



Hewlett Packard
Enterprise

Cray ClusterStor E1000 Administration Guide (4.x) (S-2761)

Cray ClusterStor E1000 Administration Guide 4.x

© Copyright

Notices

The information contained herein is subject to change without notice. The only warranties for Hewlett Packard Enterprise products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. Hewlett Packard Enterprise shall not be liable for technical or editorial errors or omissions contained herein.

Confidential computer software. Valid license from Hewlett Packard Enterprise required for possession, use, or copying. Consistent with FAR 12.211 and 12.212, Commercial Computer Software, Computer Software Documentation, and Technical Data for Commercial Items are licensed to the U.S. Government under vendor's standard commercial license.

Links to third-party websites take you outside the Hewlett Packard Enterprise website. Hewlett Packard Enterprise has no control over and is not responsible for information outside the Hewlett Packard Enterprise website.

Acknowledgments

Intel®, Itanium®, Optane™, Pentium®, Xeon®, Intel Inside®, and the Intel Inside logo are trademarks of Intel Corporation or its subsidiaries.

AMD and the AMD EPYC™ and combinations thereof are trademarks of Advanced Micro Devices, Inc.

Microsoft® and Windows® are either registered trademarks or trademarks of Microsoft Corporation in the United States and/or other countries.

Adobe® and Acrobat® are trademarks of Adobe Systems Incorporated.

Java® and Oracle® are registered trademarks of Oracle and/or its affiliates.

UNIX® is a registered trademark of The Open Group.

All third-party marks are property of their respective owners.

Table of contents

1	About Cray ClusterStor E1000 Administration Guide	
2	Command Prompt Conventions	
3	ClusterStor E1000 Terms, Abbreviations, and Definitions	
4	Cray ClusterStor E1000 Architecture	
4.1	Hardware Architecture	
4.2	Software Architecture	
5	Components and Hardware Supported by this Release	
6	Functionality Supported in Each Cray ClusterStor Release	
6.1	Qualified Functionality in Cray ClusterStor 4.4	
6.2	Qualified Functionality in Cray ClusterStor 4.3	
6.3	Qualified Functionality in Cray ClusterStor 4.2	
6.4	Qualified Functionality in Cray ClusterStor 4.1	
7	Software Versions and Requirements	
8	Power On ClusterStor E1000 System	
9	Power Off ClusterStor System	
10	Spectre/Meltdown Mitigation on Cray ClusterStor Systems	
11	Log In to CSM	
12	RAS – Reliability, Availability, and Serviceability	
12.1	Configure Service Console Access	
12.2	Configure SMTP Relay, RAS Email, Remote Support, and Service Console	
12.3	Configure RAS System Settings	
12.4	Configure the REST API	
12.5	Interpret RAS Email Notifications	
12.6	Use Guided Replacement Procedures with the Service Console	
12.7	service_console Commands	
13	Customize Lustre Configuration	
13.1	Prepare the Local Configuration	
13.2	Lustre Server: Recommended Tuning Parameters	
13.3	Create Custom LNet Configuration for ClusterStor	
13.4	Save/Restore Lustre Filesystem Parameters	
14	Change Network Settings in ClusterStor 4.x	
14.1	ean Command	
14.1.1	ean apply Subcommand	
14.1.2	ean show Subcommand	
14.1.3	ean dns Subcommand	
14.1.4	ean dns clear Subcommand	
14.1.5	ean dns set Subcommand	
14.1.6	ean ipaddr Subcommand	
14.1.7	ean ipaddr set Subcommand	
14.1.8	ean ipaddr show Subcommand	
14.1.9	ean mmu Subcommand	
14.1.10	ean mmu add Subcommand	

- 14.1.11 ean mmu delete Subcommand
- 14.1.12 ean mmu show Subcommand
- 14.1.13 ean ntp Subcommand
- 14.1.14 ean ntp clear Subcommand
- 14.1.15 ean ntp set Subcommand
- 14.1.16 ean primary Subcommand
- 14.1.17 ean primary add Subcommand
- 14.1.18 ean primary delete Subcommand
- 14.1.19 ean primary show Subcommand
- 14.1.20 ean route Subcommand
- 14.1.21 ean route add Subcommand
- 14.1.22 ean route clear Subcommand
- 14.1.23 ean route delete Subcommand
- 14.1.24 ean route load Subcommand
- 14.1.25 ean route set Subcommand
- 14.1.26 ean route show Subcommand
- 14.1.27 ean secondary Subcommand
- 14.1.28 ean secondary add Subcommand
- 14.1.29 ean secondary delete Subcommand
- 14.1.30 ean secondary show Subcommand

14.2 Change Externally Facing IP Addresses

14.3 Change LDAP or NIS Settings in Daily Mode

- 14.3.1 lustre users ldap set Subcommand
- 14.3.2 lustre users nis set Subcommand

14.4 Change Bonding Mode Type

15 Check File Systems Using e2fsck Command

16 Support Bundles

- 16.1 Use CSM GUI to Collect Support Bundles Manually
- 16.2 Collect Support Bundles Manually via CSCLI
- 16.3 Import a Support File
- 16.4 Download a Support File
- 16.5 Delete a Support File
- 16.6 View Support Files
- 16.7 Interpret Support Bundles
- 16.8 support_bundle Command
 - 16.8.1 support_bundle collect Subcommand
 - 16.8.2 support_bundle show Subcommand
 - 16.8.3 support_bundle export Subcommand
 - 16.8.4 support_bundle set Subcommand
 - 16.8.5 support_bundle delete Subcommand

17 CSCLI Command Reference Overview

- 17.1 CSCLI Reference (Release 4.4)
- 17.2 CSCLI Reference (Release 4.3)
- 17.3 CSCLI Reference (Release 4.2)

17.4 CSCLI Reference (Release 4.1)

17.5 CSCLI Reference (Release 4.0)

18 GEM CLI Commands

18.1 Log In to GEM (GOBI)

18.2 Run GEM Commands from the Node with fwdataloader

18.3 ddump

18.4 getboardid

18.5 getvpd

18.6 help

18.7 ipmi_power

18.8 logdump

18.9 report_faults

18.10 ver

19 Websites

20 Support and other resources

20.1 Accessing Hewlett Packard Enterprise Support

20.2 Accessing updates

20.3 Remote support

20.4 Warranty information

20.5 Regulatory information

20.6 Documentation feedback

About Cray ClusterStor E1000 Administration Guide

The Cray ClusterStor E1000 Administration Guide (4.x) S-2761 provides procedures to help administrators maintain Cray ClusterStor E1000 storage systems running software release 4.4 and subsequent 4.x releases.

Scope

This publication contains procedures and reference information to support Cray ClusterStor E1000 storage systems running software release 4.4 and subsequent 4.x releases.

IMPORTANT: About CCLI information in this document: CCLI command changes are summarized in this document. For a comprehensive list of CCLI commands, see the ClusterStor CCLI Command Reference Guide S-9922.

Record of Revision

Publication	Date	Description
Cray ClusterStor E1000 Administration Guide (4.x) S-2761	September 2021	Updated sections: <ul style="list-style-type: none">Power Off ClusterStor SystemCheck File Systems Using <code>e2fsck</code> Command
Cray ClusterStor E1000 Administration Guide (4.x) S-2761	June 2021	Publication for software release 4.4 and subsequent 4.x releases on Cray ClusterStor E1000 systems

Audience

The procedures presented in this publication are for administrators of Cray ClusterStor E1000 storage systems.

Related Publications

Publication Number	Description
Cray ClusterStor E1000 Administration Guide (4.3) S-2760	Publication for software release 4.3 on Cray ClusterStor E1000
Cray ClusterStor E1000 Administration Guide (4.2) S-2758	Publication for software release 4.2 on Cray ClusterStor E1000
ClusterStor E1000 Administration Guide (4.1) S-2757	Publication for software release 4.1 on Cray ClusterStor E1000
ClusterStor and Sonexion System Snapshot Analyzer (SSA) User Guide S-2561	<p>Refer to the SSA User Guide to download, install, and use Cray System Snapshot Analyzer (SSA) software.</p> <p>Cray system snapshot analyzer (SSA) software is support analytics technology that securely collects, analyzes, and uploads (if upload is enabled) product health, configuration, or triage information about a ClusterStor system to the Cray service organization. After being captured and uploaded by SSA, the data is analyzed using a sophisticated analytics platform to detect and enumerate changes over time, detect changes in the health state of various aspects of a system, or process triage information to assist with case resolution. Through automation, SSA improves the overall customer experience by reducing the manual effort and time required to report and resolve support issues.</p> <p>The SSA shepherd is the client software that manages the collection, first-level analysis, and secure transport of support telemetry information.</p>
ClusterStor Data Services Administration Guide S-1237	Procedures to help users and administrators use, configure, and maintain ClusterStor Data Services (CDS) deployments
ClusterStor Data Services Installation Guide S-1238	Procedures to install the ClusterStor data services software onto compatible servers at a customer site
Cray ClusterStor Data Services User Guide S-1239	Useful information and procedures for users of Cray ClusterStor systems with ClusterStor data services
ClusterStor CCLI Command Reference Guide S-9922	CCLI command reference from release 2.x to present

Publication Number	Description
View for ClusterStor Administration Guide S-3026	Procedures for using the View for ClusterStor graphical user interface, which provides a visual look at metrics and health information, collected at different points in the past, through data dashboards and workflows
Cray ClusterStor Software Release Installation Guide 4.x S-2598	Installation instructions for ClusterStor software release 4.x
Cray ClusterStor E1000 Field Installation Guide H-6203	Instructions for setting up, configuring, and testing Cray ClusterStor E1000 storage systems
Cray ClusterStor E1000 SSU and MDU Additions H-6204	Instructions for adding additional SSU-F, MDU (Metadata Unit), and SSU-D enclosures to a storage system in the field. These additional enclosures are supported by Cray ClusterStor E1000 systems running software release 4.x.

Typographic Conventions

Monospace	A <code>Monospace</code> font indicates program code, reserved words or library functions, screen output, file names, path names, and other software constructs
Monospaced Bold	A bold monospace font indicates commands that must be entered on a command line.
<i>Oblique or Italics</i>	An <i>oblique or italics</i> font indicates user-supplied values for options in the syntax definitions
Proportional Bold	A proportional bold font indicates a user interface control, window name, or graphical user interface button or control.
<code>Alt-Ctrl-f</code>	<code>Monospaced</code> hyphenated text typically indicates a keyboard combination

Copyright and Trademarks

© 2021 Hewlett Packard Enterprise LP. All trademarks used in this document are the property of their respective owners.

Command Prompt Conventions

Host Names and Accounts in Command Prompts

The host name in a command prompt indicates where the command must be run. The account that must run the command is also indicated in the prompt.

- The `root` or super-user account always has the `#` character at the end of the prompt.
- Any non-`root` account is indicated with a `$`. A user account that is neither `root` nor `admin` is referred to as `user`.

MGMT0#	Run the command on the primary ClusterStor management node as <code>root</code> .
MGMT0\$	Run the command on the primary ClusterStor management node as <code>admin</code> .
MGMT1#	Run the command on the secondary ClusterStor management node as <code>root</code> .
MGMT1\$	Run the command on the secondary ClusterStor management node as <code>admin</code> .
OSS#	Run the command on an OSS node as <code>root</code> .
OSS\$	Run the command on an OSS node as a non- <code>root</code> user.
MGS#	Run the command on an MGS node as <code>root</code> .
MGS\$	Run the command on an MGS node as a non- <code>root</code> user.
MDS#	Run the command on an MDS node as <code>root</code> .
MDS\$	Run the command on an MDS node as a non- <code>root</code> user.
client\$	Run the command on a Lustre client node as any non- <code>root</code> user.
client#	Run the command on a Lustre client node as <code>root</code> .
user@hostname\$	Run the command on the specified system as a non- <code>root</code> user.
mgmt-sw#	Run the command on a management switch as an admin-level user.
mgmt-sw\$	Run the command on a management switch as a non-admin user.

Sample commands and command output used throughout this publication are shown with a generic file system name of `cls12345`.

Lustre file system names

The name of the Lustre file system seen in command examples is `cls12345`, and is mounted by Lustre clients at `/lus`, as demonstrated by the command below:

```
client$ lfs df -h
UUID                               bytes      Used  Available  Use% Mounted on
cls12345-MDT0000_UUID              2.0T      59.1G    1.9T      4% /lus[MDT:0]
cls12345-MDT0001_UUID              2.0T      97.6M    1.9T      1% /lus[MDT:1]
cls12345-OST0000_UUID             112.0T      5.1T   105.8T      5% /lus[OST:0]
cls12345-OST0001_UUID             112.0T      5.1T   105.8T      5% /lus[OST:1]
cls12345-OST0002_UUID              15.3T    425.8G    14.8T      3% /lus[OST:2]
cls12345-OST0003_UUID              15.3T    423.4G    14.8T      3% /lus[OST:3]

filesystem_summary:              254.6T    11.0T   241.1T      5% /lus
```

Directory Path in Command Prompt

Example prompts do not include the directory path, because long paths can reduce the clarity of examples. Most of the time, the command can be executed from any directory. When it matters which directory the command is invoked within, the `cd` command is used to change into the necessary directory.

For example, here are actual prompts as they may appear on the system:

```
client:~ # cd /etc
client:/etc# cd /var/tmp
client:/var/tmp# ls file
```

or




```
[root@cls12345n000 ~]# cd /etc  
[root@cls12345n000 etc]# cd /var/tmp  
[root@cls12345n000 tmp]# ls file
```

And here are the same prompts as they appear in this publication:

```
client# cd /etc  
client# cd /var/tmp  
client# ls file
```

or

```
cls12345n000# cd /etc  
cls12345n000# cd /var/tmp  
cls12345n000# ls file
```

ClusterStor E1000 Terms, Abbreviations, and Definitions

This glossary of key ClusterStor terms relates to software release 4.x.

Term	Definition
5U84 G2 / 5U84 G3 (introduced in 3.0 SU006)	The second and third generation versions of the standard 5U84 enclosure, used for ClusterStor SSUs and ESUs. The 5U84 G2 chassis has several hardware changes, including an improved LED display and drawer release, redesigned side card cover, reduced number of sensors, and a different fan module latch color. The 5U84 G3 enclosure increases the amount of backend SAS bandwidth to ~12 GB/sec vs. the 5U84 G2, which has a backend SAS bandwidth of ~8.4 GB/sec.
Base MDU, Base Metadata Management Unit	The base MDU is the MDU that is always installed in the base rack. The base MDU provides two (2) MDS nodes along with two (2) MDTs. MDT0 functions as the default MDT and as the root MDT for DNE phase 1. MDT1 requires DNE for use.
Base Rack	The first rack in a ClusterStor storage cluster that contains the SMU and base MDU along with the rack networking infrastructure and from 1 to 6 SSUs
CCI, Cray ClusterStor Installer	ClusterStor software used for manufacturing and installing ClusterStor systems
CCMS, Cray ClusterStor Management Server	ClusterStor MGMT node; the primary and secondary instances of the CSM software and all associated components and services running on the CMU
CLI, Command Line Interface	A text-based interface, used to operate software and operating systems
ClusterStor Rack	See Storage Rack.
CMA, Cable Management Arm	A bidirectional, swiveling cable arm that allows both server modules and PSUs to be serviced. The CMA also allows the system to slide out to access the fan lid and HDDs. Note that the CMA can be difficult to maneuver when full of cables (especially copper).
CMU, Cluster Management Unit	See SMU.
CMU Storage	Deprecated term. A storage enclosure dedicated to the CMU on ClusterStor platforms that preceded the E1000
Critical, Critical Array State	The state of a GridRAID or MDRAID array where the subsequent failure of one more storage components may lead to the data becoming inaccessible
CSI, Cray Sonexion Installer	ClusterStor software used for manufacturing and installing ClusterStor systems
CSM, ClusterStor Manager Manager	ClusterStor platform, software and hardware management system
CTU, Customer Test Unit	CTU features are enabled in a release for customer evaluation, early access, and acceptance testing, but CTU features are not supported for production systems.
Data Block	A component of a “parity group” (or “stripe”) containing actual user data, also referred to as a “data chunk” or “data unit”
Degraded, Degraded Array State	The state of a GridRAID or MDRAID array operating with one (1) failed storage component
Distributed Spare, Distributed Spare Volume	The aggregate collection of distributed spare data blocks in a GridRAID array that comprises a single logical spare volume for the specific GridRAID array that contains it. Each distributed spare contains the equivalent of one physical drive’s worth of distributed spare space and is used as the target of the GridRAID reconstruction process and the primary data source for the GridRAID rebalance process.

Term	Definition
DMN, Dual LMN	Refers to the “Dual Local Management Networks” (or “Dual Management Networks”) feature supported beginning with the ClusterStor 1.5.0 release
DNE, Distributed Namespace	Lustre DNE Phase 1 feature supported in Lustre 2.5 and after, which allows multiple MDS/MDT components to operate within a single Lustre file system
EAC, Embedded Application Controller	See Server Module.
EAN, External Administration Network	Customer administration network, external to the ClusterStor system. Connected to the storage system's MGMT nodes to provide access to the CSM software.
ECN, Enterprise Client Network	Refers to the 10GbE or 40GbE data network connecting non-Lustre enterprise clients to the optional CIFS NFS Gateway (CNG). The CNG feature is not supported in ClusterStor E1000 systems.
ESM, Embedded Server Module	Deprecated term for an Embedded Application Controller (EAC) because it implies general server functionality that is not supported on the dedicated ClusterStor EACs. See Server Module.
ESU, Expansion Storage Unit	See HDD Disk Array.
Expansion Rack	See Storage Rack.
Failed, Failed Array State	The state of a GridRAID array that has been failed by the system
Faulted	(Definition forthcoming)
GB/sec, GigaBytes per Second	10 ⁹ Bytes per second
Gbit/sec, Gb/sec, Gigabit per Second	10 ⁹ bits per second
GbE, Gigabit Ethernet	Ethernet standard that transmits at 1 gigabit per second.
GridRAID	ClusterStor implementation of parity declustered RAID. A RAID level organization that combines RAID 6 data protection with a declustering methodology. GridRAID overcomes single drive throughput bottlenecks by distributing parity groups and spare space across all storage components in an array.
HDD Disk Array	(Definition forthcoming)
HHHL	half-height, half-length
ICL, Inter-Controller Link	A link that connects two controllers or two servers together. Used in ClusterStor as a dedicated HA communication path.
ISL, Inter-Switch Link	A connection between two related switches
KiB, Kibibyte	1024 bytes
LCN, Lustre Client Network	High speed customer data network through which Lustre clients may connect to the ClusterStor Lustre file system, which is also connected to the LCN via the storage system's Local Data Switches (LDS).
ldiskfs, Lustre Disk File System	Lustre version of a patched Ext4 file system.
LDN, Local Data Network	A dual InfiniBand or 100/200GbE network with switches installed in all racks, connecting all servers and enclosures as needed and used as uplink points to the end user client infrastructure.
LDS, Local Data Switch	An InfiniBand or 100/200GbE network switch installed in a ClusterStor rack as part of the LDN and used for providing high speed data connectivity. Used as uplink points to the end user client infrastructure.
LMN, Local Management Network	A private 1GbE network connecting all ClusterStor servers and enclosures

Term	Definition
LMS, Local Management Switch	A 1GbE switch installed in a ClusterStor rack as part of the LMN and used for providing private management network connectivity for all ClusterStor servers and enclosures.
Lustre®	Open source clustered file system
Lustre Servers	The set of Lustre servers that comprise the Lustre file system; includes the MGS, MDS, and multiple OSS servers
MDS, Metadata Server	Lustre server component that manages the Lustre file system metadata
MDT, Metadata Target	Lustre component, a storage volume that holds the Lustre file system metadata
MDU, Metadata Unit	(Definition forthcoming)
MGMT, Management Server Node	One (1) of two (2) ClusterStor management servers, operating as an HA pair that provide management functions for the storage cluster
MGMT0	The primary ClusterStor management server, typically used for web access and SSH logins for managing the storage cluster
MGMT1	The secondary management server, typically used to provide boot services to nodes in the storage cluster
MGS, Management Server	Lustre server component that manages the Lustre MGT
MGT, Management Target	Lustre component, the storage volume holding the Lustre file system management data that allows clients to discover, mount, and operate the file system
MMU, Metadata Management Unit	See MDU.
NIS, Network Information Service	Maintains and distributes a central directory of user and group information in a network
Normal, Normal Array Activity	Characterizes the activity of a GridRAID or MDRAID array that is engaged in processing I/O only and is not conducting any recovery, sync, or RAID checking activities.
Offline, Array Is Offline	The array is not available.
Optimal, Optimal Array State	The state of a GridRAID or MDRAID array where all drives in the array are operational without the involvement of spare volumes or dedicated hot spares. For GridRAID, this is equivalent to the “Redundant 0/2” terminology.
OSS, Object Storage Server	Lustre server component that operates and manages the Lustre OSTs
OST, Object Storage Target	Lustre component, a storage volume that holds Lustre file system data
Parity Block	(GridRAID) Component of a parity group that contains protection information for the group derived from the set of data blocks in the parity group. Also referred to as a “parity chunk” or “parity unit.”
Parity Group	(GridRAID) The set of “data blocks” and derivative “parity blocks” that together comprise a protected data set. Also referred to as a “stripe.”
PB, Petabyte	10 ¹⁵ bytes
PDU	Power Distribution Unit
RAID Check, RAID Consistency Check	In this process, the system periodically checks that the parity information is consistent for every “parity group” (stripe) in the array. This process is sometimes referred to as “parity scrubbing.”
RAS (Reliability, Availability, Serviceability) System	ClusterStor feature providing system RAS diagnostic and repair features

Term	Definition
Rebalance, Rebalance Process	Phase 2 of the 2-phase GridRAID recovery process: a GridRAID array essentially copies reconstructed data from a distributed spare volume in the array to a physical replacement drive, freeing the distributed spare volume when complete for future reuse.
Rebalancing, Rebalancing Array Activity	Characterizes the activity of a GridRAID array that is engaged in the rebalance phase of the recovery process
Rebuild, Rebuild Process	In this single-phase recovery process, an MDRAID array reconstructs data for a failed drive and copies it to a dedicated replacement drive.
Reconstructing, Reconstructing Array Activity	Characterizes the activity of a GridRAID array that is engaged in the reconstruction phase of the recovery process
Reconstruction, Reconstruction Process	Phase 1 of the 2-phase GridRAID recovery process: a GridRAID array reconstructs the data from a missing storage component onto one of the distributed spare volumes.
Recovering, Recovering Array Activity	Characterizes the activity of a GridRAID or MDRAID array that is engaged in the recovery process
Recovery, Recovery Process	In this process, a GridRAID or MDRAID array recovers from a storage component failure.
SED, Self-Encrypted Drive	A disk drive that automatically encrypts/decrypts data to/from the media
Server Module	The server canister in a 2U24 PCIe Gen 4 NVMe storage enclosure. The configuration of the Server Modules are different in the various use cases (SMU, MDU, SSU-F, SSU-D1, SSU-D2, and SSU-D4.)
SMU, Storage Management Unit	A 2U24 enclosure with dual EACs that provides two MGMT nodes and associated storage. There is always only one SMU in a ClusterStor file system cluster, and it is always installed in the base rack. In conjunction with the base MDU, the SMU replaces the functionality of the earlier ClusterStor CMU component.
Spare Volume, GridRAID Spare Volume, or Distributed Spare Volume	The aggregation of the equivalent of one drive's worth of distributed spare space considered collectively as a logical spare drive or volume and used as the target of the GridRAID repair operation
SSU, Scalable Storage Unit	For E1000, SSU refers to a 2U24 NVMe head connected to zero (0) or more 4U106 EBODs. Provides dual OSSes. Also see SSU-D, SSU-F.
SSU Addition	Refers to the process of increasing the storage capacity of a ClusterStor file system by incorporating additional SSUs into the cluster.
SSU-D, Scalable Storage Unit-Disk (SSU-D1, SSU-D2, SSU-D4)	(Definition forthcoming)
SSU-F, Scalable Storage Unit-Flash	(Definition forthcoming)
Storage Component	Refers to an individual drive when considered as part of a configured GridRAID or MDRAID array
Storage Rack	The additional racks (to the base rack) in a ClusterStor storage cluster that contain the rack networking infrastructure and some number of SSUs. Sometimes called "expansion rack."
Stripe	See "Parity Group."
TB, Terabyte	10 ¹² bytes
zpool scrub (ZFS), GridRAID	See Raid Check.

