



WLCG Status Report

19th November, 2008
LHCC Open Session



Ian Bird
LCG Project Leader





Agenda

- WLCG service status and CCRC'08
- Planning for shutdown
- Applications Area status
- Procurement issues
- Resource Scrutiny and requirements
- Planning for 2010 (EGI etc)
- Planning for CERN infrastructure



Status of MoU signing and federations

- All originally planned Tier 1 and Tier 2 sites have now signed the MoU
 - Czech Republic earlier this year, Final signature – Austria – last week.
- Additional recent Tier 2s
 - Korea (CMS)
 - France – Strasbourg (ALICE+CMS)
 - Germany – Goettingen (ATLAS), DESY (LHCb)
 - Canada – Canada East+West (ATLAS)
- Brazil is still planning to sign the MoU as a Tier-2 supporting all 4 experiments
 - Letter sent by J. Engelen in June 2008
 - WLCG MoU wording and future commitment being examined by their legal experts

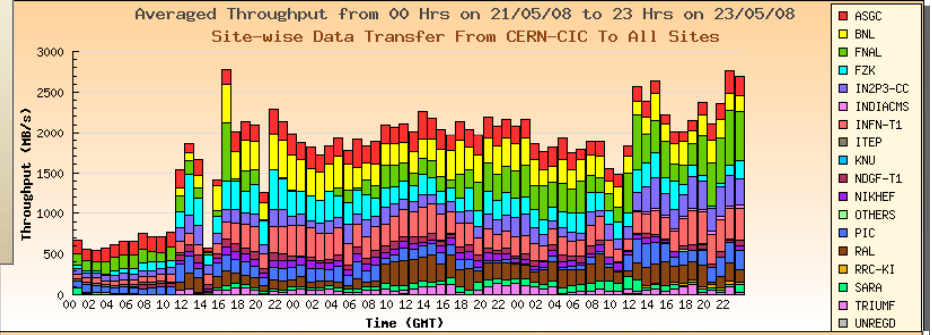
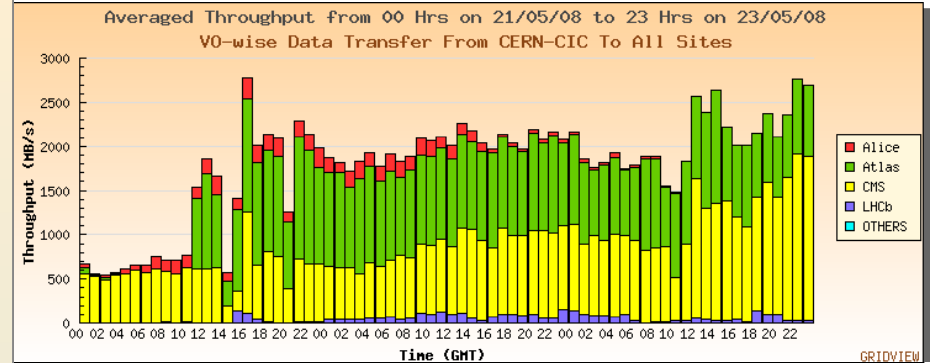
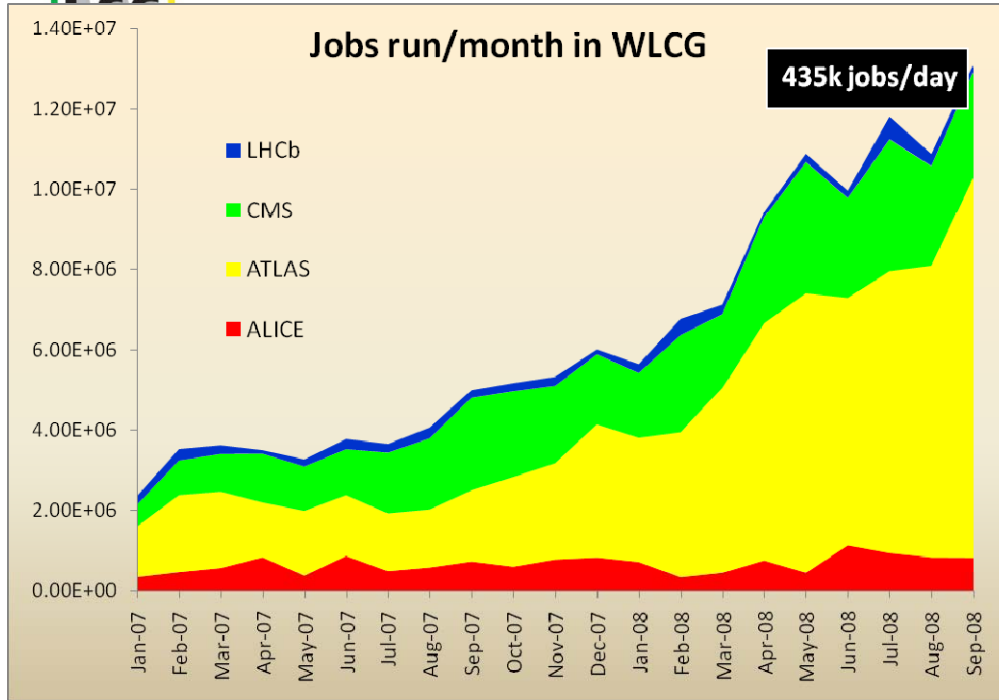


