# WLCG Status Report

**WLCG**
**Collaboration Board**
**13th November 2009**

**Ian Bird**
**LCG Project Leader**

Huge $10 billion collider resumes hunt for 'God particle'

By **Elizabeth Landau**, CNN
November 11, 2009 8:12 a.m. EST

one
the LHC

**Crumbs! Large Hadron Collider suffers snack-related bird mishap**

By **Rich Trenholm** on **06 November 2009**, 12:58pm

**Large Hadr**

**Introduction**

**STORY HIGHLIGHT**

• The LHC will circula
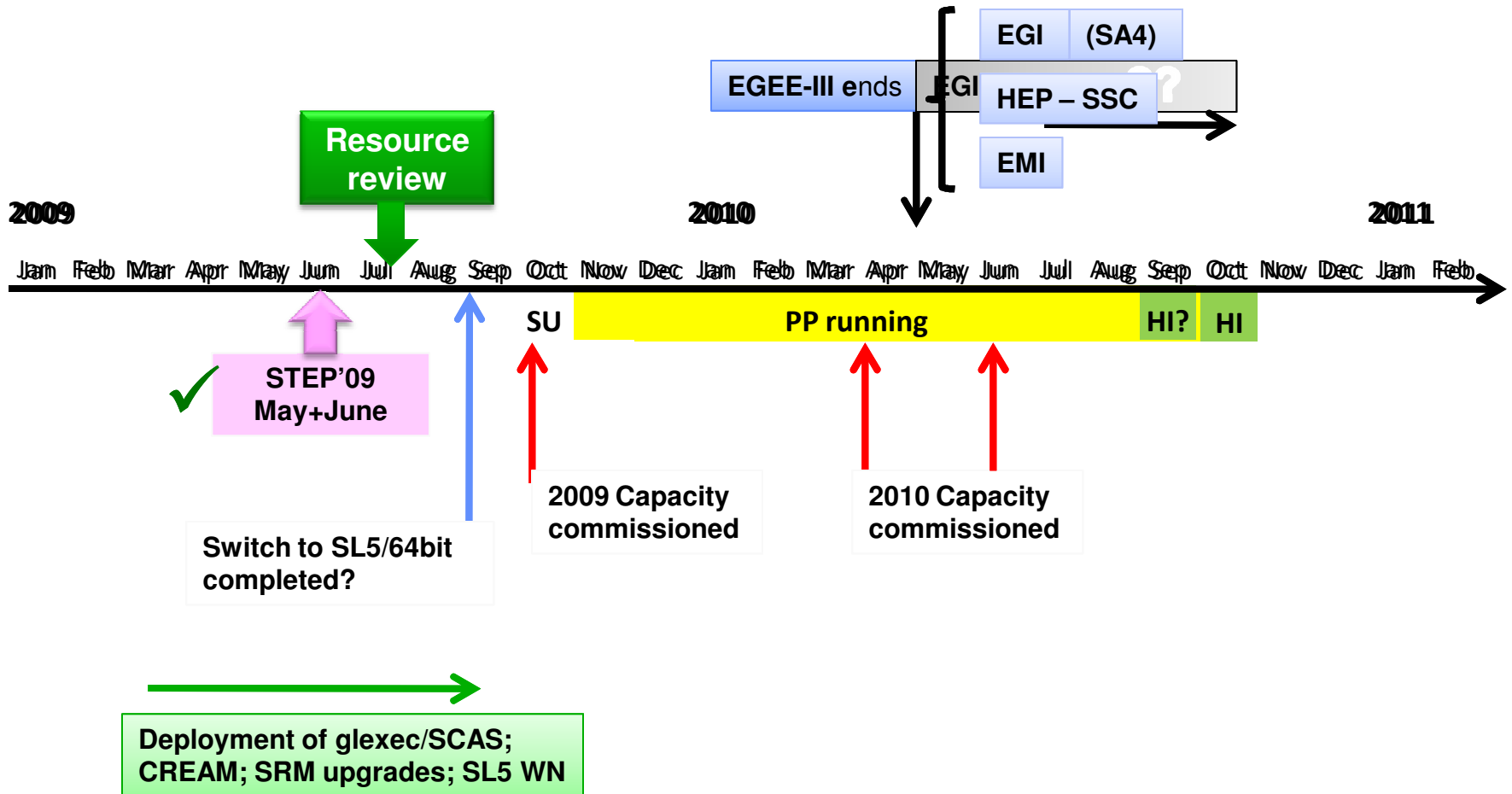around the tunnel in

