



WLCG Status Report

23rd February 2009
Overview Board

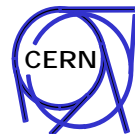
Ian Bird
LCG Project Leader



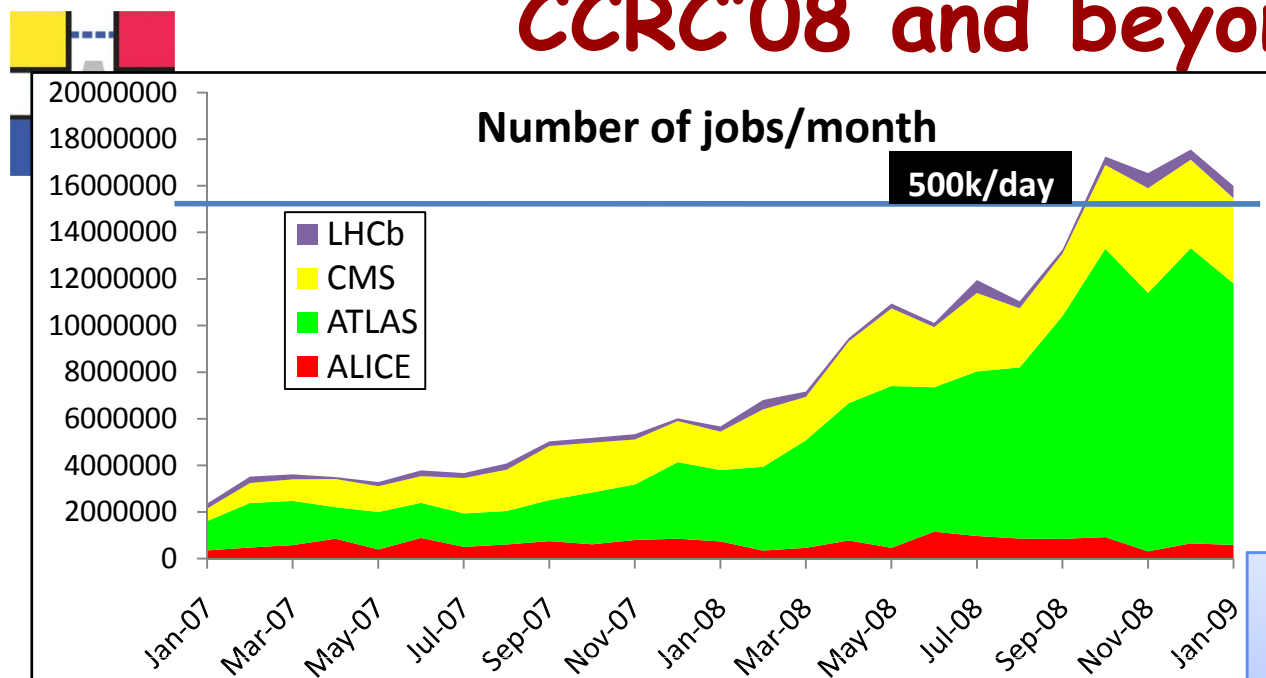


WLCG MoU Signature Status (Nov C-RRB)

- Since the April 2008 C-RRB the Czech Republic has signed the WLCG MoU as a Tier-2 supporting ALICE and ATLAS
- All Tier-1s have now signed
- All of the Tier-2s have signed except Austria
 - Signature expected before mid November 2008 (done)
- A new MoU will be signed on 11/11/08 with Korea as a Tier-2 supporting CMS (done)
- 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

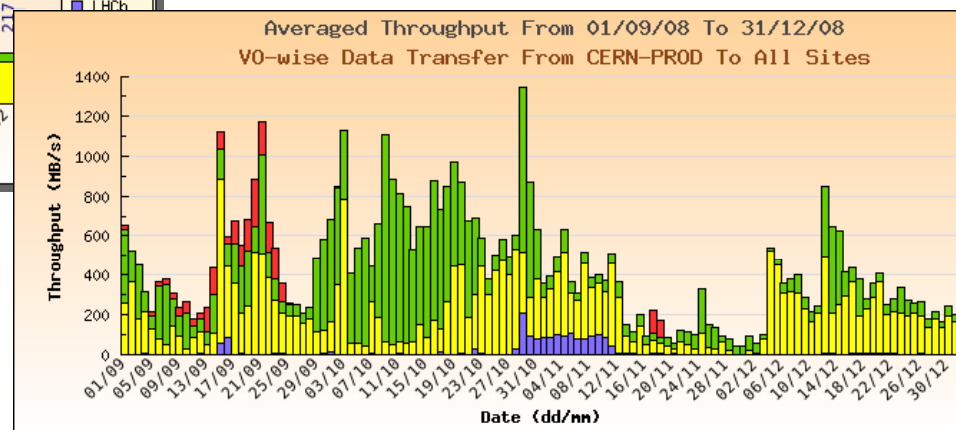
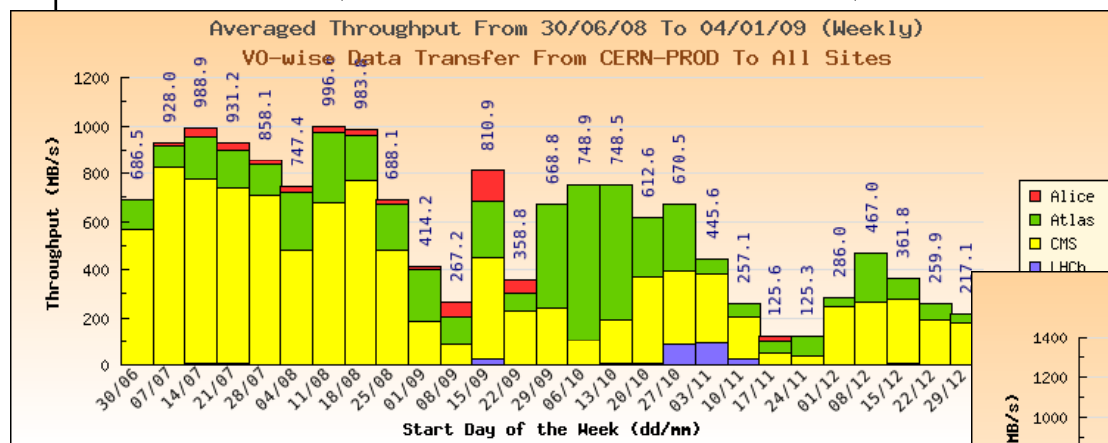


