LHC Computing Grid Project

Quarterly Status and Progress Reports

June 2008 – September 2008

20 October 2008

Table of Contents

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%													

		Completed / Cancelled High Level Milestones													
WLCG-07-09	Mar 2007	**3D Oracle Service in Production** Oracle Service in production, and certified by the Experiments													squid frontier
WLCG-07-10	May 2007	**3D Conditions DB in Production** Conditions DB in operations for ATLAS, CMS, and LHCb. Tested by the Experiments.													squid frontier
		Site Reliability - June 2007													
WLCG-07-12	Jun 2007	**Site Reliability above 91%** Considering each Tier-0 and Tier-1 site (Note: orange means > 90% of target)	Apr 88%												
			May 88%												
			Jun 91%												
			Jul 91%												
			Aug 91%												
			Sept 91%												
WLCG-07-13	Jun 2007	**Average of Best 8 Sites above 93%** Eight sites should reach a reliability above 93%		**Averages of the 8 Best sites Apr-Sept 2007 Apr 92 - May 94 - Jun 87 - Jul 93 - Aug 94 - Sept 93**											
		Procurement													

ID	Date	Milestone		ASGC	CC IN2P3	CERN	DE-KIT	INFN CNAF	NDGF	PIC	RAL	SARA NIKHEF	TRIUMF	BNL	FNAL
WLCG-07-16	1 Jul 2007	**MoU 2007 Pledges Installed** To fulfill the agreement that all sites procure the 2007 MoU pledged by July 2007													
		FTS 2.0													
WLCG-07-18	Jun 2007	**FTS 2.0 Tested and Accepted by the Experiments** In production at CERN and accepted tested by each Experiment		ALICE			ATLAS			CMS			LHCb		
WLCG-07-19	Jun 2007	**Multi-VO Tests Executed and Tested by the Experiments** Scheduled at CERN for last week of June		(will be part of CCRC in February and May 2008)											
WLCG-07-20	Sept 2007	**FTS 2.0 Deployed in Production** Installed and in production at each Tier-1 Site													
		BDII													
WLCG-07-21	Jun 2007	**BDII Guidelines Available** On how to install BDII on a separated node		EGEE - SA1 (not requesteded)											
WLCG-07-22	Jun 2007	**Top-Level BDII Installed at the Site** For each Tier-1 site													
		glexec													
WLCG-07-24	Jul 2007	**Decision on Usage of glexec and Guidelines to Follow**		GDB											
		MSS Main Storage Systems													
WLCG-07-28	Sept 2007	**Demonstrated Tier-0 Performance (Storage, DM)** Demonstration that the highest throughput (ATLAS 2008) can be reached.		CERN Tier-0											
WLCG-07-28b	Sept 2007	**Demonstrated Tier-0 Export to Tier-1 Sites** Demonstration that the highest throughput (ATLAS 2008) can be reached.		CERN Tier-0											
WLCG-07-29	Feb 2008	**SRM: CASTOR 2.1.6/dCache in Production at T1** From the SRM Roll-Out Plan (SRM-20 to -21a)													
WLCG-07-30	Dec 2007	**SRM Implementations with HEP MoU Features** With features agreed in HEP MoU (srmCopy, etc).		CASTOR			DCache			DPM					
		WN and UI													
WLCG-07-31	Jun 2007	**WN Installed in Production at the Tier-1 Sites** WN on SL4 installed on each Tier-1 site, with the configuration needed to use SL4 or SL3 nodes							n/a				n/a		
WLCG-07-32	Jun 2007	**UI Certification and Installation on the PPS Systems**		EGEE - SA1-PPS done: Jul 2007											
WLCG-07-33	Aug 2007	**UI Tested and Accepted by the Experiments**		ALICE			ATLAS			CMS			LHCb		
		xrootd													
WLCG-07-41	Jul 2007	**xrootd Interfaces Tested and Accepted by ALICE**		ALICE											
		SAM Vo-Specific Tests													
WLCG-07-39	Sept 2007	**VO-Specific SAM Tests in Place** With results included every month in the Site Availability Reports.		POSTPONED TO 2008 AND REPLACED BY A NEW MILESTONE (WLCG-08-08)											
		Site Reliability - Dec 2007													
WLCG-07-14	Dec 2007	**Site Reliability above 93%** Considering each Tier-0 and Tier-1 site (Note: orange means > 90% of target)	Aug 91%												
			Sept 91%												
			Oct 91%												
			Nov 91%												
			Dec 93%												
			Jan 93%												
			Feb 93%												
WLCG-07-15	Dec 2007	**Average of Best 8 Sites above 95%** Eight sites should reach an average > 95%		**Averages of the 8 Best sites Sept 2007 - Jan 2008** **Sept 93 - Oct 93 - Nov 95 - Dec 96 - Jan 95 A59- Feb 96**											

16 October 2008

WLCG Sites Reliability

April 2007 - September 2008

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

LCG Services Report June – September 2008

Jamie Shiers

1 October 2008

This extended quarterly report covers the period from immediately after the 2nd phase of the Common Computing Readiness Challenge 2008 (CCRC'08) up to the excitement surrounding the LHC First Beam day. Unfortunately, this was soon followed by the news of an extended machine shutdown, bringing the curtains firmly down on the eagerly awaited pp data taking runs and associated challenges. The most relevant question that this section can answer is "was – after many years of meticulous preparation – the WLCG service ready for the long-awaited startup of this world-class machine?"

Whilst a re-assessment of the situation is mandatory, it is important to stress that the WLCG production service must continue as precisely that, whilst continuing to evolve and expand in a non-disruptive manner with service issues being addressed both rapidly and systematically.

A number of events occurred during this period that trigged the production of a "post-mortem" – required when one of the service targets specific in the WLCG Memorandum of Understanding is not met. Formal thresholds for the production of such reports have not been defined but are associated with prolonged (several hours or often days) of significant service degradation or even downtime. As such problems are followed on a daily basis by the WLCG operations meeting, the requirement for the production of a post-mortem (typically in time for the weekly report to the Management Board, even if the problem is not yet fully resolved) can be readily confirmed there. The events that triggered post-mortems are described in more detail below.

Whilst it was agreed that CCRC'08 successfully met the target goals and metrics – albeit with less than desirable overlap between the activities of the different experiments, as well as the various activities of individual ones – some specific functional blocks (notably reprocessing and end-user analysis) were not fully tested. Furthermore, it was frequently stated that "real data taking will be different!" Finally, many service changes (middleware, storage-ware and database upgrades, replacement of aging hardware, machine room moves) took place in a largely non-disruptive manner.

As in previous years, the "summer effect", with many absences and a noticeably negative impact on the time taken to resolve problems, was clear. This can be attributed to the perceived unlikelihood of data taking from pp collisions during this period. However, the regular service degradation and/or prolonged service or site downtimes would have had obvious consequences on data taking had the machine been operating and adequate service coverage must be foreseen for future years. This covers not only the period of accelerator operation but also outside when the experiments will also be using the service in full production mode.

Post-Mortems

During the period covered by this report there were roughly 4 incidents per month – rather evenly spread in time – that triggered a post-mortem. Some of these were associated with the same service and not all could have been avoided (or limited) by adding redundancy or improving monitoring. Slightly more of these events can be

attributed to software than hardware and there were a similar number that were close to the somewhat informal threshold of that would have defined them as “major”. We must therefore expect similar incidents during data taking and be prepared how to respond to them both at the sites and global WLCG service levels.

The incidents for which a post-mortem was prepared are listed below:

- CERN-PROD blockage of all SRMv22 LHC experiment endpoints on Saturday morning 24/5/2008: 5 hours.
- In 9 days from 20-30 June there were three electrical outages of the CNAF Tier-1 lasting several days each and causing considerable damage to several components.
- IN2P3 Cooling compressor failure reduced the number of worker nodes by half for 8 hours on 21 June.
- CNAF had two separate partial network switch failures from 5-8 July with several long downtimes.
- CERN ATLAS conditions data capture failed from 26-30 July: for 3.5 days there was no replication of conditions data to the Tier1
- RAL CASTOR failures 14 August – 27 August mostly affecting ATLAS. Several long downtimes, the final one being of 7 days. At the time of writing this issues is still not resolved and further logical corruptions and loss of data have occurred. This is without doubt the most persistent and severe service problem encountered during the entire quarter.
- CERN ATLAS transfers/access to first beam data blocked on 10 September for 3 hours after run closed.
- The network link to PIC on 24 September due to a major incident in a Madrid data centre. A post-mortem analysis will be prepared.
- On September 7 (Sunday) a complete loss of service for CASTOR at CNAF. This problem was fixed on September 8 in the morning and was caused by an known Oracle bug (affecting the version 10.2.0.3) causing the Oracle management agent to consume up to 100% CPU time and subsequently to hang. Due to lack of response from the management agent, Oracle starts spawning new agent processes which degrade in the same way.