**LHC Cooldown Status**

# Agenda

- Status – STEP'09 report & follow-up
- Service performance – reliabilities etc
- Middleware and storage systems
- Resource situation – following RRB
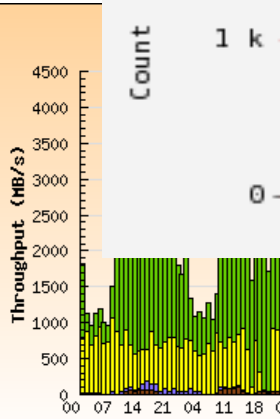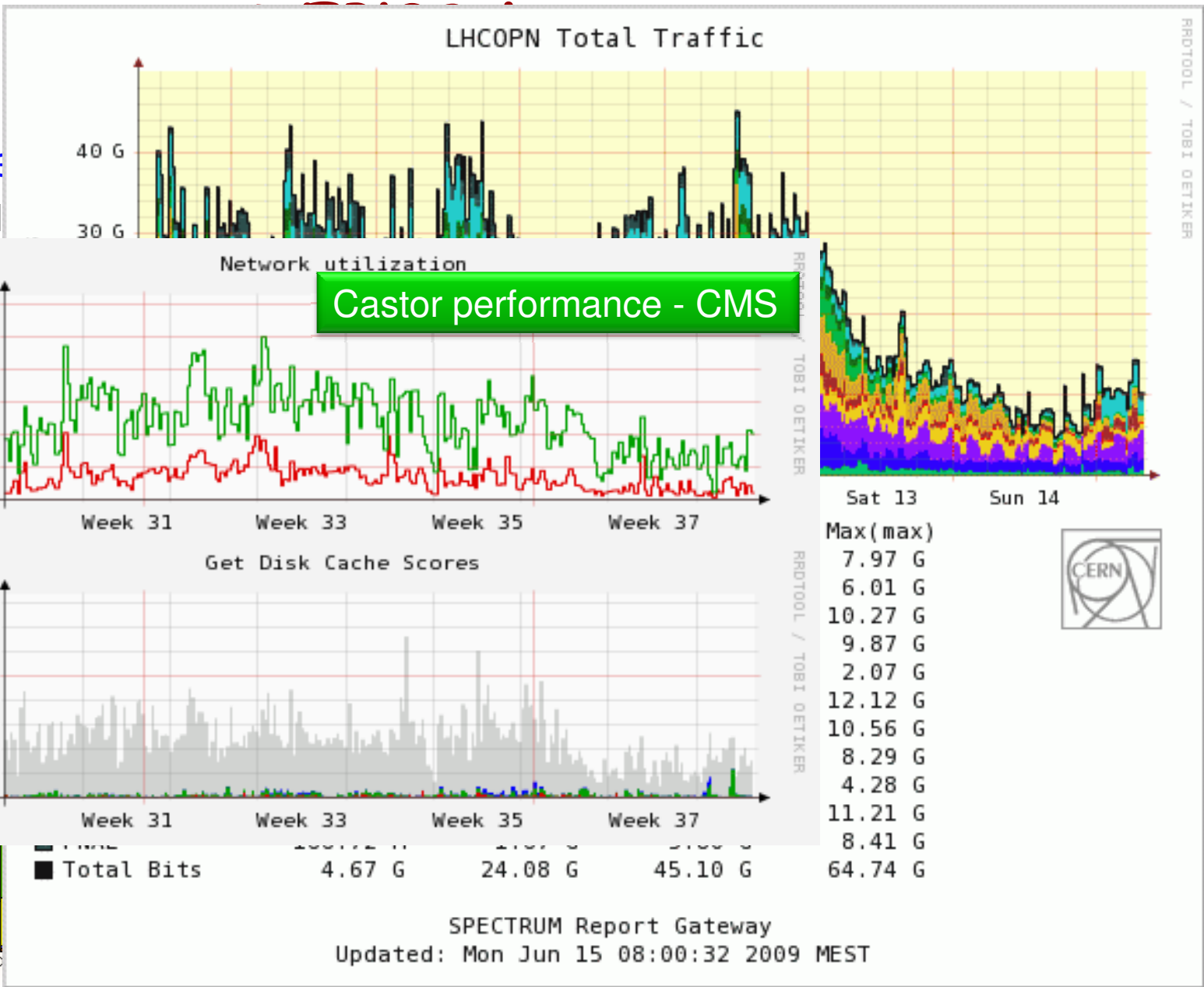- Status of preparations for EGEE → EGI transition

# WLCG timeline 2009-2010

EGI    (SA4)

EGEE-III ends    EGI    HEP – SSC    ?

EMI

**Resource review**

2009    2010    2011

Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec Jan Feb

SU    PP running    HI?    HI

✓    **STEP'09 May+June**

**Switch to SL5/64bit completed?**

**2009 Capacity commissioned**

**2010 Capacity commissioned**

**Deployment of glexec/SCAS; CREAM; SRM upgrades; SL5 WN**

Gene

M

Castor performance - CMS

**LHCOPN Total Traffic**

**Network utilization**

**Get Disk Cache Scores**

| | | Max(max) |
|---|---|---|
| | | 7.97 G |
| | | 6.01 G |
| | | 10.27 G |
| | | 9.87 G |
| | | 2.07 G |
| | | 12.12 G |
| | | 10.56 G |
| | | 8.29 G |
| | | 4.28 G |
| | | 11.21 G |
| | | 8.41 G |

Total Bits        4.67 G        24.08 G        45.10 G        64.74 G
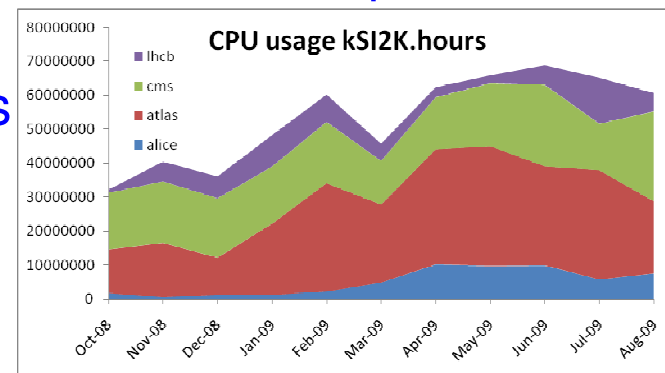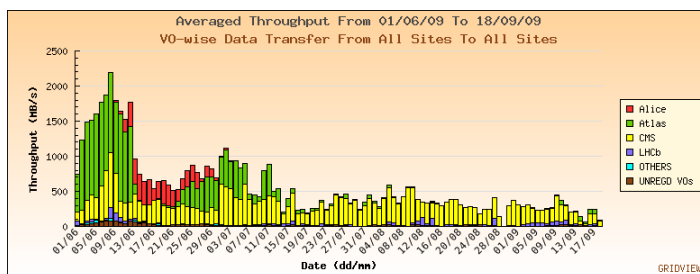
SPECTRUM Report Gateway
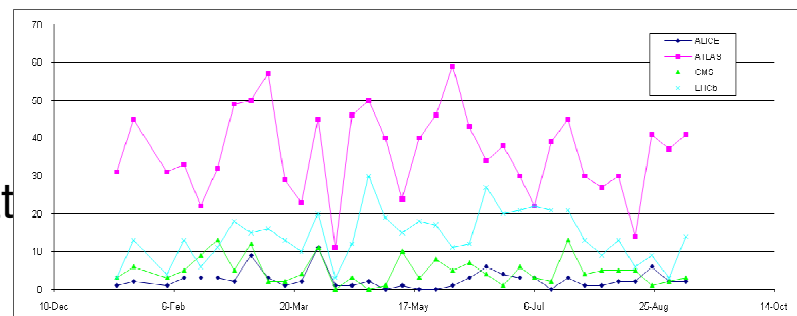Updated: Mon Jun 15 08:00:32 2009 MEST

