

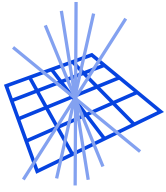
GridPP
UK Computing for Particle Physics

UK Tier-2 experiences of dCache

Greig A Cowan

University of Edinburgh

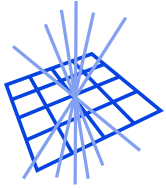




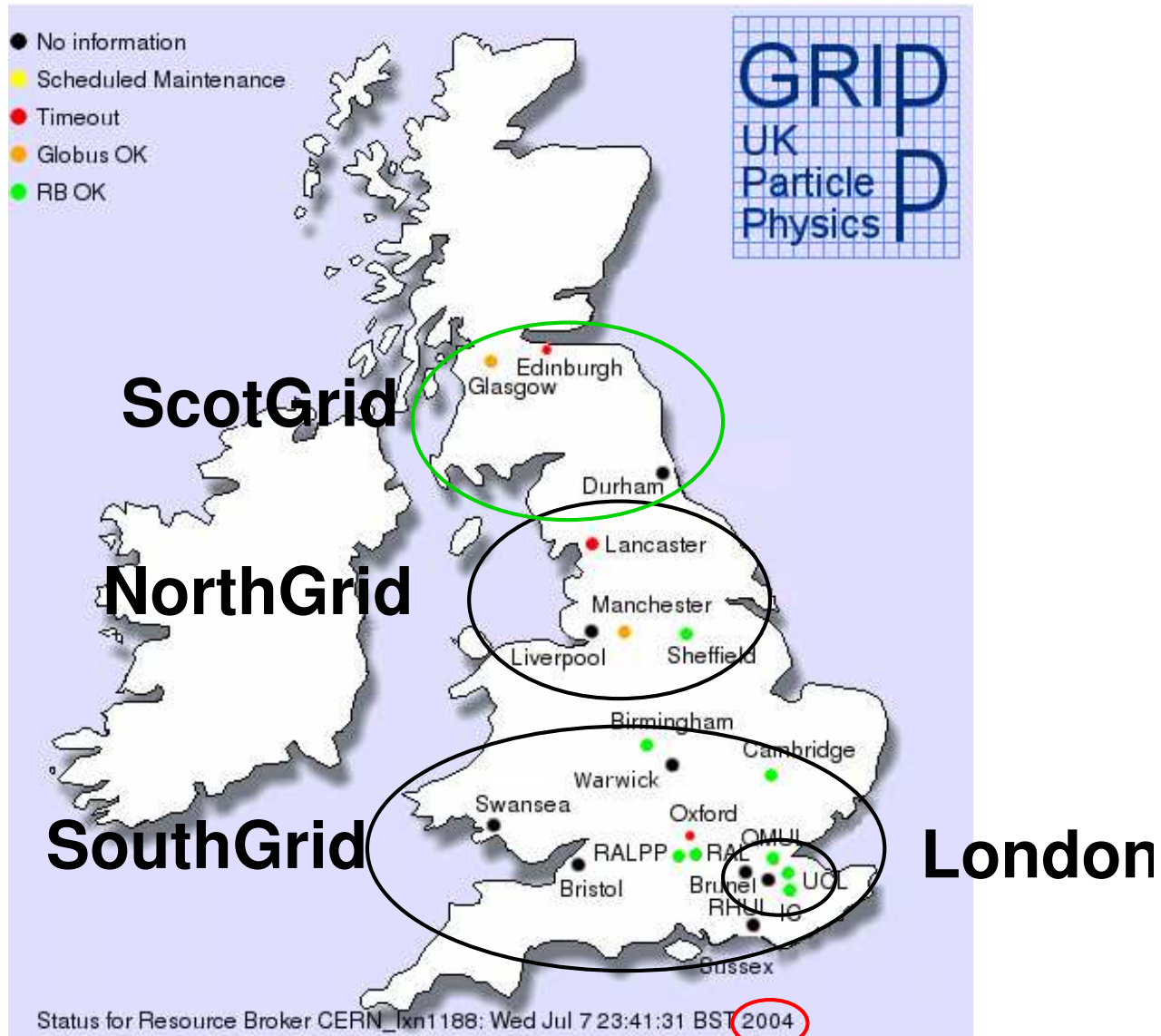
GridPP
UK Computing for Particle Physics

