass 5325 Comm Fail Pass  
Pass 5325 MIB RAM Pass  
Pass 5325 MIN MEM Pass  
Pass 5325 BUFF Con Pass

-----RESUME CSUM ERRORS-----

Pass Typhoon x Csum Error Pass  
Pass PS x Csum Error Pass  
Pass CB x Csum Error Pass

Pass Midplane Csum Error Pass

-----PEER PRESENCE CONTRADICTIONS-----

Pass Peer PS Contradict Pass

Pass Peer CB Contradict Pass

Pass Peer TY Contradict Pass

Pass Peer IO Contradict Pass

-----COLDFIRE ALARMS-----

Pass P On Self Flash Fail Pass

Pass P On Self RAM Fail Pass

Pass P On Self FEC Fail Pass

Pass P On Self 5325 Fail Pass

Pass Run Self Flash Fail Pass

Pass Run Self RAM Fail Pass

Pass Diag 5325 Fail Pass

Pass Encl Resume Block Pass

Pass Peer Rs232-A traffic Pass

Pass Peer Rs232-B traffic Pass

Pass MidPlane ID Read Pass

Pass I2C PSA Arbit Error Pass

Pass I2C PSB Arbit Error Pass

Pass I2C PS A Bus Error Pass

Pass I2C PS B Bus Error Pass

-----STATUS CONDITIONS-----

Valid Peer Typhoon Status Valid

Valid LM93 Status Valid

Valid Fault Expander Valid

Valid Data Mover Status Valid

Valid Power Supply Status Valid

-----LM93 ALARMS-----

Pass CB A Overtemp Pass

Pass CB B Overtemp Pass

Pass CB A Vcc Pass

Pass CB B Vcc Pass

Pass Vtt Alarm Pass

Pass V 1.8 Alarm Pass

Pass V 1.5 Alarm Pass

Pass V 1.0 Alarm Pass

Pass CPU Thermal Trip Pass

Pass PDN LM93 P2 Trip Pass

-----FRU STATUS-----

Pass FRU CPU DIMM 0 Pass

Pass FRU CPU DIMM 1 Pass

Pass FRU CPU DIMM 2 Pass

Pass FRU CPU DIMM 3 Pass

Pass FRU CPU DIMM 4 Pass

Pass FRU CPU DIMM 5 Pass

Pass FRU CPU DIMM 6 Pass

Pass FRU CPU DIMM 7 Pass

Pass FRU CPU Module Pass

Pass FRU CPU IO Mod 0 Pass

Pass FRU CPU IO Mod 1 Pass

Pass FRU IO Annex Pass

Pass FRU Enclosure Pass

Pass FRU Blower Mod A Pass

Pass FRU Blower Mod B Pass

Pass FRU Blower Mod C Pass

Pass FRU Blower Mod D Pass

Pass FRU Typhoon Pass

Pass FRU Power Suply Pass

-----SYSTEM VARIABLES-----

```
02    Slot ID      03
00    Enclosure ID 00
00    BackPlane ID 01
61    Post Code     61
05    Reason Code   05
2C    Blade Status Code 2C
80    Post Middle 8 bits 80
00    I2C PSA Error Mask 00
00    I2C PSB Error Mask 00
00    I2C Error Mask 00
```

**# /nas/sbin/setup\_enclosure -showAll | -checkCable | -checkSystem | -probeSystem | -readConfig | -addMgmtswitches** → used to initialize enclosureIDs and verify cabling

**# /nas/sbin/enclosure\_status** → Status of hardware components in the Enclosure [4-Fans, 2-Pwr Supplies, 2-Mgmt Switches, 2-Blades, 2-I/O Modules]

**# /nas/sbin/enclosure\_status -e 0** (or -a 192.168.1.50 mgement switch ip address)

```
Typhoon A MAC Addr = 00 60 16 05 79 4E
```

```
Typhoon B MAC Addr = 00 60 16 05 78 A4
```

```
DEVICE A  DEVICE B
```

-----Status Flags-----

```
32 90 00 00 -- 32 90 00 00
```

```
0F FF 00 00 -- 0F FF 00 00
```

```
00 00 00 00 -- 00 00 00 00
```

```
80 00 00 00 -- 40 00 00 00
```

```
00 00 00 00 -- 00 00 00 00
```

```
00 00 00 00 -- 00 00 00 00
```

```
00 00 00 00 -- 00 00 00 00
```

```
00 00 00 00 -- 00 00 00 00
```

```
00 00 00 00 -- 00 00 00 00
```

```
00 00 00 00 -- 00 00 00 00
```

```
00 00 00 00 -- 00 00 00 00
```

```
00 00 00 00 -- 00 00 00 00
```

```
00 00 00 00 -- 00 00 00 00
```

```
00 00 00 00 -- 00 00 00 00
```

```
00 00 00 00 -- 00 00 00 00
```

```
00 00 00 00 -- 00 00 00 00
```

```
00 00 00 00 -- 00 00 00 00
```

```
00 00 00 00 -- 00 00 00 00
```

```
00 00 00 00 -- 00 00 00 00
```

## **RESUME CSUM ERRORS & AMBER LEDs ON ENCLOSURE EXAMPLE:**

Enclosure\_status command was returning following error

**# /nas/sbin/enclosure\_status -e 0 -v**

-----RESUME CSUM ERRORS-----

```
Failed    Typhoon x Csum Error Failed
```

```
Pass     PS x Csum Error Pass
```

```
Pass     CB x Csum Error Pass
```

```
Pass     Midplane Csum Error Pass
```

**Corrective Action:** Reset the Management Switches by rebooting the switches

**# /nas/sbin/setup\_enclosure -resetMgmtswitches**

Run the appropriate reset checksum command to the affected switch:

**# /nas/sbin/t2net\_test -SetResumeCksum 4 20**

**Note:** The first number “4” is slotID [between 2-9] & second number is the operand--“20” is Hex for decimal value 32, which is Management Switch A. You may want to check with EE before using this command—see emc153229

**# /nas/sbin/t2net\_test -SetResumeCksum 4 34**

Resume cksum for this device has been updated. New Cksum = 0x6a8f57d9

t2net\_test: Command succeed

**Note:** Another syntax example, not sure which one is the correct one

### **operand values:**

1(Blade A Smbus Slave Port)

2(Blade B Smbus Slave Port)

3(Blade A Fault Expander)

4(Blade B Fault Expander)

5(Power Supply A)

6(Power Supply B)

7(LM93 A)

8(LM93 B)

**Following Decimal values work on all NAS 5.4 and 5.5.28.0+ code (AR80391)—convert to Hex for earlier versions of 5.5:**

32(Mgmtswitch A) →Hex = 20, etc.  
33(Mgmtswitch B)  
34(Blade A)  
35(Blade B)  
36(I/O Annex A)  
37(I/O Annex B)  
38(Blade A I/O A)  
39(Blade A I/O B)  
40(Blade B I/O A)  
41(Blade B I/O B)  
42(Power Supply A)  
43(Power Supply B)  
44(Midplane)  
45(Blade A - I/O Blade 0 SFP0)  
46(Blade A - I/O Blade 0 SFP1)  
47(Blade A - I/O Blade 0 SFP2)  
48(Blade A - I/O Blade 0 SFP3)  
49(Blade A - I/O Blade 1 SFP0)  
50(Blade A - I/O Blade 1 SFP1)  
53(Blade B - I/O Blade 0 SFP0)  
54(Blade B - I/O Blade 0 SFP1)  
55(Blade B - I/O Blade 0 SFP2)  
56(Blade B - I/O Blade 0 SFP3)  
57(Blade B - I/O Blade 1 SFP0)  
58(Blade B - I/O Blade 1 SFP1)

**Example of getreason -e codes and comparison to Getsyscode output for Slot\_2:**

# /nas/sbin/getreason -e

5 - slot\_2 contacted (Post Code = 0x61 ; Mid Code = 0x80 ; Blade Code = 0x2c)

**Note:** Post Code is operand value 3; Mid Code is operand 8; Blade Code or fault is operand 6

# /nas/sbin/t2net\_test -Getsyscode 2 3

The returned system value is 0x61

# /nas/sbin/t2net\_test -Getsyscode 2 8

The returned system value is 0x80

# /nas/sbin/t2net\_test -Getsyscode 2 6

The returned system value is 0x2C

**Getsyscode Operand Values are different from SetResumeCksum:**

# /nas/sbin/t2net\_test -Getsyscode

**operand values:**

0(Blade ID),  
1(Enclosure ID),  
2(Backplane ID),  
3(POST code),  
4(Reason code),  
6(Blade Status or Fault Code),  
8(Middle Byte of 24-bit POST Code),  
9(I2C PS0 Device Error Mask),  
10(I2C PS1 Device Error Mask),  
11(I2C Device Error Mask)

**ENCLOSURE RESUME PROM CHECKSUM ISSUES:** TMobile & Nationwide & Amber LED issues on Enclosures

**TMOBILE FIX:**

**Force flashing Coldfire Mgmt Switch & Resetting Cksums on Typhoon Enclosure Resume Prom:**

# cat /etc/hosts |egrep MGMT

```
192.168.1.50 mgmt_2_3      #ENCLOSURE-0_MGMT_A
192.168.1.51 mgmt_4_5      #ENCLOSURE-1_MGMT_A
192.168.1.52 mgmt_6_7      #ENCLOSURE-2_MGMT_A
192.168.1.53 mgmt_8_9      #ENCLOSURE-3_MGMT_A
192.168.2.50 mgmt_2_3b     #ENCLOSURE-0_MGMT_B
192.168.2.51 mgmt_4_5b     #ENCLOSURE-1_MGMT_B
```

```
192.168.2.52 mgmt_6_7b      #ENCLOSURE-2_MGMT_B  
192.168.2.53 mgmt_8_9b      #ENCLOSURE-3_MGMT_B  
# /nasmcd/sbin/setup_enclosure -flashFirmware force 192.168.1.50
```

```
# /nas/sbin/enclosure_status -v -e 1
```

-----RESUME CSUM ERRORS-----

Failed Typhoon x Csum Error Failed

```
# /nas/sbin/t2net_test -SetResumeCksum 4 34 Resume cksum for this device has been updated. New Cksum = 0x6a8f57d9
```

#### **NATIONWIDE FIX:**

```
# /nas/sbin/setup_enclosure -initSystem force
```

Run setup\_slot on blades in the Enclosure that was replaced or faulted

#### **AERA ENERGY FIX—NS80 INSTALL:**

**Amber LEDs on Both Enclosures on NS80, along with Cksum Error on enclosure status -v -e output—fixed with following:**

```
# /nas/sbin/t2net_test -SetResumeCksum 4 20
```

**Note:** The 20 value is in hex because of NAS 5.5.25 code

#### **ADDING ENCLOSURE:**

→Add to next available slot and connect to Ethernet networks

→Run setup\_enclosure –addEnclosure, -checkSystem, and –checkCable to verify installation

#### **ENCLOSURE COMMAND OUTPUT:**

```
# /nas/sbin/setup_enclosure
```

Usage:

```
/nas/sbin/setup_enclosure [-q]
```

```
| -checkCable  
| -checkSystem  
| -probeSystem  
| -readConfig
```

```
# /nas/sbin/setup_enclosure -checkCable
```

Executing -checkCable option

System discovery on both subnets ... OK

Pair up all discovered Mgmtswitches ... OK

Verify no cross-connected cabling error ... OK

Collect Enclosure cabling topology ..... OK

Examine Enclosure cabling ...

Cabling of Enclosure at position 0 ... OK

Cabling of Enclosure at position 1 ... OK

checkCable Completed

```
# /nas/sbin/setup_enclosure -checkSystem
```

Executing -checkSystem option

Checksum verification on ENCL\_DB & DHCPD\_CFG ... OK

Verify ENCL\_DB generation version ... OK

Current Enclosure database (ENCL\_DB) info ...

Enclosure ID# 0:

MgmtSwitch-A IP address = 192.168.1.50

MgmtSwitch-A MAC address = 00:60:16:05:79:4e

F/W Version: Mgmtswitch=01.63, Powersupply=00.27, Bootblock=01.51

MgmtSwitch-B IP address = 192.168.2.50

MgmtSwitch-B MAC address = 00:60:16:05:78:a4

F/W Version: Mgmtswitch=01.63, Powersupply=00.27, Bootblock=01.51

Enclosure ID# 1:

MgmtSwitch-A IP address = 192.168.1.51

MgmtSwitch-A MAC address = 00:60:16:05:78:ac

F/W Version: Mgmtswitch=01.63, Powersupply=00.27, Bootblock=01.51

MgmtSwitch-B IP address = 192.168.2.51

MgmtSwitch-B MAC address = 00:60:16:05:78:42

F/W Version: Mgmtswitch=01.63, Powersupply=00.27, Bootblock=01.51

/etc/nas\_enclosure.map:

Generation = 4

md5sum = fb21780e176946966e5def25a63548f1 →128 bit output

Checksum verification = OK

```
/etc/encl_dhcpd.conf:
    Generation = 4
    md5sum = b08fb8369c5f979048af002ef38bc108
    Checksum verification = OK
Performing Detail System check ...
System discovery on both subnets ... OK
Detail System check OK
checkSystem Completed
# /nas/sbin/setup_enclosure -probeSystem
Executing -probeSystem option
Discovering Primary subnet (A) ... OK
Enclosure ID = 0, BackplaneID 0 from IP 192.168.1.50
MgmtSwitch-A MAC = 00:60:16:05:79:4e, Peer MAC = 00:60:16:05:78:a4
F/W Version: Mgmtswitch=01.63, Powersupply=00.27, Bootblock=01.51
Enclosure ID = 1, BackplaneID 0 from IP 192.168.1.51
MgmtSwitch-A MAC = 00:60:16:05:78:ac, Peer MAC = 00:60:16:05:78:42
F/W Version: Mgmtswitch=01.63, Powersupply=00.27, Bootblock=01.51
Discovering Secondary subnet (B) ... OK
Enclosure ID = 0, BackplaneID 1 from IP 192.168.2.50
MgmtSwitch-B MAC = 00:60:16:05:78:a4, Peer MAC = 00:60:16:05:79:4e
F/W Version: Mgmtswitch=01.63, Powersupply=00.27, Bootblock=01.51
Enclosure ID = 1, BackplaneID 1 from IP 192.168.2.51
MgmtSwitch-B MAC = 00:60:16:05:78:42, Peer MAC = 00:60:16:05:78:ac
F/W Version: Mgmtswitch=01.63, Powersupply=00.27, Bootblock=01.51
# /nas/sbin/setup_enclosure -readConfig
Executing -readConfig option
Current Enclosure database (ENCL_DB) info ...
Enclosure ID# 0:
    MgmtSwitch-A IP address = 192.168.1.50
    MgmtSwitch-A MAC address = 00:60:16:05:79:4e
    F/W Version: Mgmtswitch=01.63, Powersupply=00.27, Bootblock=01.51
    MgmtSwitch-B IP address = 192.168.2.50
    MgmtSwitch-B MAC address = 00:60:16:05:78:a4
    F/W Version: Mgmtswitch=01.63, Powersupply=00.27, Bootblock=01.51
Enclosure ID# 1:
    MgmtSwitch-A IP address = 192.168.1.51
    MgmtSwitch-A MAC address = 00:60:16:05:78:ac
    F/W Version: Mgmtswitch=01.63, Powersupply=00.27, Bootblock=01.51
    MgmtSwitch-B IP address = 192.168.2.51
    MgmtSwitch-B MAC address = 00:60:16:05:78:42
    F/W Version: Mgmtswitch=01.63, Powersupply=00.27, Bootblock=01.51
/etc/nas_enclosure.map:
    Generation = 4
    md5sum = fb21780e176946966e5def25a63548f1
    Checksum verification = OK
/etc/encl_dhcpd.conf:
    Generation = 4
    md5sum = b08fb8369c5f979048af002ef38bc108
    Checksum verification = OK
```

## **NSX X-BLADE 60 DATA MOVER:**

- Consists of CPU Board, a Fibre I/O Module, & GbE I/O Module
  - Data Mover ‘Blades’ are 64-bit capable, with 16GB memory, but code won’t be available to utilize this until end of FY06
  - Only 4GB memory currently supported
  - 3.4/3.8GHz P4 processor with HT (Hyper-Threading) Technology, 800MHz FSB
- Data Mover comes with (2) I/O Modules, one for BackEnd [4-FC ports] and one for FrontEnd [2-Fibre GbE & 6-Copper GbE ports]  
(4) – (8) Data Movers with 2-DM blades for each enclosure, 2-Mgement switches, 2-Pwr Supplies, & 2-Fans each  
Each DM blade will be able to support up to 16TB of storage, for total of 112TB per Cabinet (typo? 128TB?)

**Note:** No Serial connection presently built-into the data mover, but access to DM Console can be done via USB Serial port & cable from respective blade Management Module—requires special USB-to-Serial Cable #038-003-157. NSX will support 4-8 Blades (DM's). Connect to upper middle USB port with USB-to-Serial cable.

**DATA MOVER BLADES USE SFP (Small Form Pluggable) FIBRE CONNECTIONS→NO MIAs**

**NSX BLADE ENCLOSURE COMPONENTS: (2) Blades, (2) Mgmt Switches, (2) Pwr Supplies, (4) Fans, (2) I/O Modules per Blade**

LH side Management Switch B & Power Supply B, (2) Fibre I/O Modules with 4-ports per DM, DM A & B Modules located behind Ethernet/Fibre Boards, (2) Ethernet Modules with 2-Fibre GigE ports and 6-Copper GigE ports per DM, RH side Management Switch A & Power Supply A

**DM Fibre I/O Module {Tsunami}: →Tape Devices and/or Storage**

Left-to-Right, BE0, BE1, AUX0, AUX1

Top Module, with (4) @GB Fibre ports that require connection to FC Switch

Aux ports connect to tape devices

BE0 ports connect to Clariion and/or Symmetrix Storage

**DM GbE I/O Module {Blizzard}: →Customer Networks**

Bottom Module with (2) Fibre 2GB GbE ports, left-to-right, fge0 & fge1--  
and (6) Copper 10/100/1000 Ports top l-to-r cge0, cge1, cge2; btm l-to-r cge3, cge4, cge5

**Note:** Troubleshoot the fge interfaces using “bcm” in .server\_config commands

**DM BLADE REPLACEMENT:**

→Remove I/O Modules, Remove & replace Blade, Re-install I/O Modules, run setup\_slot -i 2 [Blade ID]

**ADDING DM BLADE:**

→Install blade, cable to Ethernet & Fibre

→Run setup\_slot and follow script

→Add WWN's into Storage Groups

→Configure Zoning to match other Blades

→Complete setup\_slot instructions

**NSX CHIVAS CONTROL STATION (100-560-688):**

--No quad Serial port between CS and DMs [Management & monitoring now conducted via Management Switch modules]

--No attached (8) port Ethernet Switch--now built into CS

--2GB RAM for CS

--Dual CS configurations standard, with crossover IPMI cable connecting CS0 to CS1

**CS Ports:**

eth3 = External network; eth0 = Internal network to Management Switch A Enclosure 0;

Gb1 = IPMI interface to eth1 of alternate CS;

Gb2 = Internal Network to Management Switch B Enclosure 0;

Com1 = (1) Serial Modem for CallHome

Video Port connection

Green power LED and Yellow Fault LED & Blue enclosure LEDs

**ATS (Automatic Transfer Switch) UNITS:**

Each CS has its own ATS backup power supply to deal with short-term power loss, and each ATS has a power path to both UPS units

**CONTROL STATION REPLACEMENT:**

→Replace failed CS and load same NAS code using CS recovery option [restores CS from backend backup]

→Run rebuild script to integrate to Mgt Network

**#nas/sbin/setup\_enclosure -rebuildConfig**

**Note:** Follow CS recovery procedures from Avatar

**NSX MANAGEMENT SWITCHES {Typhoon}:**

(1) Switches per Enclosure, Switch A is primary and on the right, Switch B is secondary and on the left

→Each Switch contains (3)Broadcom 5325M Ethernet ports, Motorola Micro-Controller, Firmware, I2C Bus connections,

Interconnection to peer module through backplane on Enclosure, LED Indicators for Enclosure ID, (2) USB Serial ports for Blade console connection and debugging

**Note:** Each Enclosure contains two data movers, as well as two management switches. Observe Data Mover bootup from special DB9-to-USB serial cable (is actually a Serial port) called Serial Port 1, inside usb-style port on switch, with right-hand switch A for data mover 2 and left-hand switch for data mover 3, etc. Upper leftmost USB port is used for connection.

**RS232 Ports 1 & 2:**

Port 1, or upper left USB-style RS-232 serial port is used to connect to Laptop to observe and troubleshoot Data Mover—the other Port is not used—only manufacturing.

**Note:** Laptop connection to Data Mover requires special db9-pin-to-USB converter. Hyperterm session uses 9600 8 N 1 None.

### **LED INDICATORS:**

Amber LED for Fault

Green LED for Power

LED's 0-3 are Blue to indicate Enclosure ID

### **RIGHT SIDE MANAGEMENT SWITCH A CABLING ENCL 0:**

Switch consists of (3) ports on top, (2) ports on bottom, amber Fault, green Power and blue Enclosure ID LEDs

→Top left Port 0 Ethernet—Connects from Switch A to Enclosure 1's Port 3 on Switch A

→Top right are two USB-style RS-232 Serial ports—only the leftmost USB port is used, for connecting Laptop to Data Mover to observe bootup process

→Lower left Port 3 Ethernet—Connects from Switch A (Enclosure 0) to 10/100 Port on CS0

→Lower right Port 4 Ethernet—Connects to 10/100 Port on CS1

### **LEFT SIDE MANAGEMENT SWITCH B CABLING ENCL 0:**

→Top left Port 0 Ethernet—Connects from Switch B to Enclosure 1's Port 3 on Switch B

→Top right are two USB-style RS-232 Serial ports—only the leftmost USB port is used, for connecting Laptop to Data Mover to observe bootup process

→Lower left Port 3 Ethernet—Connects from Switch B to Gb#2 Port on CS0

→Lower right Port 4 Ethernet—Connects from Switch B to Gb#2 Port on CS1

### **ENCLOSURE 1 SWITCH A:**

→Port 4 connects to Ethernet Port on UPS 0

→Port 3 already mentioned above, connects to Port 0 on Encl 0 Switch A

### **ENCLOSURE 1 SWITCH B:**

→Port 4 connects to Ethernet Port on UPS 1

→Port 3 already mentioned above, connects to Port 0 on Encl 0 Switch B

### **CROSSOVER IPMI CABLE BETWEEN CONTROL STATIONS:**

Gb#1 on CS0 connects to Gb#1 on CS1

### **DEFAULT INTERNAL MANAGEMENT NETWORK NSX:**

|                          | Primary Network | Secondary Network                  |
|--------------------------|-----------------|------------------------------------|
| Enclosure 0 Mgmt Switch: | 192.168.1.50    | 192.168.2.50                       |
| Enclosure 1 Mgmt Switch: | 192.168.1.51    | 192.168.2.51                       |
| Enclosure 2 Mgmt Switch: | 192.168.1.52    | 192.168.2.52                       |
| Enclosure 3 Mgmt Switch: | 192.168.1.53    | 192.168.2.53                       |
| CS0                      | 192.168.1.100   | 192.168.2.100                      |
| CS1                      | 192.168.1.101   | 192.168.2.101                      |
| Blade 2                  | 192.168.1.2     | 192.168.2.2                        |
| Blade 3                  | 192.168.1.3     | 192.168.2.3 [etc., for Blades 4-9] |
| UPS0 Right Side          | 192.168.1.90    | 192.168.2.90                       |
| UPS1 Left Side           | 192.168.1.91    | 192.168.2.91                       |

### **CELLERRA NSX MANAGEMENT SWITCHES:**

Control Network IP's are preconfigured at install from DHCP Server running on CS0

Management Switches are required to detect, query, and reboot DMs

### **INTERNAL MANAGEMENT IP ADDRESSES:**

Enclosure 0 Management Switch A 192.168.1.50 mgmt\_2\_3

Enclosure 0 Management Switch B 192.168.2.50 mgmt\_2\_3b

Enclosure 1 Management Switch A 192.168.1.51 mgmt\_4\_5

Enclosure 1 Management Switch B 192.168.2.51 mgmt\_4\_5b

Enclosure 2 Management Switch A 192.168.1.52 mgmt\_6\_7

Enclosure 2 Management Switch B 192.168.2.52 mgmt\_6\_7b

Enclosure 3 Management Switch A 192.168.1.53 mgmt\_8\_9

Enclosure 3 Management Switch B 192.168.2.53 mgmt\_8\_9b

### **TROUBLESHOOTING MANAGEMENT SWITCH ISSUES:**

#### **# /nas/sbin/t2net\_test**

Usage:

t2net\_test

-Discover {subnet}

| -Getreason {slotID}

| -Getstatusflags {slotID}

| -Getsyscode {slotID operand}

| -GetmgmtswitchMAC {enclosureID}

```
| -Readfwversion {ipaddr}  
| -ReadallenclosureByIP {ipaddr}  
| -Readallenclosure {enclosureID}  
| -StatusCodeList
```

### # /nas/sbin/t2net\_test -Discover 0

```
t2net_test: Enclosure ID = 0, BackplaneID 0 from IP 192.168.1.50  
t2net_test: Mgmtswitch MAC A = 00:60:16:0a:f3:94 MAC B = 00:60:16:0a:f3:8c  
t2net_test: Enclosure ID = 1, BackplaneID 0 from IP 192.168.1.51  
t2net_test: Mgmtswitch MAC A = 00:60:16:0a:f3:b9 MAC B = 00:60:16:0a:f6:ed
```

### # /nas/sbin/t2net\_test -Discover 1

```
t2net_test: Enclosure ID = 0, BackplaneID 1 from IP 192.168.2.50  
t2net_test: Mgmtswitch MAC A = 00:60:16:0a:f3:94 MAC B = 00:60:16:0a:f3:8c  
t2net_test: Enclosure ID = 1, BackplaneID 1 from IP 192.168.2.51  
t2net_test: Mgmtswitch MAC A = 00:60:16:0a:f3:b9 MAC B = 00:60:16:0a:f6:ed
```

### # ./t2net\_test -StatusCodeList

T2NET status code list:

```
(0 = OK) - Task completion  
(-1 = UNKNOWN) - Unknown error  
(-2 = ENOBUF) - Require buf is empty  
(-3 = ENOMEM) - Cannot allocate memory  
(-4 = ERANGECHECK) - Invalid Range given  
(-5 = ECOMM) - Network I/O error  
(-6 = ESOCK) - Network socket error  
(-7 = EIPADDR) - Invalid IP address  
(-8 = ESLOTID) - Invalid Slot ID  
(-9 = ESUBNET) - Invalid subnet  
(-10 = EACCESS) - System access error  
(-11 = ETIMEOUT) - Network timeout  
(-12 = EUSAGE) - Invalid API usage  
(-13 = EBUSY) - System busy  
(-14 = EOPERAND) - Invalid operand given  
(-15 = EBLADE) - Invalid Blade ID  
(-16 = EENCLOSUREID) - Invalid Enclosure ID  
(-17 = EFILENAME) - Bad filename  
(-18 = EI2CBUS) - Mgmt I2CBUS error  
(-19 = EPOST) - System Post error  
(-20 = EDISCOVER) - Discovery stop  
(-21 = ERECEIVENAK) - Mgmt NAK Error  
(-22 = NAK_PEER_NOT_PRESENT) - Mgmt Peer not present  
(-23 = NAK_PEER_NOT_RESPONDING) - Mgmt Peer not respond  
(-24 = NAK_ILLEGAL_CMD) - Mgmt illegal command  
(-25 = NAK_ILLEGAL_OPERAND) - Mgmt illegal operand  
(-26 = NAK_ILLEGAL_ENCL_ID) - Mgmt illegal encl ID  
(-27 = NAK_UNABLE_TO_EXEC_CMD) - Mgmt unable to run cmd  
(-28 = NAK_ILLEGAL_RMCP_MSG) - Mgmt illegal RMCP msg  
(-29 = NAK_ILLEGAL_TARGET_BLADE) - Mgmt illegal target  
(-30 = EFILEACCESS) - File access error
```

### # ./t2net\_test -Getstatusflags 5

t2net\_test: Status Flags for Slot 5:

```
[0]=0x32 [1]=0x90 [2]=0x00 [3]=0x00 [4]=0x7F [5]=0xF8 [6]=0x00 [7]=0x00  
[8]=0x00 [9]=0x00 [10]=0x00 [11]=0x00 [12]=0x40 [13]=0x00 [14]=0x00 [15]=0x00  
[16]=0x00 [17]=0x00 [18]=0x04 [19]=0x00 [20]=0x00 [21]=0x00 [22]=0x00 [23]=0x00  
[24]=0x00 [25]=0x00 [26]=0x00 [27]=0x00 [28]=0x00 [29]=0x00 [30]=0x08 [31]=0x00
```

### # ./t2net\_test -Getreason 5

Reason Code for slot 5 = 5

### # ./t2net\_test -GetmgmtswitchMAC 0

Mgmtswitch MAC A=00:60:16:05:79:4E, MAC B=00:60:16:05:78:A4

### # ./t2net\_test -Readfwversion 192.168.1.50

t2net\_test: Mgmtswitch FW=01.63, Powersupply FW=00.27, Bootblock FW=01.51 (Management switch firmware info)

# ./t2net\_test -ReadallenclosureByIP 192.168.1.50 (where .50 = management switch IP)

t2net\_test: Blade A: DataMover PRESENT

t2net\_test: BladeA Status Flags:

```
[0]=0x32 [1]=0x90 [2]=0x00 [3]=0x00 [4]=0x0F [5]=0xFF [6]=0x00 [7]=0x00
[8]=0x00 [9]=0x00 [10]=0x00 [11]=0x00 [12]=0x80 [13]=0x00 [14]=0x00 [15]=0x00
[16]=0x00 [17]=0x00 [18]=0x00 [19]=0x00 [20]=0x00 [21]=0x00 [22]=0x00 [23]=0x00
[24]=0x00 [25]=0x00 [26]=0x00 [27]=0x00 [28]=0x00 [29]=0x00 [30]=0x00 [31]=0x00
```

t2net\_test: SlotID=2 EncID=0 BackPlaneID=0

POST=0x61 ReasonCode=0x05

BladeCode=0x2C MiddleByteforPOST=0x80

I2C\_PS0\_Err\_DevAddr=0x00 I2C\_PS1\_Err\_DevAddr=0x00 I2C\_DevErr=0x00

MAC=00:60:16:05:79:4E

t2net\_test: Blade B: DataMover PRESENT

t2net\_test: BladeB Status Flags:

```
[0]=0x32 [1]=0x90 [2]=0x00 [3]=0x00 [4]=0x0F [5]=0xFF [6]=0x00 [7]=0x00
[8]=0x00 [9]=0x00 [10]=0x00 [11]=0x00 [12]=0x40 [13]=0x00 [14]=0x00 [15]=0x00
[16]=0x00 [17]=0x00 [18]=0x00 [19]=0x00 [20]=0x00 [21]=0x00 [22]=0x00 [23]=0x00
[24]=0x00 [25]=0x00 [26]=0x00 [27]=0x00 [28]=0x00 [29]=0x00 [30]=0x00 [31]=0x00
```

t2net\_test: SlotID=3 EncID=0 BackPlaneID=1

POST=0x61 ReasonCode=0x05

BladeCode=0x2C MiddleByteforPOST=0x80

I2C\_PS0\_Err\_DevAddr=0x00 I2C\_PS1\_Err\_DevAddr=0x00 I2C\_DevErr=0x00

MAC=00:60:16:05:78:A4

**#./t2net\_test -SetResumeCksum 2 34** (Used to set Resume checksum on slot\_2 for Blade A)

**# ./t2led**

t2led - Version 5.4 - 12/08/04

controls a given NSX enclosure LEDs

usage: t2led enclosure led state

    enclosure = 0, 1, 2 or 3

    led = 0 (blade A)

        1 (blade B)

        2 (I/O Annex A)

        3 (I/O Annex B)

        4 (Mgmt Switch B)

        5 (Mgmt Switch A)

        6 (Enclosure)

    state = 0 (off)

        1 (on)

        2 (flashing)

**# /nas/sbin/t2vpd -s 4**

Hammerhead T2VPD utility - BoxMask 0x10 Enclosure -1

SLOT\_4 RESUME\_INFORMATION布莱德:

```
EMC_PART_NUMBER="100-560-538"
EMC_ARTWORK_REVISION=""
EMC_ASSEMBLY_REVISION="A05"
EMC_SERIAL_NUMBER="LKE00051903985"
VENDOR_NAME=""
LOCATION_OF_MANUFACTURE="Apex, NC USA"
YEAR_OF_MANUFACTURE="2005"
MONTH_OF_MANUFACTURE="5"
DAY_OF_MONTH_OF_MANUFACTURE="17"
ASSEMBLY_NAME="HH NAS WF3 DM W(8) 512MB DIMMS"
```

SLOT\_4 RESUME\_INFORMATION\_IO\_MODULE\_0:

```
EMC_PART_NUMBER="100-560-177"
EMC_ARTWORK_REVISION=""
EMC_ASSEMBLY_REVISION="A08"
EMC_SERIAL_NUMBER="LKE00050401436"
VENDOR_NAME=""
LOCATION_OF_MANUFACTURE=""
YEAR_OF_MANUFACTURE=""
MONTH_OF_MANUFACTURE=""
DAY_OF_MONTH_OF_MANUFACTURE=""
ASSEMBLY_NAME="HH NAS 2GB FC I/O MODULE"
```

SLOT\_4 RESUME\_INFORMATION\_IO\_MODULE\_1:

```
EMC_PART_NUMBER="100-560-178"
EMC_ARTWORK_REVISION=""
EMC_ASSEMBLY_REVISION="A08"
EMC_SERIAL_NUMBER="LKE00050300939"
VENDOR_NAME=""
LOCATION_OF_MANUFACTURE=""
YEAR_OF_MANUFACTURE=""
MONTH_OF_MANUFACTURE=""
DAY_OF_MONTH_OF_MANUFACTURE=""
```

ASSEMBLY\_NAME="HH NAS GBE I/O MODULE"  
SLOT\_4 RESUME\_INFORMATION\_IO\_MODULE\_0\_SFP\_MODULES

RESUME\_INFORMATION\_SFP1:

IDENTIFIER="3"  
CONNECTOR="7"  
VENDOR\_NAME="E2O COMMS INC "  
VENDOR\_PART\_NUMBER="EMA2G-LD3TA-MT "  
VENDOR\_REV="2 "  
VENDOR\_SERIAL\_NUMBER="4B9S509 "

RESUME\_INFORMATION\_SFP2:

IDENTIFIER="3"  
CONNECTOR="7"  
VENDOR\_NAME="E2O COMMS INC "  
VENDOR\_PART\_NUMBER="EMA2G-LD3TA-MT "  
VENDOR\_REV="2 "  
VENDOR\_SERIAL\_NUMBER="4B9S510 "

RESUME\_INFORMATION\_SFP3:

IDENTIFIER="3"  
CONNECTOR="7"  
VENDOR\_NAME="E2O COMMS INC "  
VENDOR\_PART\_NUMBER="EMA2G-LD3TA-MT "  
VENDOR\_REV="2 "  
VENDOR\_SERIAL\_NUMBER="4B9S502 "

RESUME\_INFORMATION\_SFP4:

IDENTIFIER="3"  
CONNECTOR="7"  
VENDOR\_NAME="E2O COMMS INC "  
VENDOR\_PART\_NUMBER="EMA2G-LD3TA-MT "  
VENDOR\_REV="2 "  
VENDOR\_SERIAL\_NUMBER="4B9S501 "

SLOT\_4 RESUME\_INFORMATION\_IO\_MODULE\_1\_SFP\_MODULES

RESUME\_INFORMATION\_SFP1:

IDENTIFIER="3"  
CONNECTOR="7"  
VENDOR\_NAME="E2O COMMS INC "  
VENDOR\_PART\_NUMBER="EMA2G-LD3TA-MT "  
VENDOR\_REV="2 "  
VENDOR\_SERIAL\_NUMBER="49PP680 "

RESUME\_INFORMATION\_SFP2:

IDENTIFIER="3"  
CONNECTOR="7"  
VENDOR\_NAME="E2O COMMS INC "  
VENDOR\_PART\_NUMBER="EMA2G-LD3TA-MT "  
VENDOR\_REV="2 "  
VENDOR\_SERIAL\_NUMBER="49PP679 "

## **USING T2NET TEST TO POWER UP A BLADE:**

**# /nas/sbin/t2net\_test -Bladecontrol**

Usage: t2net\_test -Bladecontrol {slotID operand}

where:- slotID = 2 to 9

operand = 0(RST),  
1(Hold in RST),  
2(Release from RST),  
3(RST+Hold in POST),  
4(RST+Hold in POST+PXEBoot),  
8(Power Off),  
9(Power On),  
11(Power On+Hold in POST),  
12(Power On+Hold in POST+PXEBoot)

**# /nas/sbin/t2net\_test -Bladecontrol 2 9**

**Note:** Similar to t2reset pwron -s 2

## **EXAMPLE OF MANAGEMENT SWITCH ISSUE:** Problem was cabling

1. cannot ping data movers or CS
2. cannot see contents of /nbsnas or /nas
3. # ./setup\_enclosure -checkCable

Executing -checkCable option

System discovery on both subnets .....