CCRC'08 & the WLCG Service

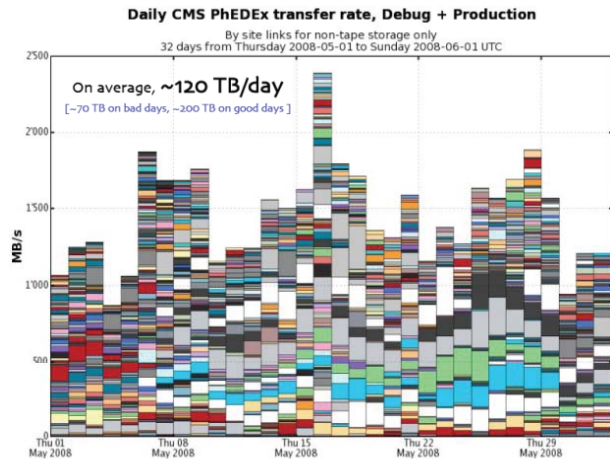
- **CCRC'08:**
 - Was run as planned in May; not all resources were in place
 - Experiments and WLCG service met or exceeded most of the targets
 - Not tape recall/re-processing for >1 experiment at Tier 1s
 - Analysis not at a scale of 100's of users
 - Effort was sustainable (and is sustained), daily operations meeting, post mortems
 - Sites have response/alarm procedures in place (outstanding milestone on 24x7 support now completed)
 - Software process shown to work well – deployment of security and other patches as part of usual process

- **After May:**
 - Service continues as standard production service
 - Continually increasing workloads – simulations, cosmic data, functional tests of all aspects