Cray ClusterStor E1000 Architecture

Cray ClusterStor systems consist of a unique scale-out storage architecture that consolidates Lustre servers, RAID controllers, disk enclosures, and the operating system into an integrated storage platform that can be easily expanded using modular storage node building blocks. The Cray ClusterStor E1000 architecture is based on a No (Hardware) Single Point of Failure (NSPF) design.

The basic hardware components are represented by a single rack system and consist of the following hardware building blocks:

- Storage Rack
- Power Distribution Unit (PDU)
- High Speed Network (HSN)
- Management Network(s)
- Processor Control Module
- Storage Enclosure, Flash, NVMe
- Storage Enclosure, Disk, HDD

The first Cray ClusterStor E1000 hardware platform ships with the Neo 4.x code stack.

View for ClusterStor 1.3 is used to manage the system from end to end, beyond hardware monitoring to see specific jobs running on the system for troubleshooting and problem solving. (See the View for ClusterStor Administration Guide S-3026.)

ClusterStor data services moves data from flash storage to HDD storage and back, as needed. (See the ClusterStor Data Services User Guide S-1237.)

Hardware Architecture

Lustre is the first filesystem use case for ClusterStor E1000 hardware.

ClusterStor E1000 hardware is a configurable, rack-level, expandable storage cluster. The common building block is a pair of high availability (active/active) servers that are configured in a single 2U-rack-height chassis. Both servers are connected via a common backplane to provide symmetric access to enclosure disks. If one server fails, the second server will provide the compute function for the SSDs/HDDs that were assigned to the failed server, providing seamless high availability. The dense HDD enclosure contains 106 drives in four (4) rack units (4U106). A single 4U106 disk enclosure provides over a petabyte of usable storage with 14TB drives. ClusterStor E1000 comes in multiple base configurations: all flash, all HDD, hybrid flash and HDD,

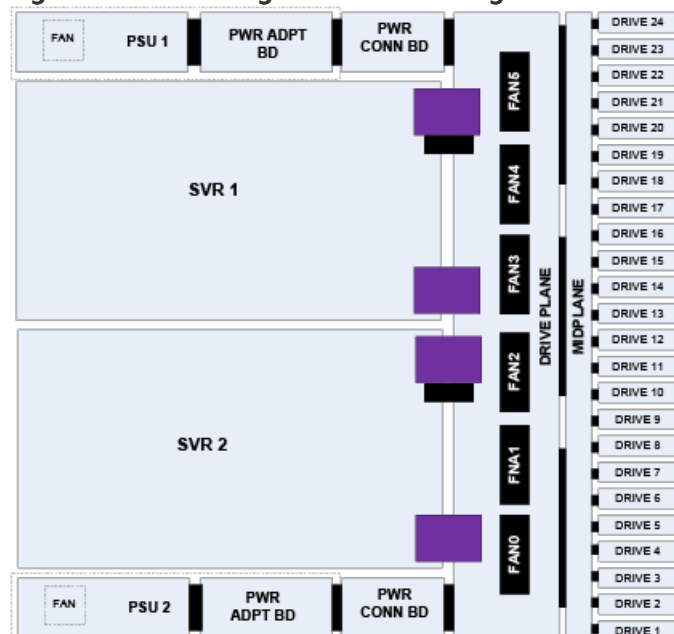
2U24 Storage Server

Each 2U24 storage server enclosure has up to 24 NVMe drives. The 2U24 server chassis are configured with variations of CPU and memory configurations. There are two (2) PCIe generation 4 (4x16) HHHL and one (1) OCP 3.0 (4x16) for each active server within an enclosure. The active server PCIe slots are populated with Slingshot 200 Gbit Ethernet or Mellanox 200 Gbit HDR Connectx-6 InfiniBand / 200 Gbit Ethernet network interface cards. The PCIe slots will be configured with 12Gbit SAS host bus adapters when supporting the HDD use case.

The 2U24 storage server is used as the system management unit (SMU) and as the metadata unit (MDU). It is also used as the disk storage controller (SSU-D) and the flash scalable storage unit (SSU-F). Depending upon the use case, the configuration for the 2U24 storage server will vary slightly. Overall, it will have the following:

- Two (2) embedded storage server modules with one (1) PCIe 4.0 AMD Rome CPU each
- Up to six (6) 100/200 Gbps PCIe 4.0 NICs
- Up to 24 x 2.5" NVMe PCIe 4.0 SSDs in two (2) rack units
- Up to 60 GB/s write and 80 GB/s read throughput in 2U
- Up to 230 TB useable capacity in 2U
- Six (6) fan modules
- Two (2) power supply units (PSUs)
- Flash-optimized Lustre

Figure 1: 2U24 Storage Server Block Diagram



4U106 Disk Enclosure

In four (4) 4U106 disk enclosures, there are up to 424 HDDs attached to a single 2U24 server acting as the controller. This active/active JBOD enclosure has 106 3.5" HDDs up to 16TB each and SAS 12G architecture. There are 96 drives on four (4) 24-drive baseplanes. There are an additional 10 drive on a 10-drive baseplane on the left side of the enclosure in the controller channel. In the controller channel, there are two (2) IO modules. On the right side of the enclosure, there are eight (8) SAS expander cards with 4+4 redundancy. In addition, there are two (2) PSUs with 1+1 redundancy and four (4) system fan modules, plus two (2) controller channel fan modules. A fully-loaded 4U106 disk enclosure is about 320 pounds (~145 kilograms).

Racks

Two (2) racks are used with the ClusterStor E1000:

1. Cray-designed rack for HDD-based systems to support up to 10 4U106 disk enclosures' weight
This rack has a 3,000-pound dynamic load capability and 4,000 static load capability. The rack measures 42U: 600 x 1300mm (600 x 1200 mm base cabinet with a 1000 mm extension). It includes an anti-tip plate. The front and rear doors are compatible with the Shasta Air-cooled APC rack doors.
2. Shasta Air-cooled APC 42U and 48U racks
*These Shasta racks DO NOT support the 4U106 disk enclosure. They are only for all-flash server systems.

These racks support PDU servicing with the ClusterStor PDU options installed.

Each rack option fits on two (2) 24-inch data center floor tiles. All have have perforated air-cooled front and rear doors, casters and leveling feet, overhead or underfloor power routing, and the option option for Motivair M8 and M12 chilled rear doors.

When constructing a ClusterStor E1000 storage system, the first rack is the base rack. This rack includes that management and data network switches at the top; then the SMU and the MDU, followed by the appropriate SSUs. There are three (3) main base rack options:

1. Capacity disk: eight (8) 4U106 disk enclosures behind two (2) SSU-Ds in a single 42U rack in an SSU-D+4 configuration
2. Performance disk: six (6) 4U106 disk enclosures behind three (3) SSU-Ds in a single 42U rack
3. All flash: 16 SSU-Fs in a 42U rack

Other hybrid options can exist.

For more about the Addition hardware, see ClusterStor E1000 SSU and MDU Additions H-6204 .

2U24 Storage Server Use Cases

System Management Unit (SMU)	ClusterStor management (customer access, boot, telemetry, and more) Required component
Metadata Unit (MDU)	Stores data about the data Scalable One (1) is required per system; additional are based on scaling needs
Flash Scalable Storage Unit (SSU-F)	Flash option to store data Highest throughput and IOPs performance for ClusterStor
Disk Scalable Storage Unit (SSU-D #)	Disk option to store data Controller for up to four (4) 4U106 disk enclosures # = amount of 4U106 disk enclosures in the SSU

Networks

The ClusterStor E1000 has five (5) networks:

- External administration network (EAN) – Customer administration network to the ClusterStor system. The EAN provides customer access to the ClusterStor and leverages the customer's authentication solution. It also provides access for the administrator using the ClusterStor Manager GUI.
- Local management network (LMN) – Dual-bonded GigE local management private network that provides the management connections within the ClusterStor system. Is is also referred to as the low speed network (LSN).
- Local data network (LDN) – Provides high speed connections for data movement withing the ClusterStor and to/from any supercompute or other systems. The LDN is also referred to as the high speed network (HSN). It is a dual fault-tolerant local data network with external uplinks. Server Module (EAC) 0 in each SMU, MDU, and SSU is connected to LDN 0 switch, and Server Module (EAC) 1 is connected to LDN 1 switch.
- Lustre client network (LCN) – Customer Lustre client network—an abstraction layer over the LDN for the Lustre file system.
- Data management network (DMN) – The serial attached SCSI (or SAS) connection from the SSU-D to the 4U106 disk enclosures.

Also referred to as the SAS network.

Switch Options

The ClusterStor E1000 has options for the switches use for both the LDN and the LMN.

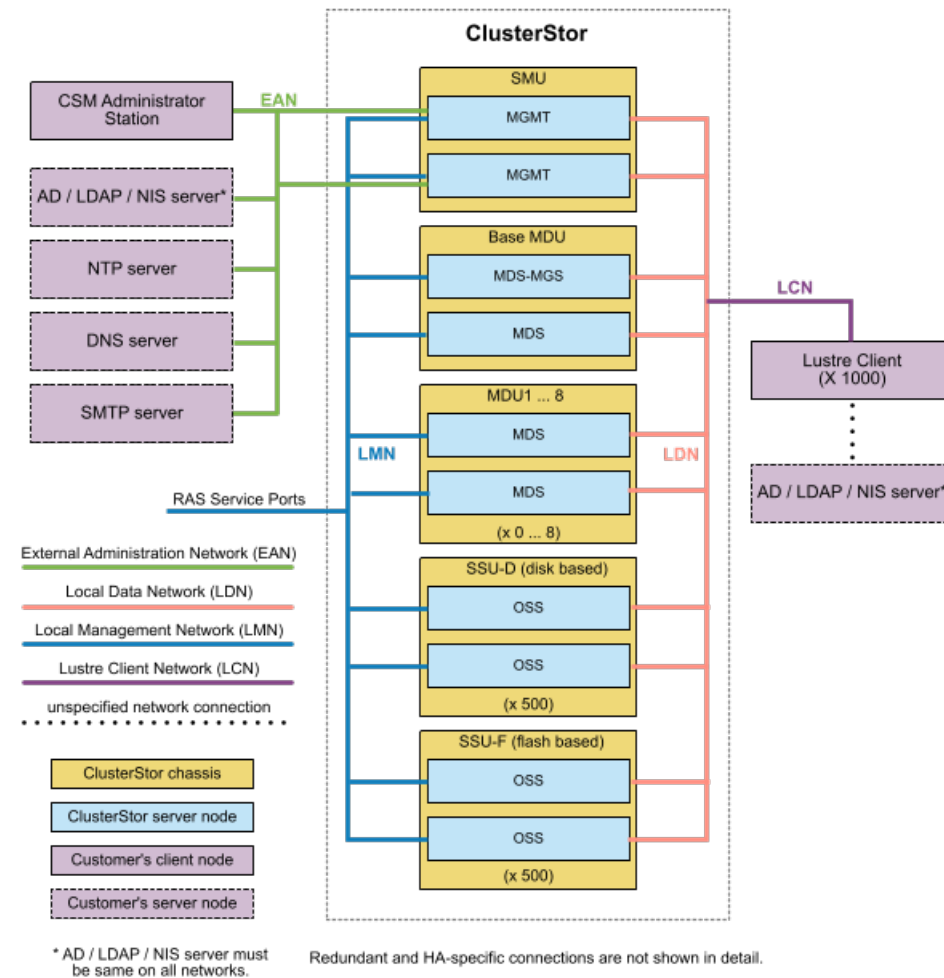
LDN/HSN Switch Options	LMN/LSN Switch Options
HDR Infiniband: Mellanox QM8790 1U HDR 200 Gb/s Infiniband switch	Dell PowerSwitch S3048-ON 48-port GbE switch (initial hardware release only; no longer standard)
Slingshot Top of Rack (ToR) switch (delayed availability)	Dell EMC Networking S4112F-ON half-width 10 GbE switch (initial hardware release only; no longer standard)
	Dell EMC Networking S4112T-ON half-width 10 GbE switch (initial hardware release only; no longer standard)

Switches are required in the base rack for both network types, but switches are optional in the expansion racks. A system needs additional switches and expansion racks when there are no more available ports on the existing switches.

Software Architecture

The system software architecture consists of an integrated, multilayer software stack that includes ClusterStor Manager (CSM), management (MGMT) nodes, Lustre file system, Lustre management server (MGS), Lustre metadata server (MDS), Lustre object storage server (OSS), Lustre clients, data protection layer (RAID), unified system management software (GEM), and Linux OS.

Figure 2: ClusterStor E1000 Network Overview



Management (MGMT) Nodes

Cray ClusterStor systems use primary and secondary MGMT nodes to manage the CSM software, ClusterStor boot services, and associated web services that provide cluster installation, configuration, baseline testing, management, and monitoring functions.

Lustre File System

The Lustre file system is a client/server based, distributed architecture that offers extreme I/O performance and unparalleled scalability, making it a popular choice as a sitewide global file system in the HPC sector, serving dozens of clusters. Integrating Lustre into the Cray ClusterStor system enables users to take full advantage of the application performance of an HPC cluster environment without the traditional challenges of architectural complexity, system integration, and solution cost.

A Lustre cluster is an integrated set of servers that process metadata, and servers that store data objects and manage free space. Together, the metadata and object storage servers present Lustre file systems to clients. A Lustre cluster includes the following components:

- Lustre Management Server (MGS)
- Lustre MetaData Server (MDS)
- Lustre Object Storage Servers (OSS)
- Local Data Network (LDN)
- Lustre Client Network (LCN)
- Lustre Clients
- Data Protection Layer (RAID)

- Unified System Management Software (GEM)
- Linux OS

ClusterStor 4.x systems run patched Lustre 2.12 software, supporting LDISKFS (and eventually ZFS).

(Also see [Functionality Supported in Each Cray ClusterStor Release](#).)

Lustre Management Server (MGS)

The MGS stores configuration information for a single Lustre file system in a cluster. Each Lustre server contacts the MGS to provide information. Each Lustre client contacts the MGS to retrieve information.

MetaData Server (MDS)

The MDS (sometimes co-located with the MGS) makes metadata available to Lustre clients from one (1) or more Metadata Target (MDTs). The MDT stores file system metadata (filenames, directories, permissions, and file layouts) on disk and manages the namespace. The MDS provides network request handling for the file system.

Object Storage Servers (OSS)

The OSS provides file I/O data services and network request handling for one or more local Object Storage Targets (OSTs). The OST stores data (files or chunks of files) on a single LUN (disk drive or drive array).

Lustre Clients

Lustre clients are computational nodes that use the Lustre file system to read and write data to files. Multiple clients can simultaneously read and write to different parts of the same file (distributed across multiple OSTs) maximizing the collective bandwidth of network and storage components.

Data Protection Layer (RAID)

ClusterStor E1000 systems use GridRAID to provide different data protection layers for the Lustre OSTs. GridRAID adds performance to the RAID protection by replacing the physical Hot Spare with the Distributed Spare. The Distributed Spare is a virtual device; its space is spread across all the drives in the array. The distributed virtual nature of GridRAID allows the repair process to write to multiple concurrent drives, reducing the amount of time that an array spends in a degraded state. The Lustre MGS and MDT both use RAID10 data protection with either double or triple mirroring, while other ancillary devices (for example, for OST external journals and GridRAID write-intent buffers) use RAID 10 or RAID1 data protection, also with double or triple mirroring.

Unified System Management Software (USM)

For systems using one (1) or more 4U106 EBODs, EBOD diagnostics are managed by Generalized Enclosure Management (GEM) through Unified System Management (USM). GEM is an advanced management software that controls and monitors the EBOD hardware infrastructure and environmental condition. It manages EBOD system health and provides power control of major subsystems. It monitors fans, thermals, power consumption, etc. and offers extensive event capture and logging mechanisms to enable post-failure analysis of hardware components.

Linux OS

ClusterStor E1000 systems run Lustre 2.12 software in a CentOS Linux 7.6 base environment.

Multiple MDT Feature

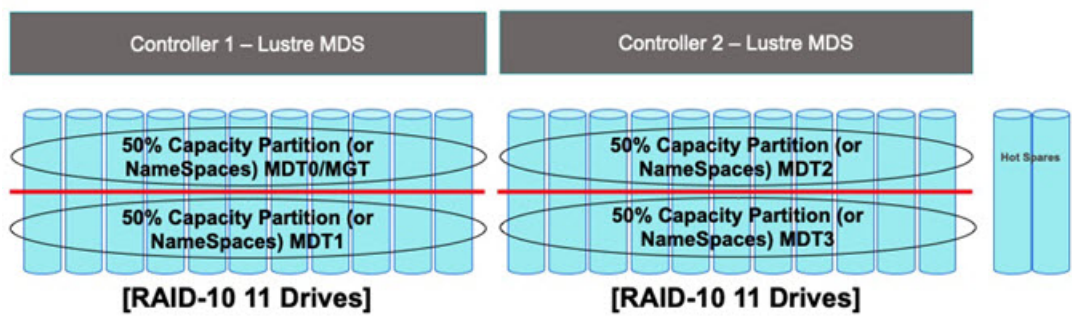
Multiple MDT Purpose:

- `ext4/ldiskfs` is limited to 4 billion inodes for a Lustre metadata target (MDT). Some users require as many as 150 billion inodes in their ClusterStor file systems. The standard MDT layout would require 19 MDUs to support that many inodes.
- The Multiple MDT (per MDS) feature doubles the number of MDTs possible from an MDU and the inode requirement can be met with 10 MDUs, reducing hardware cost.

Multiple MDT Layout

The typical MDU layout consists of two (2) MDT arrays—one (1) per MDS node in the MDU. The Multiple MDT layout partitions these sets of drives in half, each of which is a separate array and MDT. This doubles the MDT count for the MDU from two (2) to four (4).

Figure 3: Multiple MDT



Also see `configure_mds` Command and `configure_hosts` Command in the ClusterStor CSCLI Command Reference Guide S-9922.

Components and Hardware Supported by this Release

Cray ClusterStor E1000 provides a next-generation HPC storage platform that delivers industry-leading performance and durability. The Cray ClusterStor E1000 platform builds on a history of HPC excellence by offering substantial upgrades and enhancements to system components and hardware. The 4.x releases support the Cray ClusterStor E1000 platform.

Component Category	Description
Racks	Systems consist of one (1) or more racks installed with SMU, MDU, SSU and components, along with management and network switches. Valid types: 42U (alternate: 48U if using the Shasta Air-cooled APC rack)
Servers	System components include a number of different servers that host the Lustre and system nodes. Valid types: Embedded Server Node contained within 2U24 SMU, 2U24 MDU, and 2U24 SSU enclosures
Storage Hardware	The SMU, MDU, and SSU-F all use the same type of storage hardware (2U24 enclosures); the SSU-D uses a 4U106 EBOD enclosure. Valid types: 2U24 Disk Array (for the SMU, MDU, and SSU-F) and 4U106 (for the SSU-D)
Storage Media	Systems contain three (3) types of storage media: HDDs, SSDs, and NVMe. Valid types: HDD, SSD, NVMe
Power & Cooling	System components contain several types of power supplies and modules. Valid types: Power Supply Unit (PSU), Power Cooling Module (PCM)

For detailed descriptions of the LEDs and their functions, see the latest revision of the Cray ClusterStor E1000 Field Installation Guide at pubs.cray.com.



Functionality Supported in Each Cray ClusterStor Release

This section displays features and capabilities included in the Cray ClusterStor releases covered in this publication. Always check the **Release Notes** for qualified functionality on subsequent releases.

Definitions

Certified

The gold standard for features. Hewlett Packard Enterprise certifies that the feature is useful and is either integrated into the Cray ClusterStor management stack and/or can be used directly from a supported user client. In addition, comprehensive design review and testing have completed and all serious issues have been addressed. Users can confidently deploy certified features in any production environment.

Available

Features that do not meet the Hewlett Packard Enterprise gold standard but are otherwise considered useful and generally safe to use. Hewlett Packard Enterprise will support use of the feature, but it may not be tightly integrated into the Cray ClusterStor management stack, or its use may come with a small risk of serious defects.

Not Supported

Feature is not supported by Hewlett Packard Enterprise. Either the code is not present or it has been actively disabled because risk of serious defects is too great to consider use or support.

Support Category Definitions

Support Category	What It Means	Expected Level of Support
Certified	<ul style="list-style-type: none">Hewlett Packard Enterprise engineering has tested functionality (including at scale)Feature functions/performs reliably, even under adverse failure conditions and edge casesThe feature is usable through the Cray ClusterStor management stack and/or directly from a supported clientNo known blocker or critical issues either from Hewlett Packard Enterprise or the community	<ul style="list-style-type: none">L1/L2 support will triage any CAST issuesHewlett Packard Enterprise HPC R&D will provide L3 support
Available	<ul style="list-style-type: none">Hewlett Packard Enterprise has tested functionality and has found it to be safe to operate, including in production environmentsThe customer may need to take special care or appropriate mitigation steps for use in some circumstancesKnown serious issues and risks will be described in release notes	<ul style="list-style-type: none">L1/L2 support will triage any CAST issuesHewlett Packard Enterprise HPC R&D will provide L3 support
Not Supported	<p>The feature will not function on the Cray ClusterStor system for a variety of reasons:</p> <ul style="list-style-type: none">The feature is incompatible with Cray ClusterStorThe feature requires special integration that is not available in Cray ClusterStorThe feature has serious defects. Do not use the feature in production environments	<ul style="list-style-type: none">L1/L2 will not triage any CAST issues and close case as "will not fix" or reasonable categoryHewlett Packard Enterprise HPC R&D will not provide L3 support for these features

Qualified Functionality in Cray ClusterStor 4.4

Hardware

- 2U24 E1000-F NVMe servers
- 2U24 E1000-F PCIe Gen4 NVMe enclosures
- 4U106 E1000-D COBRA+ SAS JBODs
- Mellanox IB QM87xx HDR switches
- Mellanox IB SB77xx EDR switches
- Arista 100Gb Ethernet connectivity
- Aruba 6300M management switches
- Aruba 8325 10G management switches
- Broadcom 10Gb SFP+ MGMT NICs
- Mellanox 200Gb CX-6 HCAs
- Samsung 1733 NVMe SSDs
- Kioxia CM6 NVMe SSDs
- Titan3 5U84 JBOD
- Slingshot 10 direct attach network connectivity

Certified Features

- Support for the triple-way mirror RAID configuration
- Lustre 2.12
- LNet Health
- Multirail LNet
- LNet Dynamic Discovery
- Multirail routing
- Multiple MDT
- Lazy Size on MDT
- DMMP
- LockAhead
- Project Quotas
- Progressive File Layouts (PFLs)
- Subdirectory mounts
- Large Bulk IO
- T10-PI supported on HDD
- Multiple modify RPCs per client
- System Updater for providing basic rolling updates and support for rollback capability
 - Live updates are NOT supported from software version 4.1 directly to version 4.3 or a subsequent 4.x release.
 - Live updates are supported from software version 4.2, 4.3, or an earlier 4.4 release.
 - Nonlive updates from any prior 4.x release are supported.
 - The Lustre file system must be stopped to perform a rollback operation.

