WLCG Service Report

Dirk.Duellmann@cern.ch
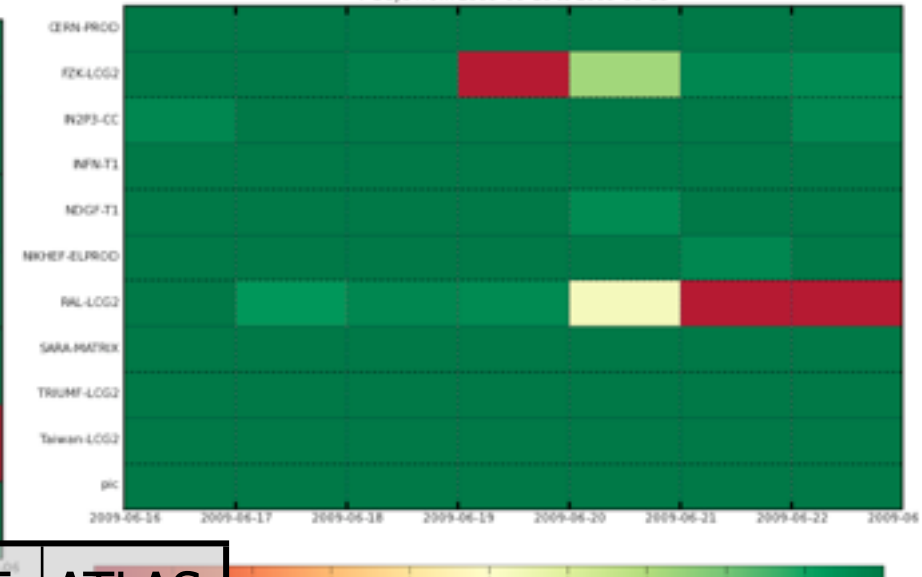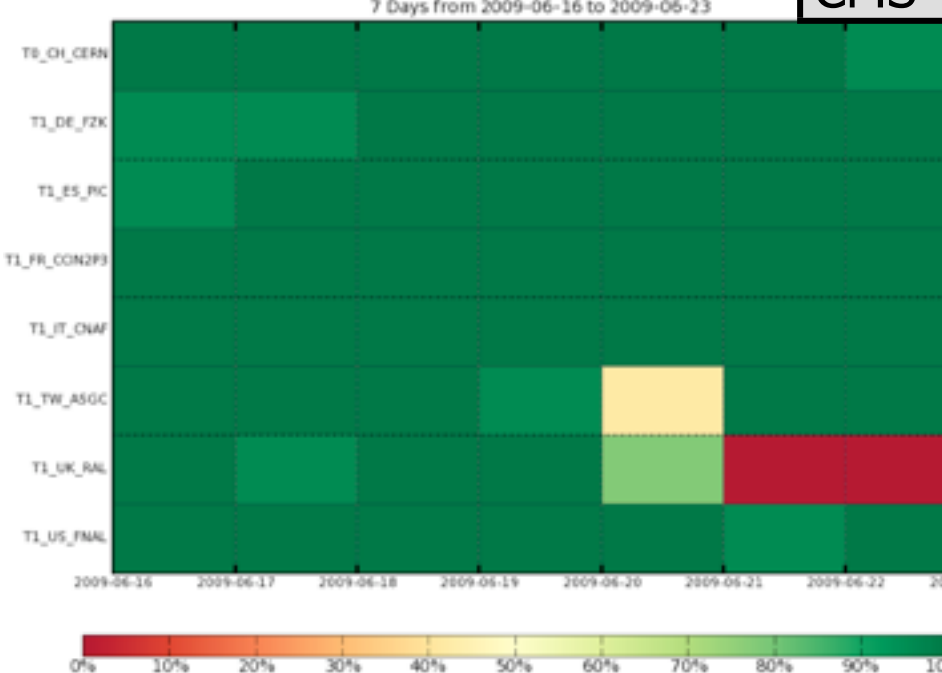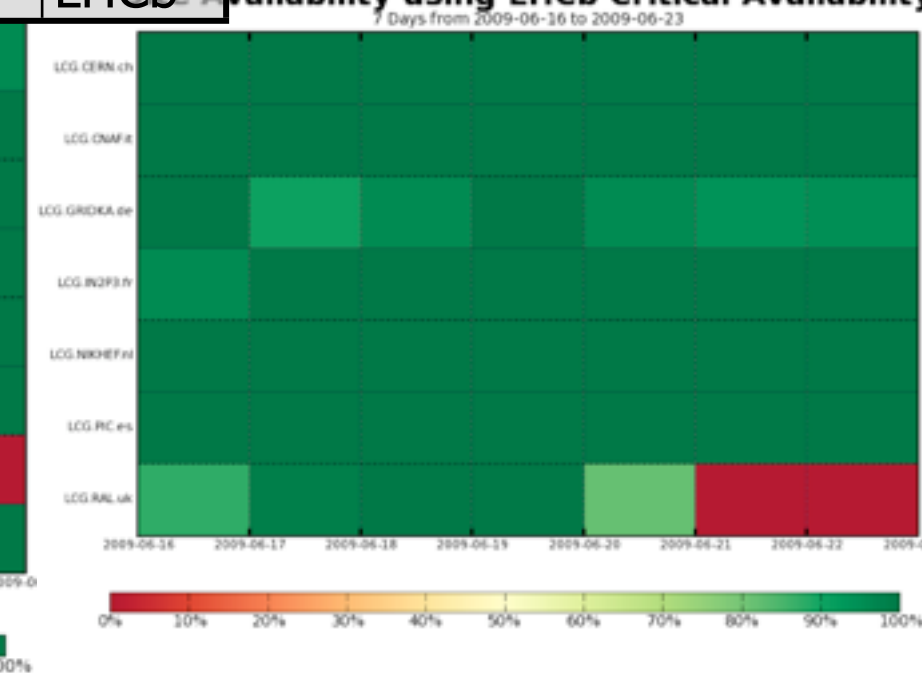
~~~

WLCG Management Board, 23rd June 2009

Introduction

- This report covers the service since last week's MB

- Ramping down after main STEP'09 activity
  - Service incident report for
    - CERN power cut
    - CORAL/LFC issues affecting LHCb

- Attendance in ops meeting ramping down too :-(
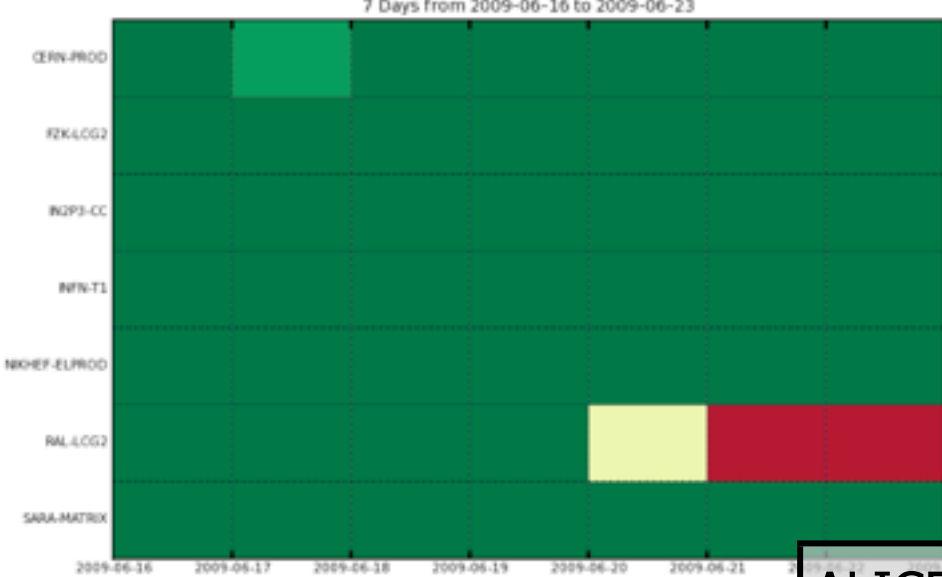
- Back to smooth production running...

2

GGUS Summaries – last week

VO concerned	USER	TEAM	ALARM	TOTAL
ALICE	3	0	1	4
ATLAS	16	22	**0**	38
CMS	1	0	0	1
LHCb	0	20	**0**	20
Totals	20	42	**1**	63

| ALICE | ATLAS |
| CMS | LHCb |

Tuesday, 23 June 2009

CORAL/LFC degradation (LHCb)

Date: June 11
Description: Scalability issues with CORAL credential look up from LFC
Impact: LHCb LFC degraded

Timeline of the Incident

11 June: The R/O instance of LFC (1 node) for LHCb was degraded. Connections were refused because the server was busy servicing over 60 connections and there were no more threads available for new requests.
