Worldwide LHC Computing Grid Project
Project Status Report
Resource Review Board – 12th Oct 2010

Ian Bird
Version 1.0
4th October 2010

This status report covers the period from April 2010 – September 2010. Further details on progress, planning and resources, including accounting and reliability data for CERN and the Tier 1 centres, and detailed quarterly progress reports, can be found in the documents linked to the LCG Planning Page on the web.

1. The WLCG Service

This reporting period represents the first 6 months of LHC data at a collision energy of 7 TeV. During this time the WLCG service has shown itself to be able to manage, process, and analyse the data as planned, thanks in large part to the significant testing activities and data challenges of the past several years. The net result of this success is the ability of physicists to have access to physics data within hours of the data being acquired and early publication of physics results. The key points to note from this early experience are:

- The resources are being used in accordance with the planned computing models, with slight variations in the light of experience;
- The networks are really being used at the scales anticipated;
- Data has been available at Tier 2s for analysis within hours;
- Tier 2 sites are used heavily for analysis, and provide the majority of the analysis power;
- The number of people doing analysis is significant: ~200 for each of LHCb and ALICE, and 500 or more for CMS and ATLAS;
- The Tier 0 and Tier 1 sites CPU capacity is so far less used than anticipated for a nominal year of data taking. Clearly this will change as the luminosity continues to increase. This has nevertheless allowed for multiple reprocessing of the data as the calibrations have improved, with direct benefit for the physics output.

First 6 months experience with data

During this period the luminosity delivered by the accelerator has been increasing dramatically, and the experiments have been acquiring data at rates close to those planned. To date several Petabytes of raw data have been archived at the Tier 0 and distributed. Figure 1 shows the data rates into and out of the Tier 0 Mass Storage system, Castor. As can be seen the system has been accepting data at a continuous rate of over 2 GB/s with peaks in excess of 5 GB/s, and has been simultaneously serving data at rates of 7 GB/s with peaks of 14 GB/s or more. Figure 2 shows the rate of writing data to tape (in GB/day). Individual tape capacities are 1 TB, thus the number of tapes written daily is on average 50 with

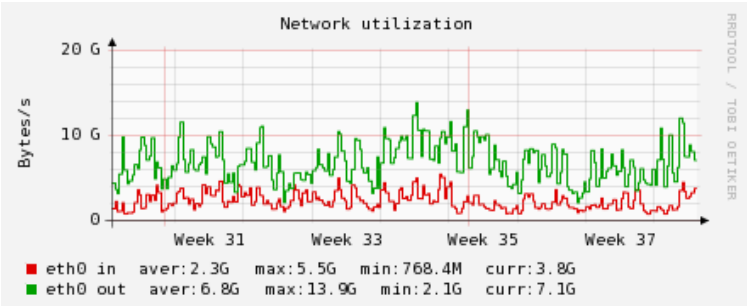


Figure 1: Castor data throughputs

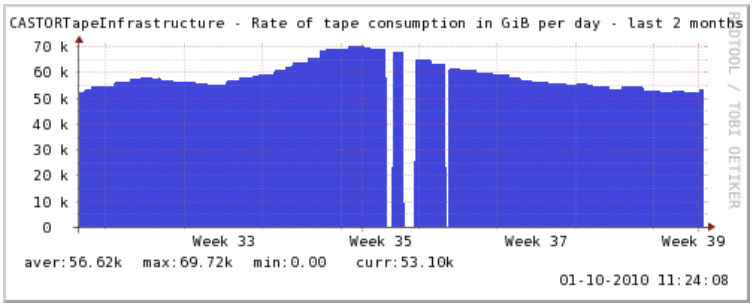


Figure 2: Rate of writing data to tape (GB/day) 50 k GB/day = 50 tapes/day

peaks up to 70. The total volume of data written to tape since data taking began in April is around 7 PB, as can be seen from Figure 3 which shows the total data volumes stored in the Tier 0 system.

The Castor system has been moving close to 1 PB of data per day during this time.