Middleware

No significant middleware enhancements were deployed in production for WLCG during this period. Unfortunately, services based on a number of middleware components that had been extremely stable for a long period suffered uncharacteristic instability during this time (in particular the LFC and also FTS). This should be taken as a warning signal and sufficient attention paid to these important components in the future. The on-going work on improving service reliability “by design” has highlighted a variety of weaknesses, as presented by Maarten Litmaath during the WLCG session at the recent EGEE’08 conference in Istanbul. Besides further data management and monitoring middleware improvements detailed elsewhere, the coming months should be put to good use with respect to the following issues:

- Deployment of the CREAM-CE, a robust and scalable successor of the lcg-CE "house of cards". The CREAM-CE has passed certification, but there is a major issue affecting its deployment: for proxy renewal it currently relies on the WN having a small configurable range of open ports to which the CE can connect. This is unacceptable for many sites, for reasons of security and/or complications in their firewall configurations. The developers are looking into how the CE can be made a proxy renewal service with the least amount of development effort. No timescale yet, but a few months seems doable. In the meantime the CREAM-CE can be deployed at sites that do not mind the port range, or it could be tested with long-lived proxies. This way experience can be gained on the production system. The CEs would be advertised such that they do not match the standard WMS requirements, i.e. they will not pick up jobs from unsuspecting users.
- The CREAM-CE also allows for job requirements to be forwarded to the batch system, allowing for a more efficient use of the WNs, a simpler configuration of the batch system, and a better approximation for the estimated response time (ERT). Here it is also important to mention the splitting of the batch cluster properties from the CE proper. The former information will be published by a logically separate service, allowing for more efficient use of heterogeneous clusters. Furthermore, recent issues with ATLAS reconstruction jobs underline the importance of consistent publication of WN properties in the information system, such that jobs can count on receiving the quality of service that is advertised for a CE.
- Deployment of gLExec, the setuid root utility allowing for proper identity switching by multi-user pilot jobs. There still are a few open bugs related to gLExec, but their resolution seems to be a matter of weeks rather than months. In parallel further progress is needed in the Multi-User Pilot Job Frameworks Review. The LHCb DIRAC system has been deemed satisfactory. The CMS glideinWMS system so far has had a favorable review as well, but further deployment details are expected if and when CMS decide to use this system on a wider scale. The review of the ATLAS PanDA system is ongoing and appears to be converging. Details of the ALICE AliEn system are being awaited.
- For various reasons the WMS has not received any patches since more than half a year. A big patch has just entered certification and further fixes and enhancements are eagerly awaited. For example, occasionally the subjobs of a collection end up without their input sandboxes; the next update should cause such collections to abort right at submission time, but a true fix will still be needed. Various WMS components should become more robust and their overall memory footprint should be reduced.
- ATLAS has a replica of the CERN VOMS servers at BNL. A similar setup should eventually be established for each of the other experiments.
- The experiments have been using long-lived proxies (lately up to 7 days) since the early days of WLCG because of various issues with proxy renewal and limitations on the proxy chain length. The MyProxy server use has steadily increased (e.g. myproxy.cern.ch has received delegation requests at almost 3 Hz for many weeks), but a remaining issue is the lack of a fail-over mechanism. If a particular MyProxy server is unavailable, a large number of jobs may fail. A feature request has been submitted to the MyProxy developers.

- The mapping of the software managers. In each experiment there are multiple software managers that traditionally have been mapped, by the CE and the VOBOX, to a single "sgm" account per experiment. Using a separate group of a pool accounts would be desirable from the audit trail perspective, but complicates the software installation procedure and the running of VOBOX services. A "sudo" approach could deal with both issues. It would have to be developed in collaboration with the experiments. At least 1 big site has reported problems with sharing the software areas over many WNs through NFS. Here a partitioning of the batch system may be a viable work-around. Otherwise we may need to revive the idea of a software installation service updating individual WNs.

Storage-ware

Experiments and sites continued to experience similar problems with storage-ware as described in previous reports. Not only did the services display instability on numerous occasions, but continued configuration problems led to some overload conditions. Confusion on the required configuration at the various sites persisted, as well as lack of clarity on what had actually been implemented. These problems were compounded by continued “bug-fix” releases of at least some of the implementations that contained features described in the release notes as “new, changed or removed”. More discipline in this key area is urgently required. This is far from a question of rocket science – simply the use of well understood standard practices.

A variety of steps have been taken to attempt to improve this situation, including better documentation, more extensive testing and more robust release procedures, as well as focused meetings between developers, sites and / or experiments.

It is nevertheless essential that problems with the storage services and progress on their resolution are also reported through the standard channels. This includes the daily and weekly operations calls, as well as regular LCG Service Coordination meetings.

Database Services

The Distributed Database Services have been in “operations mode” since shortly before the May phase of CCRC’08. As such, regular conference calls between the Tier0 and Tier1 database administrators continue, covering the operations aspects of the databases required for the associated services, such as FTS, LFC and ATLAS and LHCb conditions. They have recently been expanded to include also CASTOR and related services at the corresponding Tier1s (CNAF, RAL and ASGC).

Attendance at the daily CCRC’08 – now WLCG operations – meetings has been on-going since these were established at the beginning of the year. A number of incidents that triggered post-mortems (see above) are related to these services. This underlines the importance of a homogeneous and systematic approach to these services across the various WLCG sites. The on-going problems seen with the CASTOR services at RAL, strongly reminiscent of those seen by LFC and VOMS services at CERN two years ago, remain unresolved after many weeks. Work-arounds, such as moving to separate database instances, may not be compatible with the manpower available but in any case do not really excuse a database providing access to “someone else’s data” – a fundamental *“no-no”.* Additional issues – related both to streaming of data between online and offline as well as to Tier1s – occurred on a regular basis and were typically resolved within a matter of hours.

Towards the end of this period Oracle Data Guard stand-by databases were put in place for all the LHC experiments production databases (using hardware going out of warranty by the end of the year). This provides additional protection against human errors and potential security attacks, whilst facilitating disaster recoveries.

The extension of the service into the online systems of the experiments adds additional exposure and responsibility to these teams. Some issues related to hardware interventions on systems not located in the computer (as simple as a system reset out of hours) and console access / password changes need to be resolved if a service level equivalent to that offered for centrally managed systems is to be achieved.

Given the still unresolved problems affecting the database services used for the CASTOR services at RAL, clear mechanisms for escalating such issues within the Oracle account team – such as the Management and Technical Review Boards previously held with major suppliers to CERN – could be resurrected and used in parallel to the standard (more technical) support lines.

Monitoring, Logging and Reporting

In today's production Grid environment, monitoring is a vital part of daily operations. Given the complexity of the infrastructure, this implies a corresponding degree of complexity in the monitoring procedures and tools. As was noted in the post-mortem of the CCRC'08 exercise, experiment "shifters" rely on the dashboards, experiment-specific tests executed in the SAM framework, as well as other experiment-specific monitoring tools to run and monitor their activities on a daily basis. The level of success achieved is high-lighted by the following statement (admittedly by an expert): *"using [the above tools] it basically takes me 15 minutes a day to check the status of all CMS sites and corresponding services in the world".* The use of these tools is now spreading into much wider "non-expert" communities. One of the main outstanding issues is the provision of site-oriented views of the experiment production activities. This was one of the key areas of work during this period and a well-advanced prototype now exists that will shortly be beta-tested at some key sites. It is clear, however, that work across this entire area will continue for some time to come and in particular reacting to needs only exposed by real data taking and production.

Readiness for LHC First Beam

The service report for the previous quarter boldly claimed that we were *"ready to face the challenges of real data taking."* This, coupled with the successful results from CCRC'08 – albeit with the caveats mentioned above – should have placed us in a good position for the data taking scheduled for the latter part of this year. In reality, the service performed well in all aspects, including – perhaps most importantly, its ability to respond within the agreed time periods to problems of varying criticality and to resolve them to the satisfaction of the affected VO(s). However, some aspects of the service are still rather fragile and an unsustainable degree of attention is required to keep the services running at the needed level. Some of the experiments' work-flows are still not performing at the required level for standard data-taking: the previously mentioned reprocessing and user analysis being clear examples in this respect.

Daily Operations Meeting

Although not unanimously shared, the opinion that the daily meeting and its associated minutes are at least valuable, if not (currently) necessary, are supported by generally

good attendance on a daily basis from representatives from most to all VOs, as well as service providers from CERN and some of the Tier1 sites. Currently less than one third of all Tier1s participate on a regular basis. On the other hand – judging purely by the feedback (rapidly received) on mistakes or omissions in the notes from the meetings – these minutes have a wide audience.

Concrete examples of where the meeting has helped resolve issues include: helping to move resolution of BNL OPN failover problems forward, getting DBA intelligence for the database problems affecting the RAL CASTOR service, pushing for CMS elog service and accelerating fast replacement of new CMS web servers after the cmsmon hacking incident.

The meetings will continue and wider participation in the future is encouraged, particularly during periods of major production, data taking and challenges and when sites encountered specific problems. In this vein, (close to) full attendance on a regular basis could be a milestone for the CCRC'09 exercise, continuing through to data taking from pp collisions and beyond.

Liaison between WLCG Service and LHC Operations

Informally established two years back, the recent incidents concerning the LHC machine underline the need for a well-defined information channel between the LHC operations group and the WLCG service. Regular reports from the LHC operations group to the WLCG service meetings will be established (WLCG service coordination meetings – typically held twice per month – and, when necessary, the daily operations meeting for significant updates), together with presentations at the key WLCG planning events and workshops (as already occurs today).

