Introduction to OSG Storage

Suchandra Thapa
Computation Institute
University of Chicago

Overview

- Storage in OSG
  - Storage for VDT
  - Storage Extension
  - Support
  - Certification
  - Community tools
  - Documentation
  - Support
  - GIP & Storage Element
- Bestman
  - Overview
  - “full-mode”/“gateway”
  - “gateway”/Xrootd installation
- dCache
  - Overview
  - Monitoring tools
  - Opportunistic storage
  - Gratia Probes

Why do we care about storage?

- Sites are growing to the point where putting a bunch of disks and nfs exporting it no longer works
  - Too much space needed to make it economical
    - LHC projects will generate several petabytes of data a year
    - Typical Tier 2 sites have on the order of 500TB – 1PB of storage space
  - I/O contention and loads will bring down a normal NFS server

Why care about storage (the sequel)?

- Would like additional information on storage
  - Let users easily get space information (available/ used) without checking sites individually
- Advertise space availability and allow disk space to be reserved
  - If your site is sponsored by a VO like ATLAS, you don't want to have another VO come in and use up all the space

Solutions for OSG Storage

OSG Storage for VDT is a well integrated distributed project between Wisconsin and Fermilab. Activities:

- Packaging storage software for VDT
  - Srm/dCache
  - BeStMan
  - BeStMan-gateway/Xrootd (NEW)
- Simplify configuration/installation for OSG
- Help VOs to use storage on OSG sites
- Develop and run validation tests
- Develop/maintain/package accounting and monitoring tools
- Support/test/package community tools
- Provide users and admins support
- Perform troubleshooting and debugging
- OSG liaison to storage developer groups
- Educate OSG community about storage, provide documentation
- Participate in grid schools organized by OSG

OSG Storage Extension Project

- Tightly related to OSG Storage for VDT
- Goals for this year:
  - Develop software to match storage attributes of a job to a Storage Element
  - Develop web interface
    - for SE discovery and authorization check
    - for creating and managing space reservations
  - Work on consolidation of monitoring functionalities of Storage Elements into one interface.
  - Provide electronic means for sites to announce SE maintenance period

Support Challenges

- Complicated, highly distributed services
- Huge variety of configuration options (software and hardware)
- Widely diverse utilization patterns
- dCache is known for poor error diagnostics, exception handling and propagation
- We do not enough experience with Xrootd
- Lack of monitoring/diagnostic tools
- Support team does not have access to the service. Support personnel
  - Often are not authorized to use the service as user
  - Can not access site logs and configuration
  - Often can not access storage monitoring pages on the site

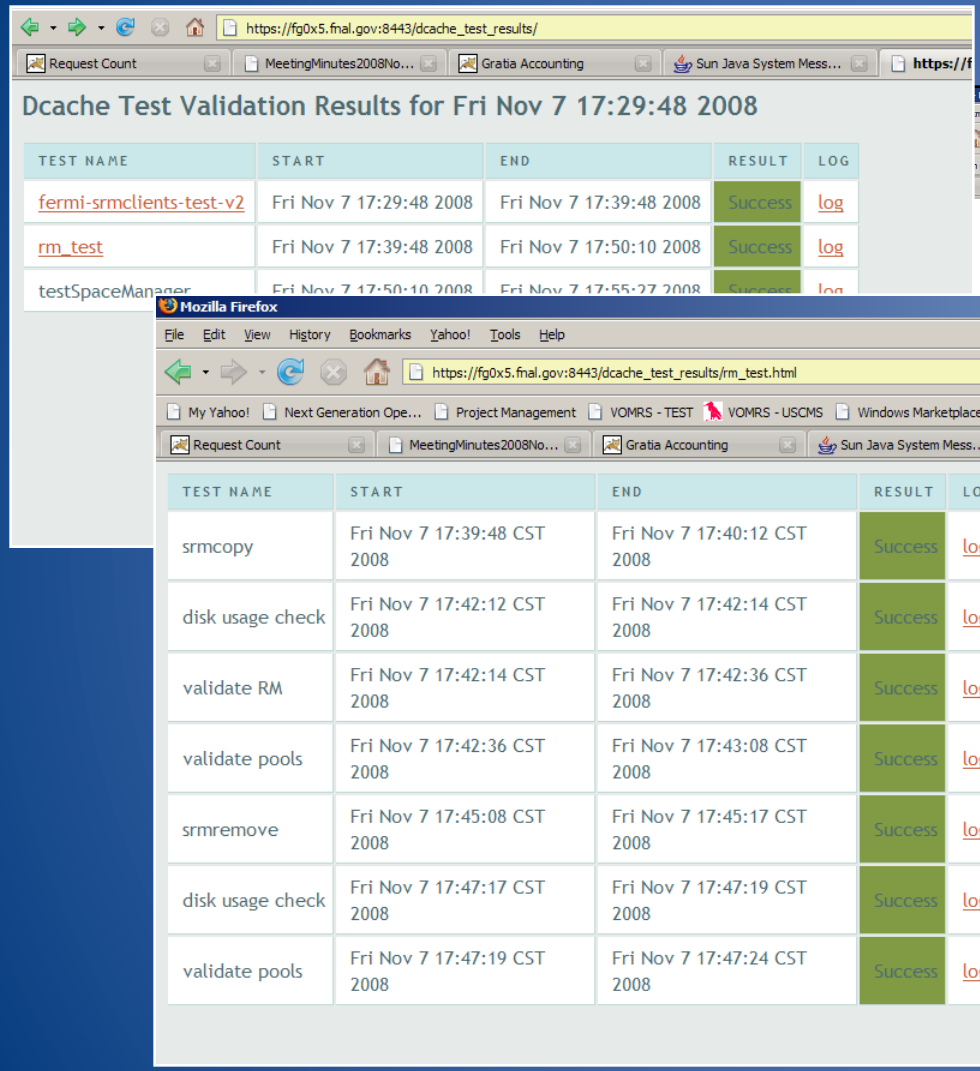
We would like to ask storage administrators for cooperation in:

- Notifying us about the reoccurring problems
- Provide us access to log files, configuration files

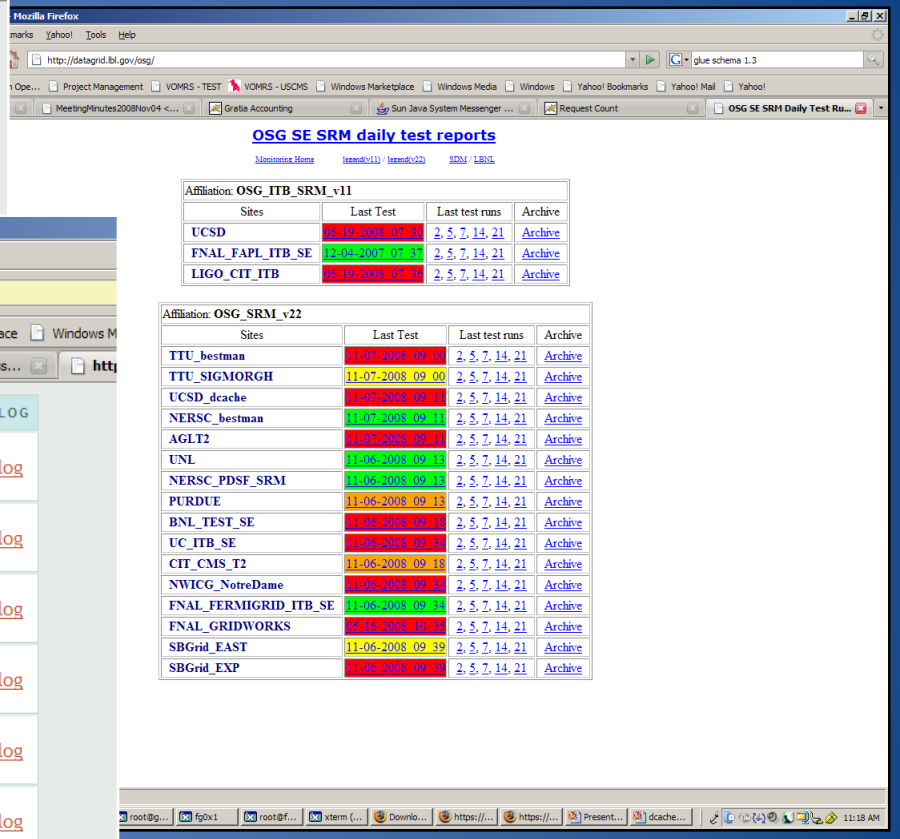
Certification

- Maintain test stands
  - 6 nodes test stand for dCache
  - Planning to have 5 nodes test stand for BeStMan-gateway/xrootd
- Develop/run validation test suites before software is released to VDT
  - dCache test suite covers:
    - all srm-fermi-client commands
    - data replication
    - space management
    - load tests
  - BeSTMan testing provided by LBL
    - Site registration, daily test results - http://datagrid.lbl.gov/osg
    - Site could run tests with srm-tester-2 – instructions at https://twiki.grid.iu.edu/twiki/bin/view/Storage/BeStMan
  - BeStMan-gateway/Xrootd covers:
    - all supported srm-lbl-client /srm-fermi-client commands – work in progress at Fermilab