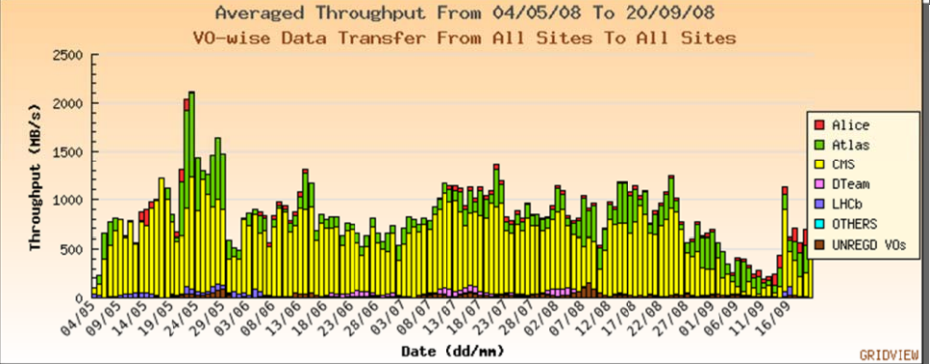
CCRC'08 and beyond

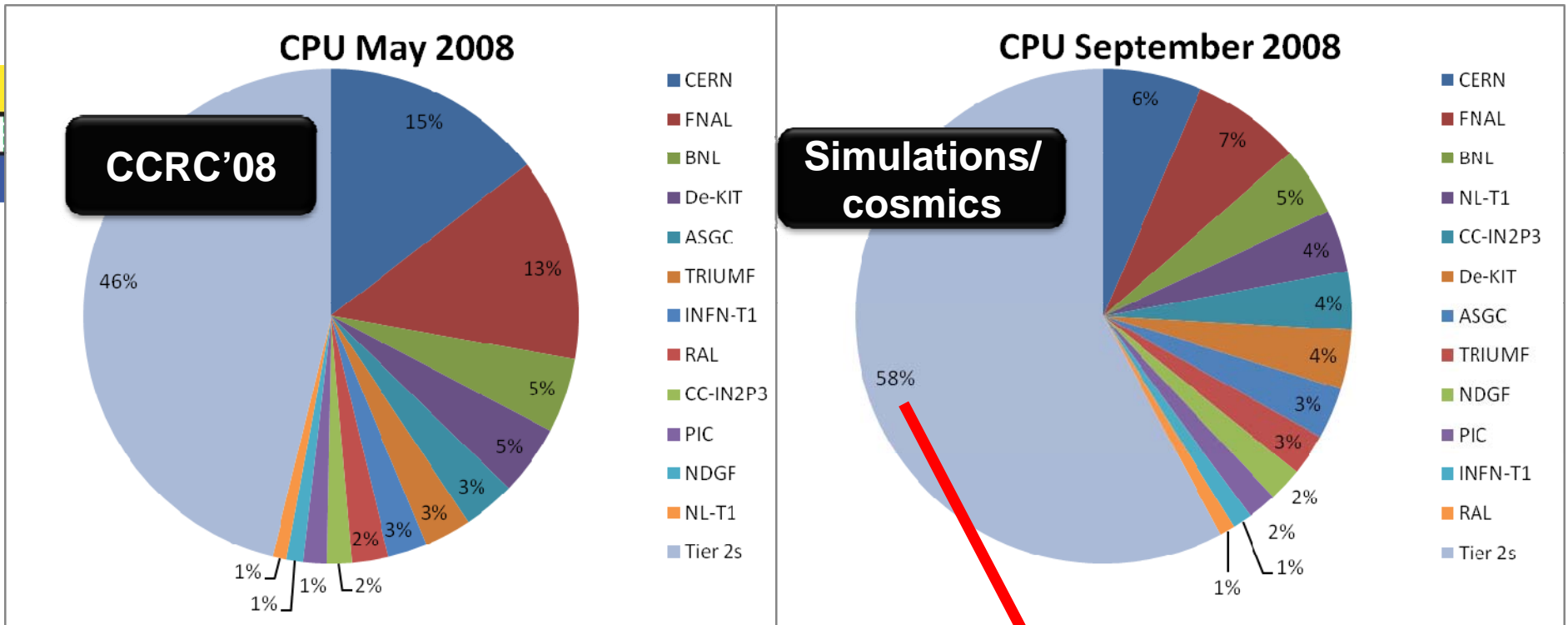


Tier-x to Tier-x in CCRC'o8/phase-2



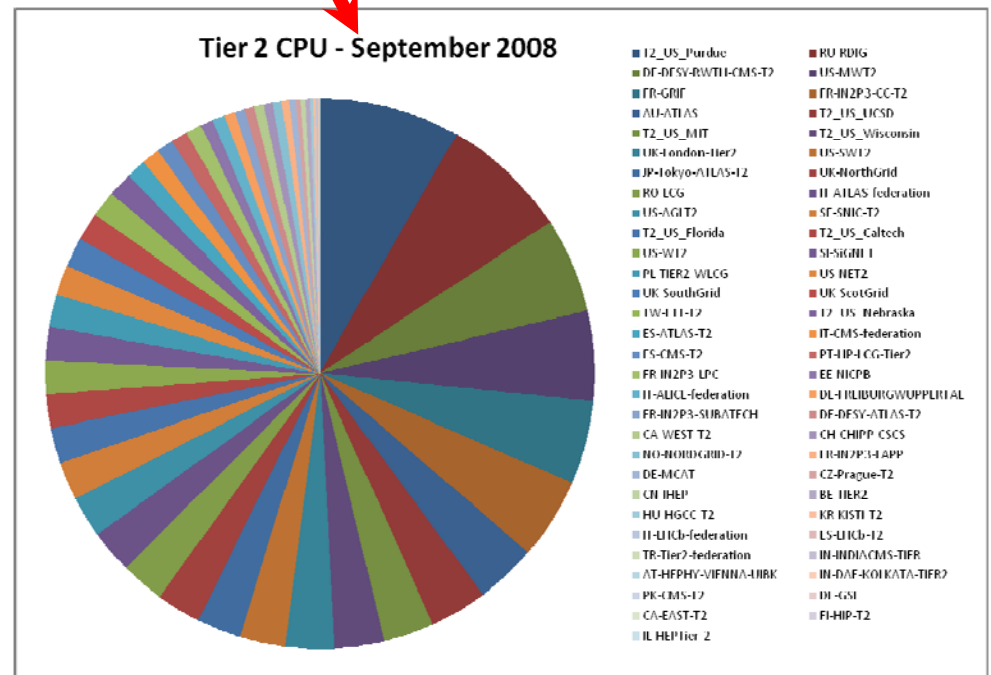
Impressive list of few hundreds of links...