Discover 0 Mgmtswitches in Primary subnet A, 0 Mgmtswitches in Secondary subnet B

Cannot find any Mgmtswitch in this system

ZERO Mgmtswitch detection may due to one or more of the following cases:

- Miswired cabling on Mgmtswitches that causes switch looping
- Loose/bad cable between Enclosure ID# 0 and CS
- Bad Ethernet port on CS
- Bad Mgmtswitch

Error: checkCable FAIL (EDISCOVER)

Error: CABLECHECK\_CMD retval = -4 (EDISCOVER)

4. Getreason -e returns error code 25:

**# /nas/sbin/getreason -e**

- 10 - slot\_0 primary control station
- 25 - slot\_2 management switch error
- 25 - slot\_3 management switch error
- 25 - slot\_4 management switch error
- 25 - slot\_5 management switch error

**Note:** Came across this in the lab for 5.6.46 CS replacement exercise--needed to run -initSystem force to get internal switches back

#### **Management Switch Failure/Replacement:**

1. Identify Failing Module:

**# /nas/sbin/setup\_enclosure -checkSystem**

2. After replacing new Switch, reconfigure:

**# /nas/sbin/setup\_enclosure -replaceMgmtswitch enclosure-id**

3. Check Hardware:

**# /nas/sbin/setup\_enclosure -checkSystem**

1. Verify Cabling:

**# /nas/sbin/setup\_enclosure**

### **CELERRA PLATFORM MANAGEMENT SWITCH TYPES:**

#### **Scorpion Switch:**

→Used in Sledgehammer systems (NS20 & NS40), mgmtfirmware\_sc.s19; bootblock\_sc.s19 (aka sh.s19)

#### **Typhoon Switch:**

→Used in Hammerhead systems (NS80, NSX), mgmtfirmware\_ty.s19; bootblock\_ty.s19 (aka hh.s19)

#### **Earthquake Switch:**

→Used in Foxglove systems (NS960, NS-G8), mgmtfirmware\_eq.s19; bootblock\_eq.s19 (aka fg.s19)

#### **Nimitz Switch:**

→Used in Argonaut systems (VG2/VG8), mgmtfirmware\_nm.s19; bootblock\_nm.s19 (aka ag.s19)

### **NSX APC UPS POWER SUPPLIES:** located at bottom of Titan cabinet

UPS0 connected to Management Switch B and Power Supply B on all DM enclosures, as well as to each CS

UPS1 connected to Management Switch A & Power Supply A on all enclosures, as well as to each CS

Both UPS's connect to Management Switches on Enclosure 1 via 10/100 Ethernet

Serial connection for direct connect to workstation, and initial setup

**Note:** ACTS (AC Transfer Switch) provides AC pwr from both UPS devices to CS in the event of a power failure. Lower ATS0 is for CS0, and ATS1 is for CS1. UPS's maintain power directly to Data Movers (BEs).

→Connect to UPS's via special UPS serial cable to laptop or via Telnet after initial setup to IP 192.168.1.90 for UPS 0 and 192.168.2.90 for UPS 1

### **PROBLEM WITH POWER LOSS TO UPS UNITS emc154409:**

1. Power loss to one of the two UPS legs:

#### **sys\_log:**

UPSMonitor 12 The UPS has been switched off by a management station →CallHome is immediate

UPSMonitor:4:13 emcnasUPS\_i1: UPS: Entered sleep mode.

**BoxMonitor:2:519** enclosure 1 power supply B fault →BoxMonitor is the Facility, Severity 2 indicates Critical, 519 is the EventID  
CallHome:6:200 Successful error file transfer via email

**# nas\_event -I -a callhome**

UPSMonitor 3 The UPS has failed its internal self-test

UPSMonitor 12 The UPS has been switched off by a management station

UPSMonitor 17 The UPS batteries require immediate replacement

UPSMonitor 18 An Environment contact closure has faulted

UPSMonitor 20 The UPS is on bypass due to an internal fault

UPSMonitor 24 The base module bypass power supply needs repair

```
UPSMonitor 25 The base module fan needs repair
UPSMonitor 51 The battery charger has failed
UPSMonitor 53 The battery temperature threshold has been violated
UPSMonitor 77 An abnormal condition has been detected
# nas_event -l -f -i |grep UPS
id facility
140 UPSMonitor
# nas_event -l -f UPSMonitor
id description
12 The UPS has been switched off by a management station
# cat /nas/site/cron.d/nas_sys|grep apc
0-59/5 * * * * nasadmin /nas/sbin/get_apc_ups_status /nas/log/ups_status.xml >/dev/null 2>&1
1 0-23/2 * * * nasadmin /nas/sbin/get_apc_ups_status -resume /nas/log/ups_resume.xml >/dev/null 2>&1
# cat /nas/sys/nas_eventlog.cfg
UPS MOnitor
#
facilitypolicy 140, 7
    disposition range=0-1000, logfile "/nas/log/sys_log"
    disposition range=3-3 severity=1-2, callhome immediate
    disposition range=12-12 severity=1-2, callhome immediate
    disposition range=17-18 severity=1-2, callhome immediate
    disposition range=20-20 severity=1-2, callhome immediate
    disposition range=24-25 severity=1-2, callhome immediate
    disposition range=51-51 severity=1-2, callhome immediate
    disposition range=53-53 severity=1-2, callhome immediate
    disposition range=77-77 severity=1-2, callhome immediate
    disposition range=0-1000 severity=1-4, trap "/nas/site/trap.cfg 2"
```

2. Trap sent by UPS unit to Control Station creates .acloss0 or .acloss1 (UPS 0 and UPS 1, respectively) in /nas/log directory:

```
# cat /nas/log.acloss1
```

100

3. Current problem is that a restore to the UPS power leg fails to remove the .acloss\* file due to missing trap action on CS

**Note:** During proper operation, when A/C power is lost to either power leg for the UPS units (there are two UPS Battery Backup supplies for an NSX system), an SNMP trap is sent from the UPS unit to the Celerra. The Celerra will log an event in the sys\_log, generate a CallHome, and write a temporary .acloss0 or .acloss1 file to the /nas/log directory. Later, when power is restored, an SNMP trap is generated and sent to the Control Station, which is then supposed to remove the /nas/log/.acloss0 (or .acloss1 for UPS leg 1) file as part of its recovery and cleanup from the event. With NAS 5.4/5.5, this recovery does not happen and the .acloss file is not removed, meaning that if the other UPS leg subsequently lost A/C power for any reason, the Celerra would think that both UPS units were faulted and power down the Celerra.

4. Workaround:

a. Using a vi editor, add the following entry to the /nas/sys/snmptrapd.conf file:

```
traphandle .iso.3.6.1.4.1.318.0.8 /nas/sbin/apc_ups_trap ACRESTORE
```

The new /nas/sys/snmptrapd.conf file will look like this:

```
# cat /nas/sys/snmptrapd.conf
```

```
traphandle .iso.3.6.1.4.1.318.0.6 /nas/sbin/apc_ups_trap WARNING
```

```
traphandle .iso.3.6.1.4.1.318.0.8 /nas/sbin/apc_ups_trap ACRESTORE
```

```
traphandle .iso.3.6.1.4.1.318.0.12 /nas/sbin/apc_ups_trap SEVERE
```

b. Force the snmptrapd process to re-read the newly edited snmptrapd.conf file using HUP:

```
# ps -ef |grep snmp
```

```
root 17201 1 0 Mar06 ? 00:00:00 /usr/sbin/snmptrapd -c /nas/sys/
```

```
# kill -HUP 17201
```

**Note:** Issue the HUP against the process ID for the snmptrapd daemon—notice that the PID does not change since the process is not actually killed—the HUP forces the snmptrapd.conf file to be re-read with the new configuration

c. Check for and remove any .acloss0 or .acloss1 files from the /nas/log & /nbsnas/log directories

```
# ls -la /nas/log/.acloss*
```

```
-rw-r--r-- 1 root root 0 Mar 12 07:40 /nas/log/.acloss0
```

```
# rm -f /nas/log/.acloss*
```

## **UPS APC FILES, COMMANDS:**

**apc\_ups\_cfg (to config.snmp & traps); apc\_ups\_cfg.exp (config.snmp mib & traps to CS); apc\_ups\_cron\_setup (cron jobs); apc\_ups\_trap (handles traps received from UPS unit and logs toCS eventlog)**

# telnet 192.168.1.90

Trying 192.168.1.90...

Connected to 192.168.1.90.

Escape character is '^]'.  
User Name : apc  
Password : \*\*\*

American Power Conversion Network Management Card AOS v2.6.4  
(c) Copyright 2004 All Rights Reserved Smart-UPS & Matrix-UPS APP v2.6.0

---

Name : Unknown Date : 08/10/2005  
Contact : Unknown Time : 16:51:33  
Location : Unknown User : Administrator  
Up Time : 2 Days 6 Hours 5 Minutes Stat : P+ N+ A+  
Environment : Thresholds Disabled, Contact Alarms Disabled, Relay Disabled  
Smart-UPS RT 7500 RM XL named UPS 0 : On Line, No Alarms Present

----- Control Console -----

- 1- Device Manager
  - 2- Network
  - 3- System
  - 4- Logout
- <ESC>- Main Menu, <ENTER>- Refresh, <CTRL-L>- Event Log> 3

----- System -----

- 1- User Manager
  - 2- Identification
  - 3- Date/Time
  - 4- Tools
  - 5- RADIUS
  - 6- Modem
  - 7- About System
- <ESC>- Back, <ENTER>- Refresh, <CTRL-L>- Event Log> 7

---

About System

Model Number : AP9619  
Serial Number : JA0430000550  
Manufacture Date : 07/20/2004  
Hardware Revision : A10  
MAC Address : 00 C0 B7 6F 48 B1  
Flash Type : AMD A29DL322DB  
Press <ENTER> to continue...

---

Module Information

Description : Smart-UPS & Matrix-UPS APP

---

Name : sumx Type : StatApp  
Version : 260 Sector : 12  
Date : 11/24/2004 Time : 10:04:23  
CRC16 : A69E

Press <ENTER> to continue...

Description : Network Management Card AOS

---

Name : aos Type : APC OS  
Version : 264 Sector : 51  
Date : 11/12/2004 Time : 19:40:33  
CRC16 : C476

Press <ENTER> to continue...

----- System -----

- 1- User Manager
- 2- Identification
- 3- Date/Time
- 4- Tools

5- RADIUS

6- Modem

7- About System

<ESC>- Back, <ENTER>- Refresh, <CTRL-L>- Event Log> 2

----- Network -----

1- TCP/IP

2- DNS

3- Ping Utility

4- FTP Server

5- Telnet/SSH

6- Web/SSL/TLS

7- WAP

8- SNMP

9- Email

10- Syslog

11- Paging

<ESC>- Back, <ENTER>- Refresh, <CTRL-L>- Event Log> 1

----- TCP/IP -----

Network started. Acquired a DHCP lease on 08/08/2005 at 10:46.

System IP : 192.168.1.90 MAC Address : 00 C0 B7 6F 48 B1

Subnet Mask : 255.255.255.0 DHCP Server : 192.168.1.100

Default Gateway : 0.0.0.0 Lease Expires : 09/12/2006 10:46

Host Name : APC

Domain Name : somedomain.com

1- Boot Mode : DHCP only

2- Advanced...

<ESC>- Back, <ENTER>- Refresh, <CTRL-L>- Event Log

## **MISCELLANEOUS NSX SERIES COMMANDS:**

# **./t2485net**

|           |           |            |            |         |
|-----------|-----------|------------|------------|---------|
| Device    | Device    | Device     | Device     | Network |
| Operating | On        | Responding | Responding | Bus     |
| From      | Poll List | Network A  | Network B  | Master  |

-----

CS 0 BOOT SLOW MASTER MASTER B UNLOCKED

CS 1 Not detected by SMB

DM 2 MAIN FAST YES YES

DM 3 MAIN FAST NO YES

DM 4 MAIN FAST YES YES

DM 5 MAIN FAST YES YES

DM 6 Not detected by SMB

**# ./t2tty →NSX sends communications between DM and CS using RPC over the internal net**

Celerra\IP DM Command Utility....

Usage:

Force Data Mover in slot # to PXE boot

t2tty -p 2

Check IPMI connection to other CS

t2tty -m hostNameOrIP

Send dart command "cmd" to data mover in slot # over RPC Interface

t2tty -c # "cmd" or t2tty -C # "cmd"

for example: t2tty -c 2 "ifconfig"

**# ./t2tty -c 4 "ifconfig"**

Devices:

mge0 dmtu=9000, dmac=0:60:16:5:36:e9

mge1 dmtu=9000, dmac=0:60:16:5:36:e8

loop dmtu=65536, dmac=0:0:0:0:0:0

Interfaces:(3)

e130 on mge0 l=192.168.1.4 n=255.255.255.0 b=192.168.1.255 DNIF UP

```
mtu=1500, dmtu=9000, vlid=0, mac=0:60:16:5:36:e9 dmac=0:60:16:5:36:e9
e131 on mge1 l=192.168.2.4 n=255.255.255.0 b=192.168.2.255 DNIF UP
    mtu=1500, dmtu=9000, vlid=0, mac=0:60:16:5:36:e8 dmac=0:60:16:5:36:e8
loop on loop l=127.0.0.1 n=255.0.0.0 b=127.255.255.255 DNIF UP
    mtu=32768, dmtu=65536, vlid=0, mac=0:0:0:0:0:0 dmac=0:0:0:0:0:0
1123687593: ADMIN: 4: Command succeeded: ifconfig
```

### # ./t2tty -c 4 "fcp bind show"

\*\*\* Persistent Binding Table \*\*\*

```
Chain 0000: WWN 50060160106007bb HBA 0 SP-a0 Bound
Chain 0016: WWN 50060168106007bb HBA 0 SP-b0 Bound
Chain 0032: WWN 50060169106007bb HBA 1 SP-b1 Bound
Chain 0048: WWN 50060161106007bb HBA 1 SP-a1 Bound
Chain 0064: WWN 5006048000000000 HBA 2 N_PORT Bind Pending
Chain 0080: WWN 5006048000000000 HBA 3 N_PORT Bind Pending
Existing CRC: c15e8961, Actual: c15e8961, CRC Matchs
```

\*\*\* Dynamic Binding Table \*\*\*

```
Chain 0000: WWN 50060160106007bb HBA 0 ID 0 Inx 00:00 Pid 0000 S_ID 031c00 Sys
Chain 0016: WWN 50060168106007bb HBA 0 ID 0 Inx 01:01 Pid 0016 S_ID 031f00 Non
Chain 0032: WWN 50060169106007bb HBA 1 ID 1 Inx 02:01 Pid 0032 S_ID 031e00 Non
Chain 0048: WWN 50060161106007bb HBA 1 ID 1 Inx 03:00 Pid 0048 S_ID 031d00 Sys
Chain 0064: WWN 0000000000000000 HBA 2 ID 2 Inx 04:81 Pid 0064 S_ID 000000 Non
Chain 0080: WWN 0000000000000000 HBA 3 ID 3 Inx 05:81 Pid 0080 S_ID 000000 Non
```

# /nas/sbin/t2tty → send console commands to DART over internal network using RPC vs. serial communications [T2net Library]

Note: Regular NS product sends the commands over the serial connection between CS and DM, which we do not have for NSX

### # /nas/sbin/t2tty -t

ttyS4 on → Means serial port is up and connected between Server and Control Station

ttyS5 on

ttyS6 on

ttyS7 on

Note: ttyS0 is used for Modem access; ttyS1 is used for Serial access

## GETREASON CODES NAS 5.4 NSX SYSTEMS:

# /nas/sbin/getreason | -e → new reason codes (getreason -e)

Note: Run getreason -e to obtain the complete error codes for the Reason Code failure

15 Blade flashing firmware, BIOS or POST, cannot be reset

16 Blade PXE booted

17 Blade Hardware fault detected

18 Memory test failure, BIOS detects memory error

19 POST test failure, general POST error (hardware issue)

20 POST NVRAM test failure, invalid NVRAM content error

21 POST invalid peer Blade type

22 POST invalid Blade part number

23 POST FC test failure, error in controller, fibre discovery, etc.

24 POST network test failure, error in Ethernet controller

25 NSX DM T2NET error, unable to get reason code due to management switch problems

## MISCELLANEOUS SECTION:

### # /nas/sbin/t2cab

3 - CELERRA IP CABINET

The Cabinet child is : 11 - a CELERRA HAMMERHEAD

# /nas/sbin/getboxmask -e [Enclosure IDs and error codes]

### # /nas/sbin/getreason -e

10 - slot\_0 primary control station

5 - slot\_2 contacted (Post Code = 0x61 ; Mid Code = 0x80 ; Blade Code = 0x2c)

5 - slot\_3 contacted (Post Code = 0x61 ; Mid Code = 0x80 ; Blade Code = 0x2c)

5 - slot\_4 contacted (Post Code = 0x61 ; Mid Code = 0x80 ; Blade Code = 0x2c)

5 - slot\_5 contacted (Post Code = 0x61 ; Mid Code = 0x80 ; Blade Code = 0x2c)

## REASON CODES:

15 DM flashing firmware, BIOS, or POST, cannot be reset

16 DM PXE booted from CS  
17 DM hardware fault  
18 Memory test failure, BIOS detected memory error  
19 Post test failure  
20 Post NVRAM failure  
21 Post invalid peer DM type  
22 Post invalid DM part number  
23 Post fibre channel test failure, error in blade fibre channel connection  
24 Post network test failure, error in Ethernet controller  
25 T2NET error, unable to get reason code due to management switch problems

**Note:** If DM stuck at reason code 17-24, use getreason -e to obtain specific POST error code using getreason -e. Lookup the 2<sup>nd</sup> hex number in the Post Code to determine error—appendix C NSX Recovery Procedures.

# **/nas/sbin/t2reset** [Use to power-off, power-on, reset any blade]

# **/nas/sbin/t2led** [Use to control enclosure LEDs]

### **PXE BOOTING NSX:**

# **./t2pxe**

Usage:

To stop pxe services: t2pxe -e

To start pxe services: t2pxe -s {-I|-R} cnt# [x.y.z] [a.b.c]

cnt#: number of data movers in the cabinet

-I : used for starting PXE interactively and/or when slot IDs are not set

-R : used for starting PXE when slot IDs are present

x.y.z and a.b.c are the primary and backup internal subnet addresses.

Subnet definitions will be extracted from /etc/nas\_device.map if not provided

To start tftp service alone: t2pxe -tftp [-setroot dir] start

To stop tftp service alone: t2pxe -tftp stop

**Note:** -tftp start also starts PXE services, and you may want to stop # t2pxe -e

### **CANNOT PXE BOOT DM WITH BOX MONITOR RUNNING:**

# **./t2tty -p 5**

Attempting to force data mover PXE boot:

Error, Can't PXE boot when BoxMointor is running.

Please retry after stopping the BoxMonitor (service nas stop)

Slot-5 Forced PXE Boot Failed.

### **PXE BOOTING CONFIGURED DATA MOVER WILL BRING TO RC 14 HARDWARE MISCONFIGURED STATE:**

# **./t2tty -p 5** -->Forcing PXE boot of configured DM brings it to RC 14 slot\_5 hardware misconfigured—boots from CS

Attempting to force data mover PXE boot:

Resetting Slot-5 with t2reset...

Waiting for DM network connectivity after PXE booting

### **PXE BOOTING DATA MOVER ON NSX:**

1. Stop NAS Services and manually remount /nbsnas /nas /nas/dos /nas/var

2. Start PXE boot services on Control Station: # **/nas/sbin/t2pxe -s -R 8** [2 4 8 dm's]

3. Force blade to PXE boot from CS: # **/nas/sbin/t2tty -p 2** [where 2 is slot number]

**Note:** Data Mover will boot to RC14 if already configured—just reboot and it will come up normally

4. Stop PXE Boot Services on CS: # **/nas/sbin/t2pxe -e**

**Note:** NSX series requires onsite DB9-to-USB connection via laptop to the Data Mover in order to watch PXE boot. Requires use of special serial cable with USB-type connector to connect from COM1 port on Laptop to middle USB connection on Management Switch A, or B (Data Mover 2 is A, DM 3 is B)

### **HYPERTERMINAL CONFIGURATIONS FOR NSX SYSTEMS:**

I. CONTROL STATION →COM1 or 2 19,200 8 N 1 Hardware Flow Control

II. DATA MOVER-----→” “ 9,600 8 N 1 None

III. UPS A & B-----→” “ 2,400 8 N 1 Hardware

### **NORMAL BOOT SEQUENCE FOR HAMMERHEAD/POST & FCCBOOT MENUS:**

roadcom NetXtrem Ethernet Boot Agent v7.0.3

Copyright(C)2003.Broadcom Corporation

All rights reserved.eon(TM) CPU 3.40GHz

Press Ctrl-S to Enter Configuration Menu ...

Copyright (c) EMC Corporation , 2005

Disk Array Subsystem Controller

Model: Hammerhead: NAS

DiagName: Extended POST

DiagRev: Rev. 01.52

Build Date: Wed Jun 15 09:12:17 2005

StartTime: 08/10/2005 21:51:23

SaSerialNo: LKE00050501550

ABAbcdefCDEabcdFabcGHabIabcJabcKabLab →Press ‘esc’ key twice and enter SHIP\_it to access POST or DB\_key to access FCCBOOT menu

## **SUCCESSFUL FCCBOOT SCAN OF LUN0 FROM BE0:**

POST 01.59

- |                     |                    |
|---------------------|--------------------|
| 1) Scan for Targets | 4) Set Port Speed  |
| 2) Enter Target ID  | 5) Set AG          |
| 3) Set Boot Target  | 6) Set Port Access |
| 0) Exit             |                    |

Enter Option : 1

Reinitializing loop

BE0 FCC SFP Inserted

Scanned Targets Found: 1

Number of Targets: 01

TARGETS FOUND

| Idx | Handle | Addr_id | WWN Port | Valid_bits |
|-----|--------|---------|----------|------------|
|     |        |         | WWN Node |            |

|  |                   |                |
|--|-------------------|----------------|
| 00 3AB25410 000000EF 50060160:41E08DF5 | 50060160:C1E08DF5 | LOG, TUR, LUN0 |
|--|-------------------|----------------|

Success

## **UNSUCCESSFUL FCCBOOT SCAN FROM DE0:**

- |                     |                    |
|---------------------|--------------------|
| 1) Scan for Targets | 4) Set Port Speed  |
| 2) Enter Target ID  | 5) Set AG          |
| 3) Set Boot Target  | 6) Set Port Access |
| 0) Exit             |                    |

Enter Option : 1

Reinitializing loop

BE0 FCC SFP Inserted

Cannot initialize the loop, status: 1

ErrorCode: 0x00000181

ErrorDesc:

FRU: All System FRUs

Device: DIAG MENU

Description: Targets are not responding to TURs Error!

Rev: 01.59

Fibre Controller: BE0 FCC

Status Returned: 0x0181

Configuration Type: 2

EndError:

ErrorTime: 09/30/2009 12:56:20

Hit any key to continue

## **BIOS/POST MENU:**

→Reboot Data Mover using t2reset or use PXE Boot procedure

→Press esc once or twice during “AabcfdefgBa.....” message

....Storage System Failure...

**Note:** The new method for accessing the POST menu is to use the “ctrl + c” keys instead of the “Esc” key. This allows POST to complete what it’s doing. If ctrl + c fails to work, however, use the “esc” key sequence.

ABCab << Stopping after POST >> DEabcdefgFGHabcdnIJK

....Storage System Failure – Contact Your Service Representative...

→enter password: SHIP\_it [accesses POST Diagnostic Menu]

Diagnostic Menu

POST 01.52

- |                            |                             |
|----------------------------|-----------------------------|
| 1) Reset Controller        | 13) Power Supply A Sub-Menu |
| 2) Enter Debugger          | 14) Power Supply B Sub-Menu |
| 3) Display Warnings/Errors | 15) System Test Sub-Menu    |
| 4) Boot OS                 | 16) Image Sub-Menu          |
| 5) MFG Boot OS             | 17) Disk Sub-Menu           |

- 6) POST Sub-Menu 18) Resume PROM Sub-Menu
- 7) Display/Change Privilege 19) DMI Log Sub-Menu
- 8) Enclosure Sub-Menu 20) Voltage Margin Sub-Menu
- 9) Motherboard Sub-Menu 21) Information Display
- 10) Memory Sub-Menu 22) ICA Sub-Menu
- 11) I/O Module 0 Sub-Menu 23) DDBS Service Sub-Menu
- 12) I/O Module 1 Sub-Menu 24) FCC Boot Sub-Menu

### **CALLHOME EVENTS:**

Fan Fault; Power Supply Fault; Management Switch Fault; Blade HW Fault; IO Module HW Fault; DM Boot Fault; DM RC Fault

### **NSX TROUBLESHOOTING:**

```
# /nasmcd/sbin/setup_enclosure -showAll | -v | -checkCable | -checkSystem (Checks connectivity to Mgmt Switches) | -rebuildConfig | -resetMgmtswitches | -flashFirmware | -dhcp stop start addDynamicIPRange | -addEnclosure | -removeEnclosure | -replaceEnclosure | -replaceMgmtswitch | t2net_test [Main script used to troubleshoot management switch issues]
```

**Note:** T2net Library is a UDP-based network relay between Management modules and Tier 2 for control and info commands

### **REPLACING FAILED MANAGEMENT SWITCH MODULE:**

```
# /nasmcd/sbin/setup_enclosure -v -checkSystem [Use this to verify system before replacing management switch]
```

```
# /nasmcd/sbin/setup_enclosure -v -replaceMgmtswitch enclosure-id [Use this to setup new replacement switch]
```

1. Identify failing module for replacement:

```
# /nas/sbin/setup_enclosure -checkSystem
```

2. Physically replace module

3. Configure new switch module:

```
# /nas/sbin/seup_enclosure -replaceMgmtswitch enclosure-id
```

4. Verify:

```
# /nas/sbin/setup_enclosure -checkSystem
```

```
# /nas/sbin/setup_enclosure -checkCable
```

### **REPLACING SINGLE BLADE NS20/NS40 SYSTEM:**

1. Log into system as nasadmin and su to root, then cd to /

2. Stop nas services

```
# /sbin/service nas stop
```

3. Manually remount NAS partitions

```
# mount /nbsnas
```

```
# mount /nas
```

```
# mount /nas/dos
```

4. Halt the Data Mover

```
# server_cpu -halt now
```

**Note:** This hangs the console session, User will need to Putty into CS again, where /nasmcd/sbin/getreason shows slot\_2 powered off

5. Physically swapped out slot\_2 with a new blade

6. Powered on the slot

```
# /nasmcd/sbin/t2reset pwron -s 2
```

7. Ran switch replacement commands for the replacement blade:

```
# /nasmcd/sbin/setup_enclosure -replaceMgmtswitch 0 [no errors, completed]
```

```
# /nasmcd/sbin/setup_enclosure -checkSystem
```

8. PXE boot the system in order to be able to successfully run the setup\_slot on the replacement blade:

```
a.) /tftpboot/bin/t2pxe -s -R 2 128.221.252 128.221.253
```

```
b.) /tftpboot/bin/t2tty -p 2
```

```
c.) /tftpboot/bin/t2pxe -e
```

```
d.) /nas/sbin/setup_slot -i 2
```

e.) Rebooted server\_2 to verify that all was still well

f.) Rebooted Control Station to verify NBS connectivity and let NAS services start normally

9. /nas/bin/nas\_checkup

### **ADDING NEW BLADE:**

1. Install and cable new blade

2. Conduct setup\_slot -i 3

**Note:** Direct connected FC Enabled models should just use -g syntax for setup\_slot to add WWNs to StorageGroup, etc.

3. Add new WWN to Storage Group

4. Configure Zoning
5. Complete setup\_slot

### **ADDING NEW ENCLOSURE:**

1. Install new enclosure and connect cabling
2. Add enclosure and verify:  
# /nas/sbin/setup\_enclosure –addEnclosure enclosure-id  
# /nas/sbin/setup\_enclosure -checkSystem  
# /nas/sbin/setup\_enclosure -checkCable

### **REPLACING NSX ENCLOSURE:**

**# /nasmcd/sbin/setup\_enclosure –v –replaceEnclosure enclosure-id** [Use to replace an Enclosure after stopping NAS services first, then update storage backend configuration with WWN values for new enclosure blades]

1. Stop NAS services
2. Remove old enclosure and replace with new hardware

**Note:** Please be aware that a special procedure may need to be run to convert the base SAN personality Enclosure to NAS personality

3. Run following to remove old ID and then generate new ID for replacement:

```
# /nas/sbin/setup_enclosure –removeEnclosure enclosure-id  
# /nas/sbin/setup_enclosure -checkSystem  
# /nas/sbin/setup_enclosure –addEnclosure enclosure-id  
# /nas/sbin/setup_enclosure -checkCable
```

4. Configure new WWNs for blades into Zoning and Storage Groups
5. Restart NAS

### **REPLACE ENCLOSURE & DM:**

1. **# /nasmcd/sbin/setup\_enclosure –v –addEnclosure**
2. Use getboxmask –e to verify that new enclosure can be seen
3. Power up DM and run setup\_slot
4. Update storage backend configuration for new DM

### **REPLACING CS:**

1. Remove old CS and insert new hardware
2. Load same NAS code version
3. During install, will ask if this is replacement CS—answer Yes and let rebuild continue
4. Rebuild management switch configuration

