



Version 1.0

Amendments history:

<i>Name</i>	<i>Area</i>	<i>Date</i>
<i>James Casey</i>	<i>First version</i>	<i>July 20</i>
<i>Kors Bos</i>	<i>Small overall changes</i>	<i>July 21</i>

Minutes of the meeting

CERN, 20th July 2005

Agenda: <http://agenda.cern.ch/fullAgenda.php?ida=a045324>

Minutes: James Casey

1. Introduction (Kors)

...

2. Service Challenge 3 (Jamie Shiers)

...

3. SC3 Status (James Casey)

James outlined the current state of SC3 – It is running at 500-600MB/s, which is the same as SC2, but well below the requirements for SC3.

He presented three questions which need to be addressed:

- How to tune long distance links?
- How to recover better from site outages
- How to move towards single stream transfers with low number of files

These need to be addressed with the specific sites that see the problems.

Jamie asked should we suspend transfers now, and start debugging. There was a general feeling we should continue on the schedule. Les asked how this work would be allocated, scheduled and tracked. James said that he would like agreement with specific sites today on what they would get involved in debugging.



4. Castor Status (Vlado Bahyl)

Vlado summarised the current deployment of CASTOR2 for the service challenge. Problems have been discovered during the deployment which led to a rapid deployment cycle of 2 releases per week. This seems to have settled down and the system is more stable now.

The focus has now turned onto monitoring and alarms. Operation procedures are being drafted.

A question was asked regarding lack of native support for gridftp? It currently leads to some performance problem (double scheduling). It will be implemented in the future.

The number of jobs running was 150 – is that a hard limit? No.

5. POOL LFC integration (David Smith)

David presented a report on the current status of the POOL LFC integration. He showed the features that are currently implemented, and integrated into the recent POOL 2.1.1 list.

Mirco asked had the CMS comments on the usage of LFC as a local catalog been addressed. David said they had been, and were in POOL 2.1.1. David Stickland commented that this just came out a few days ago, and they will do the testing now to validate this.

6. Experiment Support for Service Challenges (Patricia Mendez)

The GD EIS team provide support to the experiments, in terms of gathering their requirements, and feeding that into the middleware. They also act as shift personnel for SC3 second level support.

They have provided resources like UI and LFC/FTS pilot nodes for the experiments to test integration of their software frameworks with these systems.

David Stickland pointed that the box level monitoring system replicated some of the function of BOSS. Markus says this does replicate the tool, and provides functionality to any user. Mirco commented that BOSS also does more, since it parses the output and understands CMS specific output. He thought it didn't conflict with BOSS, and BOSS could be adapted.



David said that he felt it was typical that LCG implemented something new. Ian Bird responded that this came out of use cases directly from VOs that have met these problems.

7. GRIDVIEW Monitoring (Phool Chand)

Phool Chand presented the GRIDVIEW tool, which is visualization and monitoring tool for usage by the GOCs and ROCs. The first implementation phase focuses on gridftp monitoring for the Service Challenge.

8. Site Status Reports

ASCC (Eric Yen)

BNL (Bruce Gibbard)

Bruce reported that they've put a lot of effort in get the system stable. They now see that the effort has decreased once it's stable. They were able to get 1.5Gb/s with iperf using 160 streams, but they can't reproduce this with gridftp. Need > 50 concurrent transfers at 10 streams. Tier-2 participation is going slowly.

Andrew asked did they see crashing dCache pool nodes like the other sites. Bruce didn't think they had.

FNAL (Lothar Bauerdick)

Lothar gave a brief overview of their SC3 participation. Currently don't have a need for FTS, and haven't tested it yet. They are working on Globus RLS as local catalog at FNAL. Now have stopped on the SC3 setup to allow for FTS debugging, but are continuing a lot of work for Tier-2 setups.

GRIDKA (Andreas Heiss)

GRIDKA seem strange network behaviour, where their traffic ramps up and down in a 4 hour cycle – under investigation. Also a problem with dCache where disks get filled faster than they can delete files – will be fixed in next dCache release.

Reda asked about how big the pools were – 4TB – they see the same problem with their pools which are 1.7TB. The workaround is to delete continuously using a cron job.

IN2P3 (Fabio Hernandez)

Services installed (FTS + LFC, including LFC for egeode VO). Some problems getting to the throughput goal of 80MB/s using FTS – had demonstrated it using srmcp. Wiki is



LCG Grid Deployment Board Meeting



very useful for seeing what is going on. Fabio reported that their interaction with experiments seems to be very useful.

INFN (Luca dell'Agnello)

They have a FTS set up running 4 Tier-2 sites, getting data sent from the Tier2s to the Tier1. They have a separate setup now for phedex traffic.

Kors: Do you see any stability problems? We see some problems with FTS/SRM eating memory – to be investigated. We also had some problems due to infrastructure problems at the start. This meant we lost time.

