

Tier-2 storage in GridPP

Communications, monitoring and tools

Greig A. Cowan

University of Edinburgh

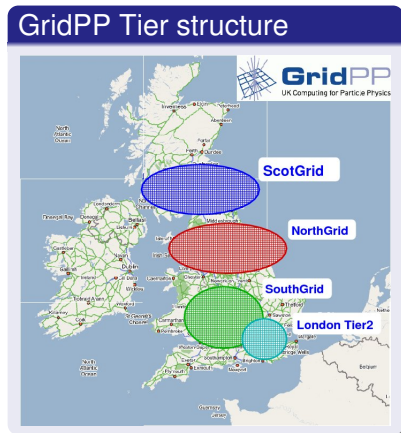
GDB 8th Oct 2008



- 1 Introduction
- 2 How we organise storage in GridPP
 - Communications
 - GridPP storage operations monitoring
 - Site monitoring of storage
 - SE administration tools
- 3 Wish list

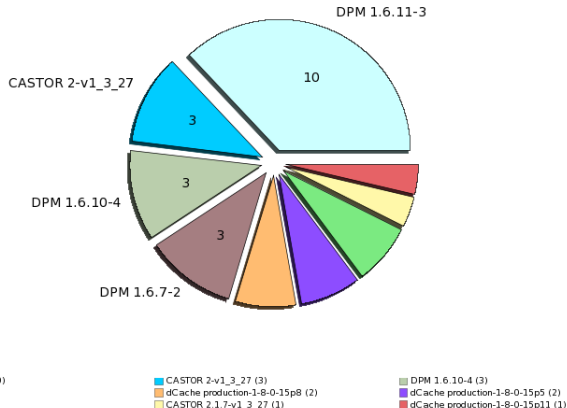


- **GridPP** is a project to construct a data-Grid for particle physics in the UK.
- I have worked for GridPP for 3.5 years.
- Focus of work has been the deployment/operations of storage and data management technologies at UK sites (19 in total).
- In particular, looking at the Grid middleware.
- GridPP3 recently funded with $\sim 40\text{M}$ Euro over 3 years.



Current status of deployed SRMv2.2 endpoints (Sum: 27 Count)

24 Hours from 2008-10-06 16:19 to 2008-10-07 16:19 UTC



- DPM and dCache at Tier2s; CASTOR at Tier1.
- Two test StoRM instances (GPFS and Lustre).
- Sites have been moving away from dCache over past year.

gridpp-storage@jiscmail.ac.uk

- 65 subscribers; 6 from outside UK.
- Discussions about how best to operate Grid storage.
- Experiment representation.

Weekly 30min meetings

- Sites can report problems.
- Information about new releases and bugs.
- Discussions about upcoming CCRC/CSA/FDR...
- Hardware procurement news.

Workshops and meetings

- SRM2.2 workshop.
- GridPP collaboration meetings (~ 6 months).

Blog

<http://gridpp-storage.blogspot.com>

GGUS ticket RSS feed

https://gus.fzk.de/pages/rss/rss_open_uki_sites.php

gridpp-storage@jiscmail.ac.uk

- 65 subscribers; 6 from outside UK.
- Discussions about how best to operate Grid storage.
- Experiment representation.

Weekly 30min meetings

- Sites can report problems.
- Information about new releases and bugs.
- Discussions about upcoming CCRC/CSA/FDR...
- Hardware procurement news.

Workshops and meetings

- SRM2.2 workshop.
- GridPP collaboration meetings (~ 6 months).

Blog

<http://gridpp-storage.blogspot.com>

GGUS ticket RSS feed

https://gus.fzk.de/pages/rss/rss_open_uki_sites.php



gridpp-storage@jiscmail.ac.uk

- 65 subscribers; 6 from outside UK.
- Discussions about how best to operate Grid storage.
- Experiment representation.

Weekly 30min meetings

- Sites can report problems.
- Information about new releases and bugs.
- Discussions about upcoming CCRC/CSA/FDR...
- Hardware procurement news.

Workshops and meetings

- SRM2.2 workshop.
- GridPP collaboration meetings (~ 6 months).

