

ScotGrid and the LCG

G. A. Cowan*, P. J. Clark, University of Edinburgh, UK

S. Thorn, National eScience Center, UK

L. Heck, M. Nelson, University of Durham, UK

J. Ferguson, D. Martin, F. Speirs, G. Stewart, University of Glasgow, UK



Department of Physics & Astronomy

Abstract

ScotGrid is a distributed Tier-2 computing centre formed as a collaboration between the Universities of Durham, Edinburgh and Glasgow, as part of the UK's national particle physics grid, GridPP. This paper describes ScotGrid's current resources by institute and how these were configured to enable participation in the LCG service challenges. The emphasis is placed on the management of storage resources detailing ScotGrid's role in the successful deployment of Storage Resource Managers (SRMs) at all GridPP Tier-2 sites. We outline future plans for ScotGrid, in particular the optimisation of available resources that is necessary to enhance the quality of service that is provided to Grid users.

Introduction

The UK's Grid for particle physics (GridPP) [1] started in 2001 with the aim of creating a computing grid that would meet the needs of particle physicists working on the next generation of particle physics experiments (i.e. the LHC). To meet this aim, participating institutions were organised into a set of Tier-2 centres according to their geographical location (see Figure 1). Here, we discuss the setup of ScotGrid with GridPP and the wider LCG project, emphasising its role in deployment, monitoring and testing of storage resources as are suitable for a Grid environment.

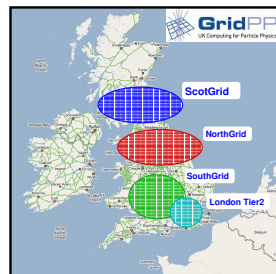


Figure 1: Federated Tier-2 centres within the GridPP collaboration.

Storage on the Grid

Due to the large volume of data that the LHC will produce, it is essential for the operation of the LCG that there is both sufficient storage capacity across the Grid and that the heterogeneous collection of storage resources is accessible to middleware applications via a common application protocol interface (API), the storage resource manager (SRM) [3]. GridPP is responsible for the deployment of SRM interfaces to storage at all UK Tier-2 centres involved in LCG. ScotGrid plays a key role in this deployment due to the work of personnel in the storage management and data management fields. GridPP has chosen to use two different middleware products to enable Tier-2s to manage their distributed collection of disk servers under a single namespace, with one method of accessing this space being through the SRM interface:

- **dCache** - Jointly developed by DESY and Fermilab to provide a highly configurable and scalable mechanism for managing a set of disk pools and tertiary storage. The SRM version 1 interface to dCache has been developed by Fermilab, enabling dCache to be used in a distributed Grid environment [4].
- **DPM** - Developed at CERN to provide a scalable method for managed storage of disk servers, particularly at Tier-2 sites. Provides an SRM v1 interface to the storage, as well as some SRM v2 functionality [5].

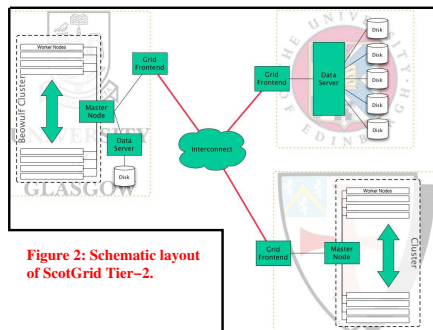


Figure 2: Schematic layout of ScotGrid Tier-2.

ScotGrid Resources

Edinburgh dCache

Edinburgh provides the second largest Tier-2 disk resource within GridPP, serving 31TB by use by LCG VOs. The storage is split such that 28TB is managed by dCache and 3TB managed by DPM, both providing SRM v1 interfaces for the VOs to use. The use of two storage resource managers to control access to the disk is key to Edinburgh being able to act as a knowledge base regarding storage within GridPP. The dCache setup can be seen in Figure 3. A single dCache administration node hosts the main dCache services (PNFS databases, logging, SRM interface) while the filesystems (dCache pools) are hosted on a single pool node. Using fibre channel connections, the pool node is attached to an IBM Dual FASTT900 22TB RAID (level 5) disk array which provides a level of resiliency against common malfunctions, protecting against data loss. An additional 3TB of storage is NFS mounted on the pool node from the University Storage Area Network (SAN).

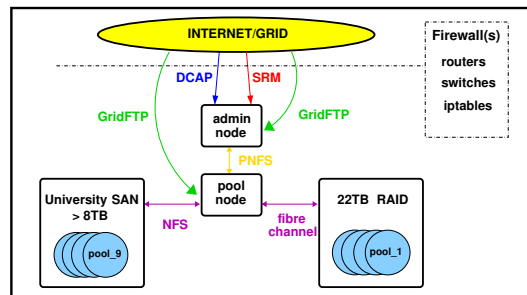
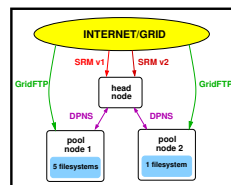


Figure 3: dCache setup at Edinburgh. See text for description.

Edinburgh DPM

Edinburgh also runs a production level DPM over two nodes, one of which was previously the GridFTP [6] server for the site, but has now been migrated into our DPM instance. Additional storage is NFS mounted from the SAN onto the node running the core DPM services (DPM namespace, SRM interfaces). Operating both dCache and DPM servers at production level places Edinburgh in a good position within GridPP, allowing for testing to be carried out, using at the interaction of the two SRM services.