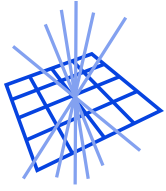
Outline

1. Structure of GridPP
2. Role of the Tier-2s in SC3
3. Initial experiences
4. Edinburgh as a Tier-2 site
 - setup
 - performance
 - problems
5. Other Tier-2 sites
6. Conclusions



Structure of GridPP

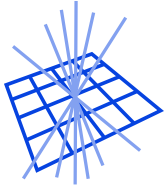




GridPP
UK Computing for Particle Physics

LCG Service Challenge 3 (SC3)

- 3 Tier-2 sites to deploy an SRM as part of SC3.
 - Edinburgh, Lancaster and Imperial College
- dCache chosen as SRM.

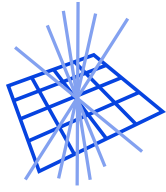


LCG Service Challenge 3 (SC3)

- 3 Tier-2 sites to deploy an SRM as part of SC3.
 - Edinburgh, Lancaster and Imperial College
- dCache chosen as SRM.
- Manchester, Glasgow and RAL-PP have also installed dCache.
- Large scale deployment without direct support from developers.
- Support group (mailing list, phone conferences, Savannah) created to facilitate this deployment.

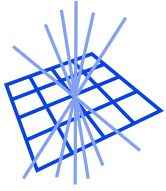
<http://wiki.gridpp.ac.uk/wiki>

<http://storage.esc.rl.ac.uk/documentation/html/D-Cache-Howto/index.html>



GridPP
UK Computing for Particle Physics

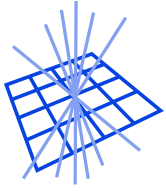
My initial experiences



GridPP
UK Computing for Particle Physics

My initial experiences

- Difficult to install (using apt-get and rpms).
- Not helped by lack of documentation/experience within GridPP.



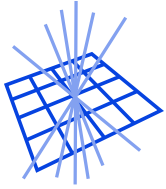
GridPP
UK Computing for Particle Physics

My initial experiences

- Difficult to install (using apt-get and rpms).
- Not helped by lack of documentation/experience within GridPP.

But . . .

- I did manage to deploy a working system after playing around with it, help from RAL (Tier-1) and the storage mailing list.
- Now finding dCache very easy to install (using YAIM).



My initial experiences

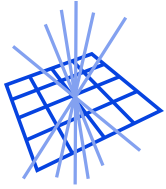
- Difficult to install (using apt-get and rpms).
- Not helped by lack of documentation/experience within GridPP.

But ...