Blog

<http://gridpp-storage.blogspot.com>

GGUS ticket RSS feed

https://gus.fzk.de/pages/rss/rss_open_uki_sites.php

Two types of interest

- 1 GridPP operations (across all T2s)
- 2 Site operations (within a T2)

Multiple sources of useful information

- 1 BDII (i.e., SRM endpoints, space information)
- 2 SAM database of test results
- 3 The SRMs themselves

- Developed some basic tools (or modified existing ones) to collect this information and visualise it for operations staff and sites.
 - Python scripts insert information into an SQL DB.
 - **GraphTool** (B. Bockleman) then used to query the DB and visualise the information.



Two types of interest

- 1 GridPP operations (across all T2s)
- 2 Site operations (within a T2)

Multiple sources of useful information

- 1 BDII (i.e., SRM endpoints, space information)
- 2 SAM database of test results
- 3 The SRMs themselves

- Developed some basic tools (or modified existing ones) to collect this information and visualise it for operations staff and sites.
 - Python scripts insert information into an SQL DB.
 - **GraphTool** (B. Bockleman) then used to query the DB and visualise the information.



Two types of interest

- 1 GridPP operations (across all T2s)
- 2 Site operations (within a T2)

Multiple sources of useful information

- 1 BDII (i.e., SRM endpoints, space information)
- 2 SAM database of test results
- 3 The SRMs themselves

- Developed some basic tools (or modified existing ones) to collect this information and visualise it for operations staff and sites.
 - Python scripts insert information into an SQL DB.
 - **GraphTool** (B. Bockleman) then used to query the DB and visualise the information.



GridPP operations monitoring

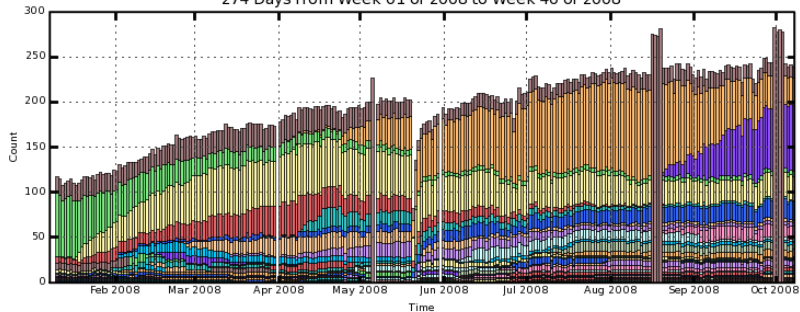


Site	SE type	Version
http://dcache01.tier2.hep.manchester.ac.uk:8443/srm/managerv2	dCache	production-1-8-0-15p5
http://cerb-mds.bris.ac.uk:8444/srm/managerv2	StoRM	
http://pc55.hep.ucl.ac.uk:8446/srm/managerv2	DPM	1.6.7-2
http://lcgse0.shef.ac.uk:8446/srm/managerv2	DPM	1.6.11-3
http://bohr3223.tier2.hep.manchester.ac.uk:8446/srm/managerv2	DPM	1.6.10-4
http://se2.ppgrid1.rhul.ac.uk:8446/srm/managerv2	DPM	1.6.11-3
http://fal-pygrid-30.lancs.ac.uk:8446/srm/managerv2	DPM	1.6.11-3
http://epgse1.ph.bham.ac.uk:8446/srm/managerv2	DPM	1.6.10-4
http://se01.esc.qmul.ac.uk:8444/srm/managerv2	StoRM	
http://srm-cms.gridpp.rl.ac.uk:8443/srm/managerv2	CASTOR	2-v1_3_27
http://heplnx204.pp.rl.ac.uk:8443/srm/managerv2	dCache	production-1-8-0-15p11
http://svr018.gla.scotgrid.ac.uk:8446/srm/managerv2	DPM	1.6.11-3
http://hepgrid5.ph.liv.ac.uk:8443/srm/managerv2	dCache	production-1-8-0-15p8
http://srm-alice.gridpp.rl.ac.uk:8443/srm/managerv2	CASTOR	2-v1_3_27
http://srm-atlas.gridpp.rl.ac.uk:8443/srm/managerv2	CASTOR	2-v1_3_27
http://se1.pp.rhul.ac.uk:8446/srm/managerv2	DPM	1.6.7-2

- Obtain list of SRM endpoints from **BDII** each day.
- Use StoRM SRM client to perform an srmPing. INSERT to MySQL DB.
- **GraphTool** allows you to pick out the endpoints of interest (i.e., *.ac.uk).
- Allows us to quickly identify sites that are falling behind.