# After STEP'09 ...

- Since July has been fairly quiet
- Experiments pursue 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

# T1 and T2 Site Readiness

## The Site readiness is closely monitored:

- Reports and follow-up during weekly Facility Operations meetings
- Additional meetings to focus on Asian and Russian&Turkish sites



Site Readiness Status for CMS Tier-1 sites



Site Readiness Status for CMS Tier-2 sites
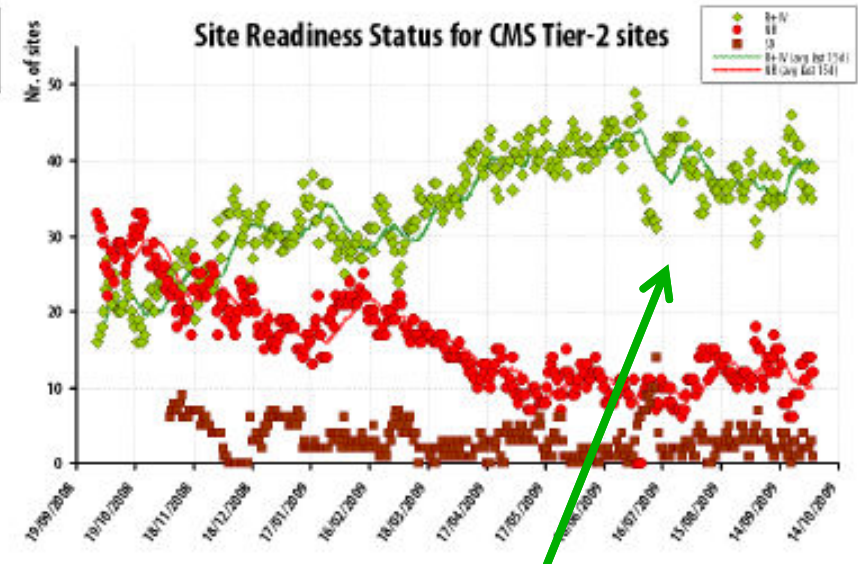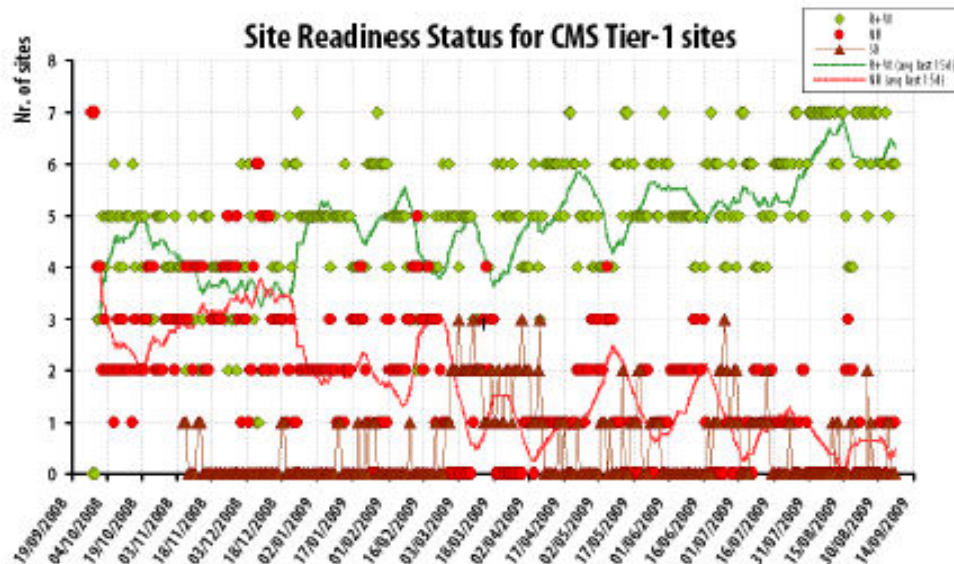
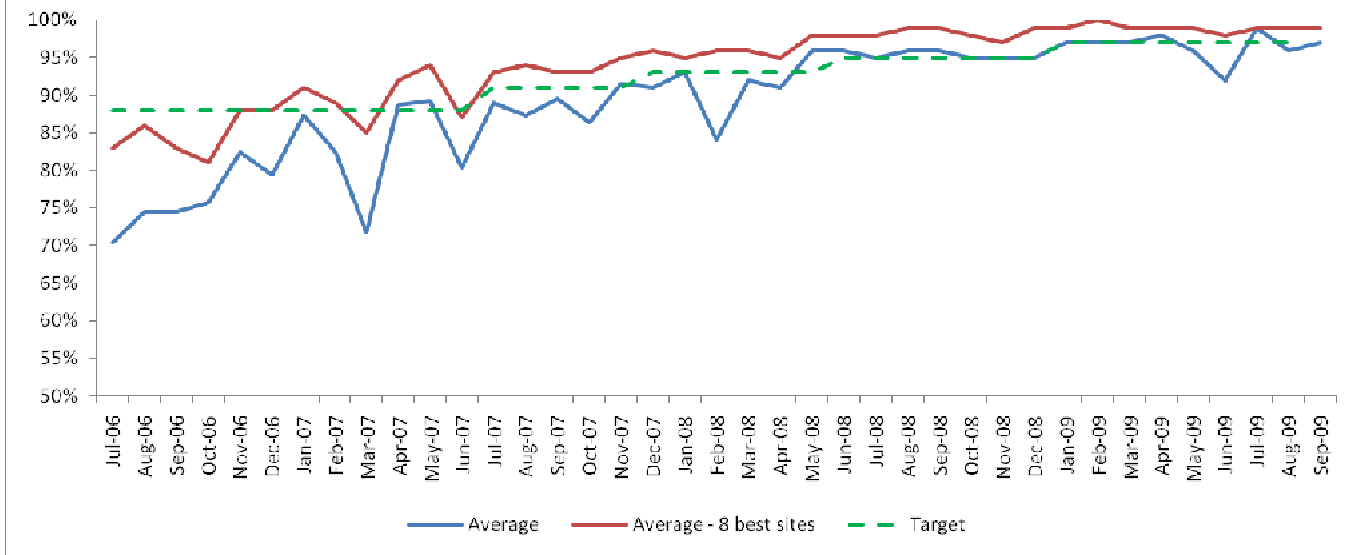## Substantial improvement is observed for large number sites.