Test Suites Results



Fermilab Validation Test Suite Results



LBL SRM Tester Results

Community-contributed toolkit

- Goal is to provide a boost in effectiveness and efficiency ofoperating the deployed storage
- Download from http://datagrid.ucsd.edu/toolkit
- Is packaged as rpms
- Collected/packaged by A. Rana

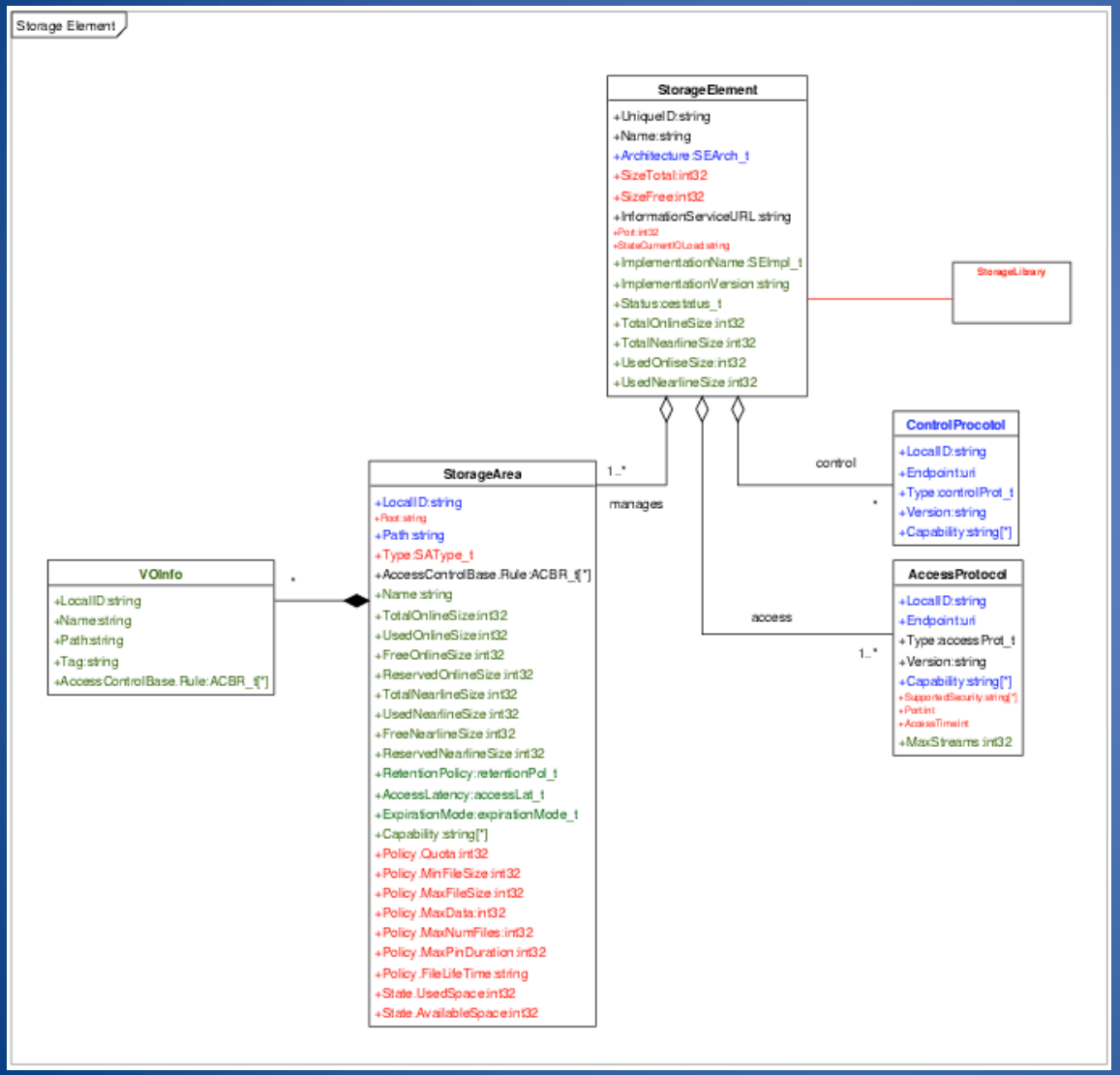
	Utility Name	R	E	C	A
RPM 1	osg_dc_srm_space_reclaimer	Y	Y		Y
RPM 2	osg_dc_analyze_disk_usage_of_all_pools		Y		
	osg_dc_analyze_disk_usage_of_pnfs_dir				
	osg_dc_analyze_num_replicas_disk_pnfs_of_pnfs_dir		Y		
	osg_dc_cancel_stuck_restores				
	osg_dc_cleanup_broken_transfers				
	osg_dc_cleanup_disk_files_not_in_pnfs	Y	Y		
	osg_dc_compare_checksum_disk_pnfs_for_all_files		Y		
	osg_dc_find_files_in_pnfs_not_on_disk		Y		
	osg_dc_find_num_replicas_for_all_files		Y		
	osg_dc_find_path_for_pnfsid				
	osg_dc_find_pnfsid_for_path				
	osg_dc_find_pools_with_file				
	osg_dc_find_real_size_of_a_pnfsid				
	osg_dc_get_html_dump_of_fs_usage_of_pnfs_dir		Y		
	osg_dc_move_files_from_full_pool_to_another_pool	Y	Y		
	osg_dc_move_pnfsids_across_preferred_pools	Y	Y		
	osg_dc_pnfs_register_all_disk_files_in_pools				
	osg_dc_query_all_transfer_rates				
	osg_dc_query_srm_transfer_rate				
	osg_dc_verify_discrepancy_disk_usage_of_all_pools				
	osg_dc_write_protect_almost_full_pools				Y
RPM 3	osg_dc_PFM	Y	Y		
	osg_dc_retire_a_pool	Y	Y		
RPM 4	osg_dc_automatic_restart_offline_pool				Y
	osg_dc_pool_usage_nagios_plugin				
	osg_dc_analyze_local_disk_usage_of_pool				

Storage Documentation

- Revised documentation
- Main Page:

  https://twiki.grid.iu.edu/twiki/bin/view/Documentation/WebHome
- Useful links under Storage Element administrators:
  - Opportunistic Storage/Space Reservation
  - Opportunistic Storage Model for USCMS
  - Gratia Storage Probes
  - Tools, Tips, FAQs
  - dCache Installation/references
  - BeStMan references