**# /nas/sbin/setup\_enclosure –v –rebuildConfig**

5. Verify management switch configuration

**# /nas/sbin/setup\_enclosure –v -checkSystem**

**# /nas/sbin/t2net\_test -Discover | -Getreason | -Getstatusflags | -Getsyscode | -GetmgmtswitchMAC | -Readfwversion | -ReadallenclosureByIP <internal ip addr> | -Readallenclosure | -StatusCodeList**

**# ps -wef |grep dhcp** [verify that dhcp daemon is running]

getreason code of 25 indicates management switch problem

### **COLLECTING LOGS FOR NSX HARDWARE ISSUES:**

1. General HW issues, Pwr Supply, Management Station issues, collect all sys\_logs & following outputs:  
\$tar -chf ~/sys\_logs.tar sys\_log\*; \$/nasmcd/sbin/getreason –e > ~/getreason.out; enclosure\_status –e <x> -v  
>~/enclosure\_status\_x.out; t2net\_test -ReadallenclosureByIP 192.168.1.5x > ~/t2net\_test\_1.5x.out; t2net\_test -ReadallenclosureByIP 192.168.2.5x > ~/t2net\_test\_2.5x.out; server\_log server\_x -a -s > ~/slogx.out
2. Voltage issues, Fan faults, Temp alarms—collect following info & sys\_logs:  
\$/nas/sbin/nsx\_cs\_sensors > ~/nsx\_cs\_sensors.out  
\$Scp /nas/log/local\_hw\_status\_slot\_0.xml
3. UPS Issues:  
--Collect all /nas/log/sys\_logs and /var/log/messages files  
--Test that UPS subsystem is accessible:  
\$ping emcnasUPS\_i0; \$ping emcnasUPS\_i1  
--Check /etc/hosts file if ping fails  
--Telnet to UPS subsystem:  
\$telnet emcnasUPS\_i0; login using “apc” and password “apc” → Show see UPS management console  
--Run snmpwalk against each UPS Hostname:  
\$snmpwalk -c public -m /nas/sys/powernet361.mib emcnasUPS\_i0 upsBasicOutputStatus > /emcnasUPS\_i0\_snmpwalkBasicStat.out  
--Additional snmpwalk output commands:  
\$snmpwalk -c public -m /nas/sys/powernet361.mib emcnasUPS\_i0 upsBasicStateOutputState >  
/emcnasUPS\_i0\_snmpwalkBasicStateoutputState.out

```
$snmpwalk -c public -m /nas/sys/powernet361.mib emcnasUPS_i0 upsAdvStateAbnormalConditions >
```

```
/emcnasUPS_i0_snmpwalkAdvStateAbnormalConditions.out
```

## **SAMPLE NSX CONFIGURATION FILES:**

### **# cat /etc/hosts**

```
# Do not remove the following line, or various programs
# that require network functionality will fail.
127.0.0.1      localhost.localdomain localhost
10.241.168.86    hammer2 hammer2..local
192.168.3.100    emcnas_ipmi
192.168.3.101    emcnasotherIPMICS_i3
# Internal Management Switch on Primary Network
# Internal Management Switch on Secondary Network
# UPS on Primary & Secondary Networks
10.241.168.52  A_APM00030600872 SPA # CLARiiON SP
10.241.168.53  B_APM00030600872 SPB # CLARiiON SP
# Internal DART Server Primary Network
192.168.1.1 server_1 #DART_data_mover_1
192.168.1.2 server_2 #DART_data_mover_2
192.168.1.3 server_3 #DART_data_mover_3
192.168.1.4 server_4 #DART_data_mover_4
192.168.1.5 server_5 #DART_data_mover_5
192.168.1.6 server_6 #DART_data_mover_6
192.168.1.7 server_7 #DART_data_mover_7
192.168.1.8 server_8 #DART_data_mover_8
192.168.1.9 server_9 #DART_data_mover_9
192.168.1.10 server_10 #DART_data_mover_10
192.168.1.11 server_11 #DART_data_mover_11
192.168.1.12 server_12 #DART_data_mover_12
192.168.1.13 server_13 #DART_data_mover_13
192.168.1.14 server_14 #DART_data_mover_14
192.168.1.15 server_15 #DART_data_mover_15
192.168.1.16 server_16 #DART_data_mover_16
# Internal Management Switch Primary Network
192.168.1.50 mgmt_2_3      #ENCLOSURE-0_MGMT_A
192.168.1.51 mgmt_4_5      #ENCLOSURE-1_MGMT_A
192.168.1.52 mgmt_6_7      #ENCLOSURE-2_MGMT_A
192.168.1.53 mgmt_8_9      #ENCLOSURE-3_MGMT_A
# Internal UPS Primary Network
192.168.1.90 emcnasUPS_i0    #UPS_0
192.168.1.100 emcnas_i0     #C-S_Internal_NW_0
192.168.1.101 emcnasotherCS_i0
# Internal DART Server Backup Network
192.168.2.1 server_1b #DART_data_mover_1
192.168.2.2 server_2b #DART_data_mover_2
192.168.2.3 server_3b #DART_data_mover_3
192.168.2.4 server_4b #DART_data_mover_4
192.168.2.5 server_5b #DART_data_mover_5
192.168.2.6 server_6b #DART_data_mover_6
192.168.2.7 server_7b #DART_data_mover_7
192.168.2.8 server_8b #DART_data_mover_8
192.168.2.9 server_9b #DART_data_mover_9
192.168.2.10 server_10b   #DART_data_mover_10
192.168.2.11 server_11b   #DART_data_mover_11
192.168.2.12 server_12b   #DART_data_mover_12
192.168.2.13 server_13b   #DART_data_mover_13
192.168.2.14 server_14b   #DART_data_mover_14
192.168.2.15 server_15b   #DART_data_mover_15
192.168.2.16 server_16b   #DART_data_mover_16
# Internal Management Switch Backup Network
192.168.2.50 mgmt_2_3b     #ENCLOSURE-0_MGMT_B
```

```
192.168.2.51 mgmt_4_5b      #ENCLOSURE-1_MGMT_B  
192.168.2.52 mgmt_6_7b      #ENCLOSURE-2_MGMT_B  
192.168.2.53 mgmt_8_9b      #ENCLOSURE-3_MGMT_B  
# Internal UPS Backup Network  
192.168.2.90 emcnasUPS_i1  #UPS_1  
192.168.2.100 emcnas_i1    #C-S_Internal_NW_1  
192.168.2.101 emcnasotherCS_i1
```

**# cat /etc/nas\_device.map**

```
DOSDSK=/dev/nda  
OS1DSK=/dev/hda  
OS2DSK=/dev/hda  
VERDSK=/dev/hda  
VARDSK=/dev/ndf  
NBSDSK=/dev/nde  
ENET_INTO=eth0  
ENET_IPMI=eth1  
ENET_INT1=eth2  
ENET_EXT=eth3  
ENET_SP=
```

**# ./t2tty -c 4 "nbsid list"**

```
1123689401: NBS: 4: nbs add name=1 vol=NBS1 rw=192.168.1.100:192.168.1.101:192.1  
68.2.100:192.168.2.101 share exclusive raw  
1123689401: NBS: 4: nbs add name=5 vol=NBS5 rw=192.168.1.100:192.168.1.101:192.1  
68.2.100:192.168.2.101 share exclusive raw  
1123689401: NBS: 4: nbs add name=6 vol=NBS6 rw=192.168.1.100:192.168.1.101:192.1  
68.2.100:192.168.2.101 share exclusive raw
```

**REPLACING DATA MOVER FOR GATEWAY SYSTEMS:**

**# /nas/sbin/setup\_slot -i -g 3**

**ADDING GATEWAY CLARIION BACKEND:**

**# /nas/sbin/add\_clariion -init**

**CONFIGURING INTEGRATED SYSTEM OR ADDING NEW DRIVES:**

**# /nbsnas/sbin/setup\_clariion -init APM00023801040**

```
Found CLARIION(s) APM00023801040  
Setup CLARiiON APM00023801040 storage device...  
System 192.168.1.200 is up  
System 192.168.1.201 is up  
Clariion Array: APM00023801040 Model: CX600 Memory: 2048  
Enabling cache...  
The following 5 template(s) available:  
1. CX_All_4Plus1_Raid_5  
2. CX_Standard_Raid_5  
3. CX_Standard_Raid_1  
4. User Defined  
5. None
```

```
Configuration for APM00023801040  
Please select a template in the range of 1-5 or 'q' to quit: 2  
Configuration for APM00023801040  
Please select a template in the range of 1-5 or 'q' to quit: 2  
Template: CX_Standard_Raid_5  
Summary:  
11 disk group(s) are created. 8,9,10,11,12,13,14,15,16,17,18  
2 spare(s) are created. 201,202  
Enclosure(s) 0_0,1_0,0_1,1_1,0_2 are installed in the system.  
Enclosure info:
```

---

```
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14
```

---

```
0_2: 72 72 72 72 72 72 72 72 72 72 72 72 72 72 72
```

-----abridged-----

Do you want to continue and configure as shown [yes or no]?: yes

Enclosure 0\_0.

Created disk group 8, luns 18,19

Enclosure 1\_0.

Created disk group 9, luns 20,21

Created disk group 10, luns 22,23

Created disk group 11, luns 24,25

-----abridged-----

Configuration completed!

Setup of CLARiON APM00023801040 storage device complete.

Discovering storage (may take several minutes)

### ***NS500 INTEGRATED & GATEWAY CELEERRA (Exterminator):*** [mid NAS 5.3 product]

→GA Sep 10, 2004, as low cost NS600 replacement with Gateway (SAN) & Integrated (Captive) versions

→Captive NS500 backend is Tarpon-Aux [DAE2 enclosure replacing LCC cards] modified CX500 array

→Clarion Tarpon technology [1.6Ghz processors on DMs; 533 FSBus; 2GB RAM; (4) Ethernet 10/100/1000 ports per DM]

→3U DME Front End, 3U Backend, 1U Falcon Control Station, 1U SPS [Optional DAE2's]

→Supports CX300/CX400/CX500/CX600/CX700/Symm 5.x/Symm DMX

→Tarpon Front End "Aspen" NAS personality card using Katana DAE2 enclosures

→Gateway DMs will ship with optical GBIC in port BE1 and MIA in port BE0

→Requires FLARE 14 or higher

→Supports 501/501G, 502/502G, limited to two Data Movers only!

#### **CAPTIVE CABLING:**

BE0 on both DMs use copper HSSDC to copper SFP cable to SPA

BE1 on both DMs use copper SFP to SFP cable to SPB

MIA and optical cable can be plugged into AUX0 for backups

#### **SAN GATEWAY CABLING:**

BE0 on both DMs use MIA with optical fibre cable to connect to primary switch

BE1 on both DMs use SFP GBIC for optical SFP cabling to secondary fibre switch

MIA optical cable on AUX0 for backups

**Note:** emc139623 documents an issue with POST versions less than 02.25 when replacing Data Movers, especially with RoHS compliant part numbers that are not recognized by the old POST version, causing Data Mover not to boot properly.

### **NS704G 4-WAY:**

--New 4-way high performance mid-tier system that includes (4) NS700 3.1Ghz DMs and (2) Falcon Control Stations (Optional)

--Model is always a fabric-connected Gateway system cabled to FC Switch using fiber optic cables and MIA's

--Crossover cable and subnet between Control Stations called IPMI (Intelligent Platform Management Interface)—192.168.3.x

--Dual CS configurations should use IP Aliasing to provide consistent IP address for access to Primary CS

**Note:** Torrey Pines Intel motherboard and Transcend memory problems on bus

--NAS 5.3 minimum

--I/O card with (6) Copper ports and (2) GigE optical ports

**Note:** The six copper Ethernet network ports on the Data Movers are labeled **cge0 - cge5**. The port labeled **cge6/mgmt** is reserved for the private LAN connection to CS1, when installed. All ports are 10/100/1000 & use standard RJ-45 connectors.

#### **SWITCH CONFIGURATIONS:**

→Use two switches for High Availability, with all BE0 ports to one switch, BE1 ports to the other switch

→Can have (4) fibre channel connections from each Data Mover if desired

→Set switch port speeds to highest setting (1Gb or 2Gb) but NOT autonegotiate

→Set DM Fibre Channel ports to autonegotiate, default setting

#### **TYPICAL NS704G vs. NS700 BIND TABLES:**

**\$ .server\_config server\_2 -v "fcp bind show"**

\*\*\* Persistent Binding Table \*\*\*

Chain 0000: WWN 500601603060090b HBA 0 SP-a0 Bound

Chain 0016: WWN 500601683060090b HBA 0 SP-b0 Bound

Chain 0032: WWN 500601693060090b HBA 1 SP-b1 Bound

Chain 0048: WWN 500601613060090b HBA 1 SP-a1 Bound

Chain 0064: WWN 500104f0008a3496 HBA 2 N\_PORT Bound

Chain 0080: WWN 5006048000000000 HBA 3 N\_PORT Bind Pending

**\$ .server\_config server\_2 -v "fcp bind show"**

\*\*\* Persistent Binding Table \*\*\*

Chain 0000: WWN 5006016710602235 HBA 0 SP-a7 Bound  
 Chain 0016: WWN 5006016f10602235 HBA 1 SP-b7 Bound  
 Chain 0032: WWN 10000e00222b2af HBA 2 N\_PORT Bound  
 Chain 0048: WWN 5006048000000000 HBA 3 N\_PORT Bind Pending

## **TYPICAL LUN ASSIGNMENTS NS600/NS700:**

**# nas\_disk -l**

| id | inuse | sizeMB | storageID-devID             | type  | name       | servers                            |
|----|-------|--------|-----------------------------|-------|------------|------------------------------------|
| 1  | y     | 4095   | APM00040303779- <b>0000</b> | CLSTD | root_disk  | 1,2                                |
| 2  | y     | 4095   | APM00040303779-0001         | CLSTD | root_ldisk | 1,2                                |
| 3  | y     | 2047   | APM00040303779-0002         | CLSTD | d3         | 1,2                                |
| 4  | y     | 2047   | APM00040303779-0003         | CLSTD | d4         | 1,2                                |
| 5  | y     | 2047   | APM00040303779-0004         | CLSTD | d5         | 1,2                                |
| 6  | y     | 2047   | APM00040303779-0005         | CLSTD | d6         | 1,2                                |
| 7  | y     | 252793 | APM00040303779- <b>0010</b> | CLSTD | d7         | 2,1 [Lun is Hex 0010 = Decimal 16] |
| 8  | y     | 252793 | APM00040303779-0011         | CLSTD | d8         | 2,1                                |

**# server\_devconfig server\_2 -p -s -a**

chain= 16, scsi-16  
 symm\_id= APM00040303779 celerra\_id= APM0004030377900  
 tid/lun= 0/0 type= disk sz= 0 val= -5 info= DGC RAID 5 02060000004429CL  
 tid/lun= 0/1 type= disk sz= 0 val= -5 info= DGC RAID 5 02060100004436CL  
 tid/lun= 0/2 type= disk sz= 0 val= -5 info= DGC RAID 5 02060200004443CL  
 tid/lun= 0/3 type= disk sz= 0 val= -5 info= DGC RAID 5 02060300004451CL  
 tid/lun= 0/4 type= disk sz= 0 val= -5 info= DGC RAID 5 0206040000445FCL  
 tid/lun= 0/5 type= disk sz= 0 val= -5 info= DGC RAID 5 0206050000446ECL  
 tid/lun= **1/0** type= disk sz= 0 val= -5 info= DGC RAID 5 02061000008CD9CL [Lun 1/0 = 16/0 = Decimal 16]  
 tid/lun= **1/1** type= disk sz= 252793 val= 8 info= DGC RAID 5 020611000084AFCL [Lun 1/1 = 16/1 = Decimal 17]

## **CELERRA RESERVED LUNS/CELERRA LUN PROTECTION/CELERRA ILLEGAL HLUs:**

HLU lun decimal numbers 6-15 are considered “reserved luns” for special use. As such, Data LUNs should never be assigned an HLU value that falls in this range. NAS code changes have eliminated this problem, as outlined below:

### **HLU values 6-15 for CLARiiON Systems:**

Beginning with NAS 5.5.27 & 5.6.36, NAS code will not diskmark illegal HLU luns on CLARiiON backends—operation will fail. Beginning with NAS 5.6.38, NAS code will allow diskmark operation to complete, but will auto reassign a valid HLU number.

## **ILLEGAL USE OF CELERRA RESERVED LUNS:**

### **DETERMINING HLU/ALU LUN ASSIGNMENTS & CORRECTING:**

**Caution:** The following procedure is meant to address only NSxxG SAN-attached systems. Captive NS platforms will require a different procedure involving data migration to new luns, or data backup, then unbind and rebind of luns with correct HLU assignments, since captive systems do not have Storage Groups or use Access Logix. Typically, the setup\_clariion script will build the backend config with the correct luns bound during the install.

#### **1. First step is to find WWN numbers for all Celerra HBA's:**

**\$ server\_log server\_3 |grep -i wwn |grep -v ONLINE →NS600G**

2004-11-08 10:48:19: CAM: 4: FCP ONLINE HBA 0: ALPA 000001 WWN: **50060168006007b9**

2004-11-08 10:48:23: CAM: 4: FCP ONLINE HBA 1: ALPA 000001 WWN: 50060169006007b9

**\$ server\_log server\_2 |grep -i WWN |grep -v ONLINE →NS700G**

2004-10-01 08:48:34: FCP: 4: ONLINE HBA 0: S\_ID 6a1113 WWN: **50060160106025c4**

2004-10-01 08:48:39: FCP: 4: ONLINE HBA 1: S\_ID 6a1013 WWN: 50060161106025c4

**Note:** Repeat for DM\_3

#### **2. Locate Matching WWN in Storagegroup List output to obtain HLU/ALU List:**

**\$ /nas/sbin/navicli -h 10.64.25.148 storagegroup -list** (AccessLogix is in use)

Storage Group Name: **NS600G**

Storage Group UID: 80:84:D6:A0:64:6D:D8:11:80:2C:AA:9C:3F:37:5E:FC

HBA/SP Pairs:

| HBA UID | SP Name | SPPort |
|---------|---------|--------|
| -----   | -----   | -----  |

50:06:01:60:80:60:07:B9:**50:06:01:68:00:60:07:B9** SP A**HLU/ALU Pairs:**

HLU Number ALU Number

| ----- | ----- |
|-------|-------|
| 0     | 50    |
| 1     | 51    |
| 2     | 52    |
| 3     | 53    |
| 4     | 54    |
| 5     | 55    |
| 16    | 0     |
| 17    | 1     |
| 18    | 3     |
| 19    | 2     |
| 20    | 5     |

**\$ /nas/sbin/navicli -h spa storagegroup -list |more**

Storage Group Name: laip2

Storage Group UID: B9:C3:E8:F9:5B:1C:D8:11:80:1E:C8:E8:0C:53:D4:D2

**HBA/SP Pairs:**

HBA UID SP Name SPPort

|  |
|--|
| 50:06:01:60:90:60:25:C4: <b>50:06:01:60:10:60:25:C4</b> SP A 1 |
| 50:06:01:60:90:60:25:C4:50:06:01:61:10:60:25:C4 SP A 1         |
| 50:06:01:60:90:60:25:C4:50:06:01:68:10:60:25:C4 SP A 1         |
| 50:06:01:60:90:60:25:C4:50:06:01:69:10:60:25:C4 SP A 1         |
| 50:06:01:60:90:60:25:C4:50:06:01:60:10:60:25:C4 SP B 1         |
| 50:06:01:60:90:60:25:C4:50:06:01:61:10:60:25:C4 SP B 1         |
| 50:06:01:60:90:60:25:C4:50:06:01:68:10:60:25:C4 SP B 1         |
| 50:06:01:60:90:60:25:C4:50:06:01:69:10:60:25:C4 SP B 1         |

**HLU/ALU Pairs:**

HLU Number ALU Number

|   |    |
|---|----|
| 5 | 26 |
| 4 | 25 |
| 3 | 24 |
| 2 | 23 |
| 1 | 22 |
| 0 | 21 |

6   **30 →Convert to Hex 0x1E**7   **29 →Convert to Hex 0x1D****3. Examine HLU Column (decimal) for any LUNs between 6-15.****Note:** Above example shows HLU LUNs 6 & 7, which are illegal for Celerra configurations, and must be moved to LUNs >than 16**4. From nas\_disk -list output, find respective ALU number for the illegal HLU's (see following example):****\$ nas\_disk -l**

```
id inuse sizeMB storageID-devID type name servers
1 y 11263 APM00030600872-0015 CLSTD root_disk 1,2
2 y 11263 APM00030600872-0016 CLSTD root_ldisk 1,2
15 y 48332 APM00030600872-001D CLSTD d15 2,1 →001D is ALU in Hex, decimal=29 on “d15”
16 n 48332 APM00030600872-001E CLSTD d16 2,1 →001E is ALU in Hex, decimal=30 on “d16”
```

**Note:** d15 and d16 will need to be deleted, but determine what file systems are in use, if any, via following commands:**\$ nas\_volume -i d15****\$ nas\_volume -i slice2****\$ nas\_volume -i ext2**

id = 120

name = ext2

acl = 0

in\_use = True

type = meta

volume\_set = slice2

disks = d8

**clnt\_volume = fs02**

5. Unmount File Systems and ensure that they are no longer in use.
6. Remove HLU LUNs 6 & 7 from StorageGroup using Navisphere
7. Add original ALUs 29 & 30 back to StorageGroup, but assign an HLU above 15 decimal
8. Reboot Data Mover
9. Run diskmark utility:

**#/nas/bin/nas\_diskmark -m -a**

11. Remount file systems if applicable and verify access

## **CORRECTING WRONG HLU LUNS ON PRODUCTION FILE SYSTEMS (SHORT VERSION):**

1. Identify all file systems built on the incorrect HLU numbers 0-15

**Note:** HLU's 0-15 are reserved only for Control Volume usage, when applicable

2. Permanently unmount all affected File Systems
3. Reassign HLU numbers using Navisphere and the following steps:
  - Connect to SPA/SPB using NAVISPHERE Manager>SPA>Storage Groups>rightclick and 'Select LUNs'
  - Select the incorrectly assigned LUNs and click on arrow to remove from Storage Group, then click apply
  - Select same LUNs from left pane and readd to Storage Group using arrow
  - Click on Host column (blank), rightclick, select any of the valid HLU numbers displayed that are > than or = LUN 16
  - Click apply to commit the change for the new HLUs
4. Reboot the affected Data Movers
5. Re-diskmark volumes:  
# nas\_diskmark -m -a
6. Remount all production file systems & verify access

## **MULTIPLE LUN 0 ISSUES:**

When adding LUNs to an existing system, it's possible to configure multiple HLU LUN 0's—this can cause boot problems for the data mover, but usually only after a binding table clear and rediscover. Quick way to check for LUN 0's on a system:

**\$ nas\_disk -query:"\*"-fields:Name,SymmId,ServerTable -format:"%s %s %q\n" -query:"\*"-fields:address -format:"%s " |grep t0l0**

root\_disk APM00023800390 c2t0l0 c18t0l0 c0t0l0 c16t0l0 c2t0l0 c18t0l0 c0t0l0 c16t0l0

→See emc109343 for more discussion and examples of the LUN 0 problem

## **CORRECTING BIND TABLE PROBLEM WITH CHAIN 0/2 LUN 0 CONFIGURED TO SEE SPB:**

1. Various arraycomm path settings were different among the SP ports—these were corrected in accordance with Primus emc95145 by issuing commands:

**# /nas/sbin/navicli -h 10.1.10.140 storagegroup -sethost -host nswebmgr2\_dm2\_0 -arraycomm path 0**

WARNING: Changing configuration options may cause the array to stop functioning correctly. Do you wish to continue (y/n)? y

**# /nas/sbin/navicli -h 10.1.10.141 storagegroup -sethost -host nswebmgr2\_dm2\_1 -arraycomm path 0**

2. Data mover needed to have Chain 16 restored. Fibre cables and zoning were checked, customer took action to shutdown Celerra and then reboot NS500 array, after which all LUNs were on the correct SP and devices were visible down Chain 16 on the Data Mover. In most cases, connectivity can be restored via less drastic measures!

3. Binding Table cleared

4. Setup\_slot conducted and bind table rebuilt with Chain 0 HBA0 seeing SPA and Chain 16 HBA1 seeing SPB

5. Do devconfig probe and “fcp bind show” to verify

### **Comments:**

LUNs 0-5 are used for Control Volumes. User ‘data’ LUNs begin at HLU 16.

**ALU—Array Luns are identified using \$ nas\_disk -l and are stated in Hex—convert to decimal for LUN number**

**HLU—Luns are identified using \$server\_devconfig -p -s -a and display in Decimal**

**Note: StorageGroup HLU data LUNs for Celerra cannot be less than 16. Verify HLU's from Navisphere, Server\_Devconfig, or Navicli StorageGroup list**

Create RAID Groups for Hot Spares as needed

**Caution:** In situations where Clariion Devices are added to StorageGroups and the HLU's are not manually assigned, the system will assign HLUs using LUNs 06 -15, which are reserved and not allowed for Data volumes. See Primus emc86551 to see if this issue can be resolved. Procedure entails unmounting any file systems that might be built on these luns, removing LUNs from storagegroup, adding back with correct HLU assignments, rebooting Servers, running nas\_diskmark -m -a and remounting all file systems.

**Note: NAS 5.3 no longer allows the assignment of HLU's 6-15 (decimal)(0x6—0xF hex) on the Celerra when creating Data Volumes—see primus emc94879.**

## # nas\_volume -n test2 -c d16

Error 3147: volume d16 resides on a reserved disk

## LOCATION OF BIOS & POST VERSIONS FOR NS PLATFORMS :

/nas/dosfs

NS600 → C2\_BIOS.rom

C2\_POST.bin

NS500 → TARPBIOS.rom

TARPPPOST.bin

NS700 → CUDABIOS.rom

CUDAPOST.bin

NS40, NSX, NS80, NS20, NX4, NS-120, NS-480, NS-960 → BIOS & POST incorporated into DART

## BIOS & POST VERSIONS BY NAS CODE :

5.1.26.0 BIOS 3.26 POST 3.12 [Only NS600 series]

5.2.18+ BIOS 3.42 POST 3.25 NS600 BIOS 3.30 POST 1.75 NS700

5.3.17.1 BIOS 3.42 POST 3.37 NS600 BIOS 3.30 POST 2.14 NS700 BIOS 3.28 POST 1.95 NS500

5.4.14.3 BIOS 3.42 POST 3.37 NS600 BIOS 3.30 POST 2.14 NS700 BIOS 3.28 POST 1.95 NS500

## NAS 5.5 BIOS AND POST LEVELS:

# cat bios

```
<DataMoverBIOSInfo Host="server_3">
<VERSION_INFORMATION BIOS_VERSION="3.42" POST_VERSION="Rev. 03.37" />
```

# server\_sysconfig server\_2 -P

server\_2 :

Processor = Intel Pentium 4

Processor speed (MHz) = 2000

Total main memory (MB) = 4024

Mother board = Chameleon II XP

Bus speed (MHz) = 400

Bios Version = 3.42

Post Version = Rev. 03.37

## VERIFYING DATA MOVER SERIAL NUMBERS:

Note: NAS 5.3 introduces a new serial number validation entry in the /nas/server/slot\_x/start file for NS systems, but the entry is not used nor populated in the start file for single DM systems:

uniqueid validate=LKE00024100134

# nas/sbin/get\_data\_mover\_status -resume server\_2 5080 /tmp/dm2.resume

# grep EMC\_SERIAL\_NUMBER /tmp/dmx.resume

Comment: Data Mover will boot to State 14 to indicate a misconfigured hardware state if this entry is incorrect or if a Data Mover has been changed without a setup\_slot, etc.

## USING EXPECT TO VERIFY DM SERIAL NUMBER:

# expect -f encl\_mon.exp resume server\_3 5080 |grep RESUME\_INFORMATION\_MIDPLANE

IDS exp7 1 RESUME\_INFORMATION\_MIDPLANE

```
<RESUME_INFORMATION_MIDPLANE EMC_PART_NUMBER="005047766" EMC_ARTWORK_REVISION=""  
EMC_ASSEMBLY_REVISION="A14" EMC_SERIAL_NUMBER="APM00025002080" VENDOR_NAME="" L  
OCATION_OF_MANUFACTURE="Apex, NC USA" YEAR_OF_MANUFACTURE="2002" MONTH_OF_MANUFA  
CTURE="12" DAY_OF_MONTH_OF_MANUFACTURE="15" ASSEMBLY_NAME="XPE NAS CHASSIS" WORL  
D_WIDE_NAME_SEED="10600278" />
```

## VERIFYING NS DATAMOVER HARDWARE:

# /nas/sbin/get\_data\_mover\_status -resume server\_2 5080 /home/nasadmin/dm2\_status

Note: Port 5080 is used prior to NAS 5.6. With NAS 5.6, use port 5082. AR95956.

# cat dm2\_status

<?xml version="1.0"?>

```
<DataMoverResume Host="server_2">  
<RESUME_INFORMATION_XP EMC_PART_NUMBER="005048243" EMC_ARTWORK_REVISION=""  
EMC_ASSEMBLY_REVISION="A19" EMC_SERIAL_NUMBER="APM00042604969" VENDOR_NAME=""  
LOCATION_OF_MANUFACTURE="" YEAR_OF_MANUFACTURE="2004" MONTH_OF_MANUFACTURE="6"  
DAY_OF_M<RESUME_INFORMATION_PSA EMC_PART_NUMBER="118031924" EMC_ARTWORK_REVISION=  
"000" EMC_ASSEMBLY_REVISION="A03" EMC_SERIAL_NUMBER="AC103042500939" "  
VENDOR_PART_NUMBER="API1FS06-000" VENDOR_ARTWORK_REVISION="000"  
VENDOR_ASSEMBLY_REVISION="M08" VENDOR_SERIAL_NUM="AC1042500939" V
```

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!

```

ENDOR_NAME="ACBEL POLYTECH INC." LOCATION_OF_MANUFACTURE="MANILA, PHILIPPINES"
YEAR_OF_MANUFACTURE="2004" MONTH_OF_MANUFACTURE="07" DAY_OF_MONTH_OF_MANUFACTURE="29"
ASSEMBLY_NAME="DUAL +12V P/S W/ FANS"
<RESUME_INFORMATION_MIDPLANE EMC_PART_NUMBER="005048242" EMC_ARTWORK_REVISION=""
EMC_ASSEMBLY_REVISION="A04" EMC_SERIAL_NUMBER="APM00042700097" VENDOR_NAME=""
LOCATION_OF_MANUFACTURE="Apex, NC USA" YEAR_OF_MANUFACTURE="2004"
MONTH_OF_MANUFACTURE="6" DAY_OF_MONTH_OF_MANUFACTURE="27" ASSEMBLY_NAME="NS700 XPE
CHASSIS - NA
<RESUME_INFORMATION_PERSONALITY EMC_PART_NUMBER="005047111" EMC_ARTWORK_REVISION="A02"
EMC_ASSEMBLY_REVISION="A16" EMC_SERIAL_NUMBER="LKE00023108310" VENDOR_NAME="BENCHMARK"
LOCATION_OF_MANUFACTURE="HUDSON,NH,USA" YEAR_OF_MANUFACTURE="2002"
MONTH_OF_MANUFACTURE="08" DAY_OF_MONTH_OF_MANUFACTURE="05" ASSEMBLY_NAME="ASSY
<VERSION_INFORMATION BIOS_VERSION="3.25" POST_VERSION="Rev. 02.14" /></DataMoverResume>
# /nas/sbin/get_local_hw_status xml_status
# cat xml_status
<?xml version="1.0"?>
<LocalHardwareStatus>
<Voltage Name="3.3V" Current="+3.28 V" Min="+0.00 V" Max="+0.00 V" Alarm="ALARM"></Voltage>
<Voltage Name="5V" Current="+5.10 V" Min="+0.00 V" Max="+0.00 V" Alarm="ALARM"></Voltage>
<Voltage Name="12V" Current="+11.62 V" Min="+0.00 V" Max="+0.00 V" Alarm="ALARM"></Voltage>
<FanSpeedStatus>
<FanSpeed Name="Fan1" Current="8132 RPM" Min="-1 RPM" Divisor="2" Alarm=""></FanSpeed>
<FanSpeed Name="Fan2" Current="8231 RPM" Min="-1 RPM" Divisor="2" Alarm=""></FanSpeed>
</FanSpeedStatus>
<Temperature Current="+35.00C" Limit="+1280C" Hysteresis="+1280C" Alarm=""></Temperature>
</LocalHardwareStatus>
```

## **DATA MOVER HARDWARE INFORMATON/PART NUMBERS:**

### **1. Obtain Part number info from file:**

```
# cat /nas/log/data_mover_resume.server_3.xml|head
```

```
<?xml version="1.0"?>
<DataMoverResume Host="server_3">
<RESUME_INFORMATION_BLADE
  EMC_PART_NUMBER="100-562-150"
```

### **2. Run get\_data\_mover\_status to collect information on particular server:**

```
# /nas/sbin/get_data_mover_status -resume server_3 5082 /home/nasadmin/server_3Resume.xml
>/dev/null 2>&1
```

```
# cat /home/nasadmin/server_3Resume.xml|head
<?xml version="1.0"?>
<DataMoverResume Host="server_3">
<RESUME_INFORMATION_BLADE
  EMC_PART_NUMBER="100-562-150"
```

**Note:** Port number 5082 is the new port used with NAS 5.6—no longer using port 5080

## **CURRENT HARDWARE STATUS:**

```
$ cat /nas/log/data_mover_status.server_2.xml
```

```
<?xml version="1.0"?>
<DataMoverStatus Host="server_2">
<ENCLOSURE_INFO PSA_INS_N="1" PSB_INS_N="1" PSA_FLT="0" PSB_FLT="1" PSA_SHTDN="0
" PSB_SHTDN="1" PSA_AC_FAIL="0" PSB_AC_FAIL="1" PSA_AMB_OT="0" PSB_AMB_OT="0" FA
N_FLT_M1="0" FAN_FLT_M2="0" FAN_FLT_M3="0" MULT_FAN_FLT="0" ENCL_LED_ON="1" SFP1
_STATE="0" SFP2_STATE="0" />
```

## **CELLERRA MANAGER ALERTS:**

Hardware Status alerts are queried by the APL Manager, which performs an inventory query, and logged to /nas/log/webui/alert\_log  
System Status alerts are generated by processes or components and also logged to alert\_log and sys\_log

**Note:** Deleting or Accepting ‘Alerts’ to cleanup the display is problematic. You can highlight all entries using ctrl + a, but if any of the underlying alerts are exact duplicates, the dupes may not be removed. So, you may need to do this a few times to cleanup all of the alerts. AR119686.