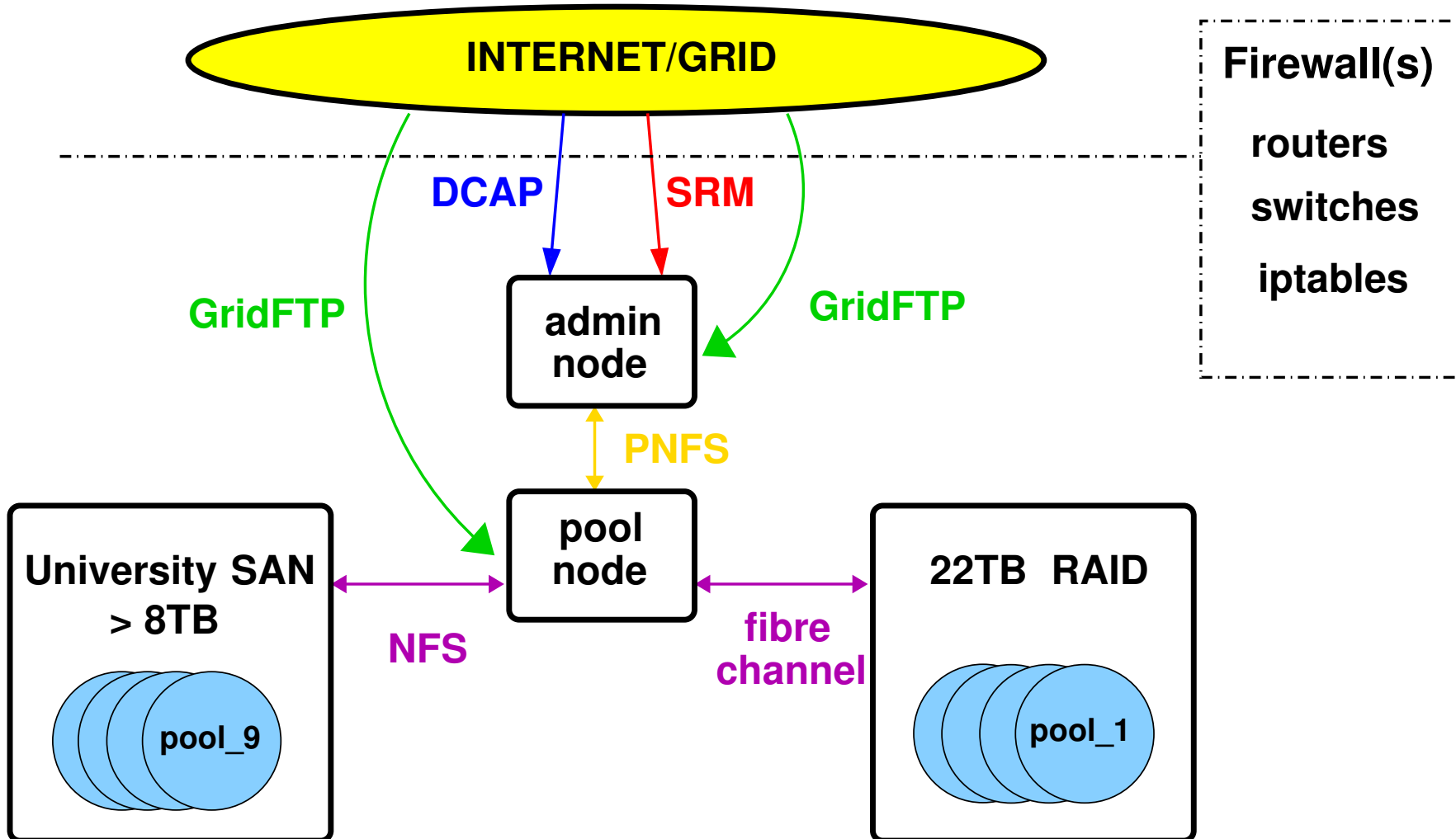
- I did manage to deploy a working system after playing around with it, help from RAL (Tier-1) and the storage mailing list.
- Now finding dCache very easy to install (using YAIM).

However ...

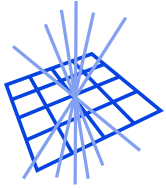
- Managing / maintaining / configuring of dCache is proving difficult.
- Lack of sysadmin documentation. GridPP trying to improve the situation (wiki).