- Rollback to CSL 3.x is not supported.
- CSCLI – ClusterStor Command Line Interface
- ClusterStor Manager (CSM) browser-based GUI
- Kioxia CM6 SSD Support
- Viking BIOS to enable NPS=2 for increased performance
 - Also drives LNET CPT=8 configuration change
- Lustre Overstriping (limit: 256 stripes)
- EvansBP support for 10, 12, 14, 16TB capacity points
- SquashFS Root FS Support for diskless nodes

Available Features for Cray ClusterStor 4.4

- Self-Extending Layout (SEL) (aka Spillover Space)
- Data on Metadata (DoM)
- File Level Redundancy (FLR) Delayed Resync
- UID/GID Mapping
- Server Side Advice and Hinting
- DNE phase 2 (striped directories)
- DNE Phase 2b (cross-target rename/links and async updates)
- DNE 3: Migrating Striped Directory
- `lfsck` Phase 4
- SELinux preliminary client support
- Shared Key Crypto

Features NOT Supported in ClusterStor 4.3

- Kerberos updates
- T10-PI (client to HDD)

Qualified Functionality in Cray ClusterStor 4.3

Hardware

- 2U24 E1000-F NVMe servers
- 2U24 E1000-F PCIe Gen4 NVMe enclosures
- 4U106 E1000-D COBRA+ SAS JBODs
- Mellanox IB HDR switches
- Aruba management switches
- Broadcom 10Gb SFP+ MGMT NICs
- Mellanox 200Gb CX-6 HCAs
- Samsung 1733 NVMe SSDs
- Titan3 5U84 JBOD
- Arista 100Gb Ethernet connectivity
- Slingshot direct attach

Certified Features

- Support for the triple-way mirror RAID configuration
- Lustre 2.12
- LNet Health
- Multirail LNet
- LNet Dynamic Discovery
- Multiple MDT
- Lazy Size on MDT
- DMMP
- LockAhead
- Project Quotas
- Progressive File Layouts (PFLs)
- Subdirectory mounts
- Large Bulk I/O
- T10-PI supported on HDD
- Multiple modify RPCs per client
- System Updater for providing basic rolling updates and support for rollback capability
 - Live updates are NOT supported from software version 4.1 to version 4.3 or later.
 - Live updates are supported from software version 4.2 or an earlier 4.3 release.
 - Nonlive updates from any prior 4.x release are supported.
 - The Lustre file system must be stopped to perform a rollback operation.
 - Rollback to CSL 3.x is not supported.
- CSCLI – ClusterStor Command Line Interface
- ClusterStor Manager (CSM) browser-based GUI

Available Features for Cray ClusterStor 4.3

- Self-Extending Layout (SEL) (aka Spillover Space)
- Data on Metadata (DoM)
- Multirail routing
- File Level Redundancy (FLR) Delayed Resync
- UID/GID Mapping
- Server-Side Advice and Hinting
- DNE phase 2 (striped directories)
- DNE Phase 2b (cross-target rename/links and async updates)
- DNE 3: Migrating Striped Directory
- `lfsck` Phase 4
- SELinux preliminary client support
- Shared Key Crypto

Features NOT Supported in ClusterStor 4.3

- Kerberos updates

Qualified Functionality in Cray ClusterStor 4.2

Hardware

- 2U24 E1000-F NVMe servers
- 2U24 E1000-F PCIe Gen4 NVMe enclosures
- 4U106 E1000-D COBRA+ SAS JBODs
- Mellanox IB HDR switches
- Aruba management switches
- Broadcom 10Gb SFP+ MGMT NICs
- Mellanox 200Gb CX-6 HCAs
- Samsung 1733 NVMe SSDs
- Titan3 5U84 JBOD
- Arista 100Gb Ethernet connectivity
- Slingshot direct attach

Certified Features

- Support for the triple-way mirror RAID configuration
- Lustre 2.12
- LNet Health
- Multirail LNet
- LNet Dynamic Discovery
- Multiple MDT
- Lazy Size on MDT
- DMMP
- LockAhead
- Project Quotas
- Progressive File Layouts (PFLs)
- Subdirectory mounts
- Large Bulk I/O
- T10-PI supported on HDD
- Multiple modify RPCs per client
- System Updater for providing basic rolling updates and support for rollback capability
 - Immediate previous version only
 - Rollback to CSL 3.x is not supported
 - Lustre file system must be stopped to perform a rollback operation
 - Any data written to root partition on either management node after the update (and prior to rollback) must be saved somewhere, otherwise it will be lost
 - `/mnt/nfsdata` and `/mnt/mgmt` are saved on per-file basis; everything else uses LVM snapshots to backup data for rollback
 - `/boot` is not part of the rollback; previous version of kernel is automatically selected during the rollback operation
 - Find more System Updater content in the [Cray ClusterStor E1000 Software Release Installation Guide S-2598](#).

- CCLI – ClusterStor Command Line Interface
- ClusterStor Manager (CSM) browser-based GUI

Available Features for Cray ClusterStor 4.2

- Data on Metadata (DoM)
- Multirail routing
- File Level Redundancy (FLR) Delayed Resync
- UID/GID Mapping
- Server-Side Advice and Hinting
- DNE phase 2 (striped directories)
- DNE Phase 2b (cross-target rename/links and async updates)
- DNE 3: Migrating Striped Directory
- `lfsck` Phase 4

Features NOT Supported in ClusterStor 4.2

- Self-Extending Layout (SEL) (aka Spillover Space) (Lustre 2.13.0)
- Shared Key Crypto
- SELinux preliminary client support
- Kerberos updates

Qualified Functionality in Cray ClusterStor 4.1

Hardware

- 2U24 E1000-F NVMe servers
- 2U24 E1000-F PCIe Gen4 NVMe enclosures
- 4U106 E1000-D COBRA+ SAS JBODs
- Mellanox IB HDR switches
- Dell 10Gb SFP+ MGMT switches
- Broadcom 10Gb SFP+ MGMT NICs
- Mellanox 200Gb CX-6 HCAs
- Samsung 1733 NVMe SSDs

Certified Features

- Support for the triple-way mirror RAID configuration
- Lustre 2.12
- Multiple MDT
- Lazy Size on MDT
- DMMP
- LockAhead
- Project Quotas
- Progressive File Layouts (PFLs)
- Subdirectory mounts
- Large Bulk I/O
- `lfsck` Phase 4
- T10-PI supported on HDD
- Multiple modify RPCs per client
- System Updater for providing basic rolling updates and support for rollback capability
 - Immediate previous version only
 - Rollback to CSL 3.x is not supported
 - Lustre file system must be stopped to perform a rollback operation
 - Any data written to root partition on either management node after the update (and prior to rollback) must be saved somewhere, otherwise it will be lost
 - `/mnt/nfsdata` and `/mnt/mgmt` are saved on per-file basis; everything else uses LVM snapshots to backup data for rollback
 - `/boot` is not part of the rollback; previous version of kernel is automatically selected during the rollback operation
 - Find more System Updater content in the ClusterStor E1000 Software Release Installation Guide S-2598.
- CSCLI – ClusterStor Command Line Interface
- ClusterStor Manager (CSM) browser-based GUI

Available Features for ClusterStor 4.1

- LNet Health
- Multirail LNet

- LNet Dynamic Discovery
- Multirail routing
- File Level Redundancy (FLR) Delayed Resync
- UID/GID Mapping
- Server-Side Advice and Hinting
- Support for Rosetta ToR network switch
- DNE phase 2 (striped directories) (enabled for 4.1-030 and subsequent releases)
- DNE Phase 2b (cross-target rename/links and async updates) (enabled for 4.1-030 and subsequent releases)
- DNE 3: Migrating Striped Directory (enabled for 4.1-030 and subsequent releases)

Features NOT Supported in ClusterStor 4.1

- Data on Metadata (DoM) (disabled by default)
- DNE phase 2 (striped directories) (not enabled for 4.1-010 and 4.1-020)
- DNE Phase 2b (cross-target rename/links and async updates) (not enabled for 4.1-010 and 4.1-020)
- DNE 3: Migrating Striped Directory (not enabled for 4.1-010 and 4.1-020)
- Self-extending Layout (aka Spillover Space)
- Self-Extending Layout (SEL) (aka Spillover Space)
- Shared Key Crypto
- SELinux preliminary client support
- Kerberos updates

Software Versions and Requirements

This section provides information about the environment and software required for the Cray ClusterStor software release 4.4-010. Check the Release Notes for subsequent 4.x release version information.

CSM Version – release 4.4-010

ClusterStor Manager (CSM) 4.4 Build v.0.1.4.4-6.701261

SMU/MDU/SSU: stx_package-3.0-1.40.1.x86_64

Lustre Server (x86_64 Architecture) – release 4.4-010

Operating System

CentOS 7.6 Linux base

Kernel

3.10.0-957.1.3957.1.3.x4.4.22

File System

2.12.4.3_cray_104_g58e03c1-127.el7

Required User-Supplied Network Infrastructure – release 4.4-010

Manual workarounds may be available for environments without these servers. Contact Hewlett Packard Enterprise support for more information.

DHCP Server

Provides the MGMT node IP addresses for browser connections (user can choose to use a static IP address configuration for the “public” interfaces on the MGMT nodes)

NTP Server

Synchronizes clocks across the cluster nodes

DNS Server

Services DNS lookups for all nodes in the ClusterStor appliance

Power On ClusterStor E1000 System

Use this procedure to power on a ClusterStor system, using the CCLI interface. Nodes should always be powered up in the following order:

1. MGMT nodes
2. MGS/MDS nodes
3. OSS nodes

Nodes for E1000 systems:

- MGMT nodes
 - SMU – 2U24 NVMe chassis
- MGS/MDT nodes
 - MDU – 2U24 NVMe chassis
- OSS nodes
 - Flash SSU – 2U24 NVMe chassis
 - Flash SSU head node (2U24 NVMe chassis) with zero (0) or more 4U106 SAS enclosures attached

Also note that ClusterStor systems can contain multiple MDU enclosures.

Prerequisites

- **System access requirements:**
Root user access is required to perform this procedure on a ClusterStor system. If user does not have root access, contact Hewlett Packard Enterprise support.
- **Estimated time required**
(information forthcoming)
- **Service interruption level:**
This procedure requires taking the Lustre file system offline.
- **Tools and equipment:**
 - ESD strap/garments or other approved protection
 - Console with monitor and keyboard (or PC with serial port configured for 115.2 Kbps, 8 data bits, no parity, and one stop bit)

Procedure

1. At the back of the rack, confirm that power is off to the 2U24 NVMe enclosures. Each 2U24 has two (2) power supply units (PSUs).
2. Verify that the power cords are plugged in to the 4U106 EBOD enclosures. The 4U106 SAS EBOD has two (2) power cooling modules (PCMs).
3. Verify that the power is connected to the PDUs in the rack.
 - If necessary, plug the PDU cords into power receptacles.
 - If using a Raritan PDU, place each PDU power switch in the ON position.
Raritan PDUs have three (3) lines, with one (1) power switch for each line. Each line must be powered on. Some Raritan PDUs also have two (2) line inputs, and a new 60A PDU will only have one (1) line input.
 - ServerTech PDUs do not have power ON/OFF switches.
4. If the cluster contains 4U106 Expansion Storage Units (ESUs), power on the 4U106 enclosures as follows:
 - a. Verify that the power cords are plugged in to the power supplies.
 - b. Wait 30 seconds for the drives to spin up after PDU power is applied.
5. On ClusterStor E1000 systems: Verify that the 2U24 MGMT enclosures are plugged in.

Make certain the primary and secondary MGMT nodes are connected to the public network of the user.

6. Both management servers are normally configured to automatically power on. If this is not the case for the user system, power on manually using the pinhole power switch on the rear of the controllers. Check the power indicator LED to verify the servers are on.

Figure 4: Rear View with FRU Locations



Figure 5: IOM Canister

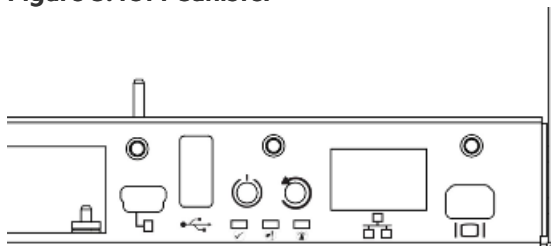
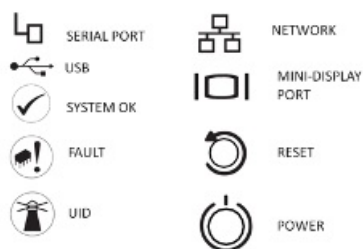


Figure 6: Canister Icon Key



7. Optionally, the user may connect a keyboard and monitor to the primary MGMT node and log in to the console using the admin account credentials. On E1000 systems, users may choose between connecting a monitor (using mini-HDMI cable) and keyboard (USB) and serial console or logging in over the public LAN. The systems may take 10 minutes to become available.
8. Check that the shared storage targets are available for the management nodes:

```
admin@n000$ pdsh -g mgmt cat /proc/mdstat
```

For example:

```
admin@cls12345n000$ pdsh -g mgmt cat /proc/mdstat | dshbak -c
-----
cls12345n000
-----
Personalities : [raid1] [raid6] [raid5] [raid4] [raid10]
md64 : active raid10 sda[0] sdc[3] sdw[2] sdl[1]
1152343680 blocks super 1.2 64K chunks 2 near-copies [4/4] [UUUU]
bitmap: 2/9 pages [8KB], 65536KB chunk
md127 : active raid1 sdy[0] sdz[1]
439548848 blocks super 1.0 [2/2] [UU]
unused devices: <none>
-----
cls12345n001
-----
Personalities : [raid1] [raid6] [raid5] [raid4] [raid10]
md67 : active raid1 sdi[0] sdt[1]
576171875 blocks super 1.2 [2/2] [UU]
bitmap: 0/5 pages [0KB], 65536KB chunk
md127 : active raid1 sdy[0] sdz[1]
439548848 blocks super 1.0 [2/2] [UU]
unused devices: <none>
```

9. Check HA status once the node is completely up and HA configuration has been established:

```
admin@n000$ sudo crm_mon -lr
```

Correct output indicates that all resources have started and are balanced between two nodes. If not, use one of the following, and recheck `crm_mon` output after five (5) minutes.

```
Last updated: Thu Aug 7 01:30:36 2014
Last change: Wed Aug 6 23:58:18 2014 via crm_resource on cls12345n001
Stack: Heartbeat
Current DC: cls12345n001 (0828104e-8d91-44ad-892a-13dbd1fd7c6c) - partition
with quorum
Version: 1.1.6.1-6.el6-0c7312c689715e096b716419e2ebc12b57962052
2 Nodes configured, unknown expected votes
53 Resources configured.
=====
Online: [ cls12345n000 cls12345n001 ]
Full list of resources:
cls12345n000-1-ipmi-stonith (stonith:external/ipmi): Started cls12345n000
cls12345n001-1-ipmi-stonith (stonith:external/ipmi): Started cls12345n001
cls12345n000-2-ipmi-stonith (stonith:external/ipmi): Started cls12345n000
cls12345n001-2-ipmi-stonith (stonith:external/ipmi): Started cls12345n001
Clone Set: cln-diskmonitor [diskmonitor]
Started: [ cls12345n000 cls12345n001 ]
Clone Set: cln-last-stonith [last-stonith]
Started: [ cls12345n000 cls12345n001 ]
Clone Set: cln-kdump-stonith [kdump-stonith]
Started: [ cls12345n000 cls12345n001 ]
prn-httpd (lsb:httpd): Started cls12345n000
prn-mysql (lsb:mysql): Started cls12345n000
prn-nfslock (lsb:nfslock): Started cls12345n001
prn-bebundd (lsb:bebundd): Started cls12345n000
Clone Set: cln-cerebrod [prn-cerebrod]
Started: [ cls12345n001 cls12345n000 ]
prn-conman (lsb:conman): Started cls12345n000
prn-dhcpd (lsb:dhcpd): Started cls12345n001
Clone Set: cln-syslogng [prn-syslogng]
Started: [ cls12345n001 cls12345n000 ]
Clone Set: cln-dnsmasq [prn-dnsmasq]
Started: [ cls12345n001 cls12345n000 ]
prn-nodes-monitor (lsb:nodes-monitor): Started cls12345n000
Clone Set: cln-ses_mon [prn-ses_monitor]
Started: [ cls12345n001 cls12345n000 ]
Clone Set: cln-nsca_passive_checks [prn-nsca_passive_checks]
Started: [ cls12345n001 cls12345n000 ]
Resource Group: grp-icinga
prn-icinga (lsb:icinga): Started cls12345n000
prn-nsca (lsb:nsca): Started cls12345n000
prn-npcd (lsb:npcd): Started cls12345n000
(output continues here...)
```

10. Power on the MGS and MDS nodes:

```
admin@n000$ cscli power_manage -n mgs_node,mds_node --power-on
```

Example:

```
root@cls12345n000# cscli power_manage -n cls12345n[002-003] --power-on
power_manage: processing cls12345n002 ...
power_manage: processing cls12345n003 ...
power_manage: Operation performed successfully.
```

Depending upon the configuration, there could be multiple MDU nodes to power on.

11. Power on the OSS nodes:

```
root@cls12345n000# cscli power_manage -n oss_node_names --power-on
```

Example:



```
root@cls12345n000# cscli power_manage -n cls12345n[004-007] --power-on
power_manage: processing cls12345n004 ...
power_manage: processing cls12345n005 ...
power_manage: processing cls12345n006 ...
power_manage: processing cls12345n007 ...
power_manage: Operation performed successfully.
```

Hewlett Packard Enterprise advises having no more than 60 nodes powering on at the same time. If there are more than 60 nodes, repeat this step as necessary (powering on 60 nodes each time) until all nodes are powered on.

12. Check the status of the nodes:

```
admin@n000$ pdsh -a date
```

The correct output includes the date for each host in the cluster.

Example:

```
admin@cls12345n000$ pdsh -a date
cls12345n000: Thu Aug 7 01:29:28 PDT 2014
cls12345n003: Thu Aug 7 01:29:28 PDT 2014
cls12345n002: Thu Aug 7 01:29:28 PDT 2014
cls12345n001: Thu Aug 7 01:29:28 PDT 2014
cls12345n007: Thu Aug 7 01:29:28 PDT 2014
cls12345n006: Thu Aug 7 01:29:28 PDT 2014
cls12345n004: Thu Aug 7 01:29:28 PDT 2014
cls12345n005: Thu Aug 7 01:29:28 PDT 2014
```

This completes the power on procedure. Mount the file system(s) provided by this cluster as required.

Power Off ClusterStor System

Use this procedure to power off a Cray ClusterStor E1000 system using the CCLI interface.

- IMPORTANT:** Unmount the Lustre file system from all clients before starting the power off procedure when powering down both storage and compute. Failure to do so may cause clients to stop responding.

If clients are not unmounted first, then they could have changes in-flight at the moment of power-off. Unmounting the server targets while clients are connected will force the clients to retransmit any in-flight work when the servers are powered back. If the clients are powered off in the meantime, then those in-flight changes will be lost. Therefore, it is a good practice to unmount clients when powering down both storage and compute.

When powering down ClusterStor servers, Hewlett Packard Enterprise recommends unmounting clients first. Otherwise, clients and servers will go through recovery procedures once the servers come back online. Lustre recovery can cope with failures during recovery, but any further failures may put data at-risk. Regardless of whether full recovery is possible, Lustre always makes the file system coherent. However, expected files and/or writes could be missing if not all clients can complete replay. See:

https://doc.lustre.org/lustre_manual.xhtml#lustre recovery for recovery details.

It may be desirable in some cases to leave clients connected and mounted during server shutdown (for example, if the servers are planned to restart soon or if it would be difficult to stop jobs, marshal administration to unmount, etc.). In this case, clients simply wait for the servers to resume.

Prerequisites

- System access requirements:**
Root user access is required to perform this procedure on a ClusterStor system. If user does not have root access, contact Hewlett Packard Enterprise support.
- Estimated time required**
(information forthcoming)
- Service interruption level:**
This procedure requires taking the Lustre file system offline.
- Tools and equipment:**
 - ESD strap/garments or other approved protection
 - Console with monitor and keyboard (or PC with serial port configured for 115.2 Kbps, eight data bits, no parity, and one stop bit)

Procedure

1. Log in to the primary MGMT node with SSH:

```
client$ ssh -l admin primary_MGMT_node
```

2. Change to root user:

```
admin@n000$ sudo su -
```

3. Stop the Lustre file system:

```
root@n000# cscli unmount -f filesystem_name
```

4. Verify that resources have been stopped by running the following on all even-numbered nodes:

```
root@n000# ssh nodename crm_mon -r1 | grep fsys
```

Sample output showing the nodes are stopped:

```
MGMT0# ssh cls12345n006 crm_mon -r1 | grep fsys
cls12345n006_md0-fsys (ocf::heartbeat:XMNTR): Stopped
cls12345n006_md1-fsys (ocf::heartbeat:XMNTR): Stopped
cls12345n006_md2-fsys (ocf::heartbeat:XMNTR): Stopped
cls12345n006_md3-fsys (ocf::heartbeat:XMNTR): Stopped
cls12345n006_md4-fsys (ocf::heartbeat:XMNTR): Stopped
cls12345n006_md5-fsys (ocf::heartbeat:XMNTR): Stopped
cls12345n006_md6-fsys (ocf::heartbeat:XMNTR): Stopped
cls12345n006_md7-fsys (ocf::heartbeat:XMNTR): Stopped
```

5. Log in to the MGS node with SSH:

```
root@n000# ssh MGS_node
```

6. To check the MGS/MDS nodes to determine whether Resource Group **md65-group** is stopped, use the `crm_mon` utility to monitor the status of the MGS and MDS nodes and to check that their resources have stopped:

```
MGS# crm_mon -lr | grep fsys
```

Example of `crm_mon` output showing the MGS and MDS nodes in a partial stopped state:

```
MGS# crm_mon -lr
cls12345n003_md66-fsys (ocf::heartbeat:XYMNTR): Stopped
cls12345n003_md65-fsys (ocf::heartbeat:XYMNTR): Started
```

If the node is not stopped, issue the `stop_xyraid` command:

```
MGS# stop_xyraid nodename_md65-group
```

Sample `crm_mon` output showing the MGS and MDS nodes in a stopped state:

```
MGS# crm_mon -lr
cls12345n003_md66-fsys (ocf::heartbeat:XYMNTR): Stopped
cls12345n003_md65-fsys (ocf::heartbeat:XYMNTR): Stopped
```

7. Power off the diskless nodes:

```
root@n000# cscli power_manage -n csnodes[XX-YY] --power-off
```

where `csnodes[XX-YY]` is a `pdsh` style for node hostnames (for example, `cls12345n[002-011]`).

8. Check the power-off status of the diskless nodes:

```
root@n000# pm -q
```

Repeat this step until all non-MGMT nodes have been powered down.

Example:

```
root@cls12345n000# pm -q
on: cls12345n[000-001]
off: cls12345n[002-011]
unknown:
```

9. From the primary MGMT node, power off the secondary MGMT node:

```
root@n000# cscli power_manage -n csnodes01 --power-off
```

10. From the primary MGMT node, power off the active MGMT node:

```
root@n000# cscli power_manage -n csnodes00 --power-off --force
```

11. Once all the nodes are shut down, users can physically power off the enclosures by turning off the PDU receptacles.
12. Power off the cabinet circuit breaker.

Spectre/Meltdown Mitigation on Cray ClusterStor Systems

For Spectre/Meltdown security vulnerabilities, a CCLI command can be used to configure Spectre/Meltdown mitigation on the Cray ClusterStor system.

In ClusterStor 3.x and 4.x, Spectre/Meltdown configuration changes require rebooting the nodes for the changes to take effect.

By default, the system is configured with the Spectre/Meltdown mitigation mechanisms enabled on the management nodes, disabled on the diskless nodes.

Syntax:

```
admin@cls12345n000$ cscli security spectre
```

Usage:

```
cscli security spectre [-h] (--enable | --disable | --auto | --show)
```

Example Output of `--show` Option with Default Configuration in Place

```
# cscli security spectre --show
Meltdown/Spectre mitigation = auto
```

Optional arguments:

- Run `cscli security spectre -h`, `--help` to show help message and exit
- Run `cscli security spectre --enable` to enable the Spectre/Meltdown patch on all nodes
- Run `cscli security spectre --disable` to disable the Spectre/Meltdown patch on all nodes
- Run `cscli security spectre --auto` to enable Meltdown and Spectre mitigation on MGMT nodes and disable on OSS and MDS nodes
- Run `cscli security spectre --show` to display Meltdown and Spectre configuration



NOTE: Enabling the Spectre/Meltdown patch on the diskless nodes will adversely impact performance of the file system.

Log In to CSM

IMPORTANT:

Use either of the following procedures to begin using the ClusterStor Manager (CSM), the primary administrative interface for the Cray ClusterStor system.