Outlook for the remainder of 2008 and beyond

The clear message for the future is that the production service must continue without interruption or degradation. Some additional service upgrades that had been delayed or put on hold can now be foreseen but the general rule of not breaking the production service must be adhered to. The experiments will continue a mix of data taking from cosmics with functional tests together with on-going Monte Carlo production and will also be considering which internal enhancements or upgrades are possible.

Given the scope of the foreseen changes, the need of a formal CCRC'09 challenge is becoming clearer. The exact dates when this should be performed clearly depend on the as-yet unknown accelerator schedule but also on other events, such as CHEP 2009. Tentative dates are currently mid-February to mid-March, although this may be revised as external schedules become clearer. The planning workshop, scheduled for November 13-14 at CERN, as well as the WLCG Collaboration workshop during the weekend prior to CHEP (20-21 March), are maintained. In the longer term, events co-located with the EGEE'09 conference and OGF 27, to cover experience with 2009 data taking and processing in these regions, are being considered.

Grid Deployment Board Report Quarterly Report

15 October 2008

June - September 2008

John Gordon

Summary of Past Quarter

The Grid Deployment Board continues to be the WLCG forum where technical discussions can take place in depth between WLCG sites, LHC experiments, middleware developers and service providers. Looking back over a year there has been a marked change in the type of discussions held. No longer are there heated debates on the functionality of the middleware, its readiness or lack of it. The imminence of data in 2008 concentrated people's minds and I believe we were ready for data. A pity it did not arrive. At the end of the quarter a number of closed issues had been reopened and the middleware changes and service upgrades that might be feasible this winter were discussed.

The GDB met three times in this quarter and the agenda and papers are available

Among the issues discussed were:

Storage. Work is underway to converge the implementations of SRM2.2 as an addendum to the MoU. The main issues in this quarter were the observed behaviour of the implementations in CCRC08 and following. Many of the dCache issues were addressed at regular phone conferences and all Tier1 sites were encouraged to participate. The database experts at Castor sites were encouraged to collaborate directly and not just through their castor service providers.

Emergency Tickets The experiments were frustrated by the delays in the GGUS helpdesk service. Tickets they raised about an issue with a known site were first triaged and then assigned to the relevant ROC before reaching the site involved. This could take more than 24 hours. The request was for experiment experts to be able to alert sites directly. Moving to email would allow this but would lose the experiments the ability to track the problem and its resolution. The EGEE group managing the GGUS service proposed a solution which takes the site contact details from GOCDB. For now, while the majority of tickets come from experts no authorisation will be required for this. This can be added later if badly-directed tickets from general users become a problem.

Middleware Fewer new releases as might be expected approaching data taking but outstanding issues with CREAM, SCAS, Glexec, and FTS. WMS still has issues but it is enough of an improvement over the RB that the migration is thought worthwhile.

Monitoring. The SAM monitoring framework is much used by WLCG both to spot problems quickly and to evaluate metrics for availability and reliability. The tests are mainly run in jobs at a site which is submitted from a central place. EGEE have a plan to run the tests under the control of Nagios, locally at sites. Nagios would then push the results to a central database as before but the local Nagios can be much more flexible in running the tests and in raising alarms for failures.

Megatable Work started on automatic collection of statistics on installed capacity at sites. This will put the accounting reports into context. It is hoped that sufficient information can be gathered from the Information system. Further development of the information providers may be required, particularly for storage.

Security. Most of the desirable security policies are now approved and some have even been round the circle of revision. We have a reasonable network of security contacts in place. There have been several challenges of sites to test the responses and procedures.

Pilot Jobs A working group has started to review the pilot job frameworks of all experiments. The review of LHCb is complete. They attempted a trial of their framework with glexec but were thwarted by WMS problems.

Benchmarking. The HEPiX Group circulated a report on their investigations. They proposed to use a variety of the SPEC2006 benchmark suite called SPEC2006-cpp which has the advantages that it is a good match for the integer/floating point mix of particle physics and there are no published results forcing sites to run it on their computers and not to use results published, perhaps from different ones.

Plans for the Next Quarter

In the next quarter:-

There will be continued discussions on the changes to be allowed over the winter shutdown.

In addition:

I hope to get firm recommendations on CPU benchmarking which will then lead to much work on re-measuring the resources at sites and the recasting of the experiment requirements in new units.

The WN configuration should be rolled out so that sites can properly advertise their heterogeneous clusters so that work can be more efficiently targeted at them by WMS and other resource discovery tools.

Glexec with SCAS should be certified and deployed so that pilot job frameworks which pass scrutiny will convince sites to deploy them.

Of course I realise that these were all issues which I hoped would make progress in the quarter being reported on. They have but are not yet at a conclusion. .

GDB Meeting During the Quarter

The GDB met three times in this quarter and the agenda and papers are available at:-

June http://indico.cern.ch/conferenceDisplay.py?confId=20230

July http://indico.cern.ch/conferenceDisplay.py?confId=20231

September http://indico.cern.ch/conferenceDisplay.py?confId=20245

QUARTERLY STATUS REPORT				
Project Name			**Date**	
Applications Area			3.10.2008	
Report Period			**Author Name**	
2008Q3 (Jun-Sept)			**Pere Mato**	
Milestones for the Quarter			**Status**	**Comments**
SPI				
SPI-13	**31.12.07** **30.04.08** **30.09.08**	Provide an integrated web-based information system about LCG software. This includes information about software dependencies, build information, LCG configurations, then nightly build system, etc.	Done	The system has been put into production (see http://lcgsoft.cern.ch)
SPI-15	**31.03.08** **30.09.08**	Integrate the Geant4 build and test procedure with the LCG AA nightly build system, this will require some adaptation as Geant4 uses different procedures for software check out using tag collectors.	Done	The nightly build system has been adapted to the needs for Geant4 builds. As a proof of concept first test suites are being executed with in the LCG/AA nightlies. Any further work to make the test suite complete will be done by a newly joined Technical student of the Geant4 collaboration (with help of SPI if needed).
SPI-16	**30.06.08** **31.12.08**	Deployment of a web content management system, after a quick survey, needed for the restructuring of the SPI web to provide a coherent and complete source of information of all services for users and maintainers.	In progress. Rescheduled.	Investigations about the amount and distribution of pages are in progress. After analyzing those results we will have a clearer picture about the distribution of SPI pages over web servers, content types, etc. which will ease the decision on a new web content management system to be implemented to server as many of the different types of pages as possible.
SPI-17	**30.06.08**	Development of a tool to bootstrapping LCG-AA software infrastructure. This tool should download the essential ingreedients to be able to install the rest of the LCG-AA software stack.	Done	The tool "boostrapLCG" will download and deploy the necessary boostrapping ingredients of LCG software (i.e. LCGCMT configuration, python, cmt) and setup a proper environment. The tool was developed and tested all LCG/AA platforms.
SPI-18	**30.09.08** **31.03.09**	Migration of the current SPI web contents to the newly deployed content management system. This will require the manual inspection and possibly correction, re-writing of the pages.	In progress. Rescheduled.	This milestone depends on SPI-16 which is currently in progress.
ROOT				
ROOT-19	**30.06.08** **31.12.08**	Implementation of the complex data schema evolution in ROOT	Rescheduled.	The schema evolution will be available in version 5.22 in December, most of it already released with 5.21/04.
ROOT-20	**30.06.08**	Improvements of the ROOT test suite as part of the nightly build system to improve the robustness of the system in general.	Done	Done but of course new tests are regularly added.
ROOT-21	**30.06.08**	Development of the Event Display library (first version).	Done	The Event Display Library (EVE) now used by several event displays, in particular AliEve for Alice and FireWorks for CMS. This area requires continuous developments.
POOL				
POOL-11	**30.04.07** **31.08.07**	Complete the porting of the POOL data regression tests into the nightly build system	Done	The work has been completed.

