Project status report

WLCG LHCC
Referees' meeting
21st September 2009

Ian Bird
LCG Project Leader

Agenda

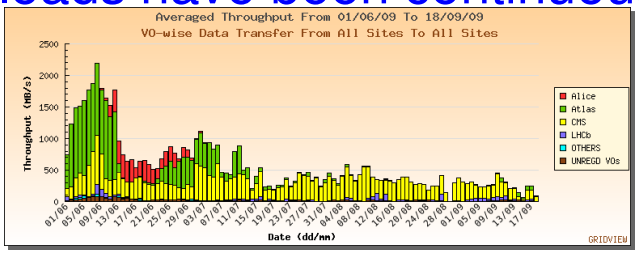
- Overall status
- Planning and milestones
- Follow up actions from STEP'09
- Status of Tier 0
- Status of Tier 1 sites (seen via milestones etc)
- Resource requests and installation status
  - Comments on CRSG report & process

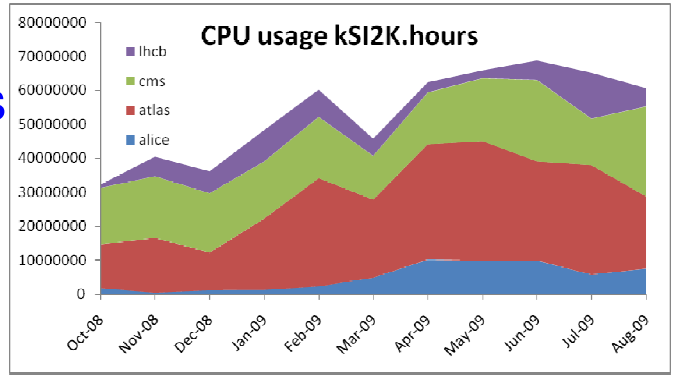
- Status of preparations for EGEE → EGI transition

- Experiment reports
  - NB LHCb external meeting, thus no LHCb representative today

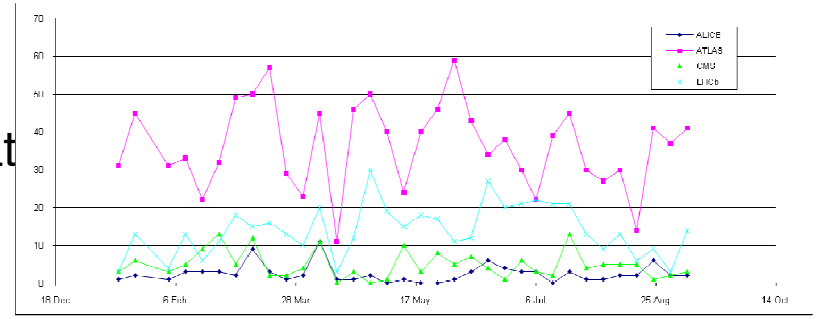
Overall summary

- Since July has been fairly quiet
- Experiments did follow-up tests with some sites where problems had been seen during STEP'09 exercise
- In general workloads have been continuous



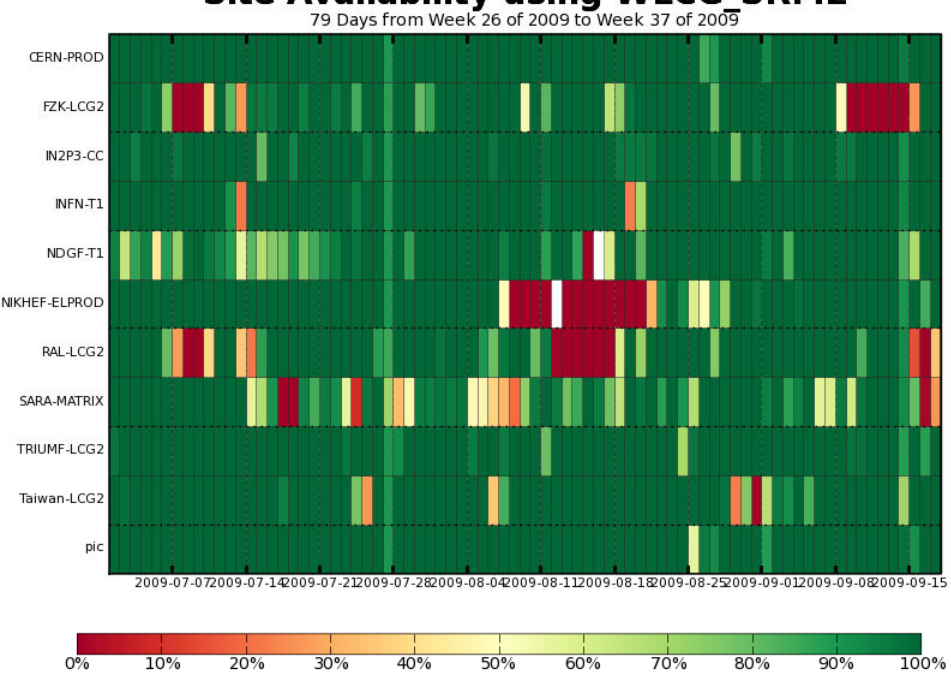


- WLCG service has been running according to the defined procedures
  - Reporting and follow up of problems at same level

