

User Board Meeting, EVO, 12th September 2007

Present: *Matt Hodges, Alan Irving, Mike Kordosky, Dave Newbold (Associate Chair), Raja Nandakumar, Andrew Sansum, Glenn Patrick (Chair), Peter Richardson, Alexandre Sousa, Fergus Wilson.*

Apologies: *Jeremy Coles, Roger Jones, James Ferrando, Peter Hobson, Nick West*

1. PREVIOUS MINUTES

The previous minutes of the UB meeting held on 20 June 2007 were accepted without alteration. Concerning actions:

- a) *Provide costing for tape storage outside of robot.* It was reported that an initial quote had been sent to Nick West regarding SNO.
- b) *Experiments to list what is included in their disk requests from 2007/Q3.* Glenn had changed the quarterly request to break down the disk requirements.
- c) *Experiments to establish what compression (if any) is applied to data.* CMS uses gzip in ROOT using the default compression level of 1 (9 levels possible), but no further significant compression is seen on tape drives. LHCb also used the embedded compression in ROOT. Andrew Sansum commented that the Tier 1 planning is based on storing the experiment bytes, assuming a very conservative compression factor.
- d) *Experiments to provide 2008 profile for CPU/disk/tape.* This had been initiated by Glenn as part of the bidding process for this quarter. Most experiments had replied, but there were a few outstanding.

2. TIER 1 USAGE FOR 2007/Q3

Matt said that the occupancy had been at ~90% for June and July, but had dipped to 60% in August and was attributed to people being on holiday.

CPU efficiencies had improved in June and July, but had then dipped to 64% in August. This was probably due to the usual holiday factor with key people being on holiday. It was also noted that the low August efficiencies for LHCb (16.6%) and CMS (7.6%) were attributed to data access problems using dCache and Castor respectively. The BaBar efficiency (69.3%) was also lower than for earlier months where >90% was usual.

Batch Job Efficiencies	Jan 2007	Feb 2007	March 2007	April 2007	May 2007	June 2007	July 2007	Aug 2007
Efficiency (%)	63.8%	60.2%	53.4%	75.8%	71.5%	76.9%	82.7%	64.1%

Concerning fair-shares, the LHC experiments had largely underused their target allocations, but BaBar had taken advantage especially in July.

% of CPU allocation used	Jan 2007	Feb 2007	March 2007	April 2007	May 2007	June 2007	July 2007	Aug 2007
ALICE	103.5%	75.3%	72.3%	34.4%	31.0%	44.3%	18.5%	30.4%
ATLAS	25.7%	22.5%	11.9%	16.3%	44.5%	49.6%	18.8%	30.6%
CMS	65.1%	95.7%	72.6%	200.5%	13.8%	11.4%	0.3%	1.3%
LHCb	74.6%	35.2%	22.6%	21.8%	56.5%	185.0%	29.9%	16.9%
BaBar	60.9%	76.2%	83.6%	67.7%	148.4%	114.9%	267.9%	73.9%

Regarding disk, the deployment for ATLAS had not kept pace with the monthly allocations, but dCache disk had been provided just in time. LHCb had been fully provisioned in dCache and had an extra 27TB in Castor ready for migration. A disk server was reserved for UKQCD and was awaiting the fabric team. After discussion with Mark Lancaster, Glenn had set the CDF resources to zero (experiment just using FNAL resources).

On the tape front, CMS had removed a lot of data and repack should recover more space.

3. STATUS OF TIER 1 RESOURCES

Andrew reported on hardware procurements. GridPP now had twelve T10K tape drives after purchases in March and August. A Framework purchasing agreement was now in place to purchase additional tape drives. ~300TB of tape media had arrived in July. Another Framework agreement for tape media was almost in place and another ~800TB was expected in November.

The restricted EU tender for disk capacity was running ~6 weeks later than the schedule given at the last meeting. It was anticipated that delivery would be made in the Christmas/New Year window. This would bring the disk capacity to 1282TB.

A restricted EU tender for additional CPU capacity was about to go out with the expectation that it would be delivered in January and come on stream in March.

The difficulties of moving disk between Castor pools was noted and 5TB was considered the smallest quantum for increasing/decreasing Castor disk.