## **BIOS/POST UPGRADE PROCEDURE FOR NAS 5.5:**

1. Connect to the Primary Control Station (either via SSH or serial console)
  2. Rename /nas/sbin/recover\_slot to /nas/sbin/recover\_slot.orig to prevent data mover from getting rebooted during this process.
  3. Start PXE services by running the command “/nas/sbin/t2pxe -s -R <Number of data movers>”
  4. Copy the 5.5.30.5 Dart image onto the Control Station /tftpboot/nas.exe (overwrite existing file)
  5. Connect to the Data Mover’s serial console (details: 9600 baud, 8-N-1)
  6. Reboot the Data Mover using “server\_cpu server\_X –r now” where X is the slot id of the data mover.
  7. On the serial console of the data mover, press Esc key twice when the characters “ABabcdefgCDE...” start appearing, to break into the Extended POST
  8. Type the password “SHIP\_it” without the quotes to get into the Extended POST menu
  9. Enter the ICA submenu (option 24) and use option 2 to PXE boot the blade
- Note:** The blade will now boot from the new Dart image in min config mode (none of the filesystems will be mounted). As part of this process, it will also update the POST from 2.42 to 2.57 (and BIOS from 5.10 to 5.14). This process can take 2-5 minutes. After flashing the new firmware, the data mover will reboot. The Data Mover will now load the Dart image from the backend (5.5.28.1)
10. Wait for the reason code of the data mover to change to 5
  11. Stop PXE services by running the command “/nas/sbin/t2pxe –e”
  12. Remove /tftpboot directory
  13. Rename /nas/sbin/recover\_slot.orig to /nas/sbin/recover\_slot to allow failover to function properly.
  14. Repeat steps 5 through 12 on all data movers that need the new BIOS/POST.

## **CHANGING FIBRE CHANNEL CARD SPEED ON DM USING FCCBOOT OPTION:**

1. Connect null modem cable from service laptop to DM port
2. Reboot DM and press esc during Aabcde... string
3. Enter “FCCBOOT” and enter
4. Enter 2) FCC Boot Sub-Menu
5. Enter 3) BEO FCC Boot
6. Enter 4) Set Port Speed
7. Enter 0=1Gb
8. Repeat options for second port
9. Enter 1) Reset Controller

## **USING MINICOM TO CHANGE FCCBOOT SPEEDS ON NAS 5.3 SYSTEMS:**

1. Run minicom program on Control Station to setup profiles on ttyS4 and ttyS5 for DM2 and DM3, respectively [A, E, E, F, Save profile using a name such as dm2]: #minicom -s
2. Turn off NAS Services prior to resetting Server
3. Use t2reset to boot data mover  
#/nas/sbin/t2reset reboot –s 2
4. Initiate minicom serial session to DM2 from the Control Station:  
#minicom dm2
5. At following sequence, press the “esc” key [Aabcdef.....]
6. Type ‘FCCBOOT’ (Bios Menu for SP’s on Backend) and scan for targets or set port speed

**Note:** Password = SHIP\_it

FCC Boot Sub-menu

BEO FCC Boot

1.) Scan for target

2.) Set port speed

## **FCCBOOT SCREENS:**

**AabcdefBCDEabFabcdGHabIab** [Type the “ctrl + c” keys and let POST finish]

EndTime: 12/22/2004 19:14:46

.... Storage System Failure - Contact your Service Representative ...

\*\*\*\*\*

\*\*\*\*\* Aborting!!!! \*\*\*\*\*

**-->Type FCCBOOT**

Diagnostic Menu

1) Reset Controller      2) FCC Boot Sub-Menu

Enter Option : 2

FCC Boot Sub-Menu

- 1) Restore Def Port Settings 4) BE1 FCC Boot
- 2) Display Port Settings 5) AUX0 FCC Boot
- 3) BE0 FCC Boot 6) AUX1 FCC Boot
- 0) Exit

Enter Option : 2

PORT SETTINGS

| Port | B/E WWN Port      | WWN Primary                          |                             |
|------|-------------------|--------------------------------------|-----------------------------|
| Num  | B/E WWN Node      | Name::Bus:Dev:Func                   | WWN Secondary Port Settings |
| 000  | 00000000:00000000 | BE0 FCC::02:04:01 00000000:00000000  | 2Gb, AG ON, ENA, DEF        |
|      | 00000000:00000000 | 00000000:00000000                    |                             |
| 001  | 00000000:00000000 | BE1 FCC::02:04:00 00000000:00000000  | 2Gb, AG ON, ENA, DEF        |
|      | 00000000:00000000 | 00000000:00000000                    |                             |
| 002  | 00000000:00000000 | AUX0 FCC::01:04:01 00000000:00000000 | 2Gb, AG ON, ENA, DEF        |
|      | 00000000:00000000 | 00000000:00000000                    |                             |
| 003  | 00000000:00000000 | AUX1 FCC::01:06:01 00000000:00000000 | 2Gb, AG ON, ENA, DEF        |
|      | 00000000:00000000 | 00000000:00000000                    |                             |

Success

Hit any key to continue

FCC Boot Sub-Menu

- 1) Restore Def Port Settings 4) BE1 FCC Boot
- 2) Display Port Settings 5) AUX0 FCC Boot
- 3) BE0 FCC Boot 6) AUX1 FCC Boot
- BE0 FCC Boot
- 1) Scan for Targets 4) Set Port Speed
- 2) Enter Target ID 5) Set AG
- 3) Set Boot Target 6) Set Port Access
- 0) Exit

Enter Option : 1

Scan for ALL targets [Y]? y

Initializing back end FIBRE...

PCI Config Reg: 2.4.1 0x0157

FCDMTL 1 [2.4.1] Dual Mode Fibre init - OSW DB PTR 0x3A5F3E20

AG: init DMD to FC\_SPEED\_2\_GBPS

Target 0 is online

WARNING: Disk 1 Failed TUR

Scanned Targets Found: 2

Number of Targets: 02

TARGETS FOUND

| Idx | Handle   | Addr_id           | WWN Port          | Valid_bits     |  |
|-----|----------|-------------------|-------------------|----------------|--|
|     |          |                   | WWN Node          |                |  |
| 00  | 3A4FDAE4 | 00050F00          | 50060160:30600533 | LOG, TUR, LUN0 |  |
|     |          | 50060160:B0600533 |                   |                |  |
| 01  | 3A4FEA12 | 00050E00          | 50060169:30600533 | LOG            |  |
|     |          | 50060160:B0600533 |                   |                |  |

## CELERRA IP & CELERRA CX600 PRODUCTS:

First Disk Enclosure Units are all configured with (1) 4+1 RAID 5 Disk Group, (1) 8+1 RAID 5 Disk Group, & (1) Hot Spare  
Celerion II consists of public SAN Storage CX600 using AccessLogix on BackEnd [Setup methodology much different than C1]  
CX600 is being marketed as EMC Celerra Clustered Network Server (CNS) and CLARiiON Storage. CX600 CLARiiON Storage  
consists of (4) front-end Fibre Channel Ports per SP, (4) Back-End Fibre channel ports per SP, 4/8GB Cache Memory per SP, & an  
Ethernet management port. CX600 supports Synchronous Remote Mirroring using MirrorView, Snapshot Point-in-time Copies, Non-  
Disruptive Upgrades, Access Logix.

### UNBINDING CONTROL LUNs:

# /nas/sbin/navicli –h 192.168.1.200 unbind 0 –o [Repeat for LUNs 0 – 5]

### BINDING CONTROL LUNs:

#/nas/sbin/navicli –h 192.168.1.200 bind r5 0 –rg 0 –rc 1 –wc 1 –aa 1 –sp a –sq gb –cap 4

```
#/nas/sbin/navicli -h 192.168.1.200 bind r5 0 -rg 1 -rc 1 -wc 1 -aa 1 -sp a -sq gb -cap 4  
#/nas/sbin/navicli -h 192.168.1.200 bind r5 0 -rg 2 -rc 1 -wc 1 -aa 1 -sp a -sq gb -cap 2  
#/nas/sbin/navicli -h 192.168.1.200 bind r5 0 -rg 3 -rc 1 -wc 1 -aa 1 -sp a -sq gb -cap 2  
#/nas/sbin/navicli -h 192.168.1.200 bind r5 0 -rg 4 -rc 1 -wc 1 -aa 1 -sp a -sq gb -cap 2  
#/nas/sbin/navicli -h 192.168.1.200 bind r5 0 -rg 5 -rc 1 -wc 1 -aa 1 -sp a -sq gb -cap 2
```

### **VERIFYING BIND STATUS & OUTPUT:**

```
# /nas/sbin/navicli -h 192.168.1.200 getlun 39 -owner  
# /nas/sbin/navicli -h 192.168.1.200 getlun 39 -state -bind
```

State: Binding  
Prct Bound: 3

### **TROUBLESHOOTING HYPER TERMINAL & COM PORTS ON CONTROL STATION:**

#### **Verify use of null modem cable and settings:**

COM Port: Bits per second: 19200 Data bits: 8 Parity: none Stop bits: 1 Flow control: hardware

#### **Verifying Console Redirection:**

1. Connect null modem cable to CS laptop and reboot
2. BIOS screen press F2 to enter setup
3. In Boot menu, verify that “Serial Console Redirection” is enabled and that “COM2 2FB IRQ3” is selected for “Serial Port”
4. Remove null modem and reconnect using Hyper Terminal

### **WHAT IS PXE BOOT?**

PXE stands for "Preboot Execution Environment" and is an open industry standard that allows PCs to boot over the network from a "boot image file" using a PXE-enabled Network Interface Device. PXE uses the Dynamic Host Configuration Protocol (DHCP) to assign the Client PC a temporary IP Address to work in conjunction with the Network Interface Card (NIC). Data Movers must use the PXE Boot method for installation and for troubleshooting when they cannot boot-up from the Array—in this situation, Data Movers boot from a NAS image on the CS over a private LAN connection using PXE & DHCP services, and a boot.cfg file.

### **PXEBOOT PROCESS (Revised PXE Boot steps):**

#### **Data Mover Blade Boot Rules:**

- Blade will try to boot from backend via BE0
- If Blade cannot boot from backend BE0, will try to PXE boot
- If PXE boot fails, the Blade will then try to boot from backend via BE1
- Additionally, the Control Luns need to be available on Chains 0 or 16 in order to boot

#### **1. Use /tftpboot/bin commands for PXE process:**

**Note:** If tftpboot directory does not exist, extract from /nas/tools. There is a command that creates the tftpboot directory, which also turns on tftp services, but also turns on the PXE services, which you may not want to do without specifying the subnets as shown below: # /nas/sbin/t2pxe -tftp start

```
# gzip -d /nas/tools/tftpboot.tar.gz  
# cd /  
# tar -xvf /nas/tools/tftpboot.tar
```

#### **2. Stop NAS Services and then mount the /nas partition:**

```
# /sbin/service nas stop  
# mount /nas
```

#### **3. Start PXE Services**

```
# /tftpboot/bin/t2pxe -s -R 2 128.221.252 128.221.253 [where -R x could be 2, 4, 6, or 8 depending on number of even blades in the cabinet]
```

```
=====  
EMC NAS PXE-BOOT Setup  
=====
```

```
Using Local Subnet 128.221.252.x for primary network  
Using Local Subnet 128.221.253.x for backup network
```

Starting PXE Services...

Done.

#### **4. Verify slot 2 & 3 config files that are created by PXE services:**

```
# cat /tftpboot/slot_2.cfg  
-serial swapserial  
flashupg post=upgrade bios=upgrade  
param config cs_ip=128.221.252.100:128.221.253.100
```

```
bufcache
device pci pci-0
pciautoconfig
ifconfig el30 dev=mge0 local=128.221.252.2 broadcast=128.221.252.255 netmask=255.255.255.0
ifconfig el31 dev=mge0 local=128.221.253.2 broadcast=128.221.253.255 netmask=255.255.255.0
hostname server_2 selfid=2
file initialize nodes=65536 dnlc=262144
transport start
tcp
rpc
volume disk NBS1 c0t0l0
volume disk NBS1 c16t0l0
volume disk NBS5 c0t0l4
volume disk NBS5 c16t0l4
volume disk NBS6 c0t0l5
volume disk NBS6 c16t0l5
nbs add nbsid=1 vol=NBS1 rw=128.221.252.100:128.221.253.100 exclusive raw share
nbs add nbsid=5 vol=NBS5 rw=128.221.252.100:128.221.253.100 exclusive raw share
nbs add nbsid=6 vol=NBS6 rw=128.221.252.100:128.221.253.100 exclusive raw share
nbs start
mac allow=128.221.252.100:128.221.253.100 httpport=5082
```

## **5. PXE Boot the desired Blade by slot:**

```
# /ftpboot/bin/t2tty -p 2 [slot_2 in this example]
```

Attempting to force data mover PXE boot:

Resetting Slot-2 with t2reset...

Waiting for DM network connectivity after PXE booting..

Data mover has PXE booted successfully.

You can now stop pxe services on CS by running t2pxe -e

Slot-2 Forced PXE Boot Succeeded.

## **6. Stop PXE Services**

```
# /ftpboot/bin/t2pxe -e
```

```
=====
EMC NAS PXE-BOOT Setup
=====
```

```
=====
Stopping PXE Services...
```

## **7. Verify that NBS client access has been restored to Backend:**

```
# /sbin/fdisk -l
```

```
-----output abridged-----
Device Boot Start End Blocks Id System
/dev/ndal * 1 17 136521 6 FAT16
/dev/ndal3 654 1435 6281415 8e Linux LVM
```

## **8. Restart NAS Services & run healthcheck:**

```
# /sbin/service nas start
```

**Note:** NAS services may take 5 minutes or more to completely start. Use ps -ef |grep nas\_m and ps -ef |grep boxm to verify that NAS Services & Box Monitor processes are running before running the healthcheck.

```
# nas_checkup
```

## **NAS 5.2/5.3/5.4/5.5 PXEBOOT RECOVERY ON NS SYSTEMS:**

**Note:** If Data Movers are not running, or are failing, and Control Station cannot see backend Control Luns using NBS protocol, then the only choice is to PXE Boot the data movers so that they boot from an image on the local IDE drive on CS0 using dhcp & TFTP protocol.

### **CONFIGURING MINICOM & USING PXEBOOT TO RECOVER NBS SERVICES: NAS 5.2 – 5.5**

**Note:** Minicom only works with NS600/500/700 Series Celerras—there is no serial connection between CS & DM Blades for NSX

1. Connect to Control Station via normal modem dialup, webex session, telnet, ssh, or serial port using hyper-terminal client

2. Login as User nasadmin, then root

3. Stop NAS Services

```
# /sbin/service nas stop
```

**Note:** If the /nas partition is in use by the system or a user, NAS Services may not stop cleanly. Always verify that all nas\_mcd processes have been stopped before continuing, and kill any remaining nas\_mcd processes. As a last resort, reboot Control Station and try stopping NAS Services again.

**Example:**

**# /sbin/service nas stop**

```
NAS:/etc/init.d/nas: ERROR: Failed to stop NAS services
/nas is in use by the following processes(uid)
/nas:      15477e 15478e 18172e 18177e 23209e 23210e
/nbsnas is mounted
# ps -ef |grep nas_m
root   1198  1 0 Sep26 ?    00:00:00 /nasmcd/nas_mcd -h /nasmcd /nas/
root   1453 1198 0 Sep26 ?    00:00:00 /nasmcd/nas_mcd -h /nasmcd /nas/
root   1656 1453 0 Sep26 ?    00:00:00 /nasmcd/nas_mcd -h /nasmcd /nas/
# pkill nas_mcd (stops all nas_mcd processes except parent ID 1198)
# kill -9 1198
```

4. If this procedure is being used to configure a Minicom profile and then access the Data Mover's POST or BIOS Systems (NBS Services are running normally and the backend partitions are available to the CS), simply remount the partitions:

```
# mount /nbsnas # mount /nas # mount /nas/dos # mount /nas/var
```

5. Setup Minicom profiles for each Data Mover (as required):

**#minicom -s**

Select 'Serial port setup'

Select "A" & manually change /dev/ttyS1 to /dev/ttyS4 for DM2, /dev/ttyS5 for DM3, /dev/ttyS6 for DM4, /dev/ttyS7 for DM5, then press the "enter" key

Select "E" twice to configure Speed/Parity for 9600 8N1 & press "enter"

Select "F" key to change Hardware Flow Control to No & press "enter"

'Save setup as...' and give a name to the profile, such as "dm2" & repeat for other Data Movers if required

Exit from Minicom

6. After minicom profiles have been configured, proceed with PXE Boot Recovery, stopping any existing PXE Boot Services:

**#/nas/sbin/t2pxe -e**

```
=====
          EMC NAS PXE-BOOT Setup
=====
Stopping PXE Services...
```

7. Restart PXE services in one-install mode:

**#/nas/sbin/t2pxe -s -R 2** [-R 2 for (2) Data Mover systems or -R 4 for (4) DM systems]

```
=====
          EMC NAS PXE-BOOT Setup
=====
```

Starting PXE Services...

Done.

8. Verify that the DHCP Daemon is running on the Control Station:

```
# ps -ef |grep -i dhcp
root   31302  1 0 13:36 ?    00:00:00 /usr/sbin/dhcpd -cf /tftpboot/dh
```

9. Force the Data Mover to PXE boot using t2tty -p command:

**#/nas/sbin/t2tty -p <slot\_x> &**

Attempting to force data mover PXE boot:

Resetting Slot-2 with t2reset...

Slot-2 Reset Successfully.

Waiting for DM To Come Up To PXE Boot Option....

Selecting DM PXE boot option...

Data Mover Is PXE Booting.

Waiting for DART to start after PXE boot...

Data mover has PXE booted successfully.

You can now stop pxe services on CS by running t2pxe -e

Slot-0 (ttyS2) Forced PXE Boot Succeeded.

**Note 1:** Forces the DM to boot from a nas.exe image located locally on the CS in the /nas/tools/tftpboot.tar.gz file. If this file is corrupt, upload a new file based on the same NAS version. There may be instances where PXE fails because of a problem with HBA0. If this happens, either unplug HBA0's fibre channel connection, or reboot Data Mover to access POST Menu & force PXE Boot.

```
# tar tzvf /nas/tools/tftpboot.tar.gz |head  
drwxr-xr-x 4294967294/4294967294 0 2006-08-30 18:35:03 tftpboot/  
-rwxr-xr-x 4294967294/4294967294 25208729 2006-08-30 18:34:55 tftpboot/nas.exe  
10. Alternatively, to watch the bootup process, ctrl + z the PXE Boot process and open a Minicom session
```

**# minicom dm2** (“dm2” represents minicom profile created in earlier step for Data Mover 2)

**Note:** Use “ctrl + A”, then enter “x” to exit minicom session. Be aware that the ctapty logs record each Data Mover boot and can be easily verified by going to the appropriate log for each Server:

/var/log/ctaptyS4.log (ctapty log for Server\_2)

11. If the Data Mover successfully boots, NBS Services should allow the Control Station to see its Backend Volumes—try remounting the volumes manually & continue troubleshooting your specific issue:

**# mount /nbsnas # mount /nas # mount /nas/dos # mount /nas/var**

12. Stop PXE Services:

**# t2pxe -e**

13. Reboot Data Mover using server\_cpu

**# server\_cpu server\_2 -r now**

14. Restart NAS Services

**# /sbin/service nas start**

### **TROUBLESHOOTING PXE BOOT DIFFICULTIES:**

→May need to reboot CS0 & retry PXE boot if it is failing

**# reboot -n -f**

→Have DM physically reseated and try again

→Check CTAP Logs for errors, etc: **/var/log/ctaptyS4.log**

→Verify Internal IP Address scheme in /etc/hosts file

→Access BIOS program & move PXE boot order to top of list (remember to set back to original position later)

→Verify status of Allied Telesyn 8-Port Internal 10/100 Ethernet Switch in back of Celerra Unit (power off, power on, cables, etc)

→Use FCCBOOT program to verify BE0, BE1 port speed settings, Scan for Backend Targets, Reset the HBA Controllers

**# /nasmcd/sbin/t2reset -s 2**

# minicom dm2 (or Hyper Term sessions via direct serial connection)

At following BIOS screen, “Aabcde...” type “esc” to exit bootup: Type “FCCBOOT” to enter FCCBoot Program

Select 1) to Reset Controller, Select 2) for FCC Boot Sub-menu and then 2) to Display Port Settings or 3) to enter BE0 FCC Boot menu, 1) Scan for target

### **FORCING PXE BOOT BY BYPASSING NORMAL BIOS BE0, PXEBOOT, BE1 BOOT ORDER:**

1. Stop NAS Services
2. Setup minicom profile for data mover from control station
3. Stop PXE Services and then restart in one-install mode
4. Reboot data mover using t2reset command
5. Initiate minicom session to DM from CS
6. At following message, press “esc” key [Aabc.....]
7. Enter password: SHIP\_it
8. Locate appropriate ICA Sub-Menu
9. Select 2) PXE Boot from the ICA Sub-Menu to force immediate PXE Boot that bypasses the normal BE0, PXE, BE1 Boot Order
10. Troubleshoot issue at hand
11. Stop PXE Services
12. Reboot Data Mover
13. Restart NAS Services

**Note:** There are references to older procedures that recommend changing the Boot Order by accessing the BIOS program upon Data Mover bootup using F2, then moving the PXE-2.0 boot order up above BE0 temporarily so as to be able force a PXE Boot. The above procedure is much simpler and does not muck with any permanent BIOS settings.

### **ACCESSING BIOS AND CHANGING BOOT ORDER WITH NAS 5.5:**

1. Stop NAS Services
2. Configure minicom profile
3. Reboot Data Mover
4. At following screen, press F2 repeatedly until BIOS menu appears:

Phoenix ServerBIOS 3 Release 6.0.

EMC BIOS Release 3.42

CPU = 4 Intel(R) Xeon(TM) CPU 2.00GHz

637K System RAM Passed

4031M Extended RAM Passed

512K Cache SRAM Passed

System BIOS shadowed

ERROR

0211: Keyboard error

Press <F1> to resume, <F2> to Setup

5. Press arrow key twice to scroll over to the “Boot” menu (→)

6. Highlight PXE-2.0 using down arrow key

7. Move Boot Order up or down using <shift +> or - key, respectively

8. After changes are made, arrow over to Exit Menu and select appropriate item:

Exit Saving Changes

Exit Discarding Changes

**Note:** Data Mover will reboot. Exit minicom using <ctrl + a >, then <x>

## **USING MINICOM TO TROUBLESHOOT & INITIATE PXE BOOT RECOVERY:**

**Purpose:** Minicom is a terminal program that runs on the CS and uses serial ports to connect to the Data Mover console, but only applies to old-style NS products: NS500, 600, 700, etc. Program is useful to observe Data Mover boot process. Also, this program is a necessary ‘backdoor’ to the CX600 backend if both data movers are down and /nbsnas is not mounted. You would use minicom to change the boot order if you needed to PXE Boot the data movers, which means that the data movers would boot a nas image directly from the Control Station’s internal IDE drive using the iscsi & PXE protocols.

The Linux Control Station for NS Series Celerra uses an internal serial modem program called 'Minicom' by which a user can connect to the Data Mover 'Console' from the Linux command shell or via a HyperTerminal session, to observe reboots, PXE boots, to access BIOS, POST, or FCCBoot menus, or to initiate commands directly to a Data Mover using the t2tty facility.

In order to use the Minicom Modem facility, a 'profile' must first be setup on the Data Mover, and given a name that is later used to invoke the actual 'Minicom' session between Control Station and Server. The following procedure explains how to setup a profile and how to initiate a Minicom session, and then exit the application.

### **I. Setup the Minicom profile for a Data Mover:**

a. Stop NAS Services and manually remount /nbsnas, /nas, /nas/dos, & /nas/var

# /sbin/service nas stop # mount /nbsnas /nas /nas/dos /nas/var

b. Start Minicom and setup profile:

**# minicom -s**

**Select ‘Serial port setup’**

**Select “A” & change /dev/ttyS1 to /dev/ttyS4 for DM2,/dev/ttyS5 for DM3, /dev/ttyS6 for DM4, /dev/ttyS7 for DM5, enter**

**Select “E” twice to configure Speed/Parity to 9600 8N1 & press enter**

**Select “F” key to change Hardware Flow Control to No & press enter**

**Save setup as... and give a name to the profile, such as “dm2”, enter [repeat to create other profiles: e.g., “dm3”, etc.]**

**Exit from Minicom** [to quit the profile session]

### **II. Establish Minicom session for Server\_2:**

a. Reboot or PXEBoot Data Mover

b. During system reboot, start Minicom Session by typing "minicom dm2" [just type to screen, even if Linux prompt is not visible]

c. Observe Data Mover reboot, enter Data Mover commands at CONSOLE> once system is up, etc.

**Note:** Alternatively, check the appropriate Server Log to observe bootup entries: /var/log/ctapptyS4.log (Server\_2)

d. To shutdown the Minicom Session, use “ctrl + a” then “x” to exit

e. If NAS Services were previously stopped, start NAS

#/sbin/service nas start

## **PXE BOOTING WITH NAS 5.2 & 5.3:**

**# /nas/sbin/t2pxe -e (Stops PXE boot services)**

**# /nas/sbin/t2pxe -s (Starts PXE boot services)**

**Note:** t2pxe is a shell script designed to assist data movers in creating a PXE boot configuration file (slot\_x.cfg) and in starting DHCP and TFTP services

**# /nas/sbin/t2pxe -s -R 2 [slot\_# 1, 2, or 4, etc]**

**# /nas/sbin/t2pxe -p <slot>** (forces PXE boot)

## **COMMAND SYNTAX FOR T2TTY :**

**# /nas/sbin/t2tty**

Celerra\IP Serial Connection Tester....

Usage:

Count active serial connections on ttyS4 thru ttyS9

t2tty -n

Check serial connection signal on ttyS4 thru ttyS9

t2tty -t

Check serial connection signal to a given slot (2-5)

t2tty -s 2

Force Data Mover in slot (2-5) to PXE boot

t2tty -p 2

Send dart command "cmd" to data mover in slot (2-5)

t2tty -c # "cmd" [e.g., t2tty -c 2 "ifconfig"]

## **NORMAL BIOS BOOT ORDER FOR NS DATA MOVERS :**

BE0

PXE

BE1

## **PXE BOOT RECOVERY PROCEDURE FOR CELERRA NS600:**

**Note:** # /nas/sbin/t2pxe -s 192.168.1.2 → New NAS 5.2 to setup PXE Booting, requires /nas/tools/tftpboot.tar.gz file—use IP address after -s if Internal IP address has changed.

1. Connect Serial cable from PC to CS0 and setup Hyper Terminal session

2. Establish serial port minicom session to DM2: #minicom dm2

### **CONSOLE>**

**Note:** Stop NAS Services prior to using Minicom [/sbin/service/nas stop]. Manually remount /nas, /nbsnas, /nas/dos, & /nas/var, if applicable.

3. Reboot data mover locally, or reseat server. With minicom session open, observe boot up process and access DM bios to change the Boot Order.

4. Press the following key combinations in order to access the Bios Setup Menu—run the key combinations during System RAM and at end of Extended RAM Checks:

Phoenix Server BIOS 3 Release 6.0

CPU=4 Intel ® XEON ™ CPU 2.00GHZ

637K System RAM Passed

4031M Extended RAM Passed

512I Cache...

**esc shift + O shift + Q**

**Note:** On some keyboards, this can be achieved using F2

5. Press right arrow key twice to move over to “Boot” Menu options. Change Boot Order of PXE in the “Boot” Menu by highlighting “Intel UNDI, PXE-2.0 (build 079)”. Use the – or + key to move entry up or down. Place ‘PXE-2.0’ entry just below the “Extended POST” line and above the “Fibre Port BE0” line:

PXE BOOT:

```
PhoenixBIOS Setup Utility
Main Advanced Boot Exit
                                Item Specific Help
· Intel UNDI, PXE-2.0 (build 079)
· Removable Devices
· Extended POST          . Keys used to view or
· Fibre Port BE0           . configure devices:
· Fibre Port BE1           . <Enter> expands or
· +Hard Drive              . collapses devices with
· Removable Devices        . a + or -
```

6. Save changes to Boot Menu by pressing “esc”, “esc” and then “enter” when “Exit Saving Changes” is highlighted. Press enter again until program exits.

7. Start the PXE Service [See step 2 of above minicom procedure for tftpboot file info]: #/etc/rc3.d/s95nas -p

=====

EMC NAS PXE-BOOT Setup

=====

[Nov 17, 17:22:13]

Detecting movers in cabinet: 2

Starting PXE Services...

/etc/rc3.d/s95nas: chkconfig: command not found

Done.

8. Verify PXE Services: # ps -ef |grep dhcp

root 5544 1 0 15:10 ? 00:00:00 /usr/sbin/dhcpd -cf /tftpboot/dh

9. Reboot Server using /nas/sbin/t2reset reboot -s 2 in order to initiate PXE boot.

10. Open minicom session [#minicom dm2] to observe boot process or use cat to check data mover boot state: #cat </dev/ttys4  
Use ctrl + c to exit the cat session. Alternatively, use #/nas/tools/.server\_tty -s 2 to check on DM state.
11. Exit minicom using “Ctrl+A”, then “x”
12. Conduct recovery tasks. When completed, restore correct Boot order in Bios Setup Menu.
13. Disable PXE Boot Service by running following commands:

```
# killall dhcpcd  
# rm -rf /tftpboot
```

14. Restart NAS Services

#### **Using Minicom to Observe Backend Storage Processors:**

1. Use serial cables S3 or S4 on laptop to connect to backend storage processors using null modem cable
2. Use setup procedure referenced above and use “ttys6 or ttys7” for S3 or S4 ports, respectively
3. Save setup and exit Minicom, then perform setup as in Step 2. above, but use new connection just created to access SP.