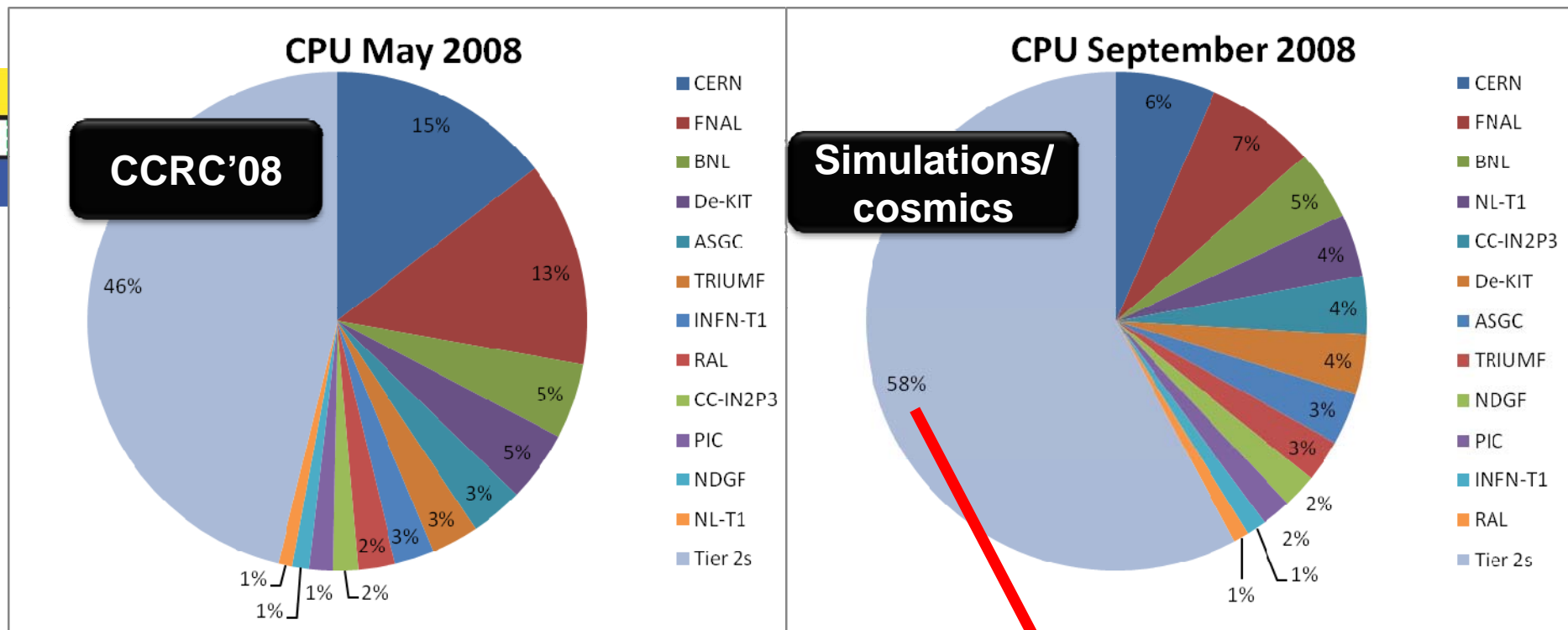
CCRC'08 and beyond



Main outstanding issues related to service reliability.

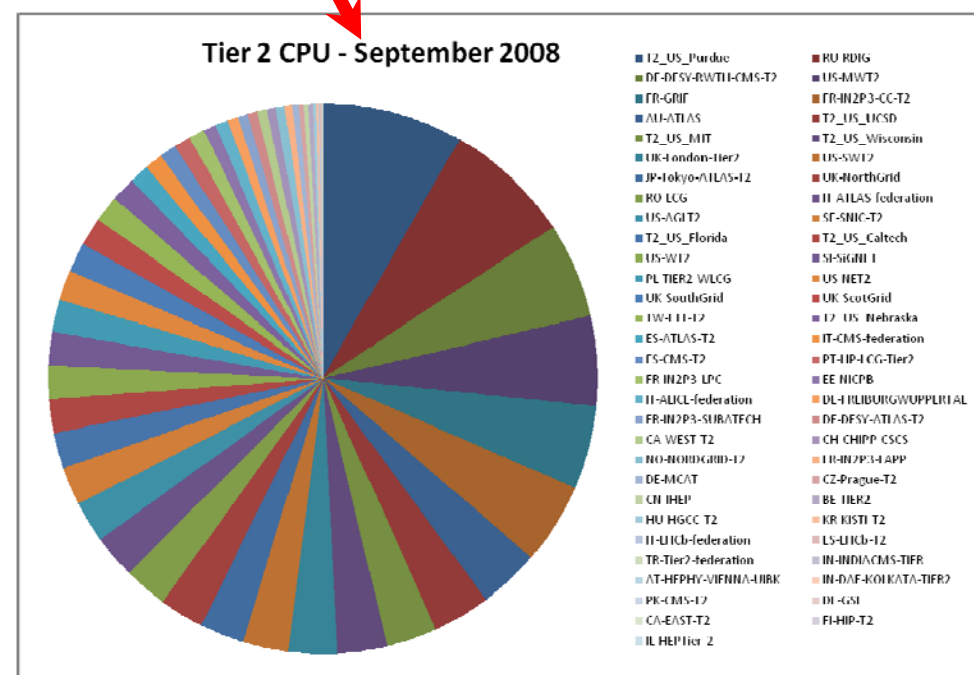
See details in following talks





Usage Patterns

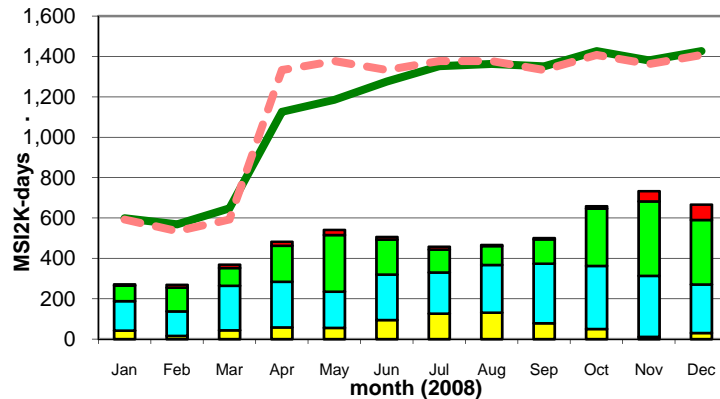
- Can change significantly e.g. between CCRC'08 in May and cosmics/simulations in September
- Tier 2s consistently deliver ~50% of total