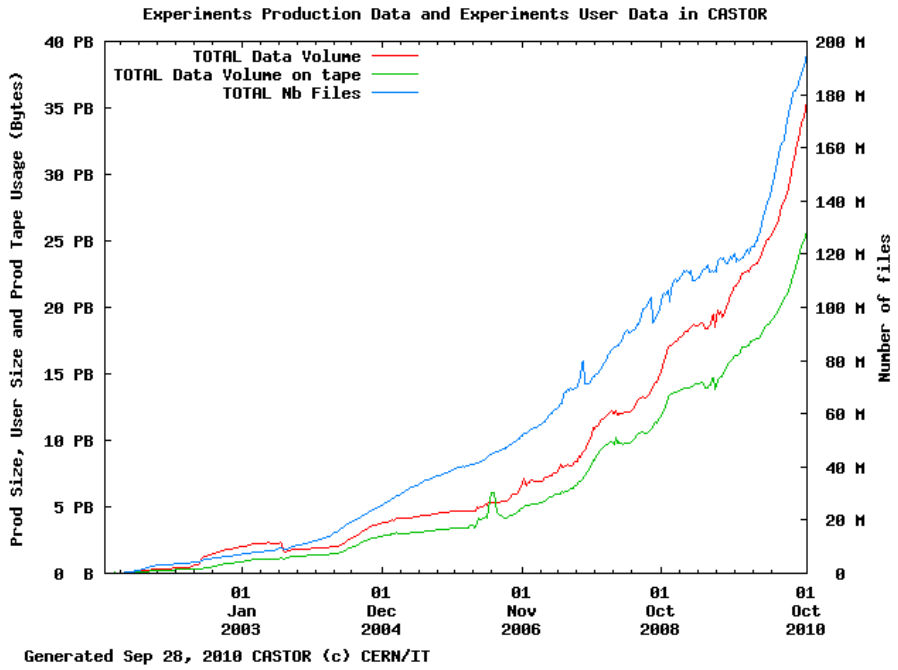


Figure 3: Total data volumes in Castor. ~7 PB have been written to tape since data taking started.

Data transfers

The data transfer rates out of the Tier 0 to the Tier 1s have been as anticipated on average, but with peaks in traffic that are significantly higher – up to 5 GB/s on occasion. This is illustrated in Figure 4 where the first plot shows the average rates over the period of April – September, while the second plot shows a period of a few days in August with significantly higher rates. The system has been shown to be able to handle these rates without problem.

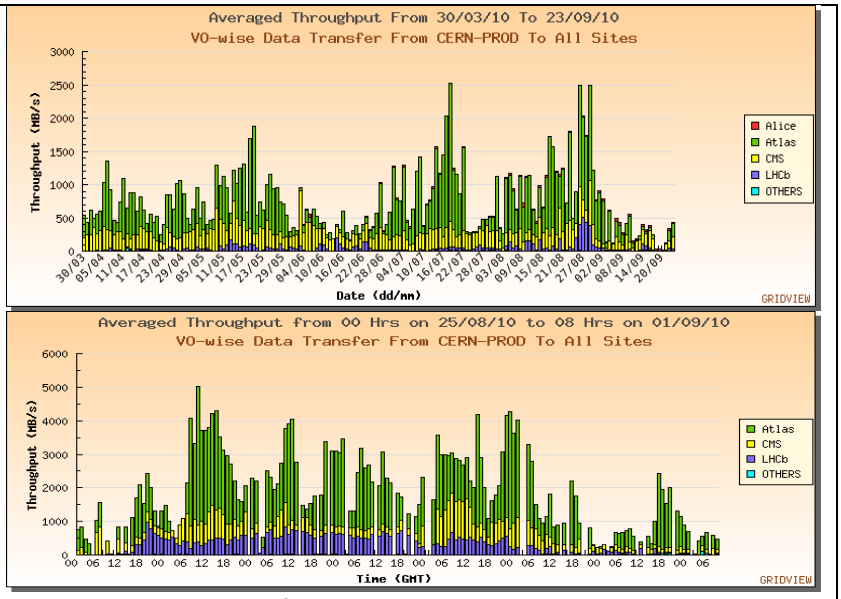


Figure 4: Tier 0 - Tier 1 transfer rates

The reliability of the OPN has been excellent, partly due to the reliability of the underlying networks, but also due to the redundancy now built in, since all Tier 1 sites now have at least 2 routes to CERN with a high bandwidth.

An example of data transfers throughout the entire infrastructure can be seen below for ATLAS, which shows the evolution of the rates, but also shows that within the system as a whole peak rates of some 10 GB/s are achieved.