Storage Middleware Release

274 Days from Week 01 of 2008 to Week 40 of 2008



Maximum: 282.00 , Minimum: 0.00 , Average: 190.27 , Current: 239.00

<https://twiki.cern.ch/twiki/bin/view/LCG/GSSDSRMv2DeploymentMonitoring>

Which sites have configured their spaces?

Site	SE, VO	Token Desc
UKI-SCOTGRID-DURHAM	('se01.dur.scotgrid.ac.uk', 'atlas')	ATLASPRODDISK ATLASDATADISK ATLASMCDISK
UKI-NORTHGRID-LIV-HEP	('hepgrid11.ph.liv.ac.uk', 'atlas')	ATLASPRODDISK ATLASUSERDISK ATLASGROUPDISK ATLASMCDISK ATLASLOCALGROUPDISK ATLASDATADISK
UKI-SOUTHGRID-CAM-HEP	('serv02.hep.phy.cam.ac.uk', 'atlas')	ATLASMCDISK ATLASPRODDISK ATLASUSERDISK ATLASGROUPDISK ATLASLOCALGROUPDISK ATLASDATADISK
UKI-SCOTGRID-GLASGOW	('svr018.gla.scotgrid.ac.uk', 'atlas')	ATLASUSERDISK ATLASLOCALGROUPDISK ATLASMCDISK ATLASGROUPDISK ATLASDATADISK ATLASPRODDISK
UKI-SCOTGRID-ECDF	('srm.glite.ecdf.ed.ac.uk', 'atlas')	ATLASDATADISK ATLASMCDISK ATLASPRODDISK

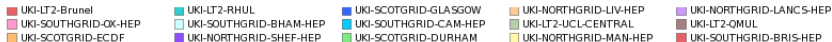
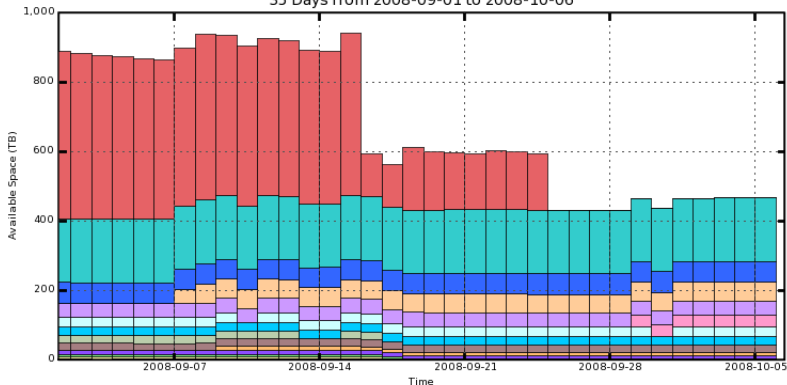
- Daily query of BDII to get list of published SRM2.2 **space token desc**s.
- Very useful for keeping track of site status.



How has the available space changed over time?

Available space of advertised SRM2.2 token descriptions (Site View)

35 Days from 2008-09-01 to 2008-10-06

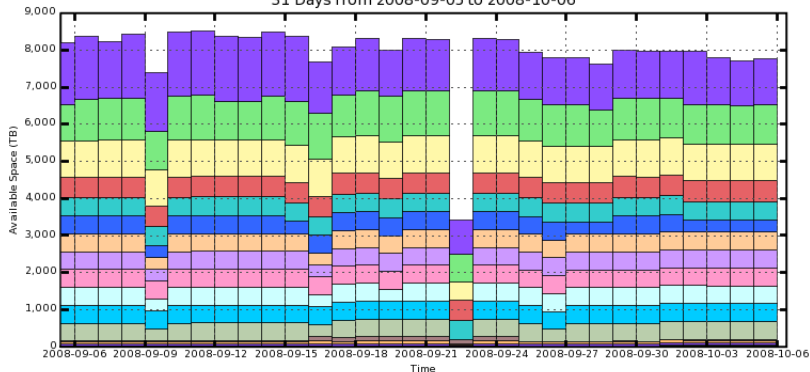


Maximum: 939.02 , Minimum: 429.64 , Average: 678.89 , Current: 464.75

How has the available space changed over time?