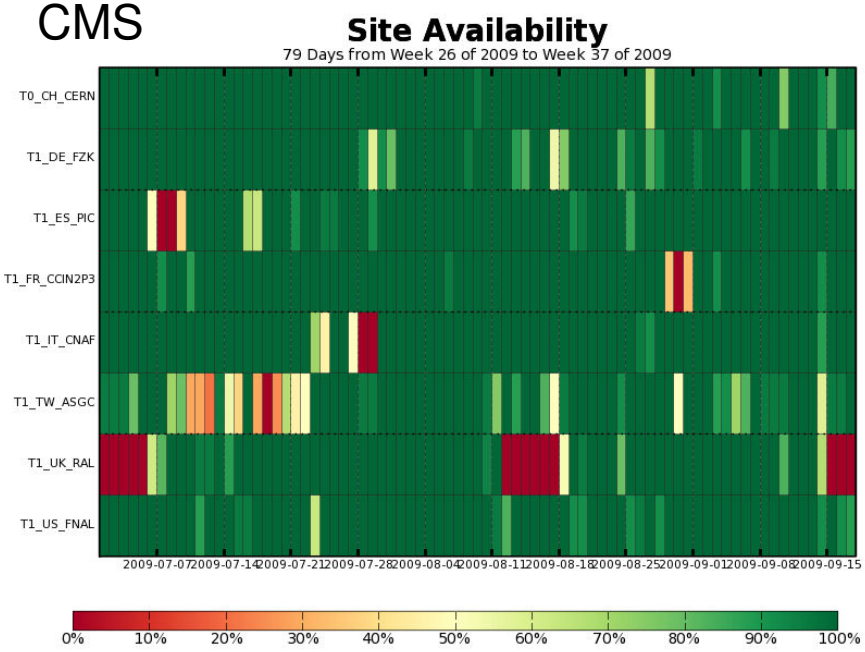


- Focus on stability and resolution of problems in preparation for data taking

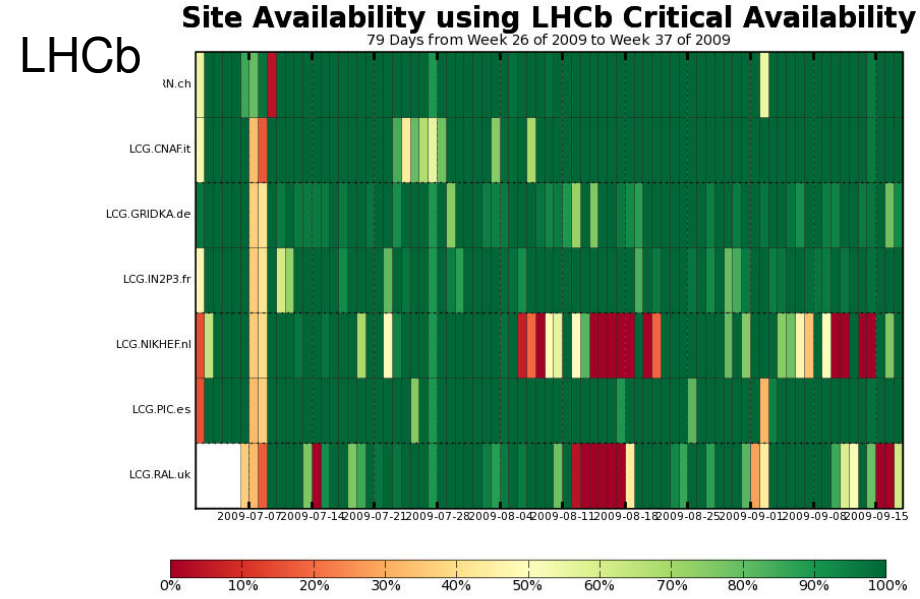
ATLAS



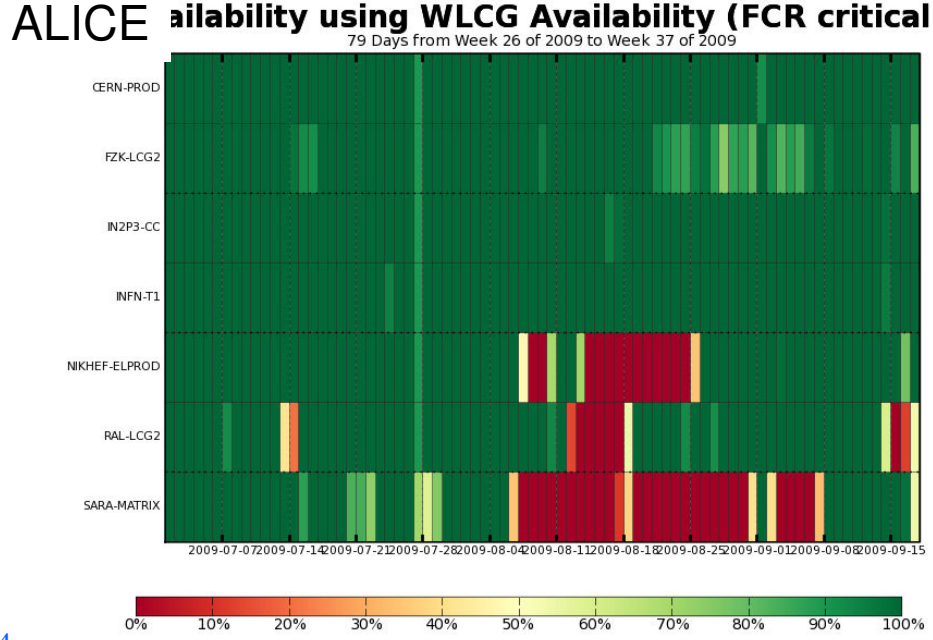
CMS



LHCb



ALICE



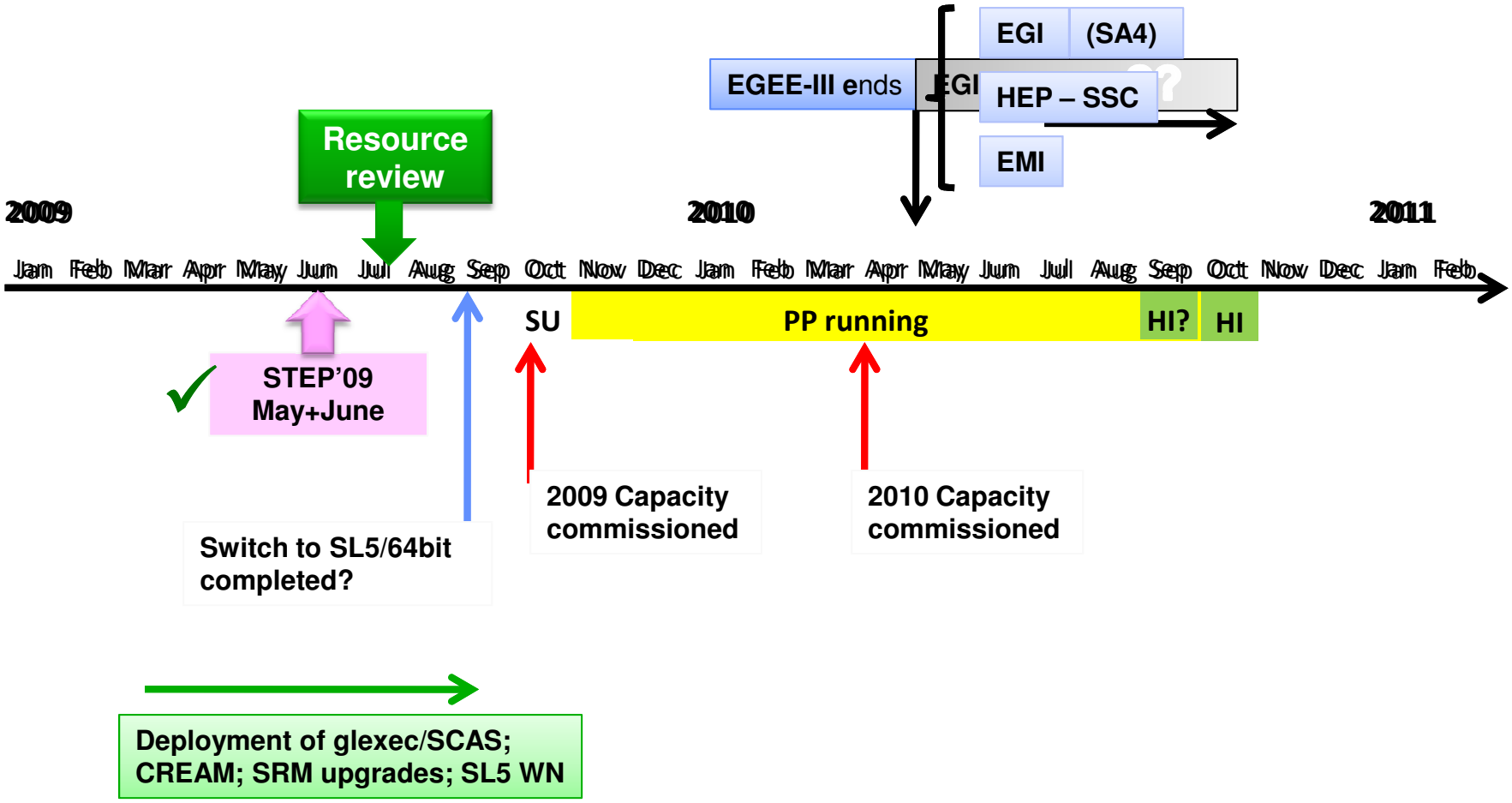
Incident reports ... fewer

WLCG Service Incident Reports

- N.B. Downtimes / degradations are always "user visible" (which is what counts...)

