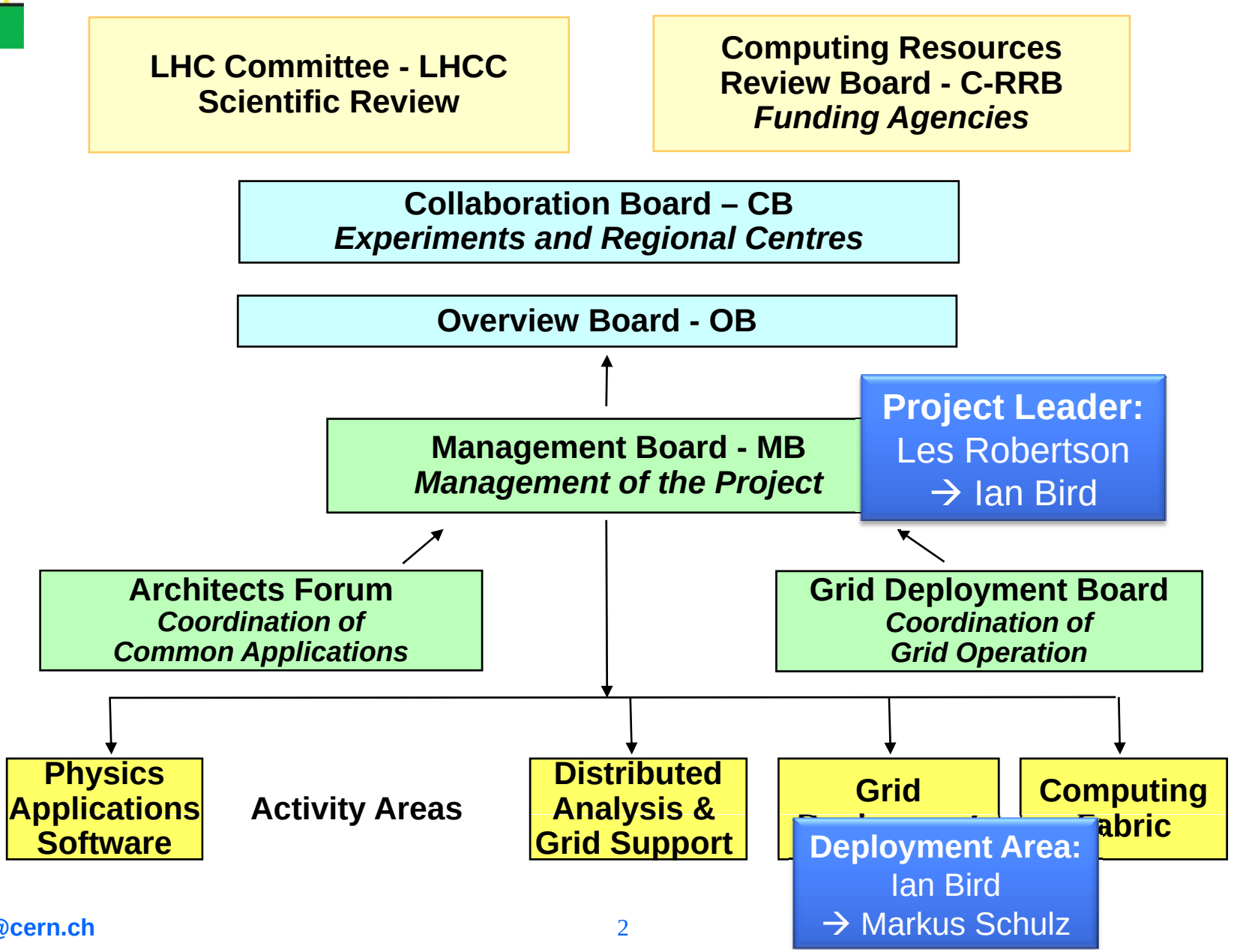
LHCC Mini-Review of WLCG

Project Status & Overview

Ian Bird
LCG Project Leader

LCG Organisation – Phase 2

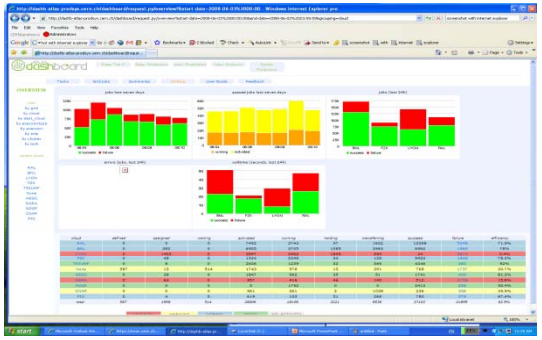


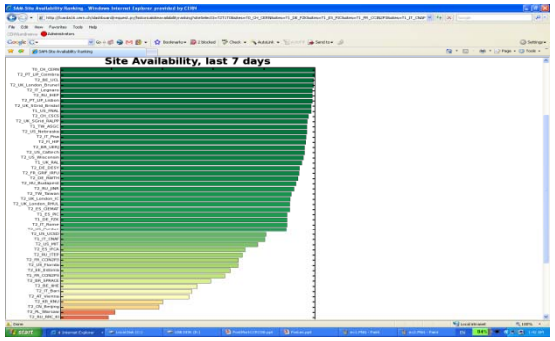
ARDA (1): Ganga

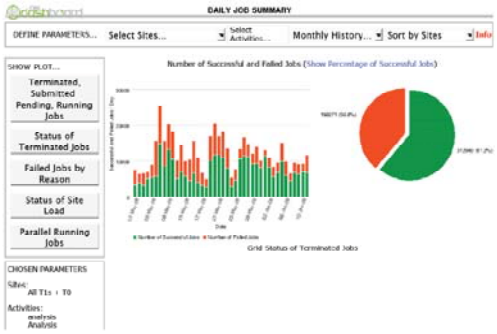
- Steady usage in the user community
  - >1400 (unique) users since Jan 2007. Regular users: 300/month (50% Atlas, 30% LHCb)
- Ganga 5 just released (beginning of June)
  - Main goal: code improved/restructured for product **maintainability**
  - User support becoming more and more important. Trying to streamline it (creation of FAQ, establishing user-support procedures, etc...)
- Actually more active users joining (e.g. access to FDR2 data)
  - In ATLAS, the GangaRobot (automatic system to help in commissioning sites for analysis) is being put in production
  - The GangaPANDA back-end (conceptually similar to the GangaDIRAC backend for LHCb) is now working and it is expected to be the main execution backend for ATLAS analysis users

ARDA(2): Dashboards

- Autonomous project, catalysing several monitoring-related activities
  - Reuse of the Dashboard toolkit
- Incomplete list of recent developments
  - Coherent access to SAM data (experiment view)
  - Monitor of the ATLAS Tier0
    - Effort from ATLAS, support and guidance from the Dashboard team
  - Monitor of the ATLAS production (PANDA based)
  - CMS site availability
  - CMS user job monitor
- Essential for ramp-up experiments activities, to commission sites etc...







