

## Deployment of a Large Scale, Parallel GridFTP Service with SRM interface

Submitted as Paper for presentation at a parallel session:

**Authors:** Derek Ross, Martin Bly, Andrew Sansum, Steve Traylen, Nick White

### Abstract

Many Grid R&D projects have deployed and tested high performance transfer tools on the wide area network [MBNG]. However until recently few projects have achieved similar data rates for production storage systems running for extended periods of time. The GRIDPP Tier-1 [GRIDPP, TIER1] facility at the Rutherford Appleton Laboratory has operated such a production service using dCache [DCACHE] since October 2004 and has also taken part in high data rate service challenges between CERN and RAL. This paper will describe the Tier-1's experience deploying and operating such a system, and will be relevant to many other projects irrespective of their choice of storage technology.

### Discussion

The Large Hadron Collider [LHC] is due to start operation at CERN in early 2007. It is expected to produce 12-14 Petabytes of data each year. The goal of the LCG project [LCG] is to meet these unprecedented computing needs by deploying a worldwide computational grid service, the backbone of which will be the Tier1 centres (5-10) worldwide. One of the principle roles of the Tier-1s will be to accept sustained data streams from the CERN Tier-0 of several Gigabits/sec while simultaneously servicing: international Tier-1s, national Tier-2s and the local CPU cluster.

The GRIDPP Tier-1 has run a production Grid based CPU service since July 2004, and has recently completed a successful deployment of a production quality Grid-enabled storage service.

A number of requirements needed to be met.

- To be compatible with LCG requirements the solution had to offer control access via an SRM [SRM] interface.
- To support wide area operation, Gridftp was required.
- To allow high bandwidth the solution had allow multiple servers to be deployed yet still present a single file-system name space.
- The solution had to have a long pedigree of reliable operation in a production environment.

Choice of software for such a storage system was limited; dCache was found to be the only likely candidate and deployment was commenced in July 2004 once a suitable release was available.

The initial production service was based on 4 disk servers. Load tests of this modest infrastructure showed that peak loads of least 3 Gbit/s could be achieved for locally generated traffic from multiple clients. Constraining factors appeared to be: the CPU load on the disk servers (about 10% per 100Mbit/s per server) and the transaction rate managed by the SRM front end (often 50% CPU busy). This service has now been in production operation for over 4 months and will soon be upgraded in both capacity and bandwidth. A number of experiments

are using this service as their main data repository – CMS in particular have now demonstrated transfer rates of real data of up to 500Mb/s over the SuperJanet 4 network, limited by access rates to tape at CERN.

To prepare for LHC start up in early 2007, the LCG project is running a series of “service challenges” increasing both the bandwidth and functionality required at the Tier-1s (and in the future at the Tier-2 centres). For LCG Service Challenge 2 (SC2) the GRIDPP Tier-1 deployed a second dCache (to avoid disruption to the production service and to allow independent performance tuning). To meet the anticipated high network loads UKLIGHT provisioned a 2Gbit layer 2 path to CERN which was dual attached to a dCache service consisting of a large parallel array of gridftp servers (Figure 1) and 4 disk servers each of 6TB capacity. Sustained transfer rates of 75MB/s were sustained over the service challenge and higher rates are expected to be achieved in early April after further performance tuning.

GRIDPP are now preparing for service challenge 3 (SC3) scheduled to run for the whole of July. The Tier-1 is expected to deliver 150MB/s to disk and a tape-tape component of at least 60MB/s.

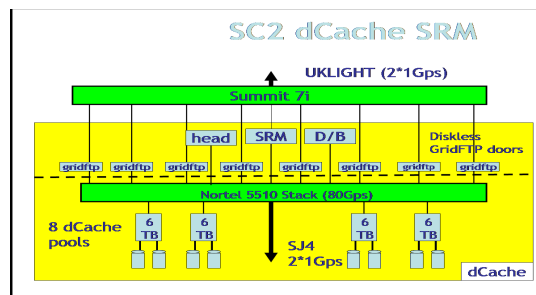


Figure 1: dCache configuration for SC2

## References

DCACHE: <http://www.dcache.org/>

GridPP – <http://www.gridpp.ac.uk/>

LHC – <http://lcg.web.cern.ch/LCG/>

MBNG - <http://www.mb-ng.net/>

SRM – <http://sdm.lbl.gov/srm-wg/>

TIER1 – <http://www.gridpp.ac.uk/tier1a/>