# CHIPP – CSCS F2F Meeting

Zürich, August 19th 2014

Sadaf Alam
George Brown
Miguel Gila
Gianni Ricciardi

# Agenda

- 9:45 - Coffee, presentation and agenda
- 10:15 - Tier-2 status and plans
  - CSCS (40')
  - UNIBE-LHEP (20')
- 11:15 - Tier-3 status and plans:
  - PSI (10')
  - UNIBE-ID (10')
  - UNIGE (10')
- 12:00 - Lunch break
- 13:15 - ATLAS on Todi/Piz Daint (Michi 10 min)
- 13:25 - Open discussion
- 13:45 - Next meeting date and place
- 14:00 - A.O.B.
- 15:00 - Coffee Break
- 15:30 - NGI_CH
  - Swiss Argus national service
  - Swiss Nagios national service
  - EGI request: SAM Nagios Probes refactoring TF
  - Update on the Swiss EGI participation (Sigve 10 min)
  - A.O.B.
- 16:30 - End of the meeting

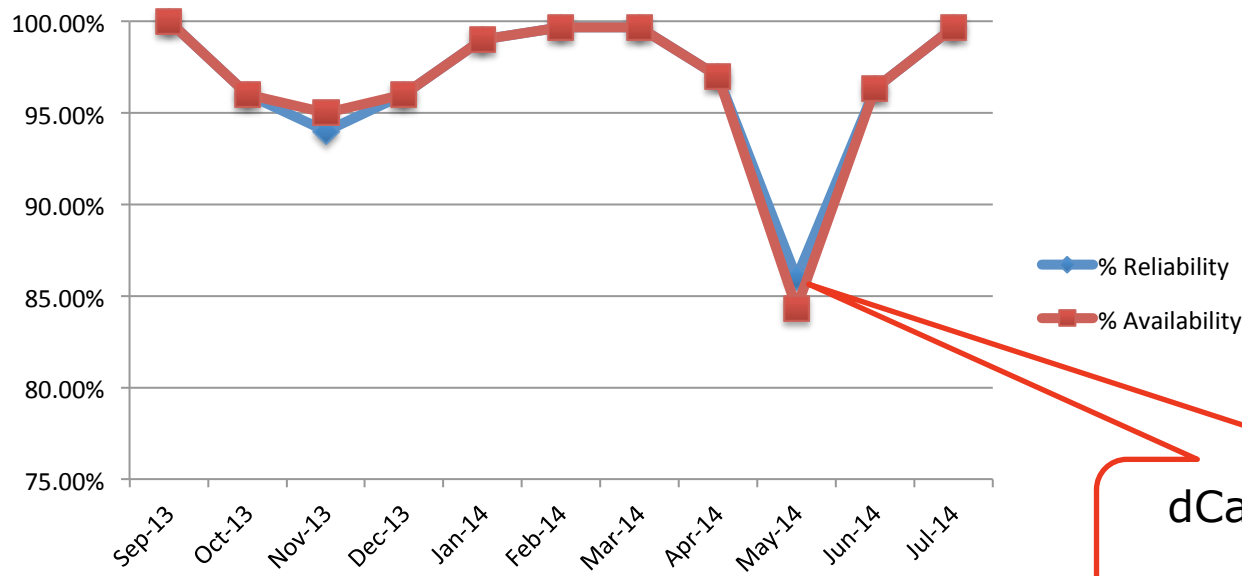# Tier 2 status and plans CSCS

# Status

# Statistics – Availability & Reliability

- **Relatively stable operation** with small hiccups:
    - GPFS: inode usage above threshold and IB cable broken
    - Middleware issues (workaround in place)
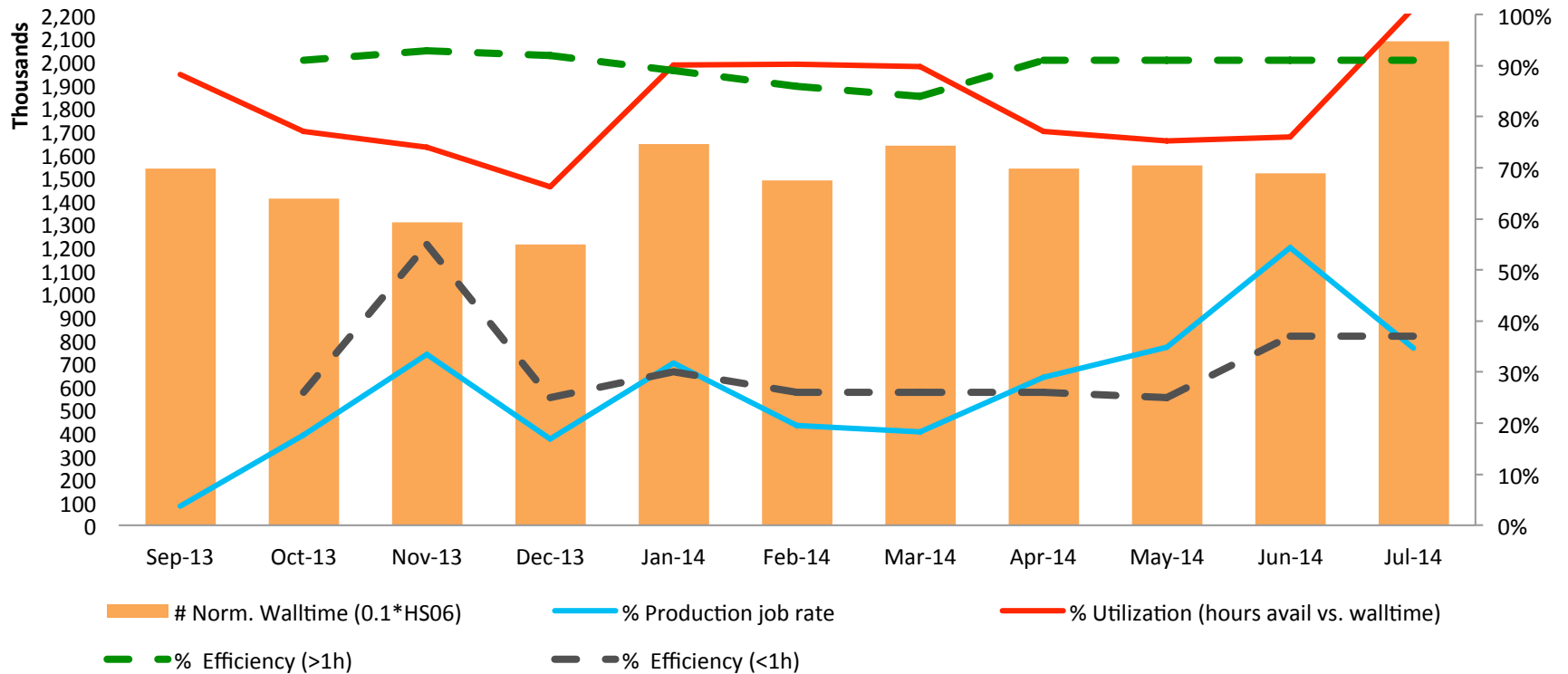


**Phoenix A/R since Sep-13**

dCache infosys errors
+
GPFS problems (inodes)