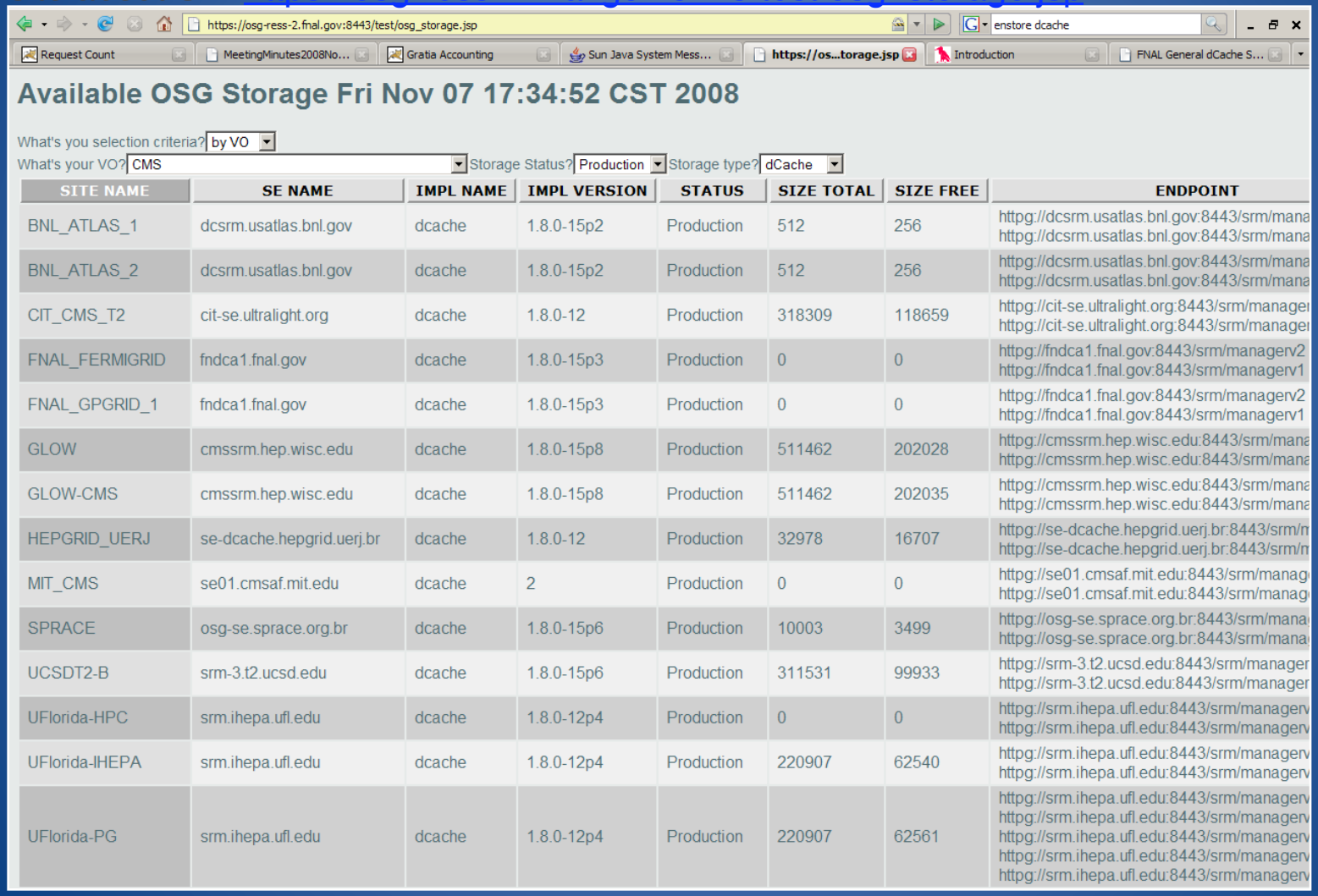
Storage Element Discovery



- Glue Schema 1.3
  - ControlProtocol
    - SRM
  - AccessProtocol
    - gsiftp
  - Storage Area
    - Groups of Pools
    - VOInfo
      - Path
- GIP/CeMon
  - Collects information on CE
  - Sends BDII/ReSS information
  - ReSS Information is used for job matching

Storage Discovery Example

You can find what sites support your VO or what information is advertised by a site about SE: https://osg-ress-2.fnal.gov:8443/test/osg_storage.jsp



https://osg-ress-2.fnal.gov:8443/test/osg_storage.jsp

Available OSG Storage Fri Nov 07 17:34:52 CST 2008

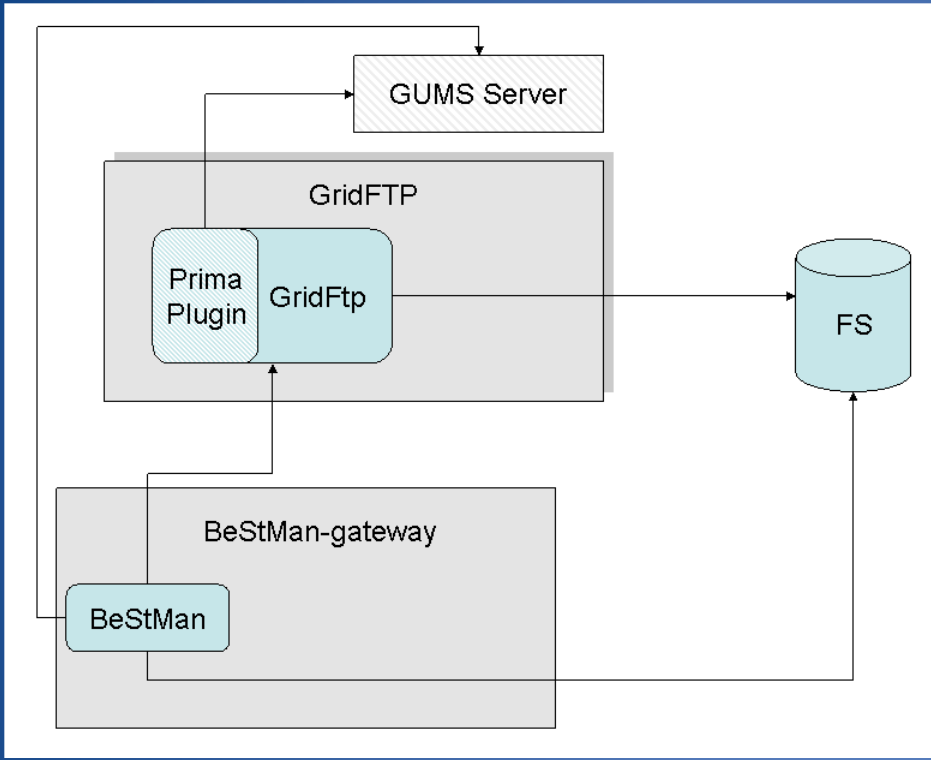
What's you selection criteria? by VO
What's your VO? CMS Storage Status? Production Storage type? dCache

SITE NAME	SE NAME	IMPL NAME	IMPL VERSION	STATUS	SIZE TOTAL	SIZE FREE	ENDPOINT
BNL_ATLAS_1	dcsrm.usatlas.bnl.gov	dcache	1.8.0-15p2	Production	512	256	httpg://dcsrm.usatlas.bnl.gov:8443/srm/mana httpg://dcsrm.usatlas.bnl.gov:8443/srm/mana
BNL_ATLAS_2	dcsrm.usatlas.bnl.gov	dcache	1.8.0-15p2	Production	512	256	httpg://dcsrm.usatlas.bnl.gov:8443/srm/mana httpg://dcsrm.usatlas.bnl.gov:8443/srm/mana
CIT_CMS_T2	cit-se.ultralight.org	dcache	1.8.0-12	Production	318309	118659	httpg://cit-se.ultralight.org:8443/srm/manager httpg://cit-se.ultralight.org:8443/srm/manager
FNAL_FERMIGRID	fndca1.fnal.gov	dcache	1.8.0-15p3	Production	0	0	httpg://fndca1.fnal.gov:8443/srm/managerv2 httpg://fndca1.fnal.gov:8443/srm/managerv1
FNAL_GPGRID_1	fndca1.fnal.gov	dcache	1.8.0-15p3	Production	0	0	httpg://fndca1.fnal.gov:8443/srm/managerv2 httpg://fndca1.fnal.gov:8443/srm/managerv1
GLOW	cmssrm.hep.wisc.edu	dcache	1.8.0-15p8	Production	511462	202028	httpg://cmssrm.hep.wisc.edu:8443/srm/mana httpg://cmssrm.hep.wisc.edu:8443/srm/mana
GLOW-CMS	cmssrm.hep.wisc.edu	dcache	1.8.0-15p8	Production	511462	202035	httpg://cmssrm.hep.wisc.edu:8443/srm/mana httpg://cmssrm.hep.wisc.edu:8443/srm/mana
HEPGRID_UERJ	se-dcache.hepgrid.uerj.br	dcache	1.8.0-12	Production	32978	16707	httpg://se-dcache.hepgrid.uerj.br:8443/srm/m httpg://se-dcache.hepgrid.uerj.br:8443/srm/m
MIT_CMS	se01.cmsaf.mit.edu	dcache	2	Production	0	0	httpg://se01.cmsaf.mit.edu:8443/srm/manag httpg://se01.cmsaf.mit.edu:8443/srm/manag
SPRACE	osg-se.sprace.org.br	dcache	1.8.0-15p6	Production	10003	3499	httpg://osg-se.sprace.org.br:8443/srm/mana httpg://osg-se.sprace.org.br:8443/srm/mana
UCSDT2-B	srm-3.t2.ucsd.edu	dcache	1.8.0-15p6	Production	311531	99933	httpg://srm-3.t2.ucsd.edu:8443/srm/manager httpg://srm-3.t2.ucsd.edu:8443/srm/manager
UFlorida-HPC	srm.ihepa.ufl.edu	dcache	1.8.0-12p4	Production	0	0	httpg://srm.ihepa.ufl.edu:8443/srm/managerv httpg://srm.ihepa.ufl.edu:8443/srm/managerv
UFlorida-IHEPA	srm.ihepa.ufl.edu	dcache	1.8.0-12p4	Production	220907	62540	httpg://srm.ihepa.ufl.edu:8443/srm/managerv httpg://srm.ihepa.ufl.edu:8443/srm/managerv
UFlorida-PG	srm.ihepa.ufl.edu	dcache	1.8.0-12p4	Production	220907	62561	httpg://srm.ihepa.ufl.edu:8443/srm/managerv httpg://srm.ihepa.ufl.edu:8443/srm/managerv httpg://srm.ihepa.ufl.edu:8443/srm/managerv httpg://srm.ihepa.ufl.edu:8443/srm/managerv httpg://srm.ihepa.ufl.edu:8443/srm/managerv