Site	Date	Duration	Service	Impact	Report
CERN	5 & 8 Sept 2009	2 * 2 hours	CASTOR LHCb	two Castor Database problems	https://twiki.cern.ch/twiki/bin/view/FIOgroup/PostMortem20090905
CERN	26 Aug 2009	18:40 - 23:30	Batch	Public and production queues closed	https://twiki.cern.ch/twiki/bin/view/FIOgroup/PostMortem20090826
ASGC	17 Jul 2009	6:00 - 10:00	Power cut	Most services went down and restarted	https://twiki.cern.ch/twiki/bin/viewfile/LCG/WLCGServiceIncidents?rev=1;filename=power_cut_ASGC.txt
ATLAS	13 Jul 2009	10:00 - 11:00	Central Catalogs	Degrade of performance	PostMortem13Jul09
NL-T1	STEP09				https://twiki.cern.ch/twiki/pub/Atlas/Step09Feedback/Post_Mortem_STEP09_NL-T1-0.4.pdf

WLCG timeline 2009-2010



01-Sep-09		
ID	**Date**	**Milestone**
Pilot Jobs Frameworks		
WLCG-08-14	**May 2008**	**Pilot Jobs Frameworks studie by the Review working group** Working group proposal comple the Experiments.
Tier-2 and VOs Sites Reliability R		
WLCG-08-09	**Jun 2008**	**Weighted Average Reliability Federation above 95% for 80%** Weighted according to the sites
WLCG-08-11	**Apr 2009**	**VO-Specific Tier-1 Sites Re** Considering each Tier-0 and T (and by VO?)

Tier2 reliabilities: publishe
40% > 95% relia

Tier 1: VO-specific reliabi

Federation	Reli-ability	Avail-ability
T2_US_Caltech	100 %	100 %
US-NET2	100 %	100 %
FR-GRIF	100 %	100 %
US-MWT2	100 %	100 %
SI-SiGNET	100 %	100 %
CZ-Prague-T2	100 %	90 %
PT-LIP-LCG-Tier2	99 %	73 %
T2_US_Wisconsin	99 %	99 %
T2_US_UCSD	99 %	99 %
CA-WEST-T2	99 %	89 %
FR-IN2P3-CC-T2	99 %	99 %
FR-IN2P3-LPC	99 %	99 %
T2_US_Florida	99 %	99 %
AU-ATLAS	99 %	99 %
CN-IHEP	99 %	91 %
FR-IN2P3-SUBATECH	99 %	95 %
DE-DESY-ATLAS-T2	99 %	98 %
JP-Tokyo-ATLAS-T2	98 %	98 %
T2_US_Nebraska	98 %	89 %
PL-TIER2-WLCG	98 %	98 %
KR-KISTI-T2	98 %	98 %
UK-SouthGrid	98 %	95 %
CH-CHIPP-CSCS	98 %	90 %
FR-IN2P3-LAPP	98 %	98 %
ES-LHCb-T2	98 %	98 %
BR-SP-SPRACE	98 %	98 %
HU-HGCC-T2	98 %	91 %
ES-CMS-T2	98 %	97 %
UK-NorthGrid	98 %	97 %
CA-EAST-T2	97 %	97 %
ES-ATLAS-T2	97 %	91 %
PK-CMS-T2	97 %	97 %
T2_US_MIT	96 %	93 %

Federation	Reli-ability	Avail-ability
TW-FTT-T2	96 %	96 %
IN-DAE-KOLKATA-TIER2	96 %	96 %
KR-KNU-T2	96 %	91 %
DE-DESY-RWTH-CMS-T2	95 %	95 %
UK-ScotGrid	95 %	93 %
DE-MCAT	94 %	94 %
US-AGLT2	94 %	93 %
US-SWT2	94 %	94 %
AT-HEPHY-VIENNA-UIBK	93 %	83 %
FR-IN2P3-IPHC	93 %	80 %
BE-TIER2	92 %	92 %
TR-Tier2-federation	92 %	92 %
IT-ALICE-federation	92 %	90 %
IT-ATLAS-federation	92 %	90 %
IT-CMS-federation	92 %	90 %
IT-LHCb-federation	92 %	90 %
US-WT2	92 %	82 %
RU-RDIG	92 %	88 %
EE-NICPB	89 %	80 %
DE-FREIBURGWUPPERTAL	89 %	81 %
RO-LCG	89 %	86 %
T2_US_Purdue	87 %	85 %
IL-HEPTier-2	84 %	84 %
DE-GSI	84 %	84 %
FI-HIP-T2	83 %	67 %
IN-INDIACMS-TIFR	81 %	79 %
UK-London-Tier2	79 %	74 %
DE-DESY-GOE-ATLAS-T2	70 %	68 %
NO-NORGRID-T2	0 %	0 %
SE-SNIC-T2	0 %	0 %
UA-Tier2-Federation	N/A	N/A

Milestones - 2

01-Sep-09		WLCG High Level Milestones												
			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
SL5 Milestones														
WLCG-09-21	Mar 2009	**SL5 gcc 4.3 (WN 4.1 binaries)Tested by the Experiments** Experiments should test whether the MW on SL5 support their grid applications	ALICE			ATLAS			CMS			LHCb		
WLCG-09-22	Jul 2009	**SL5 Deployed by the Sites (64 bits nodes)** Assuming the tests by the Experiments were successful. Otherwise a real gcc 4.3 porting of the WN software is needed.												
Tier-1 Sites Procurement - 2009														

SLC5: MB issued statement urging sites to migrate asap; clarifying no show stoppers for the experiments

Ideally would like significant resources under SL5 available before data taking.

Many large sites have migrated

Milestones - 3

01-Sep-09		WLCG High Level Milestones												
			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**
SCAS/glExec Milestones														
WLCG-09-17	**Jan 2009**	**SCAS Solutions Available for Deployment** Certification successful and SCAS packaged for deployment	**Done in March 2009**											
WLCG-09-18	**Apr 2009**	**SCAS Verified by the Experiments** Experiment verify that the SCAS implementation is working (available at CNAF and NL-T1)	ALICE n/a			ATLAS			CMS n/a ?			LHCb		
WLCG-09-19	**09-18 + 1 Month**	**SCAS + glExec Deployed and Configured at the Tier-1 Sites** SCAS and glExec ready for the Experiments.												
WLCG-09-20	**09-18 + 3 Month**	**SCAS + glExec Deployed and Configured at the Tier-2 Sites** SCAS and glExec ready for the Experiments.												

Glexec/SCAS available for deployment; deployed at several test sites; general deployment is very slow

Milestones - 4

Accounting Milestones						
WLCG-09-02	**Apr 2009**	**Wall-Clock Time Included in the Tier-2 Accounting Reports** The APEL Report should include CPU and wall-clock accounting	APEL			
WLCG-09-03	**Jul 2009**	**Tier-2 Sites Report Installed Capacity in the Info System** Both CPU and Disk Capacity is reported in the agreed GLUE 1.3 format.	~100 % of T2 Sites Reporting			
WLCG-09-04a	**Jul 2009**	**Sites publishing the User Level Accounting information**				
WLCG-09-04b	**Jul 2009**	**User Level Accounting verified and approved by the Experiments**	ALICE	ATLAS	CMS	LHCb

Almost all sites report CPU information

Comparison of published data and expected values tbd

Follow up on published storage information to be done

User-level accounting available in APEL; policy now approved; sites asked to enable the publishing of the data

Milestones - 5

01-Sep-09		WLCG High Level Milestones												
			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
STEP 2009 - Tier-1 Validation														
WLCG-09-23	Jun 2009	Tier-1 Validation 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