Middleware: Baseline Services

The *Basic* Baseline Services – from the TDR (2005)

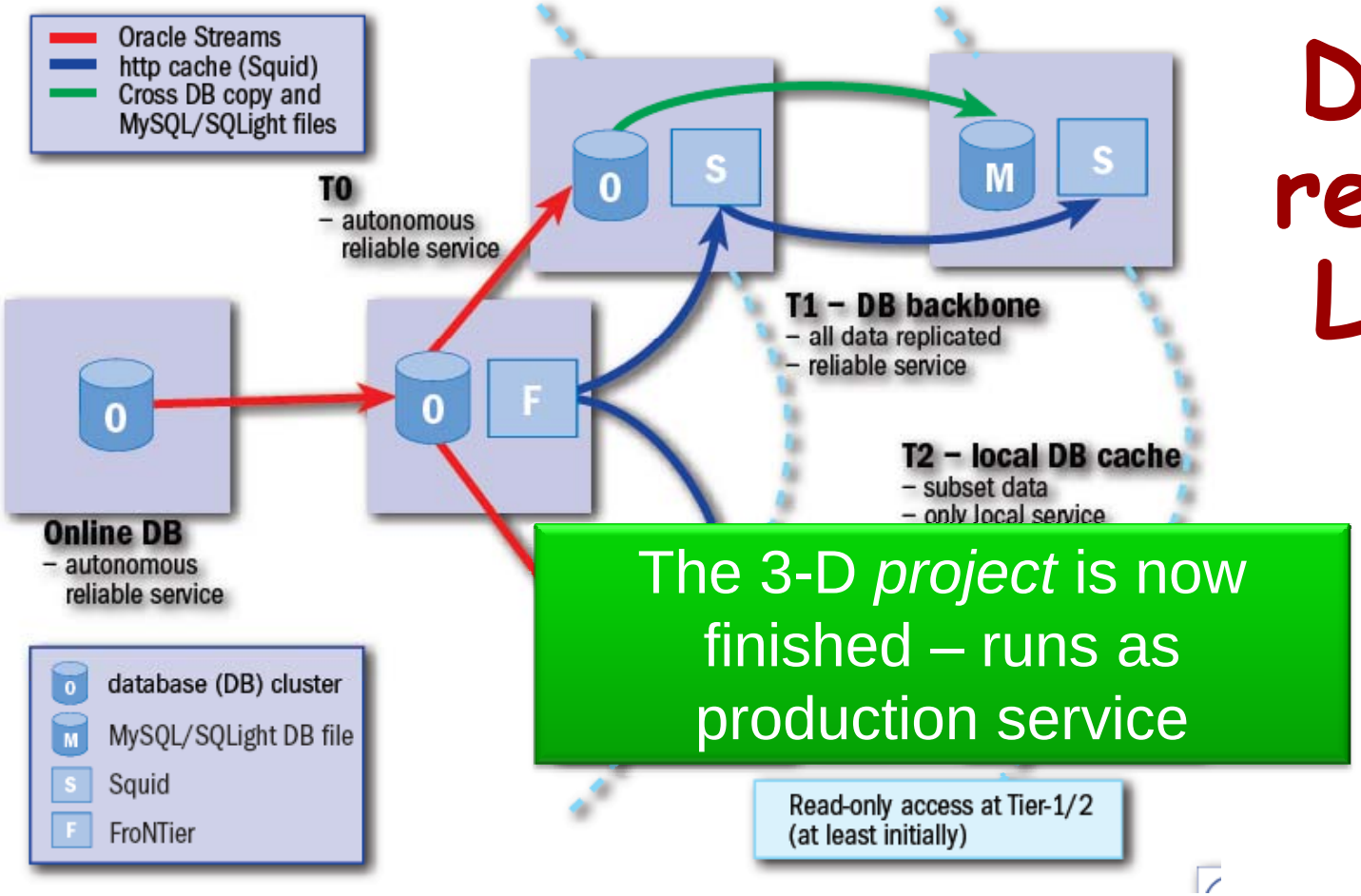
- Storage Element
  - Castor, dCache
  - Storm added in
  - SRM 2.2 – dep
    Dec 2007
- Basic transfer to
- File Transfer Service (FTS)
- LCG File Cat
- LCG data m
- Posix I/O –
  - Grid File
- Synchronis
  - 3D proje

tion System
bility improvements
e Elements
us/Condor-C – improvements
G-CE for scale/reliability
services (CREAM)
ort for multi-user pilot jobs
(glexec, SCAS)
ad Management
tion
ment System (VOMS)
software installation
ing Tools

Focus now on continuing evolution of reliability, performance, functionality, requirements

No expectation of major changes: believe this set is able to manage the levels of workload and performance required for 2008/9. Service reliability/management is an issue in some cases ...

Database replication LCG 3-D



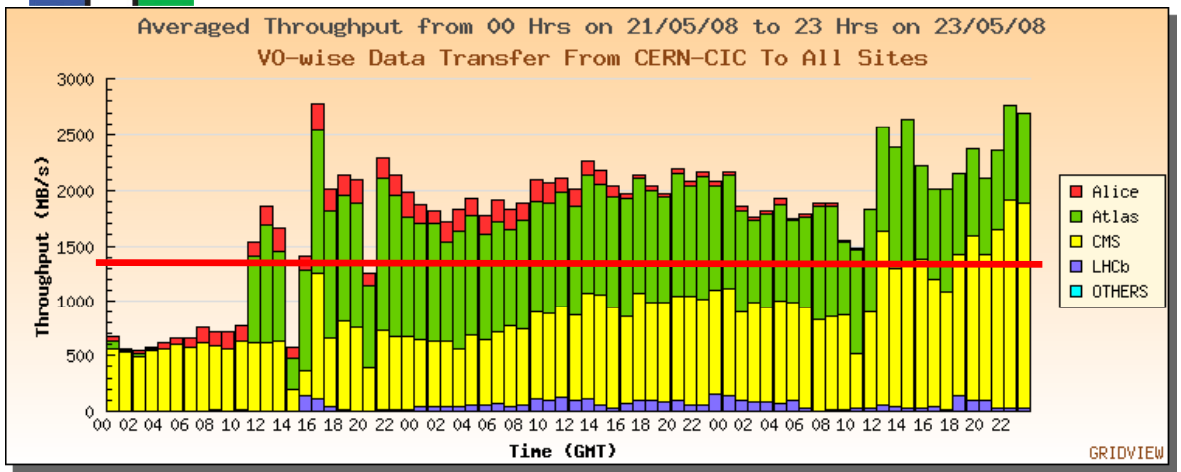
- In full production
  - Several GB/day user data can be sustained to all Tier 1s
- ~100 DB nodes at CERN and several 10's of nodes at Tier 1 sites
  - Very large distributed database deployment
- Used for several applications
  - Experiment calibration data; replicating (central, read-only) file catalogues

