





WLCG Service Report

Jamie.Shiers@cern.ch

~ ~ ~

WLCG Management Board, 16th June 2009

STEP'09 Service Summary

| | |
|--|--|
| <p>Glass half full</p>  | <p>A lot of activity in particular from ATLAS and CMS at Tier0, Tier1s and Tier2s. Results still to be analyzed (see post-mortem workshop in July: but indications encouraging. All this despite significant problems with many sites & services leading up to and during the last two weeks; A “step-up” in WLCG operations is essential if we are to sustain a similar service level over a longer period of time;</p> |
| <p>Glass half empty</p>  | <p>Despite encouraging results, many sites / services suffered serious outages or degradations during or leading up to the main two weeks of STEP'09. (Almost) everything that could go wrong did: OPN fibre cuts, major data and storage management issues, data loss, release back-out etc. Many of the rather soft WLCG operations procedures are repeatedly ignored: this is a major risk for the overall WLCG service if not addressed with priority now.</p> |

STEP'09 Site Summary

- More than ½ of ATLAS Tier1s passed the (ambitious) reprocessing metric: of those that did not, biggest concerns are (high to low):
 - **ASGC [high concern]**
 - **FZK**
 - NL-T1 (some DMF-level tuning...)
 - IN2P3: “extremely impressive” was comment from last Thursday’s minutes [low concern]
- These results must be matched against expectations: we know that problems will occur and the important thing is

ATLAS reprocessing – final numbers:

https://twiki.cern.ch/twiki/bin/view/Atlas/Step09Logbook#Saturday_13_June

PIC, NDGF, CNAF, RAL, TRIUMF, BNL all passed - in fact they all passed **with flying colours, beating the metric by x2 or more.**

LYON, SARA, FZK and **ASGC** missed the target, achieving 97%, 96%, 90% and **48%** respectively of the metric.

STEP'09 – A Real Step Up?

- ☺ **Despite the very real problems affecting sites & services, all in all STEP'09 represents a real step forward in what has been achieved and a real step up in service level**
- **We must analyse the problems transparently and thoroughly: issues of the level that continue to be seen deserve and require this!**
- **And we must seek ways to further reduce effort – could we have continued at this level for many weeks / months on end? (Particular expt OPS)**
- ☹ **Attendance at operations meeting dropped off sharply after GDB & ATLAS announcement of run-down, so most likely no...**

STEP'09: Site / Service Issues

| Site / Service | Issue |
|-------------------------|---|
| ASGC – many issues | Relocation away from temporary location foreseen in coming weeks. Need to retest once this is done; an analysis of the key problems and their resolution would be valuable. |
| FZK tape access | See STEP'09 site metric. |
| LHCb & conditions DB | CORAL/LFC – major performance problems. A “SIR” has been requested. |
| IN2P3 gridftp transfers | Transfers aborting: 03:00 – 10:30 Wednesday. A “ <u>SIR</u> ” has been received. [Alarm ticket] |
| OPN fibre cuts | CERN IT status board contained stale information (24H?) after repairs had been made. No broadcast made. [WLCG operations procedures rediscussed just days prior to start of STEP'09] |

IN2P3 – GridFTP outage

- June 10 03:00 – 10:30: SRM file transfers failed
 - 03:30 SAM tests of the SRM service start failing
 - 08:30 dCache master notified and problem investigation
 - 09:50 LHC alarm ticket received (submitted 07:40 UTC)
 - 10:00 Attempt to publish downtime notification
 - 10:30 SRM service restarted & up
- Problem occurred when GridFTP servers were overloaded. Bug in monitoring script disabled all GridFTP doors(!) SAM tests -> dCache monitoring and not to local monitoring hence no callout
- Script fixed to not disabled all doors; mid-term improvements to monitoring
- Ticket about GOCDB bug: 49383

ASGC Briefing

- problematic LTO4 drive fixed this Wed., after re-calibrating from console of the library. all new procured tape drives are full functional now. majority of the activity refer to read access and from cms, try recalling data from cmsCSAtp pool.
- around 500+ job pending in queue since yesterday, both for atlas and cms. the problem is clear earlier as two of the misconfigured wn add to the pool.