Berkley Storage Management (BeStMan)

- What is BeStMan?
  - Developed in LBNL by Scientific Data Management Research Group
  - Full implementation of SRM v2.2 for disk based storage systems and mass storage systems
  - Supports transfer services:
    - GsiFtp
    - Ftp
    - Http
    - Https
  - Supports multiple transfer servers
  - GSI security with either grid-mapfile or GUMS server
- Who would benefit from BeStMan?
  - Sites with limited hardware resources
  - Sites with limited admin effort
- BeStMan could be used in two modes:
  - Full mode
  - Gateway mode

BeStMan-gateway

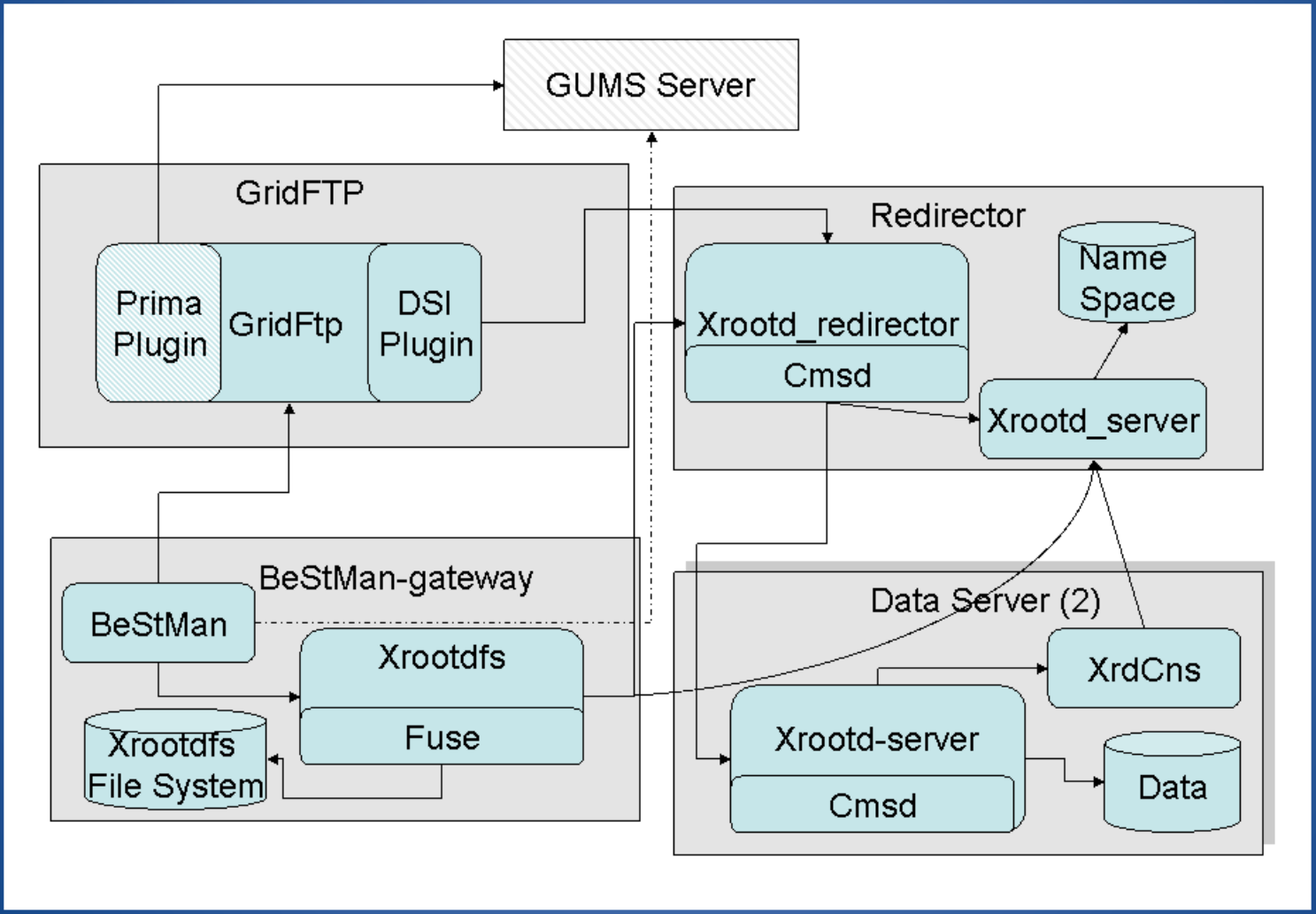


- Generic SRM v2.2 load balancing frontend for GridFTP servers
- Light-weight implementation of SRM v2.2 for POSIX file systems
  - srmPing
  - srmLs
  - srmRm
  - srmMkdir
  - srmRmdir
  - srmPrepareToPut (Status, PutDone)
  - srmPrepareToGet(Status,ReleaseFiles)
- Designed to work with any Posix-like file systems
  - NFS, GPFS, GFS, NGFS, PNFS, HFS+, PVFS, AFS, Lustre, XrootdFS, Hadoop
- Doesn't support queuing or disk space management

BeStMan-gateway/Xrootd

- Xrootd - is designed to provide POSIX-like access to files and their enclosing directory namespace
- BeStMan-gateway needs the following additional components to work with Xrootd
  - FUSE - File System in User Space http://fuse.sourceforge.net
  - XrootdFS - implements a Posix filesystem for an Xrootd storage cluster http://wt2.slac.stanford.edu/xrootdfs/xrootdfs.html
- GridFtp needs Data Storage Interface (DSI) module in order to work with Xrootd storage

BeStMan-gateway/Xrootd Architecture



Before Installing BeStMan-gateway/Xrootd

- How many nodes could be used for storage?
  - Minimum number of nodes is 3:
  - BeStMan, XroodFS, fuse, GridFtp
  - Xrootd redirector
  - Xrootd data server node
- Chose authorization mechanism do you prefer?
  - GUMS
  - gridmap-file
- Do you need to support static space tokens?
- Select name for Mount Point for XrootdFS on BeStMan node
- Decide how to partition storage areas on Xrootd redirector and data server nodes

XrootdFs Installation (I)

- Install fuse using “yum install” or rpms (e.g http://rpmfind.net/linux/rpm2html/search.php?query=fuse)
  - ***fuse-2.7.3-1***
  - ***fuse-libs-2.7.3-1***
  - ***kernel-module-fuse-2.6.9-78.0.1.EL-2.7.3-1***
- Install and configure XrootdFS

pacman -get http://vdt.cs.wisc.edu/vdt_XXX_cache/XrootdFS
***$VDT_LOCATION/vdt/setup/configure_xrootdfs ***
***--user <user> ***
***--cache <mount-point> ***
***--xrdr-host <hostname> ***
--xrdr-storage-path <path>

BeStMan-gateway Installation (II)

- Install and configure BestMan

```
pacman -get OSG:Bestman
$VDT_LOCATION/vdt/setup/configure_bestman --server y \
 --user <user> \
 --cert <service_cert> \
 --key  <service_key> \
 --http-port <public_port> \
 --https-port <secured_port> \
 --gums-host <GUMS hostname> \
 --gums-port <GUMS port number> \
 --gums-dn <Client DN for GUMS interface> \
 --use-xrootd  \
 --with-tokens-list  "<TOKEN_1_NAME>[desc:<TOKEN_1_DESC]
     [TOKEN_1_SIZE_GB];TOKEN_2_NAME[desc:TOKEN_2_DESC][TOKEN_2_SIZE]"   \
 --with-transfer-servers <GridFTP server list>
```

- Modify /etc/sudoers

```
Cmnd_Alias SRM_CMD = /bin/rm, /bin/mkdir, /bin/rmdir, /bin/mv, /bin/ls
Runas_Alias SRM_USR = ALL, !root
<user_name> ALL=(SRM_USR) NOPASSWD: SRM_CMD
```

GridFtp Installation (II)

- Install and configure GridFtp

pacman -get OSG:Xrootd-GridFTP

***$VDT_LOCATION/vdt/setup/configure_gridftp --use-xrootd ***

***--xrootd-host <hostname> ***

***--xrootd-mount-point <mount_point> ***

--xrootd-storage-path < path>

Xrootd-Redirector Installation (IV)

- Install and configure Xrootd redirector

```
pacman -get OSG:Xrootd
$VDT_LOCATION/vdt/setup/configure_xrootd \
--server y \
--this-is-xrdr \
--user <user> \
--xrdr-storage-path <path> \
--xrdr-storage-cache <cache> \
--with-tokens-list "<TOKEN_1_NAME>[desc:<TOKEN_1_DESC]
    [TOKEN_1_SIZE_GB];TOKEN_2_NAME[desc:TOKEN_2_DESC][TOKEN_2_SIZE]"   \
--public-cache-size <PUBLIC_SPACE_SIZE>
```