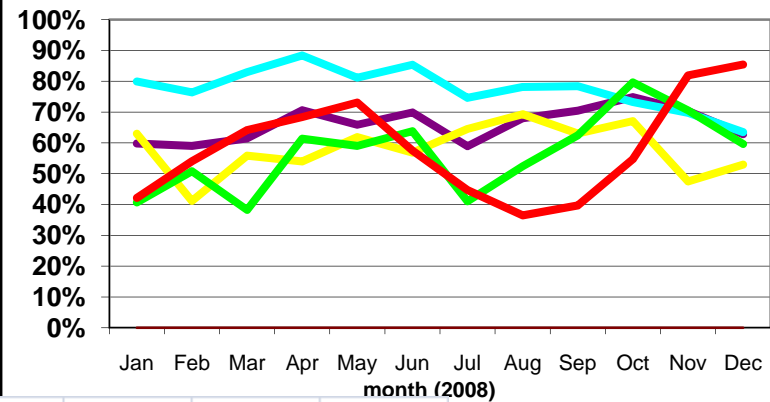


CERN + Tier 1 accounting - 2008

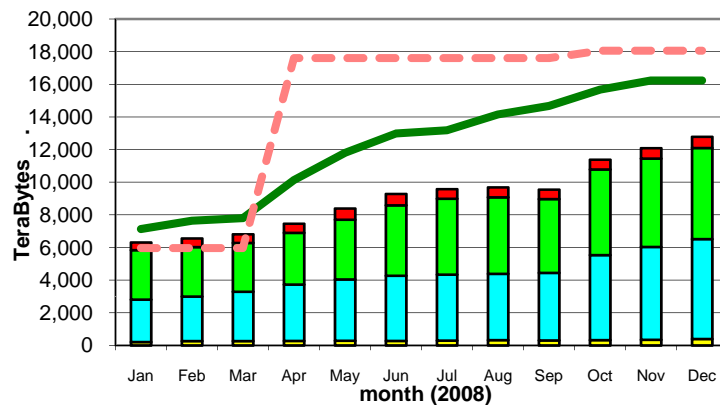
CPU Time Delivered



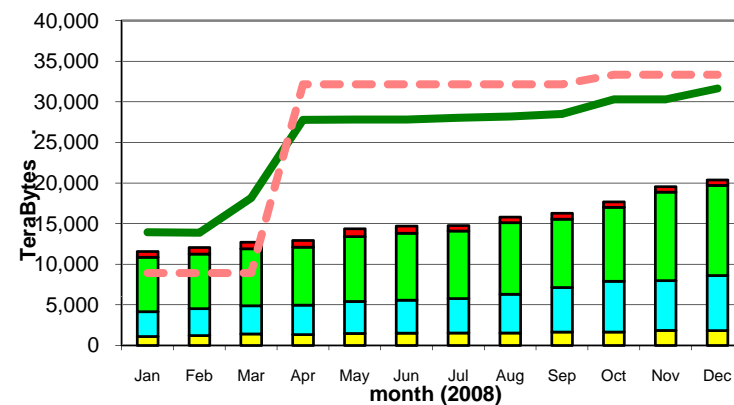
Ratio of CPU : Wall_clock Times



Disk Storage Used

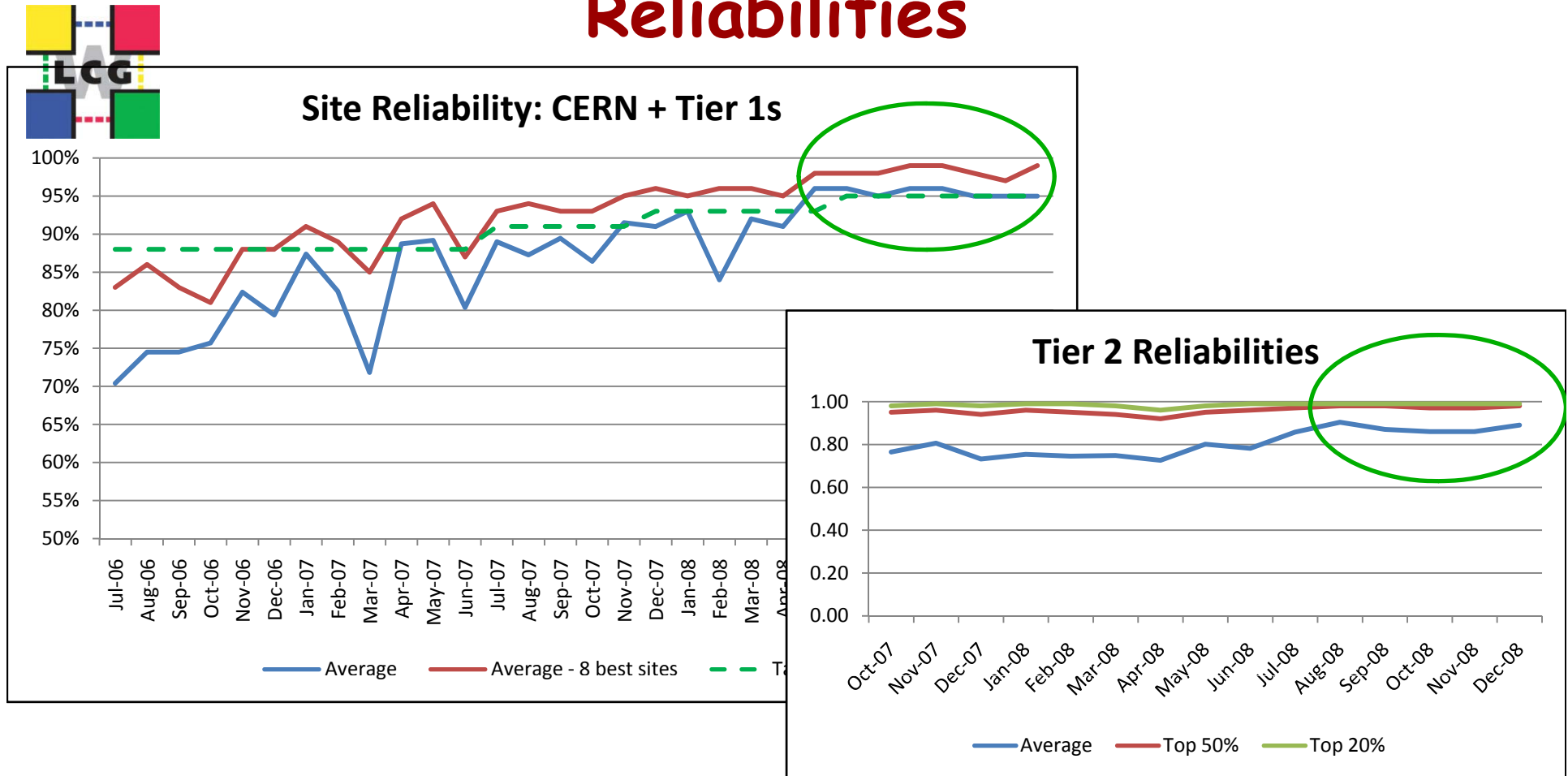


Tape Storage Used



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

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

December 2008

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

Availability = % of successful tests

Reliability = Availability / Scheduled Availability

Reliab

Colour

Reliability	Colour	Federation	Site	CPU Count	Reliability	Availability	Reliability History		
							Sep-08	Oct-08	Nov-08
Fed		AT-HEPHY-VIENNA-UIBK (Austria, Austrian Tier-2 Federation)							
T2_U			HEPHY-UIBK	262	92 %	92 %	100 %	100 %	99 %
T2_U			Hephy-Vienna	324	97 %	97 %	96 %	96 %	95 %
FR-C		AU-ATLAS (Australia, University of Melbourne)							
FI-HI			Australia-ATLAS	80	93 %	93 %	98 %	97 %	94 %
US-M			Australia-UNIMELB-LCG2	27	N/A	0 %	3 %	0 %	0 %
PT-L		BE-TIER2 (Belgium, Belgian Tier-2 Federation)							
UK-S			BEgrid-ULB-VUB	332	14 %	12 %	95 %	82 %	98 %
CZ-P			BelGrid-UCL	579	96 %	96 %	88 %	97 %	97 %
JP-Ti		CA-EAST-T2 (Canada-East Federation)							
RO-L			TORONTO-LCG2	240	41 %	41 %	95 %	97 %	45 %
DE-E		CA-WEST-T2 (Canada-West Federation)							
FR-II			ALBERTA-LCG2	44	98 %	98 %	99 %	98 %	92 %
US-M			SFU-LCG2	64	92 %	92 %	87 %	92 %	93 %
TW-F			VICTORIA-LCG2	65	95 %	95 %	84 %	93 %	91 %
SI-SiCone					97 %	97 %	98 %	97 %	97 %
UK NorthGrid					97 %	96 %	97 %	96 %	96 %
NO-NORDGRID-T2					97 %	97 %	97 %	97 %	97 %
DE-DESY-RWTH-CMS-T2					97 %	96 %	97 %	96 %	96 %
DE-MCAT					97 %	97 %	97 %	97 %	97 %
T2_US_Purdue					95 %	96 %	95 %	95 %	95 %
FR-IN2P3-LAPP					96 %	73 %	95 %	95 %	95 %
UK-London-Tier2					96 %	95 %	95 %	95 %	95 %
FR-IN2P3-IPHC					96 %	95 %	95 %	95 %	95 %
HU-HGCC-T2					95 %	95 %	95 %	95 %	95 %
AT-HEPHY-VIENNA-UIBK					95 %	95 %	95 %	95 %	95 %
CA-WEST-T2					95 %	95 %	95 %	95 %	95 %
RU-RDIG					95 %	94 %	95 %	95 %	95 %
UK-SouthGrid					95 %	90 %	95 %	95 %	95 %
FR-IN2P3-SUBATECH					94 %	94 %	95 %	95 %	95 %
US-AGLT2					94 %	94 %	95 %	95 %	95 %
IN-DAE-KOLKATA-TIER2					94 %	71 %	95 %	95 %	95 %
IL HEPTier 2					87 %	87 %	95 %	95 %	95 %
DE-FREIBURG WUPPERTAL					85 %	72 %	95 %	95 %	95 %
IN-INDIACMS-TIFR					85 %	82 %	95 %	95 %	95 %
EE-NICPB					85 %	85 %	95 %	95 %	95 %
CH-CHIPP-CSCS					84 %	83 %	95 %	95 %	95 %
ES-CMS-T2					84 %	81 %	95 %	95 %	95 %
PK-CMS-T2					80 %	77 %	95 %	95 %	95 %
KR-KNU-T2					80 %	80 %	95 %	95 %	95 %
US-SWT2					78 %	51 %	95 %	95 %	95 %
T2_US_UCSD					78 %	79 %	95 %	95 %	95 %
FR-Tier2-federation					66 %	60 %	95 %	95 %	95 %
BE-TIER2					66 %	65 %	95 %	95 %	95 %
CA-EAST-T2					41 %	41 %	95 %	95 %	95 %
KR-KISTI-T2					0 %	0 %	95 %	95 %	95 %
DE-CCI					0 %	0 %	95 %	95 %	95 %
UA-Tier2-Federation					N/A	N/A	95 %	95 %	95 %