Use CSCLI to Log In to CSM

To use the CSCLI interface, SSH into the active MGMT node as the admin user:

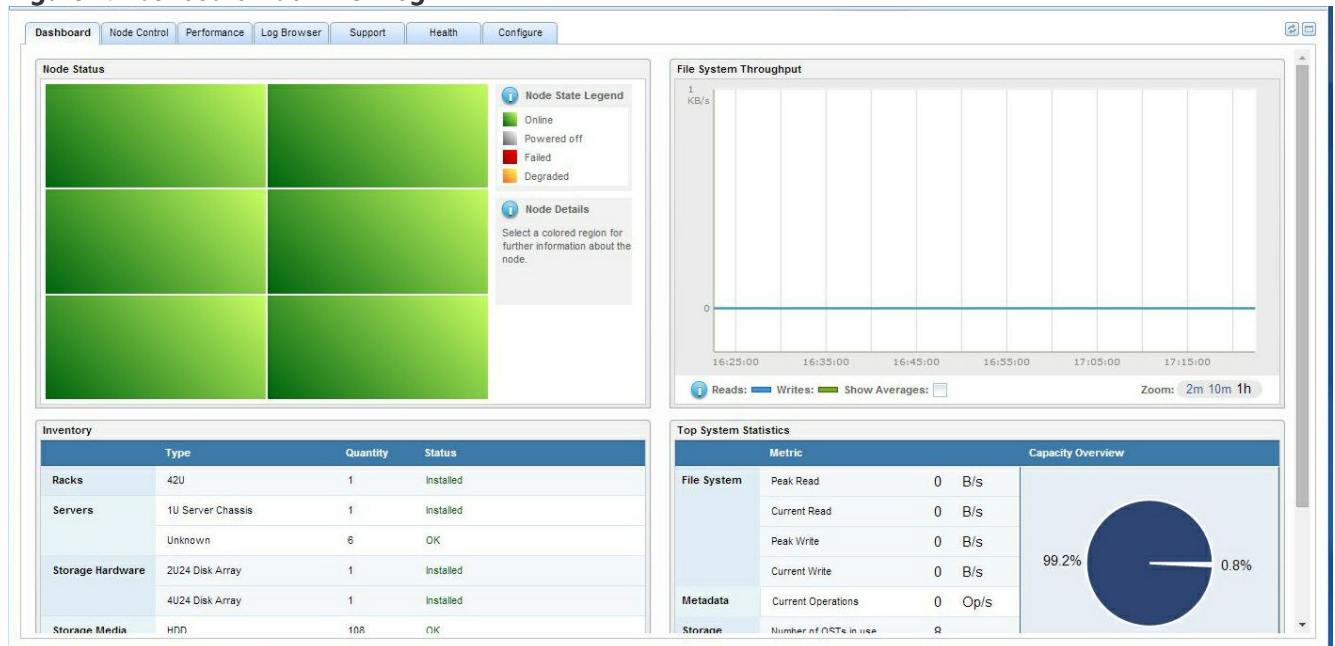
```
client$ ssh -l admin MGMT_node
```

Use the GUI to Log In to CSM

Access the CSM GUI using a direct connection between the site IP network and the Cray ClusterStor MGMT nodes. Log in to the GUI using a browser on a PC or on a terminal that is connected to the site IP network. CSM supports Chrome, Firefox, Microsoft Edge, or Internet Explorer browsers.

1. Open a browser window.
2. Enter the CSM IP address in the browser URL field, and press **Enter**.
3. Enter the administrator Username and Password used to access CSM at the site, and press **Enter**.
A What's New window will appear. Click **Close** to see the **Dashboard** tab in the browser window.

Figure 7: Dashboard Tab After Login



The Dashboard tab displays individual widgets to view status and performance data. From the Dashboard, users can survey high-level system health and resolve issues with individual nodes and take further action by quickly linking to the Node Control tab.

RAS – Reliability, Availability, and Serviceability

This section provides an overview of the RAS system to monitor and maintain Cray ClusterStor E1000 systems. There is no guided repair for the NVMe and 4U106 enclosures. If a supported enclosure is attached to a Cray ClusterStor E1000, that will have guided repair.

Reliability, Availability, and Serviceability (RAS) is a set of attributes that reflect the robustness of a hardware and/or software system.

- Reliability is a function of time that expresses the probability at a given future time that a system will still be working, given that it is working at the present time
- Availability is the measure of how often the system is available for use (such as a system's up-time percentage). Availability and reliability may sound like the same concept, but they are different, as a system can have great availability but offer no reliability
- Serviceability is a broad definition that describes the ease of system service or repair

Reliability and Availability (RA)

Reliability and availability features keep a system working when failures occur. Examples of RA:

- Hardware redundancy (power, servers, networks)
- Storage stacks (SCSI, RAID, local file system)
- HA stack
- Clustered file system

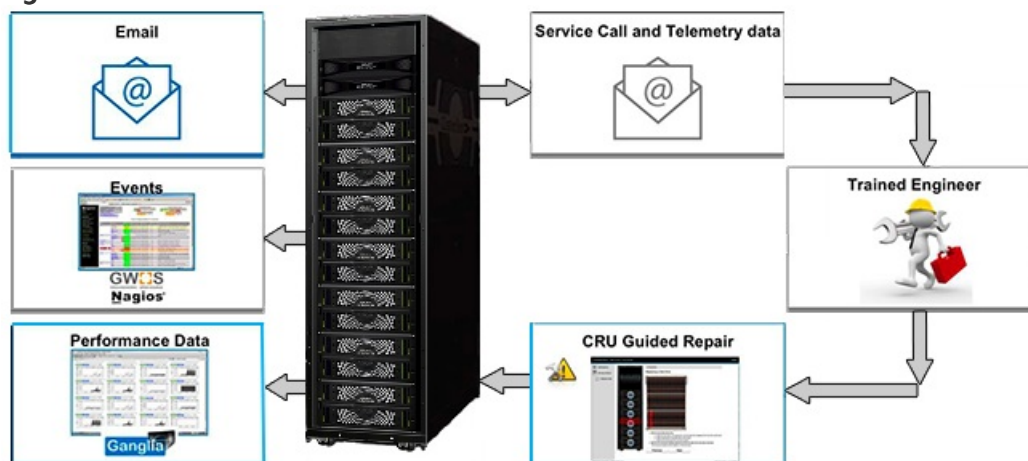
Serviceability

Serviceability consists of on-product capabilities and off-product tools, processes, and staff. Serviceability supports the needs of the product (its RA model, for example); HPC and archive products will behave differently. There are cost constraints associated with serviceability. Most important, the customer should not be impacted; there should be no loss of availability for maintenance operations. Serviceability focuses on Level 1 service personnel and eventually, end users.

In E1000 products, serviceability includes the following capabilities:

- Inventory discovery (part/serial numbers, firmware versions, location)
- Monitoring, diagnostics, fault isolation of
 - data collection from functional codes (RAID, HA, FS)
 - discovery and data collection from hardware (SES, IPMI, SNMP)
- Reporting service manager policy engine (product-specific)
- Reporting service notifications (user email, SNMP)
- Reporting live event telemetry stream (interesting events)
- Supports disks, PSUs/PCMs, and ESM controller server modules

Figure 8: RAS Service Notification Scenario



RAS for E1000 provides an "expert" system:

- Architected specifically for E1000
- With topological awareness that enables correct issue prioritization

- Designed to reduce false positives and repetitive alerts
- With event-driven data end-to-end across the system, allowing real-time updates

The E1000 RAS service console provides a real-time GUI.

Remote support includes email notifications, user alert emails, and telemetry data. RAS-based SUs are designed to be applied with no downtime/interruption.

Configure Service Console Access

The Cray ClusterStor GUI-based service console interface is a useful tool in the ClusterStor RAS solution. The interface shows system issues and conditions detected by the RAS solution. The issues can be reported to users and to Hewlett Packard Enterprise Support through email notifications, listing system repair actions to be completed by service personnel.

By default, ClusterStor systems ship with service console login capabilities enabled. To maintain a more secure environment, system administrators may opt to restrict access to the service console and to enable logins to the interface only when a service operation must be performed.

Disabling the service console only restricts logins to the GUI; no other aspect of the ClusterStor RAS system is affected. A change (enable/disable action) to the service console login takes approximately 30 seconds to be completed.

Prerequisites

Software version 3.0.0 or higher is installed.

Procedure

1. Enable login to the service console:

```
MGMT0$ cscli service_console enable
```

2. To disable login to the service console:

```
MGMT0$ cscli service_console disable
```

3. Show the status of the service console (enabled or disabled):

```
MGMT0$ cscli service_console show
```

Configure SMTP Relay, RAS Email, Remote Support, and Service Console

This section provides procedures to configure SMTP and email notifications, and to enable/disable the service console. Console-related service configuration changes also can be performed with the CCLI command-line interface.

In this process, the system administrator configures an external SMTP relay so email notifications pertaining to system events and associated system messages can be issued.

Email notifications are sent to users with email addresses configured by the system administrator. This section explains how to configure a list of users to be sent notifications and how to enable/disable an individual user notifications.

IMPORTANT: Email notifications to the specified user list are sent only if the notification functionality is enabled.

The service console login can be enabled (configured) to access guided replacement service procedures, such as replacing a disk. Systems are shipped with service console capabilities enabled, by default.

Prerequisites

This procedure assumes that RAS has been installed on the system.

Procedure

Configure SMTP

1. Show the current SMTP configuration:

```
MGMT0$ cscli service_console configure smtp show
SMTP relay: mailrelayus.us.company.com:25
```

The preceding example shows a hypothetical configuration.

2. Configure the SMTP relay:

```
MGMT0$ cscli service_console configure smtp relay \
--host=smtp_host --port=smtp_port
```

where *smtp_host* (required field) is the SMTP relay host for email notifications, and *smtp_port* is the SMTP relay port (default: 25).

Use the `--help` flag to obtain additional information for configuring the SMTP relay. For example:

```
MGMT0$ cscli service_console configure smtp relay --help
```

Configure the Email Notification List

This section explains how to configure a list of users to be sent notifications and how to enable/disable individual user notifications.

IMPORTANT: Email notifications to the specified user list are sent only if the notification functionality is enabled.

3. Display a list of users that are already in the email notification list:

```
MGMT0$ cscli service_console configure email show
```

4. Add an email address:

```
MGMT0$ cscli service_console configure email add -A user@domain.com
```

Delete an existing email address:

```
MGMT0$ cscli service_console configure email delete -A user@domain.com
```

IMPORTANT: Email notifications to the specified user list are sent only if the notification functionality is enabled.

5. Verify that the email address has been correctly established:

```
MGMT0$ cscli service_console configure email show
```

6. Enable user email support:

```
MGMT0$ cscli service_console configure email enable
```

(To disable user email support):

```
MGMT0$ cscli service_console configure email disable
```

7. Send a test email to verify that the email address has been properly added:

```
MGMT0$ cscli service_console configure email send_test_email
```

If the SMTP relay is configured properly, the configuration emails are enabled, and the specified email addresses have been set up to receive notifications, the following test email message is sent to the user list:

```
Subject: Remote Service TEST: cls12345
Body:
System Identification:
    System Serial Number: cls12345
    System Identifier:
Current System Time: Tue, 21 Oct 2017 18:05:54
System Timezone: PDT
TEST: This email is to test system SMTP relay setup only, please
ignore.
```

Example of an email successfully sent to remote support:

```
Successfully sent test email to remote support with subject
'Remote Service TEST: cls12345'
```

Example of an email successfully sent to users whose email addresses are properly configured:

```
Successfully sent test email to configured user(s) with subject
'Remote Service TEST: cls12345'
```

Enable or Disable the Service Console

8. To enable the service console:

```
MGMT0$ cscli service_console enable
```

ClusterStor systems are shipped with service console capabilities enabled, by default.

To disable the service console:

```
MGMT0$ cscli service_console disable
```

Configure RAS System Settings

This section specifies the ClusterStor installation name (identifier) for the RAS system. Use the `system` command to configure this value or show the currently specified identifier.

Prerequisites

RAS is installed.

Procedure

1. Show the current system identifier:

```
MGMT$ cscli service_console configure system show
```

2. Set the system identifier:

```
MGMT$ cscli service_console configure system identifier --name cls12345
```

where name (a required field) is the name of the ClusterStor system.

Configure the REST API

This section describes how to enable or disable RAS data streaming via the REST API, with the option to show current status.

Prerequisites

RAS is installed in the system.

Procedure

1. Enable REST API access:

```
MGMT$ cscli service_console configure rest_api enable
```

2. Disable REST API access:

```
MGMT$ cscli service_console configure rest_api disable
```

3. Show the status of data streaming via the REST API (enabled or disabled):

```
MGMT$ cscli service_console configure rest_api show
```

Interpret RAS Email Notifications

This section describes the types of RAS email notifications sent to users. The service console can be configured to send email notifications to specific users about service event messages (SEMs), which report events related to the service console's operation and include detail and location information about failed components. SEM notifications are sent to a pre-configured user email list specified by the system administrator.

SEMs include details and location information of degraded or failed components requiring repair. The notification includes a human-readable message, with instructions for accessing the service console, and a JSON-formatted machine-readable attachment.

SEM contents include the following:

- System serial number and identification information
- Date, timestamp and time zone of the service event
- Human-readable description of the failed component
 - Serial number, model, manufacturer, and firmware level (if applicable)
 - Location information (rack and enclosure containing the faulty part)
- Machine-readable attachment

Files attached to SEMs are used to provide automated FRU replacement part number lookups.

The following sample SEM (message body and JSON-formatted attachment) lists service events for a failed disk and power supply unit, along with the service console's IP address for service personnel to log in and repair the failed hardware.

SEM Body

Subject:

```
Remote Service SEM: SNX0G1GXC
```

Body:

```
This is a Sonexion service event notification:

The service console can be accessed from the service port via this URL: http://172.16.2.1:8080

System Identification:
  System Serial Number: SNX0G1GXC
  System Identifier:

Current System Time: Thu, 27 Nov 2014 18:46:54 System Timezone: EST

Current outstanding service call events:

Item 1 : "Disk drive needs replacement"

Service Code: 002005001
Time the event was first detected: Thu, 27 Nov 2014 09:31:28 EST

Details of failed component

  Disk Serial #: SOM122HN0000B40298QD
  Disk Model : ST600MM0006
  Drive Manufacturer: SEAGATE
  Firmware version of drive at time of failure: XLGD
  T10 Enabled: "Yes"

Location of failed component

  Rack Name: Rack1
  Enclosure Model: 2U24
  Enclosure Location: 36U
  Disk located in slot: 15

Item 2 : "Power supply issue detected"
```

Service Code: 002005003

Time the event was first detected: Thu, 27 Nov 2014 09:49:27 EST

Details of failed component

Power supply chassis type: OneStor Power One 580W_AC_PCM, P/N 82562.

Power supply part number: 0082562-07

Power supply product version: 0311

Power supply serial number: PMW8256200052C7

Location of failed component

Rack Name: Rack1

Enclosure Model: 2U24

Enclosure Location: 36U

Power supply located in bay: 0

SEM Attachment

Attachment:

sem-SNX0G1GXC-2014-11-27_18-46-54_074697.json

```
{
  "system_serial_number":
    "SNX0G1GXC",
  "system_identifier": "",
  "system_time": "Thu, 27 Nov 2014 18:46:54",
  "utc_timestamp": 1417132014,
  "system_timezone": "EST", "local_time": "Thu, 27 Nov 2014 18:46:54 EST",
  "service_console_url": "http://172.16.2.1:8080",
  "version": 4,
  "messages": [
    {
      "event_description": "Disk drive needs replacement",
      "re_event_code": "001001001",
      "uuid": "54ac138e-7642-11e4-8045-000c293bad8b",
      "confirmed_time": 1417098798.773304,
      "creation_time": 1417098798.844036,
      "dcs_timestamp": "1417098688",
      "completion_time": 0.0,
      "state": "SVC_CREATED",
      "location": {
        "index": "15",
        "enclosure_location": "36U",
        "enclosure_model": "2U24",
        "rack": "Rack1"
      },
      "event_code": "002005001",
      "disk": {
        "dm_report_t10": "11110111100",
        "firmware": "XLGD",
        "t10_enabled": true,
        "serial_number": "S0M122HN0000B40298QD",
        "model": "ST600MM0006",
        "manufacturer": "SEAGATE"
      },
      "type": "disk"
    },
    {
      "event_description": "Power supply issue detected",
      "re_event_code": "001003001",
      "uuid": "93c895bc-7677-11e4-ab16-000c293bad8b",
```



```
"confirmed_time": 1417121643.868457,  
"creation_time": 1417121668.054007,  
"dcs_timestamp": "1417099767", "completion_time": 0.0,  
"state": "SVC_CREATED",  
"location": {  
  "index": "0",  
  "enclosure_location": "36U",  
  "enclosure_model": "2U24",  
  "rack": "Rack1"  
},  
"event_code": "002005003",  
"type": "power_supply",  
"power_supply": {  
  "status": "failure_detected",  
  "serial_number": "PMW8256200052C7",  
  "part_number": "0082562-07",  
  "product_version": "0311",  
  "chassis_type": "OneStor Power One 580W_AC_PCM, P/N 82562."  
}  
}  
]  
}
```

Use Guided Replacement Procedures with the Service Console

Currently, there is no guided repair for the NVMe and 4U106 enclosures. If a supported enclosure is attached to an E1000, that will have guided repair.

This section provides steps to access the service console and follow the guided replacement procedures.

Prerequisites

RAS is installed and enabled.

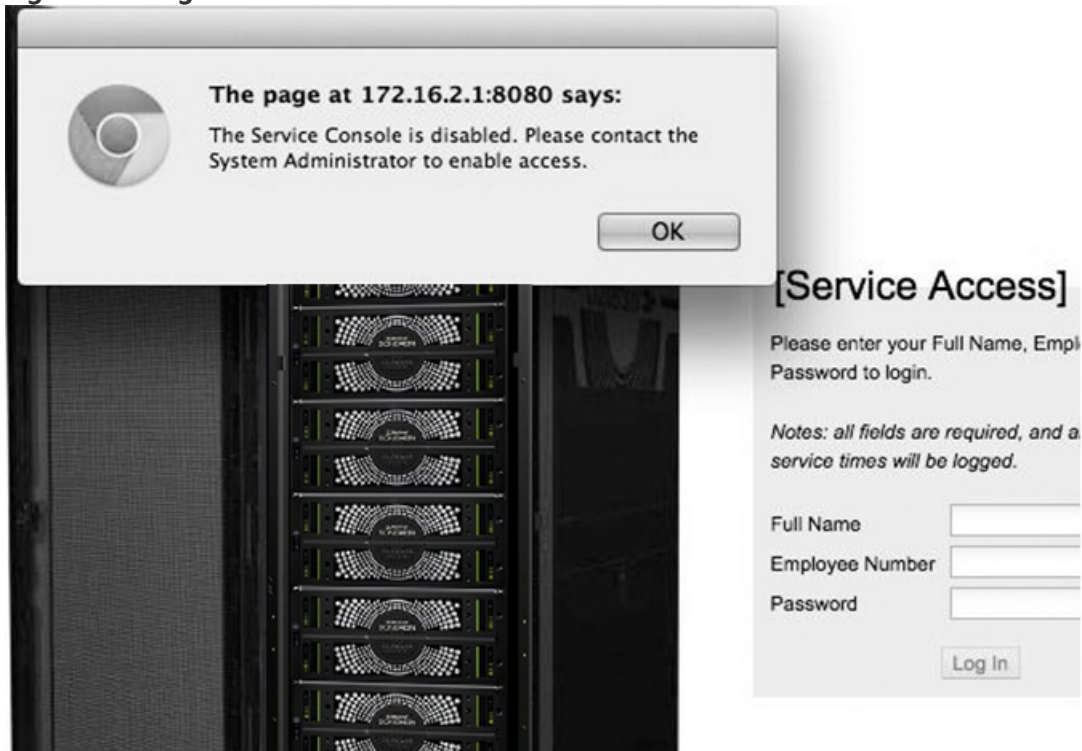
Procedure

Connect to the Service Console

1. Cable a laptop to any available port on either management switch.
2. Navigate to `https://<MGMT0 IP address>/shells/service/site/` using a web browser:
3. Enter authentication credentials provided by the customer service center.

i IMPORTANT: If when attempting to log into the service console, a message similar to the dialog box shown appears, then the system administrator has disabled access to the interface and service actions cannot proceed. If the system administrator re-enables the service console access, the Log In button becomes active within 30 seconds and logging in can continue. For more information, refer to the section, Configure Service Console Access.

Figure 9: Dialog Box Indicates Disabled Service Console



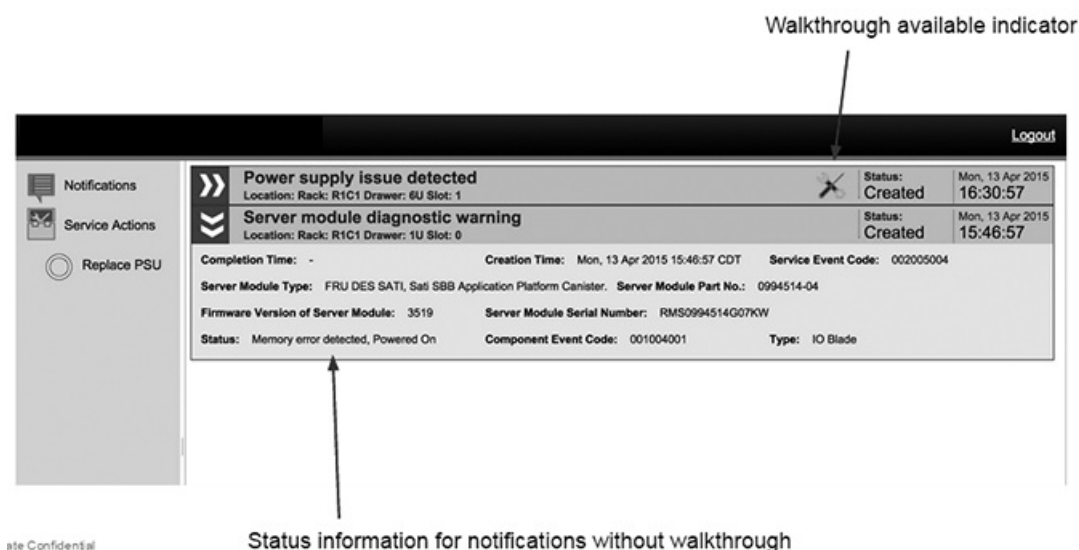
When the service console is logged into, it displays a list of repair actions (typically failed hardware replacements). The service console provides a step-by-step guide through the repairs to successfully return the system to normal operating mode.

Guided Replacement Procedures

4. Read the list of repair actions.

Diagnostic notifications and service actions are displayed. Serviceable items that can be addressed using a guided replacement procedure (a “walkthrough”), are identified with an “available walkthrough” indicator and appear in a pick list on the left side of the service console, as shown.

Figure 10: Service Console: Notification View



- To expand the view and access additional information about the service event, click on a specific notification .

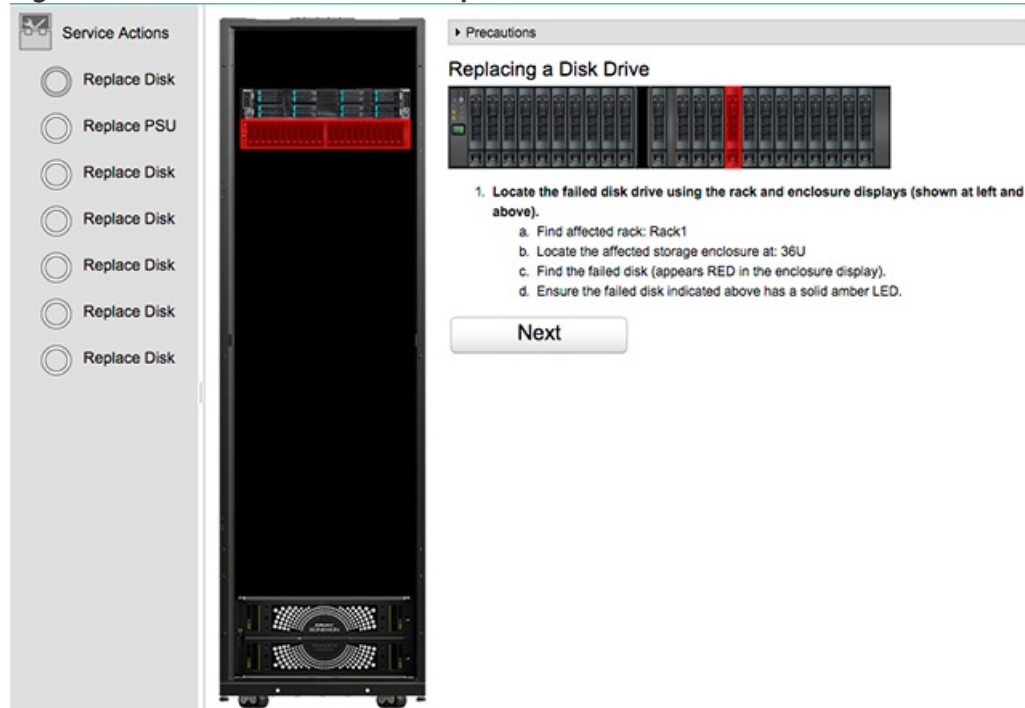
Selecting a service action from the pick list launches the guided replacement procedure for the specified failed hardware component. Walkthrough steps are displayed to the right of the pick list and feature:

- Visualization of the system
- Clear identification of the affected chassis
- Clear identification of the affected FRU
- Clear and concise instructions on how to perform the replacement procedure
- Live feedback during a procedure

Each service action provides visual cues (in red) to help the technician identify the affected enclosure and the specific location of the degraded or failed component. The following figures illustrate the guided replacement procedure to replace a failed DDIC in a 2U24 enclosure, with red indicators showing the affected enclosure in the rack and the location of the failed disk in the enclosure.