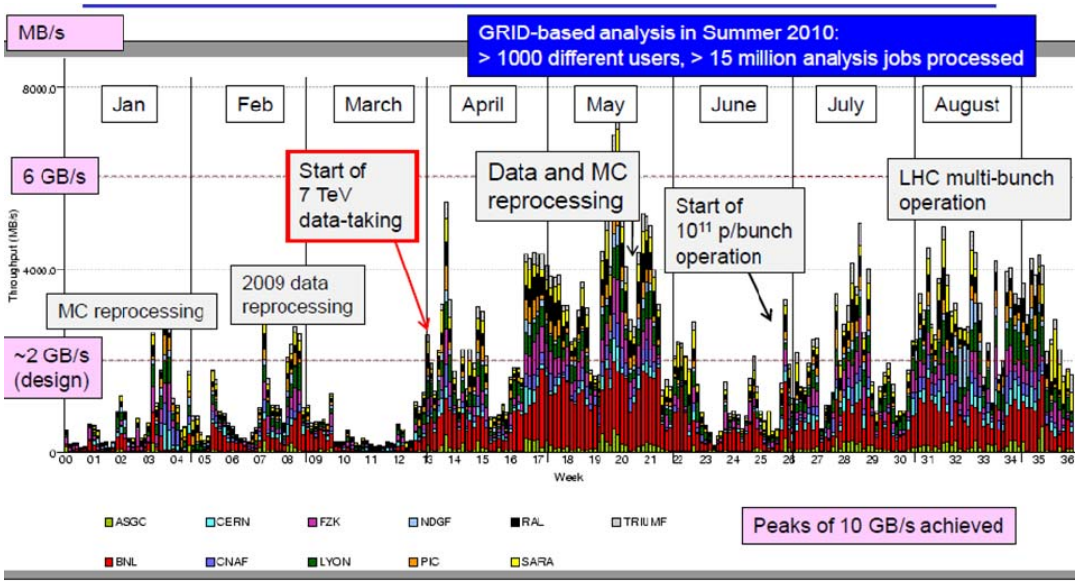


Figure 5: Example from ATLAS of data movement in the infrastructure.

Use of resources and workloads

The use of the WLCG system remains consistently high, with well in excess of 1M jobs per day being supported (Figure 6), and the equivalent of 100,000 CPU-days/day being delivered on average (Figure 7). These numbers are significant as they are the scale anticipated for the first years of data taking in the TDRs.

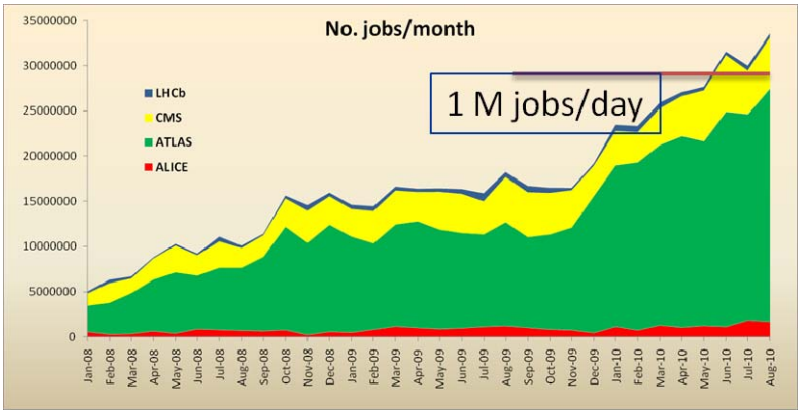


Figure 6: Number of jobs per month

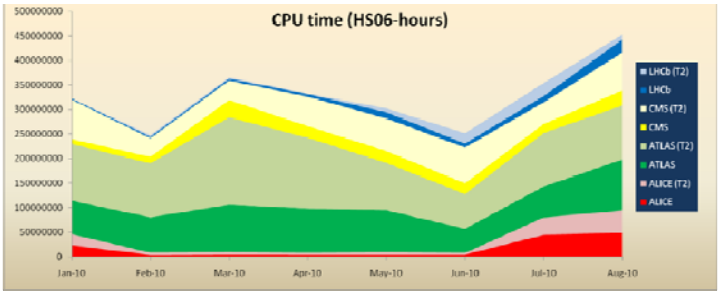


Figure 7: CPU time delivered (HS06-hours/month)

The pie-charts below (Figure 8) show an example of the distribution of CPU time between the various Tiers. For the Tier 2s they show how that was delivered by country. These shares are very much in line with the expectations based on pledges, and illustrate the significant use of the Tier 2s for real data analysis. The concern that CERN would be used for analysis rather than

the Tier 2s is not borne out.