Tier 2 reliabilities

Big improvement

Federation average is now weighted by #CPU (where avail)

Would like to fix target at 95%

- Should be achievable

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

Only 1 Federation still not reporting (Nordic started in Dec)



Reliability reporting ...

- Only Ukraine federation not reporting
- Nordic sites started reporting in December
- US Tier 1 sites:
 - Recently discovered issues in how the calculations work
 - Problem arises because they were initially treated as “EGEE” sites, and there has been a partial move to run the OSG tests
 - Due to various misunderstandings (which information is published and where it is published) the net result is that the availabilities have been calculated recently on only a subset of service
 - Now being resolved ... But likely to be a discontinuity in the reporting
 - Illustrates that interoperability is not so simple ...

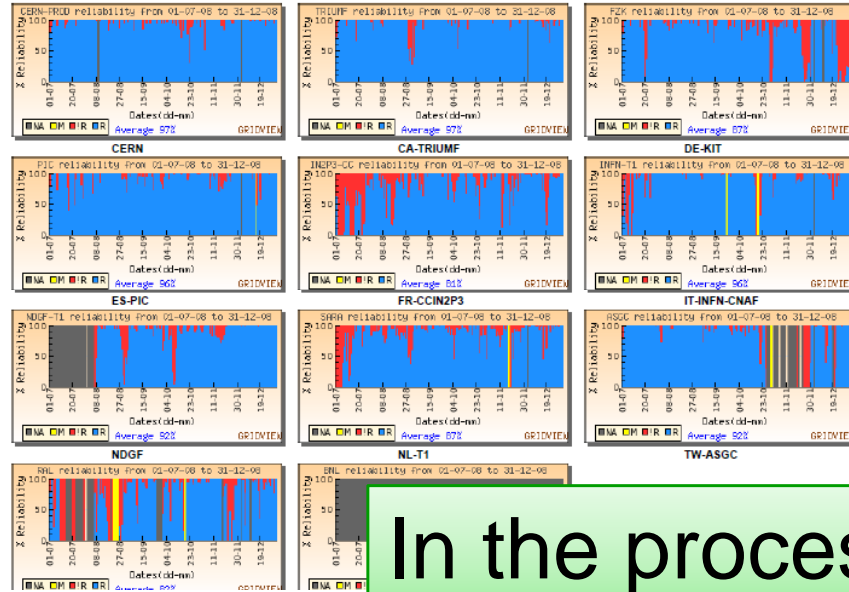
VO-specific tests



Reliability of WLCG Tier-1 Sites + CERN for ATLAS

July 2008 - December 2008

Data from SAM Monitoring. Plots show Reliability for last 6 Months
Reliability is calculated as $\text{time_site_is_available} / (\text{total_time} - \text{time_site_is_scheduled_down})$
Target reliability for each site is 95 % and Target for 8 best sites is 97 % from June, 2008



In the process of being validated

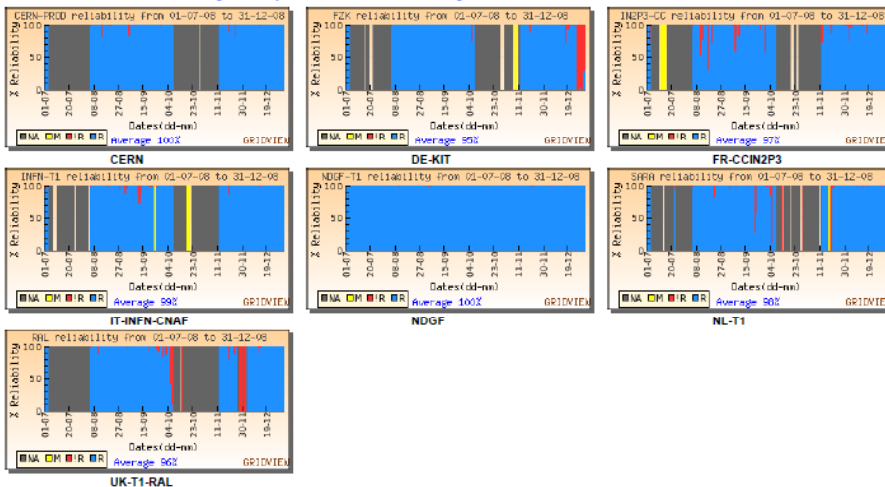
- December 2008



Reliability of WLCG Tier-1 Sites + CERN for ALICE

July 2008 - December 2008

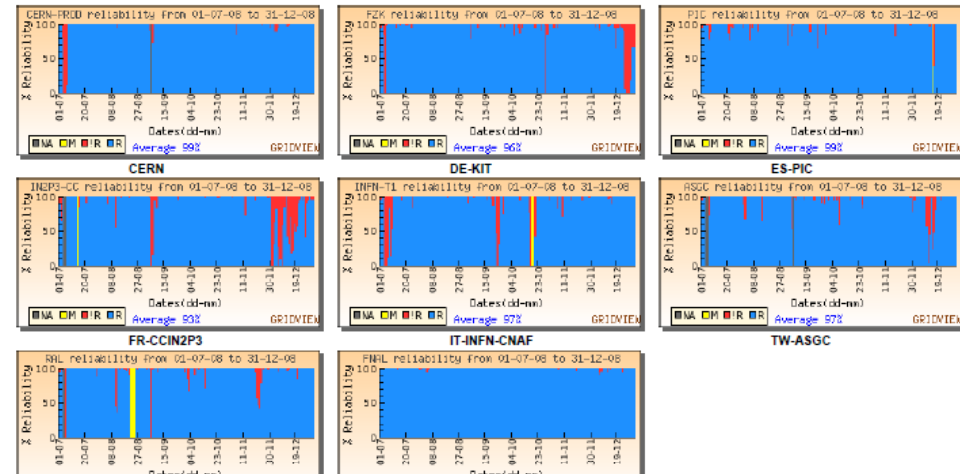
Data from SAM Monitoring. Plots show Reliability for last 6 Months
Reliability is calculated as $\text{time_site_is_available} / (\text{total_time} - \text{time_site_is_scheduled_down})$
Target reliability for each site is 95 % and Target for 8 best sites is 97 % from June, 2008



