

GridPP Tier-2 experiences of dCache

Greig A. Cowan

University of Edinburgh



1. What is a GridPP Tier-2?
2. GridPP experiences
 - (a) Configuration, administration, monitoring
3. Some comments
4. Summary

What is GridPP?

- UK Grid for particle physics.
- Large computing facility (Tier-1) at Rutherford Appleton Laboratory (RAL).
- 19 geographically distinct Tier-2 sites.

What is a Tier-2?

In terms of storage, they can typically be characterised by:

- **No tape** backend.
- Relatively small amount of RAID5 disk (~ 10 -100TB).
- Single dCache *head* node and a few pool/door nodes.
- 1GbE external + internal connectivity.
- Resources may have to be shared with non-HEP users.
- **Limited manpower** (~ 1 FTE).
 - Ease of configuration, management and monitoring are essential to maximise availability.

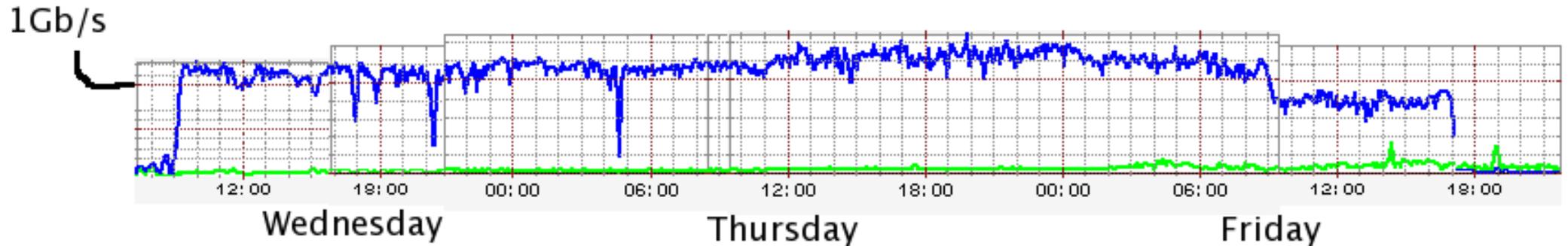
Site	Disk	Site	Disk
Edinburgh	20TB	Lancaster	60TB
RAL-PPD	20TB	Manchester	250TB
IC-HEP/IC-LeSC	50TB	Liverpool	7TB

- 12 (generally) smaller sites use DPM.
- A lot of experience:

<http://www.gridpp.ac.uk/wiki/DCache>

- Extensive testing of Tier-2 infrastructure.
 - dCache plays major part.

48 Hour Aggregate Test from RAL to Tier2s over SJ4 and UKLIGHT

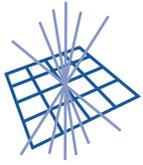


- **YAIM** used for initial basic installation.
- Admin typically performs final tweaks by hand. i.e., adding extra pools, pool groups, units, links...
- Integration of dCache with YAIM has improved greatly over the past 6 months.
 - Different pool and admin meta-packages.
 - Dedicated DESY repository.
 - Small incremental releases good.
 - * Although `apt` auto-update can break your install!

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

- Batch farm of **900 WNs**, each with >250 GB disk.
- Each WN running `dcache-pool`.
- 45 gridftp doors.
- Partitioned into **two** dCache's to ease management.
- Configuration with **cfengine** <http://www.cfengine.org>
 - Central repo of config files (`dCacheSetup`, `node_config`)
 - Node pulls in new config file if changed.
 - Not able to restart services yet.
- Resilient dCache **NOT** currently being used.
 - Testbed setup for evaluation.

- RAL-PPD has deployed the xrootd door in read only mode.
- Initial tests showed that basic functionality was working.
- Chris Brew heavily involved in BaBar computing.
- Has since included **dcap** support in BaBar software so xrootd not used significantly.

