LCG Services Report
July – September 2009

Jamie Shiers

12 October 2009

This quarterly report covers the third three months of 2009, a period during which issues revealed during the previous quarter's activities – including STEP'09 – were to have been addressed and preferably resolved. At the end of this quarter the LHC machine is close to complete cool-down and the experiments have, or are on the threshold of, entering full data taking, data processing *and analysis* mode. As opposed to previous such periods, this is expected to last for many months – perhaps as long as one year – with any inter-run periods being characterized by intensive processing, re-processing and analysis of data acquired to that date.

Whilst a number of the issues seen during STEP'09 have been resolved and with a lower frequency of major service incidents, the service still needs to survive the acid test of real data taking, expected with the LHC restart in the coming weeks.

Site Metrics

During the previous quarter a simple set of site metrics were established – repeated below – which can be used to check whether a site is providing an acceptable service. At the end of the last quarter, it was clear that a small number of sites were not consistently satisfying all of these metrics and an important milestone for this quarter was to understand and resolve the remaining issues. Whilst good progress has been made, it is premature to conclude that this is a solved problem and site visits continue into Q4 2009 to discuss these issues and their resolution. Hopefully, the next quarter will confirm that the steps taken have been successful but we will continue to monitor these criteria on a quarterly basis.

#	Metric
1	Site is providing (usable) resources that match those pledged & requested;
2	The services are running smoothly, pass the tests and meet reliability and availability targets;
3	"WLCG operations" metrics on handling scheduled and unscheduled service interruptions and degradations are met;
4	Site is meeting or exceeding metrics for "functional blocks".

Service Issues

Despite the further progress made during this quarter, a number of issues remain and were presented during the WLCG Operations sessions during the EGEE'09 conference.

Issue	Concern
Network	T0 – T1 well able to handle traffic that can be expected from normal data taking with plenty of headroom for recovery. Experience with physical cable cuts raises questions about whether sufficient redundancy exists. T1 – T1 traffic – less predictable (driven by re-processing) – actually dominates. Concerns about use of largely star network for this purpose. Tn – T2 traffic – likely to become a problem, as well internal T2 bandwidth
Storage	We still do not have our storage systems under control. Significant updates to CASTOR, dCache and FTS have been recommended by the software providers since STEP'09. Upgrade paths unclear, untested or both.
Data	Data access – particularly "chaotic" access patterns typical of analysis can be expected to cause problems – many sites configured for capacity, not optimized for many concurrent streams, random access etc.
Databases	A number of sites, including ASGC & RAL, have been unable to recover production databases from backups / recovery areas with major downtimes occurring as a result. A coordinated DB recovery validation exercise that is regularly tested should be considered to avoid such problems.
Compute	Still significant service-related issues with WMS and CE services. Long-standing scalability and robustness issues with LCG CE require advances in this area urgently, e.g. a production quality CREAM CE and WMS.
Users	Are we really ready to handle a significant increase in the number of (blissfully) grid-unaware users?

Summary of Main Service Incidents

The following table lists the main service incidents for which a "Service Incident Report" was produced. These incidents are typically characterized by a serious degradation or total loss of service for several hours or longer. Reporting of such incidents is now better, with the use of a template being increasingly adopted and the number of reports during this period was significantly lower than in previous quarters. This may reflect the reduced activity and service interventions during a period with numerous absences and vacations. Further details can be found in the weekly reports to the WLCG Management Board and on the WLCG Service Incident Report page.

Site	When	Issue
ASGC	28 Sep on	Prolonged database downtime (> 2 weeks) – report in preparation.
CERN	21 Sep	ATLAS DB Replication Tier0->Tier1 down 08:00 – 18:00
RAL	15 – 17 Sep	CASTOR Disk to Disk (D2D) transfers started failing during a planned upgrade to the CASTOR nameserver
FZK	7 – 16 Sep	ATLAS RAC – 3D Streams replication blocked then degraded
CERN	8 Sep	CASTOR DB problem affecting LHCb – 2 hours
CERN	5 Sep	CASTOR DB problem affecting LHCb – 2 hours
CERN	26 Aug	Public and production batch queues closed 18:40 – 23:30
ASGC	17 Jul	Power cut (~4 hours) – most services restarted
ATLAS	14 Jul	Degradation of performance of central catalogues for ~1 hour

Outlook for the remainder of 2009

The LHC restart is imminent and the experiments' activities that are foreseen for the immediate future mandate some changes in the overall WLCG Service Coordination. In particular, a hand-over meeting will now be held from the Service Coordinators-on-duty in out-going and in-coming weeks, and representation at the LHC Operations weekly meeting has been agreed. Given the increase in activities, it will be particularly important that experiments and sites are regularly represented – it being understood that time zones make this particularly difficult for both ASGC (in fact represented by someone located at CERN) and TRIUMF. Nevertheless, reports can always be sent via e-mail to wlcg-scod@cern.ch or added directly to the wiki. Efforts will also be made to keep the meetings as concise as possible: 15' – 20' maximum, with further discussion being handled in appropriate sub-groups. Regular WLCG Service Coordination meetings have been resumed – focusing primarily on CERN issues but also discussing potential deployment issues at other sites – on at least a monthly basis and more frequently if required.

Service changes – such as upgrades to data & storage services (CASTOR, dCache, FTS etc.) and / or those coming from ATLAS' evolving strategy for handling detector conditions data – will have to be followed carefully and planned in such a way as to minimize service interruptions. However, as we shall conclude, such changes are inevitable.

Summary and Conclusions

The WLCG service continues to deliver at a reasonably reliable and responsive level, with continued improvement as seen on the timescale of months. Well established procedures for responding to exceptions exist and are largely but (still) not always respected. Further improvement will clearly be iterative but is nevertheless required – the "site metric" as described above allows this to be measured quantitatively. A better, but not complete, understanding of how to handle larger service upgrades has been achieved, including the important realization that change is inevitable: it cannot be avoided; it needs to be planned and managed.