- Tier-1: sites readiness a concern. Improving lately.
  - Plan expert visits to improve the situation .
- Tier-2: readiness state improved significantly over the last year. Need to sustain efforts.

General observation:
Levels of effort during holidays?

**Quality Ranking for T2 Site Readiness last 30 days**

## March 2009
- 53 sites listed
- 26 sites > 80%
- 6 sites between 60-80%
- 21 sites < 60%

**Quality Ranking for T2 Scheduled Downtimes last 30 days**

**Quality Ranking for T2 Site Readiness last 30 days**

## October 2009
- 50 sites listed
- 30 sites > 80%
- 12 sites between 60-80%
- 8 sites < 60%

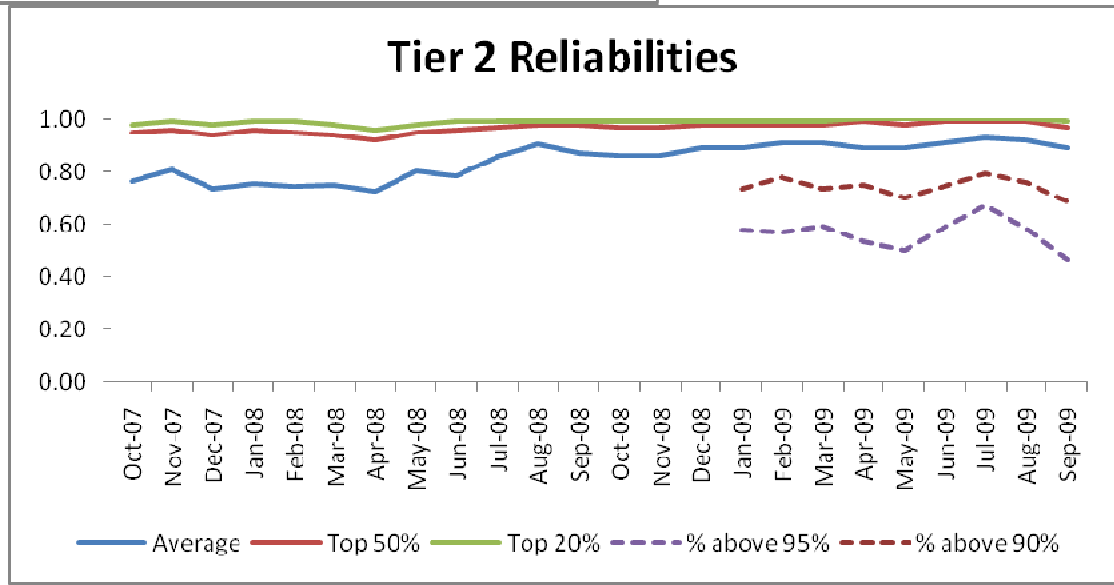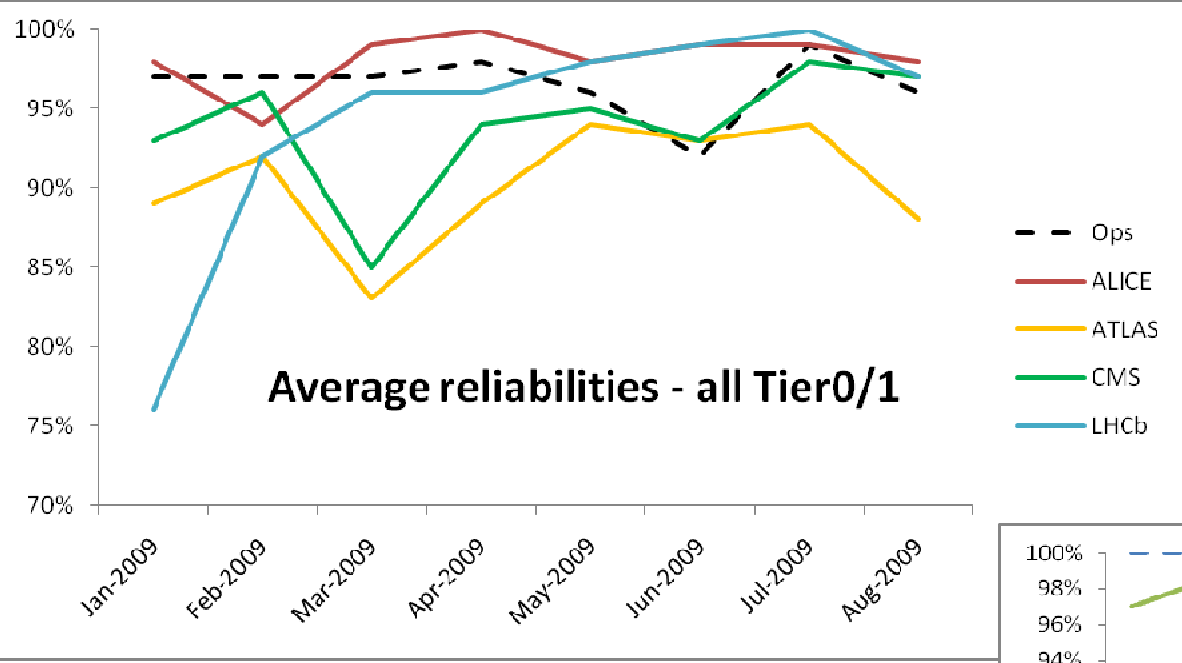**Parallel T2 jobs: Aug.-Sept. '09**