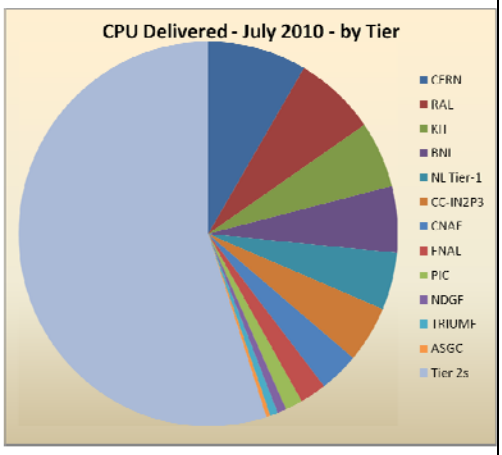
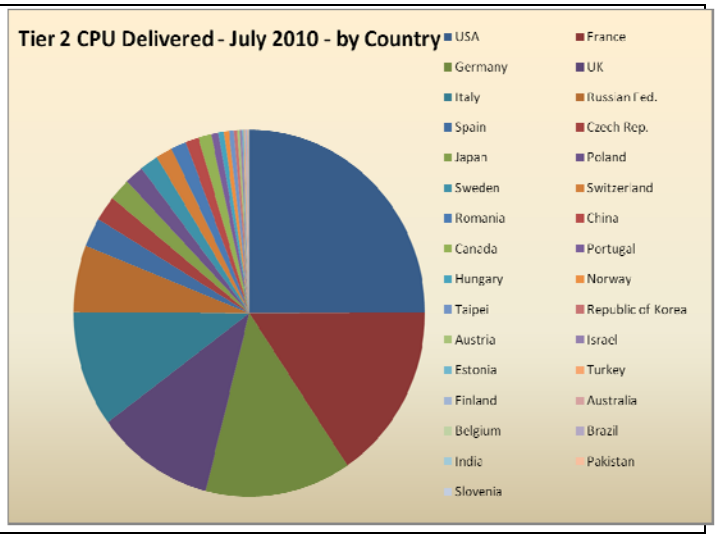


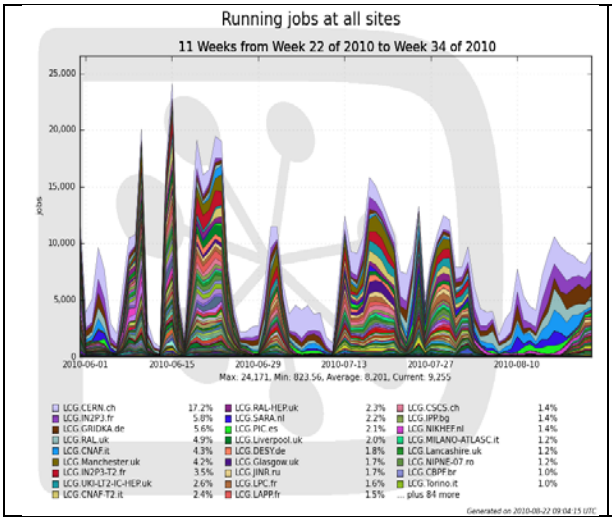
Figure 8: Distribution of CPU time in July



Analysis usage

The amount of work for analysis has remained at a high level almost from the beginning of data taking. The plots below (Figure 9) show examples for all 4 experiments. Significant numbers of users in each experiment are using the system on a continuous basis: around 200 each for ALICE and LHCb; with more than 500 users for CMS and ATLAS.

In addition to the production and analysis work for the LHC data, there are ongoing large scale simulation productions for all 4 experiments. These are essential to model the actual conditions of the accelerator and experiments with their known alignments etc.



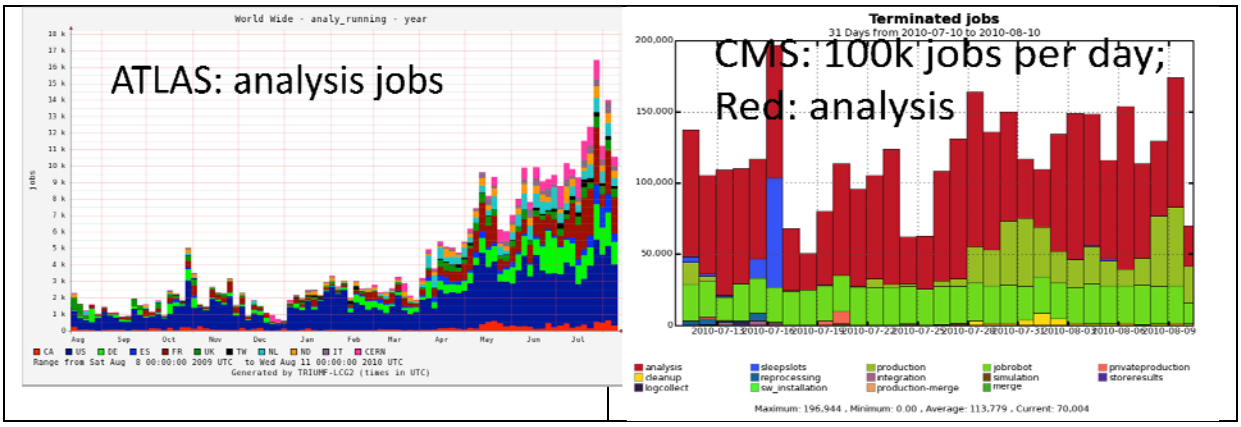


Figure 9: Examples of analysis usage by the experiments.