#### **AVM—AUTOMATIC VOLUME MANAGEMENT:**

Profiles must be selected before implementation in order to use Web UI

#### **CLARIION STORAGE PROFILES:**

clar\_r5\_economy: CLARiiON RAID 5 disk vols from 8+1 RAID group, aggregated by a stripe over an SP-balanced pair  
clar\_r5\_performance: CLARiiON RAID 5 disk vols from 4+1 RAID group, aggregated by a stripe over an SP-balanced pair  
clar\_r1: CLARiiON RAID 1 disk vols, aggregated by a stripe over an SP-balanced pair [2 Mirrored drives]  
clarata\_archive: CLARiiON ATA disks built on a Raid 5 raid group--New profile with 5.1.18 + code

#### **AVM RULES:**

- Clarion always requires an 8k stripe, whereas Symmetrix should stripe 32k for NFS, 8k for CIFS, and 256k for Highroad
- Try to allocate most important file systems first
- Do not use AVM with Timefinder environments
- Take one volume from each RAID group & alternate SP ownership
- Stripe volumes together in groups of (4) luns
- Do not spread file systems across more than one profile
- Web UI can only manage fs using profiles [i.e., fs created from CLI cannot be managed from Web UI]

#### **DETERMINING DISK POOLS FOR NAS 5.2:**

```
# /nas/bin/nas_cmd @nas_pool -l
```

| id | inuse | acl | name                |
|----|-------|-----|---------------------|
| 1  | n     | 0   | symm_std            |
| 2  | n     | 0   | clar_r1             |
| 3  | n     | 0   | clar_r5_performance |
| 4  | n     | 0   | clar_r5_economy     |
| 5  | n     | 0   | symm_bcv            |
| 6  | n     | 0   | symm_std_rdf_tgt    |
| 7  | n     | 0   | symm_bcv_rdf_tgt    |
| 8  | n     | 0   | symm_std_rdf_src    |
| 9  | n     | 0   | symm_bcv_rdf_src    |
| 10 | n     | 0   | clarata_archive     |

```
# cat /nas/sys/profiles
```

```
1:symm_std_vp:0:y:1:1:Symmetrix STD:11:8:32768:symm1:1:  
-----output abridged-----
```

#### **NAS 5.6 SYSTEM-DEFINED STORAGE POOLS:**

```
# cat /nas/volume/pools
```

```
1:symm_std:0:Symmetrix STD:1:1:n:n:n:1:y:n:  
2:clar_r1:0:CLARiiON Mirrored Pairs:1:1:y:n:n:2:y:y:  
3:clar_r5_performance:0:CLARiiON RAID5 4plus1:1:y:n:n:3:y:y:  
4:clar_r5_economy:0:CLARiiON RAID5 8plus1:1:y:n:n:4:y:y:  
5:symm_bcv:0:Backwards Compatibility:1:1:n:n:y:5:y:n:  
6:symm_std_rdf_tgt:0:Backwards Compatibility:1:1:n:n:n:6:y:n:  
7:symm_bcv_rdf_tgt:0:Backwards Compatibility:1:1:n:n:y:7:y:n:  
8:symm_std_rdf_src:0:Backwards Compatibility:1:1:n:n:n:8:y:n:  
9:symm_bcv_rdf_src:0:Backwards Compatibility:1:1:n:n:y:9:y:n:
```

10:clarata\_archive:0:CLARiiON RAID5 on S-ATA:1:1:y:n:n:10:y:y:  
 11:clarata\_r3:0:CLARiiON RAID3 on S-ATA:1:1:y:n:n:21:y:y:  
 12:cm\_r1:0:Remote Mirrored CLARiiON Mirrored Pairs:1:1:y:n:n:23:y:y:  
 13:cm\_r5\_performance:0:Remote Mirrored CLARiiON RAID5 4plus1:1:1:y:n:n:24:y:y:  
 14:cm\_r5\_economy:0:Remote Mirrored CLARiiON RAID5 8plus1:1:1:y:n:n:25:y:y:  
 15:cmata\_archive:0:Remote Mirrored CLARiiON RAID5 on S-ATA:1:1:y:n:n:26:y:y:  
 16:cmata\_r3:0:Remote Mirrored CLARiiON RAID3 on S-ATA:1:1:y:n:n:27:y:y:  
 17:clar\_r6:0:CLARiiON RAID6:1:1:y:n:n:33:y:y:  
 18:clarata\_r6:0:CLARiiON RAID6 on S-ATA:1:1:y:n:n:34:y:y:  
 19:cm\_r6:0:Remote Mirrored CLARiiON RAID6:1:1:y:n:n:35:y:y:  
 20:cmata\_r6:0:Remote Mirrored CLARiiON RAID6 on S-ATA:1:1:y:n:n:36:y:y:  
 21:symm\_ata:0:Symmetrix ATA:1:1:n:n:n:41:y:n:  
 22:symm\_bcv\_ata:0:Backwards Compatibility:1:1:n:n:y:42:y:n:  
 23:symm\_ata\_rdf\_tgt:0:Backwards Compatibility:1:1:n:n:n:43:y:n:  
 24:symm\_bcv\_ata\_rdf\_tgt:0:Backwards Compatibility:1:1:n:n:y:44:y:n:  
 25:symm\_ata\_rdf\_src:0:Backwards Compatibility:1:1:n:n:n:45:y:n:  
 26:symm\_bcv\_ata\_rdf\_src:0:Backwards Compatibility:1:1:n:n:y:46:y:n:  
 27:symm\_ssdd:0:Symmetrix SSD:1:1:y:n:n:53:y:n:  
 28:clarata\_r10:0:CLARiiON RAID10 on S-ATA:1:1:y:n:n:55:y:y:  
 29:cmata\_r10:0:Remote Mirrored CLARiiON RAID10 on S-ATA:1:1:y:n:n:57:y:y:  
 30:clarsas\_r10:0:CLARiiON RAID10 on SAS:1:1:y:n:n:59:y:y:  
 31:cmsas\_r10:0:Remote Mirrored CLARiiON RAID10 on SAS:1:1:y:n:n:61:y:y:  
 32:clarsas\_archive:0:CLARiiON RAID5 on SAS:1:1:y:n:n:63:y:y:  
 33:cmsas\_archive:0:Remote Mirrored CLARiiON RAID5 on SAS:1:1:y:n:n:65:y:y:  
 34:clarsas\_r6:0:CLARiiON RAID6 on SAS:1:1:y:n:n:67:y:y:  
 35:cmsas\_r6:0:Remote Mirrored CLARiiON RAID6 on SAS:1:1:y:n:n:69:y:y:  
 36:clar\_r10:0:CLARiiON RAID10 on Fibre Channel:1:1:y:n:n:71:y:y:  
 37:cm\_r10:0:Remote Mirrored CLARiiON RAID10 on Fibre Channel:1:1:y:n:n:73:y:y:

**Note:** Diskmarking enforces diskcount for FC clar\_r5\_economy (8+1=9 disks) and clar\_r5\_performance (4+1=5 disks) only and not for any other Storage Profile.

## **CELLERRA/CLARIION INSTALLATION TEMPLATES: AVAILABLE DISK GROUP & DISK VOLUME CONFIGURATIONS:**

### **DISK GROUP**

| <b>TYPE:</b>       | <b>DISK DRIVE TYPE</b> | <b>STORAGE PROFILE</b> | <b># LUNs/RG</b> |
|--------------------|------------------------|------------------------|------------------|
| 8+1 RAID5          | Fibre Channel          | clar_r5_economy        | 2                |
| 4+1 RAID5          | Fibre Channel          | clar_r5_performance    | 2                |
| RAID 1             | Fibre Channel          | clar_r1 (mirrored)     | 2                |
| 6+1 RAID5 (CX)     | ATA                    | clarata_archive        | 2                |
| 6+1 RAID5 (CX3)    | ATA                    | clarata_archive        | 1                |
| 4+1 RAID5 (CX3)    | ATA                    | clarata_archive        | 1                |
| 8+1 RAID5 (CX3)    | ATA                    | clarata_archive        | 2                |
| 4+1 RAID3 (CX3,CX) | ATA                    | clarata_r3             | 1                |
| 8+1 RAID3 (CX3,CX) | ATA                    | clarata_r3             | 2                |
| 6+1 RAID5 (CX3)    | LCFC                   | clarata_archive        | 1                |
| 4+1 RAID5 (CX3)    | LCFC                   | clarata_archive        | 1                |
| 4+1 RAID3          | LCFC                   | clarata_r3             | 1                |
| 8+1 RAID3          | LCFC                   | clarata_r3             | 2                |
| 4+2 RAID6          | FC                     | clar_r6                | 2                |
| 6+2 RAID6          | FC                     | clar_r6                | 2                |
| 12+2 RAID6         | FC                     | clar_r6                | 4                |
| 4+2 RAID6(CX3)     | ATA                    | clarata_r6             | 2                |
| 6+2 RAID6(CX3)     | ATA                    | clarata_r6             | 2                |
| 12+2 RAID6(CX3)    | ATA                    | clarata_r6             | 4                |
| 4+2 RAID6          | LCFC                   | clarata_r6             | 2                |
| 6+2 RAID6          | LCFC                   | clarata_r6             | 2                |
| 12+2 RAID6         | LCFC                   | clarata_r6             | 4                |

**Note:** Clariion rules say that all ATA LUNs in a RG must be owned by a single SP. You would balance SP ownership across RG's or Shelves. Fibre Channel drives can be R5, R1, or R6 for Celerra configurations. ATA drives can be R5 or R3 for CX arrays, and also R6 for CX3 arrays.

**TEMPLATE DISK GROUP REQUIREMENTS:**

- First enclosure on any system is always RAID 5
- All other enclosures can be either all RAID 5 or all RAID 1, but not both
- All disk groups must have (2) disk volumes bound for user data
- CX600/CX700 & NS600/NS700 have up to (15) disk devices per DAE
- FC4700 uses (10) disk devices per shelf
- First (6) disks NS600/NS700 are configured as a RAID5 (4+1) Disk Group with (5) control disks and (1) Hot Spare
- CNS CX600/CX700/FC4700-2 systems are configured at system install time using one of the templates
- All other Clariion systems can have remaining disk devices configured per the templates
- ATA drives cannot be configured for Bus 0 Enclosure 0

**STANDARD CELERRA TEMPLATES:** CX-Standard\_Raid\_5; CX\_All\_4plus1\_Raid\_5; CX\_Standard\_Raid\_1

**CX STANDARD RAID5 TEMPLATE**

|                   |           |           |           |
|-------------------|-----------|-----------|-----------|
| DAE 9 Bus 1/Enc 4 | RAIDS 4+1 | RAIDS 4+1 | RAIDS 4+1 |
| DAE 0 Bus 0/Enc 4 | Hot Spare | RAIDS 4+1 | RAIDS 8+1 |
| DAE 0 Bus 0/Enc 3 | RAIDS 4+1 | RAIDS 4+1 | RAIDS 4+1 |
| DAE 0 Bus 0/Enc 3 | Hot Spare | RAIDS 4+1 | RAIDS 8+1 |
| DAE 0 Bus 0/Enc 2 | RAIDS 4+1 | RAIDS 4+1 | RAIDS 4+1 |
| DAE 0 Bus 0/Enc 2 | Hot Spare | RAIDS 4+1 | RAIDS 8+1 |
| DAE 0 Bus 0/Enc 1 | RAIDS 4+1 | RAIDS 4+1 | RAIDS 4+1 |
| DAE 0 Bus 0/Enc 1 | Hot Spare | RAIDS 4+1 | RAIDS 8+1 |
| DAE 0 Bus 0/Enc 0 | RAIDS 4+1 | RAIDS 4+1 | RAIDS 4+1 |
| DAE 0 Bus 0/Enc 0 | RAIDS 4+1 | Hot Spare | RAIDS 8+1 |

**CX ALL 4PLUS1 RAID5 TEMPLATE**

|                   |           |           |  |
|-------------------|-----------|-----------|--|
| DAE 9 Bus 1/Enc 4 | RAIDS 4+1 | RAIDS 4+1 | RAIDS 4+1                              |
| DAE 0 Bus 0/Enc 4 | RAIDS 4+1 | RAIDS 4+1 | RAIDS 4+1                              |
| DAE 0 Bus 0/Enc 3 | RAIDS 4+1 | RAIDS 4+1 | RAIDS 4+1                              |
| DAE 0 Bus 0/Enc 3 | RAIDS 4+1 | RAIDS 4+1 | RAIDS 4+1                              |
| DAE 0 Bus 0/Enc 2 | RAIDS 4+1 | RAIDS 4+1 | RAIDS 4+1                              |
| DAE 0 Bus 0/Enc 2 | RAIDS 4+1 | RAIDS 4+1 | RAIDS 4+1                              |
| DAE 0 Bus 0/Enc 1 | RAIDS 4+1 | RAIDS 4+1 | RAIDS 4+1                              |
| DAE 0 Bus 0/Enc 1 | RAIDS 4+1 | RAIDS 4+1 | RAIDS 4+1                              |
| DAE 0 Bus 0/Enc 0 | RAIDS 4+1 | RAIDS 4+1 | RAIDS 4+1                              |
| DAE 0 Bus 0/Enc 0 | RAIDS 4+1 | Hot Spare | RAIDS 4+1 [Optional HS + HS + HS + HS] |

**CX STANDARD RAID1 TEMPLATE**

|                   |             |           |           |       |       |       |       |
|-------------------|-------------|-----------|-----------|-------|-------|-------|-------|
| DAE 9 Bus 1/Enc 4 | Optional HS | RAID1     | RAID1     | RAID1 | RAID1 | RAID1 | RAID1 |
| DAE 0 Bus 0/Enc 4 | Hot Spare   | RAID1     | RAID1     | RAID1 | RAID1 | RAID1 | RAID1 |
| DAE 0 Bus 0/Enc 3 | Optional HS | RAID1     | RAID1     | RAID1 | RAID1 | RAID1 | RAID1 |
| DAE 0 Bus 0/Enc 3 | Hot Spare   | RAID1     | RAID1     | RAID1 | RAID1 | RAID1 | RAID1 |
| DAE 0 Bus 0/Enc 2 | Optional HS | RAID1     | RAID1     | RAID1 | RAID1 | RAID1 | RAID1 |
| DAE 0 Bus 0/Enc 2 | Hot Spare   | RAID1     | RAID1     | RAID1 | RAID1 | RAID1 | RAID1 |
| DAE 0 Bus 0/Enc 1 | Optional HS | RAID1     | RAID1     | RAID1 | RAID1 | RAID1 | RAID1 |
| DAE 0 Bus 0/Enc 1 | Hot Spare   | RAID1     | RAID1     | RAID1 | RAID1 | RAID1 | RAID1 |
| DAE 0 Bus 0/Enc 0 | Optional HS | RAID1     | RAID1     | RAID1 | RAID1 | RAID1 | RAID1 |
| DAE 0 Bus 0/Enc 0 | RAIDS 4+1   | Hot Spare | RAIDS 8+1 |       |       |       |       |

**Note:** RAID5 & RAID1 Legacy Templates are intentionally left out.

**LEGACY TEMPLATES:**

CX\_Standard\_Raid\_5

CX\_Standard\_Raid\_1

**ATA ENCLOSURE TEMPLATE (DAE2-ATA)**

|                   |             |           |           |
|-------------------|-------------|-----------|-----------|
| DAE2-ATA          | Optional HS | RAIDS 6+1 | RAIDS 6+1 |
| DAE 0 Bus 0/Enc 0 | RAIDS 4+1   | Hot Spare | RAIDS 8+1 |

→ ATA disks first supported NAS 5.1.18.8 and higher

→ Add ATA enclosures per the above template, with exception of first Shelf, which is system shelf

→ Integrated Celerra systems configure all (15) disks in ATA Enclosure

→ Gateway Celerra NS600G/NS700G systems do not require all (15) disks per ATA Enclosure

**#/nas/sbin/setup\_clariion -init** [Command used to setup additional storage devices]

**\$nas\_storage -i -all** [Details of RAID groups, LUNs in decimal, SP]

**\$nas\_storage -info id=1l more**

### **VERIFYING NBS SERVICES ON DM & CS:**

**# cat /etc/nbs.conf** [NS700 system]

```
#  
# Simple configuration file for nbs service  
#  
# Format: devIndex:NbsId:Host1,host2,...:  
0:1:server_2,server_3:  
4:5:server_2,server_3:  
5:6:server_2,server_3:
```

**# cat /etc/nbs.conf** [NS600G system]

```
#  
# Simple configuration file for nbs service  
#  
# Format: devIndex:NbsId:Host1,host2,...:  
0:1:server_2,server_3b,server_2b,server_3:  
4:5:server_2,server_3b,server_2b,server_3:  
5:6:server_2,server_3b,server_2b,server_3:
```

**# cat /nas/server/slot\_2/nbs.cs**

```
volume disk NBS1 c0t0l0  
volume disk NBS5 c0t0l4  
volume disk NBS6 c0t0l5  
volume disk RDF7 c0t0l6  
volume disk RDF10 c0t0l9  
volume disk NBS1 c16t0l0  
volume disk NBS5 c16t0l4  
volume disk NBS6 c16t0l5  
volume disk RDF7 c16t0l6  
volume disk RDF10 c16t0l9  
nbsid add nbsid=1 vol=NBS1 exclusive raw share  
exportnbs add ip=192.168.1.100 nbsid=1  
nbsid add nbsid=5 vol=NBS5 exclusive raw share  
exportnbs add ip=192.168.1.100 nbsid=5  
nbsid add nbsid=6 vol=NBS6 exclusive raw share  
exportnbs add ip=192.168.1.100 nbsid=6  
nbsid add nbsid=7 vol=RDF7 exclusive raw ro=192.168.1.100:192.168.1.101  
exportnbs add ip=192.168.1.100 nbsid=7  
nbsid add nbsid=10 vol=RDF10 exclusive raw ro=192.168.1.100:192.168.1.101  
exportnbs add ip=192.168.1.100 nbsid=10  
nbs start
```

**# /sbin/service nbs status | nbs stop | nbs start | restart**

Configured NBS devices:

```
254: 0   NbsId: 1  
Disk_Name: nda  
Capacity: 11534272 KB  
Crnt_Server: 0X302a8c0  
Server_List: 0X201a8c0 0X302a8c0 0X202a8c0 0X301a8c0  
num_io: 8498    time 99953250 ms num_sect 41430    que 0  
Avg_blk: 2 K    throughput 0 K/s    resp_tm: 4 ms, ops/sec 0  
254: 16   NbsId: 5  
Disk_Name: nde  
Capacity: 2097088 KB  
Crnt_Server: 0X302a8c0  
Server_List: 0X201a8c0 0X302a8c0 0X202a8c0 0X301a8c0  
num_io: 2387826   time 99953260 ms num_sect 43013540   que 0  
Avg_blk: 9 K    throughput 215 K/s    resp_tm: 1 ms, ops/sec 23  
254: 20   NbsId: 6  
Disk_Name: ndf
```

Capacity: 2097088 KB

Crnt\_Server: 0X302a8c0

Server\_List: 0X201a8c0 0X302a8c0 0X202a8c0 0X301a8c0

num\_io: 19271 time 99953270 ms num\_sect 1183644 que 0

Avg\_blk: 30 K throughput 5 K/s resp\_tm: 46 ms, ops/sec 0

### # /sbin/fdisk -l

Disk /dev/nda: 255 heads, 63 sectors, 1435 cylinders

Units = cylinders of 16065 \* 512 bytes

| Device | Boot | Start | End | Blocks | Id | System |
|--------|------|-------|-----|--------|----|--------|
|--------|------|-------|-----|--------|----|--------|

|           |   |   |    |        |   |       |
|-----------|---|---|----|--------|---|-------|
| /dev/nda1 | * | 1 | 17 | 136521 | 6 | FAT16 |
|-----------|---|---|----|--------|---|-------|

Disk /dev/nde: 255 heads, 63 sectors, 261 cylinders

Units = cylinders of 16065 \* 512 bytes

| Device | Boot | Start | End | Blocks | Id | System |
|--------|------|-------|-----|--------|----|--------|
|--------|------|-------|-----|--------|----|--------|

|           |   |     |          |    |       |
|-----------|---|-----|----------|----|-------|
| /dev/nde1 | 1 | 230 | 1847443+ | 83 | Linux |
|-----------|---|-----|----------|----|-------|

Disk /dev/ndf: 255 heads, 63 sectors, 261 cylinders

Units = cylinders of 16065 \* 512 bytes

| Device | Boot | Start | End | Blocks | Id | System |
|--------|------|-------|-----|--------|----|--------|
|--------|------|-------|-----|--------|----|--------|

|           |   |     |          |    |       |
|-----------|---|-----|----------|----|-------|
| /dev/ndf1 | 1 | 230 | 1847443+ | 83 | Linux |
|-----------|---|-----|----------|----|-------|

### # ls /proc/driver/nd/devices

1 5 6

### # /usr/local/bin/nd-cfg

Usage: nd-cfg [-a action] [-d dev\_name] [-i NbsId | -l] [-s server\_name]

nd-cfg -i NbsId -s server\_name

nd-cfg -l -s server\_name

nd-cfg -a addServ -s server\_name -d /dev/nda

nd-cfg -a remServ -s server\_name -d /dev/nda

nd-cfg -a setRetry\_# -d /dev/nda

nd-cfg -a stop -d /dev/nda

nd-cfg -a disconnect -s server\_name -d /dev/nda

### # /usr/local/bin/nd-cfg -a disconnect -s server\_2 -d /dev/ndj (example of removing NBS connection to RDF volume)

#### Adding NBS devices to DM Memory:

/nasmed/sbin/t2tty -c 2 "nbs add nbsid=1 vol=NBS1 rw=192.168.1.100:192.168.1.101:192.168.2.100:192.168.2.101 exclusive raw share"

/nasmed/sbin/t2tty -c 2 "nbs add nbsid=5 vol=NBS5 rw=192.168.1.100:192.168.1.101:192.168.2.100:192.168.2.101 exclusive raw share"

/nasmed/sbin/t2tty -c 2 "nbs add nbsid=6 vol=NBS6 rw=192.168.1.100:192.168.1.101:192.168.2.100:192.168.2.101 exclusive raw share"

#.server\_config server\_2 -v "nbs add nbsid=10 vol=RDF10 exclusive raw rw=192.168.1.100:192.168.1.101

:192.168.2.100:192.168.2.101" (adding NBS connection to RDF vol. as RW Server\_2)

# /usr/local/bin/nd-cfg -a disconnect -s server\_2 -d /dev/nda

# /usr/local/bin/nd-cfg -a disconnect -s server\_2b -d /dev/nda

# /usr/local/bin/nd-cfg -a addServ -s server\_2 -d /dev/nda

# /usr/local/bin/nd-cfg -a addServ -s server\_2b -d /dev/nda [repeat for other servers]

# /nas/bin/nas\_storage -i APM00033500569

id = 1

arrayname = B-APM00033500569

name = APM00033500569

type = Clarion

model\_type = RACKMOUNT

**model\_num = 600**

db\_sync\_time = 1104849786 == Tue Jan 4 09:43:06 EST 2005

API\_version = V5.4-532

num\_disks = 30

num\_devs = 11

num\_pdevs = 1

num\_storage\_grps = 1

num\_raid\_grps = 6

cache\_page\_size = 8

```
wr_cache_mirror = True  
low_watermark = 60  
high_watermark = 80  
unassigned_cache = 0  
failed_over = False
```

**captive\_storage = False**

**# .server\_config server\_2 -v "fcp bind show"**

\*\*\* Persistent Binding Table \*\*\*

```
Chain 0000: WWN 500601601060172f HBA 0 SP-a0 Bound  
Chain 0016: WWN 500601681060172f HBA 1 SP-b0 Bound  
Chain 0032: WWN 5006048000000000 HBA 2 N_PORT Bind Pending  
Chain 0048: WWN 5006048000000000 HBA 3 N_PORT Bind Pending  
Existing CRC: 719179e2, Actual: 719179e2, CRC Matchs
```

\*\*\* Dynamic Binding Table \*\*\*

```
Chain 0000: WWN 500601601060172f HBA 0 ID 0 Inx 00:00 Pid 0000 S_ID 0000ef Sys  
Chain 0016: WWN 500601681060172f HBA 1 ID 1 Inx 01:00 Pid 0016 S_ID 0000ef Non  
Chain 0032: WWN 0000000000000000 HBA 2 ID 2 Inx 02:81 Pid 0032 S_ID 000000 Non  
Chain 0048: WWN 0000000000000000 HBA 3 ID 3 Inx 03:81 Pid 0048 S_ID 000000 Non
```

FCP Base Chain: 0 Dump Slot: 2 Dump Chain: 0 16

Adapter Chain Offset 0:0 1:16 2:32 3:48 dumpInit 1  
1105206957: ADMIN: 4: Command succeeded: fcp bind show

**#.server\_config server\_2 -v "camshowconfig"**

CAM Devices on scsi-0:

TID 00: 0:d0+ 1:d1+ 2:d2+ 3:d3+ 4:d4+ 5:d5+

TID 01: 0:d6+ 1:d7- 2:d8+ 3:d9-

CAM Devices on scsi-16:

TID 00: 0:d10- 1:d11- 2:d12- 3:d13- 4:d14- 5:d15-

TID 01: 0:d16- 1:d17+ 2:d18- 3:d19+

1105206963: ADMIN: 4: Command succeeded: camshowconfig

**Note:** NBS must be configured on Chains 0 & 16 on direct-connect captive systems or else NBS may not function correctly. With Fabric-connected systems, NBS can be on Chains 0, 16, 32 & 48.

## **CONFIGURING NBS SERVICES MANUALLY ON DATA MOVERS:**

**/nasmcd/sbin/setup\_slot -nbs -add 2**

## **VERIFYING NBS VOLUMES ON NS600 SYSTEM:**

**# /sbin/service --status-all**

Configured NBS devices:

**254: 0 NbsId: 1**

```
Disk_Name: nda  
Capacity: 4194240 KB  
Crnt_Server: 0X201a8c0  
Server_List: 0X301a8c0 0X201a8c0  
num_io: 27211 time 1201536430 ms num_sect 390684  
Avg_blk: 7 K throughput 0 M/s resp_tm: 1 ms, ops/sec 0
```

**254: 16 NbsId: 5**

```
Disk_Name: nde  
Capacity: 2097088 KB  
Crnt_Server: 0X201a8c0  
Server_List: 0X301a8c0 0X201a8c0  
num_io: 178982 time 1201536440 ms num_sect 2848996  
Avg_blk: 7 K throughput 0 M/s resp_tm: 5 ms, ops/sec 0
```

**254: 20 NbsId: 6**

```
Disk_Name: ndf  
Capacity: 2097088 KB  
Crnt_Server: 0X201a8c0  
Server_List: 0X301a8c0 0X201a8c0  
num_io: 96341 time 1201536450 ms num_sect 5731596  
Avg_blk: 29 K throughput 0 M/s resp_tm: 11 ms, ops/sec 0
```

## **NBS SERVICES ON NS600G WITH CX400 BACKEND:**

### **# /sbin/service nbs status**

Configured NBS devices:

#### **254: 0 NbsId: 1**

Disk\_Name: nda  
Capacity: 11534272 KB  
Crnt\_Server: 0X201a8c0  
Server\_List: 0X201a8c0 0X302a8c0 0X202a8c0 0X301a8c0  
num\_io: 323 time 564970 ms num\_sect 1709 que 0  
Avg\_blk: 2 K throughput 1 K/s resp\_tm: 6 ms, ops/sec 0

#### **254: 16 NbsId: 5**

Disk\_Name: nde  
Capacity: 2097088 KB  
Crnt\_Server: 0X201a8c0  
Server\_List: 0X201a8c0 0X302a8c0 0X202a8c0 0X301a8c0  
num\_io: 6235 time 564970 ms num\_sect 118848 que 0  
Avg\_blk: 9 K throughput 105 K/s resp\_tm: 2 ms, ops/sec 11

#### **254: 20 NbsId: 6**

Disk\_Name: ndf  
Capacity: 2097088 KB  
Crnt\_Server: 0X201a8c0  
Server\_List: 0X201a8c0 0X302a8c0 0X202a8c0 0X301a8c0  
num\_io: 10 time 564970 ms num\_sect 152 que 0  
Avg\_blk: 7 K throughput 0 K/s resp\_tm: 1 ms, ops/sec 0

## **UPDATING CACHE CONFIGURATIONS ON CLARIION AFTER CODE UPGRADES:**

### **#/nas/sbin/setup\_backend/nas\_raid cache configure**

**Note:** NAS 5.1M and higher recommends use of Read Cache to improve performance [32MB to 256MB max]

## **MONITORING DMs AS THEY BOOT:**

#cat /dev/ttyS4 [DM2; ttyS5=DM3] >server2 |tail -f

## **NAVICLI COMMANDS:**

**#/nas/sbin/navicli -h 192.168.1.200 getagent | getcrus |getdisk | getrg | getlun | getcache | getlog**

## **CABLING BETWEEN SPE & SPS:**

Wrong cabling may disable cache & make DMs appear to be reset or not booted from CS

**#/nas/sbin/navicli -h 192.168.1.200 getcrus** [Checks cabling status, SPS status, Disk status, SPA status, and shows faults]  
If DM LED is green, probably in contacted state

## **VERIFYING FLARE CODE:**

**#/nas/sbin/navicli -h 192.168.1.200 ndu -list | getagent**

Name of the software package: ManagementUI  
Revision of the software package: 6.4.1.0.0  
Name of the software package: Base  
Revision of the software package: 02.04.1.60.5.007  
Name of the software package: ManagementServer  
Revision of the software package: 6.4.0.5.2  
Name of the software package: Navisphere  
Revision of the software package: 6.4.0.5.2  
Commit Required: NO  
Revert Possible: NO  
Active State: YES  
Dependent packages: Base 02.04.1.60.5.007, Base 02.04.0.60.5.00  
7, ManagementServer 6.4.0.5.2  
Required packages: Base 21 OR 116  
Is installation completed: YES  
Is this System Software: NO

**\$ /nas/sbin/navicli -h 10.241.169.35 ndu -status**

Is Completed: YES

Status: Operation completed successfully

Operation: Install

**\$ /nas/sbin/navicli -h 10.241.169.35 storagegroup -status**

Data Access control: ENABLED [Access Logix is enabled on the backend]

## **CELERRA/CLARIION TROUBLESHOOTING:**

**#/nas/sbin/setup\_backend/setup\_clariion2 list config APM00034202048** [Shows overall status of Backend]

System 10.32.235.235 is up

System 10.32.235.239 is up

Clariion Array: APM00034202048 Model: CX600 Memory: 4022

Enclosure(s) 0\_0,1\_0,0\_1,1\_1,0\_2,1\_2,0\_3,1\_3,0\_4,1\_4,0\_5 are installed in the system.

Enclosure info:

|                    |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|--------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 0                  | 1   | 2   | 3   | 4   | 5   | 6   | 7   | 8   | 9   | 10  | 11  | 12  | 13  | 14  |
| -----              |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| 0_5:               | 146 | 146 | 146 | 146 | 146 | 146 | 146 | 146 | 146 | 146 | 146 | 146 | 146 | 146 |
|                    | 16  | 16  | 16  | 16  | 16  | 16  | 16  | 16  | 16  | 16  | 16  | 16  | 16  | HS  |
| -----abridged----- |     |     |     |     |     |     |     |     |     |     |     |     |     |     |

Lun info:

|                     |              |            |
|---------------------|--------------|------------|
| Lun ID: 0 RG ID: 0  | State: Bound | ??         |
| Lun ID: 1 RG ID: 1  | State: Bound | ??         |
| Lun ID: 2 RG ID: 2  | State: Bound | ??         |
| Lun ID: 3 RG ID: 3  | State: Bound | ??         |
| Lun ID: 4 RG ID: 20 | State: Bound | root_disk  |
| Lun ID: 5 RG ID: 20 | State: Bound | root_ldisk |
| Lun ID: 6 RG ID: 20 | State: Bound | roottd1000 |
| Lun ID: 7 RG ID: 20 | State: Bound | roottd1001 |
| Lun ID: 8 RG ID: 20 | State: Bound | roottd1002 |
| Lun ID: 9 RG ID: 20 | State: Bound | roottd1003 |

**\$ /nas/sbin/setup\_backend/zone -spa 10.241.168.55 -spb 10.241.168.56 showsp** (checks IP connectivity & collects WWN's SPs)

Checking network connectivity to 10.241.168.55 / 10.241.168.56 ...Done.

Collecting backend WWNs...Done.