4. REVIEW OF EXPERIMENT ACTIVITIES/REQUIREMENTS

It was noted that as from 2007/Q4 (i.e. 1 October 2007) all storage requests for the User Board are assumed to be in Terabytes (i.e. 10^{12} bytes) and NOT Tebibytes (i.e. 2^{40} Bytes). Experiments had been advised of this when revising their bids below and the Schedule produced by Matt would also be based on Terabytes from 1 October 2007.

The Tier 1 Schedule can be found at <http://www.gridpp.ac.uk/tier1a/schedule.xls>.

4.1 2008 Planning

Glenn thanked those experiments who had provided their 2008 resource profiles. These can be found at: <http://www.gridpp.ac.uk/eb/UK2008/tier1expts2008v2.xls>
<http://www.gridpp.ac.uk/eb/UK2008/tier2expts2008v2.xls>

These preliminary plans show the anticipated evolution of the experiment profiles for the T1 and T2 centres through 2008. These are effectively constrained by the GridPP3 hardware planning for the year (funded at the 85% level). One challenge will be how to meet the expectations of legacy and non-LHC experiments. These profiles will need to be tuned as the final schedules and requirements are established for the full dress rehearsal tests and common computing readiness challenges (CCRC) in February and May, LHC turn-on and finally real data taking.

The intention is that all experiments updated these profiles as soon as they have new information and certainly for each UB meeting.

4.2 Experiment Allocations for 2007/Q4

The 2007/Q4 Tier 1 requirements had been captured alongside the 2008 planning.

The CPU requests for the quarter exceeded predicted capacity for each of the three months (headroom=-405, -508, -748 KSI2K) and so CPU would be over-allocated. This would be reviewed as necessary later in the quarter as the various data challenges progressed.

Disk was fine for October, but November and December looks tight and is dependant on progress with Castor migration and release of test space.

Tape space looks fine for 2007/Q4.

The Tier 1 allocations can be found at: <http://www.gridpp.ac.uk/eb/tier1allocs/tier1allocs2007Q4.xls>

5. SL4 SERVICE

Matt reported that the plan was now to proceed as fast as possible to a full SL4 service (queue=prod4) with another 40% of capacity to be moved later that week. There would be ~10% of capacity left in SL3 (queue=prod) then this would be moved one month after final announcement. This was acceptable to the experiments present.

It was noted by MINOS that there were some problems with the fair-share plots produced by Ganglia.

6. GRID ONLY ACCESS TO TIER 1 CENTRE

A previous discussion (January 2007) can be found at: <http://www.gridpp.ac.uk/eb/040107/nongridaccess.pdf>.

Matters had now progressed and we had to now start working towards restricting access to the Tier 1 over the next few months.

6.1 Future Interactive/Batch Access

Andrew reviewed the status of restricting the Tier 1 to Grid-only access. Comments from the previous UB had been fed back to the PMB, but the view was that migration to Grid only access should not be postponed any further. At the Tier 1 Board held on 15 August a range of views were put forward, but although Castor was the holding item (as a Grid storage solution) it was thought that CPU resources and interactive login should still move forwards. The main thrusts are:

- GridPP is a Grid project and there is a need to clear the decks for LHC start-up. There is no staffing for supporting "classic" services.
- Move from using the PBS "qsub" command to Grid submission with all CPU seen as a Grid resource.
- Remove direct logins to the T1 service.
- No local file editing and job submission on the T1.

The T1 Board was sympathetic to receiving exceptions to the above, but they would have to be agreed on a case-by-case basis. For example, VO managers might still need to use direct logins, Castor debugging might need some local job submissions and experiments might argue that they are a special case and still need to use a RAL UI.

MINOS made the point that they could only switch to Castor when it provided a stable service. They would like to keep their NFS area for as long as possible and certainly for two months beyond adopting Castor. It was emphasised that there was no intent to turn off NFS or xrootd, provided they were used via the Grid.

BaBar (Fergus) commented that a large number of changes were planned and they now had virtually no support.

CMS (Dave N) said that a number of requests were emerging for local logins at Tier 2 centres. As part of the disaster planning, we should also ensure that there was the fallback of retaining the ability to revert to local job submission, GRIDFTP and qsub.

6.2 AFS Future

Unfortunately, the future of the AFS service had also got entangled with the need to minimise the user registration system.

AFS was essential to UK operations for BaBar, in particular for code distribution to institutes. Fergus commented that if the RAL AFS cell closed, then BaBar would have to close down all UK computing.