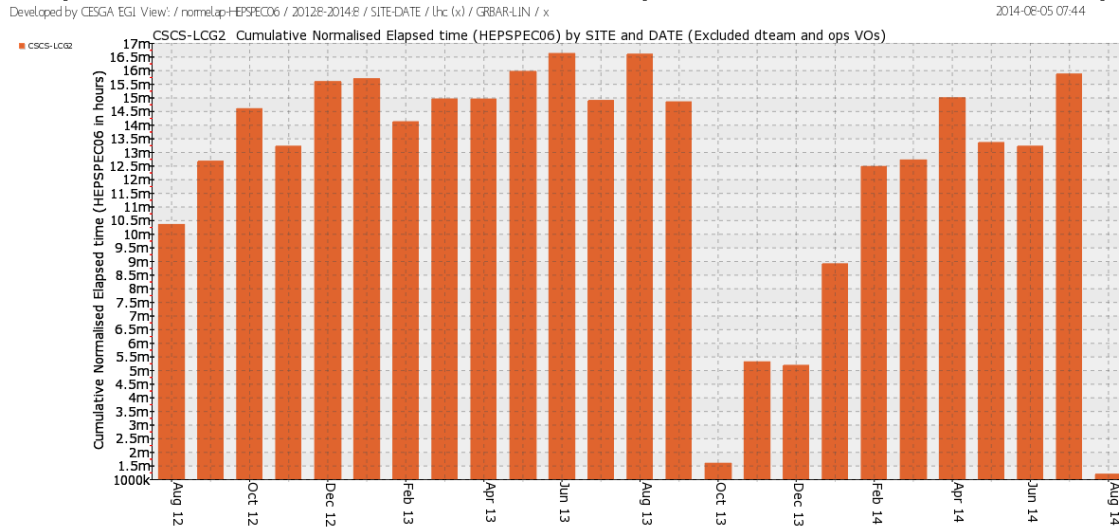# Statistics – CPU Usage

- **CPU usage increased** (specially during July)

**Phoenix usage since Sep-13**



Legend:
- # Norm. Walltime (0.1*HS06)
- % Production job rate
- % Utilization (hours avail vs. walltime)
- % Efficiency (>1h)
- % Efficiency (<1h)

# Statistics – CPU Usage from EGI perspective

- Computation hours restored to previous values over past months:



Normalized
Elapsed CPU
(HS06-hours)

Includes also failed
and cancelled jobs.

- There is still a mismatch between local accounting and EGI reports:

| | Jan 2014 | Feb 2014 | Mar 2014 | Jun 2014 | May 2014 | June 2014 | July 2014 |
|------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| SLURM | 16.529.550 | 14.945.640 | 16.461.540 | 15.453.370 | 15.594.660 | 15.240.410 | 20.964.960 |
| EGI | 8.918.772 | 12.487.696 | 12.726.256 | 15.008.996 | 13.367.080 | 13.232.896 | 15.880.340 |

Unclear as to what it includes.
TaskForce in place to identify possible
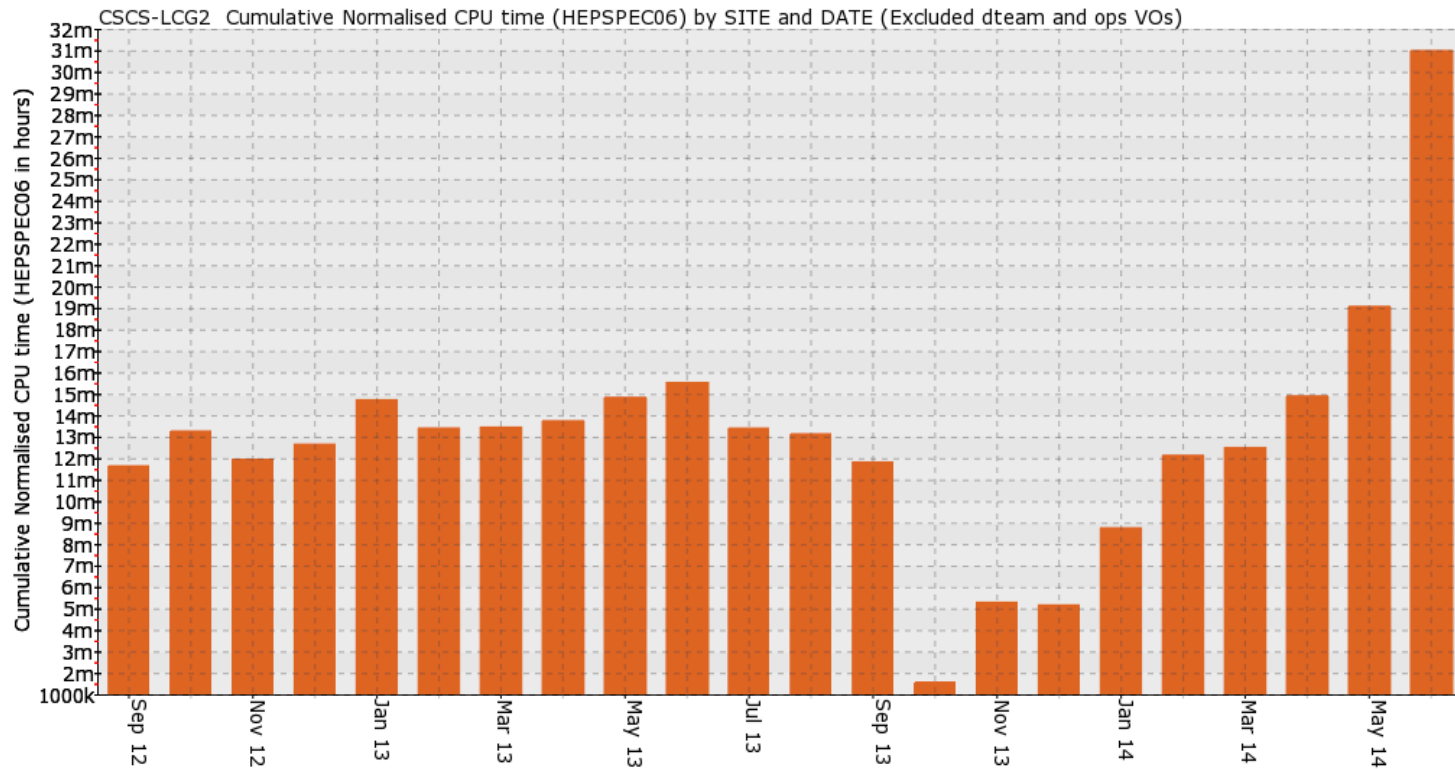improvements.