All Tier 1's tested by STEP'09

Were 3 with concern: ASGC, NL-T1, FZK

FZK and NL-T1 have since demonstrated satisfactory functionality

ASGC still of concern: meeting with ASGC management in early Oct

Milestones - 6

01-Sep-09		WLCG High Level Milestones												
			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**
CREAM CE Rollout														
WLCG-09-25	**Apr 2009**	**Release of CREAM CE for deployment**												
WLCG-09-26	**May 2009**	**All European T1 + TRIUMF and CERN at least 1 CE. 5 T2s supporting 1 CE**	n/a					n/a				n/a	n/a	n/a
WLCG-09-27	**Jul 2009**	**2 T2s for each experiment provide 1 CREAM-CE each.**	ALICE			ATLAS			CMS			LHCb		
WLCG-09-28	**Sep 2009**	**50 sites in addition to the ones above**												

Milestones - 7

01-Sep-09		WLCG High Level Milestones												
			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**
Metrics and Monitoring Milestones														
WLCG-09-08	**Nov 2008**	**Experiments Dataflows clear for the Tier-1 Sites** Experiments should present the data flows they expect to reach at the Sites (a la LHCb)	ALICE			ATLAS			CMS			LHCb Nov 2007		
WLCG-09-09	**Removed**	**Tier-1 Sites Define Their MSS Metrics** Tier-1 Sites specify which metrics are going to be collected to demonstrate the dataflow supported												
WLCG-09-10	**June 2009**	**Tier-1 Sites Publish Their MSS Metrics in SLS** Tier-1 Sites make their current MSS metrics available via SLS												
WLCG-09-11	**TDB**	**Automatic Alarms (SAM, etc) at the Tier-1 Sites** Tier-1 Sites should be able to automatically send, receive and handle alarms and problem notifications												
WLCG-09-12	**TDB**	**Monitoring of the Storage Systems** The Storage systems used provide monitoring information to Sites and Experiments	CASTOR			dCache			DPM		StoRM		BestMan	
WLCG-09-13	**TDB**	**Performance Metrics?** User Response, Services Dowtimes Operations KPI												

Milestones - 8

01-Sep-09		WLCG High Level Milestones												
			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**
CPU Benchmarks/Units Milestones														
WLCG-09-14	**Dec 2008**	**CPU New Unit Working Group Completed** Agreement on Benchmarking Methods Conversion Proposal and Test Machines	CPU New Benchmarking Unit Working Group											
WLCG-09-15	**Feb 2009**	**Sites Pledges in HEPSPEC-06** Pledged from the Sites should be converted to the new unit	LCG Office											
WLCG-09-16	**Apr 2009**	**New Experiments Requirement in HEPSPEC-06** Experiments should convert their requirements to the new unit (or by LCG Office)	ALICE			ATLAS			CMS			LHCb		
WLCG-09-24	**May 2009**	**Sites Benchmark their Capacity in the HEPSPEC-06** Resources from the Sites should be converted to the new unit												

Site benchmarking in progress

STEP09 follow-up: Tier 0 + Tier 1s

- All:
  - MSS metrics
  - Need instant real-time monitoring of throughput (and per day overview) and to view transfers per experiment (WAN in/out; LAN – WNs)
    - Tools for site and grid
- NL-T1:
  - Communication & SIRs
  - Lack of tape drives (now installed)
  - DMF tuning needed?
  - Unexplained low performance
  - LAN b/w to WN too small
  - Shared SW area problems
  - Repeat tests
- ASGC:
  - Castor version?
  - Job scheduling T2 v T1; ATLAS v CMS
  - Low efficiency for CMS reprocessing jobs
  - Repeat tests
- FZK:
  - Improve communication to world
  - SAN issues
  - Shared SW area problems
  - SRM server overload
  - dcap access problems
  - Too many lcg-cp → overload gridftp
  - Repeat tests
- NDGF:
  - No MSS performance monitoring
  - Low throughput
  - Analysis jobs overloaded network
  - No Panda pre-staging
  - What is action to improve?
- CNAF:
  - Shared SW area problems
- Site visits – planned for FZK + NL-T1

This was the summary immediately after STEP'09

Tier 2s

- Shared SW areas affect CPU efficiencies in some sites (also T1)
- ATLAS efficiencies different between WMS and PANDA
  - Likely due to different data access methods
- Data transfer timeouts (what? → see analysis summary)
- Intra-VO fairshares → how to manage? (➔ GDB?)
- VO communications
  - Need for VO-independent channel to reach all T2 admins
- Network infrastructure not good enough for data rates

