Worldwide LHC Computing Grid Project
Overview Board – 27th October 2008
Status Report

Ian Bird
27th October 2008

1. MoU Signatures

No new MoU signatures have been received since the last Overview Board meeting in July. The signature of Austria is still outstanding, although there is an expectation that this may be signed imminently. The situation with Brazil has not changed, although there are some changes within the funding agency which may eventually resolve the question of who will sign the MoU. Otherwise all of the expected signatures have been done.

2. Quarterly Status Report

The most recent WLCG status and progress report covering the period from June 2008 to September 2008 is available from the Project Planning page.

3. Project Milestones

The status of the WLCG project Level 1 milestones on 26th June 2008 is shown in Table 2. Where a milestone has been re-scheduled the new date is given, but the colour coding shows the status with respect to the original target date.

24x7 support: With the experience gained during CCRC'08 and subsequent continued production running during the summer all sites now have a procedure and mechanisms in place to respond to operational problems out of hours. This milestone is thus now complete.

VOBox SLAs: In most cases the missing actions are awaiting the experiment formal sign-off on the SLA's that have been defined, although the definitions have been created in collaboration with the experiments. Only 2 sites are delayed in this process. At NDGF VOBoxes are only for ALICE, and the functions of the different boxes need to be well defined. At SARA/Nikhef the SLA is not finalized and will need to be agreed by both SARA and Nikhef managements.

Procurement of resources: The majority of the resources for 2008 have now been installed following the significant delays experienced by many sites as explained in previous reports. The remaining discrepancies at the end of September between 2008 pledges and installed capacities are as follows:

CPU:

- ASGC: 72% installed, expect to install the remainder in October
- CNAF: 57% installed, installation of remainder ongoing
- NL-T1: 88% installed

Disk:

- CERN: All capacity is on site, but problem with racks delays installation of full capacity
- ASGC: 300 TB missing (20%)
- BNL: 1 PB missing, anticipated in November with new machine room
- IN2P3: 700 TB missing, installation ongoing, together with 50% of 2009 capacity
- CNAF: 750 TB missing (60%), delivery complete and installation ongoing
- NDGF: 200 TB missing, procurement is complete, installation ongoing
- NL-T1: 1400 TB missing (56%): lack of available power and cooling; no new estimate yet, but not before 2009.

Table 1: WLCG Project Level 1 Milestones

17-Oct-08		WLCG High Level Milestones - 2007												
			Done (green)				Late < 1 month (orange)				Late > 1 month (red)			
ID	Date	Milestone	ASGC	CC IN2P3	CERN	DE-KIT	INFN CNAF	NDGF	PIC	RAL	SARA NIKHEF	TRIUMF	BNL	FNAL
		24x7 Support												
WLCG-07-01	Feb 2007	**24x7 Support Definition** Definition of the levels of support and rules to follow, depending on the issue/alarm												
WLCG-07-02	Apr 2007	**24x7 Support Tested** Support and operation scenarios tested via realistic alarms and situations				Apr 2008	June 2008							
WLCG-07-03	Jun 2007	**24x7 Support in Operations** The sites provides 24x7 support to users as standard operations				July 2008	June 2008		Apr 2008		July 2008			
		VOBoxes Support												
WLCG-07-04	Apr 2007	**VOBoxes SLA Defined** Sites propose and agree with the VO the level of support (upgrade, backup, restore, etc) of VOBoxes	Aug 2008	Aug 2008					Aug 2008					
WLCG-07-05	May 2007	**VOBoxes SLA Implemented** VOBoxes service implemented at the site according to the SLA	Aug 2008	Aug 2008				Mar 2008	Aug 2008		Apr 2008			
WLCG-07-05b	Jul 2007	**VOBoxes Support Accepted by the Experiments** VOBoxes support level agreed by the experiments	ALICE: n/a						n/a			n/a	n/a	n/a
			ATLAS					n/a	n/a					n/a
			CMS					n/a			n/a	n/a	n/a	
			LHCb: n/a					n/a				n/a	n/a	n/a
		VOMS Job Priorities												
		VOMS Milestones below suspended until the VOMS Working Group defines new milestones.												
WLCG-07-06b	Jun 2007	**New VOMS YAIM Release and Documentation** VOMS release and deployment. Documentation on how to configure VOMS for sites not using YAIM	EGEE-SA1											
WLCG-07-06	Apr 2007	**Job Priorities Available at Site** Mapping of the Job priorities on the batch software of the site completed and information published												
WLCG-07-07	Jun 2007	**Job Priorities of the VOs Implemented at Site** Configuration and maintenance of the jobs priorities as defined by the VOs. Job Priorities in use by the VOs.												
		Accounting												
WLCG-07-08	Mar 2007	**Accounting Data published in the APEL Repository** The site is publishing the accounting data in APEL. Monthly reports extracted from the APEL Repository.												
		MSS Main Storage Systems												
WLCG-07-25	Jun 2007	**CASTOR 2.1.3 in Production at CERN** MSS system supporting SRM 2.2 deployed in production at the site	CERN Tier-0											
WLCG-07-26	Nov 2007	**SRM: CASTOR 2.1.6 Tested and Accepted by the Experiments at all Sites** From the SRM Roll-Out Plan (SRM-16 to -19)	ALICE n/a			ATLAS Nov 2007			CMS Nov 2007			LHCb Nov 2007		
WLCG-07-27	Nov 2007	**SRM: dCache 1.8 Tested and Accepted by the Experiments** From the SRM Roll-Out Plan (SRM-16 to -19)	ALICE n/a			ATLAS Nov 2007			CMS Nov 2007			LHCb Nov 2007		
WLCG-07-30b	May 2008	**SRM Missing MoU Features Implemented** With full features agreed in the HEP MoU (srmCopy, etc).	CASTOR			DCache			DPM					
		CAF CERN Analysis Facility												
WLCG-07-40	Oct 2007	**Experiment provide the Test Setup for the CAF** Specification of the requirements and setup needed by each Experiment	ALICE			ATLAS May 2008			CMS June 2008			LHCb May 2008		
		WLCG High Level Milestones - 2008												
		OSG RSV Tests												
WLCG-08-01	May 2008	**RSV Tier-2 CE Tests Equivalent to SAM** Successful WLCG verification of OSG test equivalence of RSV tests to WLCG CE tests	OSG-RSV											
WLCG-08-01b	Jun 2008	**RSV Tier-2 SE Tests Equivalent to SAM** Successful WLCG verification of OSG test equivalence of RSV tests to WLCG SE tests	OSG-RSV											
WLCG-08-02	Jun 2008	**OSG Tier-2 Reliability Reported** OSG RSV information published in SAM and GOCDB databases. Reliability reports include OSG Tier-2 sites.	OSG-RSV											
		MSS/Tape Metrics												
WLCG-08-03	April 2008	**Tape Efficiency Metrics Published** Metrics are collected and published weekly				June 2008								
		Tier-1 Procurement												