# Statistics – CPU Usage from EGI perspective (extra)

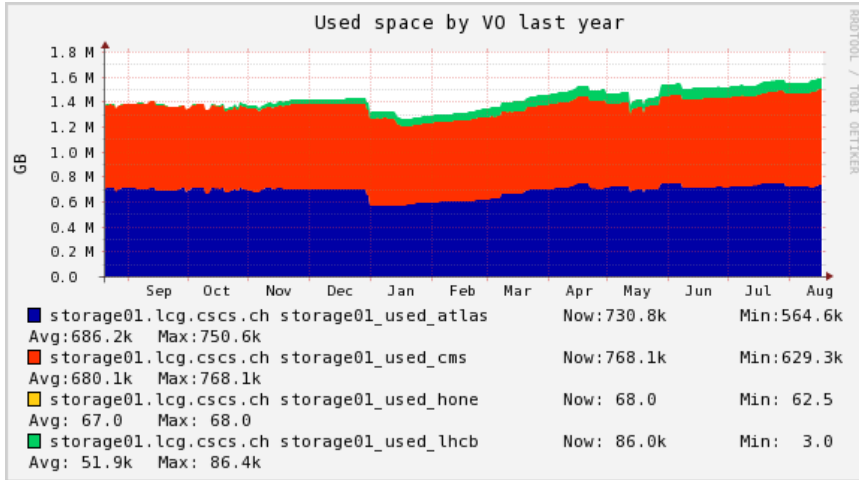- Total computation hours (HS06) (SUM)

# Statistics – Storage usage



Used space by VO last year

- ■ storage01.lcg.cscs.ch storage01_used_atlas  Now:730.8k  Min:564.6k
  Avg:686.2k  Max:750.6k
- ■ storage01.lcg.cscs.ch storage01_used_cms  Now:768.1k  Min:629.3k
  Avg:680.1k  Max:768.1k
- ■ storage01.lcg.cscs.ch storage01_used_hone  Now: 68.0  Min: 62.5
  Avg: 67.0  Max: 68.0
- ■ storage01.lcg.cscs.ch storage01_used_lhcb  Now: 86.0k  Min: 3.0
  Avg: 51.9k  Max: 86.4k