This was the summary immediately after STEP'09

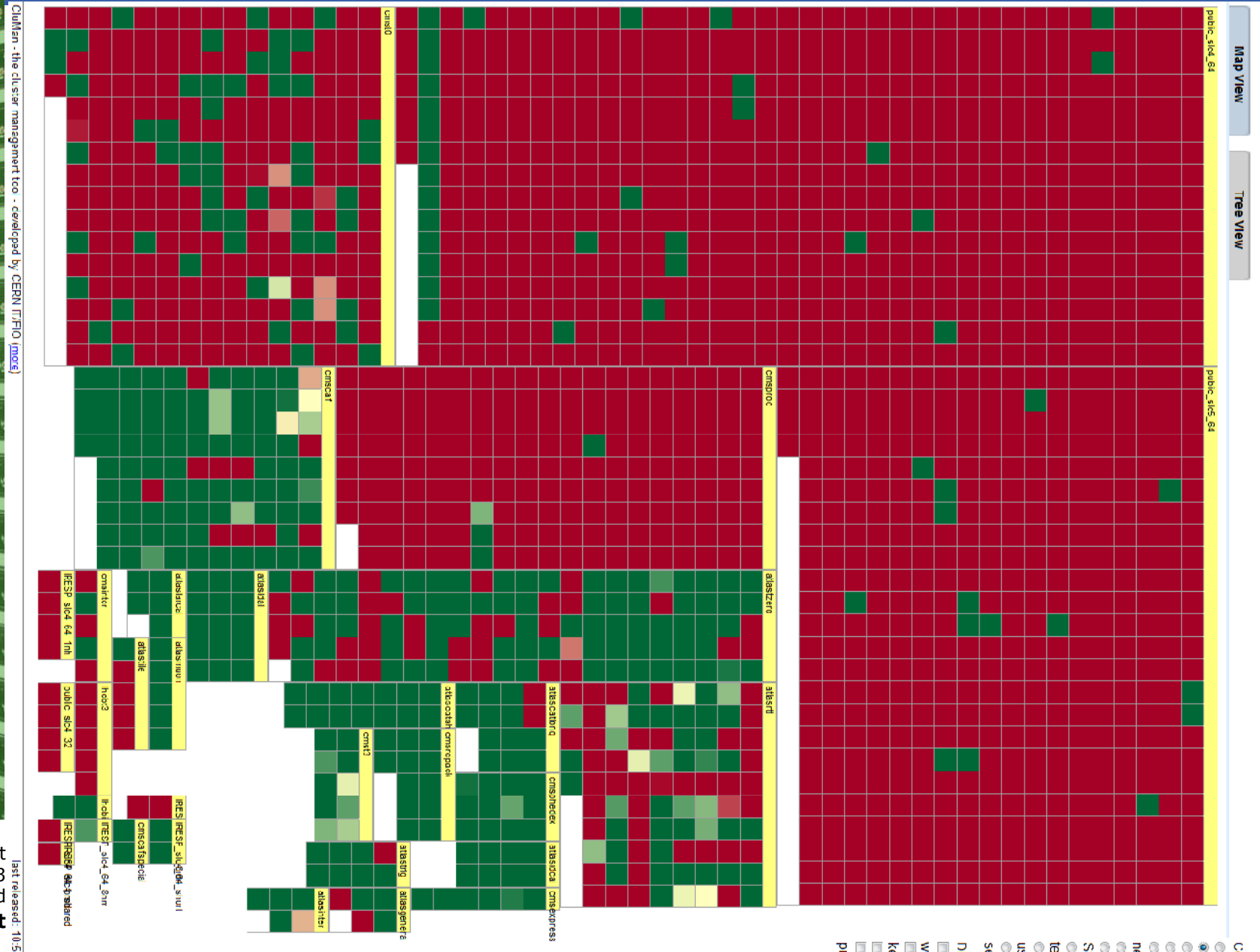
Data management

- Some detailed tuning issues; e.g.:
  - FTS: Timeouts for large files
  - FTS: Increase the number of concurrent jobs to increase bandwidth
  - LFC: Deploy bulk methods (new additions)
- dCache:
  - These were discussed in August MB:
    - Need a clear strategy for Chimera (and risk analysis)
    - Explain “golden” release v what we have (and risk analysis)
  - Site decisions, but must be based on understanding of risks
- Other issues:
  - Data transfer timeouts (dcap/rfio) at high (analysis) load
  - Better data access strategies???
    - Pre-stage vs copy locally vs Posix-io
    - General investigation of data access methods and problems
      - Work already done by ATLAS – extended to other experiments
  - Shared software areas:
    - Site issue – NFS not a good solution
  - Monitoring:
    - More real time on MSS transfers (Tier 0 and useful at Tier 1s)
    - MSS metrics
    - Improve dashboards – ATLAS vs CMS