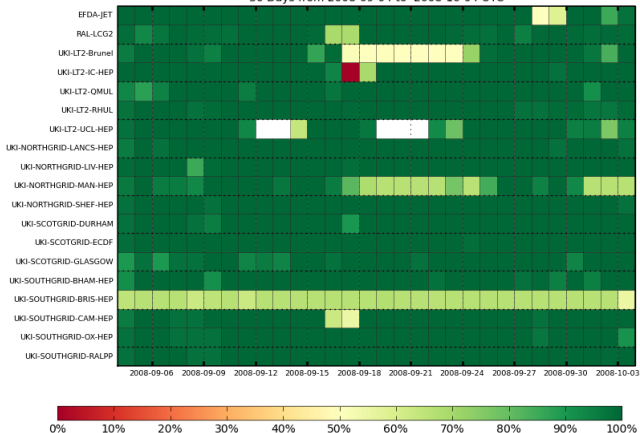
Available space of advertised SRM2.2 token descriptions (VO View)

31 Days from 2008-09-05 to 2008-10-06



Maximum: 8,536 , Minimum: 3,421 , Average: 7,951 , Current: 7,785

SAM view of GridPP site storage
30 Days from 2008-09-04 to 2008-10-04 UTC



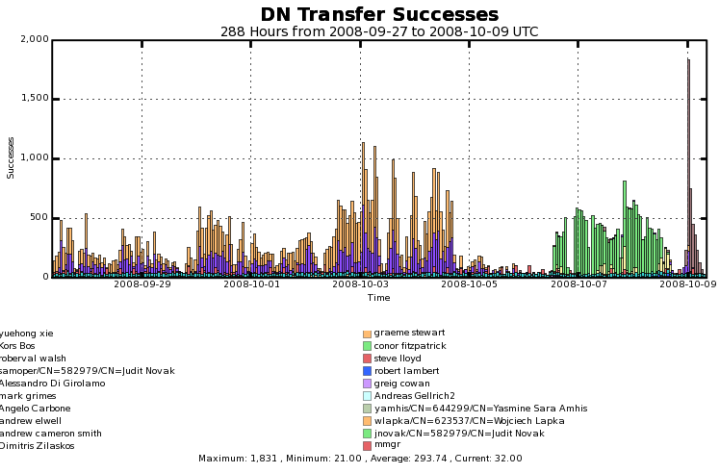
- Query the SAM DB and plot the SRM and SE results for each site.

http://www.gridpp.ac.uk/wiki/GridPP_storage_availability_monitoring



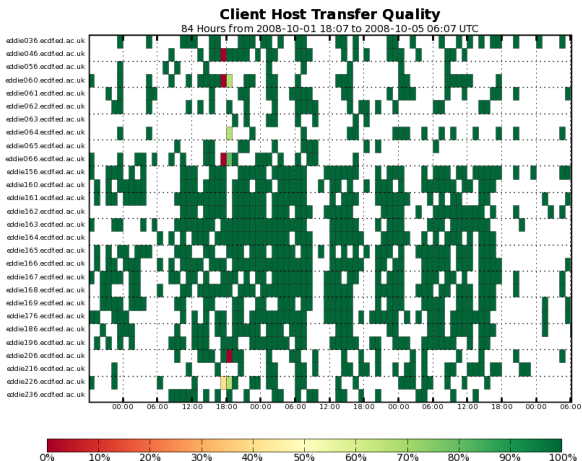
Site monitoring





- **Expose the internals** of DPM DB to the site admin in graphical form.
- Many UK sites have installed the RPM.

http://www.gridpp.ac.uk/wiki/DPM_Monitoring



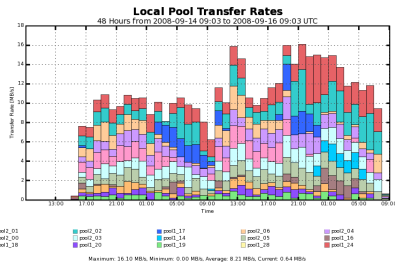
Live Edinburgh system

<http://srm.glite.ecdf.ed.ac.uk/dpm/xml/>

<http://www.ph.ed.ac.uk/gcowan1/dpm/monitoring.html>

dCache Billing Graphs

- Developed by Brian Bockleman (Nebraska).
- Provided the foundation for GridppDpmMonitor.
- Very useful for sites to see what is happening.