|       | TiB used | TiB free | TiB total |
|-------|----------|----------|-----------|
| atlas | 730.8    | 116.9    | 847.7     |
| cms   | 768.1    | 63.8     | 831.9     |
| lhcb  | 86       | 132.9    | 218.9     |
|       | **1584.9** | **313.6** | **1898.5** |
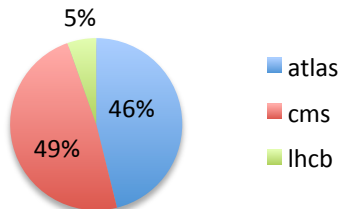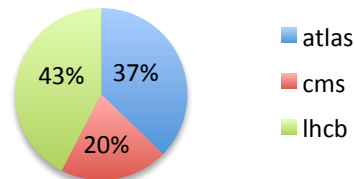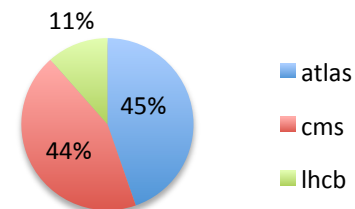
## Storage distribution CSCS-LCG2 Aug 18 2014



## TiB used



## TiB free



## TiB total

# Statistics – Swiss resources

- **Compute (8x nodes, ~3.5k HS06)**
  - Deployed and fully operational.
  - Pending to increase priority of Swiss users.

- **Storage (~375TB)**
  - Not in production.
  - HW issues have been preventing this hardware to be deployed.
  - Latest testing promising.
  - We won't put hardware in production until we are 100% sure it won't be problematic.

# Operations

# Operations

- **Deployed Virtualization servers, evaluating oVirt and/or RHEV**

- **Reserved 4 full nodes for ATLAS, 4 for CMS and 2 for LHCb**
    - A manual fix to fairshare and backfilling issues.

- **Not deployed ~375TB of Storage for Swiss users due to HW issues**

- **Deployed GPFS2 on testing environment, planning to do gradually with at least 30% WNs before the final switch over.**

- **Discovered faulty IB cable that made GPFS unstable**
    - seen tremendous performance and stability improvement since the replacement
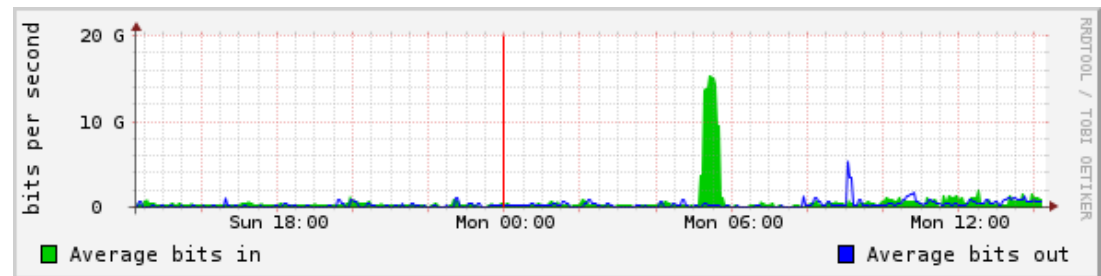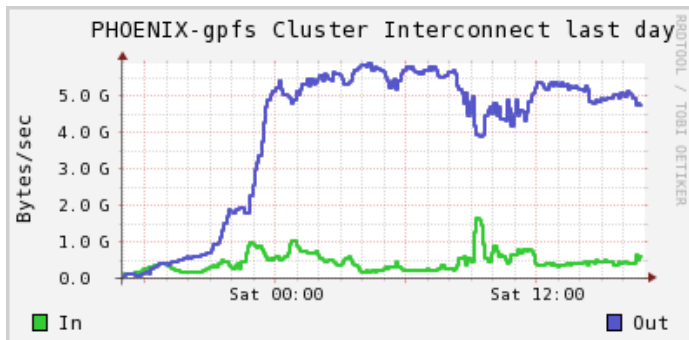
# Operations

**Next maintenance on Sept 3 or 17. Significant changes:**

- Swiss users mapping and priority
  - need to define specific mappings for CMSch and change account priorities. → reconfigure all grid services! (YAIM)

- Swiss ARGUS National service to be tested internally
  - Need to replace current setup (on CREAMs) for a set of 3 dedicated VMs on a HA configuration, similar to Site-BDII.