ID	Date	Milestone		ASGC	CC IN2P3	CERN	DE-KIT	INFN CNAF	NDGF	PIC	RAL	SARA NIKHEF	TRIUMF	BNL	FNAL
WLCG-07-17	1 Apr 2008	**MoU 2008 Pledges Installed** To fulfill the agreement that all sites procure they MoU pledged by April of every year		Sept 2008	CPU OK May Disk Sep 08	Apr 2008	Apr 2008	CPU Jul 08 Disk Sept 08	CPU OK May Disk Sep 08	CPU OK May Disk	Apr 2008	Nov 2008	Apr 2008	CPU OK Disk Nov 08	CPU OK May Disk
WLCG-08-04	Sep 2008	**Sites Report on the Status of the MoU 2009 Procurement** Reporting whether is on track with the MoU pledges by April. Or which is the date when the pledges will be fulfilled.			Tender Sept Jan Install May	Tender Sept Dec Install Apr	Tender Sept Oct Install Apr	Tender Sept Install May	Tender Sept Install Apr	Tender Oct Install Apr	Tender CPU Sep Disk Oct	Tender Sept Install TBD	Tender CPU Disk Oct	Tender CPU Sep Disk Oct	Tender Sep Install Apr
WLCG-08-05	1 Apr 2009	**MoU 2009 Pledges Installed** To fulfill the agreement that all sites procure they MoU pledged by April of every year													
		glexec/Pilot Jobs													
WLCG-08-13	May 2008	**Glexec and Pilot Jobs Implemented at the Tier-1 Sites**		Pilot Jobs Working Group evaluating the Experiments frameworks (10 June 2008)											
		Tier-1 Sites Reliability - June 2008													
WLCG-08-06	Jun 2008	**Tier-1 Sites Reliability above 95%** Considering each Tier-0 and Tier-1 site	Jan 93%					70	92		92	57		91	
			Feb 93%					20	84			84		67	85
			Mar 93%					86		88				80	
			Apr 93%					76	84			90			92
			May 93%					88							
			June 95%					86							93
WLCG-08-07	Jun 2008	**Average of Best 8 Sites above 97%** Average of eight sites should reach a reliability above 97%		Averages of the 8 Best sites Jan-Jun 2008 Jan 96 - Feb 96 - Mar 96 - Apr 96 - May 98 - Jun 96											
		SAM VO-Specific Tests													
WLCG-08-08	Jun 2008	**VO-Specific SAM Tests in Place** With results included every month in the Site Availability Reports.		ALICE			ATLAS			CMS			LHCb		
		Tier-2 Federations MIlestones													
WLCG-08-09	Jun 2008	**Weighted Average Reliability of the Tier-2 Federation above 95%** Average of each Tier-2 Federation weighted accordinging to the sites pledges		See separated table of Tier-2 Federations.											
WLCG-08-10	Jun 2008	**Installed Capacity above 2008 Pledges of the Tier-2 Federation** Capacity at each Tier-2 Federation vs. the Federation's pledges		See separated table of Tier-2 Federations.											
		Tier-1 Sites Reliability - Dec 2008													
WLCG-08-11	Dec 2008	**Tier-1 Sites Reliability above 97%** Considering each Tier-0 and Tier-1 site	Jul 95%		94			79	88			91			
			Aug 95%						43						
			Sept 95%				90	82				94			
			Oct 95%												
			Nov 95%												
			Dec 97%												
WLCG-08-12	Dec 2008	**Average of ALL Tier-1 Sites above 98%** The average across ALL Tier-1 sites should reach a reliability above 97%													