- If the guided replacement includes precautions or warnings related to the repair, the technician is prompted to acknowledge them before starting the procedure.

Figure 11: Service Console: Guided Replacement Procedure for a Disk Drive in a 2U24 Enclosure



- The system validates the new hardware.

Figure 12: Service Console: Guided Replacement Procedure to Replace a Disk Drive

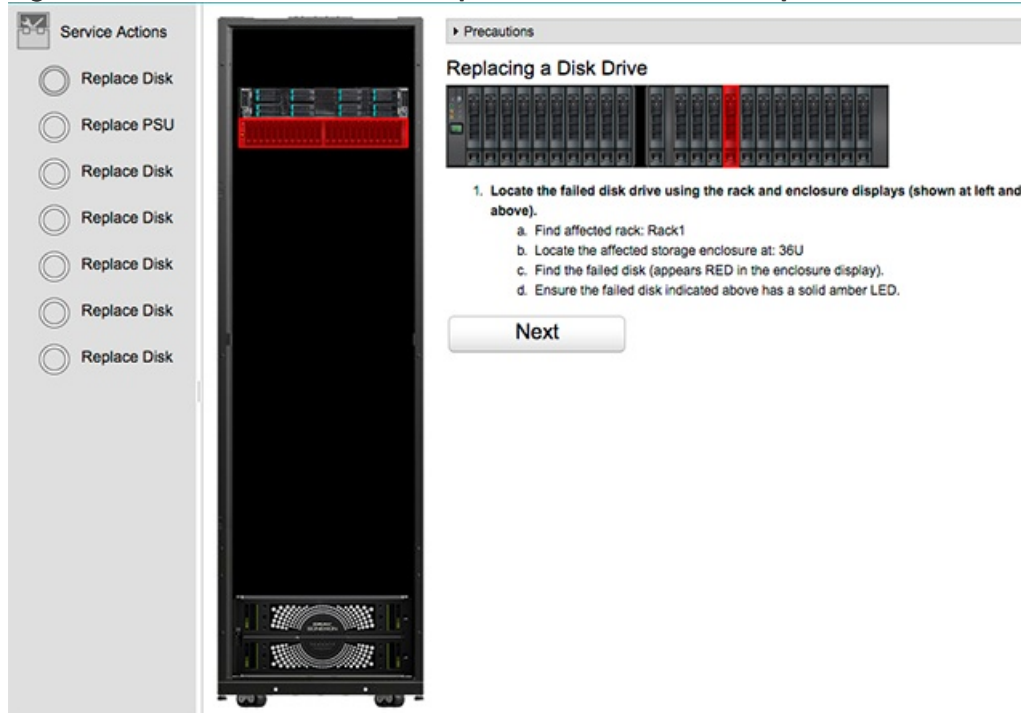
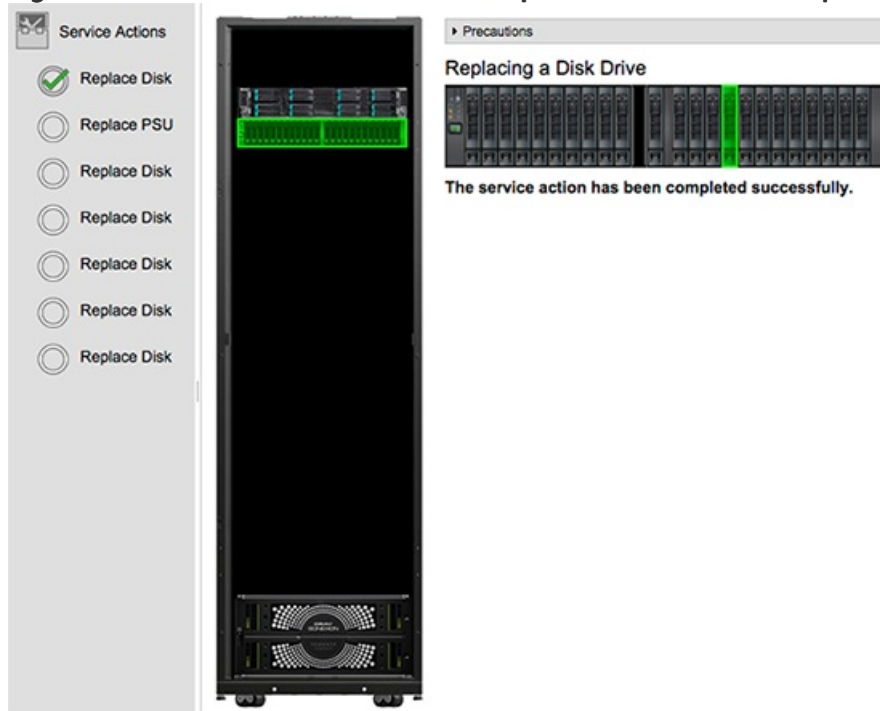


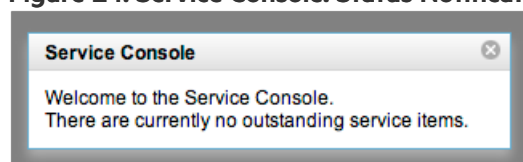
Figure 13: Service Console: Successful Completion of a Disk Drive Replacement



If the repair was completed correctly, the service action icon updates to green, in addition to visual cues appearing within the FRU replacement procedure. The following screenshot illustrates an example of a complete, successful disk drive replacement using the guided procedure on the service console.

8. When all service actions are complete, the following message appears:

Figure 14: Service Console: Status Notification



9. Log out of the service console.

If required, notify the system administrator that all repairs have been made and the service console login can be disabled.



service_console Commands

Modes: Daily

The `service_console` command lets a site administrator manage the system service console and block or allow its use for a service technician.

The service console is an alternate system management interface that is limited to maintenance functions, and is for use by authorized service personnel only. Even though it bypasses the normal administrative login, the service console cannot be used to access the root account, or to access data on the Lustre file system.

While it is not necessary to disable the service console for any system that is physically secure, administrators in high-security environments might prefer to disable it. In that case, they will need to re-enable the service console every time a service technician arrives on site to perform system maintenance, including replacement of failed disk drives and other parts.

Configuration consists of two main areas:

- Configuring the SMTP relay, which allows the system to send emails to users and service emails
- Configuring user email addresses to allow service email to go to system users

The following commands and available subcommands are used to manage the service console:

Synopsis

```
$ cscli service_console [-h] {enable,disable,show,notifications,configure}
```

where

Subcommands	Description
<code>enable</code>	Enables the service console
<code>disable</code>	Disables the service console
<code>show</code>	Show service console enabled status
<code>notifications</code>	View RAS notifications
<code>configure</code>	Configure RAS notifications

Optional Arguments	Description
--------------------	-------------

<code>-h --help</code>	Displays the help message and exits
--------------------------	-------------------------------------

Examples

Enable the service console.

```
$ cscli service_console enable
Service Console enabled.
```

Disable the Service console.

```
$ cscli service_console disable
Service Console disabled.
```

Show the status of the service console (enabled).

```
$ cscli service_console show
Service console enabled.
```

Show service console notifications (service events).

```
$ cscli service_console notifications show
Current outstanding service call events:

Item 1 : "Disk drive needs replacement"

Service Code: 002005001
Time the event was first detected: Wed, 17 Jun 2015 17:18:14 EDT
```

Details of failed component
Disk Serial #: S0M122HN0000B40298QD
Disk Model: ST600MM0006 Drive
Manufacturer: SEAGATE
Firmware version of drive at time of failure: XLGE T10 Enabled: Yes

Location of failed component

Rack Name: Rack1
Enclosure Model: 2U24
Enclosure Location: 36U
Disk located in slot: 15

Item 2: "Power supply issue detected"

Service Code: 002005003
Time the event was first detected: Wed, 02 Dec 2015 12:58:57 PST

Details of failed component

Power supply chassis type: OneStor Power One 764W_AC_PCM (Original Silver DFM).
Power supply part number: 0945768-10
Power supply product version: 0314
Power supply serial number: PMW0945768J1BDY

Location of failed component

Rack Name: R1C1
Enclosure Model: 2U24
Enclosure Location: 24U
Power supply located in bay: 1

Show service console configuration options.

```
$ cscli service_console configure
Commands:
  smtp          Configure SMTP settings
  snmp          Configure SNMP settings
  email         Configure email for sending RAS notifications
  system       Configure system settings
  rest_api      Configure REST API settings
```

notifications Subcommand

The `notifications` command is a subcommand of the `service_console` command, and is used to view notifications.

Synopsis

```
$ cscli service_console notifications [-h] {show}
```

Subcommands	Description
-------------	-------------

<code>show</code>	Show outstanding notifications.
-------------------	---------------------------------

Optional Arguments	Description
--------------------	-------------

<code>-h</code> <code>--help</code>	Displays the help message and exits.
---------------------------------------	--------------------------------------

Service Console Configuration Commands

These commands are used to configure the SMTP relay, SNMP, remote support, email addresses for RAS notifications, system and REST

API settings.

Synopsis

```
$ cscli service_console configure {smtp,snmp,remote_support,email,system,rest_api}
```

Subcommands	Description
smtp	Configures SMTP settings
snmp	Configures SNMP settings
remote_support	Configure remote support RAS notifications
email	Configure email for sending RAS notifications
system	Configure system settings
rest_api	Configure REST API settings

Optional Arguments	Description
-h --help	Display the help message and exit

smtp Commands

The `smtp` command is a second-level subcommand of the `service_console` command, and is used to configure SMTP.

Synopsis

```
$ cscli service_console configure smtp [-h]
```

Usage

```
cscli service_console [options] configure [options] smtp [options]
```

Optional Arguments	Description
-h --help	Displays the help message and exits.

Entry point for ClusterStor Manager CLI.

Subcommands	Description
relay	Configure SMTP relay for sending notifications
from	Configure SMTP from email address for sending notifications
show	Display the SMTP configuration

relay Command

The `relay` command is the third-level subcommand of the `service_console` command, and is used to configure SMTP relay.

Synopsis

```
$ cscli service_console configure smtp relay [-h] -H smtp_host -P smtp_port
```

Subcommands	Description
-P smtp_port --port smtp_port	SMTP port [default: 25]
-H smtp_host --host smtp_host	SMTP host (required)



Optional Arguments	Description
--------------------	-------------

<code>-h --help</code>	Displays the help message and exits.
--------------------------	--------------------------------------

Examples

Configure the SMTP relay.

```
$ cscli service_console configure smtp relay -H
mailrelayus.xyus.xyratex.com
OK.
```

Show the current SMTP configuration.

```
$ $ cscli service_console configure smtp show
SMTP relay: mailrelayus.xyus.xyratex.com:25
```

snmp Command

The `snmp` command is a second-level subcommand of the `service_console` command, and is used to configure SNMP settings

Synopsis

```
$ cscli service_console configure snmp [-h] {enable,disable,show}
```

Subcommands	Description
-------------	-------------

<code>enable</code>	Enables SNMP monitoring
<code>disable</code>	Disables SNMP monitoring
<code>show</code>	Shows current SNMP settings

Optional Arguments	Description
--------------------	-------------

<code>-h --help</code>	Displays the help message and exits
--------------------------	-------------------------------------

remote_support Commands

The `remote_support` command is a second-level subcommand of the `service_console` command, and is used to configure SMTP remote support.

Synopsis

```
$ cscli service_console configure remote_support [-h] {enable,disable,show}
```

Subcommands	Description
-------------	-------------

<code>enable</code>	Enables remote support
<code>disable</code>	Disables remote support
<code>show</code>	Shows current remote support status

Optional Arguments	Description
--------------------	-------------

<code>-h --help</code>	Displays the help message and exits
--------------------------	-------------------------------------

Examples

Disable remote support.

```
$ cscli service_console configure remote_support disable
```

Remote Support disabled.

Display remote support status.

```
$ cscli service_console configure remote_support show
Remote Support enabled.
```

Enable remote support.

```
$ cscli service_console configure remote_support enable
Remote Support enabled.
```

Email Notification Commands

This command configures email notifications, including enabling or disabling notification functionality, adds or deletes email addresses to the user list, shows user list entries and sends test email to determine if user notifications can be sent successfully.

Synopsis

```
$ cscli service_console configure email [-h] {show,add,delete,enable,disable,send_test_email}
```

Subcommands	Description
show	Shows all notification email addresses
add	Adds a notification email address. Also specify the email address to be added (required): <code>-A email_address</code> <code>--address email_address</code>
delete	Deletes a notification email address Also specify the email address to be deleted (required): <code>-A email_address</code> <code>--address email_address</code>
enable	Enables email notifications
disable	Disables email notifications
send_test_email	Sends test email to all enabled services

Optional Arguments	Description
<code>-h</code> <code>--help</code>	Displays the help message and exits

Examples

Show email addresses configured to receive RAS notifications (no configured addresses).

```
$ cscli service_console configure email show
No email addresses are configured to receive service notifications.
```

Show email addresses configured to receive RAS notifications (several configured addresses).

```
$ $ cscli service_console configure email show
The following email addresses are configured to receive service notifications:
sam_jones@xyzcorp.com
amy_cooper@xyzcorp.com
```

Add an email address to the user list (success).

```
$ cscli service_console configure email add -A sam_jones@xyzcorp.com
Successfully added 'sam_jones@xyzcorp.com' to receive notification emails.
```

Add an email address to the user list (fails because it is currently associated with the user list).

```
$ cscli service_console configure email add -A sam_jones@xyzcorp.com
```

```
Unable to add: 'sam_jones@xyzcorp.com' is already configured for receiving notification emails.
```

Delete an email address from the user list (success).

```
$ cscli service_console configure email delete -A sam_jones@xyzcorp.com
Successfully removed 'sam_jones@xyzcorp.com' from receiving notification emails.
```

Delete an email address from the user list (fails because it is not associated with the user list).

```
$ cscli service_console configure email delete -A sam_jones@xyzcorp.com
Unable to delete: 'sam_jones@xyzcorp.com' is not configured for receiving notification emails.
```

Enable email notifications.

```
$ cscli service_console configure email enable
OK.
```

Disable email notifications.

```
$ cscli service_console configure email disable
OK.
```

system Command

The `system` command is a second-level subcommand of the `service_console` command, and is used to configure system settings.

Synopsis

```
$ cscli service_console configure system [-h] {identifier,show}
```

Subcommands	Description
<code>identifier</code>	Sets the system identifier
<code>show</code>	Shows the current system identifier (name)

Optional Arguments	Description
<code>-h --help</code>	Displays the help message and exits

Examples

Show the current system identifier (name specified).

```
$ cscli service_console configure system show
System settings:
  System identifier name: snx-002
```

Show the current system identifier (no name set).

```
$ cscli service_console configure system show
```

identifier Command

The `identifier` command is a third-level subcommand of the `service_console` command, and is used to set the system identifier name.

Synopsis

```
$ cscli service_console configure system identifier [-h] -n id_name
```

Usage

```
$ cscli service_console [options] configure [options] system [options] identifier [options]
```

Optional Arguments	Description
--------------------	-------------

Optional Arguments	Description
<code>-h --help</code>	Displays the help message and exits
<code>-n <i>id_name</i> --name=<i>id_name</i></code>	Name (required)

Example

Set the system identifier.

```
$ cscli service_console configure system identifier --name=snx-002
System identifier name has been set to: snx-002
```

rest_api Command

The `rest_api` command is a second-level subcommand of the `service_console` command, and is used to configure REST API settings.

Synopsis

```
root@cls12345# cscli service_console configure rest_api [-h]
```

Usage

```
cscli service_console [options] configure [options] rest_api [options]
```

Subcommands	Description
<code>enable</code>	Enable REST API
<code>disable</code>	Disable REST API
<code>show</code>	Inform of REST API access status and list authorized users
<code>user_add</code>	Add a user to list of REST API authorized users
<code>user_delete</code>	Remove a user from the list of REST API authorized users

Optional Arguments	Description
<code>-h --help</code>	Displays the help message and exits

Customize Lustre Configuration

It may be necessary to customize the configuration for the ClusterStor system, to modify default parameters and/or to setup LNet routing.

For additional information on LNet routing configuration in general, see the [Lustre Operations Manual](http://doc.lustre.org/lustre_manual.xhtml) at doc.lustre.org/lustre_manual.xhtml. For more information on LNet routing configurations for systems with Cray XC clients, see the [XC™ Series Lustre® Administration Guide S-2648](#).

For notes to prepare the local configuration, see [Prepare the Local Configuration](#).

For recommended parameter settings for the Lustre server nodes, see [.](#)

For instructions to generate the config files, see [Create Custom LNet Configuration for ClusterStor](#).

A note about ARP-Flux in release 4.1:

Due to Linux routing quirks, if there are two (2) network interfaces on the same node, the HW address returned in the ARP for a specific IP might not necessarily be the one for the exact interface being ARPed. This causes problems for `o2iblnd`, because it resolves the address using IPoIB, and gets the wrong Infiniband address. To get around this problem, it is **required** to setup routing entries and rules to tell the Linux kernel to respond with the correct HW address. Also see [How to Convert ClusterStor from Bonded to Multi-rail](#) under `lustre lnet multi-rail` Command in the [ClusterStor CCLI Command Reference Guide S-9922](#).

Prepare the Local Configuration

The ClusterStor system uses puppet-generated files for the Lustre/LNet config:

`/etc/modprobe.d/lnet.conf`

`/etc/lustre/ip2nets.dat` (OPTIONAL)

`/etc/lustre/routes.dat` (OPTIONAL)

The `ip2nets` and routes details can be included in the `lnet` config file or placed in separate files.

The `lnet.conf` file is used to configure LNet and related Lustre parameters. The routes and `ip2nets` configuration entries can be included in the `lnet.conf` file or placed in separate files. Larger configurations—where the character count for the `ip2nets` or routes entry exceeds the `modprobe` limit of 1024—must use separate files for `ip2nets` and/or routes.

Create a subdirectory in `~admin` on management node 0 to store the custom configuration files, for example, `~admin/lnet_conf`. Create a custom `lnet.config` file—as well as `ip2nets` and routes files, if necessary—then copy them into `~admin/lnet_conf`.

Example `lnet.config` file:

```
## ko2iblnd parameters
options ko2iblnd timeout=10
options ko2iblnd peer_timeout=0
options ko2iblnd keepalive=30
options ko2iblnd credits=2048
options ko2iblnd ntx=2048
### NOTE peer_credits must be consistent across all peers on the IB network
options ko2iblnd peer_credits=16
options ko2iblnd concurrent_sends=16

## LNet parameters
options lnet ip2nets="o2ib3(ib0) 10.10.100.*"
options lnet routes="gni 10.10.100.[60,61,62,63,64,65,66,67]@o2ib3"
options lnet router_ping_timeout=10
options lnet alive_router_check_interval=35

## Sonexion only (if off by default)
options lnet avoid_asym_router_failure=1

## ptlrpc parameters
options ptlrpc at_max=400
options ptlrpc at_min=40
options ptlrpc ldln_enqueue_min=260
```

The `ip2nets` entry assigns the designated o2ib LNet index to the servers with the corresponding IP addr.

The `routes` entry sets up routes to the corresponding LNet network, 'gni' in this case. The hosts at 10.10.100.[100,101,102] are LNet routers connected to both o2ib3 and the gni networks.

In some configurations, there may be hosts on the InfiniBand fabric with more than one connection to the fabric.

Those connections should be configured to use separate IP subnets. For LNet to properly handle such a configuration, IP-Aliasing is used on interface `ib0` to configure additional IP subnets on the Lustre server nodes. This can be managed in the `ip2nets` specification by appropriately assigning `lnet nids` to the `ib` aliases, and in the `modprobe lnet.config` file by setting up `IB` aliases at module load time. Note that for systems using fine-grained routing, there may be separate subnets used for MDS vs. OSS nodes.

Examples of `ip2nets` and `modprobe` config entries for adding IP addresses to the `ib0` alias on the servers (10.10.102 on mds nodes, 10.10.101 on OSS nodes):



NOTE: For `ip2nets`:

```
o2ib1 10.10.100.*
o2ib1000(ib0:1) 10.10.102.*
o2ib1002(ib0) 10.10.100.*
o2ib1003(ib0:1) 10.10.101.*
o2ib1004(ib0) 10.10.8.*
o2ib1005(ib0:1) 10.10.101.*
```

For `lnet.config`, these two lines can be added at the start of the local `lnet.config` file. The first line creates the `ib0:1` aliases on `lnet` module load, the second line removes the aliases when `lnet` is unloaded:



NOTE:

```
install lnet if nodeattr mds; then /sbin/ip -o -4 a show ib0 | /usr/bin/awk '/inet/{s=$4;sub("10\\\.10\\\.100\\\.","10.10.102.",s);print "/sbin/ip address add dev i

remove lnet /sbin/modprobe -r --ignore-remove lnet && /sbin/ip -o -4 a show label ib0:1 | awk '{print "/sbin/ip address del dev ib0 label ib0:1", $4}' | /bin/sh
```

For recommended parameter settings for the Lustre server nodes, see . For instructions to generate the config files, see .

Lustre Server: Recommended Tuning Parameters

This topic provides recommended tuning parameters for Lustre servers in ClusterStor systems.

For Lustre setups using InfiniBand, the `peer_credits` setting must be consistent across all InfiniBand (IB) peers on the Lustre network. When routers and/or external Lustre clients have mlx5-based Host Channel Adapters (HCAs), `map_on_demand` must be set to 0. In addition, Cray/HPE recommends `peer_credits` and `concurrent_sends` be set to 16. `map_on_demand` must be set to 0. `peer_credits` and `concurrent_sends` should be set to 16 for all IB peers on the Lustre network.

If an IB peer must have access to an mlx4-based file system and an mlx5-based file system, the `ko2iblnd` parameters of all mlx4 peers must match the `ko2iblnd` mlx5-peer parameters to ensure shared mlx4- and mlx5-peer function.

All IB peers within the Lustre network must be able to support these same values if they are to be used.

Cray/HPE recommends an object-based disk (OBD) timeout of 100 seconds, which is the default value. Set this parameter using the `lctl conf_param` command on the management server (MGS). For example,

```
# lctl conf_param fs_name.sys.timeout=100
# cat /proc/sys/lustre/timeout
100
```

`ko2iblnd:timeout`

Default 10. The `o2iblnd` timeout in seconds. Cray/HPE recommends setting this to 10 seconds.

`ko2iblnd:peer_timeout`

Default 0. Number of seconds without aliveness news it takes to declare a peer dead. Cray/HPE recommends setting this to 0.

`ko2iblnd:keepalive`

Default 30. Idle time in seconds before sending a keepalive. Cray/HPE recommends setting this to 30.

`ko2iblnd:credits`

Default 2048. Number of concurrent sends allowed by `o2iblnd`. Shared by all CPU partitions (CPT). Cray/HPE recommends setting this to 2048.

`ko2iblnd:ntx`

Default 2048. Number of message descriptors allocated for each pool. Cray/HPE recommends setting this to 2048.

`ko2iblnd:peer_credits`

Enter the value for the `ko2iblnd` parameter `peer_credits`. This is the number of concurrent sends to a single peer. This value must be the same on all external login clients and the Lustre file system servers.

`ko2iblnd:peer_credits 16`

Enter the value for the `ko2iblnd` parameter `peer_credits`. This is the number of concurrent sends to a single peer. This value must be the same on all external login clients and the Lustre file system servers.

`ko2iblnd:concurrent_sends 16`

Determines send work queue sizing. If this option is omitted, the default is calculated based on the values of `peer_credits` and `map_on_demand`. This value must be the same on the external login clients and the Lustre file system servers.

`ko2iblnd:map_on_demand`

Default 0. Controls the use of fast memory registration (FMR). Cray/HPE recommends setting this value to 0 for InfiniBand HCAs.

`lnet:alive_router_check_interval`

Default 60. Number of seconds between alive (live or dead) router health checks. Recommendation: set this value to 35 seconds.

A value less than or equal to 0 disables pinging of alive (live or dead) routes.

`lnet:router_ping_timeout`

Default 10. Number of seconds to wait for the reply to a router health query. Cray/HPE recommends using the default value of 10 seconds.

`lnet:avoid_asym_router_failure`

Avoid asymmetrical router failures (0 to disable; 1 to enable).

`module_params:ptlrpc_at_max`

Default 400. Adaptive timeout maximum in seconds. Cray/HPE recommends setting this to 400 on a CLE client.