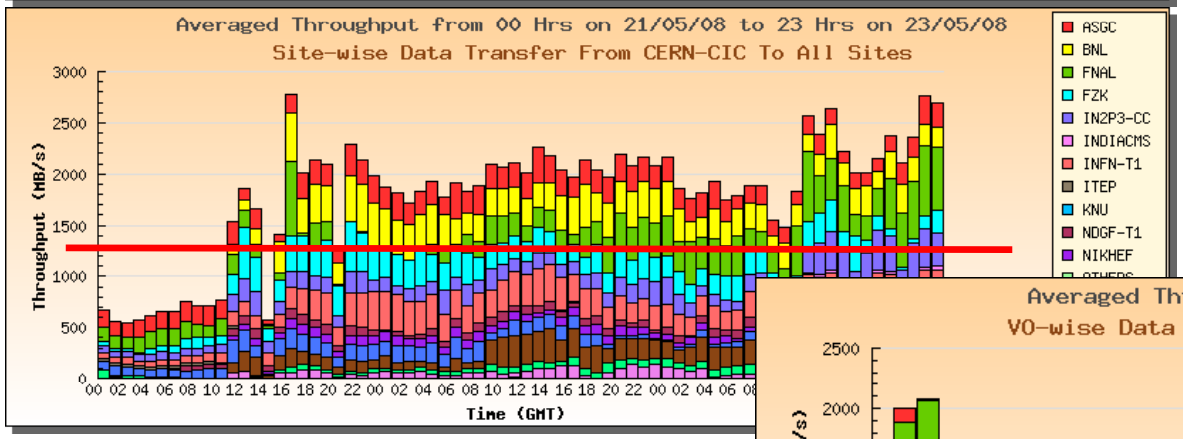
CCRC'08

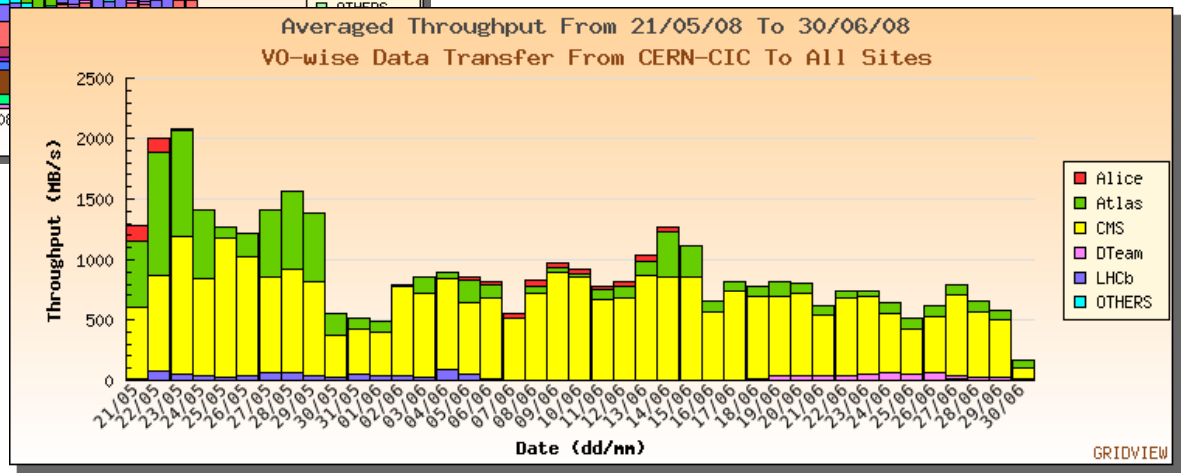
- Combined Computing Readiness Challenge – proposed in 2 phases:
  - Feb: not all 2008 resources in place; new versions of software being tested (e.g. SRM v2.2, & experiment sw)
  - May: all 2008 resources; full 2008 workloads, all aspects of experiment production chains; all experiments together

- Results: (See Jamie's talk)
  - Many sites had problems in getting full 2008 resources in place (procurement, vendor, hardware) issues
  - We have demonstrated a sustainable service model – people were not in panic mode
  - We have demonstrated full 2008/2009 workloads – at sufficient scale
  - Storage systems: SRM v2.2 was in place in time (by January), Feb phase did not show major problems
    - Some issues uncovered in May – workarounds, short term proposals
  - Middleware process works – able to update sw in production
  - Not tested: full simultaneous Tier 1 loads and reprocessing use case

Data transfer results





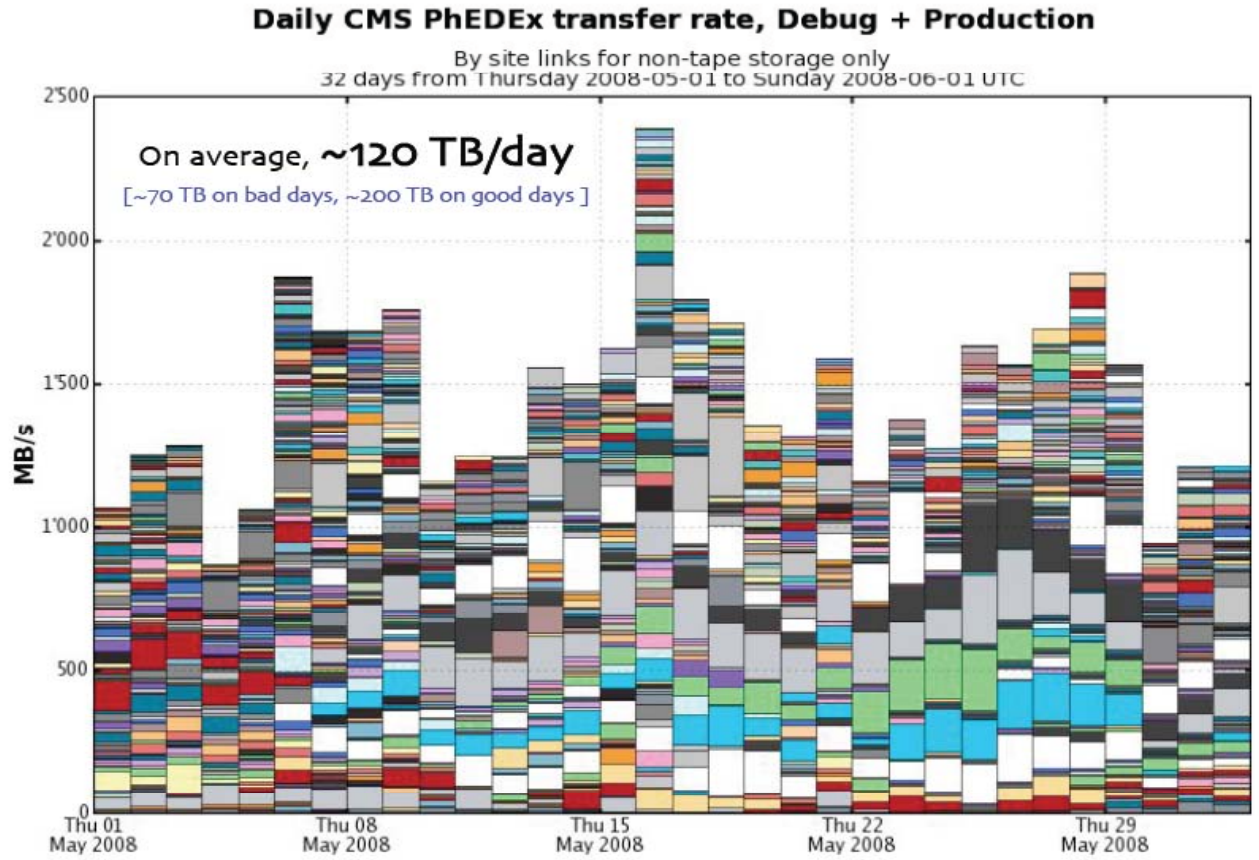


- All experiments exceeded required rates for extended periods, & simultaneously
  - 1.3 GB/s target
  - Well above 2 GB/s achievable
- All Tier 1s achieved (or exceeded) their target acceptance rates

Ian.Bird@cern.ch

Data transfers - CMS

Tier-x to Tier-x in CCRC'08/phase-2



Impressive list of few hundreds of links...