<http://t2.unl.edu/documentation/dcache/installing-the-billing-graphs/>



Storage administration tools



- Toolkit to help sites perform common operations, i.e.,
 - 1 `gridpp_dpm_dpns_to_disk`: **Maps namespace** → **physical files**
 - 2 `gridpp_dpm_disk_to_dpns`: **Maps physical files** → **namespace**
 - 3 `gridpp_dpm_dpns_du`: **Check usage of DPNS namespace**
 - 4 ...
- Scripts built using the DPM python API.
- RPM sits in HEPiX sys-admin YUM repository.
- M. Jouvin contributed his `dpm-listspaces` utility.

<http://www.gridpp.ac.uk/wiki/DPM-admin-tools>



- Toolkit to help sites perform common operations, i.e.,
 - 1 `gridpp_dpm_dpns_to_disk`: **Maps namespace** → **physical files**
 - 2 `gridpp_dpm_disk_to_dpns`: **Maps physical files** → **namespace**
 - 3 `gridpp_dpm_dpns_du`: **Check usage of DPNS namespace**
 - 4 ...
- Scripts built using the DPM python API.
- RPM sits in HEPiX sys-admin YUM repository.
- M. Jouvin contributed his `dpm-listspaces` utility.

<http://www.gridpp.ac.uk/wiki/DPM-admin-tools>



- Written and packaged by OSG (with external contributions).
- Python scripts using the dCache admin interface.
- Provides useful set of tools for performing common dCache operations, i.e.,
 - **Simple way to map between namespace and physical file.**
 - **Fix broken transfer requests.**
 - **Retire a pool (checks that replicas exist of all files on the pool).**
- I encourage all dCache sites to use this to help debug and solve problems.

<http://datagrid.ucsd.edu/toolkit/index.html>



Suggestions for middleware improvements



- Sites and experiments would like to be able to restrict the amount of space a user can write to.
 - See presentations from expts this morning.
- This is not possible with space tokens.
- User quotas are desired.
- i.e., CMS want to give local users ~ 1 TB of space at their T2.
- `gridpp_dpm_dpns_du` tool allows user to check their usage. This does not restrict the user!!



- More fine grained access desired to spaces and the SE namespace.
- Sites often asked to “run this script” in order to change the ACLs on DPM.
- Common interface?
- ACLs for dCache in 1.9.3 release (hurrah!).



- Some sites would like to see DPM support a more intelligent filesystem selection mechanism.
- Currently round robin.
- Suggestion: sites could assign a **weight** to filesystems corresponding to their size or expected level of service.
- Would allow sites to more easily balance their disk servers.



- This is always on the list of things to be improved.
- Would like all logging done to a central place, preferably using syslog.
- Readable/greppable format is essential.
- Past and ongoing security challenges require that sites can trace operations through the logs.



- Tools are being developed to allow SE namespaces to be dumped in a common format.
- How expensive will these dump operations be on the SE?
- Will this all happen automatically, or do the sites need to provide this upon request?



- Having the monitoring (i.e., GridppDpmMonitor) is an improvement upon log files/DB. . .
- . . . but it would be great to have **alarms** to tell admins when problems are occurring with particular services/daemons.
- Nagios is popular, and is now being used to report on SAM failures.
- This may require the sites to write the sensors themselves.



- See John's talk from earlier.



- GridPP has built up a successful storage **community** over past 3 years.
 - A few “experts” and a lot of input from the distributed sites.
- Storage **monitoring** tools have been developed to help:
 - GridPP-wide views
 - Site-level views
- **Administration tools** have been developed to ease day-to-day tasks.
- Please add your suggestions to the **wish list**.