MC production
Analysis
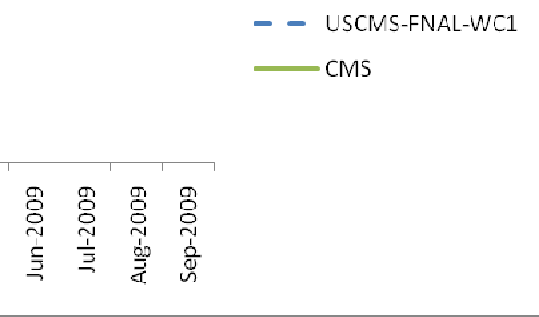
# Reliabilities
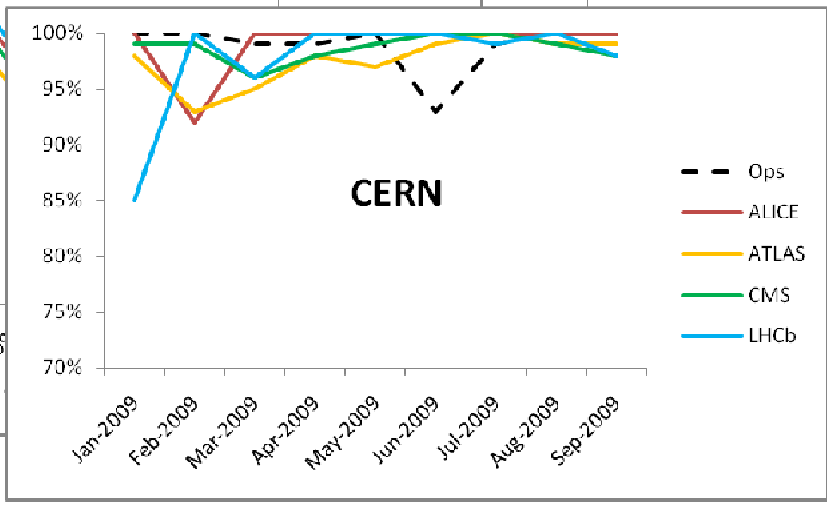


Site Reliability: CERN + Tier 1s



Tier 2 Reliabilities

# Reliabilities – per experiment



Average reliabilities - all Tier0/1

Data prior to March not always reliable

Ops tests not always a good indication of what the experiments see ...

RAL

CERN
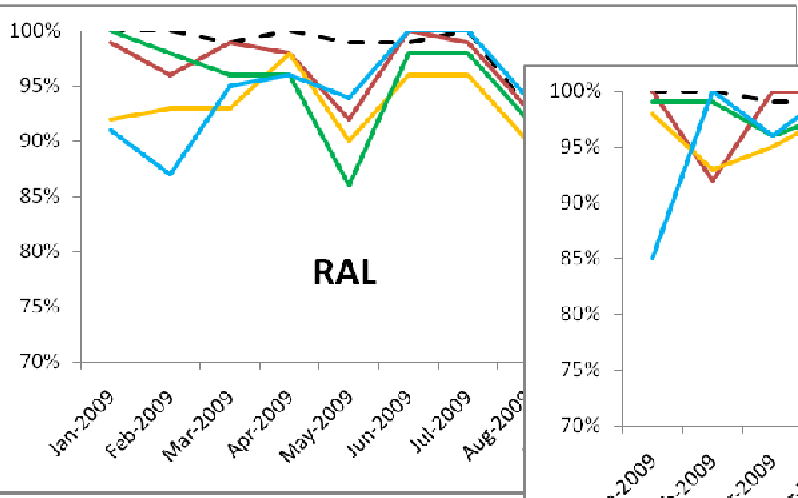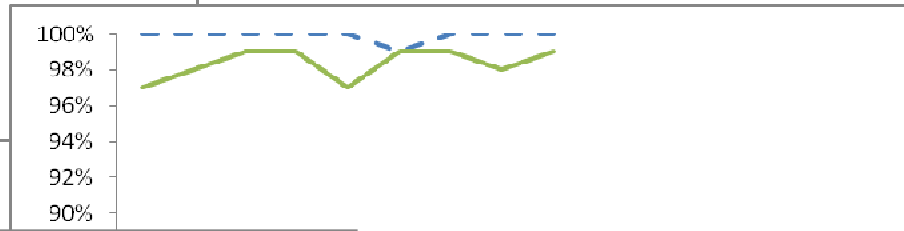
ATLAS — Site Availability using WLCG_SRM2
79 Days from Week 26 of 2009 to Week 37 of 2009
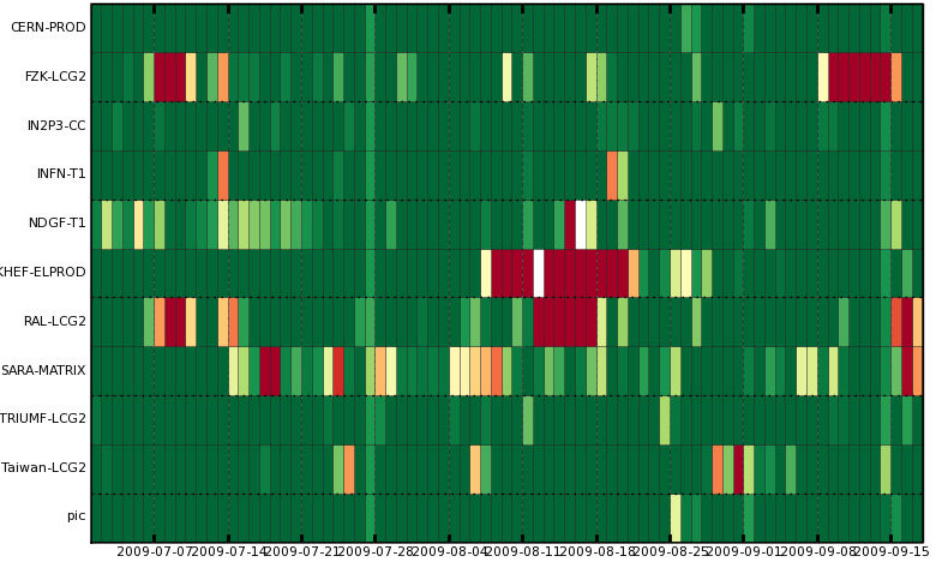
CMS — Site Availability
79 Days from Week 26 of 2009 to Week 37 of 2009