Edinburgh dCache Setup

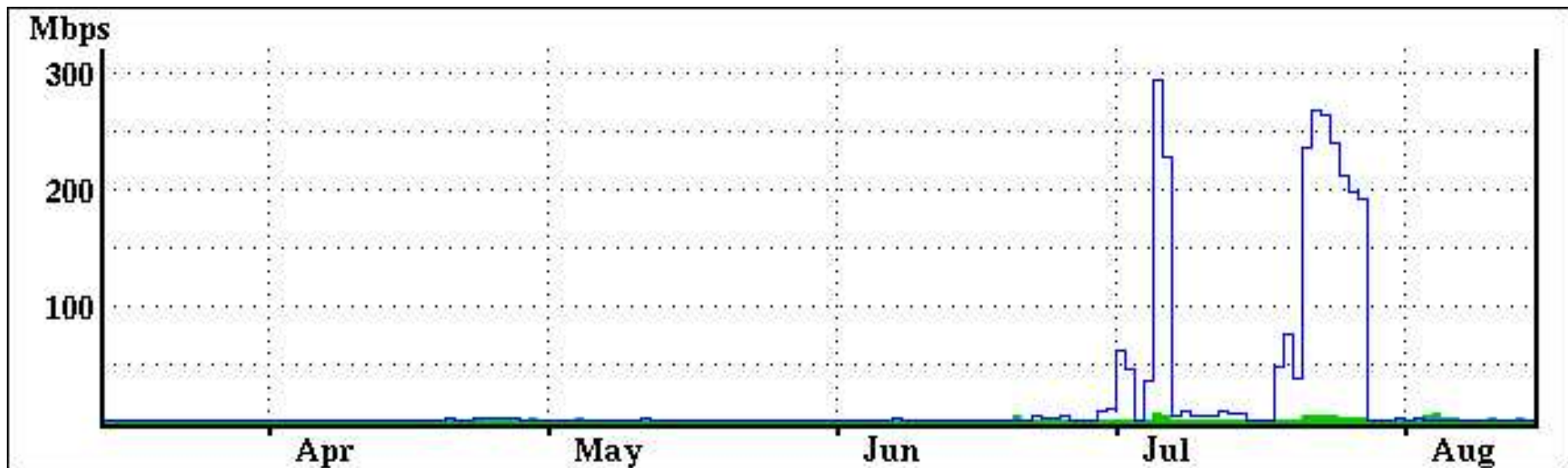


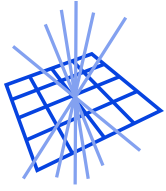
- SL3 on admin node (dual 2.8GHz); RH AS2.1 on pool node (8 * 1.9 GHz);



Edinburgh Performance

- Participated in SC3 File Transfer Service (FTS) tests.
- FTS was used by RAL to put files into our operational SRM.
- Observed sustained transfer rates of 200-250 Mbps into our dCache.
- ~ 12 TB of data transferred in during the tests.
- 30 minute average:

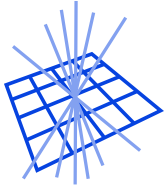




Edinburgh Issues

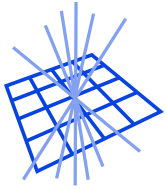
dCache related

- Initially difficult install with apt-get → improved significantly with YAIM.
- rpms installed by hand on pool node since disk server running unsupported OS. Need to hack together a solution for security.
- Successful migration of pnfs database after resolving firewall issue.
- Effective administration:



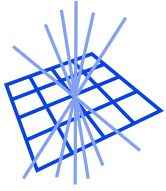
dCache related

- Initially difficult install with apt-get → improved significantly with YAIM.
- rpms installed by hand on pool node since disk server running unsupported OS. Need to hack together a solution for security.
- Successful migration of pnfs database after resolving firewall issue.
- Effective administration:
 - Difficult to understand **log-files**, making problem diagnosis difficult.



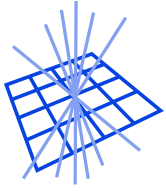
dCache related

- Initially difficult install with apt-get → improved significantly with YAIM.
- rpms installed by hand on pool node since disk server running unsupported OS. Need to hack together a solution for security.
- Successful migration of pnfs database after resolving firewall issue.
- Effective administration:
 - Difficult to understand **log-files**, making problem diagnosis difficult.
 - Need to write your own **scripts** to perform “basic” tasks i.e. draining pools.



dCache related

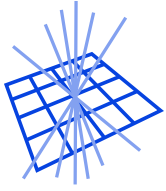
- Initially difficult install with apt-get → improved significantly with YAIM.
- rpms installed by hand on pool node since disk server running unsupported OS. Need to hack together a solution for security.
- Successful migration of pnfs database after resolving firewall issue.
- Effective administration:
 - Difficult to understand **log-files**, making problem diagnosis difficult.
 - Need to write your own **scripts** to perform “basic” tasks i.e. draining pools.
 - SRM cell was OFFLINE - required restart of dCache services. Due to faster hardware at Tier-2 sites.