`module_params:ptlrpc_at_min`

Default 40. Adaptive timeout minimum in seconds. Cray/HPE recommends setting this to 40 on a CLE client.

`module_params:ptlrpc_ldlm_enqueue_min`

Default 260. Lock enqueue timeout minimum in seconds. This is the minimum amount of time a server will wait to see traffic on a lock before it assumes a client is misbehaving and takes action to revoke the lock by evicting the client. Cray/HPE recommends setting this to 260 for CLE clients.

Create Custom LNet Configuration for ClusterStor

Follow this procedure to create a custom LNet configuration on the ClusterStor system in Daily Mode. The file system must be shut down to update the LNet configuration.

Prerequisites

For a new system, first complete the setup procedures described in the [Cray ClusterStor Field Installation Guide 4.x H-6203](#).

Procedure

1. Log in to the primary management mode.
2. Prior to stopping the Lustre file system on the ClusterStor server, verify that client systems have unmounted Lustre.

Use the MDS node to confirm that all Lustre clients are unmounted. Run the following from the primary MGMT node:

```
MGMT0$ ssh mds_nodename "lctl get_param '.*.exports.*.uuid'"
```

- Example output showing one client mounted and Lustre still running:

```
admin@cls12345n000$ ssh cls12345n003 "lctl get_param '.*.exports.*.uuid'"
mdt.testfs-MDT0000.exports.172.18.1.188@o2ib.uuid=2412630a-db8d-806a-09eb-0690c8e1e86b
```

It is recommended that every effort be made to identify and stop Lustre clients before unmounting Lustre. In the sample output above, `172.18.1.188@o2ib` (shown underscored) represents a client that still has the Lustre file system mounted and needs to be addressed. See Step if it is necessary to unmount a client manually.

- Example output showing no clients mounted and Lustre still started:

```
admin@cls12345n000$ ssh cls12345n003 "lctl get_param '.*.exports.*.uuid'"
```

- Example output showing no clients mounted and Lustre stopped:

```
admin@cls12345n000$ ssh cls12345n003 "lctl get_param '.*.exports.*.uuid'"
error: get_param: /proc/
{fs,sys}
/
{lnet,lustre}
///exports/*/uuid: Found no match
```

3. If a client system needs to be unmounted manually, run the following command:

```
Client$ sudo umount lustre_mountpoint
```

Example:

```
Client$ sudo umount /mnt/lustre
```

4. Stop the Lustre file system.

```
admin@cls12345n000$ cscli unmount
```

5. Change to `root`.

```
admin@cls12345n000$ sudo su -
```

6. Change the `o2ib` index, if necessary (for example, if clients will mount multiple Lustre file systems and another file system is using `o2ib` or `o2ib0` as the LNet index):

```
root@cls12345n000# cscli lustre lnet set_nid -i new_nid_index
```

7. Close the MySQL session.

```
mysql> quit
```

8. Run the `beSystemNetConfig.sh` script on the primary management node and wait for it to finish.

```
root@cls12345n000# /opt/xyratex/bin/beSystemNetConfig.sh \
[-c path to file/lnet.config] [-r ~path to file/routes.conf] \
[-i path to file/ipnets.conf] $(nodeattr -v cluster)
```


To find *clustername*, run the `"cscli csinfo"` command.

Each of the files in this example is optional, and the user needs to use the command according to their LNet configuration needs.

9. Verify that the customized LNet configuration has been applied, by listing the `nids`.

```
root@cls12345n000# pdsh -g lustre lctl list_nids | sort
```

10. Start the Lustre file system and wait for the targets to mount on all system nodes.

```
root@cls12345n000# cscli mount
```

Save/Restore Lustre Filesystem Parameters

Follow this procedure to save or restore Lustre configuration on a ClusterStor system in Daily mode. The filesystem must be shut down to perform the procedure.

Procedure

1. Log in to the primary management mode.
2. Prior to stopping the Lustre filesystem on the ClusterStor server, verify that client systems have unmounted Lustre.

Use the MDS node to confirm that all Lustre clients are unmounted. Run the following from the primary MGMT node:

```
MGMT0$ ssh mds_nodename "lctl get_param '*.*.exports.*.uuid'"
```

- Example output showing one client mounted and Lustre still running:

```
admin@cls12345n000$ ssh cls12345n003 "lctl get_param \
'*.*.exports.*.uuid'"
mdt.testfs-MDT0000.exports.172.18.1.188@o2ib.uuid=2412630a-db8d-806a-09eb-0690c8e1e86b
```

It is recommended that every effort be made to identify and stop Lustre clients before unmounting Lustre. In the sample output above, `172.18.1.188@o2ib` (shown underscored) represents a client that still has the Lustre filesystem mounted and needs to be addressed. See Step , if it is necessary to unmount a client manually.

- Example output showing no clients mounted and Lustre still started:

```
admin@cls12345n000$ ssh cls12345n003 "lctl get_param \
'*.*.exports.*.uuid'"
```

- Example output showing no clients mounted and Lustre stopped:

```
admin@cls12345n000$ ssh cls12345n003 "lctl get_param \
'*.*.exports.*.uuid'"
error: get_param: /proc/
{fs,sys}
/
{lnet,lustre}
///exports/*/uuid: Found no match
```

3. If a client system needs to be unmounted manually, run the following command:

```
Client$ sudo umount lustre_mountpoint
```

Example:

```
Client$ sudo umount /mnt/lustre
```

4. Stop the Lustre filesystem.

```
admin@cls12345n000$ cscli unmount
```

5. Use `cscli lustre config` commands to configure Lustre:

cscli lustre config Command

The `config` command is a subcommand of `lustre`. This command dumps and restores lustre configuration.

Synopsis

```
# cscli lustre config [-h] {dump, restore} ...
```

Positional Arguments	Description
dump	Dumps lustre configuration
restore	Restores lustre configuration from file

Optional Arguments	Description
<code>-h --help</code>	Shows help message and exits

The `dump` command should be used to save the current Lustre configuration, which would later be used to restore said configuration, primarily in the event of—but not limited to—a reset.

`cscli lustre config dump` Subcommand

The `dump` command is a subcommand of the `config` command. This subcommand makes a yaml file out of current lustre configuration and saves it as `/etc/lustre/lustre_config.yaml`.

Synopsis

```
# cscli lustre config dump [-h] ...
```

Optional Arguments	Description
<code>-h --help</code>	Shows help message and exits

Note that the existing file will be replaced when the `dump` command is run again.

`cscli lustre config restore` Subcommand

The `restore` command is a subcommand of the `pool` command. This subcommand restores the saved lustre configuration from `/etc/lustre/lustre_config.yaml`.

Synopsis

```
# cscli lustre config restore [-h] ...
```

Optional Arguments	Description
<code>-h --help</code>	Shows help message and exits

It is acceptable to manually edit the config file `/etc/lustre/lustre_config.yaml`.

6. Start the Lustre filesystem and wait for the targets to mount on all system nodes.

```
root@cls12345n000# cscli mount
```

Change Network Settings in ClusterStor 4.x

To change the DNS resolver configuration or the NTP configuration, see [ean Command](#).

To change externally-facing IP addresses, see [Change Externally Facing IP Addresses](#).

To change LDAP or NIS settings in Daily Mode, see [Change LDAP or NIS Settings in Daily Mode](#).

To change the bonding mode type, see [Change Bonding Mode Type](#)

ean Command

Introduced in Software Release: 3.2

Use the `ean` commands to configure the appliance network. These commands deprecate `network` commands (introduced with ClusterStor 2.1.0.)

Synopsis

```
csccli ean [-h] {show,apply,ntp,dns,primary,secondary,route,ipaddr,mmu}
```

Positional Arguments	Description
<code>show</code>	Show all EAN configuration settings
<code>apply</code>	Apply configuration settings
<code>ntp</code>	NTP configuration. Configuring an NTP server on the EAN will set time for the entire cluster
<code>dns</code>	DNS configuration. Configuring DNS on the EAN will provide DNS for the entire cluster
<code>primary</code>	Primary interface configuration for EAN
<code>secondary</code>	Secondary interface configuration for a split EAN
<code>route</code>	Configure custom routes for the EAN
<code>ipaddr</code>	Manage static IP addresses on EAN interfaces of the MGMT nodes
<code>mmu</code>	Interface configuration for the EAN on MMU

Optional Arguments Description

<code>-h --help</code>	Display the help message and exit
--------------------------	-----------------------------------

ean apply Subcommand

Introduced in Software Release: 3.2

The `ean apply` command is a subcommand of the `ean` command. Use the subcommand to apply the configuration changes to the management network of a ClusterStor system.

Synopsis

```
$ cscli ean apply [-h]
```

Optional Arguments	Description
--------------------	-------------

<code>-h --help</code>	Display the help message and exit
------------------------	-----------------------------------

ean show Subcommand

Introduced in Software Release: 3.2

The `ean show` command is a subcommand of the `ean` command. Use this subcommand to display all management network settings of a ClusterStor system.

Synopsis

```
$ cscli ean show [-h]
```

Optional Arguments	Description
--------------------	-------------

<code>-h --help</code>	Display the help message and exit
--------------------------	-----------------------------------

ean dns Subcommand

Introduced in Software Release: 3.2

The `ean dns` command is a subcommand of the `ean` command. Use this subcommand to configure DNS settings on a ClusterStor system.

Synopsis

```
$ cscli ean dns [-h] {set,clear}
```

Optional Arguments	Description
--------------------	-------------

<code>set</code>	Set DNS servers
<code>clear</code>	Clear DNS servers
<code>-h --help</code>	Display the help message and exit

ean dns clear Subcommand

Introduced in Software Release: 3.2

The `ean dns clear` command is a subcommand of the `ean dns` command. Use this subcommand to remove DNS servers from ClusterStor configuration.

Synopsis

```
$ cscli ean dns clear [-h] [-a]
```

Optional Arguments	Description
<code>-a</code>	Apply changes automatically (This would eliminate the need to run separate <code>apply</code> step)
<code>-h --help</code>	Display the help message and exit

ean dns set Subcommand

Introduced in Software Release: 3.2

The `ean dns set` command is a subcommand of the `ean dns` command. Use this subcommand to add new DNS servers to ClusterStor configuration.

Synopsis

```
$ cscli ean dns set [-h] [-a] dns_server [dns_server ...]
```

Optional Arguments	Description
<code>-a</code>	Apply changes automatically (this would eliminate the need to run separate <code>apply</code> step)
<code>dns_server</code>	DNS server in IP address format. ClusterStor allows the use of multiple DNS servers.
<code>-h --help</code>	Display the help message and exit

ean ipaddr Subcommand

Introduced in Software Release: 3.2

The `ipaddr` command is a subcommand of the `ean` command. Use this subcommand configure the static IP addresses the EAN interfaces on a ClusterStor system.

Synopsis

```
$ cscli ean ipaddr [-h] {show,set}
```

Optional Arguments	Description
<code>set</code>	Set the IP address of an EAN interface on a node
<code>show</code>	Show the IP address of the EAN interfaces on all nodes
<code>-h --help</code>	Display the help message and exit

ean ipaddr set Subcommand

Introduced in Software Release 3.2

The `ipaddr set` command is a subcommand of the `ean` command. Use this subcommand to configure IP addresses on the EAN interfaces.

Synopsis

```
$ csccli ean ipaddr set [-h] --node NODE --type {EAN,SecondaryEAN,MmuEAN} --address ADDRESS
```

Optional Arguments	Description
<code>--node NODE</code>	Name of the node where the IP address needs to be configured
<code>-t</code> <code>{System,EAN,SecondaryEAN,MmuEAN} </code> <code>--type</code> <code>{System,EAN,SecondaryEAN,MmuEAN}</code>	Network Type
<code>--address ADDRESS</code>	IP address. NOTE: Do not specify prefix or netmask here. See <code>ean primary secondary</code> for details
<code>-h --help</code>	Display the help message and exit



ean ipaddr show Subcommand

Introduced in Software Release: 3.2

The `ean ipaddr show` command is a subcommand of the `ean` command. Use this subcommand to display IP addresses assigned to the EAN interfaces throughout the cluster.

Synopsis

```
$ cscli ean ipaddr show [-h]
```

Optional Arguments	Description
--------------------	-------------

<code>-h --help</code>	Display the help message and exit
--------------------------	-----------------------------------

ean mmu Subcommand

Introduced in Software Release: 3.2

The `ean mmu` command is a subcommand of the `ean` command. Use this subcommand to configure secondary EAN interfaces on an MMU in a ClusterStor system.

Synopsis

```
$ csccli ean mmu [-h] {add,delete,show}
```

Optional Arguments	Description
--------------------	-------------

<code>add</code>	Add secondary EAN interface
<code>delete</code>	Remove secondary EAN interface
<code>show</code>	Show configured EAN interfaces
<code>-h --help</code>	Display the help message and exit



NOTE:

Users can only configure a secondary interface on MMU using this command.

ean mmu add Subcommand

Introduced in Software Release 3.2

The `ean mmu add` command is a subcommand of the `ean` command. Use this subcommand to add and configure secondary EAN interface on an MMU.

Synopsis

```
$ cscli ean mmu add [-h] -i IFACE [--prefix-length PREFIX] [--gateway GATEWAY]
```

Optional Arguments	Description
<code>-i --iface IFACE</code>	Name of the interface
<code>--prefix-length PREFIX</code>	Prefix of network mask
<code>--gateway GATEWAY</code>	Gateway of network (in IP address format)
<code>-h --help</code>	Display the help message and exit

ean mmu delete Subcommand

Introduced in Software Release: 3.2

The `ean mmu delete` command is a subcommand of the `ean` command. Use this subcommand to remove secondary Enterprise Access Network (EAN) interface on an MMU node in ClusterStor.

Synopsis

```
$ cscli ean mmu delete [-h] -i IFACE
```

Optional Arguments	Description
<code>-i --iface IFACE</code>	Name of the interface
<code>-h --help</code>	Display the help message and exit

ean mmu show Subcommand

Introduced in Software Release 3.2

The `ean mmu show` command is a subcommand of the `ean` command. Use this subcommand to show configured secondary Enterprise Access Network (EAN) interface on an MMU node in the system.

Synopsis

```
$ cscli ean mmu show [-h]
```

Optional Arguments	Description
--------------------	-------------

<code>-h --help</code>	Display the help message and exit
------------------------	-----------------------------------

ean ntp Subcommand

Introduced in Software Release: 3.2

The `ean ntp` command is a subcommand of the `ean` command. Use this subcommand to configure NTP settings on a ClusterStor system.

Synopsis

```
$ cscli ean ntp [-h] {set,clear}
```

Optional Arguments	Description
--------------------	-------------

<code>set</code>	Set NTP servers
<code>clear</code>	Clear NTP servers
<code>-h --help</code>	Display the help message and exit

ean ntp clear Subcommand

Introduced in Software Release: 3.2

The `ntp clear` command is a subcommand of the `ean` command. Use this subcommand to remove NTP servers from ClusterStor configuration.

Synopsis

```
$ cscli ean ntp clear [-h] [-a]
```

Optional Arguments	Description
<code>-a</code>	Apply changes automatically (This would eliminate the need to run separate <code>apply</code> step.)
<code>-h --help</code>	Display the help message and exit

ean ntp set Subcommand

Introduced in Software Release: 3.2

The `ean ntp set` command is a subcommand of the `ean ntp` command. Use this subcommand to add new NTP servers to ClusterStor configuration.

Synopsis

```
$ cscli ean ntp [-h] {set,clear} cscli ean ntp set [-h] [-a] ntp_server [ntp_server ...]
```

Optional Arguments	Description
<code>-a</code>	Apply changes automatically (this would eliminate the need to run separate <code>apply</code> step
<code>ntp_server</code>	NTP server (either in IP address or FQDN formats). NOTE: if FQDN format is used, DNS must be configured. ClusterStor allows use of multiple NTP servers.
<code>-h --help</code>	Display the help message and exit

ean primary Subcommand

Introduced in Software Release 3.2

The `ean primary` command is a subcommand of the `ean` command. Use this subcommand to configure the primary (pub0) public management interface (EAN) on a management node of the ClusterStor system.

Synopsis

```
$ cscli ean primary [-h] {add,delete,show}
```

Optional Arguments	Description
--------------------	-------------

<code>add</code>	Add primary EAN interface
<code>delete</code>	Delete primary EAN interface
<code>show</code>	Show list of primary EAN interfaces
<code>-h --help</code>	Display the help message and exit

ean primary add Subcommand

Introduced in Software Release: 3.2

The `ean primary add` command is a subcommand of the `ean primary` command. Use this subcommand to add the primary Enterprise Access Network (EAN) interface to ClusterStor configuration.

Synopsis

```
$ cscli ean primary add [-h] -i IFACE [--prefix-length PREFIX] [--gateway GATEWAY]
```

Optional Arguments	Description
<code>-i --iface IFACE</code>	Name of the interface (typically <code>pub0</code>)
<code>--prefix-length PREFIX</code>	Prefix of network mask
<code>--gateway GATEWAY</code>	Gateway of network (in IP address format)
<code>-h --help</code>	Display the help message and exit

ean primary delete Subcommand

Introduced in Software Release: 3.2

The `ean primary delete` command is a subcommand of the `ean primary` command. Use this subcommand to remove primary Enterprise Access Network (EAN) interface from ClusterStor configuration.

Synopsis

```
$ cscli ean primary delete [-h] -i IFACE
```

Optional Arguments	Description
<code>-i --iface IFACE</code>	Name of the interface (typically <code>pub0</code>)
<code>-h --help</code>	Display the help message and exit

ean primary show Subcommand

Introduced in Software Release: 3.2

The `ean primary show` command is a subcommand of the `ean primary` command. Use this command to show configured primary Enterprise Access Network (EAN) interfaces in ClusterStor.

Synopsis

```
$ cscli ean primary show [-h]
```

Optional Arguments	Description
--------------------	-------------

<code>-h --help</code>	Display the help message and exit
------------------------	-----------------------------------

ean route Subcommand

Introduced in Software Release: 3.2

The `ean route` command is a subcommand of the `ean` command. Use this command to configure the custom routes for the EAN on a ClusterStor system.

Synopsis

```
$ cscli ean route [-h] {add,delete,set,load,clear,show}
```

Optional Arguments	Description
--------------------	-------------

<code>add</code>	Add routing rule
<code>delete</code>	Delete routing rule
<code>show</code>	Show list of routes
<code>set</code>	Update routing rule
<code>load</code>	Load list for routing rules
<code>clear</code>	Clear routing rules
<code>-h --help</code>	Display the help message and exit

ean route add Subcommand

Introduced in Software Release: 3.2

The `ean route add` command is a subcommand of the `ean` command. Use this command to add custom routes to ClusterStor system.

Synopsis

```
$ csccli ean route add [-h] (-t {System,EAN,SecondaryEAN,MmuEAN} | -i IFACE) -d DEST -p PREFIX -g GATEWAY
```

Optional Arguments	Description
<code>-t</code> <code>{System,EAN,SecondaryEAN,MmuEAN} </code> <code>--type</code> <code>{System,EAN,SecondaryEAN,MmuEAN}</code>	Network type
<code>-i --iface IFACE</code>	Name of the interface
<code>-d DEST --dest DEST</code>	Destination IP address
<code>-p PREFIX --prefix PREFIX</code>	Subnet prefix length (0-32)
<code>-g GATEWAY --gateway GATEWAY</code>	Gateway IP address
<code>-h --help</code>	Display the help message and exit



NOTE: When adding a custom route, only specify one of the following arguments: `-t|--type` or `-i|--iface IFACE`.

ean route clear Subcommand

Introduced in Software Release: 3.2

The `ean route clear` command is a subcommand of the `ean` command. Use this subcommand to remove all custom routes from ClusterStor configuration.

Synopsis

```
$ cscli ean route clear [-h] -t {System,EAN,SecondaryEAN,MmuEAN} [-i IFACE]
```

Optional Arguments	Description
<code>-t</code> <code>{System,EAN,SecondaryEAN,MmuEAN} </code> <code>--type</code> <code>{System,EAN,SecondaryEAN,MmuEAN}</code>	Network type
<code>-i --iface IFACE</code>	Name of the interface
<code>-h --help</code>	Display the help message and exit

ean route delete Subcommand

Introduced in Software Release: 3.2

The `route delete` command is a subcommand of the `ean` command. Use this command to remove previously-configured custom routes from ClusterStor system.

Synopsis

```
$ csccli ean route delete [-h] -r ROUTE_ID
```

Optional Arguments	Description
<code>-r ROUTE_ID --route-id ROUTE_ID</code>	Route identifier (see <code>ean route show</code> subcommand)
<code>-h --help</code>	Display the help message and exit

ean route load Subcommand

Introduced in Software Release: 3.2

The `ean route load` command is a subcommand of the `ean` command. Use this subcommand to load a file with custom routes to ClusterStor configuration.

Synopsis

```
$ cscli ean route load [-h] -i IFACE -f FILENAME
```

Optional Arguments	Description
<code>i --iface IFACE</code>	Name of the interface
<code>-f <i>FILENAME</i> --file <i>FILENAME</i></code>	Path to file with routing rules
<code>-h --help</code>	Display the help message and exit

ean route set Subcommand

Introduced in Software Release: 3.2

The `route set` command is a subcommand of the `ean` command. Use this command to update the existing custom route in ClusterStor configuration.

Synopsis

```
$ csccli ean route set [-h] -r ROUTE_ID [-d DEST] [-p PREFIX] [-g GATEWAY]
```

Optional Arguments	Description
<code>-r ROUTE_ID --route-id ROUTE_ID</code>	Route identifier (see <code>ean route show</code> subcommand)
<code>-d DEST --dest DEST</code>	Destination IP address
<code>-p PREFIX --prefix PREFIX</code>	Subnet prefix length (0-32)
<code>-g GATEWAY --gateway GATEWAY</code>	Gateway IP address
<code>-h --help</code>	Display the help message and exit



ean route show Subcommand

Introduced in Software Release: 3.2

The `route show` command is a subcommand of the `ean` command. Use this command to show custom routes in ClusterStor.

Synopsis

```
$ cscli ean route show [-h]
```

Optional Arguments	Description
--------------------	-------------

<code>-h --help</code>	Display the help message and exit
------------------------	-----------------------------------

ean secondary Subcommand

Introduced in Software Release: 3.2

The `ean secondary` command is a subcommand of the `ean` command. Use this subcommand to configure the secondary (`pub1`) public management interface (EAN) on a management node of the ClusterStor system (if present and connected).

Synopsis

```
$ csccli ean secondary [-h] {add,delete,show}
```

Optional Arguments	Description
add	Add secondary EAN interface
delete	Delete secondary EAN interface
show	Show list of secondary EAN interfaces
-h --help	Display the help message and exit



NOTE:

For configuring secondary EAM interface on an MMU node, please refer to `ean mmu` command.

ean secondary add Subcommand

Introduced in Software Release: 3.2

The `ean secondary add` command is a subcommand of the `ean secondary` command. Use this command to add secondary Enterprise Access Network (EAN) interface to ClusterStor configuration.

Synopsis

```
$ cscli ean secondary add [-h] -i IFACE [--prefix-length PREFIX] [--gateway GATEWAY]
```

Optional Arguments	Description
<code>-i --iface IFACE</code>	Name of the interface (typically <code>pub1</code>)
<code>--prefix-length PREFIX</code>	Prefix of network mask
<code>--gateway GATEWAY</code>	Gateway of network (in IP address format)
<code>-h --help</code>	Display the help message and exit

ean secondary delete Subcommand

Introduced in Software Release: 3.2

The `ean secondary delete` command is a subcommand of the `ean secondary` command. Use this command to remove secondary Enterprise Access Network (EAN) interface from ClusterStor configuration.

Synopsis

```
$ cscli ean secondary delete [-h] -i IFACE
```

Optional Arguments	Description
<code>-i --iface IFACE</code>	Name of the interface (typically <code>pub1</code>)
<code>-h --help</code>	Display the help message and exit

ean secondary show Subcommand

Introduced in Software Release: 3.2

The `ean secondary show` command is a subcommand of the `ean secondary` command. Use this subcommand to show configured secondary Enterprise Access Network (EAN) interfaces in ClusterStor.

Synopsis

```
$ cscli ean secondary show [-h]
```

Optional Arguments	Description
--------------------	-------------

<code>-h --help</code>	Display the help message and exit
------------------------	-----------------------------------

Change Externally Facing IP Addresses

Use this procedure to change externally facing IP addresses, for customers to change one or both of the IP addresses of a ClusterStor system after it has been installed. For additional IP address information, refer to the individualized site prep guide specific to the site/system. Also see `cscli ean {primary,secondary,ipaddr}` in the ClusterStor CSCLI Command Reference Guide S-9922.