Glasgow

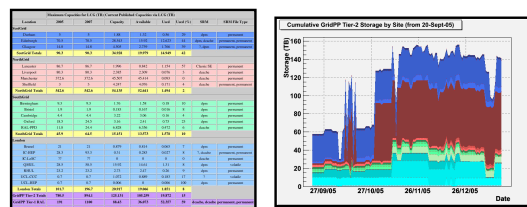
Currently providing 200 CPUs for use by LCG VOs and additional hosts for testing purposes. DPM manages the 5TB of RAID level 5 storage, spread across a two pool nodes, each allowing access to the disk via GridFTP. A third pool node with 3 filesystems is planned for deployment.

Figure 4: Glasgow DPM setup.

Durham

Provides 100 hosts, each with dual-2.2GHz Pentium 4 processors, 2GB of memory and 30GB local disk. Single 2TB filesystem (RAID level 5 disk) is accessible via the SRM interface provided by DPM, running on a single node.

Storage Deployment



Figures 5,6: Status of storage resources at each Tier-2 site within GridPP.

Monitoring

It is GridPP's aim to build a production level Grid that can be used in the next generation of particle physics experiments. Monitoring plays an essential part of this deployment as it allows a record to be kept of how sites are progressing and indicates when sites experience technical problems. This is particularly important in the era of SRMs since it is a relatively new technology of which Tier-2 sites have little experience. Through the work done at Edinburgh, ScotGrid has taken on the task of developing and operating a monitoring framework, based upon the information that sites already publish via the GLUE schema in the LCG information system [7]. By combining Perl, HTML, ROOT and C we are able to show the current status of SRM deployment in the UK while maintaining historical records which are useful for tracking long term progress. Figures 5 and 6 show the total storage capacity per Tier-2 site within GridPP on a daily basis [8]. Figure 6 clearly shows the gains that have been made in terms of available storage since the monitoring began. In particular, it should be noted that the GridPP Tier-2s have achieved greater than 150TB accessible via the SRM interface. What the Figure also shows are the fluctuations that occur in the availability to access sites storage, observed most frequently during upgrades to site SRMs.

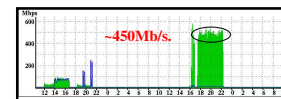


Figure 7: SC4 testing. Network monitoring during a 1TB data transfer from Edinburgh dCache to RAL. FTS was used to manage the transfer.

SRM Testing

As part of the work towards providing the reliable service that LHC requires, ScotGrid has been actively participating in the LCG Service Challenges (SC). These aim to test particular aspects of site setups in order to observe the interaction of middleware components, networking between sites and the individual site setups. Role of ScotGrid has become particularly important during SC3 and SC4 due to the presence of GridPP personnel specialising in data and storage management at the Universities of Glasgow and Edinburgh respectively. Although testing was initially done using the dCache SRM client, it is important to study how SRMs will interact with the File Transfer Service (FTS) component of the gLite suite of middleware, since this will be used to transfer files when the LCG goes into full production. An example of work done during the testing program can be seen in Figure 7, showing the achieved data transfer rate when using FTS to transfer 1TB of data from Edinburgh to the UK Tier-1 centre at Rutherford Appleton Laboratory. In addition to testing the Tier-1/Tier-2 data transfer channels, Tier-2/Tier2 transfers within ScotGrid have also been studied.

Future Work

The aim of SRM interoperation testing during the SC4 was to allow sites to understand the interactions between the LCG middleware components and their hardware setups. GridPP plans to use the SC4 period to act as a testbed for studying these interactions, finding bottlenecks and optimising available performance. Questions that will be asked:

- What is the most suitable pool filesystem to use with dCache and DPM?
- Which RAID configuration is best used for efficiency in reading/writing while also providing a suitable level of redundancy?
- What are the optimal kernel tuning parameters (sysctl) that can be used for dCache and DPM to maximise the file transfer rate?
- What alternative technologies are available to improve read and write to Tier-2 disk storage?

Using the GridPP wiki [9], ScotGrid will be able to disseminate the results of this testing to other sites, allowing GridPP to provide an optimised service to Grid users.

Conclusions

The expertise of ScotGrid in the fields of data and storage management has allowed us to take a leading role within GridPP in the deployment and testing of LCG middleware products which allow grid-transparent access to storage resources, specifically storage resource managers (SRMs). The use of a testing framework and monitoring has contributed to the successful deployment of an SRM interface to every GridPP Tier-2 site, while simultaneously feeding information back to the LCG middleware developers regarding future improvements. ScotGrid will continue in this role during the lifetime of the GridPP project, one of its aims now being the optimisation of the storage and data management middleware framework that it has helped deploy.

References

- 1 GridPP (<http://www.gridpp.ac.uk>)
- 2 ScotGrid: Scottish Grid Service (<http://www.scotgrid.ac.uk>)
- 3 SRM collaboration (<http://sdm.lbl.gov/srm-ug/in dex.html>)
- 4 dCache collaboration (<http://www.dcache.org>)
- 5 Disk Pool Manager (<http://www.mon.cern.ch/twiki/ LCG/DpmAdminGuide>)
- 6 The Globus Alliance: GridFTP (<http://www.globus.org/gri d3sfuare/data/gridftp.php>)
- 7 GLUE information Model (<http://www.fnfnforge.cern.ch/in fn.it/gua/nfom/ide/>)
- 8 GridPP storage monitoring (<http://www.gridpp.ac.uk/storage/status/gridppDis cStatus.html>)
- 9 GridPP wiki (<http://wiki.gridpp.ac.uk/wiki/ Grid.Storage>)

* g.cowan@ed.ac.uk