SP List:

SP Port:A0 WWN:50:06:01:60:90:60:05:10:50:06:01:60:10:60:05:10  
SP Port:A1 WWN:50:06:01:60:90:60:05:10:50:06:01:61:10:60:05:10  
SP Port:A2 WWN:50:06:01:60:90:60:05:10:50:06:01:62:10:60:05:10  
SP Port:A3 WWN:50:06:01:60:90:60:05:10:50:06:01:63:10:60:05:10  
SP Port:B0 WWN:50:06:01:60:90:60:05:10:50:06:01:68:10:60:05:10  
SP Port:B1 WWN:50:06:01:60:90:60:05:10:50:06:01:69:10:60:05:10  
SP Port:B2 WWN:50:06:01:60:90:60:05:10:50:06:01:6a:10:60:05:10  
SP Port:B3 WWN:50:06:01:60:90:60:05:10:50:06:01:6b:10:60:05:10

**# /nas/sbin/navisecli -h 192.1.4.221 rebootpeersp** [Rebooting SP's peer]

**# /nas/sbin/navisecli -h 192.1.4.221 rebootSP** [Rebooting SP specified by IP address]

**# /nas/sbin/navicli -h 192.168.1.200 ndu -list** [Shows Base code and SW Packages installed]

Commit Required: NO →Normal state—YES would mean that base has not been committed

**#/nas/sbin/navicli -h 192.168.1.200 -f all.txt getall**

**#/nas/sbin/navicli -h 192.168.1.200 getall >getall.txt** [Equivalent commands]

**Note:** Getall gathers configuration information on backend

**# /nas/sbin/navicli -h 192.168.1.200 inserttestevent**

**Note:** Creating test event on Clariion backend—Good way to test to see if Celerra Calls Home—see Primus emc125099 for issues on CallHomes with Clariion backends and NAS 5.3/5.4

**# /nas/sbin/navicli -h 192.168.1.200 getdisk >disk.txt**

**# navicli -h 192.189.20.16 getdisk -state -numluns -rg -capacity -lun -rb -rev -serial -type -product -vendor -hs >getdisk\_sp.log**

# /nas/sbin/navicli -h 192.168.1.200 getdisk -state

```
Bus 0 Enclosure 0 Disk 0
State: Enabled
Bus 0 Enclosure 0 Disk 1
State: Enabled
Bus 0 Enclosure 0 Disk 2
State: Enabled
Bus 0 Enclosure 0 Disk 3
State: Enabled
Bus 0 Enclosure 0 Disk 4
State: Enabled
Bus 0 Enclosure 0 Disk 5
State: Hot Spare Ready
Bus 0 Enclosure 0 Disk 6
State: Unbound
```

# /nas/sbin/navicli -h spa getdisk 1\_1\_12 (getting status of particular disk)

```
Bus 1 Enclosure 1 Disk 12
Vendor Id: ATA-MXTR
Product Id: 5A320J0 CLAR320
Product Revision: WV
Lun: 26 28
Type: 26: RAID5 28: RAID5
State: Enabled
Hot Spare: 26: NO 28: NO
Prct Rebuilt: 26: 100 28: 100
Prct Bound: 26: 100 28: 100
Serial Number: A82Z1NGE
Sectors: 533887414 (260687)
Capacity: 304169
Private: 26: 69760 28: 1868675968
Bind Signature: 0xb0fd, 1, 12
Hard Read Errors: 0
Hard Write Errors: 0
Soft Read Errors: 0
Soft Write Errors: 0
Read Retries: 0
Write Retries: 0
Remapped Sectors: N/A
Number of Reads: 0
Number of Writes: 0
Number of Luns: 2
Raid Group ID: 13
Clariion Part Number: DG118032260
Request Service Time: 0 ms
Read Requests: 0
Write Requests: 0
Kbytes Read: 0
Kbytes Written: 0
Stripe Boundary Crossing: 0
```

### **DETERMINING DISK DRIVE TYPES ON CLARIION:**

# /nas/sbin/navicli -h 10.241.168.52 getdisk -tla

```
Bus 0 Enclosure 0 Disk 0
Clariion TLA Part Number:005047892
Bus 0 Enclosure 0 Disk 1
Clariion TLA Part Number:005048129
```

# /nas/sbin/navicli -h 192.168.1.200 getlun 1 [or -owner]

# /nas/sbin/navicli -h 10.241.168.57 getlun -trespass -status

LOGICAL UNIT NUMBER 92

Default Owner: SP B

Current owner: SP A

Device Map: Valid

**# /nas/sbin/navicli -h 10.241.168.57 getlun 92 |egrep "Default|Current"**

Current owner: SP A

Default Owner: SP B

**# /nas/sbin/navicli -h 10.241.168.57 getlun 0 –capacity | getlun –capacity** (all luns on system)

LUN Capacity(Megabytes): 11264

LUN Capacity(Blocks): 23068672

**# /nas/sbin/navicli -h 10.241.168.57 getlun –bind** (good check from CLI on LUN Bind status)

LOGICAL UNIT NUMBER 43

Prct Bound: 100

**# /nas/sbin/navicli -h 192.168.1.200 getrg**

**# /nas/sbin/navicli -h 10.241.168.52 getrg 0**

RaidGroup ID: 0

RaidGroup Type: r5

RaidGroup State: Explicit\_Remove

Valid\_luns

List of disks: Bus 0 Enclosure 0 Disk 0

Bus 0 Enclosure 0 Disk 1

Bus 0 Enclosure 0 Disk 2

Bus 0 Enclosure 0 Disk 3

Bus 0 Enclosure 0 Disk 4

6 7 9

List of luns: Max Number of disks: 16

Max Number of luns: 128

Raw Capacity (Blocks): 633262480

Logical Capacity (Blocks): 506609984

Free Capacity (Blocks,non-contiguous): 456278336

Free contiguous group of unbound segments: 288506176

Defrag/Expand priority: Low

Percent defragmented: 100

Percent expanded: 100

Disk expanding onto: N/A

Lun Expansion enabled: NO

Legal RAID types: r5

**# /nas/sbin/navicli -h 192.168.1.200 getlog >spa.log** [Repeat for SPB]

**Note:** Getlogs are single most important event log file for SPs

**# /nas/sbin/navicli -h SPB getlun -default -owner | getlun –state –owner –default**

**# /nas/sbin/navicli -h 192.168.1.200 getlun –aa -state**

**# /nas/sbin/navicli -h 192.168.1.200 getcrus** [Check for faults on backend]

DAE2 Bus 0 Enclosure 0 \*FAULT\*

(Bus 0 Enclosure 0 : Faulted; Bus 0 Enclosure 0 Disk 3 : Removed)

**# /nas/sbin/navicli -h spa faults -list**

The array is operating normally.

**Output of Specific Disk after Replacing:**

**# /nas/sbin/navicli -h 10.241.168.52 getdisk 0\_2\_0 |grep State** →Bus 0, Enclosure 2, Disk 0

Equalizing →Transition state

**# /nas/sbin/navicli -h 10.241.168.52 getdisk 0\_2\_0 |grep State**

Enabled →Normal state

**# /nas/sbin/navicli -h 192.168.1.200 getagent** [Verify agents are running—look for peer signature value]

**# /nas/sbin/navicli -h 10.241.168.57 getagent -serial -spid**

Serial No: APM00023700172

SP Identifier: A

**# /nas/sbin/navicli -h 10.241.168.57 getsptime**

Time on SP A: 06/20/06 13:23:18

Time on SP B: 06/20/06 13:23:18

**# /nas/sbin/navicli -h spa getspt**

SP A

Type of Each SP: SPE  
 Signature For The SP: 938260  
 Signature For The Peer SP: 918890  
 Revision Number For The SP: 2.06.700.4.004  
 Serial Number For The SP: LKE00033800184  
 Memory Size For The SP: 3968  
 SP SCSI ID if Available: 0

**# /nas/sbin/navicli -h 10.241.168.57 getcontrol**

System Fault LED: OFF  
 Statistics Logging: ON  
 SP Read Cache State Enabled  
 SP Write Cache State Enabled  
 Max Requests: N/A  
 Average Requests: N/A  
 Hard errors: N/A  
 Total Reads: 102125  
 Total Writes: 1752819  
 Prct Busy: 2.18  
 Prct Idle: 97.8  
 System Date: 06/20/2006  
 Day of the week: Tuesday  
 System Time: 13:23:51  
 Read\_requests: 102125  
 Write\_requests: 1752819  
 Blocks\_read: 1974568  
 Blocks\_written: 34456101-----output abridged-----

**# /nas/sbin/navicli -h 192.168.1.200 storagegroup –list** [Outputs only if Access Logix installed]

Storage Group Name: Celerra\_NAS\_CS  
 Storage Group UID: 86:C0:4A:98:8A:60:D8:11:80:13:B9:62:AA:B2:26:83  
 HBA/SP Pairs:  
 HBA UID SP Name SPPort

-----  
 20:00:00:00:C9:23:47:3B:10:00:00:00:C9:23:47:3B SP A 0  
 20:00:00:00:C9:23:47:3B:10:00:00:00:C9:23:47:3B SP B 0

HLU/ALU Pairs:  
 HLU Number ALU Number  
 -----  
 0 5  
 1 4  
 2 3  
 3 2  
 4 1  
 5 0

Shareable: YES

**# /nas/sbin/navicli -h 10.241.168.57 getall –sg** (same output as storagegroup –list)**# /nas/sbin/navicli -h 10.241.168.57 -np getdisk**

**Note:** Use -np option to retrieve object info from cache and not directly query each object on array—faster

**-NP PERL SCRIPT FOR NAVICLI (From AR106276, emc175499):**

**Note:** Following script can be used to replace the symbolic links /nas/sbin/navicli and /nas/opt/Navisphere/bin/navicli, which point to nas\_navi, with a file named “navicli”—just copy the contents to notepad and then into a new file called navicli and copy to the two locations after removing the symbolic links. Script enforces the use of -np switch on array commands, retrieving objects from cache whenever possible as opposed to querying every object, making navi commands work faster.

#!/usr/bin/perl -w

```
#
# This is a hack supplied by EMC development to make sure the
# -np option is applied to every navicli command
#
# This perl script should replace the symlink navicli -> nas_navi
# in the /nas/sbin and /nas/opt/Navisphere/bin directories
#
$navi = "/nasmcd/sbin/classic_navicli";
$cmd = join(" ", @ARGV);
if( $cmd =~ m/-np/ ) {
  @out = `$navi $cmd`;
}
```

```

else {
    @out = `$navi -np $cmd`;
}

$status=$? >> 8;

print join("", @out);

exit $status;
# /nas/sbin/navicli -h 192.168.1.200 arraycommpath [Default Setting for NAS/SAN Celerra config is 0]
Current arraycommpath setting is: 0
# /nas/sbin/navicli -h 192.168.1.200 ndu -status [Checking status of upgrades in progress]
# /nas/sbin/navicli -h 192.168.1.200 ndu -commit Base [use this command to commit base code after upgrades]
# ps -ef |grep nd-clnt [Verify that NBS Services are running on Data Movers]
root  1159  1 0 Mar09 ?  00:00:00 [nd-clnt 0 1]
root  1158  1 0 Mar09 ?  00:00:07 [nd-clnt 4 5]
root  1166  1 0 Mar09 ?  00:00:01 [nd-clnt 5 6]
# mount [Following directories will not mount without at least one DM running NBS Services!]
/dev/nde1 on /nbsnas type ext3 (rw,sync)
/dev/hda5 on /nas type ext3 (rw,sync)
/dev/nda1 on /nas/dos type msdos (rw,sync,umask=002,gid=201)
/dev/ndf1 on /nas/var type ext3 (rw)
# nas_storage -i -all
```

**Note:** Extremely useful information about Array, SP's, and Raid Groups

```

id          = 1
arrayname   = APM00040601637
name        = APM00040601637
model_type  = RACKMOUNT
model_num   = 600
db_sync_time = 1093614587 == Fri Aug 27 09:49:47 EDT 2004
num_disks   = 37
num_devs    = 19
num_pdevs   = 1
num_storage_grps = 0
num_raid_grps = 7
cache_page_size = 8
wr_cache_mirror = True
low_watermark = 70
high_watermark = 90
unassigned_cache = 0
failed_over   = False
captive_storage = True
```

### Storage Processors

```

SP Identifier = A
signature     = 678828
microcode_version = 2.06.600.5.003
serial_num    = LKE00023511180
prom_rev      = 3.26.00
agent_rev     = 6.6.0 (3.1)
phys_memory   = 2048
sys_buffer    = 551
read_cache    = 32
write_cache   = 1465
free_memory   = 0
raid3_mem_size = 0
failed_over   = False
num_disk_volumes = 10 - root_disk root_ldisk d3 d4 d5 d6 d7 d9 d10 d11
```

### Disk Groups

```

id          = 0000
raid_type   = RAID5
logical_capacity = 1068997528
num_spindles = 5 - 0_0_0_0_0_1 0_0_2 0_0_3 0_0_4
num_luns    = 8 - 0000 0001 0002 0003 0004 0005 0016 0017
num_disk_volumes = 8 - root_disk root_ldisk d3 d4 d5 d6 d7 d12
spindle_type = FC
raw_capacity = 1336246910
used_capacity = 1068996608
free_capacity = 920
hidden       = False
```

## GETTING RESUME INFORMATION ON THE CLARIION:

# /nas/sbin/navicli -h 192.1.4.251 getresume -messner|more

SP A

|                          |                |
|--------------------------|----------------|
| EMC Part Number:         | 100-520-624    |
| EMC Artwork Revision:    | N/A            |
| EMC Assembly Revision:   | A03            |
| EMC Serial Number:       | CF2VZ073600136 |
| Vendor Part Number:      | N/A            |
| Vendor Artwork Number:   | N/A            |
| Vendor Assembly Number:  | N/A            |
| Vendor Serial Number:    |                |
| Vendor Name:             | CELESTICA      |
| Location of Manufacture: | THAILAND       |
| Year of Manufacture:     | 2007           |
| Month of Manufacture:    | 9              |
| Day of Manufacture:      | 7              |

**Assembly Name:** CX3-10F SAN-7  
Programmable Name: BIOS:POST:RESETPIC  
Programmable Revision: 3.58:1.44:0.21 →Output abridged

## CX3-40 ARRAY OUTPUT:

# /nas/sbin/navicli -h 192.168.1.200 getresume

SP A

|                          |                |
|--------------------------|----------------|
| EMC Part Number:         | 100-561-857    |
| EMC Artwork Revision:    | N/A            |
| EMC Assembly Revision:   | A05            |
| EMC Serial Number:       | CF2KR062900515 |
| Vendor Part Number:      | N/A            |
| Vendor Artwork Number:   | N/A            |
| Vendor Assembly Number:  | N/A            |
| Vendor Serial Number:    | N/A            |
| Vendor Name:             | CELESTICA      |
| Location of Manufacture: | THAILAND       |
| Year of Manufacture:     | 2006           |
| Month of Manufacture:    | 8              |
| Day of Manufacture:      | 16             |

**Assembly Name:** CX3-40 SAN/AUX  
Programmable Name: BIOS:POST:RESETPIC  
Programmable Revision: 3.47:1.35:0.21

## BACKENDBUS INFORMATION:

# /nas/sbin/navicli -h 10.241.168.150 backendbus -get -all

Bus 0  
Current Speed: 4Gbps.  
Available Speeds:  
    2Gbps.  
    4Gbps.  
SPA SFP State: Online  
SPB SFP State: Online  
I/O Module Slot: 0  
Physical Port ID: 0

## GETTING FRUMON CODE FROM BACKEND LCCs:

# /nas/sbin/navicli -h 192.168.1.200 getresume -messner -lcc

Bus 0 Enclosure 0  
LCC A

|                         |                          |
|-------------------------|--------------------------|
| EMC Part Number:        | 100-561-803              |
| EMC Artwork Revision:   | C04                      |
| EMC Assembly Revision:  | A05                      |
| EMC Serial Number:      | FCNBD063531833           |
| Vendor Part Number:     | N/A                      |
| Vendor Artwork Number:  | N/A                      |
| Vendor Assembly Number: | N/A                      |
| Vendor Serial Number:   | N/A                      |
| Vendor Name:            | FOXCONN, SHENZHEN, CHINA |

Location of Manufacture: LONGHUA TOWN, SHENZHEN, CHINA

Year of Manufacture: 2006

Month of Manufacture: 09

Day of Manufacture: 02

Assembly Name: 4Gb BoneSword LCC

Programmable Name: FRUMON

Programmable Revision: 7.66

**\$ /nas/sbin/navicli -h 192.168.1.200 getrus** →Also shows Frumon level on LCC cards

DAE2-ATA Bus 1 Enclosure 1

Bus 1 Enclosure 1 LCC A Revision: 1.93

Bus 1 Enclosure 1 LCC B Revision: 1.93

**# /nas/sbin/navicli -h 192.168.1.200 getresume -ps** →SPE & Enclosure Power Supply information

Enclosure SPE

Power A0

EMC Part Number: 071-000-462

EMC Artwork Revision: 000

EMC Assembly Revision: A01

EMC Serial Number: AC130062200798

Vendor Part Number: API4SG10-710L

Vendor Artwork Number: 000

Vendor Assembly Number: M00

Vendor Serial Number: AC1062200798

Vendor Name: ACBEL POLYTECH INC.

Location of Manufacture: MANILA, PHILIPPINES

Year of Manufacture: 2006

Month of Manufacture: 06

Day of Manufacture: 06

Assembly Name: 12V P/S w/FAN

Programmable Name: N/A

Programmable Revision: N/A

Bus 0 Enclosure 0

Power B

EMC Part Number: 071-000-453

EMC Artwork Revision: 000

EMC Assembly Revision: A02

EMC Serial Number: AC734063109370

Vendor Part Number: API4SG02-711L

Vendor Artwork Number: A02

Vendor Assembly Number: A02

Vendor Serial Number: AC7063104563

Vendor Name: ACBEL POLYTECH INC.

Location of Manufacture: TANG XIA TOWN, DONG GUAN, CHINA

Year of Manufacture: 2006

Month of Manufacture: 08

Day of Manufacture: 16

Assembly Name: Dual +12V P/S & Cooling Module

Programmable Name: N/A

Programmable Revision: N/A

## **DETERMINING HOT SPARE STATUS:**

**# /nas/sbin/navicli -h 10.241.169.35 getlun -default -owner**

LOGICAL UNIT NUMBER 6

Default Owner: SP B

Current owner:

**# /nas/sbin/navicli -h 10.241.169.35 getlun 6**

Name LUN 6

RAID Type: Hot Spare

## **PLACING DISK DRIVE IN FAULTED STATUS FOR TESTING:**

# /nas/sbin/navicli -h 192.168.1.200 getdisk 0\_0\_12 |grep Bus |grep -v Ticks

Bus 0 Enclosure 0 Disk 12

# /nas/sbin/navicli -h 192.168.1.200 getcrus (No faults noted)

# /nas/sbin/navicli -h 192.168.1.200 cru\_on\_off -messner 0\_0\_12 0 (this places disk 12 down)

# /nas/sbin/navicli -h 192.168.1.200 getcrus

DAE3P Bus 0 Enclosure 0 \*FAULT\*

(Bus 0 Enclosure 0 : Faulted; Bus 0 Enclosure 0 Disk 12 : Removed)

# /nas/sbin/navicli -h 192.168.1.200 cru\_on\_off -messner 0\_0\_12 1 [places drive back up]

### DETERMINING FRUMON VERSION ATA LCC's:

# /nas/sbin/navicli -h 192.1.4.220 getcrus -lcreva -lcrevb

SPE3 Enclosure SPE

DAE3P Bus 0 Enclosure 0

Bus 0 Enclosure 0 LCC A Revision: 7.71

Bus 0 Enclosure 0 LCC B Revision: 7.71

### VERIFYING SP PORTS FOR SAN-CONNECTED DM's:

\$ /nas/sbin/navicli -h 10.64.25.148 port -list

Information about each HBA:

HBA UID: 50:06:01:60:80:60:07:B9:50:06:01:68:00:60:07:B9

Server Name: ns600g\_dm30

Server IP Address: 192.168.1.3

HBA Model Description: ns600g

HBA Vendor Description: celerra

HBA Device Driver Name: N/A

Information about each port of this HBA:

SP Name: SP A

SP Port ID: 1

HBA Devicename: N/A

Trusted: NO

**Logged In:** YES

Source ID: 1

Defined: YES

Initiator Type: 3

StorageGroup Name: NS600G

### VERIFYING SP SPEED SETTINGS:

# /nas/sbin/navicli -h spa ssportspeed -get

Storage Processor : SP A

Port ID : 1

Speed Value : 2

Storage Processor : SP A

Port ID : 0

Speed Value : 2

Storage Processor : SP A

Port ID : 3

Speed Value : 2

Storage Processor : SP A

Port ID : 2

Speed Value : 2

### CHANGING SP FC PORT SPEED:

# /nas/sbin/navicli -h spa ssportspeed -set -sp a -portid 0 2 [Sets port 0 on SPA to 2GB]

### RUNNING SPCOLLECT TO GATHER CLARIION LOGS:

1. C:>Program Files\EMC\Navisphere CLI>navicli -h 192.168.1.200 spcollect -messner [password=messner]

**Note:** Script collects a variety of files and zips them.

2. Monitor spcollect using: >navicli -h 192.168.1.200 managefiles -list | -delete | -retrieve

3. Retrieve SPCollects:

C:>navicli -h 192.168.1.200 -retrieve -file SPA\_WRE00021500537\_90a1f\_07-24-2003\_14-25-05\_data.zip

### TRANSFERRING SPCOLLECT FROM SP TO HOST:

**C:>Program Files\EMC\Navisphere CLI>navicli -h 192.168.1.200 managefiles –retrieve <path> –file**

**SPA\_WRE00021500537\_90a1f\_07-24-2003\_14-25-05\_data.zip**

**Note:** SPCollects are the basic starting point files for any Clariion escalation and can be sent to following email address, along with Case #, Customer Name, and brief description of issue: GCSC@emc.com. Starting with FLARE 19, SPCOLLECTS can be setup to run automatically, so that there is always an SPCollect available. However, only a single collect is maintained and only records a finite amount of information—so if there is a lot of activity, the logs may have wrapped.

### **CONTROL STATION SCRIPT AUTOMATES SP COLLECT:**

**# /nas/tools/.get\_spcollect**

Generating spcollect zip file for Clariion(s)

Creating spcollect zip file for the Service Processor A\_APM00023801040. Please wait...

spcollect started to pull out log files(it will take several minutes)....

Wait until new \_data.zip file size becomes final(it will take several minutes)

Retrieving new \_data.zip file...

spcollect zip file APM00023801040\_SPA\_2006-04-10\_20-02-39\_a5c1b\_data.zip for the Service Processor A\_APM00023801040 was created

Creating spcollect zip file for the Service Processor B\_APM00023801040. Please wait... -----output abridged-----

Zipping all spcollect zip files in one SPCOLLECT.zip file and putting it in the

**/nas/var/log** directory...

adding: APM00023801040\_SPA\_2006-04-10\_20-02-39\_a5c1b\_data.zip (stored 0%)

adding: APM00023801040\_SPB\_2006-04-10\_20-06-50\_a34fc\_data.zip (stored 0%)

**-rw-r--r-- 1 root root 6073082 Apr 10 16:09 SPCOLLECT.zip**

### **RUNNING BACKGROUND VERIFY ON LUNS:** LRC checks

**# /nas/sbin/navicli -h 192.168.1.200 setsniffer 4 1 -bv -bvtimer asap**

**Note:** Above command runs a background verify against LUN 4 (/nas /dev/sde1), best method to check actual condition of LUNs

### **VERIFYING STATE OF LUNS ON BACKEND:**

**# /nas/sbin/navicli -h 10.241.168.57 getlun -trespass -status**

**\$ /nas/sbin/navicli -h 10.241.169.35 getlun -state -owner -default**

LOGICAL UNIT NUMBER 228

State: Bound

Current owner:

Default Owner: SP B

LOGICAL UNIT NUMBER 17

State: Bound

Current owner: SP B

Default Owner: SP B

-----abridged-----

### **VERIFYING STORAGEGROUPS ON ARRAY:** [Works only with Access Logix]

**\$ /nas/sbin/navicli -h 10.241.169.35 storagegroup -list**

Storage Group Name: NS600G

Storage Group UID: BC:C5:B1:A2:B7:42:D8:11:80:0A:91:05:4E:AD:DB:DC

#### **HBA/SP Pairs:**

| HBA UID   | SP Name | SPPort |
|---|---------|--------|
| 02:00:00:00:01:00:00:00:02:00:00:00:01:00:00:00 | SP A    | 1      |
| 50:06:01:60:90:60:15:50:50:06:01:68:10:60:15:50 | SP A    | 1      |
| 02:00:00:00:01:00:00:00:02:00:00:00:01:00:00:00 | SP A    | 0      |
| 50:06:01:60:90:60:15:50:50:06:01:60:10:60:15:50 | SP A    | 0      |
| 02:00:00:00:01:00:00:00:02:00:00:00:01:00:00:00 | SP B    | 1      |
| 50:06:01:60:90:60:15:50:50:06:01:69:10:60:15:50 | SP B    | 1      |
| 02:00:00:00:01:00:00:00:02:00:00:00:01:00:00:00 | SP B    | 0      |
| 50:06:01:60:90:60:15:50:50:06:01:61:10:60:15:50 | SP B    | 0      |

#### **HLU/ALU Pairs:**

| HLU Number | ALU Number |
|------------|------------|
| 0          | 0          |
| 1          | 1          |
| 2          | 2          |

3 3  
4 4  
5 5  
16 16  
17 17

Shareable: YES

## **COLLECTING CONFIGURATION INFO AFTER INSTALL:**

**# /nas/sbin/log\_config -outfile**

Usage: log\_config  
[ -outfile <filename> ] [ -default ] [ -nozip ] [ -callhome ]  
| -start  
| -stop  
| -version  
| -help

**#/nas/sbin/log\_config -outfile /home/nasadmin/config.log**

**# /nas/sbin/model**

**Note:** The model command runs /nas/bin/nas\_xml and looks at 2<sup>nd</sup> entry in PRODUCT\_NAME field to determine system Model

**# nas\_xml -info:ALL**

**# nas\_xml -info:server -level:3**

**Note:** Extremely useful command to obtain overall Celerra, CS, and Data Mover hardware and software information, as well as NIC status.

```
<CELERRA SRC='controlstation'>  
<CELERRA_MANAGEMENT_UNIT NAME='APM000421031830000' TYPE='Celerra IP'  
PRODUCT_NAME='Celerra NS600G' SERIAL_NO='APM123456789'  
DMSLOTS='2' CSSLOTS='1' CID='APM000421031830000'>  
<CONTROL_STATION HOSTNAME='Iprep1' VERSION='5.3.12-3' NXV='5.3.12-3'  
DATE_TIME_GMT='2005/01/13-21:31:26' DATE_TIME_LOCAL='2005/01/13-16:31:26'  
NAS_XML_LEVEL='3'>  
<DATA_MOVERS>  
<MOVER NAME="server_2" SLOT="2" ID="1" ACL="1000" TYPE="nas"  
PHYSICAL_HOST="server_2" STATUS="enabled" GROUP="" DART_VERSION="T5.3.14.2"  
MODEL="NS600" CPU_SPEED="2000" CPU="Intel Pentium 4" BUS_SPEED="400"  
MOTHERBOARD="Chameleon II XP" TOTAL_MEMORY="4024"  
-----abridged-----
```

## **VIEWING SP LOG:**

**#/nas/sbin/navicli -h 192.168.1.200 getlog >/tmp/cx600.log**

## **CREATING PPP [VPN—Virtual Private Network] CONNECTION TO CX600:**

1. Start>Settings>Network and Dial-up Connections>Make New Connection>Next>\*Guest>Connect directly to another computer
2. Select a Device: COM1>\*For all Users>Configure COM1: Max speed 115200; Enable hardware flow control
3. Finish>Enter Clariion clarion! >Connect
4. Once connected open Internet Explorer and enter IP Address to Service PC: http://192.168.1.1/setup
5. Enter IP Address, Hostname, Mask, Gateway, Peer IP Address, Management Ports
6. Next task would be to install the Navisphere CLI on the Service PC
7. Open cmd prompt to execute navicli commands to backend  
c:\Program Files\EMC\Navisphere CLI>navicli -h 192.168.25.114 ndu -list [Checks for software revision on SPs]

## **VERIFYING ACCESSLOGIX IS ENABLED ON CX600:**

**1. c:\Program Files\EMC\Navisphere CLI>navicli -h 192.168.25.114 storagegroup -status**

“Data Access Control Enabled” [Means AccessLogix is enabled]

2. c:\Program Files\EMC\Navisphere CLI>navicli -h 192.168.25.114 storagegroup -enable -o [Run this to enable AccessLogix]

## **CONFIGURING CX600:**

### **SETTING FIBRE PORT SPEEDS ON CX600 SPs:**

1. c:\Program Files\EMC\Navisphere CLI>navicli -h 192.168.25.114 spportspeed -set -sp a -portid 0 1 [Sets to 1GB/sec]

### **CREATING & SHARING STORAGE GROUPS ON CX600 WITH DMs & CS:**

1. c:\Program Files\EMC\Navisphere CLI>navicli -h 192.168.25.114 storagegroup -create -name nas\_san

2. c:\Program Files\EMC\Navisphere CLI>navicli -h 192.168.25.114 storagegroup -shareable -name nas\_san yes

## **ADDING WWNS TO CX600:**

1. c:\Program Files\EMC\Navisphere CLI>navicli -h 192.168.25.114 storagegroup -setpath -o -name nas\_san -hbauid <CS0  
WWN> -sp a -spport 0
2. c:\Program Files\EMC\Navisphere CLI>navicli -h 192.168.25.114 storagegroup -setpath -o -name nas\_san -hbauid <DM1  
WWN> -sp b -spport 0
3. c:\Program Files\EMC\Navisphere CLI>navicli -h 192.168.25.114 storagegroup -setpath -o -name nas\_san -hbauid <DM2  
WWN> -sp a -spport 0

## **SETTING CX600 CACHE:**

1. c:\Program Files\EMC\Navisphere CLI>navicli -h 192.168.25.114 setcache -wc 0 -rca 0 -rcb 0
2. c:\Program Files\EMC\Navisphere CLI>navicli -h 192.168.25.114 setcache -wsza 0 -wszb 0 -rsza 0 -rszb 0
3. c:\Program Files\EMC\Navisphere CLI>navicli -h 192.168.25.114 setcache -wsza 1497 -wszb 1497 -p 8 -I 70 -h 90
4. c:\Program Files\EMC\Navisphere CLI>navicli -h 192.168.25.114 setcache -wc 1

## **CREATING RAID GROUPS:**

1. c:\Program Files\EMC\Navisphere CLI>navicli -h 192.168.25.114 createrg 0 0\_0\_0 0\_0\_1 0\_0\_2 0\_0\_3 0\_0\_4

**Note:** Creates Raid 5 RG on System LUNS for disks 0-4

2. c:\Program Files\EMC\Navisphere CLI>navicli -h 192.168.25.114 createrg 200 0\_0\_14 [Creating Raid Group on disk 14]
3. c:\Program Files\EMC\Navisphere CLI>navicli -h 192.168.25.114 bind hs 200 -rg 200 -sp a [Binding LUN on new RG]
4. c:\Program Files\EMC\Navisphere CLI>navicli -h 192.168.25.114 bind r5 0 -rg - -rc 1 -wc 1 -aa 1 -sp a -sq gb -cap 4 -elsz 128 [Creates a 4 GB LUN]

## **MOVING CONTROL SYSTEM LUNS INTO STORAGE GROUPS:**

1. c:\Program Files\EMC\Navisphere CLI>navicli -h 192.168.25.114 storagegroup -addhlu -name nas\_san -hlu 0 -alu 0
2. c:\Program Files\EMC\Navisphere CLI>navicli -h 192.168.25.114 storagegroup -addhlu -name nas\_san -hlu 1 -alu 1
3. c:\Program Files\EMC\Navisphere CLI>navicli -h 192.168.25.114 storagegroup -addhlu -name nas\_san -hlu 2 -alu 2
4. c:\Program Files\EMC\Navisphere CLI>navicli -h 192.168.25.114 storagegroup -addhlu -name nas\_san -hlu 3 -alu 3
5. c:\Program Files\EMC\Navisphere CLI>navicli -h 192.168.25.114 storagegroup -addhlu -name nas\_san -hlu 4 -alu 4

## **CREATING USER LUNS:**

1. Create on any 4+1 Raid 5, 8+1 Raid 5, or Raid 1 raid groups [Must spread evenly across SPA & SPB]
2. Move new LUNs into Celerra Storage per previous procedure
3. Run server\_devconfig –create to update Celerra database

## **TROUBLESHOOTING CELERRA CLARIION/NS600 BACKEND ISSUES:**

### **NAS 5.3.14.2 NAVICLI COMMANDS:**

# /nas/sbin/navicli -h

### **OBTAIN IP ADDRESS OF SPA/SPB:**

Control Station: #cat /etc/hosts

Clariion → 172.24.11.110 A\_F20013800379 [SPA, etc—FC4700]

NS600 → 192.168.1.200 A\_WRE00022100964 [SPA, etc]

**NS600:** #cat /etc/hosts

|               |                       |              |
|---------------|-----------------------|--------------|
| 127.0.0.1     | localhost.localdomain | localhost    |
| 192.168.1.200 | A_APM00023700169      | SPA [Node A] |
| 192.168.1.201 | B_APM00023700169      | SPB [Node B] |

### **USE GETALL COMMAND TO GATHER INFO FOR ENG:**

# /nas/sbin/navicli -h 192.168.1.200 getall >getall.txt

Server IP Address: 192.168.1.200

Agent Rev: 6.2.0 (6.2)

Agent/Host Information

-----  
Name: K10  
Desc:  
Node: A-APM00023700169  
Physical Node: K10  
Signature: 674020  
Peer Signature: 674008  
SCSI Id: 0  
SP Identifier: A  
Revision: 2.02.0.60.5.003  
Model: 600  
Model Type: Rackmount  
Prom Rev: 3.09.00  
SP Memory: 2048  
Serial No: APM00023700169  
Cabinet: SPE

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!