Procedure

1. Connect to node 0. To set the pub0 addresses, enter the following from MGMTn000:

```
admin@cls12345n000$ cscli ean primary set --prefix-length {subnet mask} --gateway {address} -i pub0
admin@cls12345n000$ cscli ean ipaddr set --address node0 IPaddress --type EAN --node node0
admin@cls12345n000$ cscli ean ipaddr set --address node1 IPaddress --type EAN --node node1
admin@cls12345n000$ cscli ean apply
```

Example:

```
admin@cls12345n000$ cscli ean primary set --prefix-length 20 --gateway 172.30.22.1 -i pub0
admin@cls12345n000$ cscli ean ipaddr set --address 172.30.22.60 --type EAN --node cls12345n000
admin@cls12345n000$ cscli ean ipaddr set --address 172.30.22.61 --type EAN --node cls12345n001
admin@cls12345n000$ cscli ean apply
```

2. Reconnect to node 0 at the new address.

Change LDAP or NIS Settings in Daily Mode

Use the following CSCLI commands—stored in the `t0db` database table `ldap_setup`—to configure the LDAP or NIS nodes.

lustre users ldap set Subcommand

Introduced in Software Release: 3.1 (previously `lustre_users ldap set` and `set_lustre_users_ldap`)

The `lustre users ldap set` command is a subcommand of the `lustre users` command. Use the subcommand to configure LDAP settings. It allows for configuration of ldap uri (and port), base_dn, user_dn, group_dn, bind_dn, and password. In addition, it allows for configuration of TLS cert file, TLS ca cert file, and private key file. To properly configure LDAP using the `lustre users ldap set` command, directory must first be added to the nss order. This can be accomplished with the following command:

```
$ cscli lustre users order set local directory
```


Confirm the order by running the `lustre users order show` command. See the `lustre users order set` and `lustre users order show` commands for more information.

Synopsis

```
cscli lustre users ldap set [-h]
```

Usage

```
cscli lustre users ldap set [-h]
```

Optional Arguments	Description
<code>-l LDAP_URI</code> <code> --ldap-uri LDAP_URI</code>	URI, for example: <code>ldap://127.0.0.1:389</code>
<code>-m {default,rfc2307,rfc2307bis,ad,active-directory,mssfu35}</code> <code> --mapping {default,rfc2307,rfc2307bis,ad,active-directory,mssfu35}</code>	This option allows for custom attributes to be looked up instead of the default RFC 2307 attributes
<code>-b BASE_DN</code> <code> --base-dn base_dn</code>	Base Domain Name (DN)
<code>-u USER_DN</code> <code> --user-dn USER_DN</code>	User DN
<code>-G GROUP_DN</code> <code> --group-dn GROUP_DN</code>	Group DN
<code>-g GROUP_SEARCH</code> <code> --groupsearch GROUP_SEARCH</code>	LDAP search filter to use for a group map
<code>-s HOSTS_DN</code> <code> --hosts-dn HOSTS_DN</code>	Hosts DN
<code>-S {subtree,onelevel,base}</code> <code> --userscope {subtree,onelevel,base}</code>	Specify the search scope (subtree, one level, or base object)
<code>-i BIND_DN</code> <code> --bind-dn BIND_DN</code>	Bind DN
<code>-p password</code> <code> --password [PASSWORD]</code>	Bind password <div> WARNING: Entering a password on the command line may result in it being logged in plain text. Omit this parameter to be prompted for a password.</div>
<code>--pvt-key FILEPATH</code>	Private key file (pem)
<code>--tls-cert FILEPATH</code>	TLS certificate file (crt)

Optional Arguments	Description
<code>--tls-ca FILEPATH</code>	TLS ca certificate file (crt)
<code>--tls-reqcrt</code> <code>{allow, never, try, demand, hard}</code>	Specifies what checks to perform on server certificates in a TLS session, if any
<code>-h --help</code>	Show the help message and exit

lustre users nis set Subcommand

Introduced in Software Release: 3.1 (previously `lustre_users nis set` and `set_lustre_users_nis`)

The `lustre users nis set` command is a subcommand of the `lustre users` command. Use the subcommand to configure NIS settings. It allows for configuration of `nis_server` (1-3 times) IP and `nis_domain`. This subcommand is generally executed in daily mode.

Synopsis

```
$ cscli lustre users nis set [-h] [-s nis_server] [-d nis_domain]
```

Optional Arguments	Description
<code>-s <i>nis_server</i></code> <code> --nis_server <i>nis_server</i></code>	Specify this option 1 to 3 times using IP address or fully-qualified domain name
<code>-d <i>nis_domain</i></code> <code> --nis_domain <i>nis_domain</i></code>	NIS domain. Example: nisdomain
<code>-h --help</code>	Display the help message and exit

Change Bonding Mode Type

ClusterStor systems come with bonding mode enabled on the LDN. ClusterStor supports several bonding types:

- passive
- lacp
- balancealb

To change the bonding mode, use the `cscli lustre_network bonding` command.

lustre_network bonding Command

The `lustre_network bonding` command is a subcommand of the `lustre_network` command. Use the subcommand and subsequent subcommands to manage Ethernet bonding on the high speed data network.

Synopsis

```
$ cscli lustre_network bonding [-h] {show,mode,hash} ...
```

Positional Arguments	Description
show	Show bonding interfaces
mode	Manage Ethernet bonding mode
hash	Manage Ethernet bonding hash policy

Optional Arguments	Description
-h --help	Display the help message and exit

For more information about the Ethernet bonding driver, please see the official [Linux Ethernet Bonding Driver How To](#) document.

Reload the Lustre modules

When changing the bonding mode, it's important to reload the Lustre modules, to ensure proper operations.

A single command will unload all lustre modules after targets have been unmounted:

```
pdsh -g lustre "(lustre_rmmod; lctl net down; lustre_rmmod) &> /dev/null)"
```

Lustre modules are automatically loaded on filesystem mount, so no command is needed to manually load them.

Nodes with RoCE mode selected need to be rebooted to apply bonding mode changes.

Ensure Proper LDN Switch Configuration

When using LACP bonding mode, it's important to ensure the proper configuration of the LDN switches. To learn about switch configuration, see "Internal Rack Reference Information" in the ClusterStor E1000 Field Installation Guide H-6203.

If a later bonding mode changes from LACP to another mode, the switch configuration should be changed to remove the changes made.

Check File Systems Using `e2fsck` Command

IMPORTANT: This procedure must be performed by trained HPE Cray service personnel only. Do not use `e2fsck` commands, unless directed to do so by HPE Cray Support.

There may be times when it is necessary to run `e2fsck` on one (1) or more Lustre MDT or OST target devices (e.g., following a hard node failure or disk drive fallout). Some problems can result in the superblock `state` flag being marked clean with errors. This will prevent the corresponding `fsys` resource from starting until `e2fsck` has been run to verify and clean the device. Use the following procedure to run `e2fsck` on an individual MDT or OST device. The steps in the procedure describe running `e2fsck` on an OST.

Procedure

1. Log in, via SSH, to the OSS node containing the devices where the `e2fsck` command is to be run. The following example assumes the OSS node is `cls12345n004`.

```
root@cls12345n000# ssh cls12345n004
```

2. Prepare the system for manual recovery. Run commands:

```
root@cls12345n000# stop_xyraid cls12345n004_md0-group
root@cls12345n000# clean_xyraid cls12345n004_md0-group
root@cls12345n000# unmanage_xyraid cls12345n004_md0-group
```

3. Manually assemble the RAID device. Run command:

```
root@cls12345n000# mdraid-activate -d cls12345n004_md0-group
```

4. Prepare to capture the output of the `e2fsck` command. Run:

```
root@cls12345n000# screen -L -S fs_check
```

5. Run the `e2fsck` command in read-only.

```
root@cls12345n000# e2fsck -fvntt /dev/md0
```

The command from Step 4 will automatically capture the `e2fsck` output.

6. Upload the `e2fsck` output to HPE Cray Support for review.

 **NOTE:** Depending on the output, HPE Cray Support may need to provide further steps and guidance.

7. When instructed to do so by HPE Cray Support, prepare to capture the output of the next `e2fsck` command, run:

```
root@cls12345n000# screen -L -S fs_check
```

8. Run the `e2fsck` command in read-write.

```
root@cls12345n000# e2fsck -fvptt /dev/md0
```

The command from Step 7 will automatically capture the `e2fsck` output.

9. Upload the `e2fsck` output to HPE Cray Support for review.

 **NOTE:** Depending on the output, HPE Cray Support may need to provide further steps and guidance.

10. When instructed to do so by HPE Cray Support, manually deactivate the RAID, run command:

```
root@cls12345n000# mdraid-deactivate cls12345n004_md0-group
```

11. Revert control of the system to the HA software, run commands:

```
root@cls12345n000# manage_xyraid cls12345n004_md0-group
root@cls12345n000# clean_xyraid cls12345n004_md0-group
root@cls12345n000# start_xyraid cls12345n004_md0-group
```

Support Bundles

Support bundles are collections of event logs from field systems that are collected by Hewlett Packard Enterprise personnel and used to debug many ClusterStor problems. ClusterStor provides a mechanism for collecting support bundles that can be initiated manually or triggered automatically by certain events (for example, Lustre bugs or failover events). Provide these support bundles to Hewlett Packard Enterprise personnel while requesting technical support.

When a problem is encountered, use the `cscli support_bundle` command or try the following tabs in the CSM GUI to trigger a support bundle capture:

- The Health tab displays details of the host or service alerts and notifications to determine the issues.
- The Support tab shows diagnostic information from the storage cluster, including logs and configuration settings.

Support File Overview

The CSM Support tab provides support functionality for collecting diagnostic information, including logs and configuration settings, on an automatic or manual basis. When a Lustre error occurs, the system automatically collects diagnostic information. Alternatively, ClusterStor users can manually collect a diagnostic payload and browse the contents.

The three principal resources for debugging an issue are support bundles, system logs, and GEM logs.

Collection of ClusterStor Data in Support Files

When a Lustre error or a system event (such as failover) occurs, ClusterStor automatically triggers a process to collect system data/diagnostics and to bundle them in support files. The process waits two minutes before collecting the data to ensure that all consequences of the events and errors are logged. Only one collection process is active at a time.

Multiple errors do not trigger the collection of additional data if the current process is still running or within a two-hour window after the current process was triggered. For example, if a Lustre error occurs at 8:00, triggering data to be collected in support file bundle and the same error occurs one (1) hour later at 9:00, ClusterStor does not start a second data collection process related to the later error.

Contents of Support Bundles

Data related to system errors is collected in files, which are packaged together into support bundles. A support bundle is a standard UNIX-compressed file (`tar-gzip`), with files that include:

- System logs for all nodes for the 45-minute period before the error occurred
- List of all cluster nodes and information for each node:
 - Software version
 - Linux kernel and patches
 - Console logs
 - BMC and hardware logs
 - RAID or Zpool information
 - Installed RPMs
 - Lustre targets mounted on the node
 - Power states
 - Resource states
 - Process list
 - Sysrq data
- Current Apache/WSGI logs from the MGS/MDS
- Application state data (MySQL database dump)
- Diagnostic and performance test logs

Automatic vs. Manual Data Collection

When an error occurs, data collection and the bundling of support files is triggered automatically, and ClusterStor users cannot terminate or cancel the operation. The process waits two minutes after the error occurs before starting the data collection process. This ensures that all the error event consequences have been logged before collection begins.

Alternately, a user can manually start data collection and create a support bundle. Unlike the automatic process, a manual data

collection operation can be canceled. When manually starting a support bundle collection or importing a support bundle, the operator is prompted to select the nodes (defaults to “all”) and a window of time (the default is 45 minutes) for logs. After a confirmation dialog appears and is acknowledged, the process begins immediately, there is no 2-minute wait before starting the data collection process.

Use CSM GUI to Collect Support Bundles Manually

The ClusterStor user can manually start system data collection and create support files as described in this procedure. Unlike the automatic process, the manual data collection operation can be canceled.

Procedure

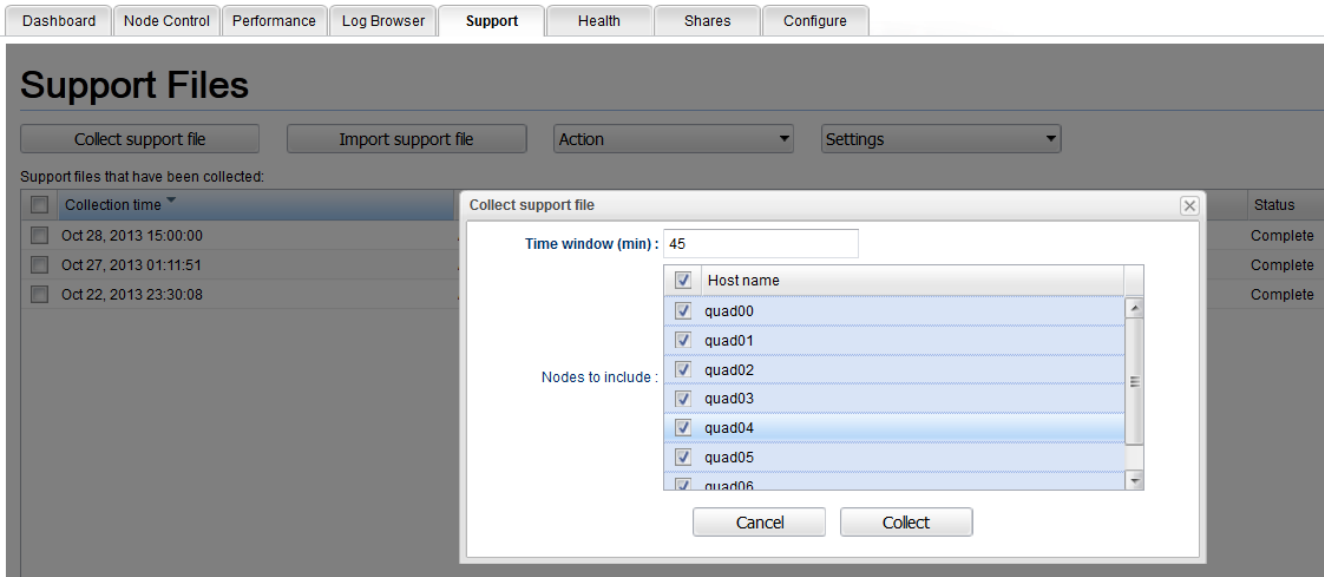
1. Click the CSM Support tab.

The Support Files screen displays.

2. Click the Collect support file button.

The Collect support file dialog window opens and lists all nodes in the cluster.

Figure 15: Support Files Screen: Collect File



3. Specify the data collection parameters for the support file:
 - a. Select the time period to look back for syslog data to be collected. The default is 45 minutes.
 - b. Select the nodes for which data will be collected (the check box next to Hostname selects all nodes).
 - c. Click the Collect button.

The data collection process starts using the specified parameters. While collecting data, it displays in the status field “Still collecting, xx% complete”. When it is complete, it states “Done.” To terminate the operation at any point, click Cancel.

When the operation is complete, the support file—which is a ZIP file containing hundreds of different log files—is created.

Collect Support Bundles Manually via CSCLI

Support bundles can be created manually by using CSCLI commands. For further information, see .

Procedure

1. Log in to the primary MGMT node via `SSH`.

```
$ ssh -l admin primary_MGMT_node
```

2. Collect the support bundle.

- To collect the bundle using the default 45-minute time period: `admin@n000$ cscli support_bundle collect`
- To collect the bundle using a different time period, *minutes*: `admin@n000$ cscli support_bundle collect -t minutes`

For example:

```
admin@n000$ cscli support_bundle collect -t 10
Collecting support bundle: id:2, nodes:all, starting at NOW-10min, time-window:10 minute(s)
```

3. Check the status of the data collection:

```
root@cls12345n000# cscli support_bundle show --bundles
ID | Completion time | Method | Reason | Status
1 | 2020-05-27 10:36:31 | Auto | Failover (STONITH of set(['cls12345n004', 'cls12345n005'])) | Complete
2 | 2020-05-27 11:52:35 | Auto | Failover (STONITH of set(['cls12345n004', 'cls12345n005'])) | Complete
3 | 2020-05-27 12:31:08 | Auto | Failover (STONITH of set(['cls12345n004', 'cls12345n005'])) | Complete
4 | 2020-05-27 13:43:36 | Auto | Failover (STONITH of set(['cls12345n004', 'cls12345n005'])) | Complete
5 | 2020-05-29 07:48:49 | Manual | Requested by 'root' | In progress (18%)
```

"In progress (18%)" indicates bundle collection is ongoing; an export (as shown in the next step) cannot start until the collection has completed.

4. Export the support bundle:

```
admin@n000$ cscli support_bundle export 2
Support_bundle with id 2 saved in file support_bundle_2019-07-05_20-14-36_199745.tgz
```

`cscli support_bundle -h` screen:

```
usage: cscli support_bundle [-h] {collect,show,export,set,delete} ...
Manage support bundles and support bundle settings.
Positional arguments:
  collect      initiate (request) collection of the support bundle on
               specified set of nodes.
  show         support_bundle show command.
  export       export support bundles as Tar.GZ archive (into the
               current folder).
  set          support_bundle set command.
  delete       delete support bundle.
Optional arguments:
  -h, --help   show this help message and exit.
```

Import a Support File

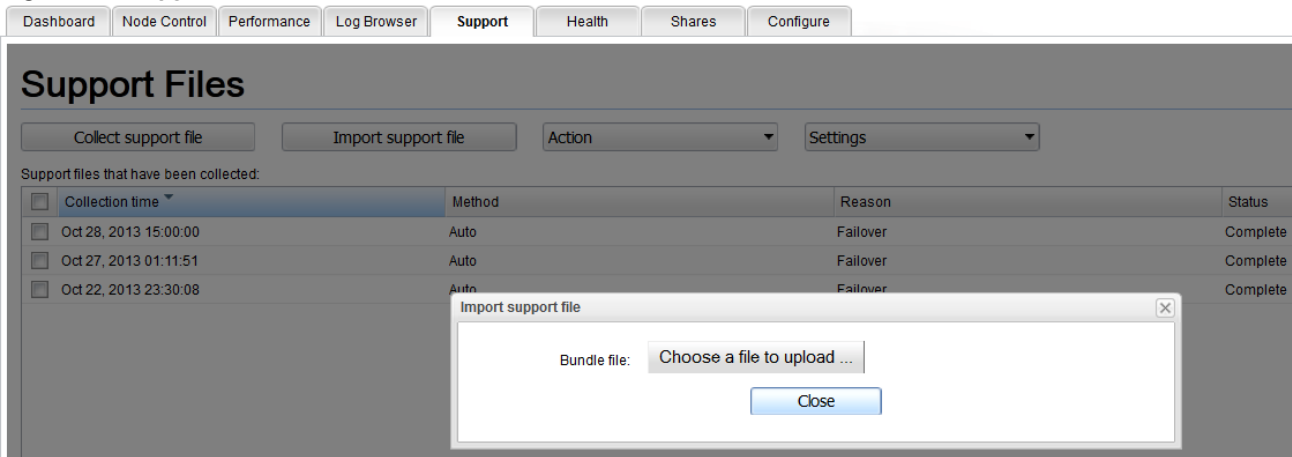
Use the Import feature to upload a single support bundle file into ClusterStor Manager to view its contents.

To import a support bundle file:

Procedure

- 1. Navigate to the Support tab.
The Support Files screen displays.
- 2. Click the Import support file button.
The Import support file window opens.
- 3. Select a support file to upload.
 - a. Click the Choose a file to upload... button.
A list of available support files displays.
 - b. Locate and select the support bundle file to upload.

Figure 16: Support Files Screen: Select File



The selected support bundle will be uploaded.

- 4. Click Close to close the Import support file window.

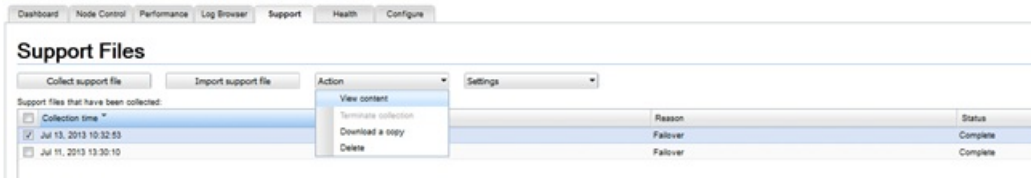
Download a Support File

Use the download feature to save a local copy of the selected support bundle file.

To download a support bundle file:

Procedure

- 1. Navigate to the Support tab.
The Support Files screen displays.
- 2. Select a support bundle file to download.
 - a. Click the checkbox to the left of the collection time of the support bundle file to download.



- b. Click the Action button.
 - c. Select Download a copy from the drop-down menu.
Depending on the local machine and browser, the support bundle file may download automatically to a default local directory. If this is not the case, proceed to the next sub-step.
 - d. Specify where to save the file on the local system.
The support bundle file is downloaded to the local machine.

Delete a Support File

Use the Delete feature to delete a selected support bundle file.

To delete a support bundle file:

Procedure

1. Navigate to the Support tab.
The Support Files screen displays.
2. Select the support bundle file to be deleted.
 - a. Click the checkbox to the left of the collection time of the support bundle file to delete.
 - b. Click the Action button.
 - c. Select Delete from the drop-down menu.A window appears prompting to confirm the deletion. Click the OK button to delete the file.

DashboardNode ControlPerformanceLog BrowserSupportHealthSharesConfigure


Support Files

Collect support fileImport support fileActionSettings

Support files that have been collected:

<input type="checkbox"/> Collection time	Method	Reason	Status
<input type="checkbox"/> Oct 28, 2013 15:00:00	Auto	Fallover	Complete
<input checked="" type="checkbox"/> Oct 27, 2013 01:11:51	Auto	Fallover	Complete
<input type="checkbox"/> Oct 22, 2013 23:30:08	Auto		Complete

Confirm

 Are you sure you want to delete the support file of Oct 27, 2013 01:11:51?

OKCancel

View Support Files

Use the Support tab to view the contents of a support file created either automatically or from manual data collection methods.

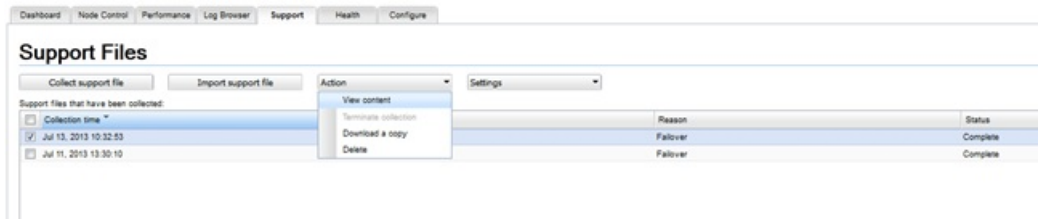
To view ClusterStor support files:

Procedure

1. Click the CSM Support tab.

The Support Files screen is displayed, listing support files that have been collected by either the automatic or manual data collection process.

2. Select a support file to view.
 - a. Check the checkbox in the row containing the desired support bundle; then click the Action button.
 - b. Select View Content.



The Support File content screen opens to display the following tabs:

- Node Information (default) – Lists information for all nodes in the cluster.
- Web Logs – Lists all web logs for the cluster when the support file was created.
- HA Logs – Shows the HA logs for all nodes at the time the support file was created.
- Application State – Shows data tracking the states of the management application, which is being transmitted to the support staff. Do not attempt to use this information, as it may change format from version to version.

Interpret Support Bundles

Support bundles contain two types of logs: systemwide logs that collect data for the entire system, and node-specific logs that collect data for an individual node. Except for `lbug_syslog.csv`, these log files are not intended for use by ClusterStor end users, but may be valuable to Hewlett Packard Enterprise personnel and OEMs to better understand system states and behavior.

Systemwide Logs

Log File Name	File Overview
<code>lbug_syslog.csv</code>	Contains syslog messages, in comma-separated value (CSV) format.
<code>logs/access.log</code>	Contains Apache HTTP access data.
<code>logs/data_tables.sql</code>	Contains a dump of MySQL database tables. The tables describe internal structures used to manage the cluster, the state of cluster resources, information about hardware/software/firmware/network configuration, a FRU inventory, etc. The database dump contains all information required to recreate the system state at the time when the support bundle was created.
<code>logs/error.log</code>	Contains the Apache error log.
<code>logs/wsgi_access.log</code>	Contains records of web service calls made from the CSM.