Ongoing Service status

As agreed, significant service interruptions require a documented follow up (Service Incident Report – SIR). The full list for this period (summarised in the Table below) including the full incident reports can be seen as a summary in each Quarterly Report, or consulted on line at https://twiki.cern.ch/twiki/bin/view/LCG/WLCGServiceIncidents. These are followed by the Management Board, with the goal being that lessons are learned and disseminated to other sites. It is still the case that often power and cooling problems are the cause of the significant downtimes.

Site	Date	Duration	Service	Impact
Q3 2010				
CERN	15 Sep	4 days	DB	Real time downstream was not set for LFC replication
CERN	13 Sep	1.5h	CMSR DB	Spontaneous reboots of nodes 2 and 4 of CMSR
SARA	August	>3weeks	DB	Replication for ATLAS conditions and LHCB conditions to SARA stopped
ASGC	31 Aug	4 days	DB	CASTOR outage due to STAGER DB problem
NL-T1	August	> week	DB	ATLAS NL-T1 cloud down, LHCb T1 site
CERN	23 Aug	35 h	Atlas conditions DB	ATLAS data streaming to Tier1 sites stopped
CERN	20 Aug	4h	CMS DB	Instability of node 3 and 4 of CMSR
CERN	9 Aug	16h	LHCb online	LHCBONR database unavailable
PIC	25 Jul	30h	CE	Service Degradation. Cooling problem causing about 50% of WNs to be shutdown (running jobs killed)
PIC	22 Jul	10h	SE	SRM service not available for ATLAS due to a problem with dCache pool costs configuration.
PIC	20 Jul	3h	CE	Computing Service not available after SD due to a wrong gridmapdir migration.
CERN	19 Jul	2h	several	Cooling failure in the vault
OSG/GOC	15 Jul	4h	GOC	GOC Service outage
CERN	13 Jul	1:30-9:15	CMS DB	Few short interruptions of replication of CMS data from

				online to offline
KIT	10 Jul	4h +	site	Outage of central and local services due to a cooling failure
NL-T1	5 Jul	1 week	SE	Reduced availability caused by data corruption
NDGF	14 Jul	16h	SE	srm.ndgf.org downtime followed by degradation
NDGF	8 Jul	3h	LFC	LFC downtime on lfc1.ndgf.org
KIT	5 Jul	18h	SE	CMS dCache SE down because of hardware failure
Q2 2010				
RAL	30 June		SE	1083 CMS files were lost.
CERN	29 June	4 h	CASTOR	CASTOR outage due to AFS
CERN	29 June	5 h	AFS	complete FC disk array - affected CASTOR and also LHC!
CERN	28 June	4+h	CASTOR	Log volume slowed down the Castor instances
ASGC	29 June	~ 15 hours	3D DB	ASGC didn't apply stream LCRs from central 3D DB for 15 hours
CERN	26 June	~50 min	ATLAS offline DB (ATLR)	9 Oracle services did not fail over properly after a node eviction
CERN	24,25 June	~10h	LHCb Streaming	Streaming of LHCb data to PIC was not working during 10 hours, streaming to other Tier1 sites not working for 40 minutes
CERN	22 June		CASTOR	LDAP high load caused CASTOR to become unresponsive
KIT/GridKa	12 June	~3:15h	CMS dCache	Service down
CERN	7 June	~3h	CREAM CE	Job submission failure
CERN	2 June		ATLAS and LHCb online and offline databases	Database access and quality of DB services compromised
CERN	31 May, 1 June		CMS online, LCGR and ATLAS offline databases	Database services unavailability during patching
CERN	26 May		CMS offline database	Hw failure affecting one node
PIC	21 May	19 hours	Whole site	Site power cut. Outage.
CERN	14 May	-	CASTOR	Data loss from incorrect recycling
GGUS	12 May	<=4.5 hours	.de domain	Domain does not exist
CERN-ASGC	12-15 May	-	LHCOPN	Reduced bandwidth

CNAF	28 and 29 April	9 hours & 12 hours	STORM	SRM blockage (hardware) followed by MCDISK full and STORM bug
IN2P3	26 Apr	17.5 hours	AFS	Distributed File System (AFS) crashed after server overload. Batch also affected.
IN2P3	24 Apr	17 hours	Batch	services location service stopped responding to requests blocking most batch system commands
IN2P3	20 Apr	9 hours & 5 days	Grid Downtime Notification	Grid downtime notifications were impossible after two consecutive incidents

However, in this period there were several occurrences where the downtimes could have been reduced, and lessons must be learned. This includes a significant, extended downtime for database services at the NL-T1 of 3 weeks. Here the database could not be restored from backup after a hardware problem. One lesson is that critical services such as this should have procedures that test either that the database restore process works, or to arrange for standby databases that can be brought into service.