In terms of capacity overall 98% of the CPU pledges are installed, while for disk this is only 76%. As pointed out in the last report this is a cause for concern and a milestone was added for the 2009 procurements as a checkpoint that they were under way in a timely manner. As can be seen in the milestone, the tendering processes are well advanced, but the anticipated delivery and installation schedules again leave little margin for error if the resources are to be fully installed for April 2009. In discussions in the Management Board recently, we have agreed that in future years we will propose a staged installation of disk capacity during the year to alleviate some of these problems. The details of this staging proposal will be discussed in the next months.

VO-Specific SAM tests: The VO-specific tests have been discussed several times in the MB in order that the sites understand what is being tested by each VO and which tests are used to determine the site availability for the VO. These VO-specific availabilities are now being regularly published and followed up by the sites and MB to validate the results. The goal is to have these published as reliable metrics by the end of the year. The underlying tests are already being used to raise alarms at the sites.

SAM testing for OSG: This is now complete. The tests in OSG have been agreed as equivalent to the set used in EGEE, and the publication of the results has been in place since August.

CERN CAF (Analysis Facility): This milestone is also now complete. The experiments have all described at a recent GDB) how they will use the CAF facilities (http://indico.cern.ch/conferenceDisplay.py?confId=20234).

Tape efficiency metrics: Tape metrics are now available for all sites with the exception of ASGC and CC-IN2P3. ASGC will implement the metrics following how this is done at CERN, while CC-IN2P3 are working on the publication of metrics from the HPSS system.

4. Applications Area

There were no major releases of the Application Area software during the last quarter. Experiments were preparing for beam and did not want any major change. On the other hand we have made substantial progress on porting the software stack (externals, and AA developed code) to other platforms such as gcc 4.3 and VC9. These ports are needed for next year's production releases. No new releases were produced for any of the Persistency Framework projects since the LCG_55 release in June 2008. Several new functionalities and performance optimizations have been prepared for COOL and are ready to be released in the upcoming COOL 2.6.0 (November 2008).

Progress was made in the development of the initial read-only implementation of the CORAL server, but a few functional and performance issues still need to be addressed before the software can be released. The addition of secure authentication and write functionalities have been postponed and rescheduled as separate milestones to be completed in 2009. A few enhancements of the POOL collections package have been prepared and will be released in Q4 2008.

The POOL project was reviewed in May 2008 to identify the steps to be taken to prepare POOL for the LHC start up and for its long term maintenance. All modules (except one that was dropped) are still used by at least one experiment and were moved to a new CVS repository.

As part of the general effort to improve ROOT documentation and tutorials the documentation of all graphics classes has being redesigned and completed. For PROOF, In addition to consolidation and debugging activities, the main developments during this quarter have been (i) the delivery of a new version of the XROOTD plug-in supporting automatic reconnections in the case of xrootd restarts; (ii) the implementation of a dynamic mechanism for "per-query" scheduling, where the master asks the scheduler the list of workers to start just before starting to process the query; and (iii) the support for memory consumption monitoring on all the workers as a function of the processing step.

A new version of the ROOT mathematical libraries has been released with improvements in the fitting and minimization. New common classes are now used for fitting all ROOT data objects, such as histograms and graphs, and various minimization algorithms can be used as independent plug-ins. The GUI fit editor has been as well improved by adding the support for multidimensional histograms and graphs.

For the simulation project during this quarter, two major achievements were made in Geant4: a preview release 9.2-Beta, in July, and a new patch to release 9.1 (9.1.p03), released in September. Most of the fixes introduced in 9.1.p03 are also part of 9.2-Beta, plus some more, including a fix in the field propagation causing a rare crash in ATLAS (about 2 per million events). Most fixes are the result of feedback received from LHC experiments and have been made promptly available to aid experiments in their production phase. ATLAS has reported great stability of their simulation based on 8.3.p02 (one failure every 500K events), and is now migrating to adopt release 9.1. The 9.2.-Beta

includes improvements in the FTF (Fritiof) hadronic model for pion incident interactions; alternative multiple-scattering models, and the first implementation of a GDML writer as part of the already existing Geant4 GDML plug-in module. The final public release 9.2 is expected for December.