- GPFS2 deployment in production
  - This is the most critical change, as it requires changing the configuration on all nodes.

- CSCS general uplink 100Gbps, Phoenix uplink still 20Gbps (2x 10Gbps)
  - Upgrade to 40Gbps possible. Next year?

# Operations

**GPFS2** ([https://wiki.chipp.ch/twiki/bin/view/LCGTier2/ServiceGPFS2](https://wiki.chipp.ch/twiki/bin/view/LCGTier2/ServiceGPFS2))

- First production service fully configured using Puppet
- Observed performance results above initial expectations:



**x2 controllers = ~13GB/s**
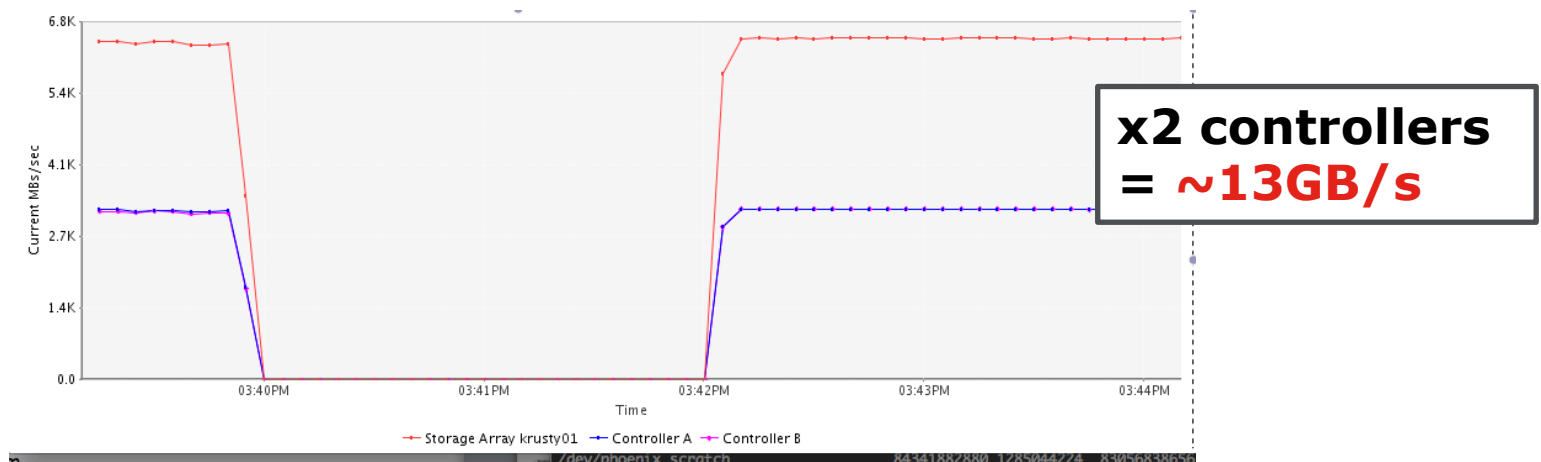
- If more storage or performance is required, we can increase it just by adding hardware.
  - Possible to increase of ~40% in bandwidth
  - Double capacity, from 79T to 150T