FIO

Tier0 Status

CERN IT Department

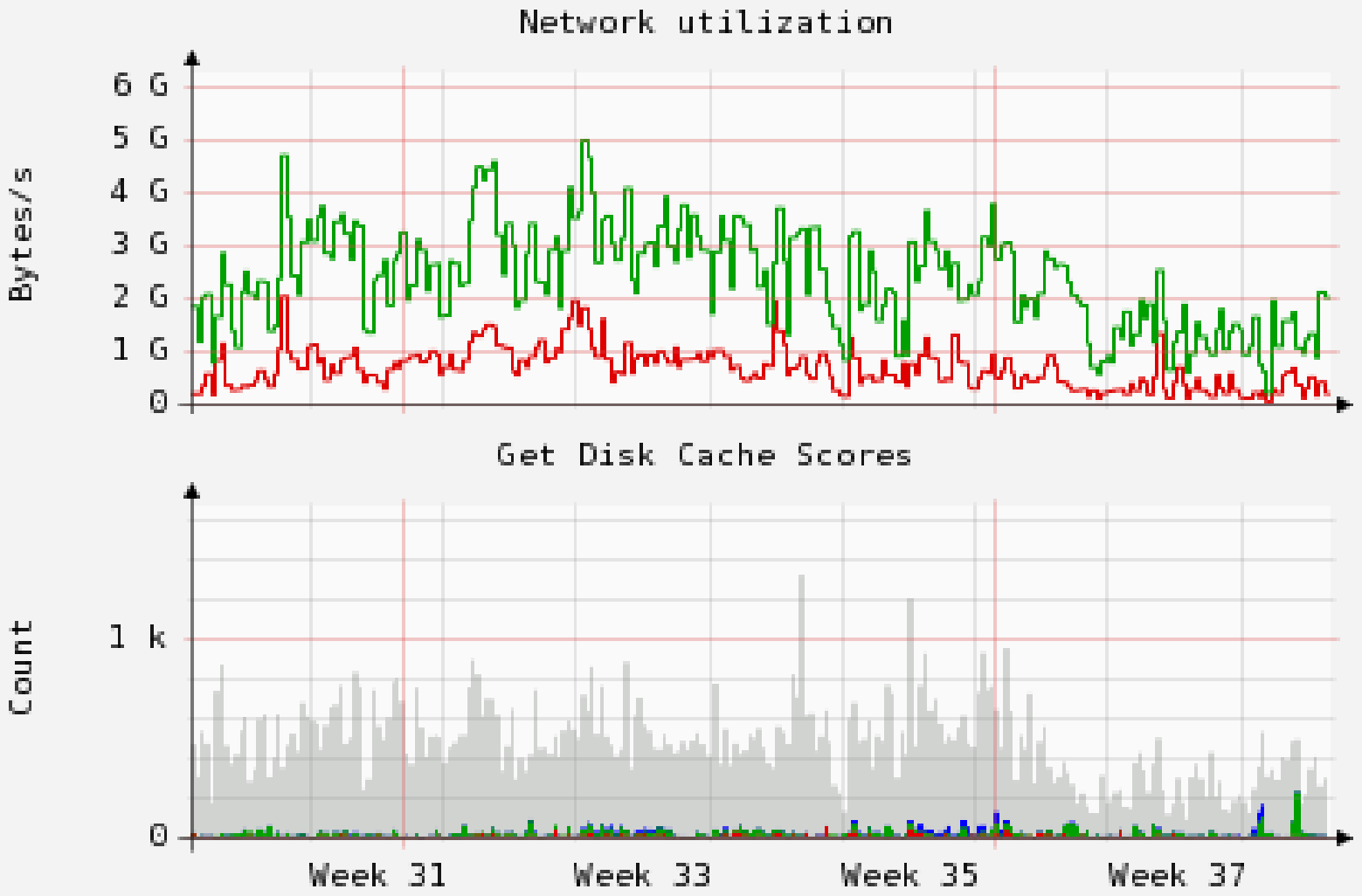


CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Tier0 Status

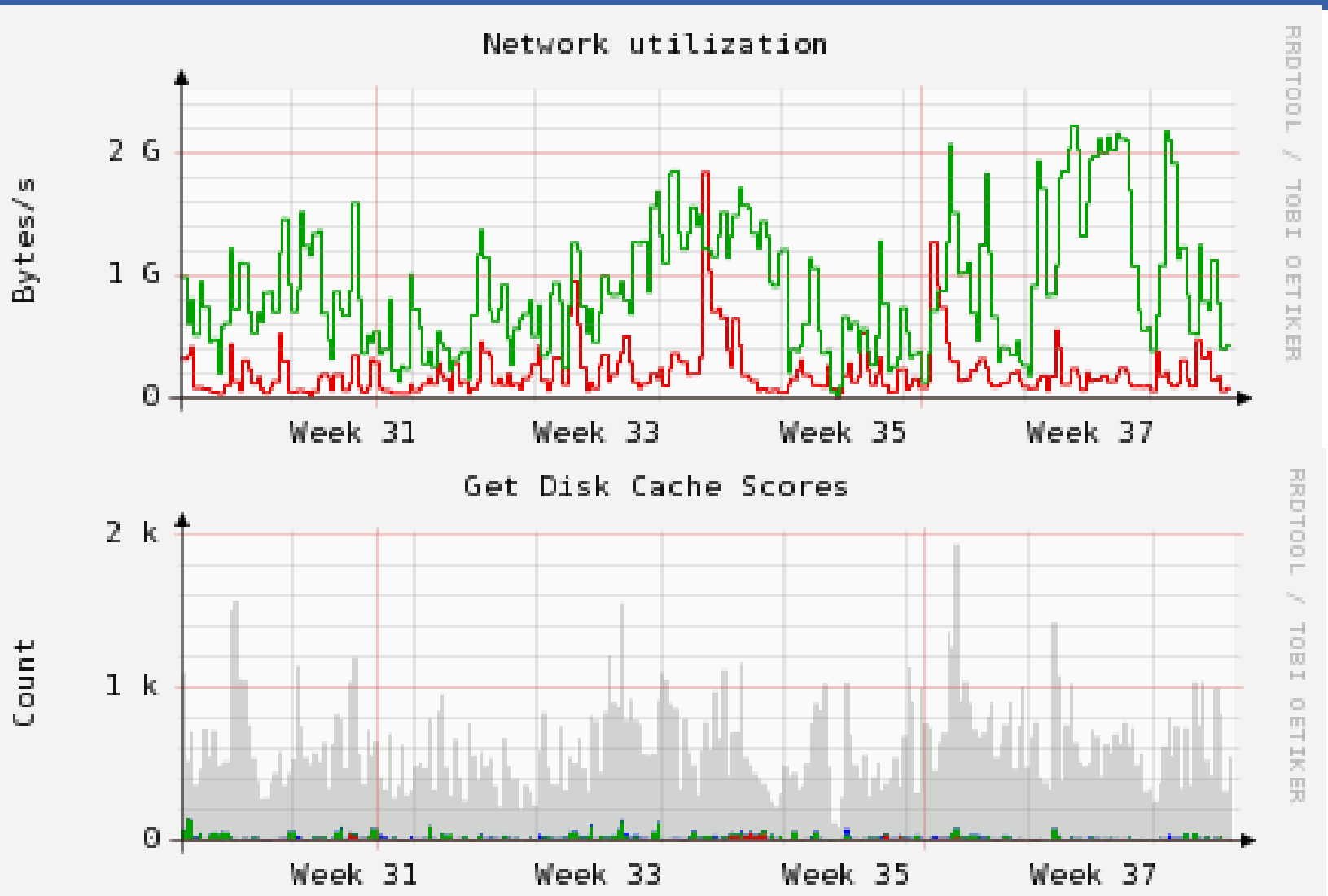
- CASTOR
  - Generally smooth running since STEP
    - Deploying 2.1.8.10 across LHC experiments; completes tomorrow with LHCb and can then have improved monitoring of migration rates
    - CASTOR 2.1.9.0 deployed on tape servers; this is the base for future tape optimisations (e.g. transfer aggregation)
  - Main issue has been service degradations due to inadvertent misuse by individual users
    - An operations priority for future developments

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it



■ The number of requests which where placed on hold waiting on other re
aver: 6.8 max: 54.1 min: 0.0 curr: 2.6
■ The number of requests which triggered a diskcopy replication
aver: 17.5 max: 224.5 min: 0.0 curr: 8.0

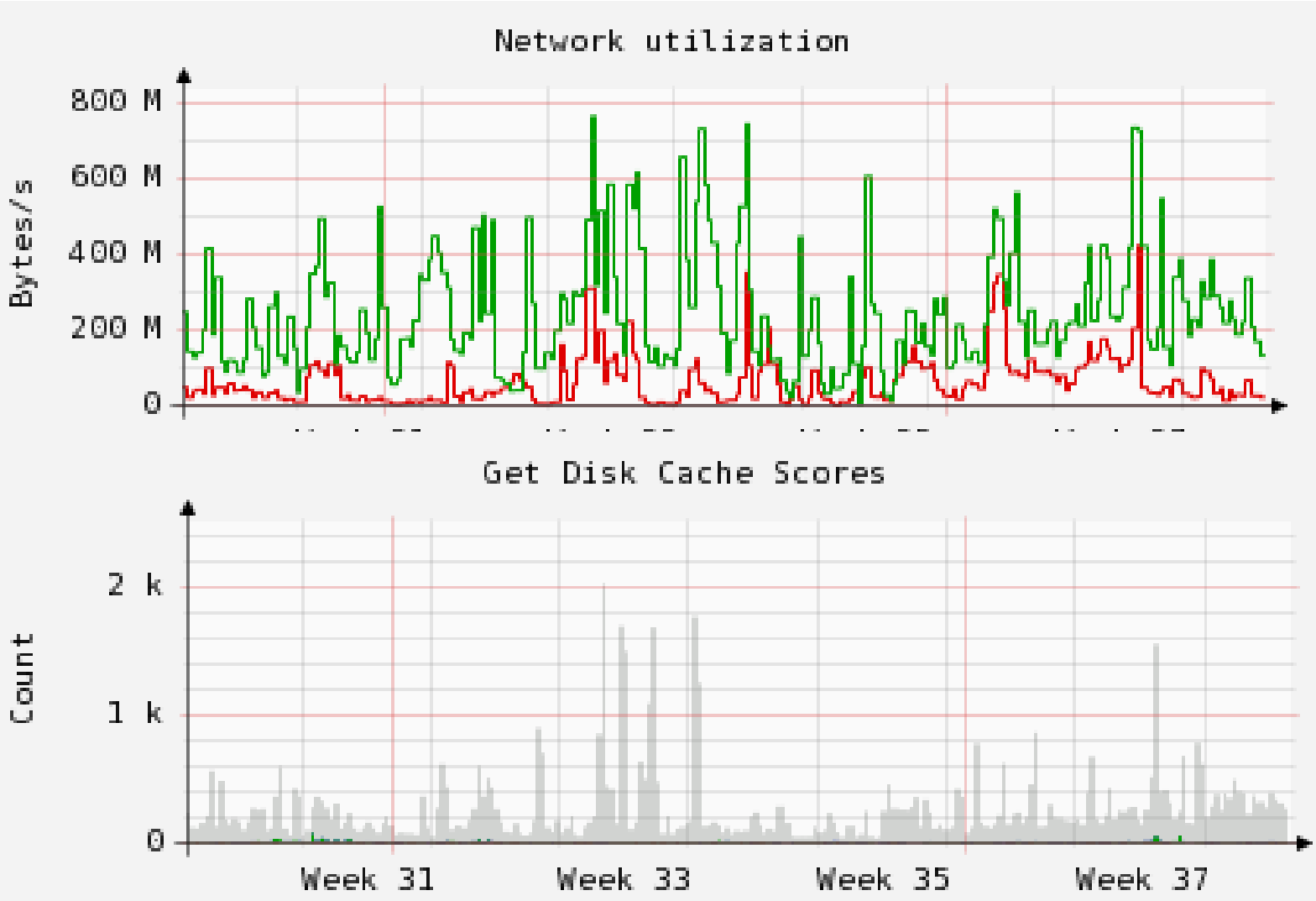
CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it



■ The number of requests which where placed on hold waiting on other re
aver:3.0 max:42.8 min:0.0 curr:180.6m
■ The number of requests which triggered a diskcopy replication