5. The WLCG Service status

Since May the WLCG has continued to run a production level service with the same operational procedures in place. The workloads have continued to be significant as illustrated in the adjacent plots showing the job workload levels and data throughput rates. The experiments have been running simulations and collecting cosmic data at significant levels. The data transfers have continued to exercise the system and to continually validate the service.

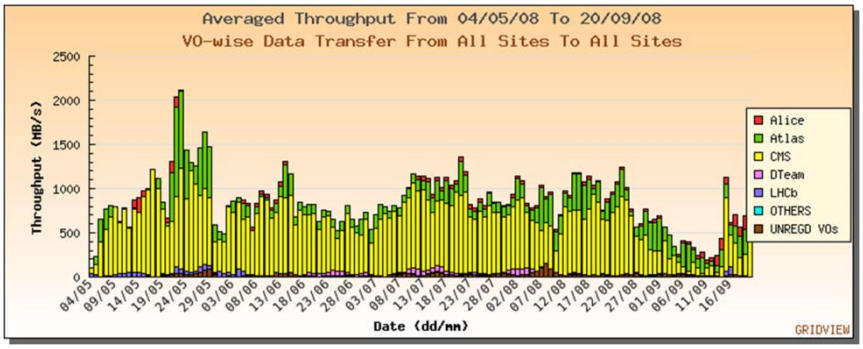


Figure 1: Data transfers Tier 0 - Tier 1

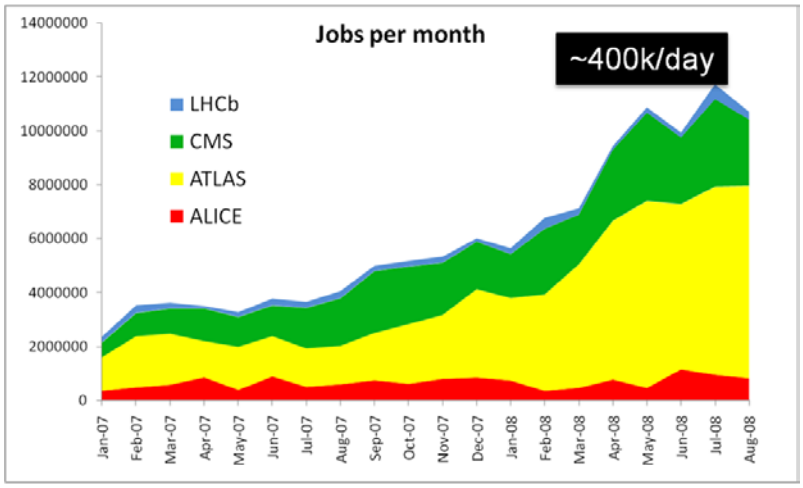


Figure 2: Job workloads over WLCG

The daily operations meetings track progress and follow up on problems. Summaries are reported weekly to the Management Board, and problems causing extended downtimes (more than a few hours) trigger post-mortems of the incidents. These are discussed in the MB and can lead to a management follow up to ensure the underlying problems are addressed. During the entire period non-disruptive updates to services and middleware took place, in a manner anticipated to be usual in the long term.

Of concern were noticeable absences of critical support staff during the summer and delays in fixing problems with experts being unavailable. This will have an impact on services if such a situation occurs during accelerator running. It is essential that Tier 1 sites address both service redundancy where feasible, and ensure that sufficient support staff are reachable.

The rate of incidents causing a post-mortem was about 4 per month. Not all could have been avoided by improved monitoring, and indeed close to half of them were due to power or cooling issues or unavoidable failures outside of our control (e.g. major failure of a network provider in Spain). Details can be seen in the Quarterly Report. The most severe service problem was the extended downtime of the Castor service at RAL, seemingly due to Oracle database problems and not the Castor software.

One of the main causes of instability were the storage services, with configuration problems causing additional instability. While regular phone conferences were organised it turned out that many sites with problems had not been joining the conferences, or reporting their problems. This situation has

improved, with additional agreement on improving the testing and release processes. It is vital that storage system managers report all the problems that they experience in order that this situation can improve. It is also apparent that many Tier 1 sites probably do not have enough staff working in the storage systems area. It is essential that this is reviewed and steps taken to improve the situation as this area is critical for successful data taking.

6. Planning during the LHC Shutdown

Following the testing earlier this year and the continued running during the summer, the WLCG service is regarded as being ready for accelerator data taking. Given the extended shutdown of the accelerator there may be a tendency to relax and slow down the service ramp up and provision. Such a move would be extremely detrimental to the quality of the service available in 2009. However, the next few months do give an opportunity to address some of the issues that had been postponed. The discussions in the Management Board in the past weeks have outlined a strategy of three items: resource procurement for 2009, middleware and service upgrades, and service validation.