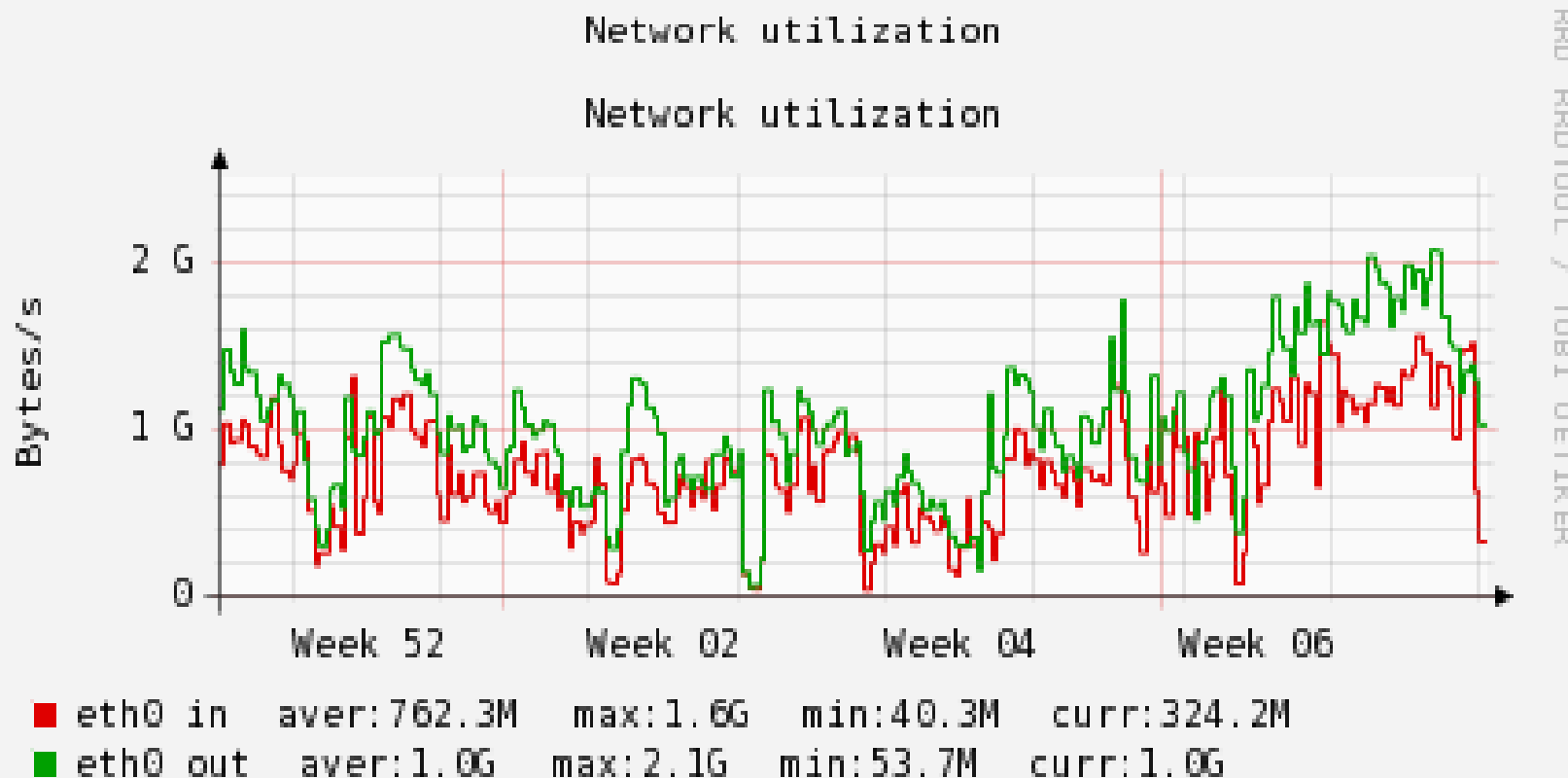
Reliability of WLCG Tier-1 Sites + CERN for CMS

July 2008 - December 2008

Data from SAM Monitoring. Plots show Reliability for last 6 Months
Reliability is calculated as $\text{time_site_is_available} / (\text{total_time} - \text{time_site_is_scheduled_down})$
Target reliability for each site is 95 % and Target for 8 best sites is 97 % from June, 2008



- Status
 - Generally quiet/good...



- fixes to problems identified by repack (main reason for deployment delays)

- Schedule: end-Feb release, in production on c2cernt3 end-March, deployment for experiment instances in April.

- Metrics have been implemented and deployed on preproduction cluster
 - Data collected in lemon
 - RRD graphs not yet implemented
- Production deployment delayed for reasons

- New metrics

Much progress, but little visible; considering how best to group metrics for display

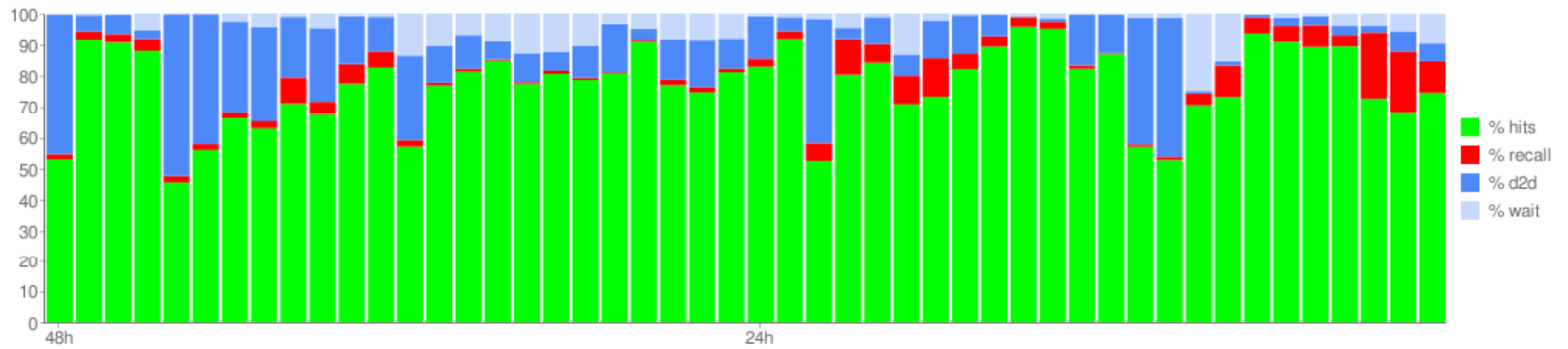
• e.g. group cache hits and garbage collection activity?

However...