NDGF (Tord Ekelof)

They use the DPM for data transfer. This has performed reasonably well, but still some issues appear for distributed deployment such as is required at NDGF.

David Stickland asked what the plans would be instead of DPM? There was no current plan for what to do when moving to tape, but there will be tape eventually. Les reminded us that we should keep the DPM simple and not overload it too much, but this use-case might come up at Tier2s.

James commented that we didn't focus on the distributed site issue, since for the first version, we assume a single administrative domain. If distributed Tier-2's are important, we can make the modifications for cross-admin domain DPMs.

PIC (Andreu Pacheco)

Good performance seen via Phedex transfers, trying to reproduce with FTS now.

RAL (Andrew Sansum)

Rate is down on what they've seen before SC3. They see some possible load related problems on the gridftp nodes. A lot of work has been done of getting Tier-2s involved in the transfers. They're now running throughput tests to Lancaster, Edinburgh and Imperial.

David asked was the FTS run in the same mode as at CERN i.e. push to T2. Yes, and they'll do some T2-T1 afterwards.

Andrew reported a funding issue where PPARC will suspend funding for more hardware until they see that the utilization rises.

Les asked to what effect the funding issue will affect SC3. Andrew thought he had enough to not impact the SC. It will come down to a resource allocation issue between the VOs. It's a new decision, and he expects better information in 3-4 weeks.



SARA/NIKHEF (Mark van der Sanden)

See 100-150MB/s. On average we see 30MB/s on the pool node but this is much lower than the IO throughput. The recounted their experience in setting up and running dCache for SC3.

Reda commented that the problem with double copies on a pool node is a dCache implementation issue, and not limited to FTS copies. Lothar mentioned that all the FNAL high performance copies worked in SC1/2/3 using srmcp, and this is one of the reasons that CMS continue to use phedex. Gavin mentioned that srmcp will be supported in the next version of FTS, scheduled for September.

TRIUMF (Reda Tafirout)

Reda reported on their progress to date. They've had some problems getting up to rate. They have a lot of network expertise, and think their network is tuned appropriately. They will carry on tests.

9. Experiment Reports

ALICE (Federico Carminati)

Federico reported on the plans for Alice involvement. He described the sample jobs which will be provided to test the system.

Les asked is the data to go to disk or tape – Federico said that given the dataset size, tape would be better. They think they can transparently zip it in ROOT. Mark asked how long we should keep the data – 1 year is the answer for the physics significant data (which will start in September)

ATLAS (Gilbert Poulard)

Jamie asked what would be used from the conditions database. Gilbert saw some problems during the Rome production, but these seem to be ok now. Andrea Valassi says they have sample data loads that they are using in testing the software. It's not clear that COOL deployment needs to be coupled to SC3.

CMS (Lothar Bauerdick)

CMS are fully committed to the SC3, and has important goals for the challenge. We realize that debugging is a priority, and need to schedule a point when SC3 is ready for CMS 'operation point' to implement realistic transfers to the regional centres.

Mirco asked what the plans are to integrate the high-level Phedex functionality on top of FTS – this is important for sites (like CNAF) which are multi-VO. Lothar responded that he would like Phedex to be a smaller software stack in the future, using less low-level tools. If higher-level file movement services, like FTS, are provided then Phedex will be



LCG Grid Deployment Board Meeting



integrated on top of it. He also mentioned that he didn't see a conflict since FTS is still maturing, and while it is getting ready, they can use the lower level tools in order to prove CMS goals, and not get in the way of the service challenge debugging.

Jamie made clear that we're not debugging FTS – it's not due to one component in particular, we're still trying to work out where in the stack the problems lie.

LHCb (Ricardo Graciani)

Minor modification of schedule- testing now up to end of September. Bulk data movement will start in October and online data transfers and processing from mid November.

10. Multi-VO support in FTS (Gavin McCance)

Gavin briefly outlined the new feature in the FTS to provide multi-VO scheduling (see slides).

Fabio wondered how he as a site admin would control all the traffic in/out of his site. James replied that currently you would have to modify all the channels sending data into you. We need to work on the topology of the connections. This isn't ideal

11. General Discussion on SC3 – open slot

Kors brought up the issue regarding the schedule – should we proceed straight away with tape. SARA and RAL expressed their agreement – it was generally agreed that this would happen tomorrow for sites who want to write to tape.

12. Preparations for SC4 (Jamie Shiers)

Jamie presented the status and preparations needed for SC4. Mirco pointed out that ROOT is not the only way to do analysis. Although the session at the ROOT workshop will be useful, there are other solutions. Jamie accepted the point.

Kors asked are the inter Tier-1 relationships defined. Not yet.

Lothar mentioned that for SC4 we need to think about CERN analysis facility.

13. AOB