Name of the software package: Navisphere  
Revision of the software package: 6.2.0.6.2  
Commit Required: NO  
Revert Possible: NO  
Active State: YES  
Dependent packages: ManagementServer 6.2.0.6.0, Base 02.02.0.60  
.5.003  
Required packages: Base 10 OR 105  
Is installation completed: YES  
Is this System Software: NO  
Name of the software package: NASBackend  
Revision of the software package: 01.19.04  
Commit Required: NO  
Revert Possible: NO  
Active State: YES  
Required packages: Base 105;  
Is installation completed: YES  
Is this System Software: NO  
Name of the software package: ManagementServer  
Revision of the software package: 6.2.0.6.0  
Commit Required: NO  
Revert Possible: NO  
Active State: YES  
Required packages: Navisphere 105  
Is installation completed: YES  
Is this System Software: NO  
Name of the software package: Base  
Revision of the software package: 02.02.0.60.5.003  
Commit Required: NO  
Revert Possible: NO  
Active State: YES  
Dependent packages: Navisphere 6.2.0.6.2, NASBackend 01.19.04  
Required packages: Navisphere >=4;  
Is installation completed: YES  
Is this System Software: NO  
Array Information

-----  
Array Name: APM00023700169-----abridged ouput-----

**TURNING CLARIION BACKEND STATISTICS OFF OR ON:**

# /nas/sbin/navicli -h spa getall |grep -i statistics

Statistics Logging: ON

# /nas/sbin/navicli -h SPA setstats -off [Note, make sure this is done for both SP's!]

# /nas/sbin/navicli -h 10.241.168.52 setstats (Use this cmd to check)

# /nas/sbin/navicli -h SPA setstats -on

**USING ENGINEERING MODE WITH NAVISPHERE MANAGER:**

To invoke Eng. Mode, use Ctrl + Shift + F12 keys. Password=messner

**VERIFY AGENT & FLARE CODE:**

# /nas/sbin/navicli -h 172.24.11.110 getagent [Agent & Flare code revisions, Serial number]

**OUTPUT FROM NS702G:**

\$ /nas/sbin/navicli -h 10.0.1.200 getagent

Agent Rev: 6.16.5 (0.0)

Name: K10

Desc:

Node: A-APM00042400864

Physical Node: K10

Signature: 1147781

Peer Signature: 1147754 →Always look for peer signature, could indicate SP problem

Revision: 2.16.700.5.022

SCSI Id: 0

Model: CX700

Model Type: Rackmount

Prom Rev: 3.54.00

SP Memory: 3967

Serial No: APM00042400864

SP Identifier: A

Cabinet: SPE

## **VERIFY FAULTED STATE (PROCESSORS, DISKS, FANS, ERRORS, ETC):**

**#/nas/sbin/navicli -h 172.24.11.110 getcrus**

SPE Enclosure SPE

SP A State: Present

SP B State: Present

Enclosure SPE Fan A State: Present

Enclosure SPE Fan B State: Present

Enclosure SPE Fan C State: Present

Enclosure SPE Power A State: Present

Enclosure SPE Power B State: Present

Enclosure SPE SPS A State: Present

Enclosure SPE SPS B State: Present

Enclosure SPE SPS A Cabling State: Valid

Enclosure SPE SPS B Cabling State: Valid

DAE Bus 0 Enclosure 0

Bus 0 Enclosure 0 Fan A State: Present

Bus 0 Enclosure 0 Fan B State: Present

Bus 0 Enclosure 0 Power A State: Present

Bus 0 Enclosure 0 Power B State: Present

Bus 0 Enclosure 0 LCC A State: Present

Bus 0 Enclosure 0 LCC B State: Present

Bus 0 Enclosure 0 LCC A Revision: 2

Bus 0 Enclosure 0 LCC B Revision: 0

Bus 0 Enclosure 0 LCC A Serial #: CE100023100379

Bus 0 Enclosure 0 LCC B Serial #: N/A

## **VERIFY STATE OF DISKS:**

**#/nas/sbin/navicli -h 172.24.11.110 getdisk >getdisk** [Observe Read or Write errors & LUNs bound to RAID groups]

Bus 0 Enclosure 0 Disk 0

Vendor Id: SEAGATE

Product Id: ST373405 CLAR72

Product Revision: 4A34

Lun: 0 1 2 3 4 5 16 17

Type: 0: RAID5 1: RAID5 2: RAID5 3: RAID5 4: RAID5 5: RAID5 16:

RAID5 17: RAID5

State: Enabled

Hot Spare: 0: NO 1: NO 2: NO 3: NO 4: NO 5: NO 16: NO 17: NO

Prcnt Rebuilt: 0: 100 1: 100 2: 100 3: 100 4: 100 5: 100 16: 100 17: 100

Prcnt Bound: 0: 100 1: 100 2: 100 3: 100 4: 100 5: 100 16: 100 17: 100

Serial Number: 3EK20TSX

Sectors: 126652160 (61842)

Capacity: 68238

Private: 0: 13099082 1: 13099082 2: 13099082 3: 13099082 4: 1309908

2 5: 13099082 16: 13099082 17: 13099082

Bind Signature: 0x4076, 0, 0

Hard Read Errors: 0

Hard Write Errors: 0

Soft Read Errors: 0

Soft Write Errors: 0

Read Retries: N/A

Write Retries: N/A

Remapped Sectors: N/A

Number of Reads: 4426198

Number of Writes: 468198

Number of Luns: 8

Raid Group ID: 0

Clarion Part Number: DG118032243

Request Service Time: N/A

Read Requests: 4426198

Write Requests: 468198

Kbytes Read: 255984885

Kbytes Written: 5376851

Stripe Boundary Crossing: 571

**\$ /nas/sbin/navicli -h spa getdisk 1\_1\_8** (looking at specific disk)

## **VERIFY RAID GROUP INFO:**

**#/nas/sbin/navicli -h 172.24.11.110 getrg >getrg** [Check for faulted Groups; LUNs are in decimal—convert to Hex]

RaidGroup ID: 0

RaidGroup Type: r5  
 RaidGroup State: Explicit\_Remove  
 Valid\_luns  
 List of disks: Bus 0 Enclosure 0 Disk 0  
                  Bus 0 Enclosure 0 Disk 1  
                  Bus 0 Enclosure 0 Disk 2  
                  Bus 0 Enclosure 0 Disk 3  
                  Bus 0 Enclosure 0 Disk 4  
 List of luns: 0 1 2 3 4 5 16 17  
 Max Number of disks: 16  
 Max Number of luns: 32  
 Raw Capacity (Blocks): 633262480  
 Logical Capacity (Blocks): 506609984  
 Free Capacity (Blocks,non-contiguous): 1344  
 Free contiguous group of unbound segments: 1344  
 Defrag/Expand priority: Low  
 Percent defragmented: 100  
 Percent expanded: 100  
 Disk expanding onto: N/A  
 Lun Expansion enabled: NO  
 Legal RAID types: r5  
  
 RaidGroup ID: 8  
 RaidGroup Type: r5  
 RaidGroup State: Explicit\_Remove  
 Valid\_luns  
 List of disks: Bus 0 Enclosure 0 Disk 6  
                  Bus 0 Enclosure 0 Disk 7  
                  Bus 0 Enclosure 0 Disk 8  
                  Bus 0 Enclosure 0 Disk 9  
                  Bus 0 Enclosure 0 Disk 10  
                  Bus 0 Enclosure 0 Disk 11  
                  Bus 0 Enclosure 0 Disk 12  
                  Bus 0 Enclosure 0 Disk 13  
                  Bus 0 Enclosure 0 Disk 14  
 List of luns: 18 19  
 Max Number of disks: 16  
 Max Number of luns: 32  
 Raw Capacity (Blocks): 1257136866  
 Logical Capacity (Blocks): 1117454992  
 Free Capacity (Blocks,non-contiguous): 656  
 Free contiguous group of unbound segments: 656  
 Defrag/Expand priority: Low  
 Percent defragmented: 100  
 Percent expanded: 100  
 Disk expanding onto: N/A  
 Lun Expansion enabled: NO  
 Legal RAID types: r5

**VERIFY LUNs & ARRAY:**

# /nas/sbin/navicli -h 10.241.168.57 getlun -tresspass -status

LOGICAL UNIT NUMBER 100

Default Owner: SP B  
 Current owner: SP A  
 Device Map: Valid

# /nas/sbin/navicli -h 172.24.11.110 getlun >getlun

**Note:** SPA owns control volumes & even LUNs. SPB owns odd LUNs

LOGICAL UNIT NUMBER 0

Prefetch size (blocks) = 0  
 Prefetch multiplier = 4  
 Segment size (blocks) = 0  
 Segment multiplier = 4  
 Maximum prefetch (blocks) = 4096  
 Prefetch Disable Size (blocks) = 4097  
 Prefetch idle count = 40  
 Variable length prefetching YES  
 Prefetched data retained YES  
 Read cache configured according to specified parameters.  
 Total Hard Errors: 0  
 Total Soft Errors: 0  
 Total Queue Length: 85572

Celerra TechGuide EMC CONFIDENTIAL--INTERNAL USE ONLY—NOT A FORMAL EMC PUBLICATION!!!

Name LUN 0x00  
Minimum latency reads N/A  
Read Histogram[0] 4384  
Read Histogram overflows 0  
Write Histogram[0] 2362  
Write Histogram overflows 0  
Read Requests: 9876  
Write Requests: 74197  
Blocks read: 242660  
Blocks written: 1188832  
Read cache hits: 9100  
Read cache misses: N/A  
Prefetched blocks: 45328  
Unused prefetched blocks: 1912  
Write cache hits: 3003  
Forced flushes: 0  
Read Hit Ratio: N/A  
Write Hit Ratio: N/A  
RAID Type: RAID5  
RAIDGroup ID: 0  
State: Bound  
Stripe Crossing: 571  
Element Size: 128  
Current owner: SP A  
Offset: 0  
Auto-trespass: DISABLED  
Auto-assign: DISABLED  
Write cache: ENABLED  
Read cache: ENABLED  
Idle Threshold: 0  
Idle Delay Time: 20  
Write Aside Size: 1023  
Default Owner: SP A  
Rebuild Priority: ASAP  
Verify Priority: ASAP  
Prct Reads Forced Flushed: 0  
Prct Writes Forced Flushed: 0  
Prct Rebuilt: 100  
Prct Bound: 100  
LUN Capacity(Megabytes): 4096  
LUN Capacity(Blocks): 8388608  
UID: 60:06:01:E4:48:0A:00:00:1A:DD:88:56:4F:A6:D7:11  
Bus 0 Enclosure 0 Disk 0 Queue Length: 441730  
Bus 0 Enclosure 0 Disk 1 Queue Length: 542825  
Bus 0 Enclosure 0 Disk 2 Queue Length: 1038682  
Bus 0 Enclosure 0 Disk 3 Queue Length: 573803  
Bus 0 Enclosure 0 Disk 4 Queue Length: 470773  
Bus 0 Enclosure 0 Disk 0 Hard Read Errors: 0  
-----  
Bus 0 Enclosure 0 Disk 1 Hard Read Errors: 0  
Bus 0 Enclosure 0 Disk 4 Soft Write Errors: 0

Bus 0 Enclosure 0 Disk 0 Enabled  
Reads: 4428189  
Writes: 468309  
Blocks Read: 512200412  
Blocks Written: 10754107  
Queue Max: N/A  
Queue Avg: N/A  
Avg Service Time: N/A  
Prct Idle 99.21  
Prct Busy 0.78  
Remapped Sectors: N/A  
Read Retries: N/A  
Write Retries: N/A

Bus 0 Enclosure 0 Disk 1 Enabled  
Reads: 400101  
Writes: 476573  
Blocks Read: 44377784  
Blocks Written: 10709572  
Queue Max: N/A  
Queue Avg: N/A  
Avg Service Time: N/A  
Prct Idle 99.59

Pret Busy 0.40  
Remapped Sectors: N/A  
Read Retries: N/A  
Write Retries: N/A

## **CELERRA ISSUE WITH TRESPASSED LUNS:**

Do not run server\_devconfig –create if any LUNS are trespassed on the Celerra [Camdisk and Volumes files become inconsistent when failed back to normal state]—see emc82523.

**# /nas/sbin/navicli -h 10.241.168.57 getlun -tresspass –status**

## **CONSIDERATIONS TO TAKE INTO ACCOUNT BEFORE RESTORING LUNS TO SP OWNER:**

**Note:** Typically, the WebUI>Storage>Systems screen will show a Storage Processor in, “Is Ready to Restore” status

1. Ensure Data Movers have both fibre channel paths “bound” by doing “fcp bind show”
2. Ensure that Data Movers can see down both backend paths by doing “server\_devconfig –p –s –a”
3. Compare camdisk to probe to make sure that no inconsistencies are noted
4. Click on the “Is Ready to Restore” icon in the WebUI during a time when the NASDB Backup is NOT running
5. If Luns fail to trespass back, make sure that LUNs are in the proper Storage group, then try trespassing back manually using CLI:

**# /nas/sbin/navicli -h 10.241.169.36 trespass lun 9**

## **SERVER LOG SHOWING TRESPASSED LUN:**

2007-09-04 13:22:23: STORAGE: 4: 5: volume 150 has moved from SPB to SPA

2007-09-04 13:22:23: CAM: 3: I/O Error: c16t1112 Irp 0xe5270e04 CamStatus 0x84 ScsiStatus 0x02 Sense 0x05/0x04/0x00

**Note:** The 05/04/00 Sense code indicates that Celerra is trying to access a lun owned by the other SP

## **FAILING OVER LUNS TO ALTERNATE SP:**

**# /nas/sbin/navicli -h 172.24.173.64 trespass all**

**Note:** This command is listed here for illustration purposes only—you would probably not want to risk panicking the Celerra if you choose the wrong SP. This command must be run to the SP that is up and running!

## **FAILING BACK CELERRA LUNS TO OWNER SP:**

\$nas\_storage –l [obtain Serial Number of system]

**\$nas\_storage -failback WRE00022000779**

## **FAILING BACK TO SPB AFTER REPAIRS:**

**#/nas/sbin/navicli -h 192.168.10.10 trespass mine** [IP for SPB, requesting LUNs be given back to SPB]

**Caution:** Do not use this command against all luns on an SP if the Clariion has other hosts besides the Celerra. Rather, trespass back each Celerra lun individually using the alu number found in the storagegroup –list. PowerPath hosts can allow for automatic failback of LUNs, Celerra does not. Also, power cycling an array will reset all LUNs back to their default owners.

## **FAILING BACK LUNS/RESTORING LUNS/TRESPASSING LUNS BACK TO OWNER SP:**

### **Option 1:**

**# /nas/sbin/navicli -h 10.241.169.36 trespass lun 16**

**Note:** Must run command from SP to which you are trespassing the lun to

### **Option 2:**

**# nas\_storage -failback id=1**

id = 1

serial\_number = APM00040400145

name = APM00040400145

acl = 0

**Note:** Purpose of this command is to failback all NAS Luns to their correct SP Owner. Check nas\_log.al for command start and completion, but run getlun –default –owner to verify that luns go back to their rightful Owner. In some cases, the specific lun must be trespassed back individually.

### **nas log.al**

2005-12-07 17:30:15.210 db:0:4332:S: nas\_storage -failback id=1

2005-12-07 17:30:15.382 db:0:4332:E: nas\_storage -failback id=1

### **Option 3:**

Celerra Manager> Click on Storage>Restore SP A>then click on o.k. to restore

### **Option 4:**

If LUNs will not remained trespassed back to their Owner SP, try running “fcp bind show”, as this command invokes a Test Unit Ready (TUR) message that forces “cam” layer to reset path failed bit—see AR74104 for more details 5.4.22.1 fix

### **Option 5:**

If a specific HBA path is down, issue the port reset command:

# .server\_config server\_2 -v "fcp portreset=0"

#### Option 6:

If LUNS constantly bounce back, try disabling the no\_trespass param, and then trespass LUNs back—this should disable DART from initiating LUN trespassing while trying to get things back to their proper configuration:

\$ .server\_config server\_2 -v “param clarion no\_tresspass=1”

#### Option 7:

Shutting down the Celerra, then the array, and bringing the array back up will restore all LUNS to their proper SP

#### Option 8:

If there is a fundamental binding table issue or a path issue where the HBA Chains have been diskmarked down the wrong chain, use one of the following two methods to restore LUNs to their rightful path:

a.) Failover to Standby → Clear Bind tables on faulted non-production slot → Do setup\_slot –init on faulted slot → Do devconfig –create to remark paths → Failback to original slot

b.) Temp. unmount all file systems on affected DM → Trespass luns back to correct SP → do devconfig –create to remark chains → Remount all file systems and verify

### **VERIFY CACHE SETTINGS:**

# /nas/sbin/navicli -h 172.24.11.110 [Write cache should be Enabled!]

|                                     |                      |
|-------------------------------------|----------------------|
| SP Read Cache State                 | Enabled              |
| SP Write Cache State                | Enabled              |
| Cache Page size:                    | 8                    |
| Write Cache Mirrored:               | YES                  |
| Low Watermark:                      | 70                   |
| High Watermark:                     | 90                   |
| SPA Cache pages:                    | 92287                |
| SPB Cache pages:                    | 0                    |
| Unassigned Cache Pages:             | 0                    |
| Read Hit Ratio:                     | N/A                  |
| Write Hit Ratio:                    | N/A                  |
| Prc Dirty Cache Pages =             | 0                    |
| Prc Cache Pages Owned =             | 49                   |
| SPA Read Cache State                | Enabled              |
| SPB Read Cache State                | Enabled              |
| SPA Write Cache State               | Enabled              |
| SPB Write Cache State               | Enabled              |
| System Buffer (spA):                | 551 MB               |
| System Buffer (spB):                | 551 MB               |
| SPS Test Day:                       | Saturday             |
| SPS Test Time:                      | 21:00                |
| SPA Physical Memory Size =          | 2048                 |
| SPB Physical Memory Size =          | 2048                 |
| Physical memory size of Front-End = | Switch not supported |
| Physical memory size of Back-End =  | Switch not supported |
| SPA Free Memory Size =              | 0                    |
| SPB Free Memory Size =              | 0                    |
| Free Memory Size of Front-End =     | Switch not supported |
| Free Memory Size of Back-End =      | Switch not supported |
| SPA Read Cache Size =               | 32                   |
| SPB Read Cache Size =               | 32                   |
| SPA Write Cache Size =              | 1465                 |
| SPB Write Cache Size =              | 1465                 |
| SPA Optimized Raid 3 Memory Size =  | 0                    |
| SPB Optimized Raid 3 Memory Size =  | 0                    |

# /nas/sbin/navicli -h 10.241.168.52 getsp -mem

SP A

Memory Size For The SP: 2048

SP B

Memory Size For The SP: 2048

### **DISABLING READ & WRITE CACHE AT ARRAY LEVEL:**

# /nas/sbin/navicli -h 172.24.11.110 setcache -wc 0 -rca 0 -rcb 0

**Note:** Write Cache could be disabled by faulted SP, faulted SPS power supply, Bad vault drive in first DAE, requires min. of (1) fully charged SPS

### **ALLOCATING 1/3RD READ CACHE & 2/3RD WRITE CACHE:**

# /nas/sbin/navicli -h 172.24.11.110 setcache -wsza 1197 -wszb 1197 -rsza 300 -rszb 300

### **ENABLING READ & WRITE CACHE AT ARRAY LEVEL:**

# /nas/sbin/navicli -h 172.24.11.110 setcache -rca 1 -rcb 1 -wc 1 [Enabling Read & Write cache]

# /nas/sbin/navicli -h 172.24.11.110 setcache -wc 1 [Enabling write cache only]

## **RE-SETTING READ & WRITE CACHE ON SP's WITH 50:50 RATIO:**

**Note:** Perform during period of relative inactivity, as these steps could impact users

### **1. DISABLE WRITE & READ CACHE:**

```
# /nas/sbin/navicli -h 192.168.1.200 setcache -wc 0 (SPA)
# /nas/sbin/navicli -h 192.168.1.201 setcache -wc 0 (SPB)
# /nas/sbin/navicli -h 192.168.1.200 setcache -rca 0 (SPA)
# /nas/sbin/navicli -h 192.168.1.201 setcache -rcb 0 (SPB)
```

### **2. ZERO OUT WRITE & READ CACHE VALUES:**

```
#/nas/sbin/navicli -h 192.168.1.200 setcache -wsza 0 (SPA)
#/nas/sbin/navicli -h 192.168.1.201 setcache -wszb 0 (SPB)
#/nas/sbin/navicli -h 192.168.1.200 setcache -rsza 0 (SPA)
#/nas/sbin/navicli -h 192.168.1.201 setcache -rszb 0 (SPB)
```

### **3. SET NEW WRITE & READ CACHE VALUES:**

```
#/nas/sbin/navicli -h 192.168.1.200 setcache -wsza 1465
#/nas/sbin/navicli -h 192.168.1.201 setcache -wszb 1465
#/nas/sbin/navicli -h 192.168.1.200 setcache -rsza 32
#/nas/sbin/navicli -h 192.168.1.201 setcache -rszb 32
```

### **4. RE-ENABLE WRITE & READ CACHE:**

```
$ /nas/sbin/navicli -h 192.168.1.200 setcache -wc 1
$ /nas/sbin/navicli -h 192.168.1.201 setcache -wc 1
$ /nas/sbin/navicli -h 192.168.1.200 setcache -rca 1
$ /nas/sbin/navicli -h 192.168.1.201 setcache -rcb 1
```

### **5. VERIFY SETTINGS:**

\$ /nas/sbin/navicli -h 192.168.1.200 getcache

|                        |         |
|------------------------|---------|
| SP Read Cache State    | Enabled |
| SP Write Cache State   | Enabled |
| SPA Read Cache State   | Enabled |
| SPB Read Cache State   | Enabled |
| SPA Write Cache State  | Enabled |
| SPB Write Cache State  | Enabled |
| SPA Read Cache Size =  | 1497    |
| SPB Read Cache Size =  | 1497    |
| SPA Write Cache Size = | 1497    |
| SPB Write Cache Size = | 1497    |

### **6. VERIFY THAT CACHE IS ENABLED ON EACH SP:**

```
# /nas/sbin/navicli -h 192.168.1.200 getlog |grep -i cache |tail
# /nas/sbin/navicli -h 192.168.1.201 getlog |grep -i cache |tail
```

```
03/16/2005 11:10:13 SP A      (639) System Cache Enabled [System cache is Write cache]
03/16/2005 11:09:59 SP B      (639) System Cache Enabled
07/27/2005 14:05:14 SP A      (714) Read cache enabled
07/27/2005 14:08:49 SP B      (714) Read cache enabled
```

### **SETTING CACHE ON BOTH SP's WITH ONE COMMAND:**

#/nas/sbin/navicli -h 192.168.1.200 setcache -wsza 1497 -wszb 1497 -rsza 1497 -rszb 1497

\$ /nas/sbin/navicli -h 10.0.1.200 getcache

|                          |        |
|--------------------------|--------|
| Low Watermark:           | 70     |
| High Watermark:          | 90     |
| SPA Cache pages:         | 190783 |
| SPB Cache pages:         | 190784 |
| Unassigned Cache Pages:  | 0      |
| Prct Dirty Cache Pages = | 1      |

**Note:** If dirty cache pages were to approach LWM, then Clariion performance would be impacted as it flushes to disk

## **WRITE CACHE AND EFD LUNS:**

See emc234527. CLARiiON best practice states that both Read and Write cache be disabled at the LUN level for EFD drives.

Celerra best practice is to have Read cache disabled, but Write cache enabled. This has caused confusion in the field. Also, NAS 5.6.47 code diskmarking will fail if Read cache is disabled. Therefore, in order to diskmark EFD luns, make sure that both Read & Write cache are enabled.

**Workaround:**

1) Enable Read & Write cache on the lun(s):

```
$ /nas/sbin/navicli -h 10.241.168.188 chglun -l 16 -c rw
```

2) Diskmark from CLI or Rescan the disk devices from Celerra Manager in order to bring them into the Celerra database:

```
$ server_devconfig server_2 -c -s -a [or nas_diskmark -m -a]
```

Discovering storage (may take several minutes)

server\_2 : done

3) Disable Read cache on the EFD Celerra LUN(s):

```
$ /nas/sbin/navicli -h 10.241.168.188 chglun -l 16 -c write
```

4) Verify:

```
$ /nas/sbin/navicli -h 10.241.168.188 getlun 16 -wc -rc
```

Write cache:       ENABLED

Read cache:      DISABLED

**DISABLING READ & WRITE CACHE ON A LUN:**

**# /nas/sbin/navicli -h 10.241.168.188 chglun -l 16 -c none**

**ENABLING READ & WRITE CACHE ON A LUN:**

**# /nas/sbin/navicli -h 10.241.168.188 chglun -l 16 -c rw**

**VERIFYING READ & WRITE CACHE STATUS ON A LUN:**

**# /nas/sbin/navicli -h 10.241.168.188 getlun 16 -wc -rc**

**ENABLING WRITE CACHE ONLY:**

**# /nas/sbin/navicli -h 10.241.168.188 chglun -l 16 -c write**

**# /nas/sbin/navicli -h 10.241.168.188 getlun 16 -wc -rc**

Write cache:       ENABLED

Read cache:      DISABLED

**CHECK SP LOGS:**

**#/nas/sbin/navicli -h 172.24.11.110 getlog >spalog** [Dump log for SPA & SPB to file and troubleshoot SP issues]

**VERIFYING FIBRE PATHS TO BACKEND:**

**\$server\_config server\_x -v "fcip bind show"** [Should show four paths to backend]

Persistent Binding Table

Chain 0000: WWN 5006016000600251 HBA 0 SP-a0 [SPA-0]

Chain 0016: WWN 5006016800600251 HBA 0 SP-b0 [SPB-0]

Chain 0032: WWN 5006016100600251 HBA 1 SP-a1 [SPA-1]

Chain 0048: WWN 5006016900600251 HBA 1 SP-b1 [SPB-1]

**NAVICLI COMMANDS FAIL WITH “Aborted”:**

Possible cause is that /etc/hosts file and /sbin/ifconfig are not up-to-date with external IP address]

Run uname -a and /sbin/ifconfig -a to obtain proper Hostname and External IP to add to /etc/hosts file

**SAMPLE:**

**# uname -a**

Linux cel9cs0 2.4.9-34.8.EMC #1 Thu Feb 6 16:34:45 EST 2003 i686 unknown

**#/sbin/ifconfig -a**

```
eth1 Link encap:Ethernet HWaddr 00:02:B3:B9:EB:0D
      inet addr:10.64.25.90 Bcast:10.64.25.255 Mask:255.255.254.0
          UP BROADCAST RUNNING MULTICAST MTU:1500 Metric:1
          RX packets:5387497 errors:0 dropped:0 overruns:0 frame:0
          TX packets:2173368 errors:0 dropped:0 overruns:0 carrier:117
          collisions:188 txqueuelen:100
          RX bytes:667840133 (636.9 Mb) TX bytes:404120327 (385.3 Mb)
          Interrupt:10 Base address:0x6000
```

**#cat /etc/hosts**

```
10.64.25.90 cel9cs0
```

**CELLERRA WITH CLARIION BACKENDS:**

**LUN AUTO-ASSIGN FEATURE FOR SP's:**

**AUTO-ASSIGN DEFINED:**

Auto-Assign is a Clariion function that controls the ownership of a LUN when an SP fails. With this feauture Enabled, if an SP that owns a LUN fails and the Host system tries to access the LUN via the second SP, the latter will assume ownership (via LUN trespass) until the failed SP can be replaced. After replacement, the ownership of the LUN needs to be trespassed back to the default owner. This mechanism was devised in those configurations where failover software is not employed.

Primus emc87328 states that all direct-attached Clariion storage and Fabric-attached systems should have auto-assign enabled so that unowned Luns trespass to partner SP—change from ‘DISABLED’ to ‘ENABLED’. This function works on the Backend but is monitored by DART. If the DM loses connectivity to LUNs owned by a specific SP, LUN failover occurs. Failover of LUN 50 from SPA to SPB could occur as an example. All Clariion systems attached to Celerra should implement auto-assign.

**Note:** New default behavior with NAS 5.2.16.x and 5.3.10.4 is to have Auto-Assign Enabled on all Celerra systems with Clariion backends that have conducted a full package upgrade or new install of 5.3

## **ENABLING AUTO ASSIGN ON NS600/700 DIRECT-ATTACHED SYSTEMS:**

1. Verify Current Status:

**\$ /nas/sbin/navicli -h <sp\_ip> getlun -owner -default -state -aa**

LOGICAL UNIT NUMBER 16

|                |          |
|----------------|----------|
| Current owner: | SP A     |
| Default Owner: | SP A     |
| State:         | Bound    |
| Auto-assign:   | DISABLED |

2. Enable auto-assign of Luns:

**\$ /nas/sbin/navicli -h <sp\_ip> chglun -l <lun ID> -a 1** [ALU lun number]

**Note:** Run this command to enable Auto-assign for each each LUN. Starting with NAS 5.4 and higher, the auto-assign value should be disabled on all Celerra systems connected to Clariion arrays.

## **ENABLING CLARIION TRESPASS OF LUNS FOR CLARIION BACKENDS:**

**param clariion no\_tresspass=0**

**Note:** Older NAS versions had trespass set to 1 by default to enable the Clariion LUN trespass feature for switched environments, but set to 0 by default for dedicated storage systems such as an NS600 and NS600G direct-attached [i.e., not switch-connected]. Param is implemented so that when external Host clients are trying to drive I/O down a specific path, or chain, an auto trespass would be initiated if a path was no longer available—trespassing from Chain 0 to Chain 16, as an example.

**Note 2:** NAS 5.3.10.4 & 5.2.16.x officially changes trespass default to clariion no\_tresspass=0, or disabled

**NASDB:** NAS configuration database, mostly located on /nbsnas partition /dev/nde1 on LUN 4

## **CONTROL LUN PARTITIONS:**

**/nbsnas:** Mountpoint for NBS partition on /dev/nde1, LUN 4 [Contains mountpoints for DOS & VAR partitions]

**/nas:** Mountpoint for NAS Partition on /dev/hda5 containing Celerra/IP Configuration files—many files are links to /nbsbas directory

**/nas/dos:** FAT16 DOS Partition containing DART OS using LUN 0 on BackEnd--/nas/dos is a symbolic link to /nbsnas/dos [CS0 sees as /dev/nda1]. DART loads OS and Configuration Files from this partition via Fibre connection.

**/nas/var:** LOG Partition for NAS\_DB Backups, Dumps, and Logs, using LUN 5 on BackEnd--/nas/var is a symbolic link to /dev/ndf1 [CS0 sees as /dev/ndf1]

## **USING FDISK TO VERIFY COMMUNICATION TO NBSNAS:**

**# /sbin/fdisk -l /dev/nda (/nas/dos or /nbsnas/dos)**

**# /sbin/fdisk -l /dev/nde (/nas or /nbsnas)**

**# /sbin/fdisk -l /dev/ndf (/nas/var or /nbsnas/var)**

## **VERIFYING PARTITIONS/VOLUMES:**

**# /sbin/fdisk -l**

**# cat /etc/fstab**

**# df -k        #df -h**

## **RECREATING DOS PARTITION USING install\_init -r:**

1. #umount /nas/dos

2. #fdisk -l /dev/nda [verify partition accessibility]

3. #/nas/sbin/install\_init -r

Creating the DOS Partition on /dev/nda1 with fdisk...

Building the DOS filesystem on /dev/nda1 with mkfs.msdos...