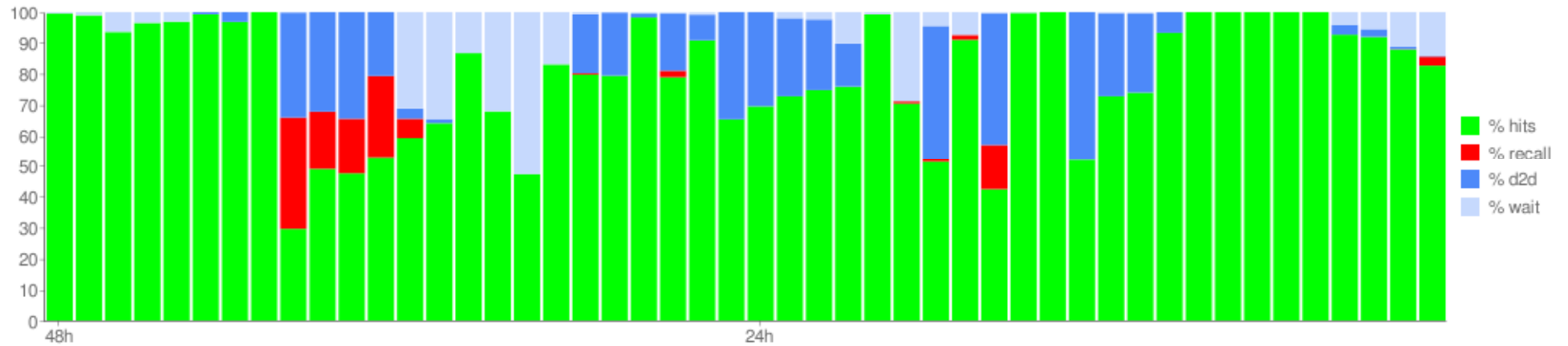
An unexpected technical dependency on the late SRM 2.7 version

Ongoing work to back-port the implementation

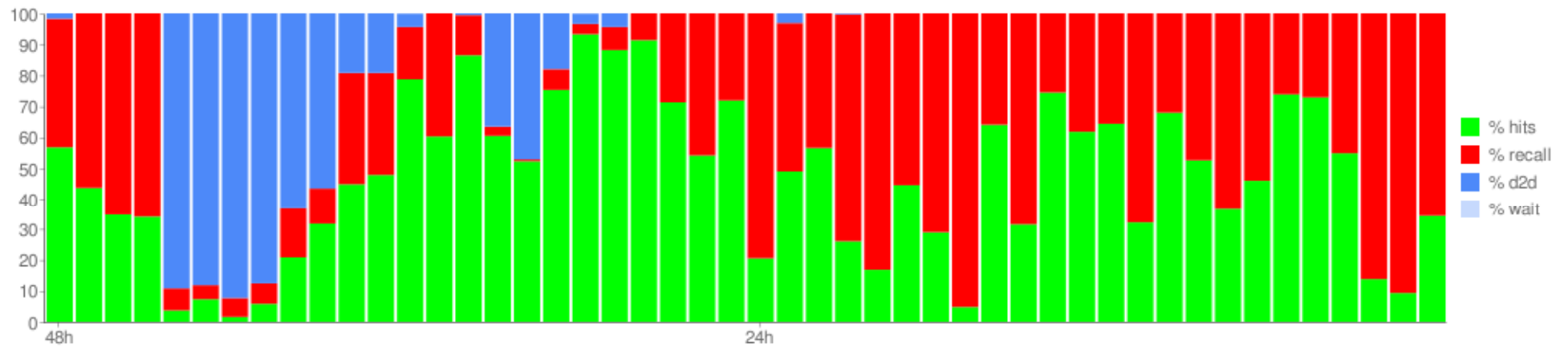
castorcms disk cache efficiency



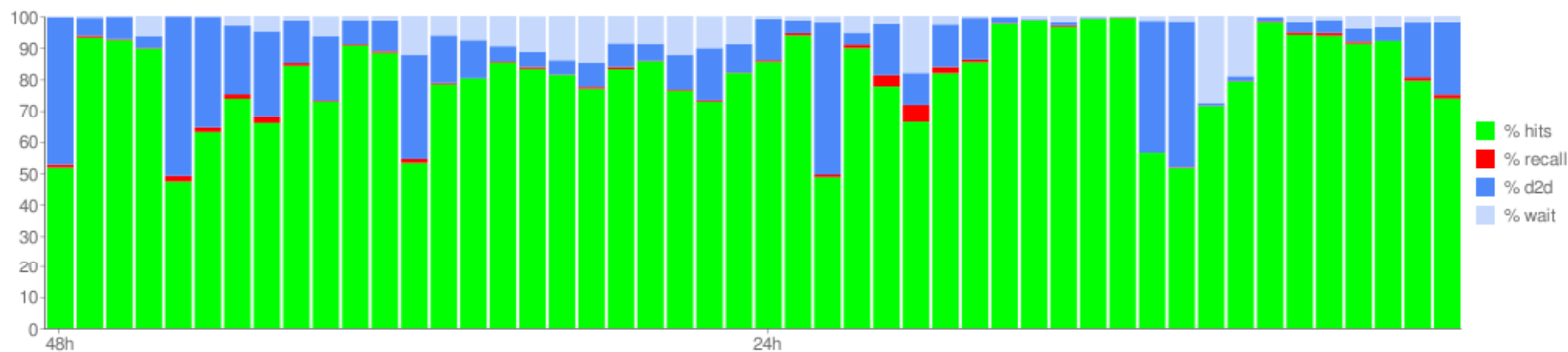
castorcms t0export disk cache efficiency



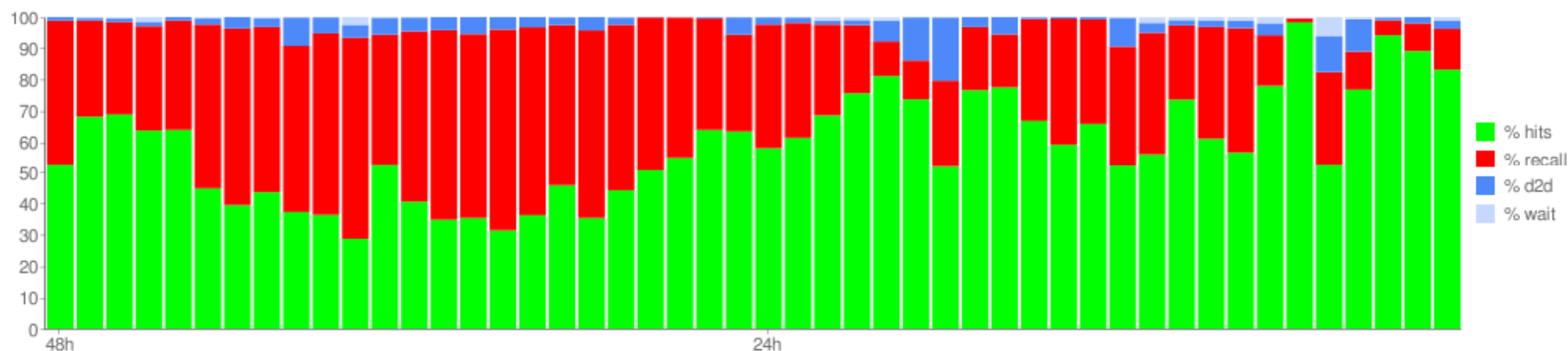
castorcms t1transfer disk cache efficiency