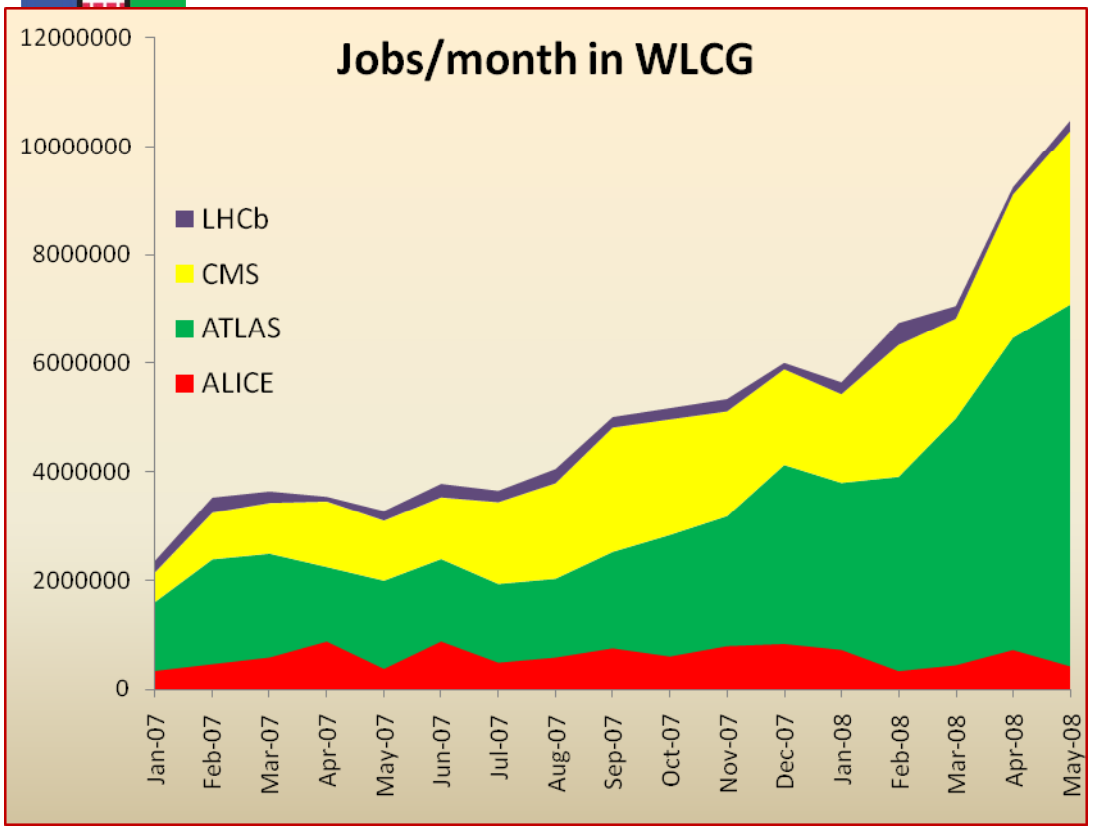
CCRC'08 post-mortem workshop - CERN, 12-13 June 08 D. Bonacorsi 16

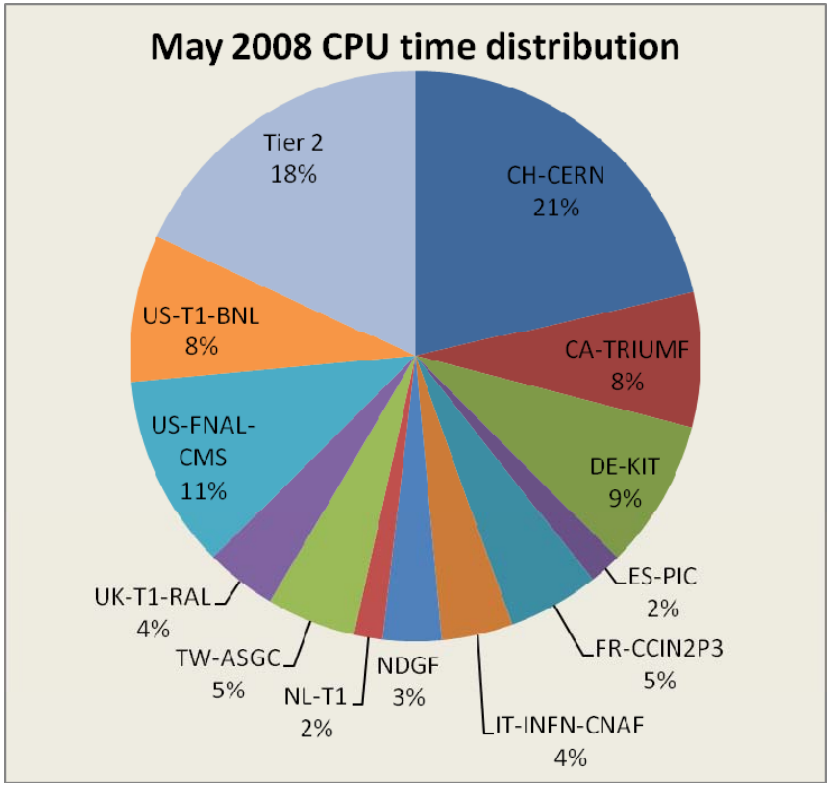
Storage Services Summary

- SRM v2.2
  - Initial deployment was achieved by end 2007; by May all instances of SE were running SRM v2.2 (Tier 1+Tier 2):
    - Castor, dCache, DPM, Storm, (+ BestMan in OSG?)
  - Various issues uncovered during May run (bugs and functional issues):
    - At MB have agreed that priorities are:
      - Bug fixes and reliability/performance issues during use
      - "Short term" functional improvements – to address specific issues found in May (subset of the SRM MoU addendum)
        - These are different issues for the various implementations
      - No other development work will be requested until a review of the situation in the light of experience (e.g. Early 2009)
- As anticipated, site configurations to support the experiment use cases are complex
  - Lessons learned in CCRC have to be addressed – (e.g. How to implement various storage classes and how these map into the functionalities possible for specific disk pools in Castor, dCache, etc.)