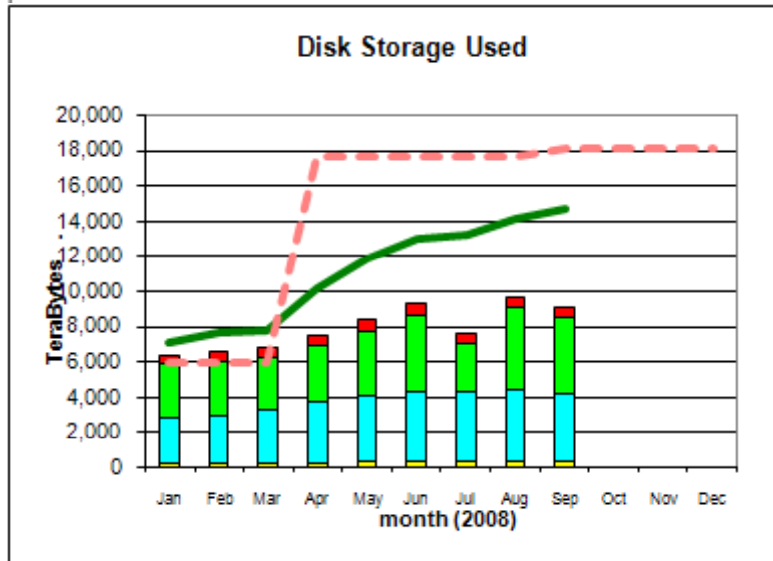
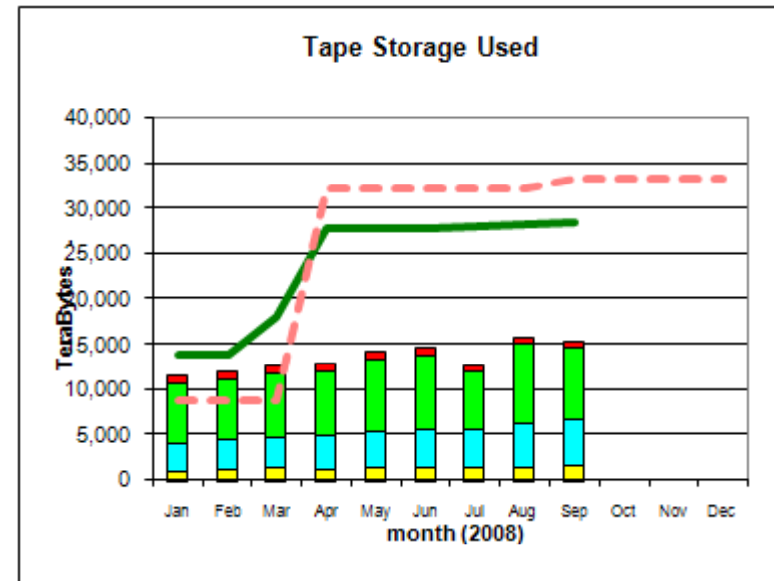
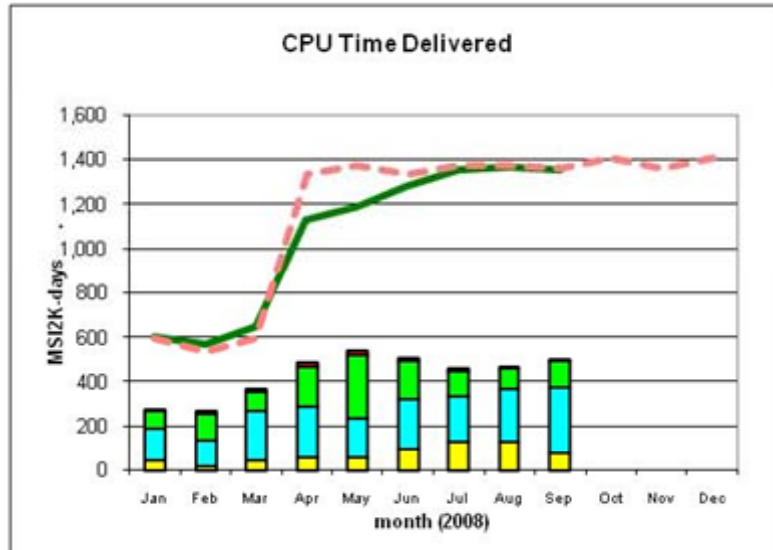
Usage Patterns

- Can change significantly e.g. between CCRC'08 in May and cosmics/simulations in September
- Tier 2s consistently deliver ~50% of total





CERN and Tier-1 Accounting January-September 2008



ALICE		CMS	
ATLAS		LHCb	
installed capacity (inc. efficiency factor)			
MoU commitment (inc. efficiency factor)			



Outstanding service issues

- Levels of support at many sites – especially during the summer
 - It is not yet clear that there are sufficient staff in place to support the level of service needed
- Continued rate of significant downtimes e.g. that trigger a “post-mortem” (Service Incident Review)
 - More than 1 per week and is not improving
 - Many are due to power/cooling issues (and recovery process!)
 - Many are due to issues with storage services – but some are surely lack of adequate service resources or support
- Main management focus is now on
 - Trying to improve service deployments
 - Understanding why we have so many incidents
 - This is not a “grid” issue, it is really a traditional service management issue
 - Absolutely essential that resolving these issues are seen as priorities by site management