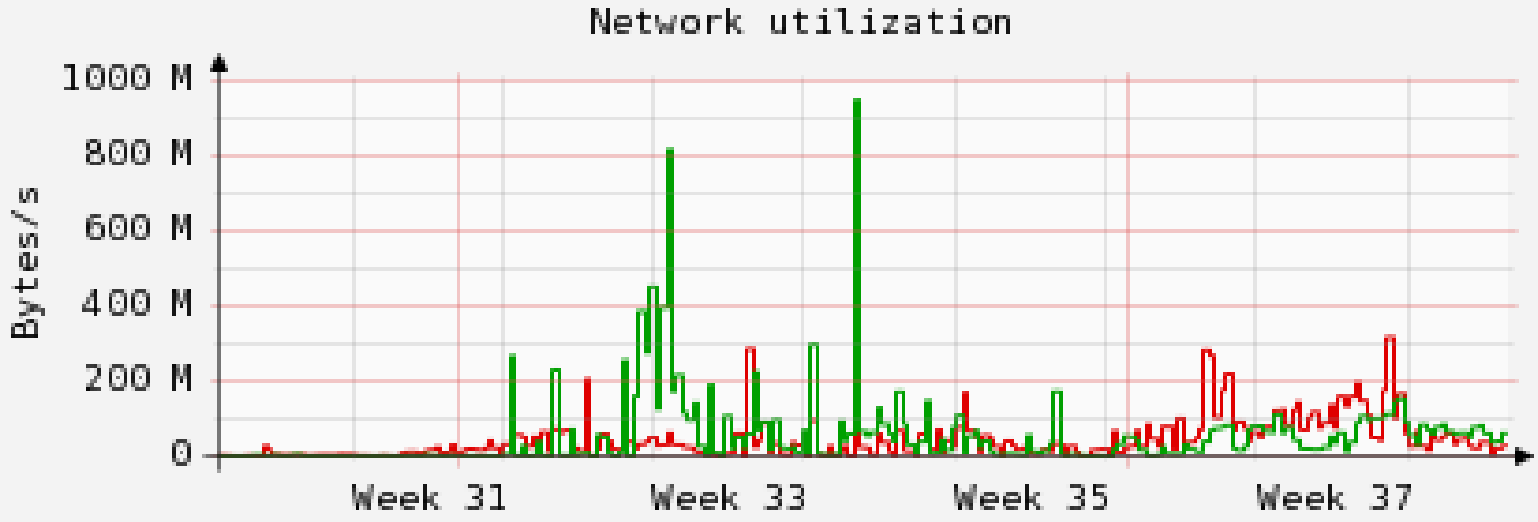
CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

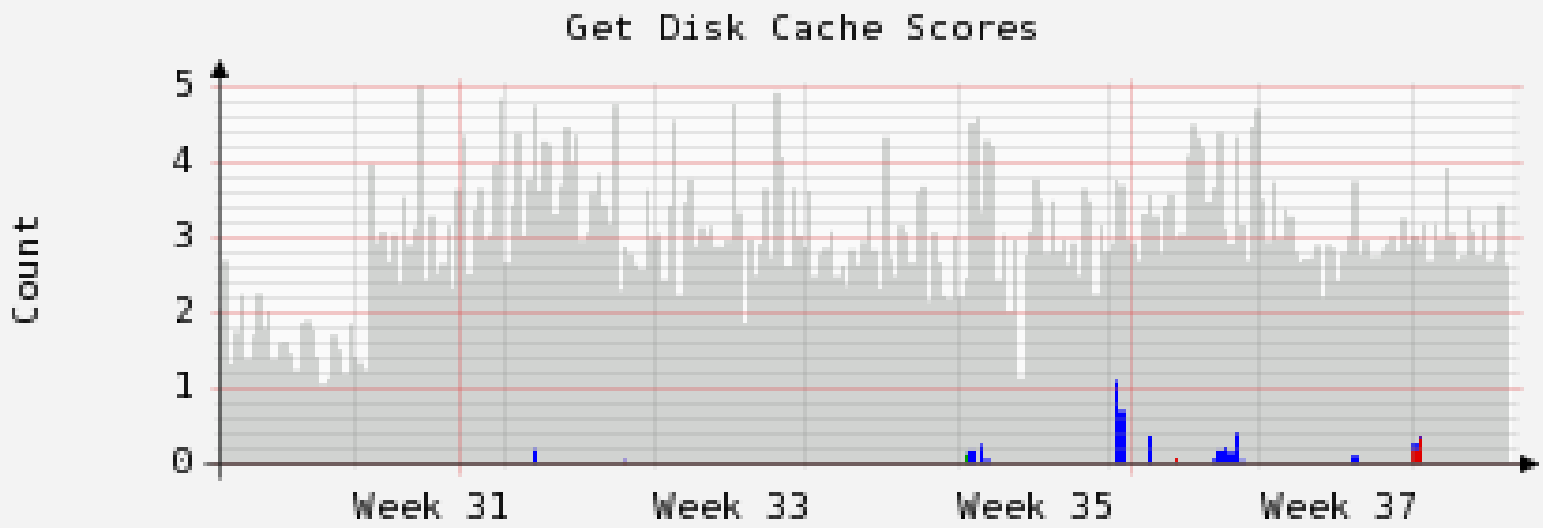
LHCb



CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

ALICE





■ The number of requests which where placed on hold waiting on other re
aver:2.4m max:331.6m min:0.0 curr:0.0
■ The number of requests which triggered a diskcopy replication

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Tier0 Status

- Computer Centre Infrastructure
  - Decision on construction of new Computer Centre in Prévessin suspended pending clearer view of
    - costs for a container based solution
    - long-term computing requirements
  - Projects now underway to
    - Reinforce “critical power” infrastructure in B513 to guarantee support for load of ~480kW.
    - Plan installation of containers in 2010 (for both “critical” and “physics” loads.

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Mass storage

- Upgrades for dCache and Castor – to get stable versions for data taking
- dCache:
  - 1.9.4 introduces ACLs to ensure file protection
  - Migration to new namespace (Chimera) from pnfs for better scalability (recommended for large sites)
    - Site decision, but had full discussion of risks/advantages
  - 1.9.5 will be "golden" version for support during 1st year of data taking
- Castor:
  - 2.1.9 (consolidation version) will be deployed at CERN
  - Encourage RAL, CNAF, ASGC to upgrade to this for better analysis support

- Resolution of open issues from SRM functionality requests
  - Remember: "addendum" of functionality that had bee requested, but put on hold in order to ensure stable versions
  - See next slides

Missing SRM features by priority

- **Extremely important**
  - Space protection
    - But at least tape protection is available everywhere
  - File pinning
    - On CASTOR is almost non-existent
    - But pinning importance is lower with file pre-staging services
- **Rather important**
  - VOMS awareness
    - Missing from CASTOR
- **Useful**
  - Target space selection
    - Missing in dCache an StoRM
- Nice to have
  - Ls returns all spaces with a copy of the file
    - Missing in CASTOR (where it makes sense)
- Scalability and stability (CMS)
  - Main issues at Tier-2's
  - It is required that the SRM front-end should guarantee that the activity of a single user could not disrupt the service