In May there was a problem in Castor at the Tier 0 that caused a loss of data. A configuration error resulted in data being directed across all available tape pools instead of to the dedicated raw data pools. For ALICE, ATLAS, and CMS this included a pool where the tapes were recycled after a certain time. This pool had been previously used for data challenges where the data was only for testing. The result of this mis-configuration was that a number of files were lost on tapes that were recycled. For ATLAS and CMS the tapes had not been overwritten and could be fully recovered. The fall-back would have been to copy files back from the Tier 1s as foreseen in the computing models. Unfortunately for ALICE, 10k files were on tapes that were recycled, including 1700 files of 900 GeV data that were important for the ALICE analysis.

The following actions were taken:

- The underlying problem was addressed, and all recycle pools were completely removed;
- The software change procedures which allowed this error to happen were reviewed;
- An effort has been made to improve significantly the user-facing monitoring in Castor, as this could have helped realise such a problem;
- Tapes were sent to IBM and SUN for recovery. Eventually 97% of the critical ALICE 900 GeV data was recovered, although only about 50% of all lost files were recoverable;
- As this was essentially a procedural problem, a review of the Castor operations procedures (software development, deployment, and operation) together with experiment representatives and experts from other large laboratories was organised to verify that the actions taken were appropriate and that measures are in place to avoid such problems in future. The report from this review is anticipated this month.

However, from this problem it became clear that ALICE is exposed to a risk of data loss during heavy-ion running. As they acquire so much data in a short time, copying it all to Tier 1s synchronously is not possible as the ALICE Tier 1 sites do not have the aggregate capacity for that data rate. Thus for several weeks while the data is copied, ALICE has only a single copy of data. To mitigate the risk of losing part of that the Tier 0 has expanded the ALICE disk pools to ensure that two full copies of the raw data exist at CERN (one on tape and one copy on disk) until a second copy is available at the Tier 1 sites.

LHCOPN

Following a workshop to look at data management for the future it became clear that network connectivity would play a more and more important role. It was agreed to set up a working group to look at experiments' likely requirements for bandwidth in the future and to work with the OPN community to communicate these to the network providers. This group is now in place and has had several discussions with the OPN group. A first report is anticipated before the end of 2010.

It has become clear that the network operation is difficult to interact with when there are operational problems (for example a link not behaving as expected). An effort is required to address this together with the networking group to confirm the operations model and support procedures. This had been set up under the aegis of EGEE, and is perhaps an area where we are now suffering after the end of that project. This is being addressed with the OPN group. A related subject is the absence of user level monitoring to allow the WLCG or experiments to see ongoing network problems. Prototype monitoring displays have been produced, but DANTE did not feel able to support them. CERN and SARA have taken over the development of this monitoring.

2. Site Reliability

The reliabilities for the last 6 months for CERN and the Tier 1 sites are shown in Table 1. The relatively low figures for September are due to the downtimes necessary to address the security risk that was a potential danger for all sites. While in many cases the downtimes were unscheduled, the fast reactions had been urged by the WLCG Management Board, and undoubtedly mitigated the risk of far longer downtimes in the event of an actual security intrusion.

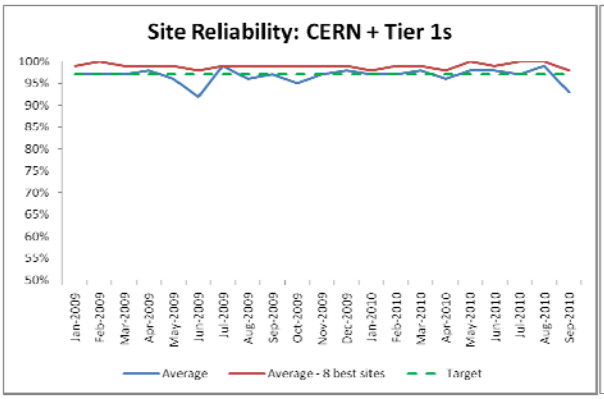
In addition to the general reliability testing reported in this table, the experiment-specific measurements are published monthly together with the general reports. The regular reporting for the Tier 2s now also provides an overall Tier 2 federation reliability which is the average of the sites in the federation weighted by the number of CPU reported in the information system where that number is published.

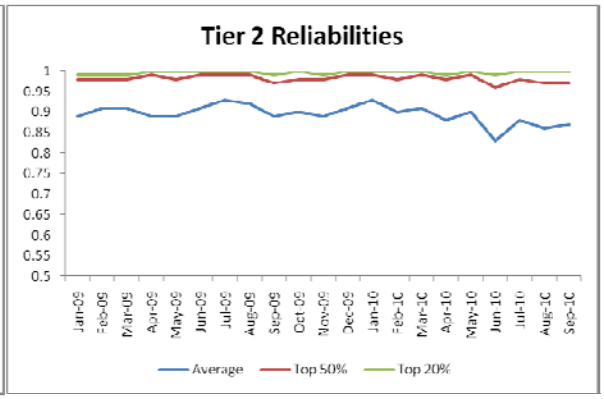
The figures below show the recent evolution of the reliabilities for the Tier 1 and Tier 2 sites. While this is now quite stable for the Tier 1 sites (with the exception of September as noted above), there is an apparent overall drop in the number of Tier 2 sites with a good level of reliability, although the top 20% of sites remain very reliable. It can be noted that the majority of resources are provided by these sites, but it is still important that the others sites address this.