POOL-13	**30.06.08** **31.12.08**	CORAL server development. COOL read-only tests for selected basic use cases pass	In progress. Rescheduled.	Developent is taking longer than expected. The first official release (read-only) is expected by December.
POOL-14	**15.08.08** **31.04.09**	CORAL server development. All CORAL integration tests (including write test) pass. This will also require some extension of the current CORAL tests suite to achieve full coverage	In progress. Rescheduled.	The functionality (update and security) of CORAL sever is expected by end April 2009.
POOL-15	**30.09.08**	CORAL Server scalability and stress tests pass.	No progress. On hold.	Waiting for POOL-13 and POOL-14
		COOL		
COOL-9b	**31.12.06** **31.03.07** **30.06.07** **30.09.07** **31.12.07** **31.03.08** **30.09.08**	Deployment of COOL database services at Tier0 (separate instances for online and offline) and Tier1 for LHCb with Streams replication.	Done	For LHCb: a test service setup was prepared with two-step Streams replication between CERN online (private LHCb test single-instance server at the pit), CERN offline (IT-PSS 'integration' RAC) and three 'phase-1' Tier1 sites (Gridka/FZK, IN2P3, RAL) by Q4 2006. One 'phase-1' (CNAF) and one 'phase-2' (Nikhef/SARA) Tier1 sites joined in Q1 2007. The last 'phase-2' site (PIC) joined in Q2 2007. The production 'LHCb-offline' RAC server replaced the 'integration' RAC in the T0 setup for LHCb in Q2 2007. The production T0 setup was finally completed in Q3 2008, with the move to the production 'LHCb-online' RAC server, installed and managed by LHCb at the pit. The delay is due entirely to LHCb.
COOL-25	**30.09.08**	Implement a 'partial' tag locking mechanism.	Done	Partial' tag locking is meant to prevent the removal but allow the addition of new IOVs or HVS nodes to partially locked tags. The generic API for partial tag locking, and its implementation for the additions of new HVS tags, have been completed in COOL 2.3.0 (January 2008). The functionality to allow also the addition of IOVs to partially locked tags has been completed and is ready to be released in the upcoming COOL 2.6.0 (Nov 2008).
COOL-26	**30.09.08** **30.12.08**	Support for the gcc4 compiler.	In progress. Rescheduled.	The port of the COOL code and configuration to support gcc4.1 has been completed in COOL 2.3.0 (January 2008). This is not an officially supported platform yet - it is expected that only gcc4.3 will be supported in the LCG AA. The port of the COOL code to gcc4.3 (stricter than gcc4.1) has started in October 2008 and is almost completed. It also depends on the completion of the CORAL port (POOL-19).
COOL-28	**30.09.08** **31.12.08**	Support for the 'CORAL server' backend.	Rescheduled.	Support for 'coral://' URLs was first prototyped in COOL 2.4.0 (February 2008), allowing simple tests against early prototypes of the CORAL server and the definition of additional constraints on its development for its integration into COOL. The COOL read-only tests are now routinely used to validate the CORAL server implementation (POOL-13). Full support in COOL depends on the release of the CORAL server (POOL-16).
COOL-29	**30.09.08** **31.12.08**	Expose transaction management in the user API.	Rescheduled.	Prototypes of the API and implementation are ready to be internally reviewed for inclusion in one of the upcoming COOL releases.
COOL-30	**30.09.08** **31.12.08**	Allow session sharing in the user API.	Not started.	Depends on COOL-29. Waiting for review of API of COOL-29 transactions. Instead: COOLxx (pyalod queries), resurrected, will be in COOL260.

SIMU				
SIMU-10	**30.06.07** **31.12.07** **31.12.08**	Application of corrections of test-beam data, for validation of stand-alone simulation, to the LHC calorimeter test-beams (VD703)	No progress. On hold.	No progress. Experiments are still working to complete their test-beam analyses.
SIMU-20	**30.11.07**	Review, redesign and debugging of the FLUGG tool (SF711)	On hold	Partially done. An important bug fix was recently provided, enabling to use FLUGG with the latest version of Geant4. A general code review has not been done due to lack of manpower.
SIMU-21	**15.12.07** **31.12.08**	Thin-target validations of Geant4 forward physics (G4712)	On hold	Work is suspended, due to lack of manpower in physics validation. Problems exist with acceptance corrections in the published HELIOS data. Awaiting man-power (a fellowship) in order to continue this work. Postponed to December 2008.
SIMU-25	**30.03.08**	4th simple benchmark for Geant4 and Fluka: diffraction of nuclei (VD801)	On hold	After first Geant4 results, also some preliminary Fluka results have been compared with data. After discussions with Fluka experts, it has been agreed that the data needs further investigation since the original analysis was based on some old, wrong assumptions. Furthermore, proton-proton data is considered important for a more complete investigation of the diffraction, therefore requiring additional analysis. The activity has been postponed, pending the assignment of new manpower.
SIMU-27	**30.04.08**	Status report on comparisons with shower shapes and relevant physics modeling (G4802)	In progress	With the addition of quasi-elastic scattering and the use of the Geant4 QGSP_BERT physics-list, starting with version 8.3, the hadronic shower shape descriptions are now much improved, resolving the issue for the LHC experiments. A short note summarizing this progress is pending and under preparation.
SIMU-29	**30.09.08**	Fluka extension to the ATLAS HEC test-beam analysis (VD804)	In Progress	The ATLAS HEC test-beam analysis is one of the main calorimeter validation tests for the hadronic physics of Geant4. The aim of this milestone is to extend this analysis to Fluka using Flugg, and the experience gained in the similar Fluka extension of the ATLAS TileCal test-beam (see Milestone VD524).
SIMU-31	**01.06.08** **31.12.08**	Extend Rivet validation to new C++ generators (GS808)	In progress. Rescheduled	Some work has started in December with Sherpa, but it is not yet concluded due to problems with Sherpa which were fixed after the GENSER integrator left. The work should resume and be concluded by December 2008.
SIMU-32	**15.09.08**	First version of System Integration Testing of Geant4 running on SPI-nightly platform (G4811)		Level-2 milestone. Provide migrated test suite for integration in the SPI-nightly facility.
SIMU-34	**30.06.08**	New release of HepMC (2.04) including new handling of units (GS815)		2nd level milestone, the new release will include optional handling of units and other minor features

Summary Of Progress

No major releases of the Application Area software was released during last quarter. Experiments were getting ready for beam and didn't require 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 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). Significant progress was made also in the port of COOL to gcc4.3 and VC9.
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.

On the CINT/Reflex merge, a version that passes all ROOT tests and that is comparable in speed with the old CINT is on schedule for ROOT's December release. ROOT can be built already now with the new and the old CINT in parallel. Still to come: speed improvements for Reflex and several bug fixes for the core of CINT7.

We released the first step in the implementation of the infrastructure for the new Data Model Evolution Scheme. The current capabilities are: assign values to transient data members, rename classes, rename data members. change the shape of the data structures or convert
one class structure to another, change the meaning of data members, ability to access the TBuffer directly when needed, ensure that the objects in collections are handled in the same way as the ones stored separately; all of those are supported in object-wise, member-wise and split modes. The following features and issues should be addressed in the next release: make things operational also in bare ROOT mode; ability to transform data before writing; support for changing the class type of nested object in a split branch; support for access to onfile version of nested objects from within the parent rule.

Graphical User Interface development was mainly focused on event displays (Alice and CMS), implementing a few new widgets and extending several existing ones. An event recorder prototype has been successfully implemented and integrated in ROOT. Focus has been given on improvements of the robustness and performance of the GUI.

A new version of the ROOT mathematical libraries has been release 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-in's. The GUI fit editor has been as well improved by adding the support for multi-dimensional histograms and graphs.
The documentation has been also improved and in particular, for RooFit a new version of the user guide has been released and many example tutorials added to the distributions.

Redesign of the TGraph family classes to separate the data handling from the graphics. Improvement of test suite to make it the more accurate as possible in early bugs tracking. As part of the general effort to improve ROOT documentation and tutorials the documentation of all graphics classes has being redesigned and completed.

Finally, for what concerns 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 plugin, 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 start processing the query; and (iii) the support for memory consumption monitoring on all the workers as a function of the processing step.

The POOL project was reviewed in May 2008 to identify the steps to be taken to prepare POOL for the LHC startup 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.

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). Significant progress was made also in the port of COOL to gcc4.3 and VC9. 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 also been prepared and will be released in Q4 2008. For all three projects, some bug fixes have also been produced in the LCG_54-patches nightlies as not all the experiments have migrated to the "deSEALed" LCG_55 releases yet.

For the simulation proejct during the 3rd quarter of 2008, two major achievements were made in Geant4: a preview release 9.2-Beta, released to public in July, and a new patch to release 9.1 (9.1.p03), released last 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. 9.1.p03 also contains corrections for two issues reported by the HARP-CDP group last March. 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. 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 plugin module. The final public release 9.2 is expected for December. Quite a lot of activity in the physics validation domain has been carried out to test and verify the new improved Fritiof model by V.Uzhinskiy, also using the HELIOS target diffraction benchmark. A new version of HepMC (2.04.00) was released last summer, adopting the new release process implying direct contertation with MC authors and developers in a dedicated forum; the release includes the announced changes in the handling of units, agreed with the experiments.

Issues During the Quarter

Milestones Changes and Actions

References and Hyperlinks

New and Next Quarter Milestones			Status	Comments
SPI-19	**30.09.08**	Exploit the full potential of multicore build servers for LCG/AA nightly builds	New Done	The nightly build systems have been adapted to exploit the full potential of the multicore build servers used for LCG/AA s/w builds. This has been achieved by multithreading the builds on different levels (i.e. packages, compilation units) - decreasing the overall build time by a factor 3.
SPI-20	**31.12.08**	Establish the software removal procedure	New	Over time the afs area of LCG/AA has been the container for many different versions of software packages. A proposal for cleanup of these has been sent out to the stakeholders of the LCG/AA software. If accepted some tools for the automatic cleanup need to be implemented.
SPI-21	**31.12.08**	Review of the LCGCMT configuration database	New	LCGCMT is the basic software configuration database for CMT based projects in LHC. A review and adaptation to new needs of this database is needed. E.g. make it easier to do cross compilation of software. A proposal with changes has been sent out to the stakeholders, i.e. members of the architects forum which will be touched by these changes.
SPI-22	**31.12.08**	Nightly builds with a "client server architecture"	New	In order to allow even more dynamic builds of the LCG/AA nightly builds a client-server architecture is envisaged. This will allow "build nodes" to connect to a client which will distribute the builds according to the capabilities of the client.
SPI-23	**31.03.09**	Automatic external s/w stack rebuild	New	The LCG/AA sofware stack is permanently adapted to new compilers, operating systems, architectures. Every time such a change happens all the external software packages need to be recompiled. With the newly introduced "Builder" system this can be done easy on a package per package basis. A tool on top of this system should allow further automatization and ease the recompilation of all software packages in one go.
COOL-14	**31.03.07** **30.06.07** **30.11.08**	Support for simple payload queries (lookup of IOVs by payload data).	*Resumed.* Ready to be released.	The implementation of payload queries will be based on the new record and field interfaces described in milestone COOL-7 and released in COOL 2.0.0 (January 2007). This milestone has been resumed after being removed in Q2 2007. The new API and its implementation are ready to be released in the upcoming COOL 2.6.0 (November 2008).
COOL-31	**30.12.08**	Reimplement and optimize all SQL queries for IOV retrieval by time, reusing the same C++ methods for different SV and MV use cases.	Ready for release.	The SQL queries needed to handle the various COOL use cases (SV, MV tags, MV user tags, MV HEAD...) were originally defined in separate C++ methods, added over time. In order to allow the future maintenance of the software and further performance optimizations, these pieces of code need to be merged together. Some improvements in this direction were added in the COOL 2.3.1 release (February 2008): the same code is used for IOV retrieval from MV tags and MV user tags. This has allowed the simultaneous performance optimizations of IOV retrieval from MV tags, and IOV insertion with MV user tags. Additional improvements were then added in COOL 2.5.0 (June 2008) to
COOL-32	**30.11.08**	Implement the 'tag cloning' functionality.	New. Ready to be release	This functionality has been requested by LHCb. Its implementation is completed and ready to be released in the upcoming COOL 2.6.0 (November 2008).