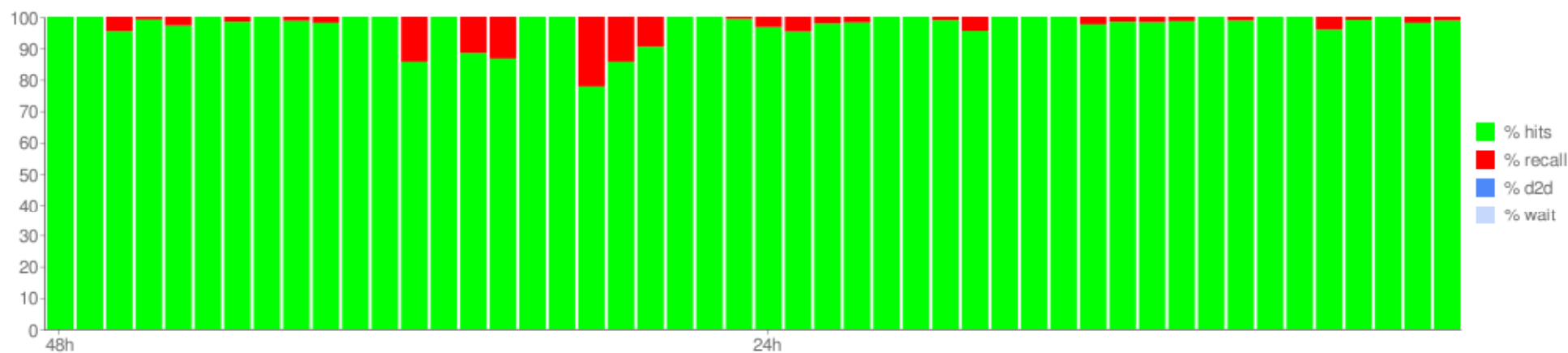
castorcms default disk cache efficiency



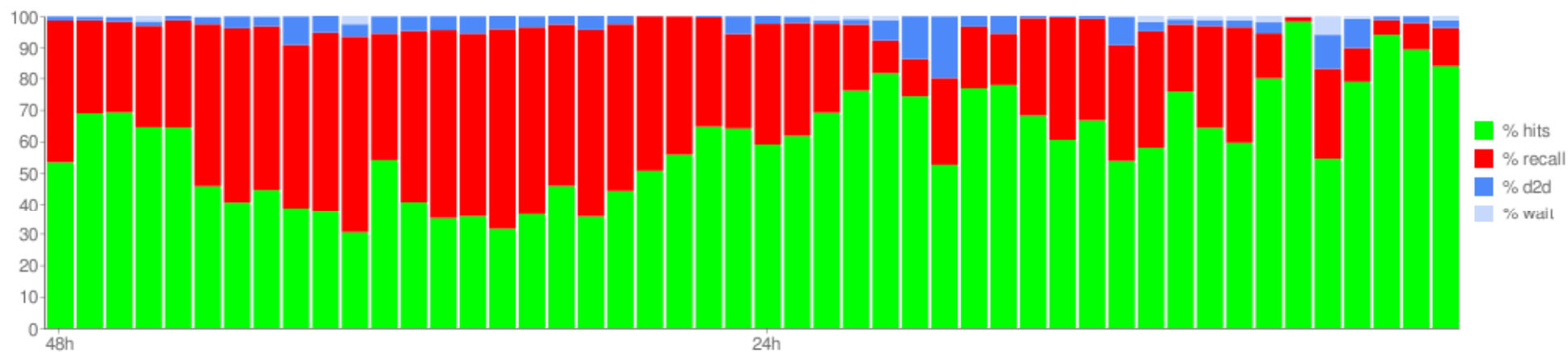
castoratlas default disk cache efficiency



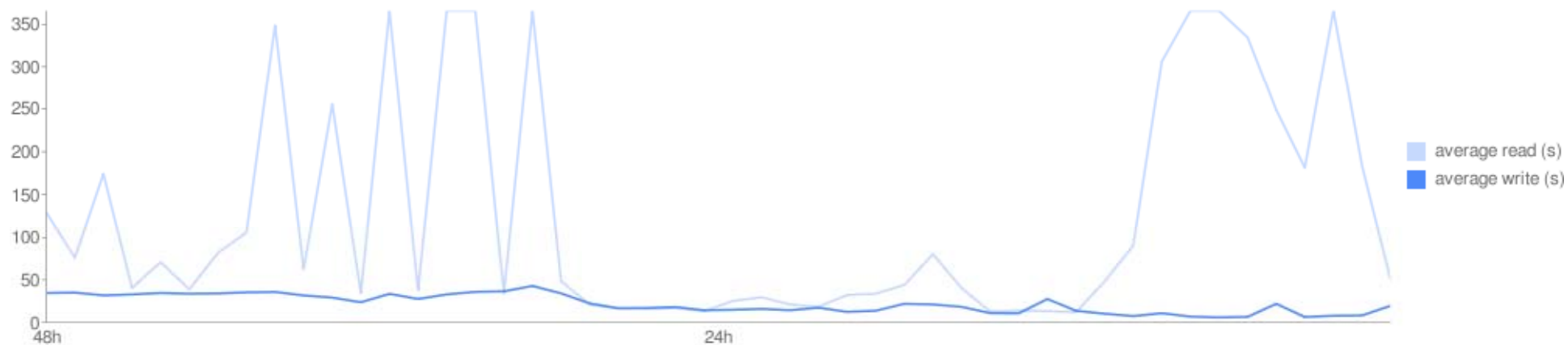
castoratlas t0atlas disk cache efficiency



castoratlask disk cache efficiency



castoratlask disk cache open latency





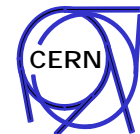
Pledge Balance in 2009

The table below shows the status at 27/10/08 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
- Following a re-organisation of some of the German Federations, pledges for 3 are still to be included
- Pledges for the new French Tier2 IPHC Strasbourg not included

% indicates the balance between offered and required.

	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	-44%	0%	-8%	-40%	-12%
T2 Disk	-44%	-20%	35%	-	-2%





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%

- Last RRB agreed that planning timescale would change from 5 years to 3 to be more realistic



Schedule for 2009 - 2010



Schedule with running in winter months

- Gains 20 weeks of LHC physics (independent of “slip”)

Year	2009												2010														
Month	F	M	A	M	J	J	A	S	O	N	D	J	F	M	A	M	J	J	A	S	O	N	D	J	F	M	
Baseline	SH	SH	SH	SH	SH	SH	SH	SH	SU	PH	PH	SH	SH	SH	SH	SH	SH	SU	PH	PH	PH	PH	SH	SH	SH	SH	
	24 weeks physics possible																										
Base ¹	SH	SH	SH	SH	SH	SH	SH	SH	SU	PH	PH	PH	PH	PH	PH	PH	PH	PH	PH	PH	SH	SH	SH	SH	SH	SH	
	44 weeks physics possible																										
Gain 20 weeks of physics in 2010 by running during winter months																											
Delay (4W)	SH	SH	SH	SH	SH	SH	SH	SH	SU	PH	PH	PH	PH	PH	PH	PH	PH	PH	PH	PH	PH	SH	SH	SH	SH	SH	
Delay (8W)	SH	SH	SH	SH	SH	SH	SH	SH	SU	PH	PH	PH	PH	PH	PH	PH	PH	PH	PH	PH	PH	SH	SH	SH	SH	SH	

From Chamonix summary: <http://indico.cern.ch/conferenceDisplay.py?confId=45433>



Likely scenario

- Injection: end September 2009
- Collisions: end October 2009
- Long run from ~November 2009 for ~44 weeks
 - This is equivalent to the full 2009 + 2010 running as planned with 2010 being a nominal year
 - Short stop (2 weeks) over Christmas/New Year
- Energy will be limited to 5 TeV
- Heavy Ion run at the end of 2010
 - No detailed planning yet
- 6 month shutdown between 2010/2011 (?) – restart in May ?



Implications for resources

- This extended run is equivalent to the original plans for 2009 + 2010
 - 2009
 - Start is delayed until October (we always planned to be ready for machine switch-on)
 - Thus should have full 2009 resources commissioned by September
 - 2010
 - Assume starts in May
 - Need to have full 2010 resources commissioned by April (as planned)
 - NB have always said will stage installation of disk during 2010: April + August (?)
- ⇒ This is close to the original plan, but with an initial delay in 2009
- Allows newer equipment (in some cases!)