Node-specific Logs

Log File Name	File Overview
<code>nodes/nodename/conman.log</code>	Contains console data captured by CONsole MANager (<code>Conman</code>), a daemon that provides centralized access to node SOL (serial over LAN, IPMI) or real serial consoles. It also provides logging, broadcasting to several consoles or shared console sessions.
<code>nodes/nodename/crm.log</code>	Contains state data for the RAID and Lustre resources as seen by <code>Pacemaker</code> , an open-source, high-availability resource manager that is suitable for small and large clusters.
<code>nodes/nodename/dmesg.log</code>	Contains a dump of kernel messages collected from the node.
<code>nodes/nodename/fru_dump.yaml</code>	Contains an inventory of FRUs for the enclosure hosting the node (DDICs, PSU, fans, power supplies, etc.). The dump file includes serial numbers for individual FRU equipment, firmware versions, and states such as <code>OK</code> or <code>Failure</code> .
<code>nodes/nodename/lspci.log</code>	Contains a list of PCI devices in a free-form text format generated by the <code>lspci</code> tool. <code>lspci</code> lists PCI devices and their characteristics. <code>lspci</code> can be run in standard or verbose (<code>-vvv</code> option) mode.
<code>nodes/nodename/mdstat.log</code>	Contains state data of the MDRAID arrays, that is, content of the <code>/proc/mdstat</code> file.
<code>nodes/nodename/processes.csv</code>	Contains a list of processes, a snapshot of <code>top</code> , which is a standard monitoring program that reports the top consumers of CPU or memory.
<code>nodes/nodename/sgmap.log</code>	Contains a list of sg devices and specifies for each device the SCSI address, firmware version, and corresponding block devices.
<code>nodes/nodename/software_versions.csv</code>	Contains a list of all installed packages with version information (<code>rpm -qa</code> output).
<code>nodes/nodename/states.csv</code>	Contains miscellaneous state data, including power, memory, uptime, CPU load, and Lustre targets.

support_bundle Command

Introduced in Software Release: 1.3.1 (updated in 3.0.0)

Use the `support_bundle` command and its subcommands to manage support bundles and support bundle settings. When a support bundle is collected, it contains extra information about RAID configuration and local Lustre users (if any are defined). Additional log files are available in the support bundle for MDRAID examine output and Lustre users/groups.

Synopsis

```
$ cscli support_bundle [-h] {collect,set,export,show,delete}
```

Positional Arguments	Description	Release
<code>collect</code>	Initiate (request) collection of the support bundle on specified set of nodes	1.3.1+
<code>show</code>	Support bundle show command	1.3.1+
<code>export</code>	Export support bundles as Tar.GZ archive (into the current folder)	1.3.1+
<code>set</code>	Support bundle set command	1.3.1+
<code>delete</code>	Delete support bundle	3.0.0+

Optional Arguments Description

<code>-h --help</code>	Display the help message and exit
--------------------------	-----------------------------------

support_bundle collect Subcommand

The `support_bundle collect` command is a subcommand of the `support_bundle` command. Use this subcommand to display help, pdsh-style node names, and the time window in minutes.

Synopsis

```
$ cscli support_bundle collect [-h] [-n nodes] [-t minutes]
```

Optional Arguments	Description	Release
<code>-n nodes --nodes nodes</code>	Display pdsh-style node names. The default is all nodes.	1.3.1+
<code>-t minutes --time-window minutes</code>	Display the time window in minutes. The default is 45 minutes.	1.3.1+
<code>-h --help</code>	Display the help message and exit	1.3.1+

support_bundle show Subcommand

Introduced in Software Release: 1.3.1

The `support_bundle show` command is a subcommand of the `support_bundle` command. Use the subcommand to display help, display triggers that initiate automatic bundle collection, display a list of support bundles collected, and/or display the purge limit.

Synopsis

```
$ cscli support_bundle show [-h] (--triggers | --bundles | --purge-limit)
```

Optional Arguments	Description	Release
<code>-t --triggers</code>	Display triggers that initiate automatic bundle collection	1.3.1+
<code>-b --bundles</code>	Display a list of support bundles collected	1.3.1+
<code>-p --purge-limit</code>	Display the purge limit. Free file system space limit in percents; after reaching, 1.3.1+ system will purge old support bundle files.	
<code>-h --help</code>	Display the help message and exit	1.3.1+

support_bundle export Subcommand

Introduced in Software Release: 1.3.1

The `support_bundle export` command is a subcommand of the `support_bundle` command. Use the subcommand to display help and the bundle ID.

Synopsis

```
$ cscli support_bundle export [-h] bundle_id
```

Optional Arguments	Description	Release
<i>bundle_id</i>	ID number of the support bundle file, which can be obtained using the <code>cscli support_bundle show</code> command	1.3.1+
<code>-h</code> <code>--help</code>	Display the help message and exit	1.3.1+

support_bundle set Subcommand

Introduced in Software Release: 1.3.1

The `support_bundle set` command is a subcommand of the `support_bundle` command. Use the subcommand to display help, display purge limit for the support bundle, and/or display the triggers that initiate automatic bundle collection.

Synopsis

```
$ cscli support_bundle set [-h]
(-p purge_limit | --trigger {lbug,Failover})
[--on | --enable | --off | --disable]
```

Optional Arguments	Description	Release
<code>-p purge_limit --purge-limit purge_limit</code>	Set the purge-limit in percentage for support bundles	1.3.1+
<code>-t --trigger {lbug,Failover}, {lbug,Failover}</code>	Display triggers that initiate automatic bundle collection	1.3.1+
<code>--on</code>	Turn the command on	1.3.1+
<code>--enable</code>	Enable the command	1.3.1+
<code>--off</code>	Turn the command off	1.3.1+
<code>--disable</code>	Disable the command	1.3.1+
<code>-h --help</code>	Display the help message and exit	1.3.1+



support_bundle delete Subcommand

Introduced in Software Release: 3.0.0

The `support_bundle delete` command is a subcommand of the `support_bundle` command. Use the subcommand to delete the support bundle identified by the bundle ID.

Synopsis

```
$ cscli support_bundle delete [-h] [--force] bundle_id
```

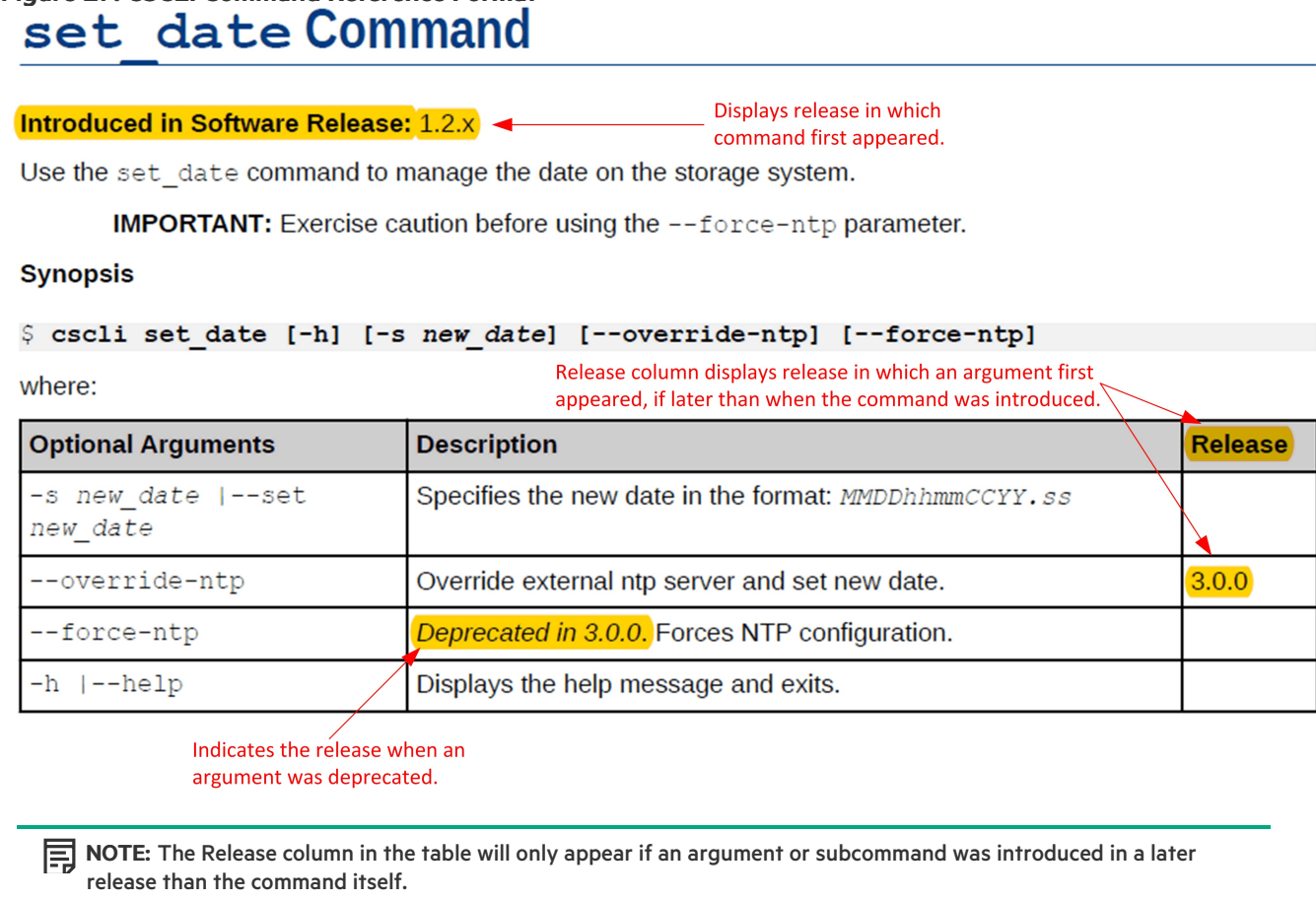
Positional Arguments	Description
<code>bundle_id</code>	ID number of the support bundle file, which can be obtained using the <code>cscli support_bundle show</code> command

Optional Arguments	Description
<code>--force</code>	Force deletion of support bundle in progress
<code>-h</code> <code>--help</code>	Display the help message and exit

CSELI Command Reference Overview

The CSELI Command Reference Guide provides ClusterStor Command Line Interface (CSELI) command syntax and usage information for Cray ClusterStor systems. The Guide provides CSELI command details for software release 4.4 and previous releases. Detailed CSELI commands and subcommands are listed alphabetically in the format shown in the following figure. At the end of the Guide, locate release-specific details for new, modified, and deprecated commands that were introduced each release.

Figure 17: CSELI Command Reference Format



The CSELI feature allows the administrative functions available in the ClusterStor Manager GUI to be run from a Command Line Interface (CLI) using regular SSH clients. CSELI commands for Node Control (for example, power management, start/stop Lustre) along with a selected number of monitoring commands are currently supported.

Since Cray ClusterStor systems and software are in active development, the CSELI documentation is a work-in-progress. As new CSELI commands become available, they will be added to this publication. Notations to deprecated commands are added upon that status change.

CSCLI Command Changes in Release 4.4

The following commands were added, changed, or removed in software release 4.4:

Addition/Change	CSCLI Command	Description	Component
Change	lustre users order set	Additional information for the order and db arguments	Lustre Users

CSLI Reference (Release 4.3)

CSLI Command Changes in Release 4.3

The following commands were added, changed, or removed in software release 4.3:

Addition/Change	CSLI Command	Description	Component
Change	<code>configure_oss</code>	Additional information for the <code>-bfs</code> , <code>-pc</code> , <code>--raid-mode-flash</code> , <code>--raid-strategy</code> , and <code>-s</code> arguments	Node Control

CSCLI Command Changes in Release 4.2

The following commands were added, changed, or removed in software release 4.2:

Addition/Change	CSCLI Command	Description	Component

CSCLI Reference (Release 4.1)

CSCLI Command Changes in Release 4.1

The following commands were added, changed, or removed in software release 4.1:

Addition/Change	CSCLI Command	Description	Component
Added	<code>lustre quota</code>	Manage Lustre quota configuration	Lustre
Added	<code>lustre lnet interfaces</code>	View current connectivity information along with PCIe card to network device mapping	Lustre
Changed	<code>remove_unit</code>	Added new flag: <code>--partially-added</code>	Node Control
Changed	<code>configure_oss</code>	Added new flags: <code>-bfs {yes,no} benchmark-file-system {yes,no}</code> and <code>--partition-count (with examples)</code> . Edited <code>-s {iops,bandwidth,draid}, --raid-strategy {iops,bandwidth,draid}</code> .	Node Control
Changed	<code>configure_mds</code>	Added new flag: <code>-bfs {yes,no} benchmark-file-system {yes,no}</code>	Node Control
Changed	<code>configure_hosts</code>	Added new flags: <code>--role {oss,mds}</code> and <code>-partition-count (with examples)</code>	System Configuration
Changed	<code>lustre_network</code>	Added <code>sm</code> subcommand	Lustre

CSCLI Reference (Release 4.0)

CSCLI Command Changes in Release 4.0

The following commands were added, changed, or removed in software release 4.0:

Addition/Change	CSCLI Command	Description	Component
Added	<code>admins {add,list,remove,modify,reset_password,policy,enable,disable,show,logout}</code>	Manage admin accounts	System Admin
Added	<code>lustre lnet multi-rail</code>	Enable Multi-Rail LNet	Lustre
Added	<code>raid check urc</code>	Manage RAID URC utility	RAID
Changed	<code>configure_oss --raid-strategy</code> renamed to <code>configure_oss --raid-mode-flash</code>	Provide additional RAID layout configuration for flash-based nodes	Node Control
Changed	<code>lustre_network bonding mode set --mode balancealb</code>	Added note to avoid this option when ROCE is enabled	Lustre



GEM CLI Commands

This section describes command line commands provided by the Generic Enclosure Management (GEM) software. While these commands are not typically used during normal operation, they can be used to help resolve issues.

The GEM software controls and monitors the hardware infrastructure and overall system environmental conditions. GEM manages system health, provides power control to power-cycle major subsystems, monitors fans, thermals, power consumption, etc., and offers extensive event capture and logging mechanisms to support file system failover and enable post-failure analysis of hardware components.

Each enclosure and FRU within has an EEPROM containing Vital Product Data (VPD) that is used by GEM to identify and control system components. GEM can parse three (3) main VPD types:

- Midplane
- Power/Cooling Module (PCM)
- Canister

Unified System Management (USM) combines the GEM firmware with firmware from other enclosure components, such as x86 subsystems, primarily BIOS, BMC, and FPGA.

Supported number bases

Numeric parameters passed into CLIs can be in different bases. Decimal is the default. Octal or hexadecimal can be supplied by using a leading code:

Decimal – Plain number

Octal – Leading '0'

Hexadecimal – Leading '0x'

For example, the decimal number 14 would be represented in the following ways:

Decimal – 14

Octal – 016

Hexadecimal – 0xE

Serial port settings

Use the following settings for using HyperTerminal or other serial communications GUI to work with the CLI:

Baud rate (bits per second): 115200	
Data bits:	8
Parity:	None
Stop bits:	1
Flow control:	None

The above settings apply to manually typed commands. If multiple commands are sent via a text file, then the baud rate needs to be reduced for all characters to be processed.

Set the baud rate in the running firmware by issuing:

```
rmon baud 0
```

Change the serial communications GUI settings to: Baud rate (bits/sec):

Baud rate (bits/sec): 9600	
Data bits:	8
Parity:	None
Stop bits:	1
Flow control:	None

To return to the higher baud rate, issue:

```
rmon baud 4
```


The complete set of supported values is:

0 = 9600

1 = 19200

2 = 38400

3 = 57600

4 = 115200

Log In to GEM (GOBI)

Follow these steps to log in to GEM (GOBI) on a specific Lustre node of a Cray ClusterStor E1000 system:

Procedure

1. Log in, using SSH, to the node's GEM (GOBI) interface:

```
root@cls12345n000# ssh gobi@cls12345n004-ipmi
gobi@cls12345n004-ipmi's password
```

2. Enter the password when prompted; then press **Enter** several times.

The GEM (GOBI) command prompt is displayed.

```
*****
GEM Telnet Command Interface. Type 'help' for a list of commands.
*****
0+17:47:31.938 M0 GEM>
```

Run GEM Commands from the Node with `fwdownloader`

Use the `fwdownloader` command to run GEM commands from the node without having to log in to GEM/GOBI. This works across all systems.

Procedure

Enter the `fwdownloader` command.

```
root@even_node# fwdownloader -d 0 -cli GEM_CLI_cmd
root@odd_node# fwdownloader -d 0 -CLI GEM_CLI_cmd

root@cls12345n004# fwdownloader -d 0 -cli report_faults
root@cls12345n005# fwdownloader -d 0 -CLI report_faults
```

 **NOTE:** The lowercase `cli` instructs the command to run from the even node; the uppercase `CLI` instructs the command to run from the odd node.

ddump

Command name	ddump
Command synopsis	Returns a system-wide diagnostic dump
Command description	Calls all commands of the command type 'diagnostic' that do not demand an argument; that is, a simple single-shot diagnostic dump.
Command arguments	None
Command type	Diagnostic
Access level	General

getboardid

Command name	getboardid
Command synopsis	Reports the local board slot ID and HA mode
Command description	Reports the local board slot ID and HA mode in human-readable and machine-readable form.
Command arguments	hex: Returns the slot ID (byte 1) and HA mode (byte 2) in hexadecimal form. If the canister is the master, then the HA mode is set to 0x0. If the canister is the slave, then the mode is 0x1.
Command type	Debug
Access level	General

getvpd

Command name	getvpd
Command synopsis	Retrieves VPD information from all enclosure FRUs
Command description	<p>The <code>getvpd</code> command displays the following enclosure VPD data:</p> <ul style="list-style-type: none">• Enclosure Vendor• Enclosure Product ID• Enclosure WWN• Enclosure Serial Number• Enclosure Part Number• Canister VPD Version• Canister Vendor• Canister Product ID• Canister SAS Address• Canister Serial Number• Canister Part Number• Midplane VPD Version• Midplane Product ID• Midplane Serial Number• Midplane Part Number• PCM VPD Version• PCM Vendor• PCM Product ID• PCM Serial Number• PCM Part Number
Command arguments	<code>getvpd</code> – No additional arguments
Command type	Debug
Access level	General

help

Command name	help
Command synopsis	Displays helpful information about the GEM commands
Command description	Provides a mechanism to discover the available commands and display the command usage information. By default (i.e. no argument supplied), the command only lists the synopsis for those commands with the access level 'general'. The argument all lists the synopsis for all commands, regardless of access level. The argument testing lists the synopsis for all commands that have the 'testing' access level. If the argument matches a command (for example help ddump) then detailed help for the specified command displays instead.
Command arguments	One optional argument – see preceding description.
Command type	Control
Access level	General

ipmi_power

Command name	ipmi_power
Command synopsis	Performs safe canister-level power control using chassis commands to the BMC
Command description	This command allows the user to request a canister-level shutdown through the BMC. The benefit of using this command is to cleanly shut down the x86 subsystem using ACPI.
Command arguments	2 "soft" – Orchestrated shutdown of x86 complex. 3 "off" – Immediate shutdown of x86 complex. 4 "cycle" – Canister power cycle. 5 "reset" – Canister reset. 6 "on" – Wake x86 complex from standby/soft-off.
Command type	Control
Access level	General Access

logdump

Command name	logdump
Command synopsis	Displays logged messages
Command description	Provides a mechanism to output logging information.
Command arguments	<p>6 optional arguments:</p> <p>Argument 1 specifies the area of memory from which to retrieve log messages from. 'r' = RAM, 'n' = non-volatile.</p> <p>Argument 2 specifies the order of the log messages. "old" = oldest first, "new" = newest first.</p> <p>Argument 3 limits the number of logged messages displayed to <i>n</i>. Set to zero (0) or omit the argument to display all logged messages.</p> <p>Argument 4 controls the generation of a <i>timestamp</i> field in the log dump messages. Set to 1 for enable; 0 for disable.</p> <p>Argument 5 controls the generation of a <i>subsystem name</i> field in the log dump messages. Set to 1 for enable; 0 for disable.</p> <p>Argument 6 controls the generation of a <i>service name</i> field in the log dump messages. Set to 1 for enable; 0 for disable.</p> <p>The default (for omitted command arguments) displays all logged messages from RAM, newest first, with all message fields enabled.</p>
Command type	Diagnostic
Access level	General

report_faults

Command name	report_faults
Command synopsis	Reports all system-wide faults
Command description	Outputs all known faults, collected from each GEM service
Command arguments	None
Command type	Diagnostic
Access level	General

ver

Command name	ver
Command synopsis	Displays version information
Command description	Displays version numbers and information for the components in the local canister, midplane and PCMs
Command arguments	None
Command type	Diagnostic
Access level	General

Websites

General websites

Single Point of Connectivity Knowledge (SPOCK) Storage compatibility matrix

<https://www.hpe.com/storage/spock>

Storage white papers and analyst reports

<https://www.hpe.com/storage/whitepapers>

For additional websites, see [Support and other resources](#).

Accessing Hewlett Packard Enterprise Support

- For live assistance, go to the Contact Hewlett Packard Enterprise Worldwide website:

<https://www.hpe.com/info/assistance>

- To access documentation and support services, go to the Hewlett Packard Enterprise Support Center website:

<https://www.hpe.com/support/hpesc>

Information to collect

- Technical support registration number (if applicable)
- Product name, model or version, and serial number
- Operating system name and version
- Firmware version
- Error messages
- Product-specific reports and logs
- Add-on products or components
- Third-party products or components

Accessing updates

- Some software products provide a mechanism for accessing software updates through the product interface. Review your product documentation to identify the recommended software update method.

- To download product updates:

Hewlett Packard Enterprise Support Center

<https://www.hpe.com/support/hpesc>

Hewlett Packard Enterprise Support Center: Software downloads

<https://www.hpe.com/support/downloads>

My HPE Software Center

<https://www.hpe.com/software/hpesoftwarecenter>

- To subscribe to eNewsletters and alerts:

<https://www.hpe.com/support/e-updates>

- To view and update your entitlements, and to link your contracts and warranties with your profile, go to the Hewlett Packard Enterprise Support Center More Information on Access to Support Materials page:

<https://www.hpe.com/support/AccessToSupportMaterials>

ⓘ IMPORTANT:

Access to some updates might require product entitlement when accessed through the Hewlett Packard Enterprise Support Center. You must have an HPE Passport set up with relevant entitlements.

Remote support

Remote support is available with supported devices as part of your warranty or contractual support agreement. It provides intelligent event diagnosis, and automatic, secure submission of hardware event notifications to Hewlett Packard Enterprise, which initiates a fast and accurate resolution based on the service level of your product. Hewlett Packard Enterprise strongly recommends that you register your device for remote support.

If your product includes additional remote support details, use search to locate that information.

HPE Get Connected

<https://www.hpe.com/services/getconnected>

HPE Pointnext Tech Care

<https://www.hpe.com/services/techcare>

HPE Complete Care

<https://www.hpe.com/services/completecure>

Warranty information

To view the warranty information for your product, see the links provided below:

HPE ProLiant and IA-32 Servers and Options

<https://www.hpe.com/support/ProLiantServers-Warranties>

HPE Enterprise and Cloudline Servers

<https://www.hpe.com/support/EnterpriseServers-Warranties>

HPE Storage Products

<https://www.hpe.com/support/Storage-Warranties>

HPE Networking Products

<https://www.hpe.com/support/Networking-Warranties>

Regulatory information

To view the regulatory information for your product, view the Safety and Compliance Information for Server, Storage, Power, Networking, and Rack Products, available at the Hewlett Packard Enterprise Support Center:

<https://www.hpe.com/support/Safety-Compliance-EnterpriseProducts>

Additional regulatory information

Hewlett Packard Enterprise is committed to providing our customers with information about the chemical substances in our products as needed to comply with legal requirements such as REACH (Regulation EC No 1907/2006 of the European Parliament and the Council). A chemical information report for this product can be found at:

<https://www.hpe.com/info/reach>

For Hewlett Packard Enterprise product environmental and safety information and compliance data, including RoHS and REACH, see:

<https://www.hpe.com/info/ecodata>

For Hewlett Packard Enterprise environmental information, including company programs, product recycling, and energy efficiency, see:

<https://www.hpe.com/info/environment>

Documentation feedback

Hewlett Packard Enterprise is committed to providing documentation that meets your needs. To help us improve the documentation, use the Feedback button and icons (located at the bottom of an opened document) on the Hewlett Packard Enterprise Support Center portal (<https://www.hpe.com/support/hpesc>) to send any errors, suggestions, or comments. All document information is captured by the process.