COOL-33	**30.11.08**	Avoid unnecessary COUNT(*) queries in IOV retrieval.	New. Ready to be released.	This performance optimization has been requested by Atlas as a result of their distributed stress tests in Q3 2008. Its implementation is completed and ready to be released in the upcoming COOL 2.6.0 (November 2008). The size of IOV iterators is now computed only on demand, avoiding unnecessary COUNT(*) queries against the database server.
COOL-34	**31.12.08**	Support for MS VC9.	New. In progress.	A significant effort was spent during Q3 2008 on the port of the COOL code and configuration to support the Microsoft Visual Studio Express 2008 (VC9) compiler. In cooperation with the SPI and ROOT teams, this resulted on good progress also in fixing several issues with gccxml, ROOT and LCGCMT. COOL can now be built but several issues still exist at runtime during tests. Support for COOL on VC9 also depends on the completion of the CORAL port (POOL-21).
COOL-35	**30.06.09**	Migration from CVS to SVN.	New.	
POOL-16	**31.10.08** **31.12.08**	First CORAL release with read-only CORAL server support, start of experiment validation	Rescheduled.	This milestone has been rescheduled and reduced in scope to functional tests of the read-only functionality. The releases of more complee CORAL server software with secure authentication and full write functionalities have been rescheduled as milestones POOL-17 and POOL-18.
POOL-17	**31.10.08** **28.02.09**	Release of CORAL Server with secure authentication. All functional tests pass.	Rescheduled.	This is a rescheduled milestone, previously expected for October 2008 as part of POOL-16.
POOL-18	**31.10.08** **30.04.09**	Release of CORAL Server with full write functionality (DML and DDL). All functional tests pass.	Rescheduled.	This is a rescheduled milestone previously expected for October 2008 as part of POOL-16.
POOL-19	**31.12.08**	CORAL support for gcc4.3.	New.	
POOL-20	**31.12.08**	POOL support for gcc4.3.	New.	
POOL-21	**31.12.08**	CORAL support for MS VC9.	New.	
POOL-22	**31.12.08**	POOL support for MS VC9.	New.	
SIMU-35	**1.12.08**	Test of MCDB in CMS large productions	New	
ROOT-22	**31.12.08**	Restructuring of the ROOT web site and documentation system.	New	We are planning to make a substantial reorganization of the ROOT web site (unchanged since many years) to reflect a more modern presentation style. The new site will include a brief description of the main ROOT functionalities and more guidance for newcomers. The class documentation system will be revisited to provide a more coherent description of the class, function parameters and side-effects. A first version of the web site is expected for June 30.
ROOT-23	**31.12.08**	Implementation of PROOF optimized to run locally on multi-core platforms (PROOF-lite).	New	This version of PROOF, PROOF-lite, will not use the xrootd daemons, but start directly the master and workers. Communication will be via local mechanism, like Unix sockets and message queues. Also this version will optimized the I/O by trying to use memory mapped I/O. Due date December 2008
Comments and Additional Information				

ALICE Report

June 2008 - September 2008

Y.Schutz

Monte Carlo Production

ALICE performed continuous production in preparation for 2009 data taking: production in Tier-1/Tier-2 sites, and data and end user analysis at the Tier-2 sites.

Fast First Physics MC production covering different scenarios of LHC energies and detector conditions with rapid changes in the code and conditions reflected immediately.

There will also be grid production of statistical samples equivalent to the expected RAW data in the first hours of LHC and analysis of data on Grid and CAF developed by dedicated Physics Working groups.

ALICE will proceed with the production of first physics including the LHC and detectors conditions of Spring 2009 and will start with productions for a standard year of data taking including pp@14 TeV and AA.

Analysis

Several kinds of analysis tools are operational: the CERN Analysis Facility (CAF) for fast analysis, the ALICE analysis train for organized analysis and end user analysis with mostly chaotic access patterns. The end user analysis is performed on the primary copy at Tier-2s SE and jobs are run at the Tier-2 holding the required data. They are single user analysis jobs yielding low CPU/Wall efficiency (I/O bound).

The ALICE Analysis Train provides organized analysis allowing many tasks processing a single data stream and thus yields nominal CPU/Wall efficiency,

The analysis on the CAF is performed using PROOF and local data. In particular, the data is imported from Grid, stored locally on CAF nodes and performed on Physics Working Group selected data sets.

Raw Data Processing

Online production of condition parameters, first pass processing @ Tier-0, replication in Tier-1s, N pass processing @ Tier-1s operational.

The online data replication to Tier-1s and the condition parameters calculation are working and stable. The conditions framework is fully operational and condition data are collected for each run. Also on line reconstruction, monitoring and QA are working.

In general the framework is ready and the detector implementations are in progress.

Data is reconstructed offline after data taking. The first pass of reconstruction is executed at the Tier-0; Data replicated to 2 Tier-1 Sites (reconstructed ESDs) and at Tier-2s on demand.

Fast reconstruction line of selected datasets at CAF under development, while the end user analysis is performed by detector experts on the Grid and on the CAF.

RAW Data replication and reconstruction at Tier-1s is currently suspended.

ALIROOT Software Releases

Stable release ready for data taking; code evaluation and some refactoring can still be done before the LHC start. A strict release policy is now implemented.

The current version is ready for MC production for first physics (pp@0.9/10 TeV, w/wo B field) with the detectors as installed. An open issue is that the Raw data format is not yet final.

In preparation of Spring 2009 ALIROOT will undergo to some code refactoring where possible.

Services

The new AliEn version is deployed routinely with effectively no downtime. Job management in all its form is now used: ALICE uses the RB (phased out but still widely used), WMS, and CREAM (very promising initial stability and scaling tests). The new version of AliEn routinely is deployed on the sites and job management in all its forms (RB, WMS, CREAM) is well under control. The xrootd-enabled SEs are working fine

A full test setup provided of CREAM was provided by GridKA. The system has been tested in the last 3 months with remarkable stability; more than 75000 jobs passed through the CREAM CE

Once officially available in gLite release, ALICE is very much in favour for a quick distribution of CREAM at the sites.

Accounting and Resources

ALICE has used 40% (6M SI2K) of allocated CPU and 53% of required. It would be important to homogenize the CPU factors used by the WLCG sites.

The SE is operational at 30 sites, the remainder being installed. Currently 27% of pledged storage is operational and 64% of that is used.

Requirements for 2008/2009 had been re-evaluated (before the LHC incident). New requirements will be prepared for 2009, depending on the LHC schedule; expect larger requirements with respect to C-TDR (CPU, disk)

In 2008 a substantial reduction of the requirements with respect to the C-TDR requirements

The requirements for 2009 have to be reevaluated once the LHC running scenario for 2009 is known. ALICE does not anticipate a reduction in the requirements as compared to the C-TDR.

Plans to cope with a possible deficit in 2009 are investigated and should stay within 10% of the allocated global computing budget.

ATLAS Report

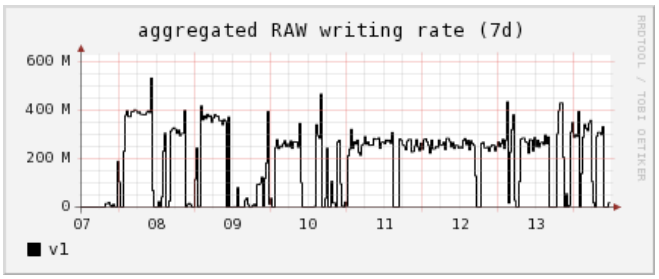
June 2008 - September 2008

D.Barberis

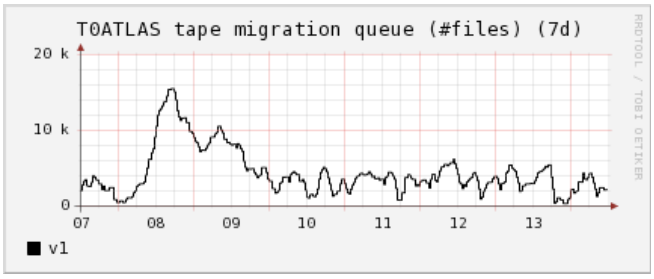
Tier-0 and Data-taking Activities

ATLAS are taking continuously cosmic ray data since several months and until 3rd November, With only short breaks for detector work (and LHC data!). The Tier-0 is coping well with nominal data rates and processing tasks. A few Castor glitches are usually sorted out with the Castor team within a very reasonable time.

