



Workshop Report

02-03 July 2009

Document identifier :	GridPP-Storage
Date:	08 July 2009
Version:	1
Document status:	DRAFT/Final
Author	J Jensen, B Davies

Introduction

This report is a writeup of the GridPP storage workshop held at RAL on 02-03 July 2009, following the hepsysman workshop on 0-1 July. It had three goals:

- Review the T2 outcome of STEP (a WLCG service challenge in June 2009, involving, in the UK, all experiments and most T2 sites);
- Reviewing state of the art in Grid storage middleware and management;
- Enabling sharing knowledge, and technical discussions.

Agenda

The agenda and the presentations can be found on the hepsysman website:

<http://hepwww.rl.ac.uk/sysman/June2009/agenda.html>

It was intended that all sites represented should give a site presentation, but some had signed up only in the last minute and were not on the agenda, or were added briefly before the meeting. In that case we had a brief oral report, with no slides.

Site representatives were asked to focus on:

- STEP *storage* experiences;
- Their site storage architecture and infrastructure, with permission to get technical, and also cover recent procurements, if any;
- Other relevant changes, e.g. changes of staff.

Occasionally a site's CPU and network setup is also pertinent to the use and performance of the SE, so to some extent WNs can be said to be part of the storage infrastructure. Some reports went a bit beyond the remit and talked about general STEP experiences (e.g. jobs). Finally, some sessions did not have slides (despite the "slides" on the agenda page.)

STEP report

[this section by Brian Davies]

Main outcome from STEP was that storage stability and setup is key to a good site performance.

For ATLAS, the main goals were

- Monte Carlo Production
- Monte Carlo Production continued with no obvious issues.
- Full Chain Data Distribution

UK sites passed distribution metrics. Transfers were 99% efficient at an average rate over the cloud being 838MB/s (only second to BNL) this compares to FZK which are a

similar size cloud for atlas who achieved 85% and 556MB/s. achieved over 1GB/s for sustained period. Sites received share of data based on datasets, not volume. This lead to some sites receiving more data than was planned for (and more than the disk allocation into respective service classes.) Hence planned overhead needs to be increased (was ~7% for). 224TB were expected to be transferred to the Tier2s, overhead accounted for up to 240TB; 246TB were moved.

ATLAS were also moving data other than STEP Data analysis data; this included "merged" files used as input for analysis jobs.

It was noted that RALLCG2 to UK Tier2 sites was five times rate of T1*-RALLCG2 traffic. This was confirmed by ATLAS central to be as expected.

Some work was needed during STEP to achieve this:

1. Increasing FTS channels
2. Re-adjusting a few SPACE token sizes

UK was one of the top clouds for both CMS and ATLAS.

In Data Distribution further work needs to be done on:

1. ATLAS data share distribution

This is to correct for increased capacity and increased reliabilities of previously under-achieving large pledged sites.) Large discrepancies between ratio of pledged disk space / ksi2k and deployed disk space/ viable job slot capacity.

2. Space Token size allocation.

(n.b Current 2009 pledges and current ATLAS central space token allocations equate to ATLAS filling DATADISK space token in ~3.75 weeks (6.5 if only one copy of AOD/DPD data is distributed within the UK c.f the two copies distributed within STEP09)

3. Check of ACL validity
4. Increased Usage of GROUPLINK space token.
5. Checking whether sites are publishing all space and whether or not this meets VO pledges.
6. Re-evaluation of FTS settings.
 - a. Review of TCP settings on storage Pools to allow large RTT WAN transfers to achieve a more acceptable rate

This led to FTS timeout changes and channel interventions during step.

- b. ATLAS now would like FTS channel setting for T0-T1 to be balanced wrt other T0-T1 channels at a level corresponding to expected shares and rates; leaving spare capacity to be added to sites acquire a backlog. CASTOR at RAL add caveat that this channel should always be limited to 30*Number of disk servers in T0Raw service class. This sets the RAL limit currently to 150. Current setting is much lower than this at the moment.

User Analysis Challenge: HammerCloud

User Analysis was a mixture of success and failures and a mixture of low and high cpu efficiency at Tier2s. With multiple bottleneck issues discovered. ATLAS centrally have also learnt much information from tests in terms of types of jobs and factors effect their success and avenues of improvement. Out of 10 Clouds, the UK was 4th in number of successes but only 7th in terms of efficiency.