Some "STEP-style" re-validation of ASGC once relocated back from IDC clearly required.

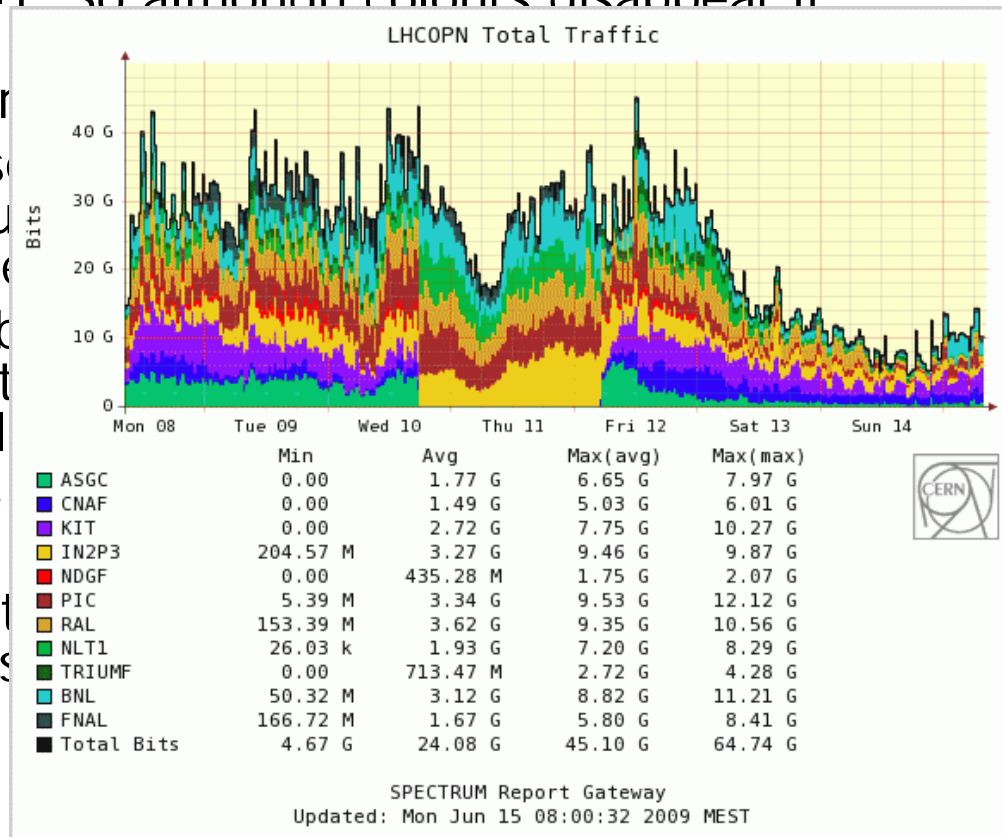
July is perhaps a little tight for this, but should be done prior to any "SEPT" or other multi-site/multi-VO Scale Tests...

mapping in different VLS from 1/1/2 able to balancing the usage.

- **draft planning for the facility relocation 11 days later. the IDC contract end at Jun 23, and we'll have one day downtime for computing/storage facility and two days for tape library. as re-installation of IBM TS3500 linear tape system need calibrating the horizontal between all frame sets.**

OPN Fibre Cuts

- The recent outage, some 36 hours on a major fiber trunk can be seen clearly in the weekly statistics here:
- <http://network-statistics.web.cern.ch/network-statistics/ext/LHCOPN-Total/?p=sc&q=LHCOPN%20Total%20Traffic>
- Note that when traffic is re-routed it is then attributed together with the traffic to another Tier-1. So although colours disappear it does not mean that the Tier-1 can be seen from the total aggregate.
- This was one of the “worse cases” identified that cut 5 Geant circuit connectivity paths to their respective providers.
- However, the use of the cross border backup paths enabled most of the overall traffic to continue with little interruption.
- So congratulations to everyone involved for their level of resiliency!
- If anyone has specific comments on our OPN, they are welcome to send them to me.
- Best Regards, David.