Xrootd Data Server Installation (V)

- Install and configure Xrootd data server

```
pacman -get OSG:rootd
$VDT_LOCATION/vdt/setup/configure_xrootd \
--server y \
--user <user> \
--xrdr-host  <hostname> \
--xrdr-storage-path <path> \
 --xrdr-storage-cache <cache> \
 --with-tokens-list "<TOKEN_1_NAME>[desc:<TOKEN_1_DESC]
     [TOKEN_1_SIZE_GB];TOKEN_2_NAME[desc:TOKEN_2_DESC][TOKEN_2_SIZE]"   \
--public-cache-size <PUBLIC_SPACE_SIZE>
```

Start/stop order for BeStMan-gateway/Xrootd

- Use vdt-control –on/--off mechanism
- Start severs in the following order (use reverse order to stop):
  - Xrootd redirector
  - Xrootd data server
  - GridFtp
  - XrootdFs/Bestman-gateway
- Test

srm-ping srm://<BeStMan_host>:8443/srm/v2/server

srm-copy file:////tmp/test srm://<BeStMan_host>.gov:8443/srm/v2/server\?SFN=<MOUNT_POINT>/test -spacetoken <TOKEN_1_NAME>

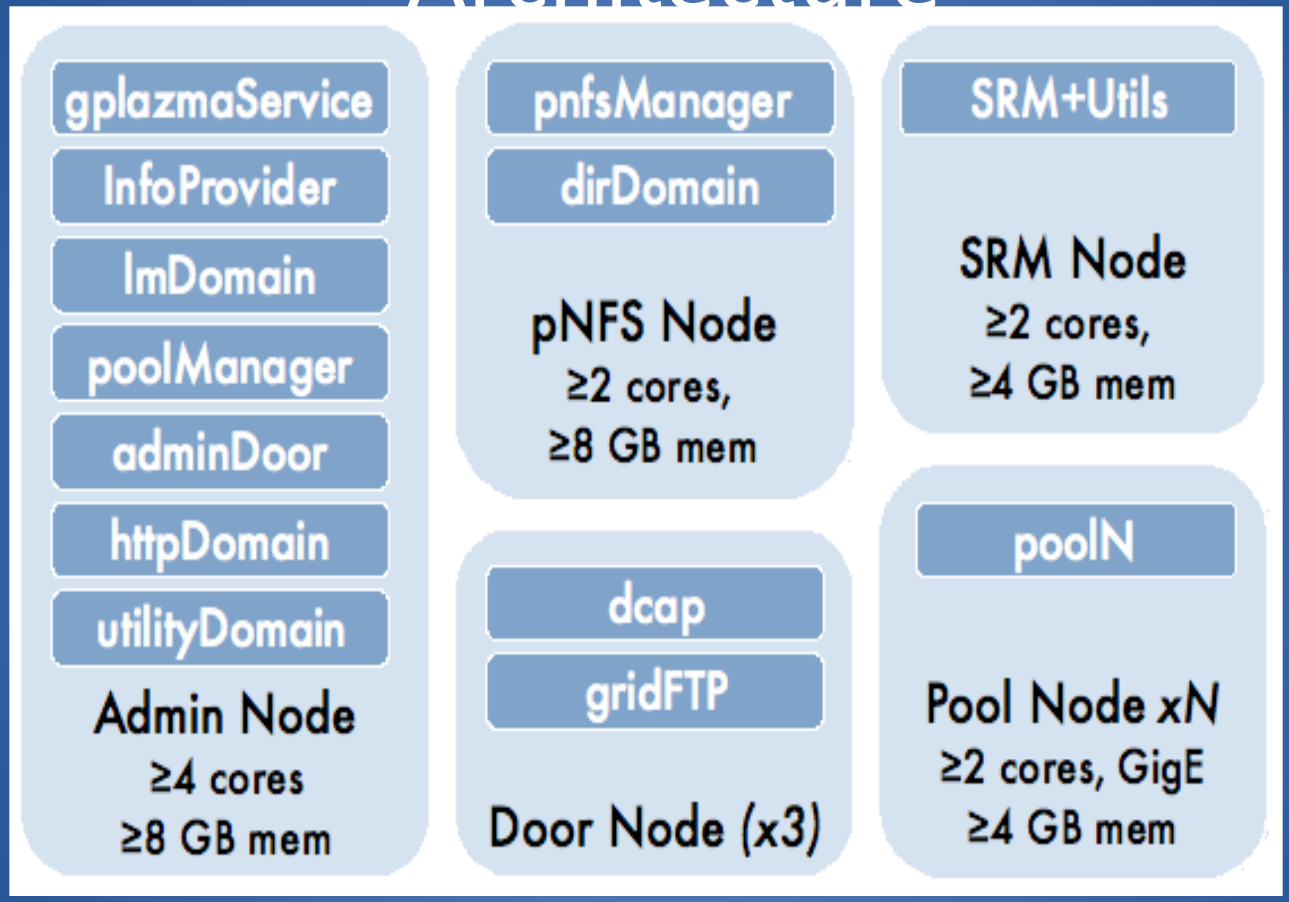
BeStMan in VDT

- BeStMan is one of data storage solutions supported by OSG. BeStMan-gateway, BeStMan-gateway/Xrootd requested by ATLAS are just released in VDT
- Current version of software
  - BeStMan - 2.2.1.2.e1
  - XrootdFS - 2.2.1.1
  - GridFTP-Xrootd ,xrootd-dsi-20080828-1632
  - Prima 0.7.1
  - Xrootd - 20080828-1632
- VDT configuration script tailored to set up BeStMan "full mode"/ gateway for Tier-2/Tier-3
- BeStMan srm-clients are distributed as a part of VDT client cache
  - Fermi client
  - LBNL client
  - LCG-utils
- There are several installations of BeStMan on OSG sites

dCache Main Features

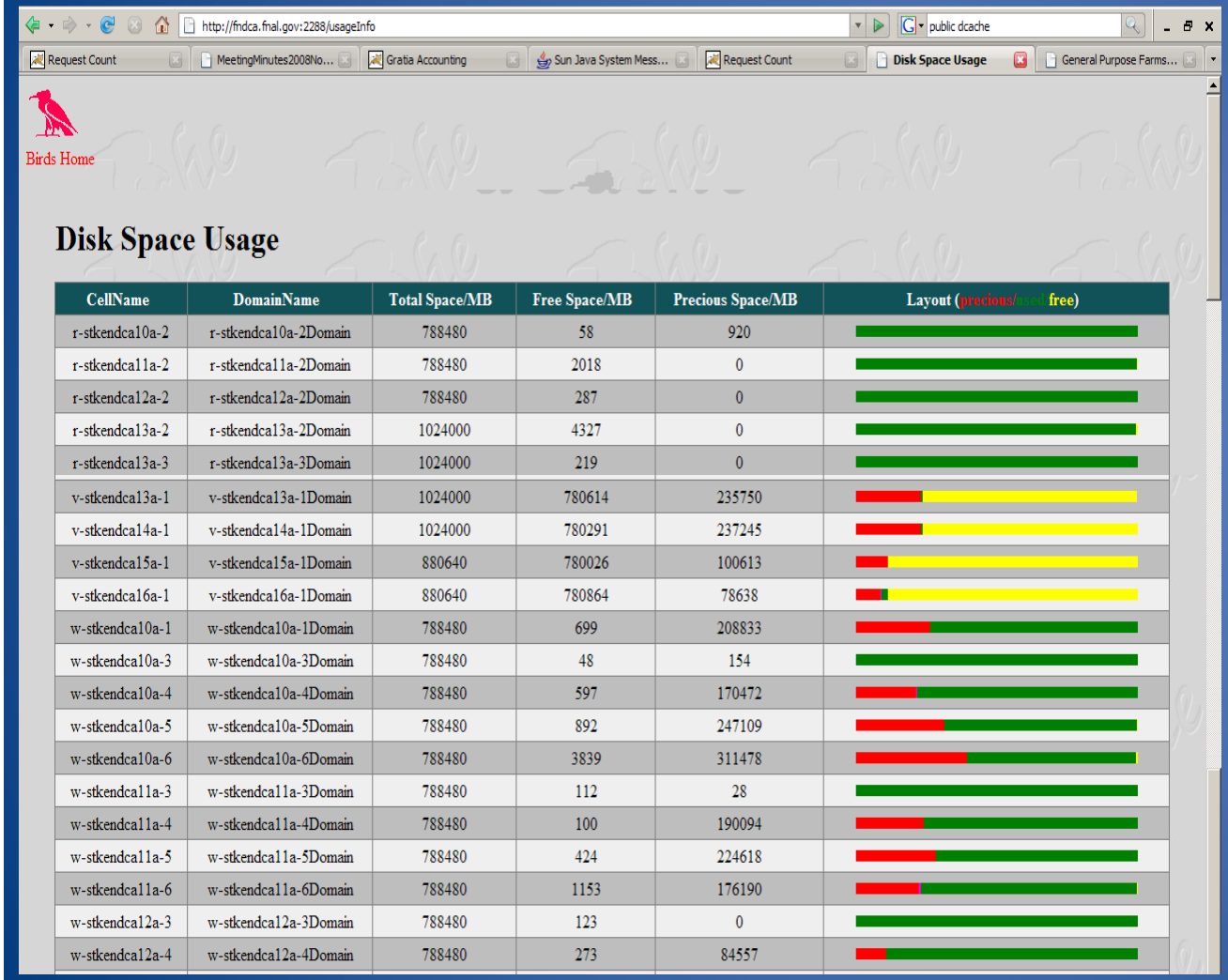
- nfs-mountable namespace
- Multiple access protocols
  - dcap (posix io), gsidcap
  - xrootd (posix io)
  - gsiftp (multiple channels)
- Replica Manager
  - Set min/max number of replicas
- Role-based authorization
  - Selection of authorization mechanisms
- Opportunistic storage
- Billing
- Admin interface
  - ssh, jython
- InformationProvider (not in production - yet)
  - SRM and gsiftp described in glue schema
- Platform, fs independent (Java)
  - 32 and 64-bit linux, solaris; ext3, xfs, zfs