GridPP
UK Computing for Particle Physics

Edinburgh Issues (2)

- How do you best **test** a new dCache setup (test suite?). i.e. transfers seem to work, but log files contain strange messages which lead to you thinking that something is wrong with system.

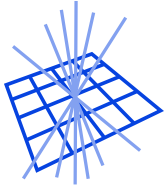


Edinburgh Issues (2)

- How do you best **test** a new dCache setup (test suite?). i.e. transfers seem to work, but log files contain strange messages which lead to you thinking that something is wrong with system.

General site/LCG issues

- Firewall configuration (gridftp problems).
- NFS mounting of SAN initially proved difficult (now using TCP). Transfer rates \sim 2-3 times lower (still need to optimise).
- Information system always seems to need tweaking.

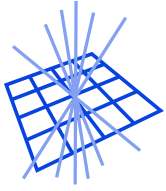


GridPP
UK Computing for Particle Physics

Other UK Tier-2's

Lancaster

- 1 admin node and 3 pool nodes.
- ~ 10 TB of disk on each pool node.
- Intend to dual home the system to connect up to the UKLight network (no experience yet).



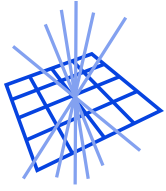
Other UK Tier-2's

Lancaster

- 1 admin node and 3 pool nodes.
- \sim 10 TB of disk on each pool node.
- Intend to dual home the system to connect up to the UKLight network (no experience yet).

Imperial College

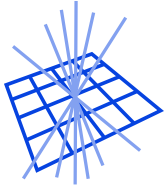
- 1 admin node and 2 pool nodes (running RHEL4 x86-64).
- 6TB of disk.
- Site firewall, but no machine firewall.
- Small pools filled up quickly: how to resolve?



GridPP
UK Computing for Particle Physics

Future problems at Tier-2's

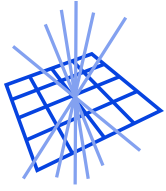
- Dual homing (GridPP have already done some work on this).



GridPP
UK Computing for Particle Physics

Future problems at Tier-2's

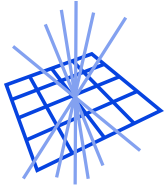
- Dual homing (GridPP have already done some work on this).
- HowTo upgrade to latest dCache release without breaking system?



GridPP
UK Computing for Particle Physics

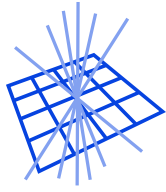
Future problems at Tier-2's

- Dual homing (GridPP have already done some work on this).
- HowTo upgrade to latest dCache release without breaking system?
- Migrating Classic SE to dCache SRM?



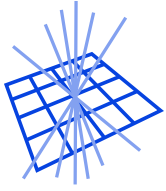
Future problems at Tier-2's

- Dual homing (GridPP have already done some work on this).
- HowTo upgrade to latest dCache release without breaking system?
- Migrating Classic SE to dCache SRM?
- Configuration and performance for sites with different/unusual hardware resources.
 - i.e. how will dCache behave/scale when deployed over a large number of pool nodes (Manchester).
 - unsupported OS on disk servers.



GridPP
UK Computing for Particle Physics

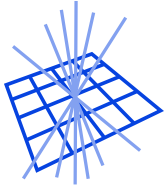
Other questions from GridPP



GridPP
UK Computing for Particle Physics

Other questions from GridPP

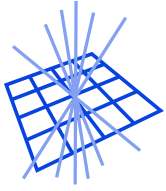
- What is the status of SRM 2.1 support in the dCache SRM?
 - File/space types (permanent, volatile, durable)
 - Space reservation



GridPP
UK Computing for Particle Physics

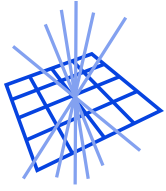
Other questions from GridPP

- What is the status of SRM 2.1 support in the dCache SRM?
 - File/space types (permanent, volatile, durable)
 - Space reservation
- Will a future dCache release support VOMs?