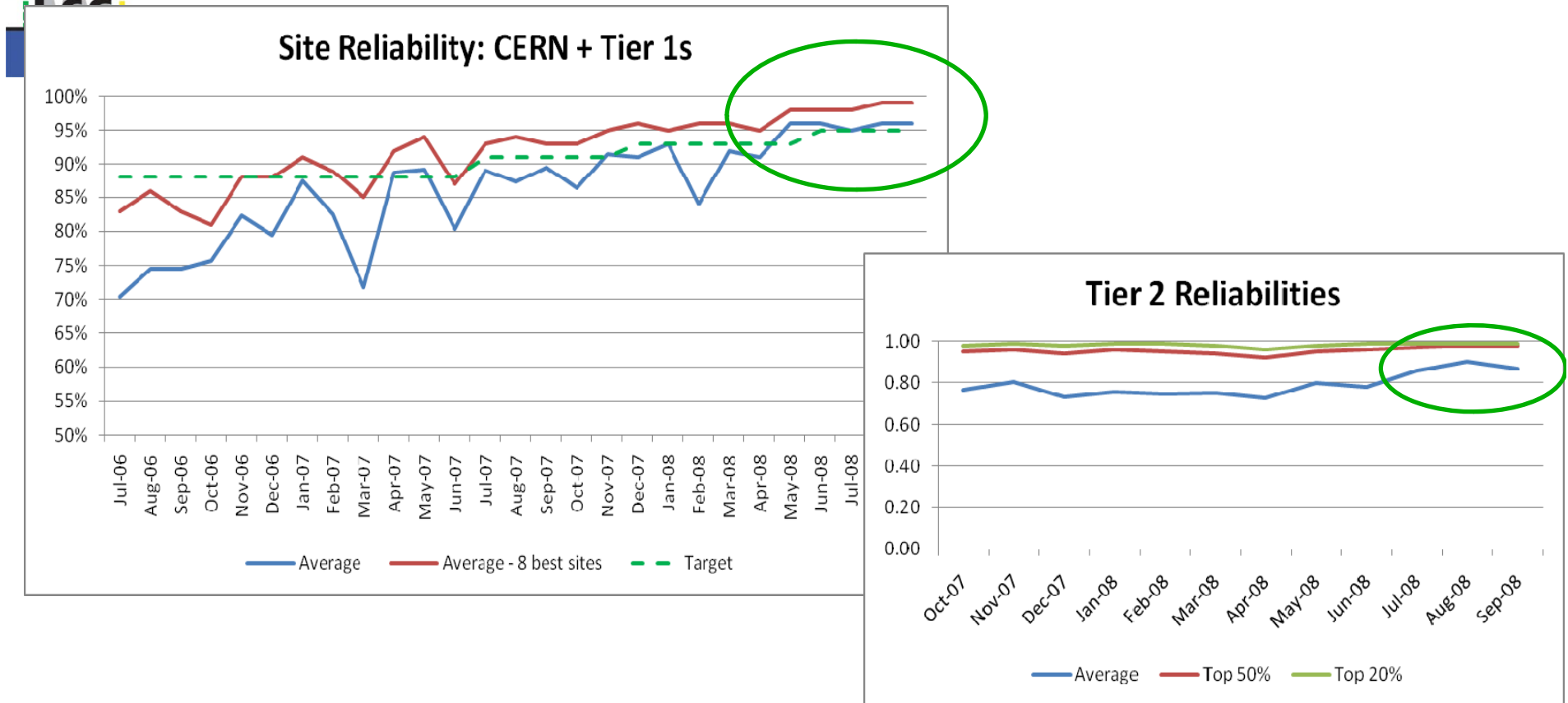


Service incidents - since May

Type of problem	#	<impact time>
Storage	8	4 days
Electrical	3	4.5 days
Cooling	2	4 days
Hardware	2	3 days
Database	2	3 days
Middleware	1	4 hours

- ~ 1 major incident per week
 - not improving
- <impact time> ~ 3.5 days!
 - this is worse now than it was
- 7/20 are hardware (electrical, cooling, network, etc)
 - May be mostly unavoidable, but recovery must be rapid and tested!
 - Power and cooling expectations
 - 1 per site per year ⇒ 1 /month at a Tier 1 !

Reliabilities



Improvement during CCRC and later is encouraging

- Tests do not show full picture – e.g. Hide experiment-specific issues,
- “OR” of service instances probably too simplistic

- publish VO-specific tests regularly;
- rethink algorithm for combining service instances



Tier-2 Availability and Reliability Report

Federation Summary - Sorted by Reliability

September 2008

Critical SAM Tests - <http://sam-docs.web.cern.ch/sam-docs/docs/htmldocs/MANUserManual/node22.html>

Availability = % of successful tests

Reliability = Availability / Scheduled Availability

Reliability and Availability for federation - average of all sites in the federation