LHCb — Site Availability using LHCb Critical Availability
79 Days from Week 26 of 2009 to Week 37 of 2009

ALICE — Availability using WLCG Availability (FCR critical)
79 Days from Week 26 of 2009 to Week 37 of 2009

# Service outages

**Q409**

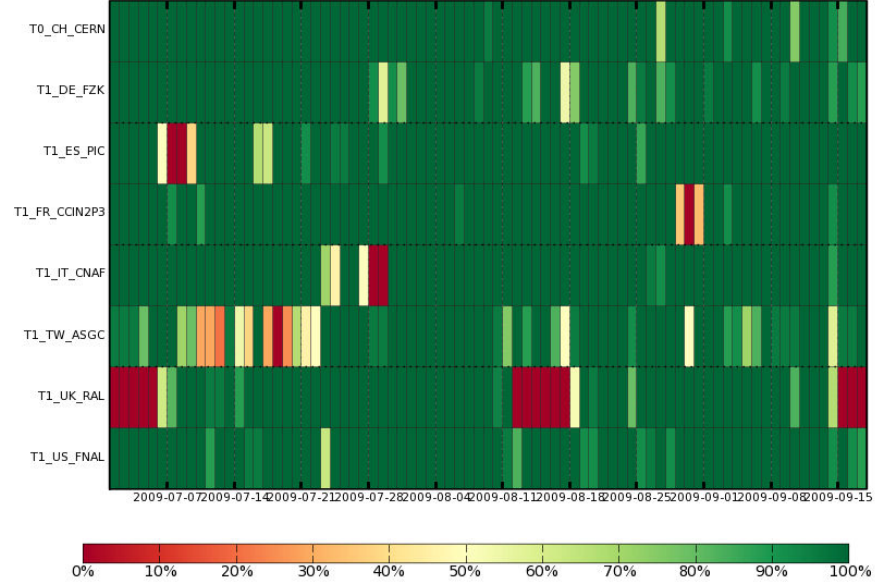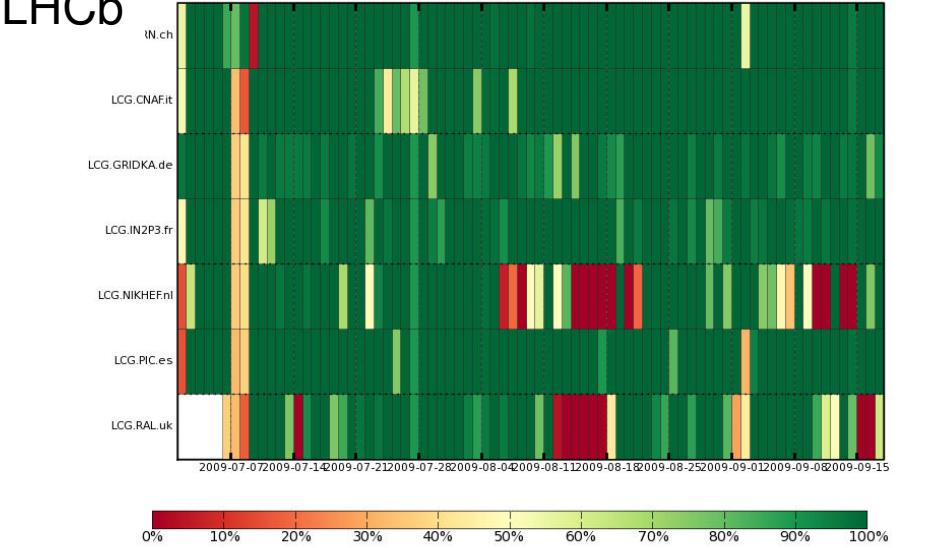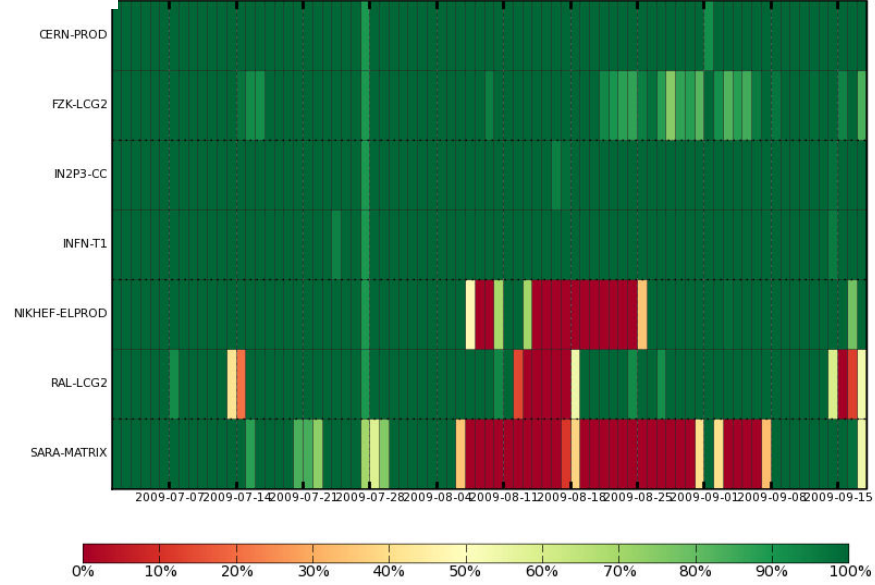| | Date | Duration | Service | Impact |
|---|---|---|---|---|
| IN2P3 | 3 Nov | 4h | Many | Many services have been disturbed due to automatic reboot of machines |
| RAL | 9 Oct | n/a | Storage (Castor) | data loss from Castor |
| IN2P3 | 14 Oct 2009 | 13h | batch | only very short jobs able to run |
| CERN | 13 Oct 2009 | 1-2h | CASTOR nameserver sick | All CASTOR services dead |
| IN2P3 | 8 & 10 Oct 2009 | 11h (8 Oct) and 6h (10 Oct) | SRM crashed | SRM service interrupted |
| RAL | 4-9 Oct 2009 | | disk failures -> Oracle problems | CASTOR, LFC and FTS services down |
| ASGC | 27 Sep - xx Oct | >3 weeks | DBs | down & out |