➢ This is ongoing effort and will undoubtedly evolve ...

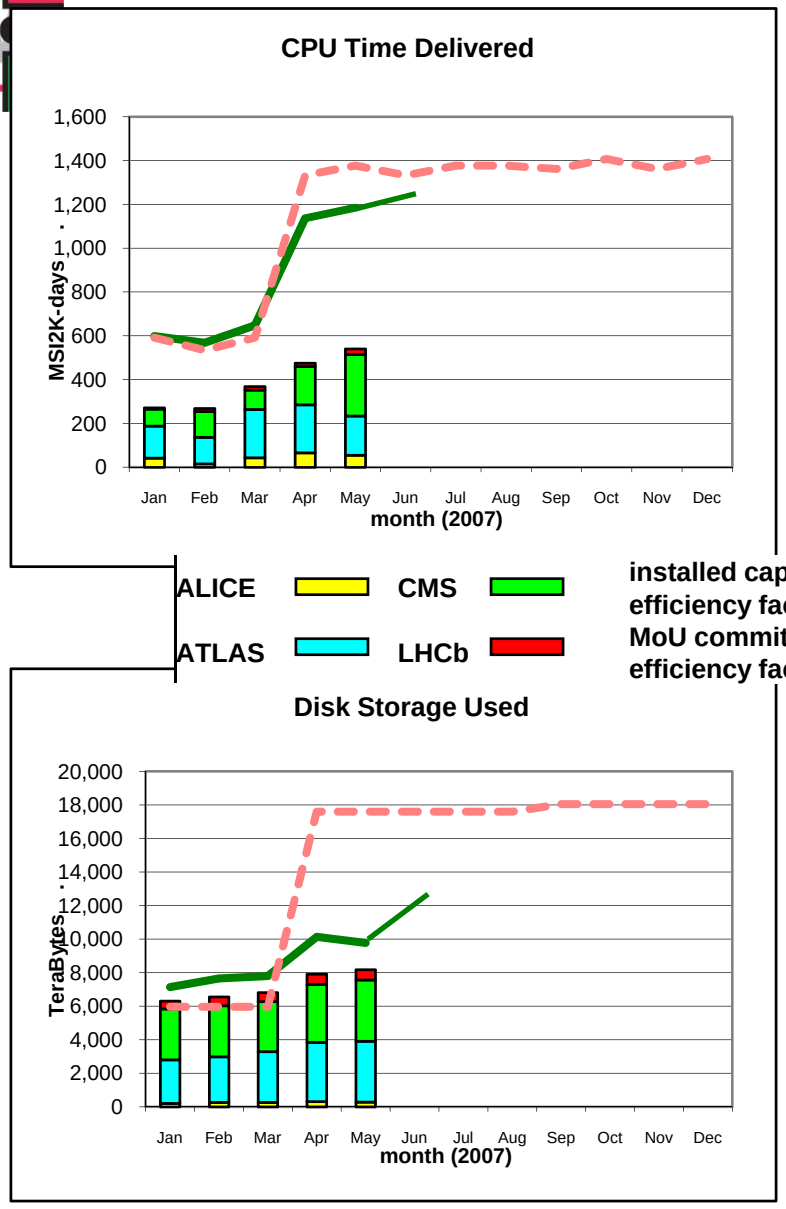
Grid Activity

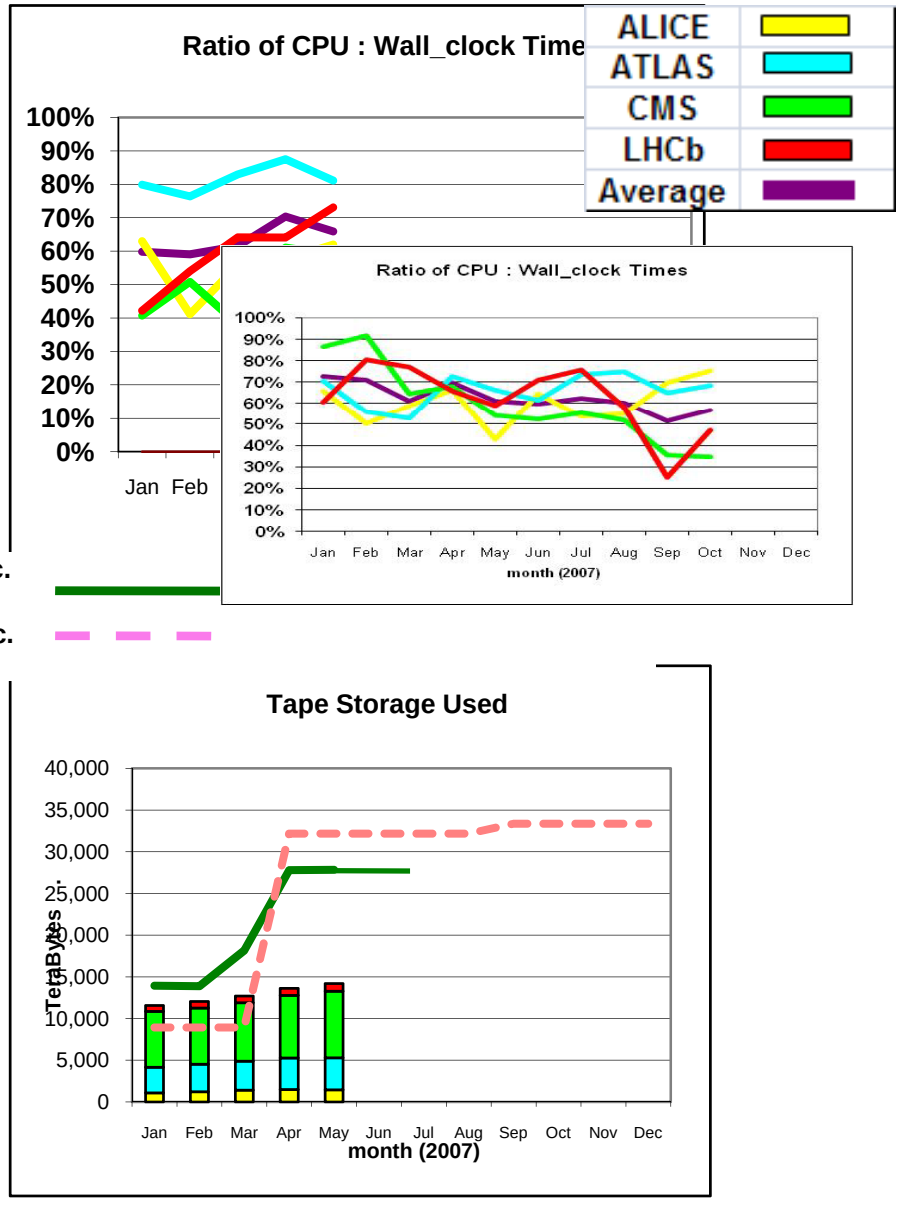




- Average over May total: (10.5 M) 340k jobs / day
  - ATLAS average >200k jobs/day
  - CMS average > 100k jobs/ day with peaks up to 200k
- This is the level needed for 2008/9