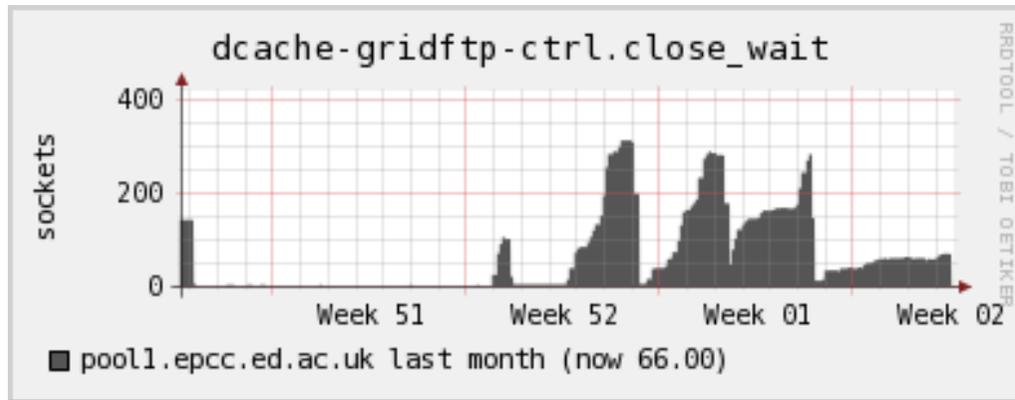


GridPP

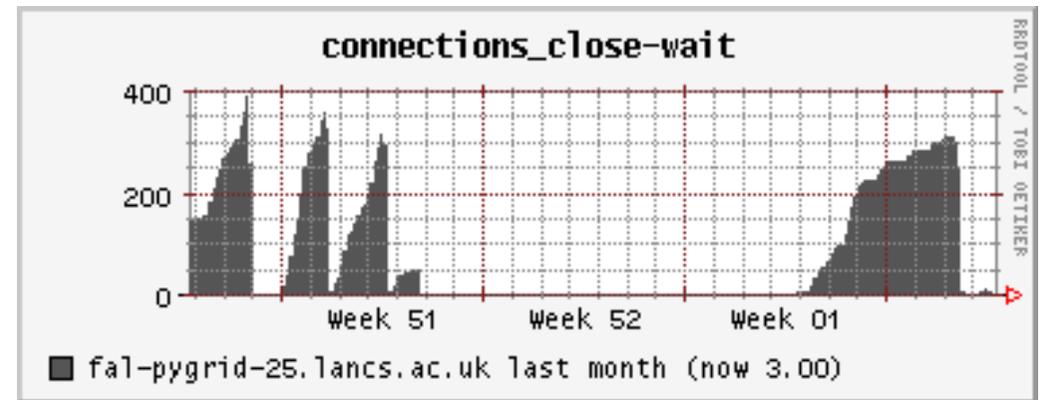
UK Computing for Particle Physics

OBSERVATIONS

Edinburgh



Lancaster



- Eventually door stops working.

```
java.lang.OutOfMemory
```

```
gPlazma diskCacheV111.services.authorization.AuthorizationServiceException
```

- Everything else is functioning.

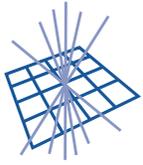
- Typical netstat output (29107 is the gridftp door process):

```
tcp 1 0 pool1.epcc.ed.ac.uk:2811 fts106.cern.ch:20009 CLOSE_WAIT 29107/java
```

- dCache logs remain **cryptic**.
 - Solution is often to restart the service. Is there a better way?
- Tomcat logs filled up root partition of Lancaster SRM node.
 - 5GB!
 - Tomcat logs in a different place from dCache and PNFS logs.

https://www.gridpp.ac.uk/wiki/DCache_Log_Message_Archive

- Namespace ↔ disk pool **synchronisation**.
 - People often find *ghost* files on disk or in PNFS.
 - Would like to identify discrepancies and fix them.
- **Admin shell** is really not user friendly.
 - Could we share scripting tools that individual sites have developed?
 - Would like to find out about the jpython interface.

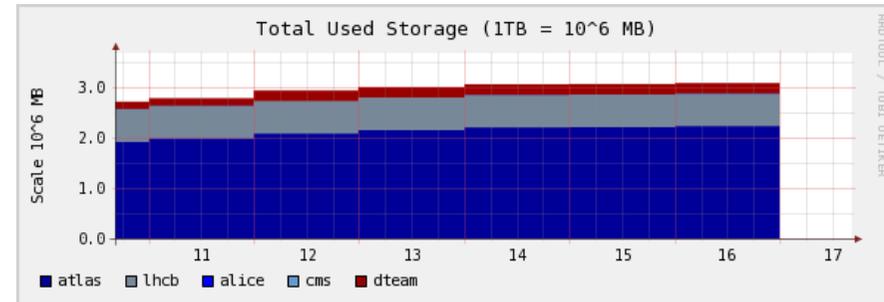


GridPP

UK Computing for Particle Physics

CURRENT WORK

- System deployed in UK. Accounting for EGEE.
- Uses information in the global **BDII**.
- Difficult to account storage if VOs share disk pools.
 - GridPP have own GIP plugin (du on /pnfs).
 - * Unable to query database. Chimera?



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

- Stress testing **dcap** access from batch farm.
 - Do we need separate read/write pools? File hopping?
- ScotGrid **distributed dCache**.
 - Storage at Edinburgh and Glasgow.
 - Use lightpath between sites.
 - Single SRM for the entire Tier-2 cluster → simpler to manage with shared support?
- **Monitoring**
 - See talk tomorrow.

- **Good understanding** within GridPP of how to setup basic Tier-2 SRM (see wiki).
 - Still gaining experience in setting up a large site (100's of nodes and TB's).
- dCache is a key component of the SRM landscape in the UK.
- Problems difficult to debug due to **logfiles**.
- Further investigation of **local access** to the storage is needed.
- Improved **monitoring** would be beneficial to community.