In November hardware detector commissioning work will restart but cosmic data-taking will carry on with partial read-out. Below is an example of the data rate from the online to the offline of ATLAS.

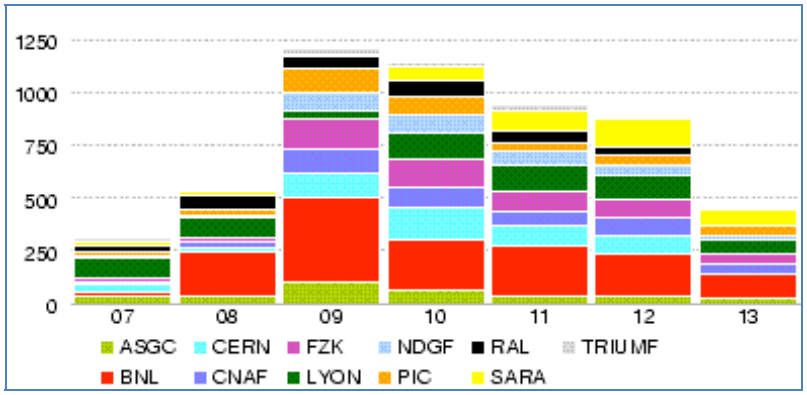


The tape queue can handle the amount of data and copy it on tape.



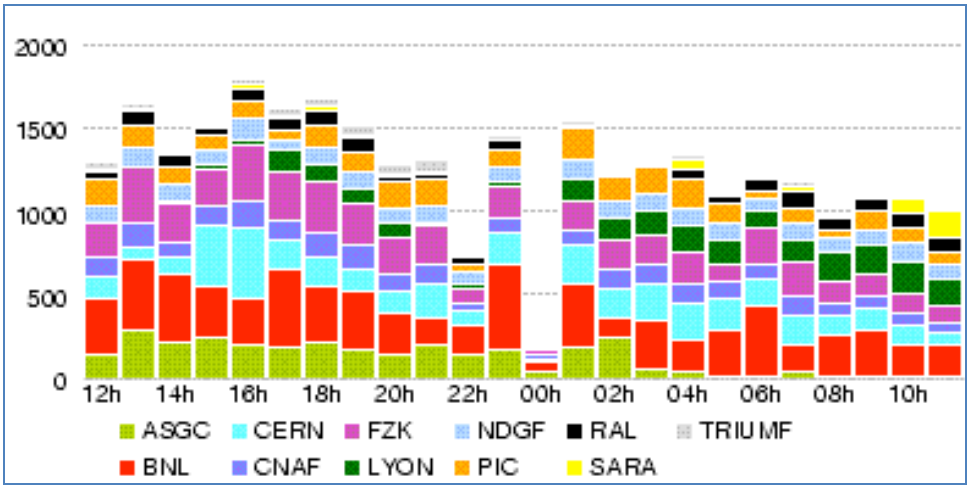
Data Export

ATLAS exports all raw and processed data from Tier-0 to Tier-1s and Tier-2s according to the computing model. The system can sustain the peak rate of 1.2 GB/s for an indefinite time.



Data distribution patterns are periodically revised as data types (triggers) and processing needs change. Above is the data export summary from CERN 7-13 October (MB/s)

For instance on the 9-10 October the 1.2GB rate is consistently maintained:



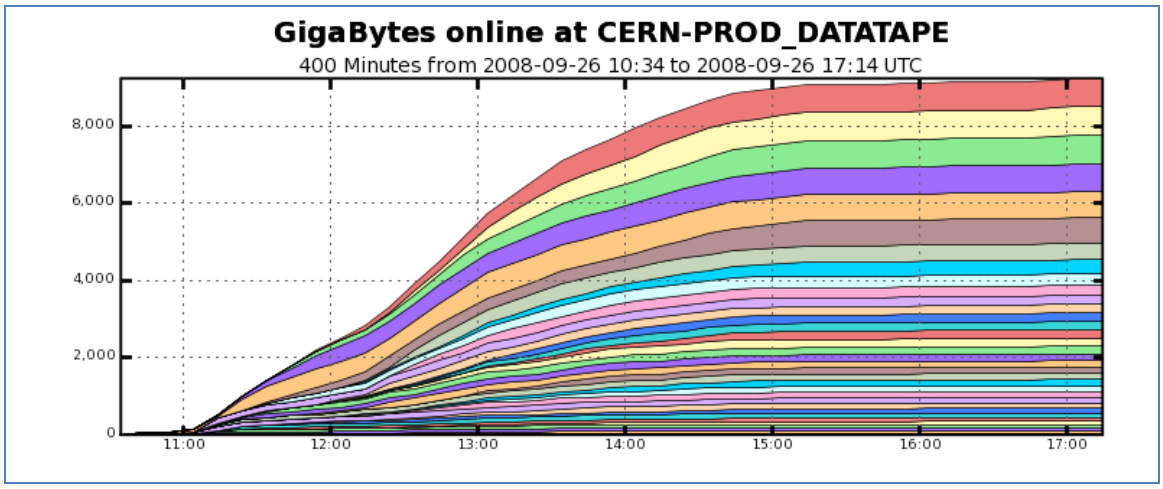
Pre-Staging Tests

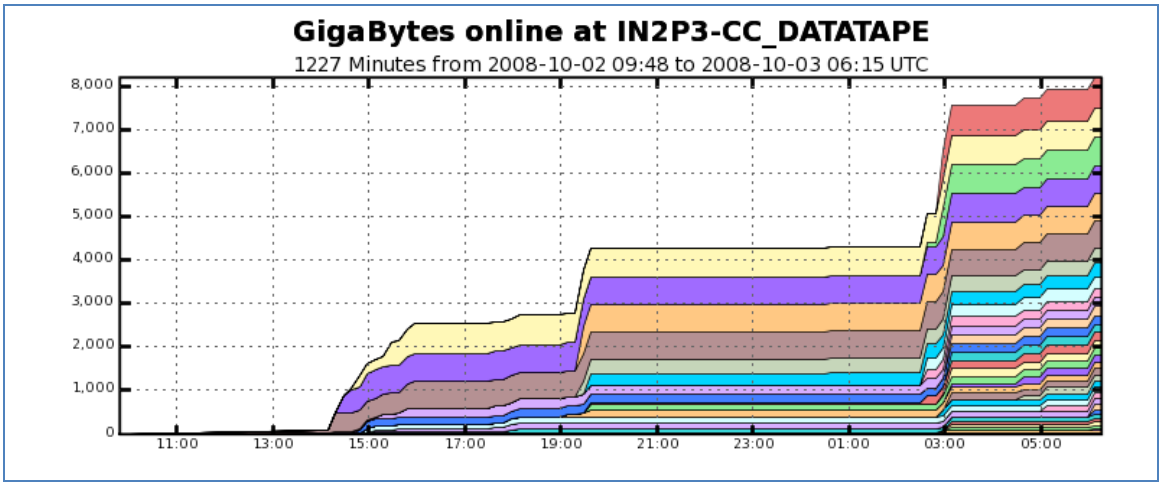
ATLAS started during the summer pre-staging tests at all Tier-1s. Recalling whole datasets at a time (up to 10 TB).

Performance varies a lot as tape back-ends are different at each site. After a few tries, most sites are mostly working even if there are outstanding (different) problems at PIC, FZK and SARA.

This exercise also showed that the number of available tape drives varies a lot from site to site. There is no point in having 1000s of processing cores if they cannot be fed at the correct rate with data.