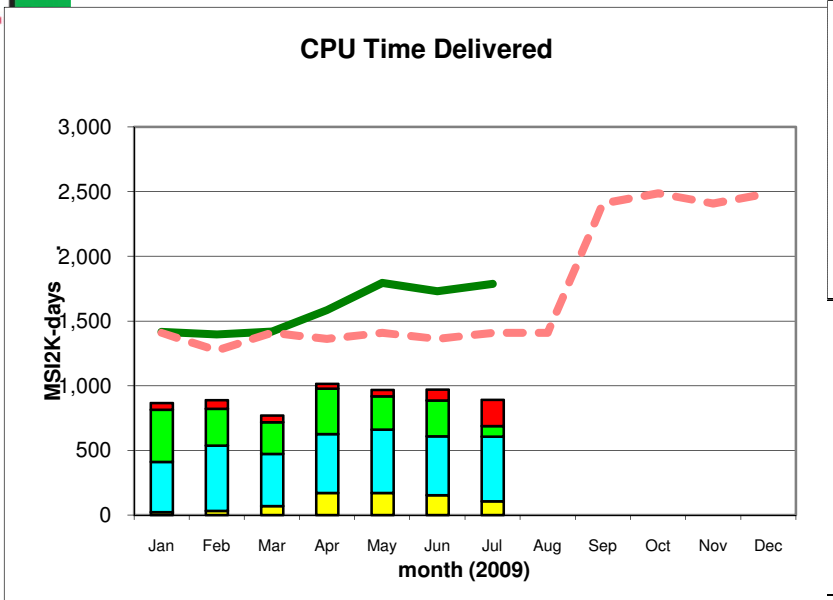
Middleware

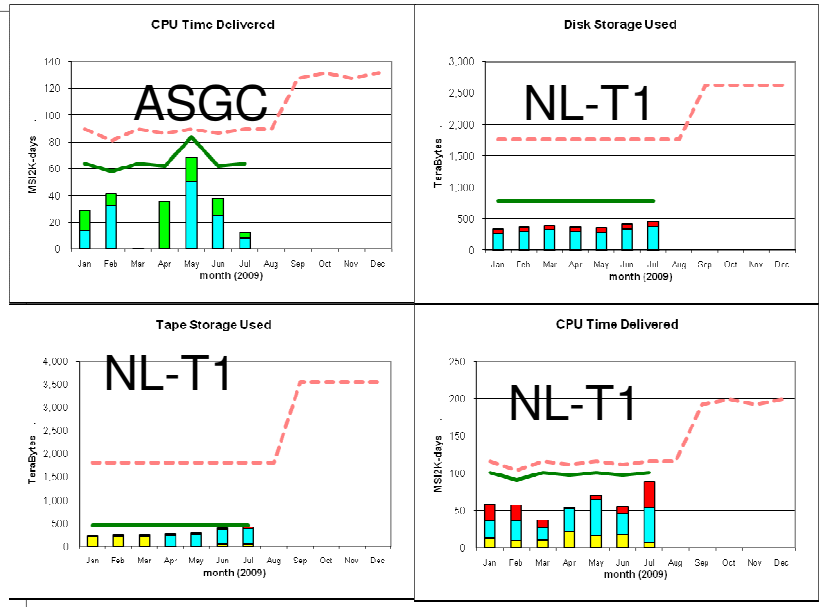
- WMS:
  - WMS 3.2 available – can submit to CREAM and to the ARC-CE
- Compute Element
  - New version of CREAM with many bug fixes – now deployable in parallel with existing CEs
  - Once the old WMS have been retired, can seriously use CREAM in production
- Glexec/SCAS (needed for multi-user pilot jobs)
  - Now ready ... But deployment take-up by sites is very slow
- Information system:
  - Latest version can also handle new schema with improved service discovery
- Generally:
  - Continuous process of patching and updating middleware as needed

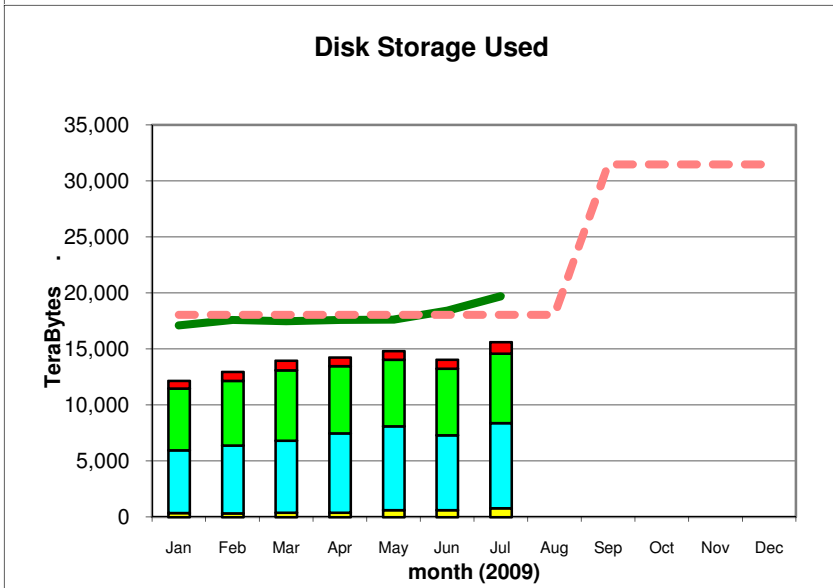
Resources

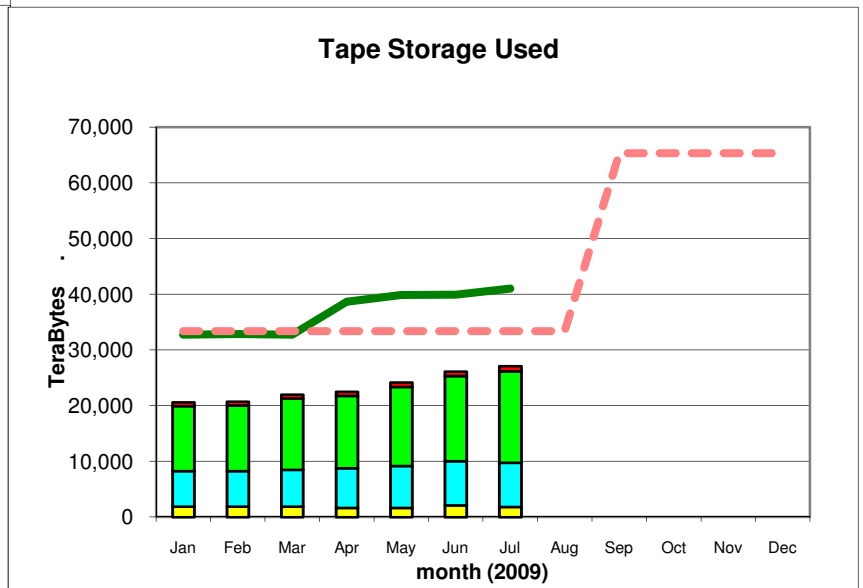
- LHCC report and CRSG report only recently available
- 2009 resources being installed (deadline October)
- 2010 resources
  - Use requests as presented in July review/listed in final CRSG report
    - Except: ALICE where RSG report has wrong numbers – updates provided by ALICE; anticipate updated CRSG report
- Message, agreed with CERN management, sent to funding agencies for RRB preparation:
  - experiment requests have been reviewed by both the LHCC and C-RSG bodies who have been working closely together. The result of their work led to agreement on the experiment requests in the attached table* to be used until data taking and the running conditions beyond 2010 are known. Their results also clearly state that computing resources should not in any way hinder the LHC physics programme.
  - Consequently at the next CRRB..., the main emphasis will be on the 2010 resource situation. In order to prepare the meeting, each site is asked to:
    - confirm that 2009 resources pledged are now available in production, or if not explain any difference
    - confirm 2010 pledges, and for sites supporting more than one experiment, indicate the split between the experiments
  - Until first data-taking and further news about the running scenario beyond 2010 is known, the CRRB Chairman has agreed that exceptionally we do not require your pledge data beyond 2010 for this meeting.

Resource installations & issues









Comments on review process/outcome

- Timeliness of report and overall timing of the process
  - Long procurement cycles are unavoidable and we must have the conclusions earlier
  - Results needed in the Spring for requests for the following year (now in Autumn RRB)
  - The report is essential well in advance of the RRB to allow discussion and feedback
- Much effort has been invested by the RSG members and others – in future it might be better to better coordinate between the LHCC and the RSG?
- Is having dedicated scrutineers per experiment most efficient? E.g. comparison of CMS and ATLAS was required.
- Actually the fundamental questions have not been answered:
  - Trigger rates
  - Amount of MC
  - # of reprocessings
- Experiment models and RSG simple models agree within 10% (level of uncertainty any way); future understanding must be based on:
  - Experience with data
  - Real answers to the above basic parameter questions
- C-RRB (and C-RSG) modelled on other RRBs; but computing is different and the process needs to reflect that:
  - Combine RSG and LHCC reviews/reviewers?
  - Change timing
  - ...