WLCG Weekly Service Report

Jamie.Shiers@cern.ch

~ ~ ~

WLCG Management Board, 10th June 2008

Introduction

- Although both February and May runs of CCRC'08 are over, a regular (weekly?) service update is useful(?)

➢ Will continue daily operations calls at 15:00, with a weekly summary, until it is decided otherwise...

- Notes from the daily meetings continue to be made available shortly after the con-call is over

- These are now in 513 R-068 for those at CERN – remote joining details remain unchanged
  - x76000 code 0119168

Core Services

- Main issue this week is CERN SRM failure following scheduled upgrade: a post-mortem is here.
- The incident started after the 9.30 SRM2 upgrades created a confusing situation.
- At 13.00 the developer understood that a bug had been introduced for requests involving tape recalls and it was decided to rollback the software.
- The lessons learned from this:
  - Even straight-forward, well understood and exercised upgrades may not be as transparent as we think.
  - Therefore: announce them better. Testing of Castor/SRM software has deficiencies. Until these are addressed, we should adopt a more conservative upgrade strategy:
    - Day-1: upgrade one endpoint, and baby-sit it
    - Day-2: upgrade the others
- Other issues: preparation of FTS data for analysis of May run

Monitoring

- We have observed that few Tier1 installed the FTS monitoring (FTM) to report Tier-1 to Tier-1 transfers back to the central gridview repository. JT queried that they do run FTS monitoring but JA said this was different being an internal monitoring of FTS and that both monitorings were needed. They will circulate advice to the Tier1 sites that are not publishing their T1 to T1 traffic to GridView.

M/W component	Patch #	Status
LCG CE	Patch #1752	Released gLite 3.1 Update 20
FTS (T0)	Patch #1740	Released gLite 3.0 Update 42
FTS (T1)	Patch #1671	Released gLite 3.0 Update 41
FTM	Patch #1458	Released gLite 3.1. Update 10
gFAL/lcg_utils	Patch #1738	Released gLite 3.1 Update 20
DPM 1.6.7-4	Patch #1706	Released gLite 3.1 Update 18

DB Services

- The issues with some DB services during the power cut have been analysed and traced to Ethernet switches of RAC6: on physics power instead of critical power due to wrong connection of the power bars (had correct “yellow label”).
- A new bar connected with critical power was installed as stop-gap and an intervention to cleanup the setup is being planned.
- The upgrade of production databases to Oracle version 10.2.0.4 has been scheduled with experiments:
  - LHCb 5.June (done), LCG 16.June, Atlas 18.June, CMS 24.June.
  - These upgrades are expected to require 2 hours of downtime.
- An upgrade of the downstream databases, ATLDSC and LHCBDSC, to Oracle 10.2.0.4 has been performed on Thursday 5th June.
- More details of CERN interventions: IT status board
- New database clusters at PIC for ATLAS joined to the Streams environment for ATLAS on Monday 2nd June using a new procedure based on Transportable Tablespaces in order to reduce the re-instantiation time.

Site Issues

- BNL (Mon) reported many CERN myproxy failures. McCance[?] explained CERN had seen problems today and replaced its myproxy server with new hardware around midday CEST but that this should have been transparent. He recommended BNL move to using proxy delegation in FTS to avoid such problems.
- BNL (Tue) M.Ernst reported that after the Friday CERN power cut, and continuing over the weekend, network connectivity from BNL to CERN was poor with a high latency. This degraded ATLAS production since the central PANDA database was hence not performing at the required level. It was also difficult to copy the byte stream data required for FDR2 directly to CERN and this had to be rerouted. He had a report that there was a problem with SURFNET but with no detail. It is important for future operation to understand this incident so it will be followed up with the LHC-OPN management.
- NIKHEF (Wed): said that R.Walker was submitting many jobs to NIKHEF that immediately exit using no cpu. SC thought these would be prestaging jobs but that he should ask R.Walker.
- NIKHEF (Thu): asked if the reason LHCb and ATLAS have stopped sending jobs to them is because of the dcache bug in their patch level 4 where an application cannot tell if a file was on tape or disk. This was confirmed and that it was not known if this bug is also in patch level 5. It is thought it came in in patch levels 3 or 4. H.Renshall will check with P.Fuhrmann.
- BNL (Thu): H.Renshall asked if their weekend SURFNET problems were now understood. ME said not and HR volunteered to follow this up with CERN CS group.

Experiment Issues

- ALICE (PM): See next slide...
- LHCb: elog summary of weekly activities expected Monday afternoon (after weekly meeting). Preliminary conclusions from May run: https://prod-grid-logger.cern.ch/elog/CCRC'08+Observations/106
- CMS: Phase 2 operations elog: https://twiki.cern.ch/twiki/bin/view/CMS/CCRC08-Phase2-OpsElog
- ATLAS: Plans for last week: https://prod-grid-logger.cern.ch/elog/CCRC'08+Observations/104

ALICE Report

- All ALICE sites VO-boxes have been updated with SLC4, gLite 3.1 and new AliEn services. In production since one week

- Data taking: detector calibration data since mid-May, replicated online to T1s quasi-online (see Gridview plots) ~50MB/sec, full detector complement data taking starting this week (alignment, calibration)

- Processing: calibration data processed and put in conditions database, user and detector expert analysis ongoing on Grid and CAF

Questions

- We should probably rename CCRC'08 mailing list
- Any objection to wlcg-operations@cern.ch ?
- A new wiki focussing on on-going service issues (alone) will be prepared in coming weeks, summarizing the key links / pages that we have learned we need to run / monitor the service
- Has anyone noticed the subtle changes in template over time?

Summary

- The service runs smoothly – most of the time
- Problems are typically handled rather rapidly, with a decreasing number that require escalation
- Most importantly, we have a well-proven “Service Model” that allows us to handle anything from “Steady State” to “Crisis” situations
- We have repeatedly proven that we can – typically rather rapidly – work through even the most challenging “Crisis Situation”
- Typically, this involves short-term work-arounds followed by longer term solutions
- It is essential that we all follow the “rules” (rather soft...) of this service model which has proven so effective...