dCache OSG Tier-2 site Architecture



Slide courtesy of Ted Hesselroth (from presentation: "Installing and Using SRM-dCache")

dCache Monitoring



Disk Space Usage

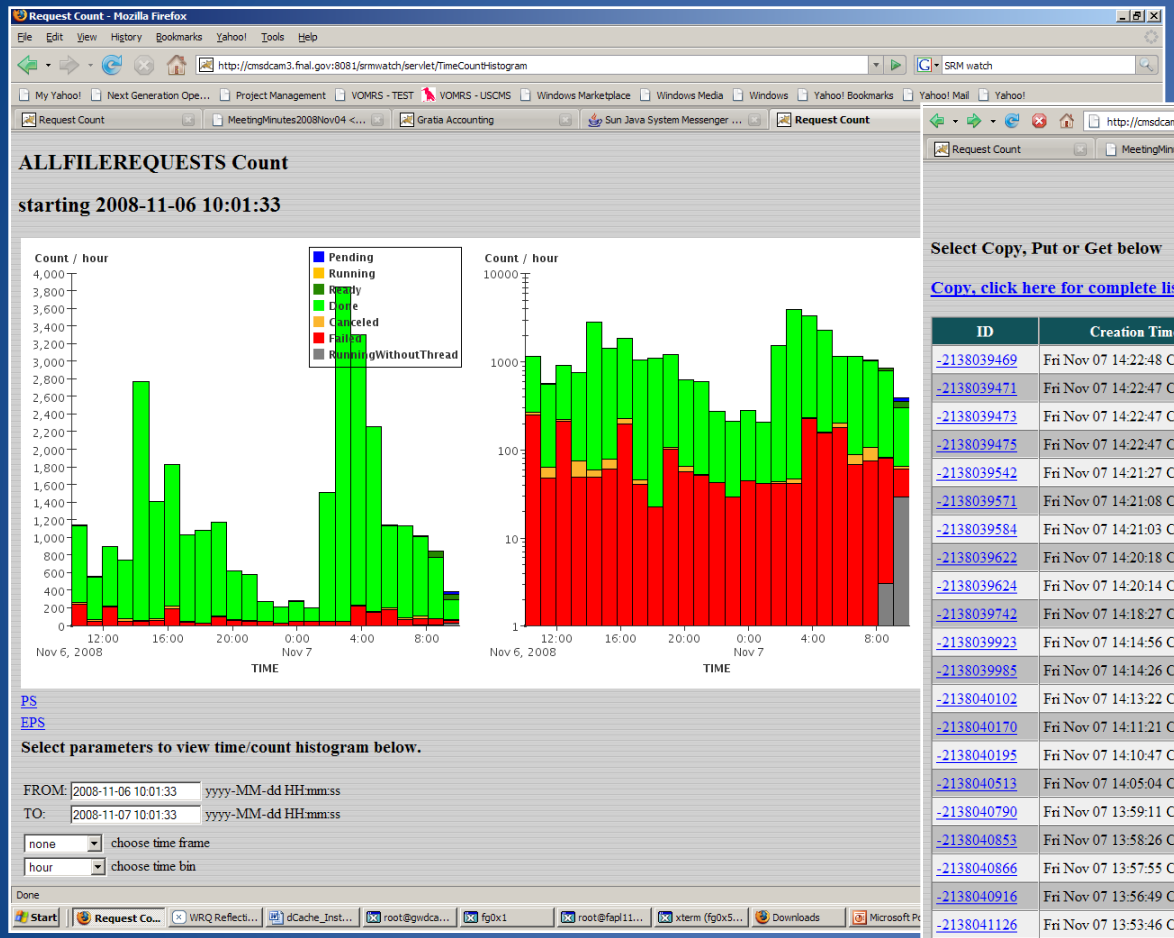
CellName	DomainName	Total Space/MB	Free Space/MB	Precious Space/MB	Layout (precious/used/free)
r-stkendca10a-2	r-stkendca10a-2Domain	788480	58	920	
r-stkendca11a-2	r-stkendca11a-2Domain	788480	2018	0	
r-stkendca12a-2	r-stkendca12a-2Domain	788480	287	0	
r-stkendca13a-2	r-stkendca13a-2Domain	1024000	4327	0	
r-stkendca13a-3	r-stkendca13a-3Domain	1024000	219	0	
v-stkendca13a-1	v-stkendca13a-1Domain	1024000	780614	235750	
v-stkendca14a-1	v-stkendca14a-1Domain	1024000	780291	237245	
v-stkendca15a-1	v-stkendca15a-1Domain	880640	780026	100613	
v-stkendca16a-1	v-stkendca16a-1Domain	880640	780864	78638	
w-stkendca10a-1	w-stkendca10a-1Domain	788480	699	208833	
w-stkendca10a-3	w-stkendca10a-3Domain	788480	48	154	
w-stkendca10a-4	w-stkendca10a-4Domain	788480	597	170472	
w-stkendca10a-5	w-stkendca10a-5Domain	788480	892	247109	
w-stkendca10a-6	w-stkendca10a-6Domain	788480	3839	311478	
w-stkendca11a-3	w-stkendca11a-3Domain	788480	112	28	
w-stkendca11a-4	w-stkendca11a-4Domain	788480	100	190094	
w-stkendca11a-5	w-stkendca11a-5Domain	788480	424	224618	
w-stkendca11a-6	w-stkendca11a-6Domain	788480	1153	176190	
w-stkendca12a-3	w-stkendca12a-3Domain	788480	123	0	
w-stkendca12a-4	w-stkendca12a-4Domain	788480	273	84557	