GridPP

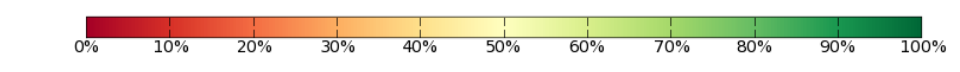
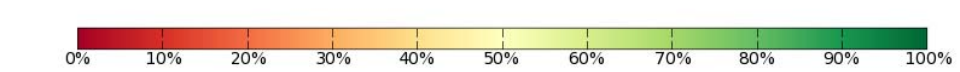
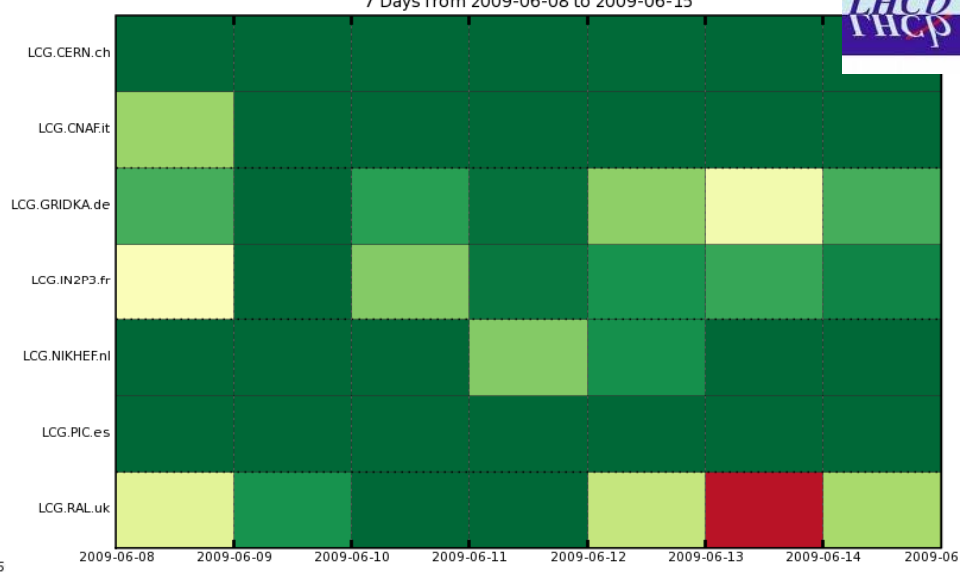
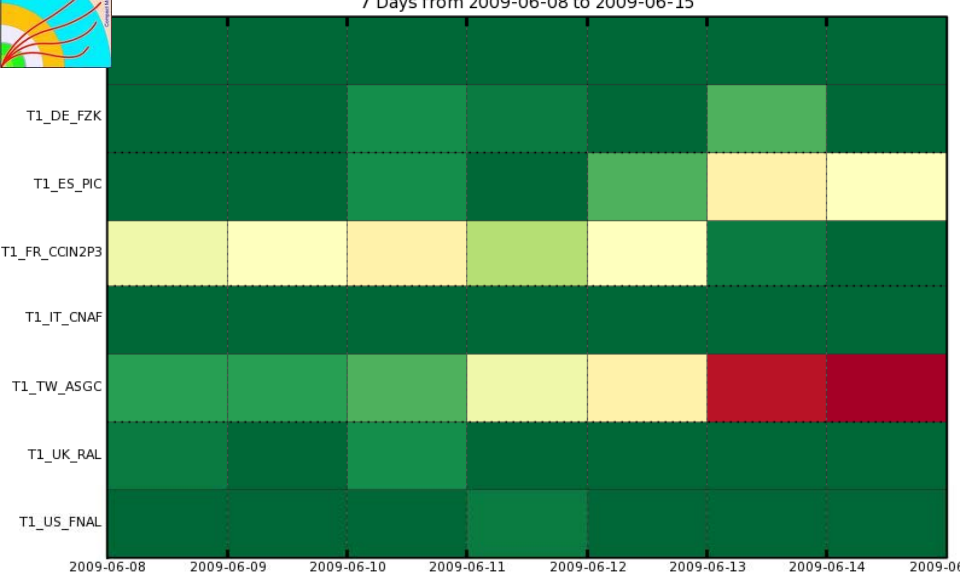
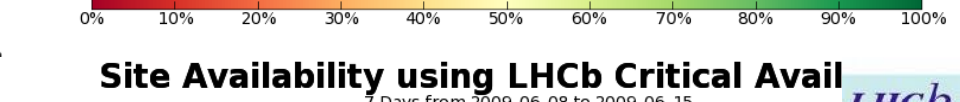
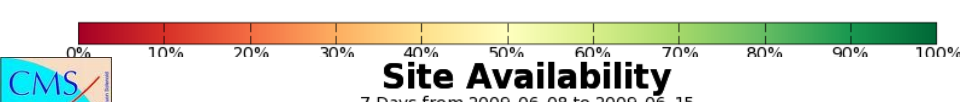
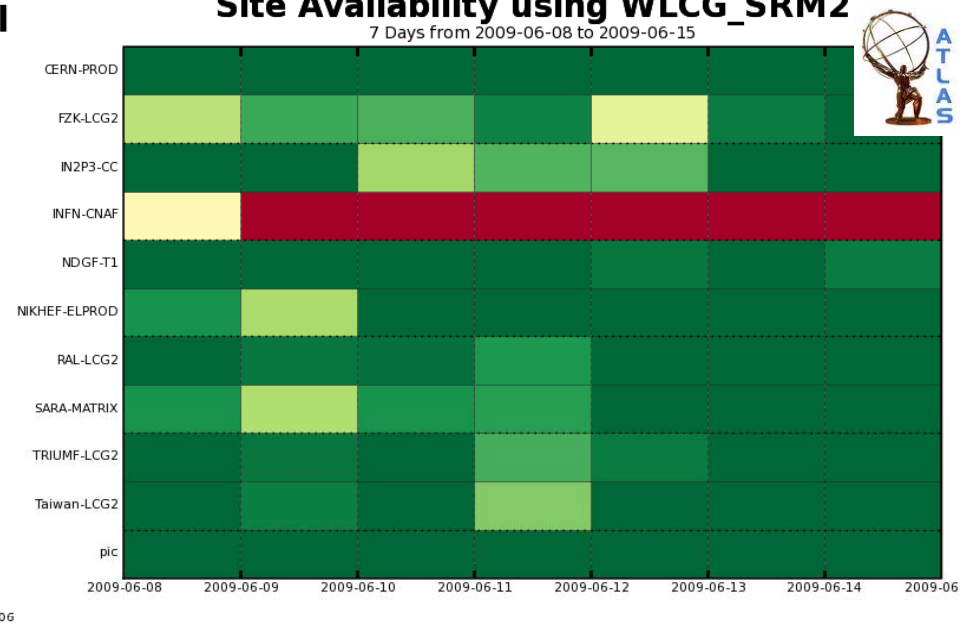
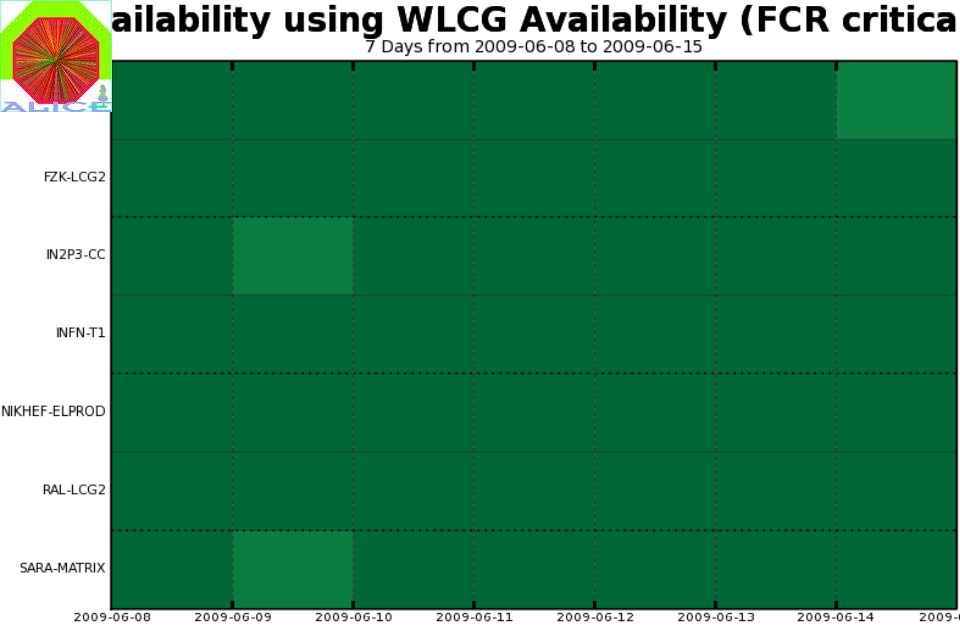
- We were very busy trying to learn from STEP, but in many ways it was a very calm 2 weeks. There was no out-of-hours page out until Saturday just gone (formally after the end of STEP09) and although there WERE problems that gave us a few long days we would hope these kind of issues would shake out pretty fast. On the whole, things look quite comfortable, though perhaps we were just lucky. We did, for example, have a Tier-1 away-day in the second week with just a skeleton crew left on site, so we were certainly not in "hero-mode".
- I would conclude that STEP09 has given the UK confidence that we can sustain a service at that level.
- David Britton