Colour coding : N/A < 30% < 60% < 90% >= 90%

Federation	Reli-ability	Avail-ability	Federation	Reli-ability	Avail-ability
US-SWT2	100 %	100 %	IT-ALICE-federation	92 %	88 %
T2_US_Wisconsin	100 %	100 %	IT-ATLAS-federation	92 %	88 %
FR-GRIF	100 %	99 %	IT-CMS-federation	92 %	88 %
AT-HEPHY-VIENNA-UIBK	100 %	100 %	IT-LHCb-federation	92 %	88 %
US-MWT2	100 %	100 %	CZ-Prague-T2	91 %	91 %
FI-HIP-T2	100 %	99 %	BE-TIER2	91 %	67 %
CN-IHEP	99 %	99 %	ES-CMS-T2	91 %	89 %
FR-IN2P3-SUBATECH	99 %	99 %	DE-FREIBURG WUPPERTAL	91 %	91 %
T2_US_UCSD	99 %	99 %	T2_US_Nebraska	91 %	93 %
US-NET2	99 %	99 %	CA-WEST-T2	90 %	87 %
UK-NorthGrid	98 %	98 %	KR-KISTI-T2	89 %	66 %
T2_US_Purdue	98 %	98 %	US-WT2	88 %	91 %
FR-IN2P3-LPC	98 %	97 %	UK-London-Tier2	88 %	74 %
FR-IN2P3-CC-T2	97 %	97 %	RO-LCG	87 %	83 %
TW-FTT-T2	97 %	97 %	FR-IN2P3-LAPP	86 %	82 %
JP-Tokyo-ATLAS-T2	97 %	95 %	ES-LHCb-T2	85 %	85 %
DE-DESY-ATLAS-T2	97 %	90 %	T2_US_Caltech	83 %	86 %
PT-LIP-LCG-Tier2	96 %	48 %	IL-RDIG	81 %	81 %
T2_US_Florida	96 %	97 %	IL-HEPTier-2	78 %	50 %
DE-MCAT	96 %	81 %	SE-NORDS	67 %	68 %
UK-ScotGrid	96 %	93 %	TR-Tier2-federation	66 %	65 %
CH-CHIPP-CSCS	96 %	93 %	PK-CMS-T2	62 %	26 %
US-AGLT2	96 %	96 %	AU-ATLAS	51 %	48 %
PL-TIER2-WLCG	95 %	94 %	IN-INDIA-CMS-TIFR	46 %	42 %
HU-HGCC-T2	95 %	95 %	IN-DIXE-KOLKATA-TIER2	1 %	1 %
CA-EAST-T2	95 %	95 %	DE-GSI	0 %	0 %
SI-SIGNET	95 %	94 %	NO-NORDGRID-T2	N/A	0 %
ES-ATLAS-T2	93 %	91 %	SE-SNIC-T2	N/A	N/A
DE-DESY-RWTH-CMS-T2	93 %	93 %	UA-	N/A	N/A
UK-SouthGrid	93 %	88 %			
T2_US_MIT	92 %	93 %			

Tier 2 reliabilities

- Big improvement
- Federation average will be (soon) weighted by #CPU
- Would like to fix target at 95%
 - Should be achievable

■ e.g. of extended scheduled downtimes (availability << reliability)

■ Federations still not reporting



Consequences of LHC shutdown

- The present shutdown of the LHC has a number of consequences for the planning of WLCG:
 - Capacities and Procurements for 2009
 - Software and service upgrades during the shutdown
 - (Re-)Validation of services for 2009 following changes



Capacities and procurements

- 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 x2 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.



Upgrade plans

- Since several software upgrades were postponed in anticipation of LHC start-up, we propose that the following changes are addressed in the coming months:
 - SRM – agreed list of “short term” changes; available by end 2008
 - FTS on SL4 (+available for SL5?) – deployment was postponed
 - WN on SL5 to be available for deployment
 - glexec/SCAS to support pilot jobs with identity changing
 - CREAM CE – make available in parallel to existing CE which is known to have scaling issues when there are many different users;
 - needs Condor_g client
 - WMS: must be able to submit to CREAM
- + a few other smaller changes ...

- Many of the above are deployments in parallel to existing production services and so non-disruptive



Re-validation of the service

- All experiments are continually running simulations, cosmics, specific tests (and have been since CCRC'08) at high workload levels – this will continue
- A full CCRC'09 in the same mode as 2008 is not regarded as useful
- But, we will perform specific tests/validations:
 - Service validation if software is changed/updated
 - 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 were discussed in the workshop last week



Applications Area Status

- No major releases of the AA software during last quarter
 - Experiments preparing for beam did not want major changes
 - Substantial progress on porting the complete software stack to other platforms such as gcc 4.3 and VC9.
- Preparing 'production' versions for ROOT and Geant4
 - ROOT 5.22 will include the new CINT based on Reflex and the support for the new 'data model schema evolution'
 - Geant4 9.2 will include the improvements in the FTF (Fritiof) hadronic model for pion incident interactions; alternative multiple-scattering models, and the first implementation of a GDML writer.
 - New version of HepMC (2.04.00) was released last summer and adopted by all experiments
- Release schedule adapted to new LHC schedule
 - ROOT and Geant4 releases for mid-December
 - The rest of the software stack should be ready by mid-January
 - Experiments will start integrating their applications immediately after



Applications Area Status (2)

- The LCG-AA nightly build and test system fully operational
 - Several configurations in parallel for all supported platforms (including MacOSX and Windows)
 - Geant4 builds and tests are also integrated
 - LHCb has cloned the system for their software
- Continuous Software Integration and Testing
 - Allow us to release new software versions (with new functionality or bug fixes) on demand from the experiments with relatively short notice (1-2 days)
- The two Theme 3 PH-R&D projects monitored by the AF
 - Virtualization: The first release of the CernVM virtual appliance to provide a portable environment for analysis is ready
 - Multi-core: Exploring in collaboration with experiments and IT (openlab) a number of possibilities to exploit multi-core architectures with the current software frameworks



Procurement Issues

- 2008 resources were not in place by April; and in some cases still not now:

CPU (100% of total pledge installed)		
ASGC	72%	Due October
NL-T1	88%	Not before mid-2009:

Disk (80% of total pledge installed)		
ASGC	300 TB missing	Installation now?
BNL	1 PB missing	Was expected November
IN2P3	700 TB missing	Ongoing, with 50% of 2009
CNAF	750 TB missing	Installation ongoing
NDGF	200 TB missing	Procurement complete
NL-T1	1400 TB missing (50%)	Power/cooling; not before mid-2009

- Many procurement issues: delays, faulty equipment, vendor failures.
- No indication that this will be any better in future years – must take into account in the procurement process
 - For 2009 have added checkpoint milestones to follow up on process
 - Little margin for error for April schedule
 - Future years (>2009) anticipate splitting disk installations into 2 – first for April, second for end of Summer (details to be discussed and agreed)



Resource Scrutiny Group

- Process: scrutinise the requirements (normally) presented in Spring RRB:
 - Resource accounting for previous year
 - Use made of the resources
 - Overall requirements for the next year (and for +2 years)
 - Examine the match between requirements and pledges
 - Make recommendations in case of apparent under-funding
- This is the first scrutiny, requirements based on the TDRs (2005) + changes since as elements of the computing models have been tested
 - In particular ATLAS doubled the CERN (Tier 0 + CAF) request for 2009
 - Next reassessment of requirements should be with 6 months of beam experience
- Scrutiny took a standard set of assumptions on beam time
 - 3 months 2008 + 7 months 2009 → 7 months 2009 (after Sep 19); includes 1 month AA
 - Implemented simplified versions of the Computing Models for comparison (therefore cannot just take the numbers from RSG at face value)



Summary of conclusions

- Given that the ...
 - Experiment models are accurate only to ~10-20% before we have more experience
 - The RSG models are quite simplified compared to the experiment models
- ... the scrutiny agrees with the experiment requirements at this level (including the increased ATLAS request at CERN)
 - ALICE:
 - Reasonable; disk requests underestimated at Tier 1s, overestimated at CERN; tape request too large as assumed 2007 start
 - Unlikely that ALICE requests will be met; recommend ALICE make statement to LHCC how physics will be affected
 - ATLAS:
 - TDR model was optimistic for event sizes, data formats, etc. Parameters of the model should be re-examined.
 - Concern over event sizes and proliferation of data formats
 - Worry that resource needs for ≥ 2010 may be hard to find
 - Agree that CERN request is justified but concern over de-emphasis of Tier 1s
 - CMS:
 - Good agreement with CMS model; CMS made good progress in understanding model details
 - RSG compute far less need for Tier 2 resources (but assumptions not agreed)
 - LHCb:
 - LHCb model is valid and solid
 - Slight reduction in disk requirement without 2008 data; similar conclusion for tapes



RSG comments for LHCC

- Most experiments propose using increased trigger rates as compared to the ones stated in the TDR reviewed by the LHCC. We feel we are not sufficiently competent to review the need or convenience to do so.
- ALICE wants to increase substantially their amount of pp data; in particular they stress the benefit of acquiring data at 10TeV. We have not assessed these needs from the physics point of view and we do not know whether such lower energies will be available in the 2009 run or anytime in the future.
- One of our conclusions is to recommend that ALICE undertakes a full assessment of how their physics reach might be affected by requested computing resources not materializing.
- The event size has a very direct impact on the computing requirements. Some experiments, such as CMS, have made an effort to reduce the event sizes by establishing a reduction profile after startup. We believe that this example should be followed by all experiments.
- We take note of potential modifications of the computing models due to the use of different data formats serving the same purposes, not always well justified.
- The realization of the computing model for ATLAS seems to differ slightly from the implementation originally envisaged in the TDR for reasons discussed in the report. This implies, in particular, heavier demands on CERN resources. We believe these demands are largely justified, however.
- Cosmic data taking is now much emphasized by experiments; while it is clear that cosmics are extremely useful in commissioning for calibration, this data is by nature transient and it seems somewhat questionable to us to support substantial requests based on cosmic runs, but we do feel we have not sufficient insight to make a definite scientific judgement on this.



Experiment requirements - next steps

- The C-RSG report validates the experiment requirements – within the limits of how well the models are currently understood
 - No useful re-evaluation is possible without ~ 6 months real data taking experience
- ATLAS request at CERN is now doubled compared to pledged resources
- How should this be managed? Require guidance on priorities.
 - The existing CERN resources (and MTP) based on the understanding at the time of the TDRs
 - Increasing the ATLAS allocation would be at the cost of other experiments ...