# Plans
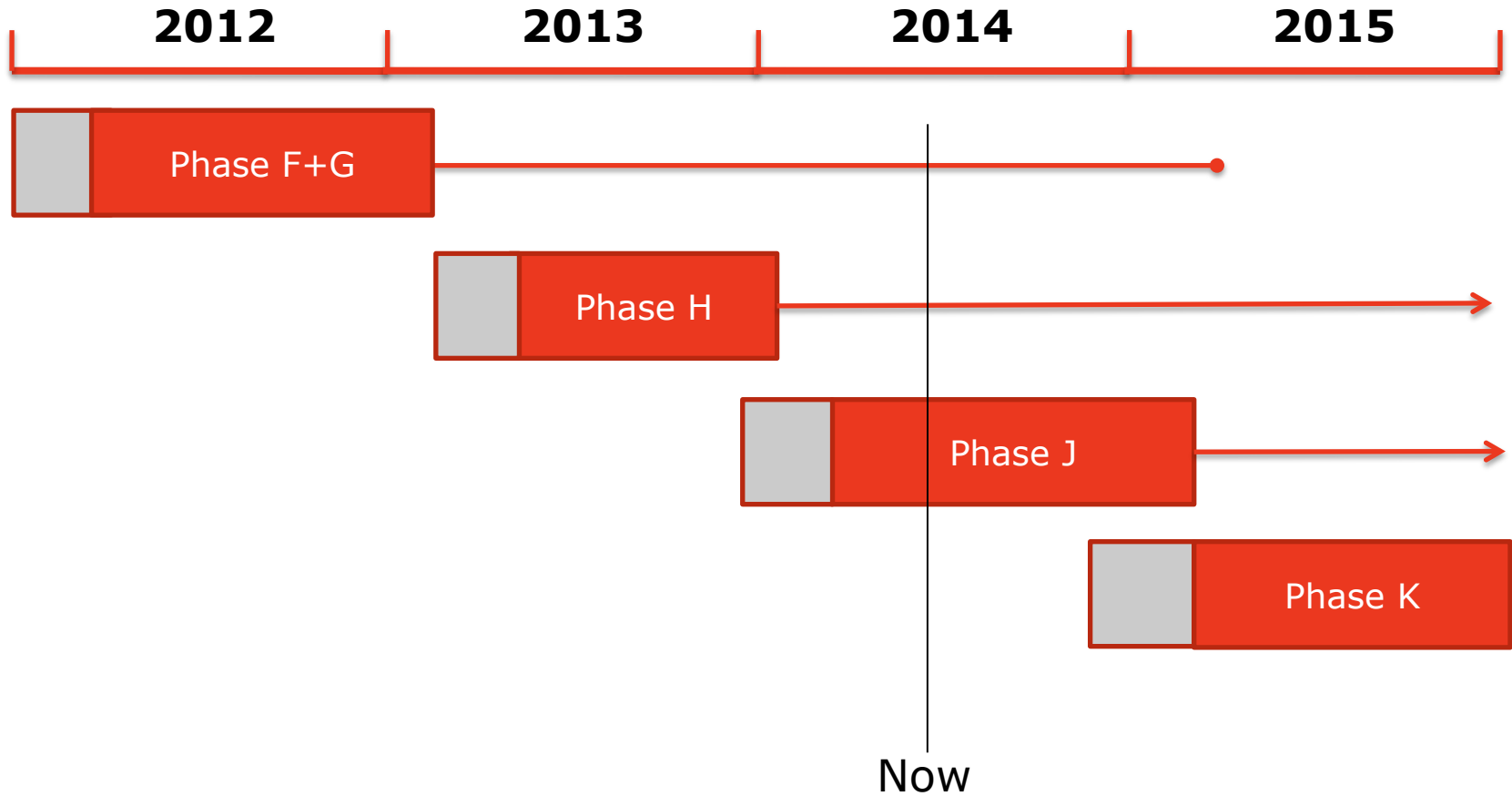
# Phases of Phoenix



| 2012 | 2013 | 2014 | 2015 |

Phase F+G

Phase H

Phase J

Phase K

Now

# Pledges

> Great increase in compute power. Sandy Bridge systems to be run until the end of their life

| Phase | Compute power actual/pledged [HS06] | Storage actual/pledged [TB] | Scratch actual/desired [GB/s] |
|-------|-------------------------------------|------------------------------|--------------------------------|
| Phase H - 2013 | 28826/26000 | 1919/1800 | 9/9 |
| Phase J - 2014 | 39000/*35000* | 2100/*2000* | 9/9 |
| Phase K - 2015 | 44000/*39000* | 2600/*2200* | 9/13 |