**Note:** Remount /nas/dos before running setup\_slot -r

4. #/nasmcd/sbin/setup\_slot -r 2

Initializing server in slot 2 as server\_2

5. #/nasmcd/sbin/setup\_slot -r 3

6. #sync

7. #/nasmcd/sbin/t2reset reboot -s 2 | -s 3

## **RECREATING /nas/var PARTITION:**

1. Delete old partition  
# /sbin/fdisk /dev/ndf  
p | d | 1 | p | w
2. Recreate partition  
# /sbin/fdisk /dev/ndf  
n | 1 | 1 | +1800M | w
3. Format Partition  
# /sbin/mkfs -t ext2 /dev/ndf1 1847443 (as seen in output of fdisk -l /dev/ndf)
4. Change to journal file system ext3  
# /sbin/mke2fs -j /dev/ndj1
5. Confirm and then create necessary directories and change ownership to nasadmin

## **CONTROL STATION RECOVERY (Prior to NAS 5.6):**

**Note:** Use this if Linux cannot boot

1. Insert CFS CD ROM and Floppy disk
2. Reboot CS # reboot or # reboot -n -f [Force, no sync reboot]
3. Prompt, type >**serialinstall rescue**

**Note:** Ignore failure to find Linux partition message

4. At prompt fsck root file system:

```
#fsck -f /dev/hda1 #fsck -f /dev/hda3 #mkdir /mnt/sysimage #mount /dev/hda3 /mnt/sysimage #chroot /mnt/sysimage #mount /dev/hda1 /boot  
5. If FSCK doesn't appear to fix any problems, then boot configuration file could be corrupt  
/mnt/sysimage/boot/grub.conf  
default=0  
timeout=10  
serial .unit=1 .speed=19200  
terminal .timeout=10 serial console  
title linux
```

6. Repair and reinstall GRUB boot loader  
#grub-install --recheck /dev/hda

7. #umount /boot #exit #umount /mnt/sysimage #exit

**CONTROL STATION RECOVERY:** Linux won't boot and serial console connection shows blank screen

1. Connect to serial port ttyS1 on CS and start hyperterminal session
2. Boot from CD and Floppy and at prompt type: serialinstall
3. Complete install and at boot backend LUNS should be detected and recovery script will attempt to run (does not touch NASDB)
4. If system contains non-default IP addresses for SPs (192.168.1.200), let script error out, then login as root
5. Change CS Internal Interface to be same subnet as SPs:

```
#ifconfig eth0 10.6.8.100 netmask 255.255.255.0 broadcast 10.6.8.255
```

6. Change SPB IP address back to default:

**#/nasmcd/tftpboot/bin/navicli -h 10.6.8.201 networkadmin -set address 192.168.1.201 -subnetmask 255.255.255.0 -gateway 192.168.1.100**

**Note:** Let SPB reboot and then repeat for SPA, then change CS internal interface back to default

7. Change DMs internal IPs back to defaults if needed

```
/tftpboot/bin/t2tty -c 2 "logsys add output console" "ifconfig", repeat for slot_3
```

```
/tftpboot/bin/t2tty -c 2 "ifconfig el30 protocol=IP device=fpx0 local=192.168.1.2 netmask=255.255.255.0 broadcast=192.168.1.255
```

**Note:** Repeat above for slot\_3, then turn off DART console: "logsys set output disk=root\_log\_2"

8. Run recovery script: #/etc/rc3.d/S95nas

9. If successful, verify nbs.conf and repair if necessary using following;

**#/nasmcd/sbin/setup\_slot -nbs -add 2 | 3**

10. Update /etc/hosts with backend array IP addresses:

**#/nas/sbin/setup\_clariion -init**

11. Manually add SPA and SPA aliases to /etc/hosts

```
192.168.1.200 A_APM00023400700 SPA
```

12. Reboot CS and verify

## **CS RECOVERY FOR NASDB PARTITION CORRUPTED:** /dev/nde1

**#/fsck -f /dev/nde1**

**Note:** Reboot to see if /nbsnas will mount

If NASDB cannot be repaired, conduct reinstall using #emcrpm -install, which recreates /nbsnas /nbsnas/dos /nbsnas/var /nas, then restore NASDB from backup

1. Script reinstalls all NASDB partitions from scratch-last resort only!
2. PXE boot data movers and ensure NBS devices are accessible
3. Save installed package to temp directory: #cp /var/sadm/pkg/emcnas/\*.rpm /root/tmp
4. Uninstall old package: #emcrpm -e emcnas
5. Install new package: #emcrpm -i /root/tmp/emcnas~...i385.rpm
6. Recover NASDB from backup and restore

### **RECOVERING ROOT CS PASSWORD:**

1. Boot to single user mode by entering “e” at grub boot loader screen
2. Add word ‘single’ at end of kernel /vmlinuz ro root=/dev/hda3 console=ttyS1,19200 line, then press enter, then ‘b’ to boot

**kernel /vmlinuz ro root=/dev/hda3 console=ttyS1,19200 single**

3. #vi /etc/passwd and change root:x: to root::
4. Reboot and then change password after logging in

### **FAILED NAS UPGRADE:**

**Note:** Retry upgrade.

1. Change /nas from a symbolic link to /nbsnas back to normal /nas /dev/hda5.
2. Delete nas mountpoint, mkdir nas, mount /nas
3. Rerun NAS Upgrade

### **RECOVERING FROM CORRUPTED BOOT CONFIG FILES:**

1. Verify /nas/dos/slot files, nas.exe, and boot.bat files
2. Run setup\_slot -i 2 to repair files
3. Find backup file and extract to /temp [tar -zxf bkup04.tar.gz]
4. Copy recovered files back to appropriate /nas/server directory:  
#cp -Rf server/server\_1/\* /nas/server/server\_1/
5. Run setup slot to regenerate boot.cfg file

### **REPLACING DATA MOVER:**

1. Ensure NAS Services are stopped and DM is halted
2. Cable up new data mover and manually mount /nas /nbsnas /nas/dos /nas/var
3. #export NAS\_DB=/nas
4. Run setup\_slot -i 2 on new data mover and start NAS Services

**PTY Ports:** Linux Serial ports used on Control Station—accessed by Reading & Writing files in /dev/ttx directories

/dev/ttys0—Serial modem access, back of CS

/dev/ttys1—Serial console access, front of CS

/dev/ttys4—Connects to Slot\_2, on 4-port serial cable

/dev/ttys5—Connects to Slot\_3, on 4-port serial cable

**PXE:** Pre-Execution Boot Environment allows CS to act as PXE server to which DMs boot DART from via network connection

### **SYSTEM READINESS:**

**/nasmcd/sbin/getreason -s 2**

getboxmask

### **NS INTEGRATED CALLHOMES:**

Celerra IP uses a single CallHome for both DART & Clariion hardware [Diagnose BackEnd using NaviCLI commands]

### **REBOOTING SYSTEM:**

/nasmcd/sbin/t2reset reboot -s 2

server\_cpu

/nas/tools/server\_peer\_powerctrl -reboot 2

### **CONFIGURING BACKEND:**

\$ ls -l /nasmcd/tftpboot/bin

\$/nasmcd/tftpboot/setup\_backend

agent.pm → Properties and subroutines for SP agent

common.pm → common definitions and subroutines

disk.pm → properties & routines pertaining to physical disks  
enclosure.pm → Enclosure subroutines  
enclosure.cfg → Config file used to build RAID volumes  
nas\_raid.log → nas\_raid command log file  
system.pm → properties for creating Celerra Clariion system

## **SYSTEM LOGS:**

/nas/log/sys\_log [Panics, Hardware Errors, Hardware Status]  
/nas/log/agent.log [Navisphere agent log]  
/nas/log/navimon.log [Navisphere monitor log]  
/nas/log/setup\_clariion2.log [BackEnd install log]

## **TESTING TTY SERIAL CONNECTIVITY & ISSUING DM COMMANDS: .server\_tty -t**

**Note:** Would use in the event the Ethernet network was not available

1. service nas stop
2. /nas/tools/.server\_tty -t [Test serial connectivity]
2. /nas/tools/.server\_tty -c 2 “**logsys add output console**” [allows command output to Console]
3. /nas/tools/.server\_tty -c 2 “ifconfig”
4. Disable Console Output: /nas/tools/.server\_tty -c 2 “**logsys set output disk=root\_log\_server2**”

## **EXAMPLES TTY COMMANDS:**

```
# /nas/tools/.server_peer_powerctrl -chkins 3
# /nas/tools/.server_tty -s 3
Slot-3 (ttyS5) on
# /nas/tools/.server_tty -t
ttyS4 on [DM2]
ttyS5 on [DM3]
ttyS6 off
ttyS7 off
ttyS8 off
# /nas/tools/.server_tty -c 2 "sib" [Check getreason status of DMs]
reason_code=05
```

## **USE T2TTY TO CHANGE GETREASON, CHECK IFCONFIG OR RUN CMDS—“fcp bind show”:**

1. Must set console output to console:

```
# ./t2tty -c 3 "logsys add output console"
```

logsys add output console

2. Verify SIB Status:

```
# ./t2tty -c 3 "sib"
```

reason\_code=05

3. Verifying NIC Configuration on DMs:

```
# ./t2tty -c 3 "ifconfig"
```

ifconfig

Devices:

fxp0 dmtu=1500, dmac=8:0:1b:43:58:7

cge0 dmtu=9000, dmac=8:0:1b:42:4a:86

loop dmtu=65536, dmac=0:0:0:0:0:0

Interfaces:(4)

el30 on fxp0 l=192.168.1.3 n=255.255.255.0 b=192.168.1.255 DNIF UP

    mtu=1500, dmtu=1500, vlid=0, mac=8:0:1b:43:58:7 dmac=8:0:1b:43:58:7

el31 on fxp0 l=192.168.2.3 n=255.255.255.0 b=192.168.2.255 DNIF UP

    mtu=1500, dmtu=1500, vlid=0, mac=8:0:1b:43:58:7 dmac=8:0:1b:43:58:7

cge0 on cge0 l=10.241.169.45 n=255.255.255.0 b=10.241.169.255 DNIF UP

    mtu=1500, dmtu=9000, vlid=0, mac=8:0:1b:42:4a:86 dmac=8:0:1b:42:4a:86

loop on loop l=127.0.0.1 n=255.0.0.0 b=127.255.255.255 UP

    mtu=32768, dmtu=65536, vlid=0, mac=0:0:0:0:0:0 dmac=0:0:0:0:0:0

4. Changing SIB Status:

# ./t2tty -c 3 "sib attention=4"

sib attention=4

reason\_code=04

5. Redirect Console Output to Log after completing T2TTY session:

# ./t2tty -c 3 "logsys set output disk=root\_log\_server\_3"

logsys set output disk=root\_log\_server\_3

## **CHANGING INTERNAL IP ADDRESSES TEMPORARILY FOR NAS INSTALL RECOVERY:**

1. #.server\_tty -c 2 "logsys add output console"

2. #.server\_tty -c 2 "ifconfig"

3. #.server\_tty -c 3 "logsys add output console"

4. #.server\_tty -c 3 "ifconfig"

5. #.server\_tty -c 2 "ifconfig el30 protocol=IP device=fxp0 local=192.168.1.2 netmask=255.255.255.0 broadcast=192.168.1.255"

6. #.server\_tty -c 3 "ifconfig el30 protocol=IP device=fxp0 local=192.168.1.3 netmask=255.255.255.0 broadcast=192.168.1.255"

7. Verify ping and then turn off console logging: #.server\_tty -c 2 "logsys set output disk=root\_log\_server2"

8. Run recovery install script → /etc/rc3.d/S95nas

**BACKING UP NAS DATABASE:** Stored in (2) locations

### **(1) BACKUP STORED LOCAL CS IDE DRIVE:**

/nasmcd/sbin/nasdb\_backup /nas /home/nasadmin 120702

**OTHER BACKUP STORED ON** /nas/var/backup on array LUN 5

## **VARIOUS NASDB BACKUP FILES:**

-rw-r--r-- 1 nasadmin nasadmin 5203347 Jan 7 12:06 \_nasbkup.12.tar.gz [One of (12) hourly backups maintained on CS]

-rw-r--r-- 1 nasadmin nasadmin 5759420 Jan 7 21:06 nasdb\_backup.1.tar.gz [Latest hourly backup + extra log files]

-rw-r--r-- 1 nasadmin root 5631868 Jan 4 20:16 nasdb\_backup.b.tar.gz [Backup after last CS reboot]

## **VERIFYING ARRAY STATUS:**

/nas/sbin/navicli -h spa getagent [Checks status of backend storage array—shows mapping of LUN numbers to Celerra partitions]

/nas/sbin/navicli -h spa getlun 05 -state [Verifies indicated LUN]

/nas/sbin/navicli -h spa getlun -aa -state [Status of all LUNs]

/nas/sbin/navicli -h spa getlun 05 -owner

/nas/sbin/navicli -h spa trespass lun 06 [Assign a LUN to specific]

# **./nas/sbin/navicli -h 10.241.169.35 getlun -state -owner -default** [Ownership of Luns and Bound state]

# **./nas/sbin/navicli -h 10.241.168.57 getlun -trespass** (use to verify trespassed luns)

## **RESTORING SP LUNs:**

\$./nas/sbin/navicli -h spa trespass mine

**Note:** When trespassing control LUNs back to original SPs, CS0 may reboot. Also note that for arrays that have other hosts connected to them, it is NOT recommended to run trespass mine against all luns on an SP. Rather, issue the command to each Celerra alu in order to trespass back.

## **ETHERNET BROADCOM CONTROLLER ISSUES:**

\$#.server\_config server\_2 -v "bcm cgeN stop"

\$#.server\_config server\_2 -v "bcm cgeN start"

\$#.server\_config server\_2 -v "bcm cgeN stat"

\$#.server\_config server\_2 -v "bcm fgeN stat"

## **CELERRA NBS NETWORK BLOCK STORAGE PROTOCOL:**

iSCSI technology that allows clients to access block storage devices on Servers [Control Station NBS Client connects to DM NBS Server on one of the following three partitions]

NBS Client runs on Control Station [Doesn't have an HBA to communicate with, so uses the NBS protocol to write to BackEnd db] NBS Servers run on Data Movers [/nas /nas/dos /nas/var]

\$ ps -ef |grep nd-clnt [Should find one instance of NBS running for /nas, /nas/dos, & /nas/var partitions]

root 6354 1 0 Jan15 ? 00:00:02 [nd-clnt 0 1]

root 6367 1 0 Jan15 ? 00:00:02 [nd-clnt 4 5]

root 6386 1 0 Jan15 ? 00:00:00 [nd-clnt 5 6]

\$ /sbin/fdisk -l /dev/nda = /nas/dos NBS [Verifies that partitions are accessible]

\$ /sbin/fdisk -l /dev/nde = /nas NBS

\$ /sbin/fdisk -l /dev/ndf = /nas/var NBS

### **MOUNTING NBS PARTITIONS:**

```
#mount /nas          [/nbsnas on /dev/nde1]  
#mount /nas/dos     [/dev/nda1]  
#mount /nas/var     [/dev/ndf1]
```

### **TROUBLESHOOTING NBS:**

# cat /etc/nbs.conf

```
#  
# Simple configuration file for nbs service  
#  
# Format: devIndex:NbsId:Host1,host2,...:  
0:1:server_2,server_3b,server_2b,server_3:  
4:5:server_2,server_3b,server_2b,server_3:  
5:6:server_2,server_3b,server_2b,server_3:
```

# ls /proc/driver/nd/devices

1 5 6

# ps -ef |grep nd-clnt

```
root  6975  1 0 May08 ?    00:01:10 [nd-clnt 0 1]  
root  6979  1 0 May08 ?    00:10:33 [nd-clnt 4 5]  
root  6983  1 0 May08 ?    00:00:28 [nd-clnt 5 6]
```

**Note:** Above represents normal NBS processes for /nas, /nas/dos, /nas/var partitions, as seen by Control Station

# /sbin/fdisk -l /dev/nda

Disk /dev/nda: 255 heads, 63 sectors, 1435 cylinders

Units = cylinders of 16065 \* 512 bytes

| Device    | Boot | Start | End | Blocks | Id | System |
|-----------|------|-------|-----|--------|----|--------|
| /dev/nda1 | *    | 1     | 17  | 136521 | 6  | FAT16  |

# /sbin/fdisk -l /dev/nde

Disk /dev/nde: 255 heads, 63 sectors, 261 cylinders

Units = cylinders of 16065 \* 512 bytes

| Device    | Boot | Start | End      | Blocks | Id    | System |
|-----------|------|-------|----------|--------|-------|--------|
| /dev/nde1 | 1    | 230   | 1847443+ | 83     | Linux |        |

# /sbin/fdisk -l /dev/ndf

Disk /dev/ndf: 255 heads, 63 sectors, 261 cylinders

Units = cylinders of 16065 \* 512 bytes

| Device    | Boot | Start | End      | Blocks | Id    | System |
|-----------|------|-------|----------|--------|-------|--------|
| /dev/ndf1 | 1    | 230   | 1847443+ | 83     | Linux |        |

**Note:** Above output means that CS can see each of the NBS partitions

**\$ .server\_config server\_2 -v "nbsstat"**

```
1147963266: NBS: 4: BlkSrvStat: Avg Rd IO size:0 kBytes, Avg Wr IO size:8 kBytes  
1147963266: NBS: 4: BlkSrvStat: RDcnt:0, RD Success:0, RD Fail:0  
1147963266: NBS: 4: BlkSrvStat: WRcnt:6, WR Success:6, WR Fail:0  
1147963266: NBS: 4: BlkSrvStat: RD:0, WR:9, RW:0 Kb/sec  
1147963271: NBS: 4: *****
```

1147963271: NBS: 4: BlkSrvStat: Avg Rd IO size:0 kBytes, Avg Wr IO size:12 kBytes-----output abridged-----

**\$ .server\_config server\_2 -v "nbs start"**

**\$ .server\_config server\_2 -v "nbs info"**

```
1147963515: NBS: 4: **** Block Server ****
```

```
1147963515: NBS: 4: T5.5.21.1-NBSv6
```

```
1147963515: NBS: 4: Listening on port 5033
```

```
1147963515: NBS: 4: 20 threads started
```

```
1147963515: NBS: 4: TCP hi watermark: 524288
```

```
1147963515: NBS: 4: Using default TCP low watermark
```

```
1147963515: ADMIN: 4: Command succeeded: nbs info
```

**\$ .server\_config server\_2 -v "nbsid list"**

1147963559: NBS: 4: nbs add name=1 vol=NBS1 rw=192.168.1.100:192.168.1.101:192.168.2.100:192.168.2.101 share exclusive raw  
1147963559: NBS: 4: nbs add name=5 vol=NBS5 rw=192.168.1.100:192.168.1.101:192.168.2.100:192.168.2.101 share exclusive raw  
1147963559: NBS: 4: nbs add name=6 vol=NBS6 rw=192.168.1.100:192.168.1.101:192.168.2.100:192.168.2.101 share exclusive raw

### \$ .server\_config server\_2 -v "volume info NBS1"

\*\*\*\* Basic Volume 1 : 0x10a6204 Information: \*\*\*\*

Total References:.....0x000b

Total Blocks:.....0x15fff80

Bytes Per Block:.....0x0200

## **MANUALLY RESTARTING NBS SERVICES ON CS:**

**\$/usr/local/bin/nd-clnt 0 1 server\_2,server\_3** | 192.168.1.2, .3 [Starting NBS on /dev/nda for /nas/dos]

**\$/usr/local/bin/nd-clnt 4 5 server\_2,server\_3** [Starting NBS on /dev/nde for /nas]

**\$/usr/local/bin/nd-clnt 5 6 server\_2,server\_3** [Starting NBS on /dev/ndf for /nas/var]

### # /usr/local/bin/nd-cfg

Usage: nd-cfg [-a action] [-d dev\_name] [-i NbsId | -l] [-s server\_name] [ -v]

nd-cfg -i NbsId -s server\_name

nd-cfg -l -s server\_name

nd-cfg -a addServ -s server\_name -d /dev/nda

nd-cfg -a remServ -s server\_name -d /dev/nda

nd-cfg -a setRetry\_# -d /dev/nda

nd-cfg -a stop -d /dev/nda

nd-cfg -a disconnect -s server\_name -d /dev/nda

## **PROCEDURE TO CORRECT NBS SERVICES FOR A SPECIFIC DATA MOVER:**

1. Run command to add missing nbs configuration for server\_3:

**#/nasmcd/sbin/setup\_slot -nbs -add 3**

2. Check that server\_3 & server\_3b have been added to the config file:

**#cat /etc/nbs.conf**

3. Stop NAS services using the command:

**#/sbin/service nas stop**

4. Recycle NBS service using:

**#/sbin/service nbs restart**

5. Check that NBS has started correctly with the added servers:

**#/sbin/service nbs status**

**Note:** Should see output for each NBS device with the following server list:

Configured NBS devices:

254: 0 NbsId: 1

Disk\_Name: nda

Capacity: 11534272 KB

Crnt\_Server: 0X202a8c0

**Server\_List: 0X301a8c0 0X201a8c0 0X302a8c0 0X202a8c0**

num\_io: 22199 time 901410 ms num\_sect 1396388 que 0

Avg\_blk: 31 K throughput 774 K/s resp\_tm: 17 ms, ops/sec 24

254: 16 NbsId: 5

Disk\_Name: nde

Capacity: 2097088 KB

Crnt\_Server: 0X201a8c0

**Server\_List: 0X201a8c0 0X302a8c0 0X202a8c0 0X301a8c0**

num\_io: 26935 time 901400 ms num\_sect 488192 que 0

Avg\_blk: 9 K throughput 270 K/s resp\_tm: 2 ms, ops/sec 29

254: 20 NbsId: 6

Disk\_Name: ndf

Capacity: 2097088 KB

Crnt\_Server: 0X202a8c0

**Server\_List: 0X302a8c0 0X202a8c0 0X301a8c0 0X201a8c0**

num\_io: 23246 time 901400 ms num\_sect 1405496 que 0

Avg\_blk: 30 K throughput 779 K/s resp\_tm: 138 ms, ops/sec 25

6. Restart NAS services using the command:

**#/sbin/service nas start**

### **TROUBLESHOOTING IF PXEBOOT FAILS TO REGAIN ACCESS TO BACKEND:**

**Note:** Many variables can affect access to the backend devices via NBS, but the following steps should eliminate most of the issues

- a) Check Control Station's NBS configuration file (Single Blade system shown)

```
# cat /etc/nbs.conf  
#  
# Simple configuration file for nbs service  
#  
# Format: devIndex:NbsId:Host1,host2,...:  
0:1:server_2,server_2b:  
4:5:server_2,server_2b:  
5:6:server_2,server_2b:  
b) Could be that the NBS Service is not running properly on the Control Station
```

```
# ps -ef |grep nd-
```

**Note:** If services are not running, do the following

- c) First verify that the Control Station can access the backend via IP and Internal network (ping IP addresses, run /ftptboot/bin/navicli)

- d) Stop NAS Services

```
# /sbin/service nas stop →If this fails to stop any NAS Services that are running, manually kill the processes
```

- e) Restart NBS on the Control Station

```
# /sbin/service nbs restart →Restarts NBS services
```

- f) Verify that NBS processes are running correctly

```
# ps -ef |grep nd-
```

```
root 4368 1 0 11:40 ? 00:00:00 [nd-clnt 0 1]  
root 4369 1 0 11:40 ? 00:00:00 [nd-clnt 4 5]  
root 4370 1 0 11:40 ? 00:00:00 [nd-clnt 5 6]  
root 4371 1 0 11:40 ? 00:00:00 [nd-snnd 4 5]  
root 4372 1 0 11:40 ? 00:00:00 [nd-snnd 5 6]  
root 4373 1 0 11:40 ? 00:00:00 [nd-snnd 0 1]
```

- g) Restart NAS Services

```
# /sbin/service nas start
```

- h) Verify backend partitions

```
# df -h →Should see NAS partitions mounted
```

```
# /sbin/fdisk -l
```

-----abridged-----

| Device    | Boot | Start | End  | Blocks  | Id | System    |
|-----------|------|-------|------|---------|----|-----------|
| /dev/nda1 | *    | 1     | 17   | 136521  | 6  | FAT16     |
| /dev/nda3 |      | 654   | 1435 | 6281415 | 8e | Linux LVM |

### **POST NAS 5.2.7.0:**

1. Stop NAS and remount partitions
2. Start PXE services: t2pxe -s -R 2
3. PXE Boot DM's: t2tty -p 2
4. Shutdown pxe services when completed: t2pxe -e

### **RESTORING A DATA MOVER'S CONFIGURATION:**

1. #/nasmcd/sbin/setup\_slot -r 2
2. #/nasmcd/sbin/t2reset soft -s 2

### **RECREATING DOS PARTITION (/nas/nda1):**

1. #umount /nas/dos
2. #fdisk -l /dev/nda
3. #/mnt/source/EMC/nas/install\_init -r or #/nas/sbin/install\_init -r
4. #/nasmcd/sbin/setup\_slot -r 2 or #/nas/sbin/build\_config /nas/server/slot\_2 /nas/dos/slot\_2
5. #sync
6. /nasmcd/sbin/t2reset reboot -s 2

### **REINSTALLING NAS DATABASE:**

1. PXE Boot Datamovers are per above procedure
2. Ensure NBS Service and Partitions are accessible

3. Uninstall old EMCNAS Package: #emcrpm -r emcnas
4. Run Setup Script: #/mnt/source/EMC/nas/setup
5. Use nasdb\_restore to restore NASDB from /home/nasadmin or /nas/var/backup

## **USING LINUX RESCUE FOR FAILED NAS UPGRADE:**

1. Press F5 key at ‘Starting MS-DOS’ prompt
2. Type ‘rescue’ to startup mini linux root file system at c:> prompt
3. Create two mount points /mnt1 & /mnt2, and mount to /dev/sdc1 & /dev/sda1, respectively
4. Replace current vmlinuz & initrd.img with aok versions
5. Reboot CS and re-run NAS setup

**Note:** See Primus emc67387 & emc63091

## **RUNNING FSCK ON LINUX PARTITION:**

**#e2fsck -f /dev/nd1**

#reboot or #reboot -n -f [Force, no-sync reboot method]

## **RUNNING FSCK ON DOS PARTITION:**

1. Unmount Partition first: #umount /nas/dos
2. Run FSCK: #e2fsck -f /dev/nd1

## **CONTROL STATION BOOT PROCESS:**

1. Motherboard BIOS initialization
2. IDE Drive Discovery via Fibre or SCSI
3. BIOS for Adapter Card loads
4. DOS boots from IDE drive /dev/hda1
5. exec autoexec.bat & exec t2slot -f programs run
6. /nas/dos/vmlinuz [Loads LINUX kernel]
7. /nas/dos/bin/initrd.img [RAM disk img file]
8. LINUX completes load via scripts: /etc/rc.d/init.d [using /etc/inittab settings]
9. Root Volumes: /dev/hda3

## **DATAMOVER BOOT PROCESS:**

1. BIOS initializes & system boots from TLUN 00
2. DOS loads and reads autoexec.bat file, which then runs exec t2slot -f program to discover proper slot
3. NAS then loads using exec flashupg.exe -rUU
4. gload program, from boot.bat, loads both NAS and the ascii configuration file (boot.cfg) into memory [NAS overlays DOS]
5. Control is passed to DART’s init assembler routine called “sysinit”, which performs hardware setup, determines physical memory size, stack & memory initialization, interrupt descriptors, clock initialization, etc. MEMinit is called by ‘sysinit’ to initialize page allocator and marks ownership of pages used by BIOS, DART, data, symbol table, etc.
6. Main program starts the config thread, which begins parsing the ‘boot.cfg’ file to load the Data Mover configuration: →boot.cfg file loads buffers, parameters, devices, and interface settings; Volumes & UFSLOG are then replayed for all file systems; filesystems mounted & exported and Quotas started; network protocols started; nbs configured & started; nfs services started

## **MISC:**

### **SYR SYSTEMS REPORTING:**

SYR has the ability to collect monthly configuration & statistical information for EMC products. Celerra uses the log\_config command to collect information, which is called home monthly according to last digit of serial number that matches day.

**/nas/sbin/log\_config -d -c** [creates CallHome with payload info]

### **LOG CONFIG SCRIPT & SYR:**

#### **Monthly CallHomes via System Cron Job:**

```
# cat /etc/cron.d/nas_syslog grep log_config  
59 3 11 * * root /nas/sbin/log_config -d -c >/dev/null 2>&1
```

→So, in the above example, the log\_config script runs at 3:59 a.m. on the 11<sup>th</sup> Day of each Month, with the –default and –callhome switches, and a postevent at the end of the script generates an XML callhome to deliver the log\_config.gz file

#### **From /nas/sbin/log\_config Bourne Shell script:**

```
# Facility = SYR (143), event ID = 5, severity = INFO (6)  
$SDIR/postevent -c 6 -f 143 -s 6 -i 5 -n src_file_path -t 8 -v "$LOG" -n dest_extension -t 8 -v 'log_config.gz'
```

**Verify that log config was sent as a Callhome:**

```
# view /nas/opt/connectemc/logs/archive/ConnectEMC.archive
```

```
-- Time : 03-12-2008 12:38:27.
-- Filename : /opt/connectemc/poll/RSC_APM00083103123_120308_123724919.xml.
```

```
<?xml version="1.0" encoding="UTF-8"?>
<ConnectHome xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" SchemaVersion="1.0.8">
  <TransType>0090</TransType>
  <TransTypeDesc>EventXML, RawData</TransTypeDesc>
  <TransID>0</TransID>
  <Node ID="APM00083103123" State="Online" Status="OK">
    <Identifier>
      <ClarifyID>APM00083103123</ClarifyID>
      <SiteName>6555</SiteName>
      <Vendor>EMC</Vendor>
      <DeviceType>CELERRA</DeviceType>
      <Model>NS-120</Model>
      <SerialNumber>APM00083103123</SerialNumber>
      <WWN></WWN>
      <Platform></Platform>
      <OS>Dart</OS>
      <OS_VER>5.6.41-2</OS_VER>
      <UcodeVer></UcodeVer>
      <EmbedLevel>0</EmbedLevel>
      <InternalMaxSize>0</InternalMaxSize>
      <Comment></Comment>
    </Identifier>
    <Connection>
      <ConnectType>Dial</ConnectType>
      <AccessType>Telnet</AccessType>
      <Version></Version>
      <InPolicy></InPolicy>
      <OutPolicy></OutPolicy>
      <RouterIP></RouterIP>
      <IPAddress>10.241.168.178</IPAddress>
      <IPName></IPName>
      <NAT_IP></NAT_IP>
      <EMC_IP></EMC_IP>
      <State></State>
      <Time></Time>
      <UserID></UserID>
      <AppName></AppName>
      <UdpSocket></UdpSocket>
      <ConnectNum></ConnectNum>
      <Port>ttyS0</Port>
    </Connection>
    <HeartBeat>
    </HeartBeat>
    <InternalData>
      <EventList>
        <Event>
          <SymptomCode>5</SymptomCode>
          <Category>Status</Category>
          <Severity>Info</Severity>
          <Status>OK</Status>
          <Component></Component>
          <ComponentID></ComponentID>
          <SubComponent></SubComponent>
          <SubComponentID></SubComponentID>
```

```
<CallHome>Yes</CallHome>
<FirstTime>2008-12-03T12:37:24</FirstTime>
<LastTime>2008-12-03T12:37:24</LastTime>
<Count>1</Count>
<EventData>
```

<![CDATA[CCMD ID: **96109264901**

Brief Description: The SYR file /nas/var/log/log\_config\_APM00083103123\_081203123724.gz with log\_config.gz extension is attached.

Full Description: The automatic monthly SYR (Systems Reporting Database) data collection completed successfully.

Recommended Actions: No user action is required.

]]>

```
</EventData>
```

```
<Description>
```

<![CDATA[COMPONENT: CS\_PLATFORM , FACILITY: SYR

]]>

```
</Description>
```

```
</Event>
```

```
<Event>
```

```
<SymptomCode>5</SymptomCode>
```

```
<Category>Configuration</Category>
```

```
<Severity>Info</Severity>
```

```
<Status>OK</Status>
```

```
<Component></Component>
```

```
<ComponentID></ComponentID>
```

```
<SubComponent></SubComponent>
```

```
<SubComponentID></SubComponentID>
```

```
<CallHome></CallHome>
```

```
<FirstTime>2008-12-03T12:37:24</FirstTime>
```

```
<LastTime>2008-12-03T12:37:24</LastTime>
```

```
<Count>1</Count>
```

```
<EventData>
```

<![CDATA[APM00083201184

]]>

```
</EventData>
```

```
<Description>
```

<![CDATA[Backend Storage Serial Number(s)

]]>

```
</Description>
```

```
</Event>
```

```
</EventList>
```

```
</InternalData>
```

```
<ExternalFiles>
```

```
<FileName>/nas/var/log/RSC_APM00083103123_120308_123724919_log_config.gz</FileName>
```

```
</ExternalFiles>
```

```
</Node>
```

```
</ConnectHome>
```

```
</ConnectHome>
```

^M

-----^M

## ADMINISTRATIVE TASKS:

### CELERRA ESRS (EMC Secure Remote Support) SUPPORT:

→Supported with NAS 5.4 and higher—deployed from the following zip

<http://www.cs.isus.emc.com/csweb2/esrs/celerra.ZIP>

→Customers want this feature for secure IP-based access that provides authorization, auditing, and authentication, and reduced costs

→Ideal configuration is to have a single Server running the ESRS Gateway and Policy Manager applications, with the Server installed inside the customer's network (i.e., inside the external firewalls)

→Need connectivity between all managed devices and the ESRS Gateway server, and to EMC

→Use of the ESRS product requires a SecureID, membership in the ESRS Group, and resident on an EMC network  
→EMC Support/Partners would connect to a customer's end device via a secure tunnel through the ESRS Gateway, but not actually to the Gateway itself  
→Connection permissions are Always Allowed, Never Allow, or Always Ask. Our preference is for Always Allowed, but with Always Ask, a customer can grant access for the device. Never Allow would require a permission change to allow and then grant access

→To connect to the Gateway Server requires Webex session

→ESRS connection is made to IPs provided by one of (4) different connection islands maintained around the world by EMC

→The Gateway Server initiates the remote sessions to EMC in response to events or customer-initiated actions

→EMC uses IP over TCP via an industry standard SSL VPN encrypted tunnel, using Triple DES encryption

#### **Deploying ESRS Gateway 1.01.0x with Celerra 5.3.31 or later:**

1. Need operational Gateway and Policy Manager Servers

2. FTP to Gateway Server to download GWExt utility to create the Celerra as a target device:

```
$ ftp <IP_address_Gateway>
```

```
esrsconfig
```

```
esrsconfig
```

```
>cd /utilities/GWExt/linux
```

```
>bin
```

```
>mget *
```

```
GWExt & GWExt.ini
```

```
$ chmod 755 *
```

```
$ ./GWExt
```

3. Configure and test ConnectHome

**Note:** ESRS Gateway forwards FTP or Email ConnectHome payloads to EMC

```
# /nas/sbin/nas_connecthome -modify -ftp_ipport <IP_addr:21> -ftp_folder incoming -ftp_user onalert -ftp_passwd  
EMCCONNECT -ftp_mode passive  
# /nas/sbin/nas_connecthome -modify -ftp_ipport_2 <IP_addr:21> -ftp_folder_2 incoming -ftp_user_2 onalert -ftp_passwd_2  
EMCCONNECT -ftp_mode_2 passive  
# /nas/sbin/nas_connecthome -modify -ftp_priority 1  
# /nas/sbin/nas_connecthome -modify -encryption_enabled yes  
# /nas/sbin/nas_connecthome -info  
# /nas/sbin/nas_connecthome -test -ftp_1
```

#### **USING WEBEX SUPPORT SESSION:**

1. Open Web Browser and go to <http://emcsupport2.webex.com> (or https://emcsupport.webex.com)

2. Click on My WebEx Navigator bar and Log In using NTlogon name and pass123

3. Click on Start a Support Session and record Session Number [#751953066]

4. Provide Session ID Number to customer

5. Customer enters ID number and customer info will be displayed on your screen

6. Click on Desktop tab to take control of Remote System to access Celerra via Telnet, SSH, etc.

**Other Passwords:** pass123

#### **CREATING FILTERS FOR WIRESHARK (formerly Ethereal) TRACES:**

Filter: ip.addr == 138.127.59.219 [To filter by a given IP Address]

```
ip.src == 10.1.1.32 ip.dst == 10.1.1.32
```

Filter: smb.nt\_status == 0xc0000022 [To filter by NT Status Error Message—‘STATUS\_ACCESS\_DENIED’]

**Note:** Click on ‘Filter’ → Add Expression → Drill down to “SMB” & select “NT Status” → Select ‘Relation’ of “==” and then choose the appropriate error code or plug in a known NT Error Status Code.

#### **MOUNTING ISO IMAGE WITHOUT CD-ROM ON WINDOWS OR LINUX:**

For Windows, download sw from [www.daemon-tools.cc/dtcc/portal/download.php](http://www.daemon-tools.cc/dtcc/portal/download.php) and you can mount iso images as virtual drives.

For Linux, do following: **#mount /tmp/toolsandapps5.3.15.3.iso /mnt -t iso9660 -o loop=/dev/loop0**