Resource Procurement for 2009

The WLCG MB has agreed that with the information currently available to us and the present understanding of the accelerator schedule for 2009:

- The amount of data gathered in 2009 is likely to be at least at the level originally planned, with pressure to run for as long a period as possible this may be close to or exceed the amount originally anticipated in 2008 + 2009 together;
- The original planning meant that the capacity to be installed in 2009 was still close to a factor of 2 with respect to 2008 as part of the initial ramp up of WLCG capacity;
- Many procurement and acceptance problems arose in 2008 which meant that the 2008 capacities were very late in being installed; there is a grave concern that such problems will continue with the 2009 procurements;
- The 2009 procurement processes should have been well advanced by the time of the LHC problem in September.

The WLCG MB thus does not regard the present situation as a reason to delay the 2009 procurements, and we urge the sites and funding agencies to proceed as planned. It is essential that adequate resources are available to support the first years of LHC data taking.

Middleware and Service Upgrades

Since several software upgrades were postponed in anticipation of LHC start-up, we propose that the following changes main are addressed in the coming months:

- Deployment of FTS/SL4.This was postponed and will now be deployed. It has been tested extensively.
- Preparation of the middleware worker nodes for SL5. There is already a 1st installation at CERN, to be tested by experiments. The goal is to make this available in parallel to SL4.
- Introduction of glexec/SCAS to support multi-user pilot jobs via glexec. SCAS is currently in testing. This is essential for analysis use cases with pilot jobs.
- Introduction of the CREAM CE in a more aggressive way in parallel with the LCG-CE as the LCG-CE is known to have a limitation on the number of simultaneous different users. Today WMS submission to CREAM is missing, it will come with ICE, on a timescale of months.
- Fix problems in the WMS that limit the number of proxy delegations. This is available now.
- Availability of multiple parallel versions of client software.

The other important area of updates are related to the agreed programme of improvements in the SRM implementations already agreed, and scheduled to be available by the end of the year. It is important that this continue and these changes are deployed before the accelerator restarts.

Service Validation

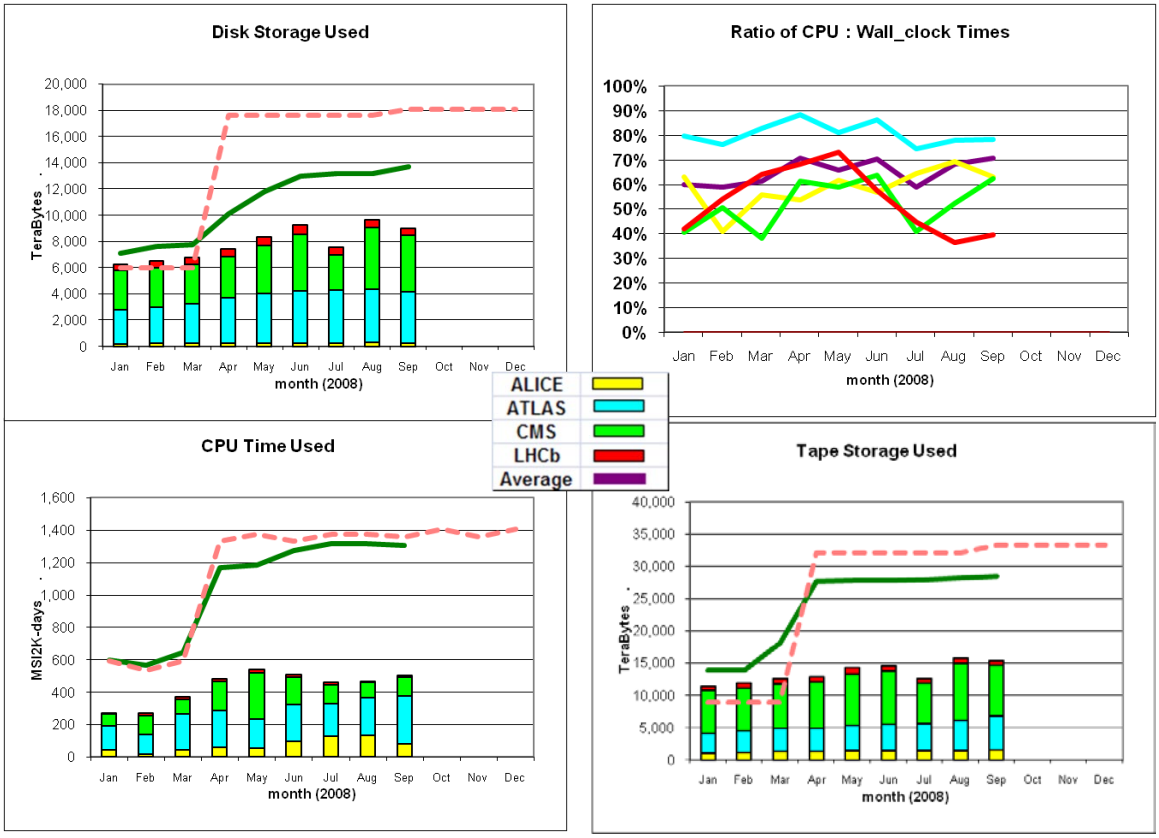
All experiments are continually running simulations, taking cosmic data, and doing specific tests (and have been since CCRC'08) at high workload levels. This will continue, and so a full-scale CCRC'09 in the same mode as 2008 is not regarded as useful. However, we will perform specific tests/validations:

- Service validation if software is changed/upgraded
- Specific tests (e.g. throughput) to ensure that no problems have been introduced
- Tests of functions not yet tested (e.g. Reprocessing/data recall at Tier 1s)

Details of the test programme will be discussed and agreed in the workshop already planned for November

7. Resource Accounting

The complete accounting reports for all WLCG sites are available directly from the EGEE accounting portal http://www3.egee.cesga.es/gridsite/accounting/CESGA/reports.html. At this URL the full reports may be seen under the "Tier 1" and "Tier 2" links. These reports are formalised, checked



and published monthly as a record of resources delivered to the experiments. The formal reports include only those sites that have an MoU agreement. In the accounting portal resource usage from all sites that publish data can be seen irrespective of the MoU agreements.

The full accounting reports for the period January to September 2008, and for the complete years 2006 and 2007 can be consulted at the WLCG web site. There are 120 Tier 2 sites now reporting. The figure below shows the situation for the Tier 1s and CERN for January – September. The solid

lines show the installed capacity relative to the pledge for 2008 (dashed line). The histograms show the used capacity from the accounting system. The ratios of CPU to wall-clock time have improved over the recent months – the average is now around 70%. While this is still cause for concern (the planning assumption was 85%) the general trend is improving.

Resource planning process

The current process foresees a 5-year planning cycle for experiment requirements and matching pledges from the funding agencies. Based on experience this now seems somewhat unrealistic from both points of view. There is little understanding of resource requirements 5 years out, and today those estimates are purely simple extrapolations of the previous year. In addition at the moment before any experience with data it is difficult to understand how to adjust the requirements. Similarly, for many funding agencies 5 years is too long a time scale. Thus in the next C-RRB we will propose a change in the MoU resource planning cycle from 5 years to 3.

The other problem is that the current pledge cycle is too late. By the Autumn RRB at which formally the pledges for the following year are approved, the procurements should ideally be already well under way in order to be able to provide equipment for the Spring. Similarly the Scrutiny process should ideally be looking at the next+1 year as input to the requirement and pledge process. These issues should be raised in the RRB.

Finally, as noted earlier, it is probably more realistic to split the storage (disk) procurements in 2 parts with Spring and late summer targets for installation. This proposal still needs some discussion, and would not affect the pledge cycle but would hopefully ease the work of the sites and allow for a slightly better cost optimisation.

Change of CPU accounting unit

The SI2K unit is now obsolete and benchmark values for new machines are not available in this unit. A working group has investigated alternatives and has recommended moving to the SPEC 2006 suite. A team is documenting the details of how this benchmark should be used, and will propose a conversion from the existing requirements and pledges to the new units without changing the existing agreements. For future procurements the new units will be used, and will require the vendors to run the specific benchmark agreed.

Reporting of installed CPU capacity and storage data

In order to completely understand the availability and use of resources the full set of information must include the pledges, the installed capacities, the used resources, the availability of the resources, and the efficiency of the usage. The gathering of data on the installed resources is presently incomplete. The gathering of the information of installed capacity for CPU and storage for the Tier 1 accounting reports is done manually. This is impractical for the Tier 2 sites. A team has been working on automating the gathering of installed CPU capacity and storage resources. Before this can be put into production a thorough validation of the gathered information is needed. This process is just beginning.

8. Site Reliability

The site reliability summary for CERN and the Tier 1 sites for the period October 2007 to September 2008 is given in Table 2. The site reliability target level was 91% until November 2007, 93% from December 2007, and 95% from June 2008.

Table 2: Reliability of CERN + Tier 1s

Average of the 8 best sites (not always the same 8)

Apr 08	May 08	Jun 08	Jul 08	Aug 08	Sept 08
95	98	98	98	99	99

Average of ALL Tier-0 and Tier-1 sites

Apr 08	May 08	Jun 08	Jul 08	Aug 08	Sept 08
91	96	96	95	96	96

Detailed Monthly Site Reliability

Site	Apr 08	May 08	Jun 08	Jul 08	Aug 08	Sept 08
CA-TRIUMF	96	98	98	98	99	96
CERN	95	100	98	99	100	100
DE-KIT (FZK)	95	97	98	96	99	90
ES-PIC	94	99	99	99	99	95
FR-CCIN2P3	98	97	96	94	95	98
IT-INFN-CNAF	76	88	86	79	99	82
NDGF	84	96	96	88	43	97
NL-T1(NIKHEF)	90	95	98	91	96	94
TW-ASGC	97	99	100	100	100	100
UK-T1-RAL	93	98	99	99	100	100
US-FNAL-CMS	92	96	93	100	99	100
US-T1-BNL	93	94	95	96	95	100
Target	*93*	*93*	*95*	*95*	*95*	*95*
Above Target (+ > 90% Target)	7 +3	11 +1	10 +1	8 +2	11 +1	9 +2