Resources


CPU Time Delivered
MSI2K-days
1,600
1,400
1,200
1,000
800
600
400
200
0
Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec
month (2007)
ALICE
CMS
ATLAS
LHCb
installed capacity (inc. efficiency factor)
MoU commitment (inc. efficiency factor)
Disk Storage Used
TeraBytes
20,000
18,000
16,000
14,000
12,000
10,000
8,000
6,000
4,000
2,000
0
Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec
month (2007)


Ratio of CPU : Wall_clock Time
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
Jan Feb
ALICE
ATLAS
CMS
LHCb
Average
Ratio of CPU : Wall_clock Times
Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec
month (2007)
Tape Storage Used
TeraBytes
40,000
35,000
30,000
25,000
20,000
15,000
10,000
5,000
0
Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec
month (2007)

WLCG MoU Signature Status

- Tier-1s : All signatures now obtained – last signature from Sweden for NDGF obtained 10/03/08
- Tier-2s: All signatures now obtained with the exception of Austria – still waiting Finance Ministry approval (confirmed 27/06/08)
- Brazil has announced for a few years their intention to join the WLCG collaboration as a Tier-2, however there is difficulty to obtain an MoU signature. J. Engelen has sent a letter on 20/06/08 which has been forwarded to the President of CNPq National Financial Support Agency on 24/06/08 to clarify the situation – awaiting feedback.

Tier-1 Accounting & 2008 Pledge status

- Monthly reporting continues with reports published on WLCG web
- Latest report currently available January-May 2008
- 2008 MoU pledge values now used since April 2008
- Not all resources were available on 01/04/08 mainly due to procurement timescale, late delivery or delivery of faulty equipment – lessons have hopefully been learnt for the future
- CPU pledges: Most Tier-1s have their CPU pledge in place by now or planned for September 2008 (ASGC, FNAL)
- Disk and Tape pledges: some Tier-1s already have all these pledges in place, others (BNL, PIC) plan for July or CC-IN2P3 plan for September
- More details in the Tier-1 status report by John Gordon

Tier-2 Accounting & 2008 pledge status

- Monthly reporting now established with reports published on WLCG web
- Most sites are now publishing accounting data apart from Norway, Sweden and Ukraine
- Latest report currently available May 2008
- 2008 MoU pledge values now used since April 2008
- All Tier-2 sites contacted on 09/06/08 to get information about their installed 2008 resource status
- Status of replies on 27/06/08:
  - 10 replies from Estonia, Germany GSI, Hungary, India, Israel, Poland, Portugal, Spain (CMS), Switzerland, UK NorthGrid either confirming resources installed or with planned schedule for September (Israel, Hungary, Portugal), November (Poland) or year end (India TIFR, Germany GSI)
  - The request stated that no reply by 01/07/08 implied all resources fully available
  - Conclusion: remaining 54 sites ?

Revised Pledges 2009-2013

- The Management Board (MB) meeting of 02/10/07 agreed that experiments would not update their resource requirements until after December 2008 based on initial data experience
- Experiments have only been requested to estimate their 2013 requirements: data either received or expected to be received by 06/07/08
  - Triggered an MB discussion on the feasibility of the WLCG MoU 5 year forward look. To be raised at next C-RRB meeting.
- In preparation for the next C-RRB meeting and in accordance with the current MoU timeframe, Tier-1s and Tier-2s were contacted on 09/06/08 and asked by 20/10/08 to:
  - confirm their pledge values for 2009
  - provide planned values for 2010-2013 inclusive
- Replies received to date confirm planned pledges for 2009. Portugal increases planned pledges for 2009 CPU: from 750 to 1600 kSI2K, Disk: from 160 to 700 Tbytes – may this increase in 2009 resources continue!

Pledge status 2008

- The table below shows the snapshot for 2008 status at 27/06/08
- % indicates the balance between offered and required

	ALICE	ATLAS	CMS	LHCb	Sum 2008
T1 CPU	-45%	6%	7%	43%	-5%
T1 Disk	-40%	2%	-23%	33%	-12%
T1Tape	-49%	-5%	-4%	39%	-13%
T2 CPU	-47%	2%	35%	-7%	-2%
T2 Disk	-20%	-17%	-11%	-	-14%

Pledge status 2008-2012

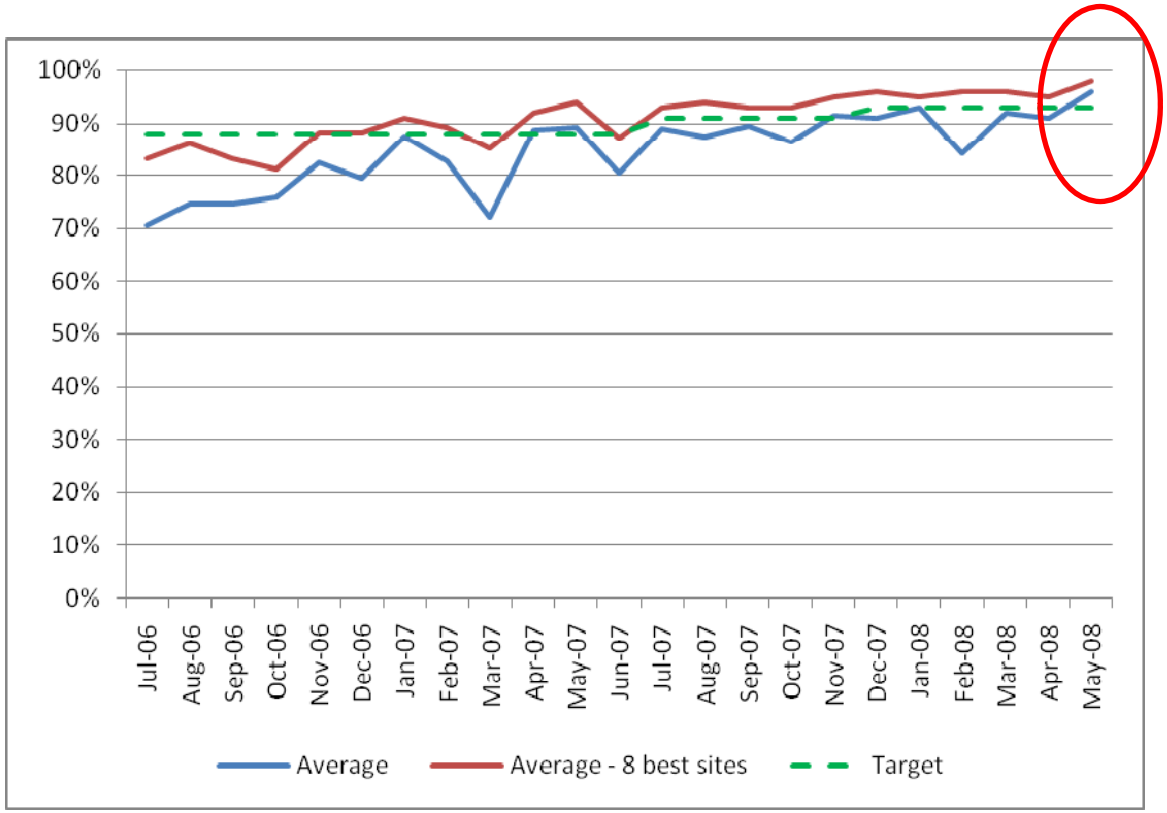
- The table below shows the snapshot for 2008-2012 status at 27/06/08
- % indicates the balance between offered and required
- Not enough data received yet to include information for 2013
- Current focus is on 2009, and it is hoped by the end of the pledge revision exercise, when all sites have confirmed their 2009 pledges that this picture gets greener