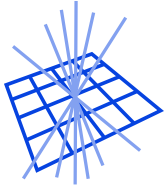
Other questions from GridPP

- What is the status of SRM 2.1 support in the dCache SRM?
 - File/space types (permanent, volatile, durable)
 - Space reservation
- Will a future dCache release support VOMs?
- Can dCache be used yet without an HSM to provide permanent persistent storage? (resilient dCache?)



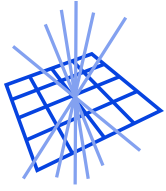
Other questions from GridPP

- What is the status of SRM 2.1 support in the dCache SRM?
 - File/space types (permanent, volatile, durable)
 - Space reservation
- Will a future dCache release support VOMs?
- Can dCache be used yet without an HSM to provide permanent persistent storage? (resilient dCache?)
- Which WSDL does dCache use?



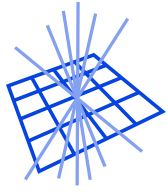
Other questions from GridPP

- What is the status of SRM 2.1 support in the dCache SRM?
 - File/space types (permanent, volatile, durable)
 - Space reservation
- Will a future dCache release support VOMs?
- Can dCache be used yet without an HSM to provide permanent persistent storage? (resilient dCache?)
- Which WSDL does dCache use?
- Is the server open source, and if so, where is the source? Some sites may want to perform security review of code before deployment.



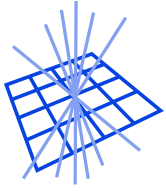
Other questions from GridPP

- The SRM client always try to use 10 parallel streams.
- dCache defaults to 10K writes. The write block size should be configurable in dCache.
- Excessive open java socket connections on the pool nodes `CLOSE_WAIT`
`10127 java.`
- Random SRM failures (because of database corruption?)



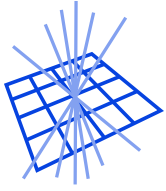
GridPP
UK Computing for Particle Physics

Conclusions + Outlook



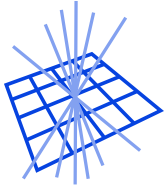
Conclusions + Outlook

- dCache is a mature product (has been successfully deployed at Tier-1).
- Tier-2 experience shows that it is now easy to deploy as a basic system (admin + small number of pool nodes):



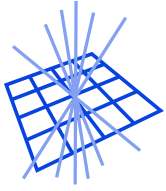
Conclusions + Outlook

- dCache is a mature product (has been successfully deployed at Tier-1).
- Tier-2 experience shows that it is now easy to deploy as a basic system (admin + small number of pool nodes):
 - Install has been greatly improved over past few months (YAIM).



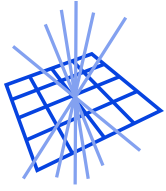
Conclusions + Outlook

- dCache is a mature product (has been successfully deployed at Tier-1).
- Tier-2 experience shows that it is now easy to deploy as a basic system (admin + small number of pool nodes):
 - Install has been greatly improved over past few months (YAIM).
 - GridPP knowledge base has grown significantly over this time.



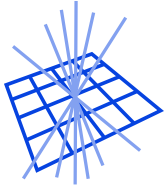
Conclusions + Outlook

- dCache is a mature product (has been successfully deployed at Tier-1).
- Tier-2 experience shows that it is now easy to deploy as a basic system (admin + small number of pool nodes):
 - Install has been greatly improved over past few months (YAIM).
 - GridPP knowledge base has grown significantly over this time.
 - Support structure already in place to help with SRM issues of all UK Tier-2 sites.