Colors: Green > Target, Orange > 90% Target, Red > 90% Target

The project target for the eight best sites was 93% until November and then 95%. The project target has regularly been achieved. The evolution of the reliabilities for the Tier 1 sites and CERN is shown in Figure 4. As noted in the previous report the reliabilities improved in May during CCRC'08 as sites were responding to problems through the agreed processes. The overall reliability has remained higher than previously as the experiments have continued to use the service at the same level, although during this time several problems have arisen in Tier 1 sites. However, we know that these generic tests do not always show the real problems that affect the experiments. For instance, during August RAL had serious problems in the database for Castor that meant that the ATLAS instance was unavailable for 10 days. This problem is not seen in the generic reliability measure as other services

were still available. This is one of the reasons why the VO-specific measures need to be regularly published in addition to these tests.

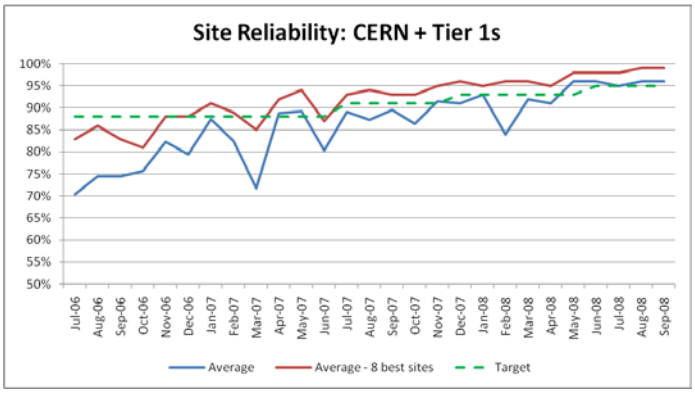


Figure 4: Site reliabilities - CERN + Tier 1s

Figure 5 shows the same measure for the Tier 2 sites. The best 50% (20%) of the sites are consistently more that 98% (95%) reliable although the average of all sites remains between 75-80%. Again there is a noticeable improvement in the overall reliability averaged over all sites since CCRC'08.

The full report of Tier 2 reliabilities is summarised by Tier 2 federation and by site is given on the web. There are some 120 sites now being reported on.

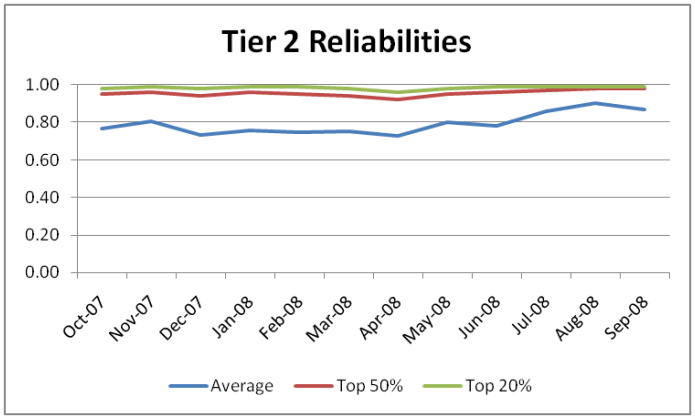


Figure 5: Tier 2 Reliabilities

The detailed comparisons of VO-specific and the general availability measures are in progress and it is anticipated that the project will start to report VO-specific reliabilities in the coming months.

Federations still not reporting include Norway and Sweden.

9. Tier 0 Capacity Planning

The capacity in the existing Computer Centre will run out in a few years (~2010) and the electrical capacity to the building cannot be further expanded beyond that which is currently foreseen (i.e. increasing the computing capacity from 2.5 to 2.9 MW). The assumption underlying this is that the resources needed by the experiments will increase by some 30% per year (once the initial ramp-up is complete in 2009). This increase is based on the estimates of the experiments at the time of the TDR (2005-2006), and is thought to be extremely conservative, especially in the light of experience in the past where computing capacity has increased far more rapidly. The estimate is also conservative in that it assumes that the experiments' software performs at the assumed levels, which is far from true today, and may not be their first priority in the next few years. In addition to the increase, it was assumed that 30% of the existing servers will be replaced every year – typically the systems have a 3 year warranty. Any future upgrade to the LHC itself and the experiments would increase the computing needs in addition to what is currently foreseen.