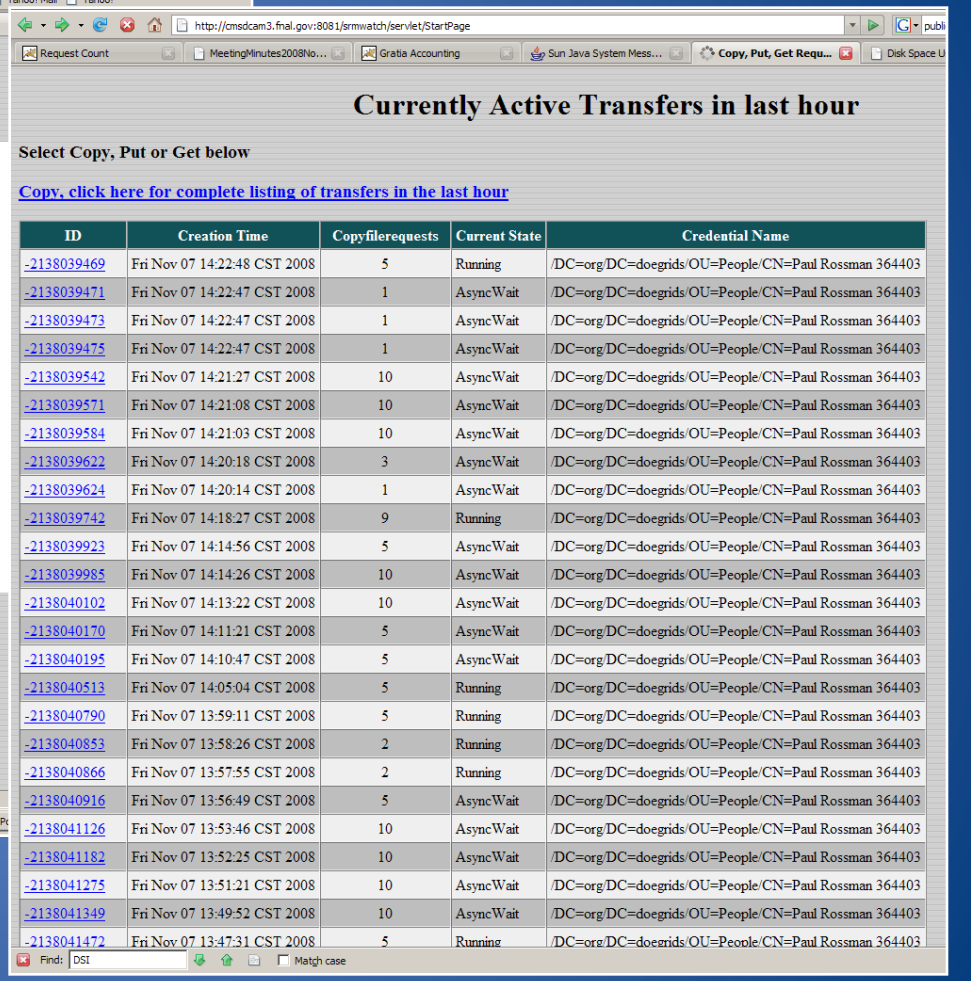
Numerous ways to monitor

dCache

- Info Pages - http://localhost:2288
- SrmWatch - http://localhost:8080/srmwatch/
- Admin Interface (UI or from command line)
- Scripts/web pages built by community
  - http://fndca.fnal.gov
  - http://fndca2a.fnal.gov:9090/lps/plots/src/plots.lzx

SRMWatch Examples





ID	Creation Time	Copyfilerequests	Current State	Credential Name
-2138039469	Fri Nov 07 14:22:48 CST 2008	5	Running	/DC=org/DC=doegrids/OU=People/CN=Paul Rossman 364403
-2138039471	Fri Nov 07 14:22:47 CST 2008	1	AsyncWait	/DC=org/DC=doegrids/OU=People/CN=Paul Rossman 364403
-2138039473	Fri Nov 07 14:22:47 CST 2008	1	AsyncWait	/DC=org/DC=doegrids/OU=People/CN=Paul Rossman 364403
-2138039475	Fri Nov 07 14:22:47 CST 2008	1	AsyncWait	/DC=org/DC=doegrids/OU=People/CN=Paul Rossman 364403
-2138039542	Fri Nov 07 14:21:27 CST 2008	10	AsyncWait	/DC=org/DC=doegrids/OU=People/CN=Paul Rossman 364403
-2138039571	Fri Nov 07 14:21:08 CST 2008	10	AsyncWait	/DC=org/DC=doegrids/OU=People/CN=Paul Rossman 364403
-2138039584	Fri Nov 07 14:21:03 CST 2008	10	AsyncWait	/DC=org/DC=doegrids/OU=People/CN=Paul Rossman 364403
-2138039622	Fri Nov 07 14:20:18 CST 2008	3	AsyncWait	/DC=org/DC=doegrids/OU=People/CN=Paul Rossman 364403
-2138039624	Fri Nov 07 14:20:14 CST 2008	1	AsyncWait	/DC=org/DC=doegrids/OU=People/CN=Paul Rossman 364403
-2138039742	Fri Nov 07 14:18:27 CST 2008	9	Running	/DC=org/DC=doegrids/OU=People/CN=Paul Rossman 364403
-2138039923	Fri Nov 07 14:14:56 CST 2008	5	AsyncWait	/DC=org/DC=doegrids/OU=People/CN=Paul Rossman 364403
-2138039985	Fri Nov 07 14:14:26 CST 2008	10	AsyncWait	/DC=org/DC=doegrids/OU=People/CN=Paul Rossman 364403
-2138040102	Fri Nov 07 14:13:22 CST 2008	10	AsyncWait	/DC=org/DC=doegrids/OU=People/CN=Paul Rossman 364403
-2138040170	Fri Nov 07 14:11:21 CST 2008	5	AsyncWait	/DC=org/DC=doegrids/OU=People/CN=Paul Rossman 364403
-2138040195	Fri Nov 07 14:10:47 CST 2008	5	AsyncWait	/DC=org/DC=doegrids/OU=People/CN=Paul Rossman 364403
-2138040513	Fri Nov 07 14:05:04 CST 2008	5	Running	/DC=org/DC=doegrids/OU=People/CN=Paul Rossman 364403
-2138040790	Fri Nov 07 13:59:11 CST 2008	5	Running	/DC=org/DC=doegrids/OU=People/CN=Paul Rossman 364403
-2138040853	Fri Nov 07 13:58:26 CST 2008	2	Running	/DC=org/DC=doegrids/OU=People/CN=Paul Rossman 364403
-2138040866	Fri Nov 07 13:57:55 CST 2008	2	Running	/DC=org/DC=doegrids/OU=People/CN=Paul Rossman 364403
-2138040916	Fri Nov 07 13:56:49 CST 2008	5	AsyncWait	/DC=org/DC=doegrids/OU=People/CN=Paul Rossman 364403
-2138041126	Fri Nov 07 13:53:46 CST 2008	10	AsyncWait	/DC=org/DC=doegrids/OU=People/CN=Paul Rossman 364403
-2138041182	Fri Nov 07 13:52:25 CST 2008	10	AsyncWait	/DC=org/DC=doegrids/OU=People/CN=Paul Rossman 364403
-2138041275	Fri Nov 07 13:51:21 CST 2008	10	AsyncWait	/DC=org/DC=doegrids/OU=People/CN=Paul Rossman 364403
-2138041349	Fri Nov 07 13:49:52 CST 2008	10	AsyncWait	/DC=org/DC=doegrids/OU=People/CN=Paul Rossman 364403
-2138041472	Fri Nov 07 13:47:31 CST 2008	5	Running	/DC=org/DC=doegrids/OU=People/CN=Paul Rossman 364403

Gratia Service

Gratia is the accounting service for OSG is provided by the Gratia

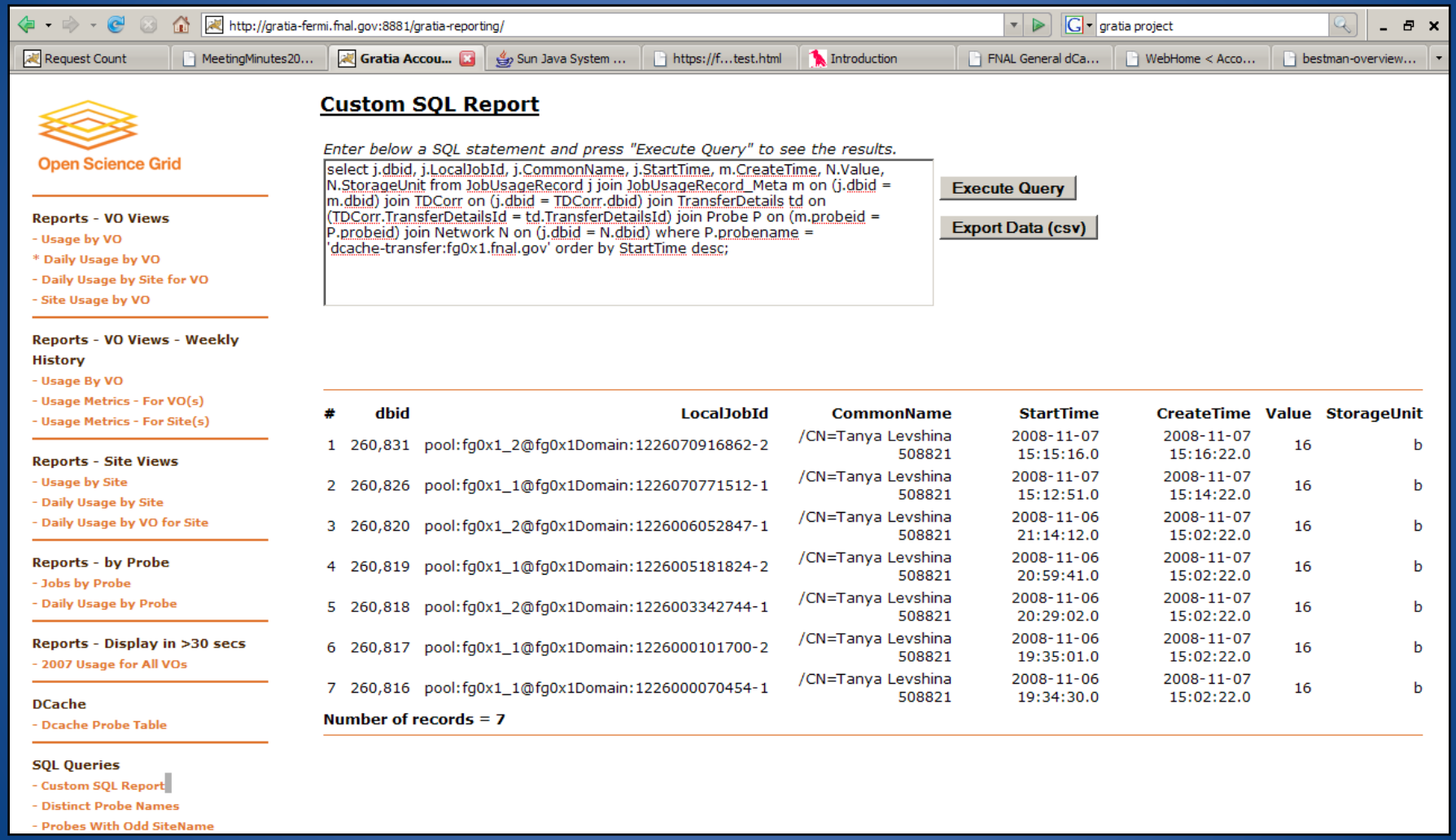
external project.