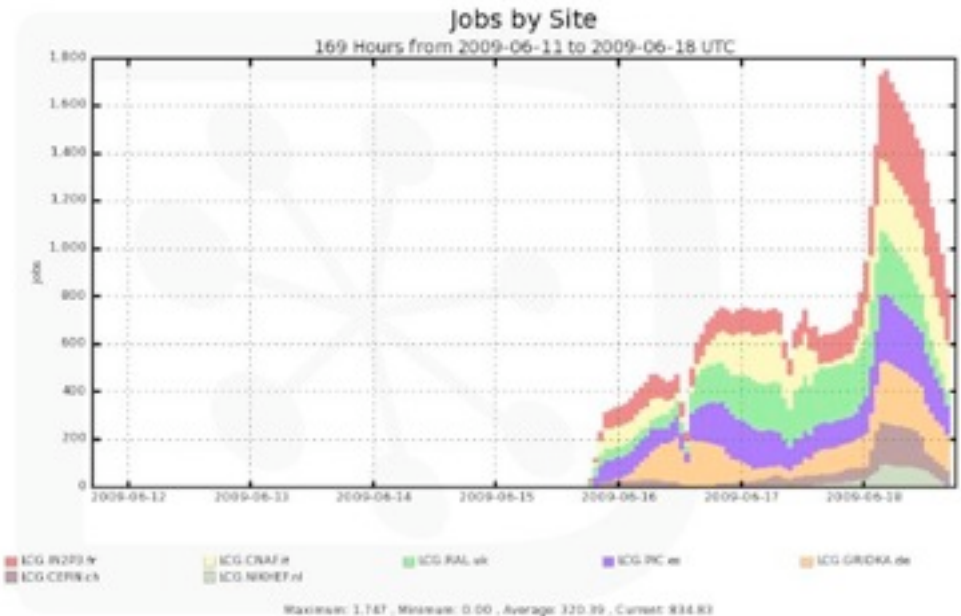12 June: LHCb moved to conditions data access via sqlite files
16 June: LHCb reverted to CORAL access to conditions, but using a different method to obtain DB credentials

- CORAL changes to avoid blocking LFC thread during LFC credential lookup agreed between CORAL and LFC team.
- Development to fix LFC look up from CORAL planned for July.

5

Full report at https://twiki.cern.ch/twiki/bin/view/LCG/WLCGServiceIncidents

Power cut @ CERN

Date: June 18
Description: CERN-wide power cut
Impact: Batch, CASTOR and FTS services interrupted for ~2 hours

Timeline of the Incident

~12:05 – Initial brief CERN-wide power cut. CC services continue on UPS battery power.
~12:40 – Second power cut; CC running with on UPS battery power
~12:45 – Power restored, but batteries for one of the 3 UPS units without diesel backup power are exhausted leading to service outage for parts of the batch, CASTOR and database services. Failure of UPS supply leads to automatic trip of associated switchboard in CC substation.
~13:10 – Substation switchboard re-enabled.
~13:15 – CERN Control Centre confirms power is stable; power-up starts for affected systems
~13:20 – batch scheduling disabled
~13:30 – Service restart commences
~14:00 – FTS database back – FTS service resumes
~14:10 – All LFC nodes have been rebooted – all LFC services probably back then
~14:26 – Load balancing service back a first time after restart by operator (probably stuck due to high number of nodes down)
~14:30 – CASTOR and SRM services restored
~14:40 – batch queues re-opened
~14:40 – Normal service is resumed, myproxy-fts service still affected
~15:45 – Load balancing service back again after being stuck again
~15:49 – myproxy-fts service restored

Preliminary analysis: believed to have been caused initially by a short-circuit followed by a thermal cut out when rerouted supplies caused a cable overload.

6

Full report at https://twiki.cern.ch/twiki/bin/view/LCG/WLCGServiceIncidents

Other Service Issues

- Twiki outage - June 15
  - Affected experiments reports and also daily operations meeting attendance as connection info was temporarily unavailable
    - Alternate copy of this information in ops list archive and at http://indico.cern.ch/conferenceDisplay.py?confId=62132
  - Twiki content recovered to 15 mins before the outage but some twiki attachments of lost (morning of 15 June)

7

Update: Oracle BIGID problem in CASTOR

Patch for Oracle internal consistency problem now available for 64-bit and 32-bit clusters

- Test deployments of the patch ongoing at CERN
  - original test case is fixed

- ASGC (who have been most significantly been affected) are suggested to confirm the patch if possible before applying the castor workaround

Other Experiment News

- CMS tape tests at CERN and RAL (w/o ATLAS)
  - CERN: more than 1GB/s, back-log digested over 12h period
  - RAL: 200 MB/s

- ALICE transfer tests successful at 6 T1s
  - initially some efficiency issues at FZK, but redistribution of data worked well

- Post-STEP deletion (ATLAS: ~5PB) ongoing

- All experiments declared end-of-STEP
  - analysing STEP results and preparing for STEP post-mortem workshop

- Returning to standard production activities

Other Site News

- RAL proceeding with scheduled outage for machine room move
  - detailed schedule at http://www.gridpp.rl.ac.uk/blog/2009/05/14/schedulemovenewbuilding/

- ASGC moving back to main site

- PIC: issues with dCache load balancing configuration
  - similar to NIKHEF before
  - should there be repository collecting deployment receipts per storage project?
- SARA:
  - DMF upgrade done in preparation for installation of T10k tape drives

10