	Sum 2008	Sum 2009	Sum 2010	Sum 2011	Sum 2012
T1 CPU	-5%	-11%	-11%	-17%	-23%
T1 Disk	-12%	-12%	-15%	-17%	-24%
T1 Tape	-13%	-13%	-17%	-22%	-28%
T2 CPU	-2%	-13%	-34%	-37%	-43%
T2 Disk	-14%	-3%	0%	-10%	-20%

Resource Scrutiny Group

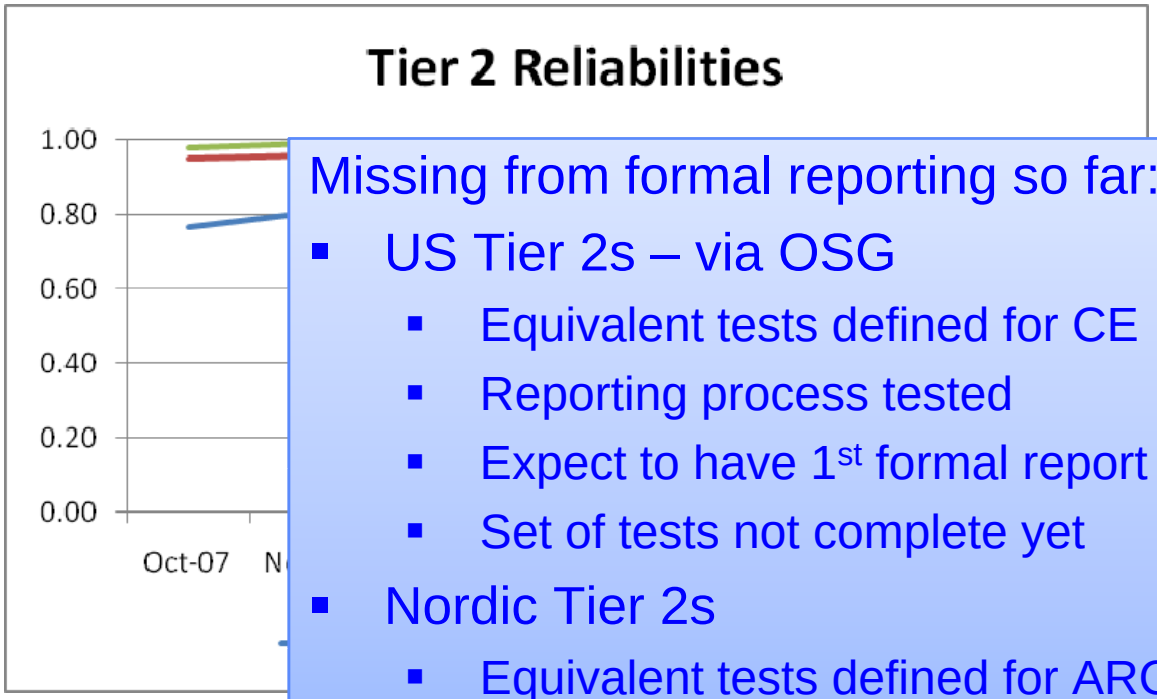
- The RSG has now been set up:
  - Chairman: Dominec Espriu (Spain)
  - Has met several times in the last few months
  - Has nominated referees to scrutinize each experiment's requirements; expect 1st report to C-RRB in November
    - 2 referees for each experiment (3 for ALICE)
  - Jürgen Knobloch as link to LHCC
- Mandate:
  - As specified in the WLCG MoU (Annex 9, items 5 and 6) every year the C-RSG shall scrutinize
    - The resource accounting figures for the preceding year
    - The use the experiments made of these resources
    - The overall request for resources for every experiment for the following year and forecasts for the subsequent two years
  - The C-RSG will also examine the match between the refereed requests and pledges from the Institutions and make recommendations concerning apparent under-funding. The C-RSG is not expected to perform the role of mediator between the experiments and the resource providers.

Site reliability: CERN+Tier 1s



	Jan 08	Feb 08	Mar 08	Apr 08	May 08
Target	93	93	93	93	93
Average – 8 best site	96	96	96	95	98
Average – all sites	90	85	91	91	96
# above target (+>90% target)	7 +3	7 +3	7 +3	7 +3	11 +1

Site Reliability: Tier 2s



Missing from formal reporting so far:

- US Tier 2s – via OSG
  - Equivalent tests defined for CE
  - Reporting process tested
  - Expect to have 1st formal report for June
  - Set of tests not complete yet
- Nordic Tier 2s
  - Equivalent tests defined for ARC (used by NDGF)
  - Many Nordic Tier 2s not yet set up

- Formal r
  - #sites
  - Overall average: 75-80%, but top 50% (20%) of sites: 95% (98%)
  - More than 70% of resources are at sites with >90% reliability

Monitoring

- More sites now integrating SAM results into site monitoring and alarming (~30)
  - Working on publicizing work more
    - Tutorials at last WLCG Workshop & upcoming EGEE'08
- EGEE-III has taken the WLCG Nagios-based prototype as a blueprint for monitoring
  - Will be deployed at all sites/ROCs over next 2 years
- Integrating dashboard visualization of SAM

LOGO — SAM VISUALIZATION

VO view: Latest Results | HistoricalView — Site View — Feedback

Sites: Tier0 + Tier1s (Tier0 + Tier1s, BNL-LCG2, CERN-CIC, CERN-PROD, CERN-SC, CERN_PPS, FZK-LCG2)
Service Types: VO critical (Select all, CE, FTS, LFC, MyProxy, OSGCE, RB)
Test Types: FCR Tests (Select All, CE-sft-brokerinfo, CE-sft-caver, CE-sft-csh, CE-sft-job, CE-sft-lcg-rm, CE-sft-softver)
Test Exit Status: All Exit Status (na, ok, info, note, warn, error, crit) — Show Results