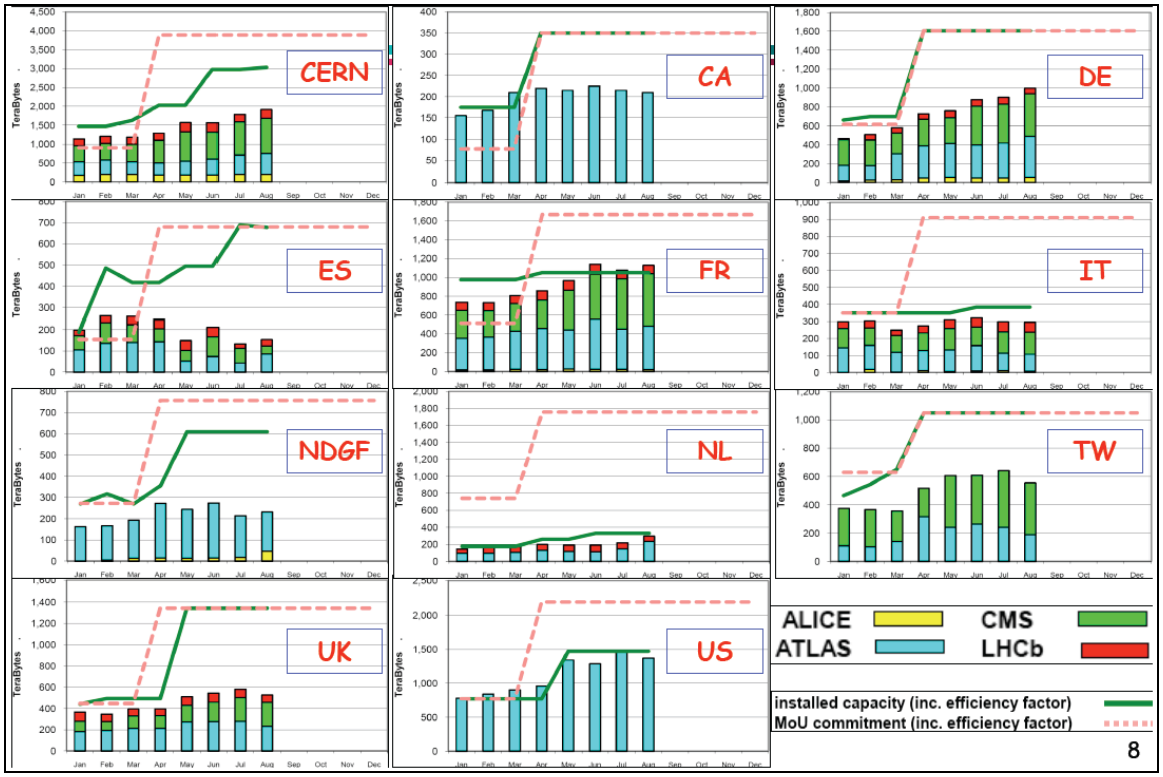
For instance CERN and IN2P3 could store all the dataset required, although with different patterns. At other sites not all datasets were correctly distributed and the issue is being studied.





ATLAS Disk Usage

Below is the disk pledges (dotted line), the installed capacity (green line) and the usage of the VOs at the Sites (ATLAS in light blue).



Database Access Issues

Early tests of database scalability did not indicate there would be any problem with reprocessing at Tier-1s.

More recent tests instead showed a serious limitation when more than a few 10s (up to 100) jobs start simultaneously, as they all access conditions data from Oracle databases. Two factors differed between these tests: (1) Oracle Streams are now used to move data from CERN to Tier-1s and (2) DCS (Detector Control System) data are now accessed by reconstruction tasks.

ATLAS started a task force to analyse data access patterns from the Oracle server side with ATLAS and CERN DBAs. It also includes the work to instrument Athena to log database access and data volumes and on detector code developers to revise and optimise their database access

Patterns.

ATLAS started to explore SQLite technology for reprocessing tasks: Dump all data for a given run to a SQLite file and use it locally for all jobs. This reduces the database access by a factor of several 100s (the number of files in a run). Oracle will still be needed at the sites for other activities.

Plans

ATLAS has several upcoming software releases:

- 14.X.Y.- will include bug fixes only for HLT/Tier-0 and Grid operations
- 15.0.0 - February 2009. Include feedback from 2008 cosmic running and will be the base release for 2009 operations.

The cosmic runs with the complete detector will continue till early November 2008 and restarting late March 2009. While with partial read-out will continue all the time at reduced rates.

For collision data: ATLAS will be ready to go from April 2009 for what concerns the ATLAS Software & Computing.

Organization News

The ATLAS Collaboration Board met last Friday and took the following decisions (among others):

- Dario Barberis was re-appointed as Computing Coordinator until February 2010
- David Quarrie was re-appointed as Software Project Leader until February 2010
- Kors Bos was elected Deputy Computing Coordinator until February 2010. And will become Computing Coordinator from March 2010 until February 2011
- Hans von der Schmitt's term in office as Database Coordinator ended on 30 Sept after 2.5 years. He is replaced by Giovanna Lehmann Miotto (CERN) for online databases and Elizabeth Gallas (Oxford) for offline databases

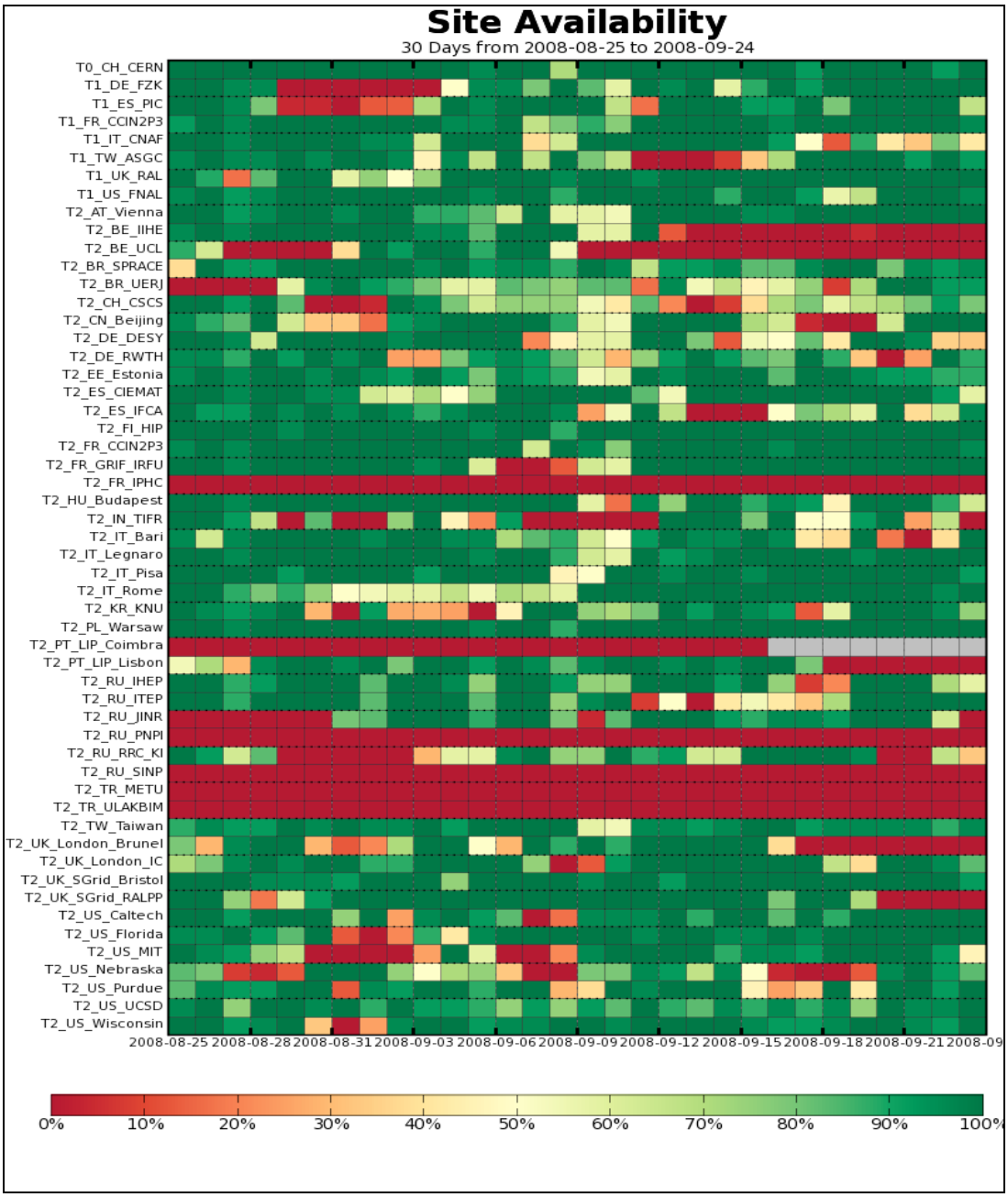
CMS Report

June 2008 - September 2008

M.Kasemann

CMS Computing Achievements

Several improvements after CSA08 and CCRC via “Integration Campaigns” which improved data handling, monitoring and analysis job submission. The effort came from developers, facility and data operations and the integration team. These “Campaigns” are successful to achieve mini-milestones.



The CMS site commissioning (SC) is one of the activities of PADA (Processing and Data Access) Task Force. The aimed objectives of the task is to guarantee that data processing workflows at Tier-1 and T2 sites can be performed efficiently and reliably.