Work will continue with subsequent HammerCloud Tests to find capabilities for efficient analysis. Measures included in study will be

1. file access protocol tests to find appropriate read ahead window/ file access protocol/ submission method (this last should not necessarily be an issue for the site but ATLAS need to find best mechanism.
2. Network configurations and bottleneck resolution
 - a. Is 1Gbps WAN connection enough for a large ATLAS site? At what point does a site need to thin about increasing WAN capacity.
 - b. What is the LAN capacity needed per job slot. Observed 200Mbps network usage for job assumed to require 2-4 MB/s in job reading.
3. IO wait on WN adjustments
4. SE Pool node configuration

Other VOs results from STEP09

CMS and LHCb results to come from STEP are:

1. A CMS issue to come from STEP09 was increased view for a need for USER accounting mechanism and quotaing for different groups (n.b ATLAS have solved this broadly with using Space Tokens, a move that CMS are reluctant to do. Also the DPM monitoring developed by Grieg et al should also provide User Accounting.
2. UKI-SOUTHGRID-RALPPD are asking whether ATLAS mind FTS channel returning to srmcp mode now that ATLAS use larger files so as to reduce internal disk to disk copies in WAN transfers.
3. CMS plan to increase using "non-dedicated GridPP Tier 2 sites" as "UK_T3_" sites in the CMS model
4. UKI-SOUTHGRID-RALPP coped with being a multi-VO Tier2 site with no obvious problems.
5. LHCb confirmed assurance that they do not really intend to use storage at Tier2s as a

Other Improvements needed are:

1. Dark data/ SE reconciliation seen as a need to clear space before data taking.

2. DPM tools for Load balancing WAN transfers/ Pool Draining / Hot File distribution felt as the tactical requirements for DPM sites
3. Usage of file:// protocol felt needed for STORM/Lustre sites rather than gsiftp:// protocol for LAN transfer/access.
4. Upgrade schedule/versions/functionality for DPM/dCache/Storm needed to obtain stable release before data taking.

Workshop Writeup

The workshop brought together Tier 2 storage admins from the UK and Ireland, some of whom are seasoned veterans, others are relatively new to storage or grid storage. This was also an excellent opportunity for Wahid Bhimji (who replaces Greig) to introduce himself to the community.

The agenda was very packed, so it was not always possible to do justice to each topic. The agenda had been reviewed several times at the storage meetings up to months in advance of the workshop, to ensure that it would remain relevant to attendees, and to ensure that all topics people wanted to discuss were at least raised. The agenda also tried to balance between site input, technical content, new technology/software, reviewing and planning ahead, and finally to leave room for discussion.

The workshop also brought in experts from Tier 1 and the National Grid Service.

Operations

Identifying Bottlenecks

The main issue is how to identify bottlenecks. If users expect site admins to tell them whether they are running effectively at the sites, admins need to be able to identify bottlenecks and have some understanding of the experiments' data models. CMS solve this by mainly running jobs at sites with local CMS people.

The SRM protocol is designed to be able to protect the underlying storage by asking the client to back off. When a file is opened for reading or writing, an unloaded storage system can return a TURL immediately, so the client doesn't have to call back later. StoRM is believed to support this "immediate" mode, but perhaps not all do – CASTOR doesn't. Also, bulk requests can be used more frequently.

Testing and Debugging

The workshop had a quick tutorial on SRM commands: how to run `srmcp` for testing, how to use `lcg-*` (and when to use one and not the other). We also had a very quick introduction (but not a demo) to S2, to see how S2 could be used to perform low level tests and even some useful things like querying space metadata.

Hot Files

Preparing extra copies of hot files can be a useful way to prevent a single disk pool or its network link from overloading. However, with DPM this must be done manually, and experiments can't always predict which files are "hot." We currently have no good solution for this.

Use of Virtualisation

Are virtual machines more likely to fail than physical ones?

Storage hardware

James Thorne from Tier1 gave a talk about verifying disk arrays. For a T2, when should the disk be returned to the vendor?

Kevin Haines from NGS gave a presentation on some of the SANs run by NGS. Kevin has set up multipath routing for them.

One question that was raised was which quality of storage is expected from the Tier2s? Specifically, if sites are expected to provide Replica storage (file can easily be replicated), can it be Raid 0? If sites are expected to provide Output storage (think job output, can be recreated but it's expensive), how is it done? Only Tier1s are expected to provide Custodial (which in practise is seen as synonymous with "tape").