All of the availability and reliability reports for all sites can be consulted at: http://lcg.web.cern.ch/LCG/reliability.htm.

Table 1: WLCG Tier0/1 Site Reliability

	Average of the 8 best sites (not always same 8)					
	Apr-10	May-10	Jun-10	Jul-10	Aug-10	Sep-10
	98	100	99	100	100	98
	Average of ALL Tier 0 and Tier 1 sites					
	Apr-10	May-10	Jun-10	Jul-10	Aug-10	Sep-10
	96	98	98	97	99	93
Detailed Monthly Site Reliability						
Site	Apr-10	May-10	Jun-10	Jul-10	Aug-10	Sep-10
CA-TRIUMF	98	98	97	100	100	100
CERN	99	100	100	100	100	96
DE-KIT (FZK)	94	100	100	93	99	88
ES-PIC	98	96	100	99	100	99
FR-CCIN2P3	95	100	100	100	100	81
IT-INFN-CNAF	98	96	96	99	99	92
NDGF	91	99	93	98	100	96
NL-T1	91	87	94	80	98	94
TW-ASGC	89	97	97	100	96	72
UK-T1-RAL	99	100	97	100	95	99
US-FNAL-CMS	99	100	100	100	99	100
US-T1-BNL	100	100	100	97	99	100
Target	*97*	*97*	*97*	*97*	*97*	*97*
Above Target (+ >90% Target)	7+5	9+2	9+3	10+1	10+2	5+5
Colours:	Green > Target	Orange > 90% Target	Red < 90% Target			





3. Applications Area

Software Infrastructure

Since April there have been 8 releases of the common software stack for LHC experiments. The changes include major upgrades of the AA developed projects (ROOT, POOL, COOL, CORAL, RELAX) and at least one upgrade of 26 external packages. Currently 14 different software platforms are being supported. The complete software stack was ported recently to Mac OSX 10.6 (snow leopard) on 32 and 64 bit architectures. Testing on the upcoming RHEL 6 is starting.

Python continues to be one of the cornerstones of AA and LHC software with an ever increasing need for new modules. This required a new deployment method to be implemented for the existing Python modules.

ROOT

A new development release, 5.27/06, has recently been made available. This is a preview of the production release scheduled on December 15th. Among the developments, it is worth to mention the work done in optimizing the I/O performance. With this version the default for streaming the content of STL containers has been changed from object-wise to member-wise. A reduction of about 10% CPU time has been observed.

Simulation

A new beta release of Geant4, 9.4-beta, was announced as scheduled on June 25th. Among the developments included are: a new revised Bertini cascade hadronic model; new physics-lists configurations utilising improved modelling for anti-baryons and hyperons (using the CHIPS hadronic model instead of parameterised LEP/HEP models); improvements in memory management for issues also reported by performance monitoring teams in ATLAS and CMS; a new geometrical shape, G4GenericTrap, requested by ALICE implementing an arbitrary trapezoid with up to 8 vertices standing on two parallel planes.

Two patch-releases, 9.2.p04 and 9.3.p02, were recently distributed, including also fixes for some issues affecting the releases currently in use by the LHC experiments.

Thanks to the contribution of Geant4 developers from Northeastern University, USA and the Geant4-SFT team, a new prototype of a multi-threaded Geant4 has been prepared, achieving excellent benchmarking results carried out on machines up to 32 cores (up to 90% efficiency) in collaboration with the CERN openlab team. This prototype, based on release 9.4, will soon be made publicly available, as a Beta version.

Good progress has been made in understanding the effects of model transition on energy resolution in Geant4. The simulation quality of new physics configurations combining the Fritiof model and the Bertini intra-nuclear cascade, has improved considerably, with rather positive feedback from validation studies made by ATLAS and CMS.

The regression testing suite for Monte Carlo Generators has been renewed and is now based on distributions and using tools like Rivet and the HepMC-analysis tool. Good progress has been also achieved in the realization of a new web interface for the tuning and validation of Generators against the latest available public data from the experiments. A new infrastructure for building Generators in the SPI nightly-build system is now in place, allowing also for a more automatic distribution of development builds to experiments.