- Main goal is to provide the stakeholders with a reliable and accurate set of views of the Grid resources usage.
- Job and other accounting information gathered by Gratia probes run on the compute element or other site nodes are reported to a Gratia collectors
  - Fermi collector: http://gratia-fermi.fnal.gov:8886/gratia-reporting
  - OSG collector: http://gratia.opensciencegrid.org:8886/gratia-reporting
- Accounting records collected by Gratia are forwarded to the EGEE accounting system, APEL:
  - http://www3.egee.cesga.es/gridsite/accounting/CESGA/osg_view.html

dCache Gratia Probes

- dCache Gratia Probes
  - Storage Probe
  - Transfer Probe
- Storage Probe
  - Is responsible for reporting storage capacity and storage usage
  - Gets the pool information from the dCache admin server
  - Gets the SRM information from the SRM tables in the SRM Postgres database
  - Runs as a cron job on the host running the Postgres database server for SRM
- Transfer Probe
  - Reports the details of each file transfer into or out of a dCache file server
  - Gets this information from the dCache "billing" database.
  - Runs as a daemon process
  - For performance reasons, sites with large dCache billing databases are advised to alter the "billinginfo" table by indexing specific tables in order speed up the search for newly added records

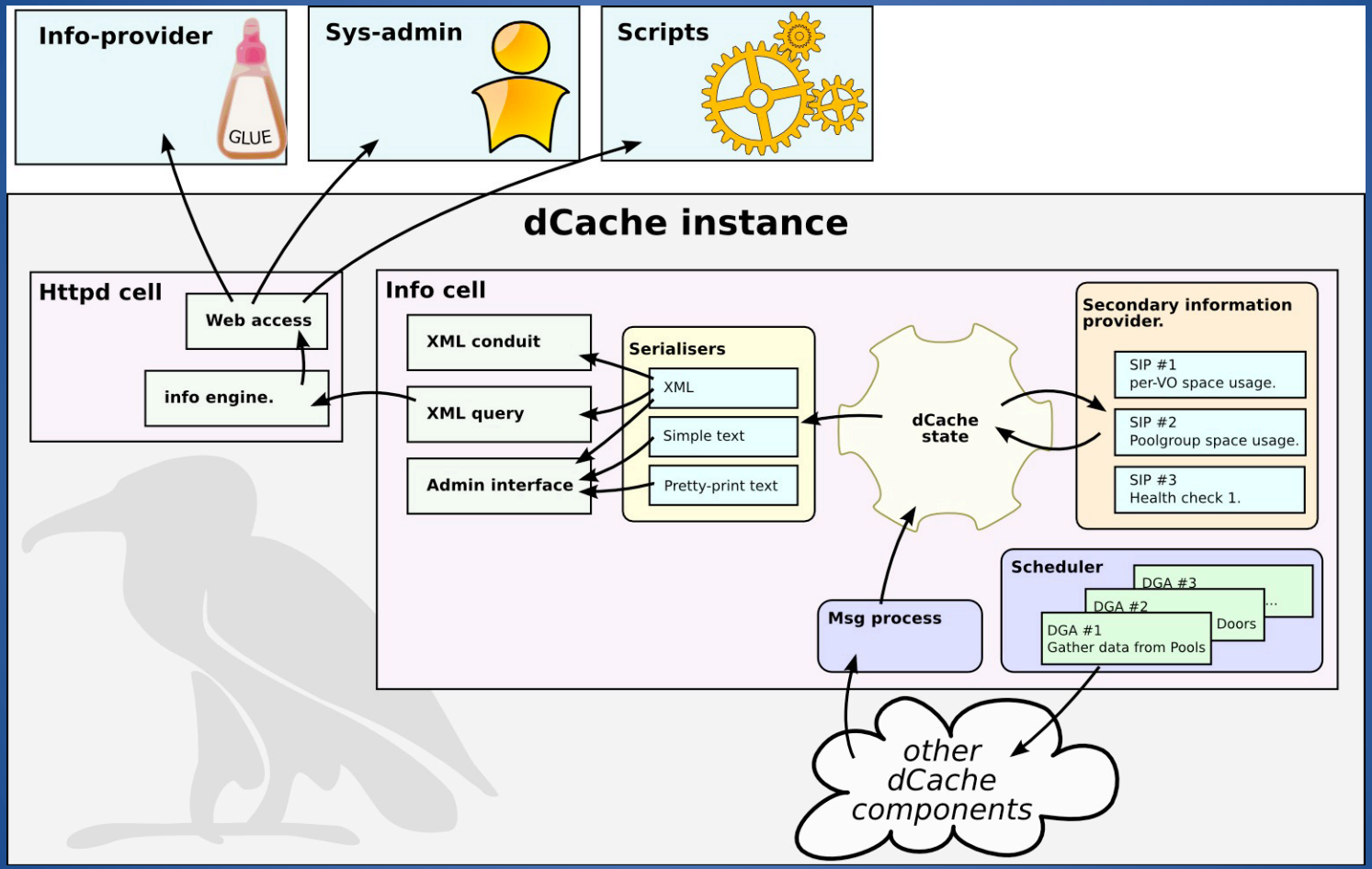
Gratia Transfer Probes Report



dCache Information Service

- A snapshot of the current status of a dCache instance for external consumption
  - Works independently of the rest of dCache.
  - Pools information for various dCache components (configurable update period).
- Doesn't provide historic data
- Not an info-provider but GIP is a customer
- Access data:
  - Via the admin interface
  - Via XML Conduit - A TCP connection get complete state as XML.
  - Web front-end
  - Supports some advanced features:
- Derived data (re-)calculated as state changes.
- Multiple output formats and transports.
- If additional metrics, data formats or transports are needed, they can be added

Information Service Architecture



From Paul Millar's presentation: "dCache seminar: introducing the info service "
http://www.dcache.org/manuals/dCache-info-20080813.pdf

Opportunistic Storage

- Opportunistic Storage in dCache 1.8 with SRM 2.2
  - Provides a capability of specifying a portion of the total storage for opportunistic use
  - Allows particular VOs and Roles a privilege to use space other than that included in opportunistic storage
  - Files created through opportunistic use will not be permanently available in the storage system
  - A storage site administrator may configure the site for opportunistic use through space reservation.
  - Creation of space reservations is controlled by use of link groups
  - The administrator may assign storage pools to link groups
  - Certain pools are designated for opportunistic use.
- Numerous documents describing how to install and operate Opportunistic Storgae on Tier-2 sites
  - https://twiki.grid.iu.edu/twiki/bin/view/Storage/OpportunisticStorageSetup
  - https://twiki.grid.iu.edu/bin/view/Storage/OpportunisticStorageModelForCMS

dCache in VDT

- dCache is one of the data storage solutions supported by OSG
- dCache could be installed from VDT
- Current version is vdt 2.2.8 (dcache 1.8.0.15 p11)
- Distribution contains dCache-server, pnfs, postgress, gratia probes rpms and a configuration script tailored to set up dCache for Tier-2/Tier-3
- Configuration script allows to do system setup, enable opportunistic storage, replication etc
- dCache-clients are distributed as a part of VDT client cache
  - Fermi client
  - LBNL client
  - LCG-utils
- There are multiple installations of dCache on OSG sites

Summary

- We will continue to work on improving storage packaging in VDT
  - Feedback is welcome!
- We are trying to make support more efficient by providing FQA, debugging the most frequently occurred problems, working with developers on improving logging and error diagnostic
  - The quality of the support depends greatly on Storage Admins cooperation!!!
- We will add BeStMan test stand and will do BeStMan/Xrootd certification the same way it is done with dCache release
- We will be glad to accept/package more community tools
- We are trying to maintain documentation up-to-date, adding new interesting references and “how to do” tips
  - Please let us know if we are missing some important topics!
- As a liaison to software developers we will be happy to pass your requests/ suggestions

Acknowledgements

- Tanya Levshina