**Q309**

| | Date | Duration | Service | Impact |
|---|---|---|---|---|
| CERN | 21 Sep 2009 | 08:00 - 18:00 | DB Replication | ATLAS Replication Tier0->Tier1 down |
| RAL | 15 - 17 Sept 2009 | 2 days | CASTOR | Disk to Disk (D2D) transfers started failing during a planned upgrade to the NS |
| FZK | 7 - 16 Sep 2009 | 10 days | ATLAS RAC | 3D Streams replication blocked then degraded |
| CERN | 5 & 8 Sept 2009 | 2 * 2 hours | CASTOR LHCb | two Castor Database problems |
| CERN | 26 Aug 2009 | 18:40 - 23:30 | Batch | Public and production queues closed |
| ASGC | 17 Jul 2009 | 6:00 - 10:00 | Power cut | Most services went down and restarted |
| ATLAS | 13 Jul 2009 | 10:00 - 11:00 | Central Catalogs | Degrade of perform |

Reports now systematically produced for outages resulting in service unavailability for a few hours or more

Followed up in daily operations meetings and weekly in MB

https://twiki.cern.ch/twiki/bin/view/LCG/WLCGServiceIncidents

# Operational Issues

- Patching & security
  - Serious vulnerabilities in Linux kernel required rapid updates
  - Security coordination worked as expected – via EGEE and OSG security contacts (ROCs etc) to sites
  - Except....
  - Updates were not done by many (most!) sites
  - ➔ posing serious risk (and embarrassment) to entire infrastructure
  - Raised to EGEE PMB (and GDB)
    - EGEE threat of site suspension unblocked things
- Issue:
  - Worry that many sites do not do normal security patching
  - (already know that it is hard to get sites to update middleware)
- We (security team) monitor with tools that access public information about a site – no special privileges
  - Sites must take this seriously and maintain regular security updates
    - When OSCT prompts for serious vulnerabilities – action is needed on a reasonable timescale

# Mass storage

- Generally very good performance
  - Some specific issues being investigated
- 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) for better scalability (recommended for large sites)
    - Site decision, but had full discussion of risks/advantages
- 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 been requested, but put on hold in order to ensure stable versions

# Middleware

- Generally:
  - Continuous process of patching and updating middleware as needed during STEP09

- Some upgrades with improved performance/functionality/stability:
  - 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
  - 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

- ➔ Middleware is not really an issue now ...

# 2009/2010 resources

- Requirements were re-assessed by experiments early this year
- Reviewed by LHCC and C-RSG July – September
- New requirements converged and presented at C-RRB in October

- "Computing should not inhibit the experiments' abilities to analyse the first LHC data"
  - Although stronger justifications should be required in future

- Overall at CERN and Tier 1s full request will be available apart from ALICE
  - Tier 2 pledges for 2010 OK for CMS, LHCb; slightly low for ATLAS, significantly low for ALICE.

# Resource situation – following RRB

| CERN Tier0 / CAF (note 1) | 2009 | 2010 | Split 2010 | ALICE | ATLAS | CMS | LHCb | SUM 2010 |
|---|---|---|---|---|---|---|---|---|
| CPU (HEP-SPEC06) | 131880 | 233400 | Offered | 46800 | 67000 | 96600 | 23000 | 233400 |
| | | | Required | 46800 | 67000 | 96600 | 23000 | 233400 |
| | | | % of Req. | 100% | 100% | 100% | 100% | 100% |
| Disk (Tbytes) | 10065 | 14790 | Offered | 5500 | 3900 | 4100 | 1290 | 14790 |
| | | | Required | 5500 | 3900 | 4100 | 1290 | 14790 |
| | | | % of Req. | 100% | 100% | 100% | 100% | 100% |
| Tape (Tbytes) | 25083 | 31600 | Offered | 6300 | 8900 | 14600 | 1800 | 31600 |
| | | | Required | 6300 | 8900 | 14600 | 1800 | 31600 |
| | | | % of Req. | 100% | 100% | 100% | 100% | 100% |
| Nominal WAN (Mbits/sec) | 140000 | 160000 | | | | | | |

| Summary Ext. Tier1s | 2009 | 2010 | Split 2010 | ALICE | ATLAS | CMS | LHCb | SUM 2010 |
|---|---|---|---|---|---|---|---|---|
| CPU (HEP-SPEC06) | 245800 | 414135 | Offered | 46458 | 217504 | 105505 | 44668 | 414135 |
| | | | Required | 57600 | 192000 | 100500 | 44000 | 394100 |
| | | | Balance | -19% | 13% | 5% | 2% | 5% |
| Disk (Tbytes) | 34890 | 44544 | Offered | 6368 | 22252 | 12510 | 3414 | 44544 |
| | | | Required | 10800 | 21900 | 13400 | 3290 | 49390 |
| | | | Balance | -41% | 2% | -7% | 4% | -10% |
| Tape (Tbytes) | 40189 | 51560 | Offered | 8720 | 15549 | 24138 | 3153 | 51560 |
| | | | Required | 16300 | 14200 | 23300 | 2400 | 56200 |
| | | | Balance | -47% | 10% | 4% | 31% | -8% |

| Summary Tier2s with Split in 2010 | 2009 | 2010 | Split 2010 | ALICE | ATLAS | CMS | LHCb | SUM 2010 |
|---|---|---|---|---|---|---|---|---|
| CPU (HEP-SPEC06) | 308524 | 511829 | Offered | 45669 | 207005 | 187678 | 41476 | 511828 |
| | | | Required | 89600 | 240000 | 195000 | 38000 | 562600 |
| | | | Balance | -49% | -14% | -4% | 9% | -9% |
| Disk (Tbytes) | 22647 | 39353 | Offered | 3431 | 20055 | 12753 | 315 | 39354 |
| | | | Required | 12600 | 24800 | 9200 | 20 | 46620 |
| | | | Balance | -73% | -19% | 39% | 1474% | -16% |
| Tape (Tbytes) | | | | | | | | |