LHCb (Raja) stated that AFS was useful, but not essential to the experiment.

After much discussion, Glenn made a proposal that there should be a staged approach instead of a single deadline for non-Grid access covering 6.1 and 6.2 as follows:

1. Experiments first need to establish that Castor is a stable Grid storage solution for their needs. The current timescale for deployment of castor 2.1.4 (i.e. the version with the increased functionality that most experiments demand) is end-October. Experiments are therefore requested to test Castor and to be in a position to make a statement about its adoption by **December 2007**.
2. Experiments start to test Grid User Interfaces (UI) on other computing clusters besides the Tier 1 centre and are in a position to report by **December 2007** on any outstanding issues or problems adopting this approach. Any use cases for special treatment (production management, VO hardship, etc) are identified at this point.
3. The PBS "qsub" command is withdrawn start of **January 2008**, obliging experiments to use Grid submission (e.g. "edg-job-submit"). Any further special cases are identified.
4. Direct logon to the Tier 1 is disabled (except for special cases) from **February 2008**. Access to the UI on the Tier 1 is restricted (except for special cases) from **February 2008**. This obliges experiments to submit work via the Grid from an external UI.
5. The "classic" User Registration Service on the Tier 1 is finally terminated (except for special cases) at the **end of March 2008**.
6. A separate, but related milestone is that final termination of the dCache storage system requires six months notice. Assuming this notice is given when Castor 2.1.4 is deployed, this is anticipated (but not confirmed) to happen in **April 2008**.
7. It was also proposed that the AFS service should be retained until the end of 2009 to at least meet the requirements of the BaBar.

Note

These recommendations were submitted to the Tier 1 Board on 13th September (see <http://www.gridpp.ac.uk/eb/ubnongridsep07.pdf>) and subsequently endorsed.

On 29 October, an email was circulated by Andrew to the UKHEPGRID list outlining this plan and requesting that consideration for any special exemptions be made known by email to support@gridpp.rl.ac.uk with copies to Andrew and Glenn.

Glenn advised BaBar on 2 October that the AFS service would continue until 2009.

7. dCACHE/ADS FUTURE

The existing dCache system works reasonably well. Once Castor 2.1.4 is installed and working, it should be easier to establish the final timescale for the termination of the dCache service. The LHCb, ATLAS and general instances of Castor should be upgraded to 2.1.4 at the end of October with the CMS instance following in November. The current projection for the end of dCache is therefore ~April 2008.

As indicated previously (see email to UB from Andrew dated 18.04.2007), RAL has no plans to terminate the Atlas DataStore (ADS) tape service, but GridPP will cease funding the service after March 2008. This means that any experiment that wishes to continue to use the service after this date will need to pay the storage costs from the experiment budget.

8. CURRENT STATUS OF CASTOR

It was reported that Castor basically works and that many of the current issues were about getting sufficient throughput. The separate instances had helped to insulate the experiments from one another's problems.

The difficulties of understanding Castor-SRM-FTS interactions and accessing the entire Castor software stack were raised.

The current Castor plans and a summary of the (weekly) technical weekly issues can be found at:
http://www.gridpp.ac.uk/wiki/RAL_Tier1_CASTOR_planning
http://www.gridpp.ac.uk/wiki/RAL_Tier1_CASTOR_Experiments_Technical_Issues

9. NEXT MEETING

The next meeting is scheduled for Tuesday 4th December at 15:00 using EVO.

ACTIONS

Action 1: *Submit UB recommendations regarding non-Grid access and AFS future to Tier 1 Board*
- Glenn.

Action 2: *Circulate final plans for non-Grid access to particle physics community (when known)*
- Andrew.

Action 3: *Establish whether Castor is stable Grid storage solution by December 2007*
- All experiments.

Action 4: *Start to test Grid User Interfaces (UI) on other computing clusters besides the Tier 1 Centre and be in a position to report at UB meeting in December 2007 - All experiments.*

Action 5: *Migrate to using Grid submission instead of the PBS "qsub" command by January 2008*
- All experiments.

Action 6: *Identify special cases for non-Grid access (direct logon, local job submission and UI) to Tier 1 Centre, by December if possible - All experiments.*

Action 7: *Provide updates to 2008 resource profiles for UK T1 and T2 for December UB meeting*
- All experiments.