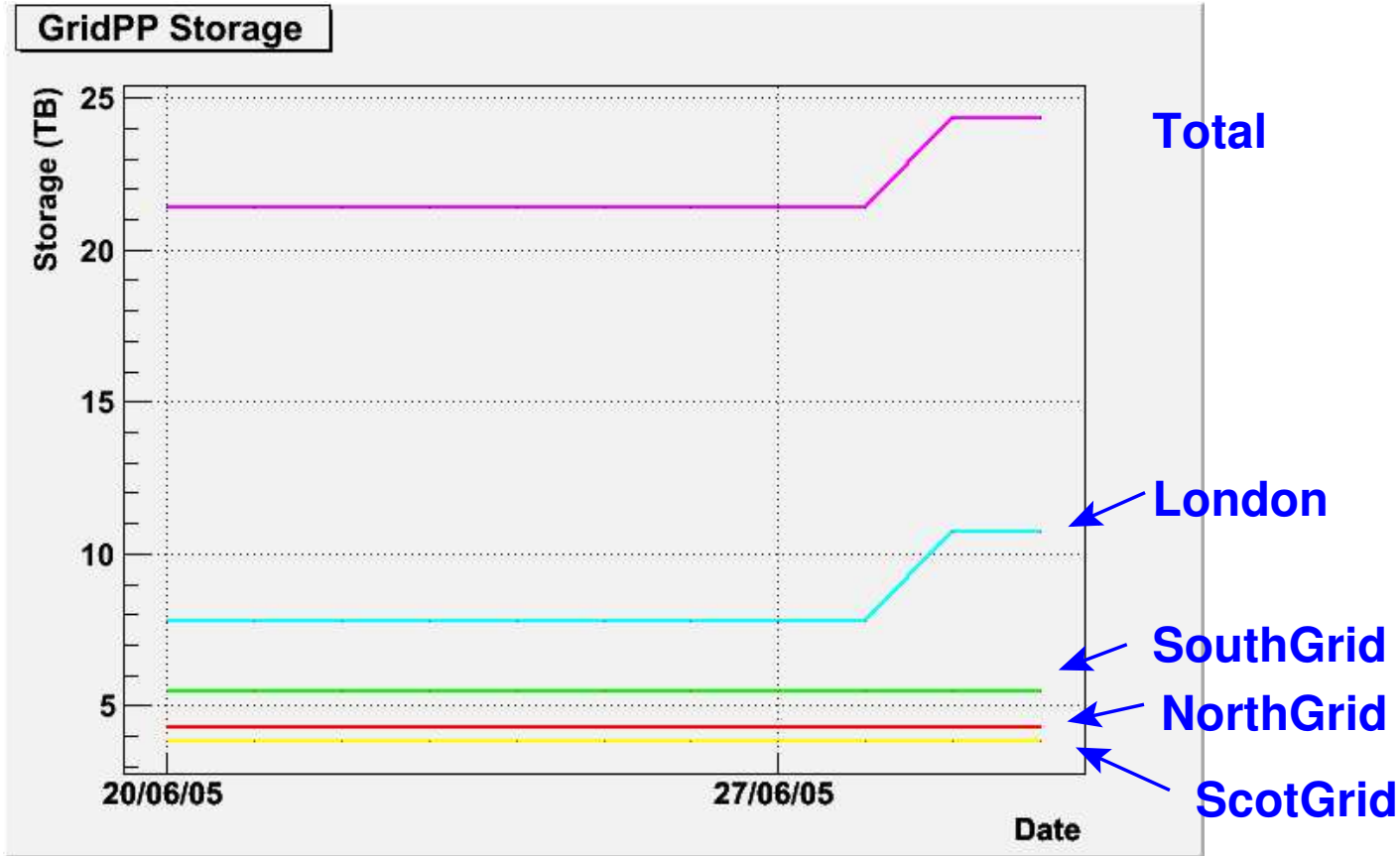


Conclusions + Outlook

- dCache is a mature product (has been successfully deployed at Tier-1).
- Tier-2 experience shows that it is now easy to deploy as a basic system (admin + small number of pool nodes):
 - Install has been greatly improved over past few months (YAIM).
 - GridPP knowledge base has grown significantly over this time.
 - Support structure already in place to help with SRM issues of all UK Tier-2 sites.
- But general system admin issues remain for Tier-2 sites. Most have limited manpower to spend on storage (compared to Tier-1's).
 - GridPP working on this issue (see wiki).



GridPP storage monitoring



<http://www.gridpp.ac.uk/storage/status/gridppDiscStatus.html>