Strategy and Status

The strategy that is being pursued to address this has several components:

- Expand the capacity of the existing building as far as possible. This includes the upgrade to 2.9 MW total usable capacity excluding cooling, and the addition of water-cooled racks in the basement;

- Aggressive removal and replacement of older equipment with new systems with lower power consumption. This includes replacement at end of warranty (3 years) rather than leaving systems in place for 4 or 5 years as had been the case previously;
- Planning for a second Computer Centre to be built on the Prévessin site;
- Investigation of stop-gap solutions for the 18 months - 2 years between running out of power in the existing building and having a new building ready to install equipment.

The first 2 items and a better estimate of the evolution of the power of new systems mean that the present expectation is that the existing building may be able to handle the capacity for the Tier 0 and CAF until the end of 2010 rather than the beginning of 2010 as initially feared.

Planning for a new building

A full in-house design and construction of a new centre is not realistic. In the first part of this year discussions have been held with designers and builders of several CCs and visits to hosting companies and to CCs have also been organized in order to understand the possible strategies to follow.

Tendering directly for turn-key design and construction of a new CC is not considered wise, either within CERN or by external experts. Instead a four-phased process has been proposed:

- Request (many) conceptual designs (~26 companies contacted);
- Issue contracts to the 3-4 companies submitting the most interesting conceptual designs to develop an outline design;
- In-house, turn a selected outline design into a tender specification;
- Single tender for the detailed design and construction of the new CC.

The Price Enquiry for the conceptual and outline designs was sent out in June. The schedule for the above approach is tight, but if maintained, could lead to the negotiation of a detailed design and construction contract by end 2009. Based on an estimated subsequent detailed design phase of ~6 months and construction phase of ~18 months, a new CC could be available for equipment installation towards the end of 2011.

The actual construction cost will depend on the final design selected and this is one of the reasons to select 4 different companies to produce outline designs each based on different cooling concepts. The cooling is considered to be the most important design area. In the price enquiry, the companies were asked to optimize the 10-year cost of ownership. Hence, it might be that we select a design that is more expensive to build but actually due to the annual operation and maintenance costs gives a cheaper 10-year cost of ownership than a design that is initially cheaper to build.

A workshop was held with the four companies in late August to ensure that they all understood the CERN environment and boundary conditions. Interim meetings are scheduled with the companies to discuss progress with the outline design to ensure that these will be compatible with CERN's needs.

The date for the delivery of the outline design is the end of November. However, we hope to keep to the rest of the schedule but this will depend on how quickly we can turn the design into tender documents.

Stop-gap Solutions

Even if this aggressive schedule is maintained there will be at least 1 year, and perhaps 2 years during which the existing infrastructure will not be able to manage the anticipated load, and before a new building is ready. During this time we will have to find alternative locations in which to deploy the needed resources. Depending on the type of facility found this could be physics resources or

general infrastructure services. A query to the Tier 1 sites for available capacity was made earlier in the year. At the moment there is an ongoing discussion a facility that may have sufficient spare capacity for us. We expect more detailed discussions with them in the next month or so, but at the moment this is not moving as fast as we had hoped. The UK facility at RAL may also be a possibility, but this is less clear, and may only be available until 2011. In either of these cases we would anticipate locating Tier 0 (or CAF) resources at the remote facilities. Models for management and costs will need to be discussed and agreed.

Other alternatives would be finding a hosting company in the local area. At the moment there seems to be no obvious facility that could host the required capacity, and the costs would be very high. Nevertheless, this option will not be excluded. In this case however, since such a facility would be providing redundant power, and consequently it might make more sense to locate other types of service in such a facility - e.g. Database services where redundant power is essential.

10. Future Infrastructure Support

The EGI Design Study (EGI_DS) project has recently published a second draft of the blueprint document that describes the functions and organisation of a future European grid infrastructure built on National Grid Infrastructures (NGIs) coordinated with a central European organisation. However, from the WLCG point of view this document has some serious shortcomings:

- It is not clear exactly what is being proposed in terms of the roles and functions of the National and central organisations;
- There is no representation of the user communities, and no description of how those communities interact with the infrastructures;
- It is not clear how the present operational infrastructure upon which WLCG depends will evolve and appear in the future;
- Very few of the NGIs are as yet established, and so how they can support the WLCG sites is not clear, in particular during a transition period;
- Given the state of the current blueprint, it seems unlikely that there will be an organisation in place in time to take over the European grid infrastructure from EGEE in early 2010 with a managed transition process during the preceding year.

Taking these considerations into account it is important that WLCG has an appropriate plan to ensure that the European Tier 1 and Tier 2 sites are able to fulfil their MoU commitments in terms of the services today provided through EGEE. These sites will clearly be dependent upon the National Infrastructures and so it is essential that the WLCG Overview Board and Collaboration Board members ensure that their NGIs are working towards the appropriate goals. They must work together with the NGI representatives in the EGI_DS project, and at the funding agency level to achieve this.