Persistency framework

Three new releases of the Persistency Framework projects have been prepared since April for LHCb (LCG_58e), ATLAS (LCG_59a) and CMS (CORAL 2.3.12). One important issue in Oracle database services was also analyzed and solved with the help of the PF team. The motivation for LCG_58e (July 2010) was to urgently fix the problems reported by LHCb in a CORAL component caused by a clash in the implementation of the gssapi library provided by Globus that is incompatible with the version provided by the system. The LCG_59a release for ATLAS (July 2010) includes several enhancements in CORAL (functional fixes for Frontier, fixes of memory leaks for Oracle), COOL (support for a new relational schema with 'vector payload') and POOL (fixes in collections and in the test infrastructure). The CORAL 2.3.12 tag prepared for CMS includes the fix for a crash observed with optimized gcc 4.3 builds, which is now understood to be caused by a bug in the OracleAccess plugin.

Server-side process crashes (ORA-07445) triggered by COOL applications were observed on the ATLAS and LHCb databases, after the Oracle security updates were applied in June. A COOL-based stress test suite has been prepared that was successful in reproducing the issue and validating the last patch proposed by Oracle to fix the issue.

4. Level-1 Milestones

A full report on milestones and progress can be found on the WLCG web at http://lcg.web.cern.ch/LCG/milestones.htm. Several of these have been mentioned in sections above.

Progress on specific milestones includes:

- Support for multi-user pilot jobs. Deployment continues, with all of the Tier 1 sites now having installed glexec and either SCAS or ARGUS, or GUMS for some US sites. A significant number of Tier 2 sites have now also installed the software and are available to be used by the experiments. At the moment there is very little reason for the experiments to change their software to use glexec, particularly in the first few months of data taking where the pressure to produce physics has meant that such software changes have not been considered.
- CREAM CE deployment. More than half of the sites now have at least one instance of the CREAM CE installed. The intention is that this CE should replace the old LCG-CE as it will provide better scalability and performance. The adaptation of Condor-g required by ATLAS in order to submit to CREAM has now been verified and is undergoing stress testing. A number of stability and scalability issues have been addressed by the developers. More production work can start to move to the CREAM CEs in the next months.
- Data Management prototypes. Following the data management workshop in Amsterdam in June and the follow up in London in early July, a number of prototype projects were proposed. Progress on these will be followed in the GDB meetings during the Autumn, with a goal of deciding which of them should be pursued by the end of Q4 2010.
- Automated gathering of installed capacity data. The mechanism and process for data publication are agreed, and most sites publish data now. The published data for the Tier 0 and Tier 1 sites has now been validated (the milestone for Sep 30 for this was met). The milestone for correct Tier 2 site reporting is the end of October. The reporting is available through an online tool that allows one to view the experiment requirements, the pledges, as well as the installed capacity. Care should be taken in interpreting some of the installed capacity data as sites publish their real installed capacity which may be more than their pledges if they provide resources to other than LHC experiments.

5. Planning and Evolution

Tier 0

In recent months the schedule of the accelerator and other factors influencing the evolution of infrastructure and power requirements has evolved. It is now planned to not run the accelerator in 2012 and 2016, and significant efforts in the past few years have benefitted the total power needs. In addition, the anticipated reduction in budgets for the Tier 0 in the new Medium Term Plan will also have implications for the total power needs. Thus, compared with the original planning, it is expected that the need for significant additional power is delayed by a year to 2014.

While a more detailed re-analysis will be performed in the coming months to refine the strategy, several key points are clear:

- The plan for a new Tier 0 building at CERN has been cancelled for the moment;
- The ongoing work to upgrade the existing building from 2.9 to 3.5 MW usable power will continue, as this is essential, particularly because this upgrade will include an additional 600 kW of diesel-backed power to resolve the main critical issue at the moment;
- The 100 kW available in a local hosting facility will continue to provide an additional source of capacity for suitable services (e.g. databases), and could potentially be expanded.

The investigation of the possibility of remote hosting as a solution for the Tier 0 in the long term will continue to be pursued. Following the preliminary offer made by one host state, several others have expressed interest. The CERN Council was invited to solicit the submission of informal bids for Data Centre computing equipment co-hosting from their respective institutes or commercial consortia. These bids should be based on estimating the usable electrical power that would be available for hosting CERN computing equipment for an annual charge of 4M CHF. Once it has been determined that the proposals are economically interesting would more formal steps be proposed. The limit for such bids is the end of November 2010.

EGEE to EGI transition

Since the last meeting, the EGEE project has ended and the EGI-Inspire project has started. For the most part this transition has been transparent to WLCG, although at the detailed level there have been several points where EGI procedures or support has been missing – in particular those for dealing with sites outside of the formal EGI member countries. These issues are now being addressed by EGI.