Middleware

DPM

It was noted that many sites still have not upgraded to 1.7. Support for DPM pool nodes on SL5 is an issue: support is supposedly better in 1.7.

dCache

Manchester confirmed they will stop running dCache. This leaves two dCache sites in the UK (both of whom were present at the workshop): Imperial and RAL PP.

StoRM and BeStMan, Hadoop and Lustre

Chris Walker had evaluated StoRM; his evaluation (which was merged with the Lustre evaluation talk) can be found on the [agenda page](#).

Unfortunately our evaluation of BeStMan hadn't progressed as far as we'd planned, mostly due to services being unavailable during the machine room migration at RAL. An [interim report](#) was presented.

A quick poll showed most people interested in knowing more about StoRM (10) than BeStMan (4). For StoRM, we believe the pros include the ACL support (on GPFS) and that it comes with an information system; the cons could be that we need to clarify the support with INFN, and supporting local GridFTP access. For BeStMan, the pros include the good experiences in US with Hadoop, and the cons again the support: we'll need support from sites in the US if we are to run BeStMan.

Configuration and Deployment

To help larger sites manage their pool nodes, we need to look at configuration and deployment management. We had three presentations, cfengine – which has been presented at hepsysman before, puppet – which is being used for CASTOR at RAL, and Quattor. Both GridIreland and Tier1 use Quattor – and so do . GridIreland no longer need to use YAIM – which could be both good or bad, many admins like YAIM because it is relatively simple (shell). There will be a Quattor session in the week of

For puppet, it would be useful to have an overview so you can see how the clusters have been configured with particular profiles. Perhaps Nagios can be used to build such an overview.

With cfengine, it occasionally has to be run twice to get the output right. We discussed keeping the configurations in SVN. Apparently no one has done it yet? SVN is being used with puppet at Tier 1.

Infrastructure

Procurement

We had two presentations, one by Duncan Tooke from the NGS and the other by Matt Doidge from Lancaster, the latter without slides so not on the agenda page. The aim was to get the more experienced – or most recent – procurers to share their tips with the less experienced or less recent procurers.

Monitoring

The monitoring sessions intentionally had more room for discussion than for presentations. Firefox Nagios plugin potentially useful. Kashif's integration of testing and monitoring can also monitor other sites, and other VOs provided it holds credentials for those VOs.

Will monitoring integrate with what the NGI is running? NGS running INCA. No clear answer at this time.

Accounting

It looks like even more recent versions of StoRM are now publishing correctly; Chris' new instance at QMUL looks sane, the old one is mad (publishing over 300 exabytes).

Filesystems

We discussed various filesystems briefly: ext4, xfs, ZFS. A quick poll showed most interest in ext4 (14), some in ZFS (6). Everybody is interested in xfs: many (most?) sites are already running xfs, and a lot of work has been done testing performance with DPM.

Distributed filesystems

For Hadoop and Lustre, see evaluation under StoRM/BeStMan above.

A quick poll showed that more people were interested in Lustre (7), closely followed by NFS4 (6), GPFS (5), AFS (4), Hadoop-as-distributed-filesystem (4), and xrootd-as-a-filesystem (1.5) coming last.

NGS have been looking at AFS as a distributed filesystem between the sites (DEISA-style) for a while. Lustre may have security issues? For Hadoop, the concern is GridFTP access to the filesystem – the way Hadoop handles files in blocks etc may not be optimal with a non-native GridFTP server.

NFS4 is already supported in recent versions of dCache; the clients are currently not fully implemented (coming in recent Linux kernels). xrootd is partially implemented in dCache (which is interesting as it's the first implementation independent of SLAC, but

it will be very difficult to implement the rest of xrootd in dCache apparently). The xrootd in current versions of DPM is old.

Planning

How should sites invest in new storage?

Project

Milestones

We need to set new milestones for the project, since we have completed all the ones we've got. We discussed this in the meeting, and decided on a few new ones:

Summary

The workshop was considered useful and there was a lot of stuff crammed into just two days. This could be a disadvantage, but it is also useful to cover different topics, as long as they are relevant to Tier 2s' storage. Workshops are also a good opportunity for people to get to know each other – we have quite a few new people. We should maybe consider having workshops more frequently, e.g. annually.