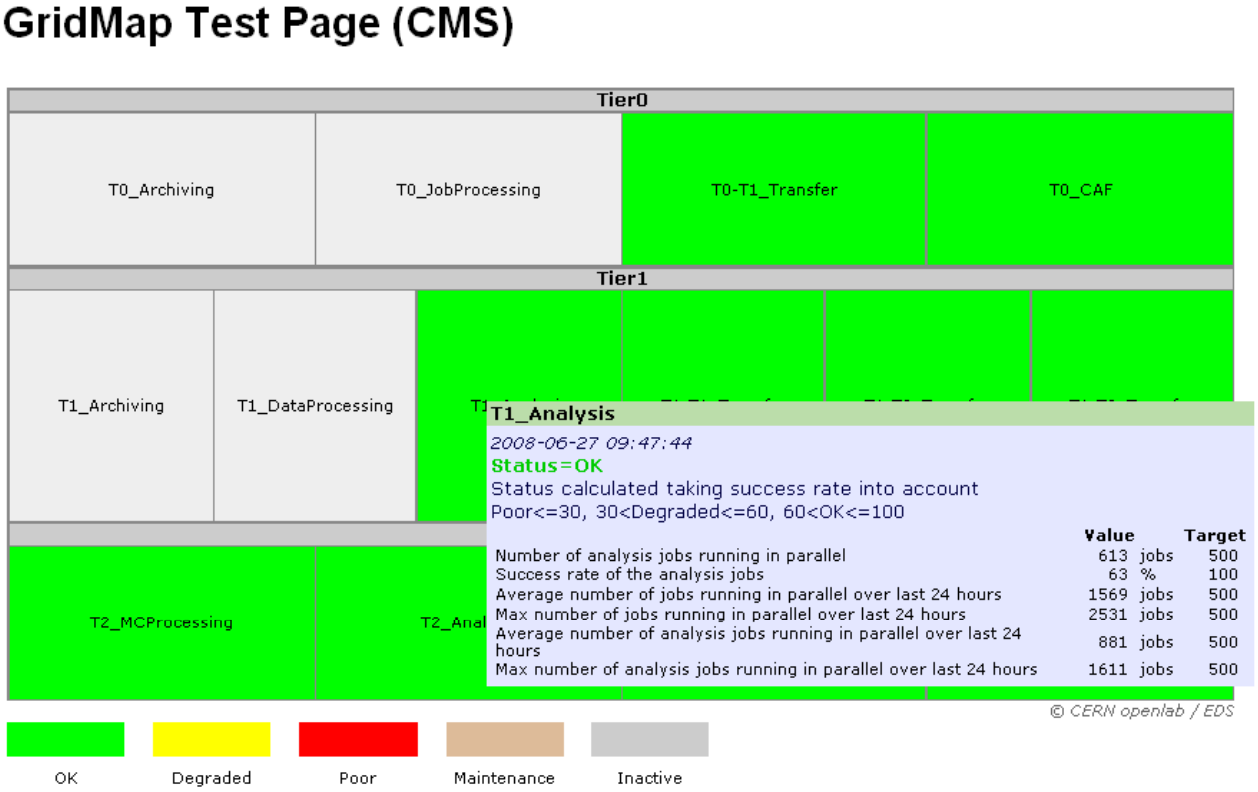
Legend: NA | OK | MAINTENANCE | ERROR | WARNING | INFO | NOTE | CRITICAL
Note: brightest colors: test is 0 - 6 hours old, ... lightest colors: test is more that 24 hours old

Link to the table

Sitename	Service Type	Service Name	ver	rm	csh	ca	bi	js	ftsinfo	ftschn	lfcwf	lfcls	js	jc	jm	put	get	del
BNL-LCG2	FTS	lcg03.usatlas.bnl.gov							ok	ok								
	SE	dcsrm.usatlas.bnl.gov														ok	ok	ok
	SRM	dcsrm.usatlas.bnl.gov														ok	ok	ok
	SRMv2	dcsrm.usatlas.bnl.gov														ok	ok	ok
CERN-PROD	CE	ce101.cern.ch	ok	ok	ok	ok	ok	maint										
		ce102.cern.ch	ok	ok	ok	ok	ok	maint										
		ce103.cern.ch	ok	ok	ok	ok	ok	ok										
		ce104.cern.ch	ok	ok	ok	ok	ok	ok										
		ce105.cern.ch	ok	ok	ok	ok	ok	ok										
		ce106.cern.ch	ok	ok	ok	ok	ok	warn										
		ce107.cern.ch	ok	ok	ok	ok	ok	warn										
		ce108.cern.ch	ok	ok	ok	ok	ok	ok										
		ce109.cern.ch	ok	ok	ok	ok	ok	ok										
		ce111.cern.ch	ok	ok	ok	ok	ok	ok										
		ce112.cern.ch	ok	ok	ok	ok	ok	ok										
		ce113.cern.ch	ok	ok	ok	ok	ok	ok										
		ce114.cern.ch	ok	ok	ok	ok	ok	error										
		ce115.cern.ch	ok	ok	ok	ok	ok	ok										
		ce116.cern.ch	ok	ok	ok	ok	ok	maint										
		ce117.cern.ch	ok	ok	ok	ok	ok	maint										
		ce110.cern.ch	ok	ok	ok	ok	ok	maint										
		ce119.cern.ch	ok	ok	ok	ok	ok	maint										
		ce120.cern.ch	ok	ok	ok	ok	ok	maint										
		ce121.cern.ch	ok	ok	ok	ok	ok	maint										
		ce122.cern.ch	ok	ok	ok	ok	ok	maint										
		ce123.cern.ch	ok	ok	ok	ok	ok	maint										
		ce124.cern.ch	ok	ok	ok	ok	ok	error										

VO Monitoring

- VO Maps
  - Extension of gridmaps to show VO workflows in gridmap format
    - Data supplied from dashboards
    - Data Transfer, Data Processing, Data Archiving
    - Shows work from all VOs to sites in a single place



Ian.Bird@cern.ch

Other Milestones

- Most Tier 1-ralated discussed in Tier 1 status talk

CAF CERN Analysis Facility						
WLCG-07-48	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 May 2008	LHCb May 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	

- In addition:
  - Reliability milestones – now on each site to be above target
  - Tier 2 sites – will start to follow up by federation, reporting at RRB: have introduced milestones for targets
  - Start to follow up on VO-specific availability

Future infrastructure

- EGEE
  - EGEE-III has been approved and began in May 2008 (until April 2010)
  - Effort ~20% less than that of EGEE-II
    - Support for specific applications (inc HEP) and middleware cut significantly
- EGI
  - Design study has produced a draft of the draft blueprint
  - First workshop to present it was held yesterday (June 30) at CERN
  - Now WLCG has to understand the implications and document how it will operate in this environment; to be used as
    - Feedback to the EGI design study, and
    - As basis for discussions with Tier1s and Tier2s – how does the model help or hinder them meet their WLCG commitments?
    - Initial fall-back plan if this turns out to be required
- OSG
  - Funded until 2010; no information about future evolutions

Summary

- CCRC'08 demonstrated
  - sustainable service support , including middleware process
  - Data transfers in excess of needed levels
  - Workloads at scale needed for data taking
  - Still to validate some parts of T1 loads in computing models
- Storage systems
  - Basic SRM v2.2 functionality in place
  - Short term workarounds and configurations, specific developments (by end of year), review status in early 2009
  - Tier 1 configurations of disk pools still evolving to meet needs
- Resources
  - Issues shown during procurement/install cycle – concern for future years when less leeway
  - Regular reporting of MoU commitments (accounting, reliability) – almost all sites now

➢ Experiments and sites in data-taking mode from now on ...