### Compute

| | Phase H HS | Phase J HS | Phase K HS |
|------|------|------|------|
| ATLAS | 10'400 | 14'000 | 15'600 |
| CMS | 10'400 | 14'000 | 15'600 |
| LHCb | 5'200 | 7'000 | 7'800 |

### Storage

| | Phase H TiB | Phase J TiB | Phase K TiB |
|------|------|------|------|
| ATLAS | 792 | 875 | 955 |
| CMS | 792 | 875 | 955 |
| LHCb | 216 | 250 | 290 |

### Compute/Storage

| | Phase H HS/TB | Phase J HS/TB | Phase K HS/TB |
|------|------|------|------|
| ATLAS | 13.13 | 16.00 | 16.34 |
| CMS | 13.13 | 16.00 | 16.34 |
| LHCb | 24.07 | 28.00 | 26.90 |

# Decommissions & purchases

**Purchases**
- Storage for a total of 720TiB
  - intended to replace 3x half-racks of IBM DC3500



- 20x compute nodes (~8.5kHS06)
  - Not decided if Ivy Bridge or Haswell, initial tests on the latter not very promising



- 2x service nodes



- 2x Mellanox FDR IB switches



- 2x CISCO Ethernet switches



**Decommissions:**
- 3x half-racks of IBM DC3500 storage, 60 disks each
- Old GPFS (2 full racks of JBODs)
- Old CFengine servers? ☺

`09:39:44 up 832 days, 22:25,  1 user,  load average: 0.00, 0.00, 0.00`

# Thank you for your attention

# Extra slides

# NetApp problems (Swiss users storage)

- Initial tests ran on the storage were successful.

- When the system was put in production **under heavy I/O, paths would drop** randomly without any particular reason.

- Spent **3 full weeks** working with **NetApp support**, changing both server and storage controller configuration and HW.

- **The result was unsuccessful, until controller firmware was replaced.** So far, it's working OK under heavy load, but need to test for a longer period with different IO patterns to make sure the problem is fixed and not masked.

- **Lessons learned:**
    - NetApp support is US-based, need to review support policies with them, as this makes solving issues difficult.

    - When a service is provided using HW from different vendors, they can start the **blame war** and CSCS/CHIPP are left in the middle. → **this needs to change**

    - We (CSCS) need **to improve our acceptance criteria** before deploying new production services.

# GPFS issues

- **Metadata inode exhaustion**
  - Due to several identified problems, inodes were exhausted on metadata servers.
  - This caused the whole cluster to limp while the policies were enforced (= massive removal of ~130million files).
  - ➢ Policies moved to several **independent** nodes running at different times.
  - ➢ Better monitoring (Ganglia) deployed.

- **Broken Infiniband cable**
  - WNs were dropping off GPFS when under heavy IO.
  - Compiled and run *netperf* between all GPFS servers.
  - One GPFS server was providing **10x times less** bandwidth than expected!
  - WNs dropping off because the GPFS server would think it could not reach clients.
  - ➢ Working on running *netperf* automatically via Nagios

# dCache issues

- **Information system not properly handling this**

- **dCache did provided an official fix for this on release 2.6.31.**
  - We run 2.6.27
  - Applying this required downtime AND fix was only available after we deployed our 'hack'.

- **Deployed the following workaround until dCache gets upgraded.**

```
# MG 16/06/14 -- hack to solve
#   https://xgus.ggus.eu/ngi_ch/?mode=ticket_info&ticket_id=327
#   until dCache support fixes it

/usr/sbin/dcache-info-provider |sed /'dn:
GlueSEAccessProtocolLocalID=NFSv41-storage02@nfs-storage02Domain/,/^$/{d}'
```

# Swiss National Argus service

- **3 KVM VMs on 3 different KVM hosts.**
- **Load-balanced with a common DNS alias: argus.lcg.cscs.ch**
  - Similar to current CSCS site-bdii configuration
  - Need to test this before opening the service to Swiss community