# Installation status – Tier 1s

- Canada:
  - 2009 + 2010 bought together;
  - CPU installed and expected in production in 2 weeks;
  - Disk 850 TB installed, rest coming in Jan
- FZK:
  - 2009 pledges are fully installed; 2010 in place by June
- IN2P3:
  - 2009 CPU installed and deployed; disk all installed and partially allocated;
  - 2010: tendering for disk; expected mid-end January; CPU will order in Q110
- INFN:
  - No tender for 2009 as announced at previous RRB – stay with 2008 capacity
  - Tenders for 2010 ongoing. All 2010 CPU in place in Q110; storage part in Q210, fully available by June
- NL:
  - All 2009 pledges delivered:
  - CPU: 2009 pledge available now, 2010 in 2 weeks
  - Disk: 1st part – next week (meets 2009), rest in January (meets 2010)

# Installation status – cont

- **NDGF:**
  - 2009: all CPU deployed, not all disks online
  - 2010: Still missing formal pledge commitment from Norway (other countries OK). Expect all available by June
- **PIC:**
  - CPU: 2009 pledge available; 2010 tendering: available early 2010
  - Disk: Full 2009 + 2010 pledges should be available within 2 weeks
- **ASGC:**
  - 2009 tenders delayed – mid-Dec.  Storage end of Dec,  For 2010 will follow schedules (June)
- **UK:**
  - 2009: CPU in place; Disk: 50% of 2009 purchases OK, rest have h/w issues – should be available early 2010
  - 2010: CPU and disk tenders ongoing – early 2010
- **BNL:**
  - 2009+2010 pledges installed according to schedule agreed with ATLAS
- **FNAL:**
  - 2009 pledges all in place; 2010: ~1/3 already procured.  Assuming FY10 budget is OK will have rest in place by June.
- **CERN:**
  - 2009 pledge delivered; h/w problems with disk servers (PS/motherboard), but now on site for evaluation: hopefully avail by mid-Dec.
  - 2010 expect to have commissioned by March.

# EGEE→EGI

- **EGI organisation set up is ongoing:**
  - EGI.eu will be hosted in Amsterdam
  - ~30 NGIs have signed MoU and will pay fees
  - EGI council has formed and met

- **Will be supported and partially funded by several projects proposed under various calls under EU FP7**
  - EGI (project named EGI-INSPIRE)
    - Including funding for "Existing multi-national heavy user communities"
  - Several Specialised Support Centres (SSC) including one for HEP
  - EMI (middleware)

| Service | EGEE provider | EGI task | EGI provider (or for WLCG) |
|---|---|---|---|
| Grid Topology - GOCDB | STFC | O-E-1 O-N-1 | STFC |
| Accounting repository – APEL | STFC + CESGA | O-E-2 O-N-2 | STFC+CESGA |
| Monitoring data repositories – SAM etc | CERN | O-E-3 O-N-3 | CERN bid |
| Operations portal –CIC portal | IN2P3 | O-E-4 O-N-4 | IN2P3 |
| Ops oversight – OCC, COD | CERN, IN2P3 | O-E-5 O-N-5 | Coord NL+PL and Tier 1s |
| Gstat | ASGC | | ASGC |
| Nagios + sensors | CERN, SRCE | | CERN ?? |
| Messaging | CERN | O-E-3 O-E-17 O-N-3 | CERN ++ |
| Dashboards | CERN | | CERN (SA4) |
| Regional ops dashboard | IN2P3 | | IN2P3 |

| Service | EGEE provider | EGI task | EGI provider (or for WLCG) |
|---|---|---|---|
| Ticketing system – GGUS | FZK | O-E-6 O-E-8 | FZK |
| Ticket triage etc – TPM | ROCs | O-N-6 O-N-7 | All NGIs |
| Middleware deployment coord | CERN | O-E-9 | ES + P |
| Interoperation coord | CERN | O-E-11 | SE |
| Network coord - ENOC | IN2P3 | O-E-12 | IT |
| Ops procedures | CERN | O-E-13 | FI |
| Policy development – JSPG | STFC | O-E-15 | UK + NL |
| Ops security coord | CERN | O-E-16 | UK +NL |
| Coord & maint of ops tools | CERN | O-E-17 | ?? |
| Apps support – EIS | CERN/INFN | ➔ SSC | HEP SSC |

**Services required for WLCG**

| Component | Developer/Maint | Component | Developer/Maint |
|---|---|---|---|
| **Data Management** | | **Operations Tools** | |
| FTS | CERN | APEL | STFC |
| DPM | CERN | Accounting portal | CESGA |
| Castor | CERN | GOCDB | STFC |
| dCache | DESY/FNAL/NDGF | SAM | CERN |
| GFAL/lcg-utils | CERN | GridView | CERN/India |
| LFC | CERN | GridMap | CERN |
| Storm | INFN | Dashboards | CERN |
| **Workload Management** | | Nagios sensors | CERN + SRCE +? |
| WMS | INFN, ElsagDatamat | MSG | CERN |
| LB | Czech Rep. | Gstat | ASGC |
| CREAM/BLAH | INFN | CIC Portal | IN2P3 |
| LCG-CE | CERN | DGAS | INFN |
| VOBox container | CERN | | |
| **AAA** | | | |
| VOMS | INFN | | |
| VOMRS | FNAL | | |
| MyProxy | VDT | | |
| Proxy renewal | CESNET | | |
| LCAS/LCMAPS/SCAS | Nikhef | | |
| gLexec | Nikhef | | |
| Delegation framework | CERN, HIP, STFC | | |
| Trustmanager | HIP | | |
| GridSite | STFC | | |
| **General** | | | |
| Information system | CERN | | |
| YAIM framework | CERN | | |

✓ In EMI

✓ In EGI

✓ In SA4

**Middleware required for WLCG**

# Conclusions

- STEP'09 carried out as planned – shows that we are ready to take data
    - Results from reprocessing and analysis tests are encouraging
    - Effort for operations seems sustainable, although some concern during holiday period

- Business as usual since STEP'09 ... With continuing heavy usage of resources
    - Operational issues now dominate

- EGEE → EGI transition – situation is encouraging, but must be aware of the potential disruption
    - Overall level of benefit to WLCG not entirely clear...