KPIs: GGUS summary & SAM results

- GGUS (“manual” search used for #s below):
 - Only 1 real alarm ticket (ATLAS to IN2P3) sent during this week! (Plus one test from LHCb to FZK)
 - TEAM tickets dominate for ATLAS & LHCb
 - ALICE & CMS have about 5 tickets / week
 - ATLAS (& LHCb) close to 1 order of magnitude higher!
- SAM results (VO tests):
 - Generally speaking, very “green”
 - Hard to match (this week’s) results with WLCG OPS minutes: no clear correlation with site problems described above!



| VO | User | Team | Alarm | Total |
|--------|------|------|-------|-------|
| ALICE | 6 | 0 | 0 | 6 |
| ATLAS | 5 | 32 | 1 | 38 |
| CMS | 4 | 0 | 0 | 4 |
| LHCb | 0 | 26 | 1 | 27 |
| Totals | 33 | 32 | 0 | 75 |



Summary

- ‡ A lot of numbers to digest – some positive results but some less positive
- 💣 Apart from the technical lessons of STEP'09, it is clear that we could not continue (experiments, sites, services) at this level of operational stress for prolonged periods;
- We need to address this issue along with the technical problems: “hero mode” will not be sufficient!
- On the other hand, this is not likely to be a 9-5 job for anyone in a key role anytime soon: is 12h week days and 8h (less?) weekends achievable in the coming months?

Post-Mortem Workshop

- Some 40 people have registered – would be good to have better site-representation
- EVO will be available
- By end of last week only 1 person other than myself had signed up for BBQ: there is a [Twiki page](#) for ensuring not everyone comes with (or without) the same things, from which a [Doodle](#) is linked for those who wish to keep it a surprise **(now there are 2!)**
- [People signing up slowly...]