The goal is to verify that CMS sites are complying with their resource pledges and are able to sustain both Data analysis and MC production activities. The site commissioning makes use of several sources of information to assess the readiness of a site to run CMS workflows:

- The average site availability according to the CMS SAM tests
- The success rate of analysis-like jobs submitted by the Job Robot
- The number of commissioned transfer links with other sites

The evaluation of the global site status relies on daily rules to be satisfied for the Tier-1 and T2 sites:

Daily Rules for Tier-1 sites
daily SAM availability ≥ 90%
daily JR-MM efficiency ≥ 95%
having commissioned the downlink with the Tier-0
having ≥ 10 commissioned downlinks to Tier-2 sites
having ≥ 4 commissioned downlinks/uplinks to other Tier-1 sites

Daily Rules for Tier-2 sites
daily SAM availability ≥ 80%
daily JR-MM efficiency ≥ 90%
having a commissioned uplink with at least 1 Tier-1
having a commissioned downlink with ≥ 2 Tier-1 sites

The global SC status is determined as follows:

- COMMISSIONED: daily rules are satisfied during the last 2 days, or during the last day and at least 5 days in the last 7
- WARNING: daily rules are not satisfied in the last day but satisfied during at least 5 days in the last 7
- UNCOMMISSIONED: daily rules satisfied for less than 5 days in the last 7

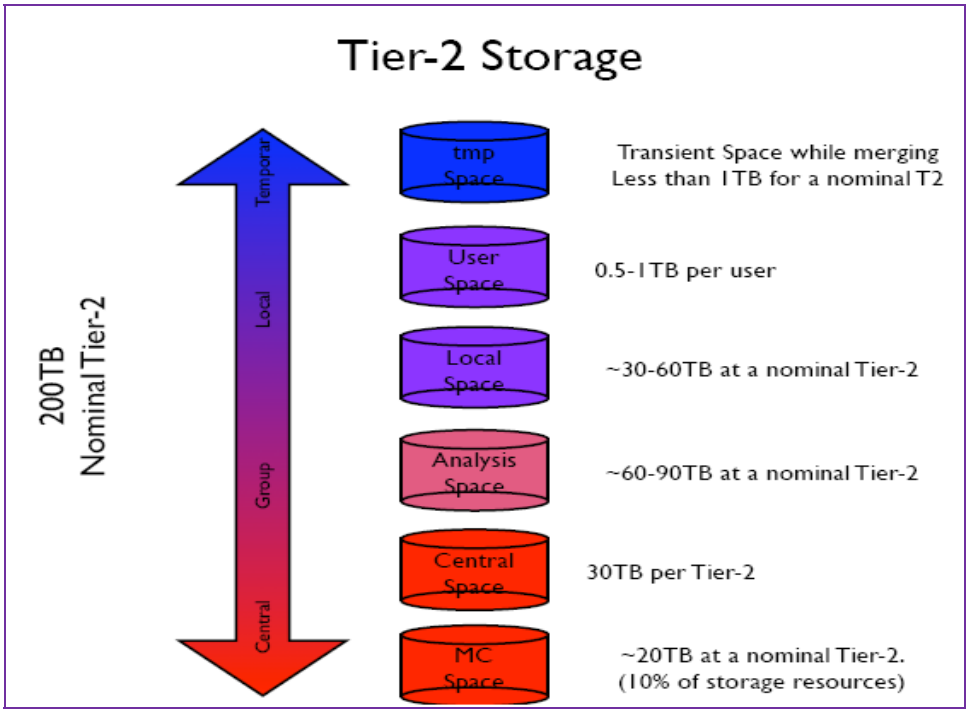
Automating Tier-0 Processing

Many parts of the Tier-0 system were deployed in the last several months. The system is designed around the Computing Model and has been stretched by current and proposed usage, e.g. one Primary Dataset for all data. Open trigger data taking who defines complete luminosity sections. CMS will have to evaluate plans given new data taking expectations.

CMS Analysis at Tier-2 Centres

CMS coordinates which Tier-2 is hosting which physics analysis; this was a delicate and long process to reach agreement.

The Tier-2 disk space is organized as described in the following figure.



Monte Carlo Production

Continuous request of MC samples produced to support commissioning and studies for Primary Dataset definitions.

A large Startup MC production (> 200M events) started with release 2_1_7/2_1_8 for start of data taking. Several samples are being prepared:

- DPG sample: 26M produced, 22M reconstructed
- Physics sample: 97M produced, 14M reconstructed

Production is running smoothly CMS were able to use > 8000 T2 slots (nominal is ~5600)

Data Processing and Distribution

Data taking took place over the summer with CMS global data taking every week (Mid-Week-Runs) and CruZet3 and CruZet4 runs. The CAF was used for the Commissioning and Calibration and Alignment tasks

Data were routinely processed and distributed while the main focus was on actualizing the Tier-0 processing step. Several reprocessing steps were performed at Tier-1 centres.

The data processing is running without problems and is closely steered and monitored by the Data Operations team. The CAF functions at Tier-1 are in operation, tested in CCRC&CSA08 and a major portion of CAF resources is deployed.

The CAF-Tier-0 is CERN-based for CMS scientists use LXBATCH/LXPLUS for individual data analysis.

For user output CMS will use the CASTOR user pool “CMS-CAF-T2” and plans to test another technology for user space in parallel. Others users perform analysis at “their” Tier-2 centre as user space is assigned and allocated mostly by proximity and association.

CMS plans to test another technology for user space in parallel. NFS space provided by BLueArc network storage devices providing storage accessible by the large array of collaboration desktop computing systems at CERN

CMS Plans for Q4/08

CMS will continue their Monte Carlo production with large (full Geant) MC samples and many Fast Simulation productions will be required for the coming months.

In addition data taking, for cosmics and commissioning, is planned until end November

The table below shows the calendar for the CMS tests:

- LR = local readout of individual components
- MWGR = Mid Week Global Run (Wed-Thu)
- GR/CRAFT = Global Run, Oct. 13-27, Nov. 3-17
- CRAFT = CMS Run At Four Tesla

Week 40	Week 41	Week 42	Week 43	Week 44 (local school holiday, Halloween)	Week 45	Week 46	Week 47	Week 48
LR MWGR	LR MWGR	GR	CRAFT	LR MWGR	CRAFT	CRAFT	LR MWGR	LR MWGR

LHCb Report

June 2008 - September 2008

Ph.Charpentier

LHCb had a stable version of its software and was ready for 2008 data taking. Gaudi is based on latest LCG configuration and the applications are ready for using survey geometry, conditionsDB. There are still difficulties in commissioning CondDB access at Tier-1s.

Plans for the shutdown are to use the opportunity for “cleaning” part of the framework, unifying some interfaces and remove obsolete parts.

Studies on multi-core support, for parallel file processing, are taking place in collaboration with LCG-AA / PH-SFT,

Commissioning of DIRAC3

Fully reengineered the DIRAC system and it is now a single framework for services, clients and agent; its main features are:

- Fully integrated Workload and Data Management Systems
- Supports production and user analysis activities. Allow to apply VO policy: priorities, quota
- Uses pilot jobs as DIRAC2. Ready for using generic pilot jobs (not switched on yet). Full scale test with generic pilots will take place in the coming weeks

Production activities

Complete simulation and stripping of MC data, so-called DC06 as was launched in 2006, with CCRC-like activity at low rate (10%). Started 2008 simulation, mainly for alignment and calibration studies, and waiting for first data for tuning generators and the detector's response.

Issues Encountered

Instability of SEs - Very good response from sites and dCache developers but permanent issues due to various causes:

- Software issues, addressed with sites and developers
- Sub-optimal hardware configuration at some Tier-1 sites
- Unavailability of files: are in the namespace at site but cannot be accessed or even get a tURL.
- Damaged tapes, unavailable servers.

The transfer rates are sufficient, also because of the low throughput needed by LHCb (70 MB/s).

Issues with WMS: WMS also caused several issues to the LHCb activities:

- Mixing up credentials of jobs submitted by the same user with different roles
- Limitation in proxy handling: too few delegations are allowed preventing some users to run jobs (e.g. from French CA)
- Misbehavior of WMS after some idle time will not find resources to run jobs.

Outlook for DIRAC 3

DIRAC3 will be fully ready for first data. All LHCB analysis will be fully migrated and by end of September: no dependency any longer on SRM v1 (legacy files).

In the future LHCb will use Glexec on worker nodes and this will allow exploiting the full power of DIRAC with ate binding of jobs, VO policy and running analysis jobs with higher priority without site intervention.

The DIRAC3 pilot jobs model was certified long ago by the GDB working group on the Experiment frameworks but is waiting for the SCAS middleware to be ready.

Commissioning of the alignment and calibration loop will be completed with the setting up an LHCb-CAF (Calibration and Alignment Facility). Its requirements are rather modest (“simple”

detector). Also the full commissioning of Conditions Database update and streaming will take place: currently very few commits to CondDB.

Is not clear whether this is a longer or advanced shutdown. LHCb cannot make much use of cosmics because its detectors are vertical.

More simulation will be needed for replacing 2008 data in commissioning for alignment and calibration and test the Full chain (including HLT)

Calibration and Alignment (CAF)

The control of the calibration is done on the online monitoring farm with a low rate (~5 Hz) calibration stream to the Tier-0.

The calibration loop runs calibration and alignment at CERN (CAF) with dedicated queues set up (currently 20 slots). The validation of calibration will also be at CERN (Tier-0) using the calibration stream.

LHCb Full Dress Rehearsal

The LHCb full dress rehearsal will include the complete commissioning of:

- HLT (for early and further data). Algorithms, memory leaks…
- Event streaming and real data online monitoring. Started with first beams but…
- Automatic transfers and reconstruction (à la CCRC)
- Calibration loop

The preparatory phase will include large samples, so that events can be reused, of minimum bias events and with online injection software (mimicking Front-End electronics).

The expected time scale for this work for January 2009 but the actual date depends on maintenance and cooling at the pit. February or March will be used for the test typically running every other week, interleaved with pit commissioning activities.

Resources for LHCb

LHCb will reassess their needs for 2009 for CERN in particular. It does not plan to change the resource needs after the LHC incident. The 2008 data was small anyhow, and therefore there has no influence on the future.
LHCb will give priority to the required activities and verify all Grid services (storage, transfers, WMS, etc) and needs to have better control on user activities in order, for example, to avoid a single user exhausting the complete VO's share.