Issues?

- No allowance for any change in experiment requirements
 - No change in budgets, but delay *in some cases* allows for more resources for same cost
 - How to handle the ATLAS request for additional Tier 0 resources?
 - How do experiment models deal with no shutdown?
 - Tier 1 issues with installation schedules for 2010?
 - Installation while supporting data taking
- Experiments now re-assessing their requirements
 - LHCC – clarified running time/efficiency
 - 20% efficiency for 1 year → $\sim 6 \times 10^6$ sec (c.f. 10^7 sec per year in original planning for 2009 and 2010)
 - However, must ensure that there are sufficient resources to rapidly exploit the data



Upgrade plans

- Since several software upgrades were postponed in anticipation of LHC start-up, we proposed 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;
- + 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 was not regarded as useful
- But, we will perform specific tests/validations:
 - Service validation if software is changed/upgraded
 - Specific tests (e.g. throughput) to ensure that no problems have been introduced
 - Tests of functions not yet tested (e.g. Reprocessing/data recall at Tier 1s)
 - Analysis scenario testing
- Details of the test programme were discussed in the workshop last November – will be refined now timescale is clearer



Resources ...

- Plans for 2009, 2010 based on existing requirements, validated last November by the C-RSG
 - Except that the increased Tier0/CAF request of ATLAS is not budgeted for
 - Need guidance on how to manage this
 - → now looks as though requirements decrease will cover this ... For the moment
 - AND – ALICE now intend to take much more p-p data
- New benchmark agreed
 - kSI2K → HEP-SPEC06 (based on SPEC06 c++ - mix of FP and Int tests)
 - Shown to scale well for LHC experiments
 - Simple conversion factor
 - Sites will benchmark existing capacity; vendors must run this benchmark suite (simple to run)
 - Process underway to convert requirements/pledges, and accounting
- Automated gathering of installed capacity
 - Process agreed between all parties – will be put in place to allow better understanding of available capacity; changes in information system will also improve normalisation between sites



Milestones

15-Dec-08		WLCG High Level Milestones – 2008/09													
			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 NIKHE	TRIUM F	BNL	FNAL	
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					Aug 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						
			CMSS					n/a		n/a	n/a	n/a			
			LHCb	n/a				n/a				n/a	n/a	n/a	

- Timescales being discussed for:
 - Pilot job deployment; glxexec/SCAS + framework reviews
 - Publication of VO-specific SAM testing
 - Accounting: improved T2 reports; installed capacity; user-level accounting
 - New CPU benchmarks – deployment
 - Middleware + SRM improvements
 - Metrics & monitoring: Tier 1 MSS metrics; storage system monitors; site monitoring/alarms; performance metrics



Middleware - issue

- Multi-user pilot jobs have been an open issue for a long term
- Nervousness of sites led to a set of requirements:
 - Review of experiment frameworks and how they are secured
 - Development of glexec – to perform identity switching from pilot
 - Development of SCAS – service to allow consistent auth service at a site
- Framework review
 - Complete for LHCb, CMS (until wider deployment);
 - ATLAS – few features requested and being worked on
 - ALICE – plans and questionnaire completed; discussion on deployment
- SCAS:
 - Development has been very delayed (inexperienced developer) – this is now the real hold-up to deploying user-switching pilot jobs securely

- Reminder: the selected strategy is to do a single tender for an overall solution
- Four phase process developed:
 1. Request (many) conceptual designs
 2. Commission 3-4 companies submitting conceptual designs to develop an outline design
 3. In-house, turn a selected outline design into plans and documents enabling
 4. Single tender for overall construction.

- Deadline: 28th November
 - Contacts with all 4 companies during design phase
 - All 4 companies say deadline will be met
- Meetings to review proposed designs scheduled in week of December 8th.
- Market Survey in preparation as first stage in selection of company for detailed design & construction.
- Discussions in Oslo on 28th November to further investigate possible remote server installation in 2011 (and beyond)
 - RAL also have power available in 2011, but not as much and for a shorter period.

- Four designs reviewed
 - No clear winner, but consensus on leading design.
- New Management supports project. Good, but...
 - New requirements --- “Green” & Prévessin heat recovery option
 - New organisation brings new players to brief
- “Single Contract for construction” agreed
- Agreement to work with one company to deliver fully acceptable design with modifications for new requirements.
 - Will lead to ~6 month delay.
- Will need to revisit option to install equipment at University of Oslo.



Planning for 2010 (end of EGEE)

- A final draft of the EGI blueprint has been produced (January)
 - Process will be discussed in the Overview Board next week
 - Document how the countries (Tier 1 + Tier 2) will provide the services and support needed for WLCG
 - Either as part of their NGI
 - Specific contribution
 - Must be no break in service at the end of EGEE-III
- EGEE-III have transition planning
- 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
 - The location of the EGI.org is being studied now – bids have been received
 - Decision 1st week of March at EGEE User Forum ?
 - Still not clear how many NGIs will really exist in 2010 to support this



Pending issues for 2009

- Plan to have visits of Tier 1 sites – to understand service issues
 - MSS
 - Databases – seems to be often a source of problems
 - Share and spread knowledge of running reliable services
- SRM performance
 - Need good testing of Tier 1 tape recall/reprocessing, together with Tier 1 tape writing – for several experiments together
 - Encapsulated tests?
- Data access with many users for analysis – need to see experiment testing of analysis
- Transition plan for 2010 – to cover services today provided by EGEE
 - May be short or long term – but is probably going to be needed



Conclusions from mini-review

- LHCC mini-review held 16th Feb – 1 day
- Conclusions + recommendations:
 - “Recommend that there is a CCRC’09 in some form:”
 - At least CMS+ATLAS – but preferably with all 4 experiments
 - Testing reprocessing at Tier 1s (recall from tape) and massive/chaotic user analysis
 - Need metrics with which to evaluate this
 - “Let’s make sure we are not limited by resources when data comes..”
 - Not obvious you can just move the schedule by one year in terms of resources”
 - “Need an official statement on 2009/2010 running time and LHC efficiency factor common for all experiments so they can provide a consistent/coherent estimation of resources needed in 2009/2010”
 - Promised urgently (~today)
 - “Experiments still suffer from SRM (MSS) performance”
 - Applications area – “Very good progress on all fronts with very mature organisation well managed giving results”



New issue: US and ion programme

- Reminder: there is no formal US commitment to WLCG – no MoU signature – for the involvement in ALICE and CMS-ions
- However:
 - We know for a long time discussions around Tier 1 and Tier 2 sites for ALICE, (there are real resources being used)
 - Recently heard from US-CMS-ions – “plans” to create a ion-Tier 1 at Vanderbilt –
 - We received request to provide connectivity to Vanderbilt (i.e. CERN was asked for money from Geant to add this capacity)
 - BUT:
 - All transatlantic networking for LHC should go via USLHCNET – there should not be an additional path (the additional needed bandwidth does not justify an additional circuit)
- Request:
 - The US formal participation in the ion programme should be resolved
 - Different technical solutions are unnecessary – and unacceptable



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 2009/10 continues:
 - Delay allows for more effective purchases in some cases
 - Resource procurements/installations were significantly delayed in 2008
 - Must be ready for the accelerator start-up, even if resources are today not saturated
- Planning for future – Tier 0/CAF and European Grid infrastructure – ongoing