Tier 0/CAF resource balance

CERN Tier0	Split 2009	ALICE	ATLAS	CMS	LHCb
CPU (kSI2K)	Offered	9000	4058	9800	1050
	Required	9000	7600	9800	1050
	% of Req.	100%	53%	100%	100%
Disk (Tbytes)	Offered	4200	265	200	991
	Required	4200	650	200	991
	% of Req.	100%	41%	100%	100%
Tape (Tbytes)	Offered	7300	5562	7300	2270
	Required	7300	8560	7300	2270
	% of Req.	100%	65%	100%	100%

CERN Analysis Facility	Split 2009	ALICE	ATLAS	CMS	LHCb
CPU (kSI2K)	Offered	2600	2562	3900	0
	Required	2600	5800	3900	0
	% of Req.	100%	44%	100%	100%
Disk (Tbytes)	Offered	300	1809	2300	0
	Required	300	3300	2300	0
	% of Req.	100%	55%	100%	100%
Tape (Tbytes)	Offered	0	651	2000	0
	Required	0	1130	2000	0
	% of Req.	100%	58%	100%	100%



Tier 1+2 Pledge Balance in 2009

- The table below shows the status at 17/11/08 for 2009 from the responses received from the Tier-1 and Tier-2 sites
 - The Total 2009 pledge from Russia is included but not the split across the experiments

	ALICE	ATLAS	CMS	LHCb	Sum 2009
T1 CPU	-49%	6%	-2%	2%	-12%
T1 Disk	-43%	-5%	-13%	-2%	-13%
T1 Tape	-50%	-7%	7%	6%	-13%
T2 CPU	-41%	0%	-3%	-33%	-8%
T2 Disk	-42%	-19%	38%	-	-1%



Pledge Balance 2008-2013

- The table below shows the global picture for 2008-2013, status as of 17/11/08. % indicates the balance between offered and required
- Some Federations have recently signalled a change to procurements for 2009, not supported by WLCG Management or Overview Boards

	2008	2009	2010	2011	2012	2013
T1 CPU	-5%	-12%	-11%	-15%	-20%	-26%
T1 Disk	-12%	-13%	-15%	-18%	-24%	-29%
T1 Tape	-13%	-13%	-16%	-22%	-24%	-23%
T2 CPU	-2%	-8%	-29%	-31%	-32%	-37%
T2 Disk	-12%	-1%	3%	-6%	-6%	-17%



Planning for 2010 (end of EGEE)

- A second draft of the EGI blueprint has been produced
- There are still some serious shortcomings in the process and in the blueprint:
 - It is not clear exactly what is being proposed in terms of the roles and functions of the National and central organisations;
 - Lack of representation of the user communities, and how those communities interact with the infrastructures; (they own many resources)
 - It is not clear how the present operational infrastructure upon which WLCG depends will evolve and appear in the future, e.g.:
 - Insufficient resources for central operations
 - Risk of discontinuation of ROCs
 - User support is being challenged
 - 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.



WLCG & EGI/NGIs

- The Tier 1 and Tier 2 sites in Europe will rely on National Grid Infrastructures being in place to provide the support and functions today provided by EGEE
 - Important for operations and middleware support (maintenance and distribution)
 - The Tier 1s have provided position statements
 - Still important that WLCG members raise this to the NGI and national funding agencies
- The Tier 0 is probably in a reasonable position – current planning does not rely on external funding; but the capability will be strictly limited to core WLCG Tier 0/CAF tasks
- It seems optimistic that EGI will be in place by the end of EGEE-3, and likely that not all NGIs will be in existence when it does start
- WLCG must have a concrete plan to operate without relying on the European level support, either for an interim period or indefinitely



Planning for EGI

- Agreed in OB that we will now document how each of the core functions that we today rely on will be managed in future
 - Start with the position statements sent by Tier 1 sites
 - Consider each of the major functions:
 - GGUS, operational support, monitoring tools, middleware support, certification, deployment support, service management, etc.
 - Work with EGEE to understand the expected status of each of these in mid-2010
 - Negotiate who will manage /contribute to each function if there is no EGI



Planning for Tier 0 infrastructure

- Capacity in CERN CC will run out in ~2010; electrical capacity cannot be extended above currently foreseen levels
- Strategy:
 - Expand the capacity of the building as far as possible (2.5 → 2.9 MW), addition of water-cooled racks: NB. This leaves no redundancy in the power supply;
 - Aggressive removal & replacement of older equipment with new lower-power units. Replace at end of warranty (3 yrs);
 - Planning for a second centre to be built on Prévessin site;
 - Investigate stop-gap solutions for 1.5 – 2 years between running out of power and having a new building available
- First 2 points + better estimate of power evolution of new systems → sufficient capacity until ~end 2010
- Planning ongoing at the level of conceptual designs
- Stop gap solutions being investigated; commercial solutions very expensive (x10 over CERN cost), some possibilities under discussion with other WLCG sites.



Summary

- CCRC'08 was successful
 - Almost all experiments' and service targets were achieved
 - Exception: user analysis with 100's of users; T1 reprocessing at full scale
- Service has continued to be used at significant levels
 - Main focus is on improving service reliability – especially storage systems
- Important that resource ramp-up for 2009 continues:
 - Must be ready for the accelerator start-up, even if resources are today not saturated
- Resource procurements/installations –
 - Were significantly delayed in 2008
 - Concern that this does not improve in future years
- Planning for future – Tier 0/CAF and European Grid infrastructure