WLCG ‘Weekly’ Service Report

Harry.Renshall@cern.ch

~ ~ ~

WLCG Management Board, 14 October 2008

Introduction

- Last week (6 to 12 October):

- Notes from the daily meetings can be found from:
  - https://twiki.cern.ch/twiki/bin/view/LCG/WLCGOperationsMeetings
  - Regular local participation from all IT physics groups.
  - Systematic remote participation by BNL, RAL, PIC, NIKHEF and GRIF.
  - (Some additional info from CERN C5 reports & other sources)

Site Reports (1/2)

- CERN:
  - LHCb LFC recataloging of srmv1 to srmv2 file names completed successfully on Tuesday. Srmv1 services were then progressively stopped definitively.
  - A new CASTOR hot fix release for the 2.1.7 version (-19-1) has been deployed on c2atlas. This release fixed the behavior of pools that have a tape backend, but which refuse new requests when they become full. A bug in this area was found on the preproduction system without having been detected during the certification phase. This is explained by the absence of such a pool in the certification system (this has been fixed). The deployment of 2.1.7-19 on Atlas had to be postponed by 24h due to this incident. Note hotfix -2 affecting updates on replicated diskpools has since been deployed for all experiments.
  - There was an unannounced upgrade of the AFS UI (job_submit scripts) as part of a voms certificate change at 15.00 on Thursday 9 October but which did not complete properly and stopped the UI from working. Fixed by 17.00.
- RAL:
  - Cosmics exports to RAL from weekend were throttled back to 50-60 MB/sec instead of 80-90 MB/sec leading to a 40TB backlog on Monday. Then CASTOR went down with a 60 TB backlog so ATLAS excluded them from cosmics distribution till they recovered.
  - LHCb LFC recataloging of srmv1 to srmv2 entries ran very slowly (5 hours instead of 10 minutes) so was suspended and a complete copy taken from CERN instead.

Site Reports (2/2)

- BNL:
  - The US PANDA services at BNL (previously the primary instance) is to be stopped by the end of 2008 and CERN will become the primary instance.
- NL-T1:
  - Amsterdam area power failure at 16.30 on Saturday 5th October. Srm.grid.sara.nl reported as partially available from 17.00 Sunday 6th but transfers were still failing so ATLAS quarantined the site until late on Monday 7th. Post-Mortem is in preparation.
- General:
  - Bug found in SLC4 FTS where large error messages crashed the server – fixed on Wednesday 8 October. More serious is bug in the area of no longer returning detailed error messages to the FTS api – apparently known since some time. A fix is now in test but we suggest a Post-Mortem on this. Meanwhile ATLAS and CMS transferred about 20TB each last week.
  - gLite 3.1 Update 33 contained a bdii configuration bug whereby sites not using Yaim had a missing file system root variable causing an incorrect chown. New version being certified to go to PPS this week. This bug took down the GRIF site. Also flickering of resources in the information system seen at several sites – the meta-rpm was hence withdrawn on Friday.

Experiment reports (1/3)

- ALICE:
  - Have been testing SLC4 FTS and CREAM-CE
  - Prepared a site requirements document for CREAM-CE deployment.
- LHCb:
  - All LFC now converted from srmv1 to srmv2 endpoints allowing sites to shut down srm1 services.
  - Over weekend of 11/12 October suffered overload of the CERN castor server after requesting 180000 prestage requests which hit a server that was being drained for hardware replacement. A Post-Mortem is in preparation.
  - Long running problem of insufficient LSF batch slots being occupied given their shares at CERN finally understood. Another major experiment was requesting unneeded memory swap resources which caused LSF to not schedule on otherwise free job slots.

Experiment reports (2/3)

- CMS:
  - Magnet ramped up to 3.8 Tesla over weekend of 11/12 October taking cosmics scheduled to run till 27 October (the CRAFT1 run).
  - Silent file corruption discovered at CERN on one disk server in CASTORCMS/CMSCAF: FIO have started to draft a postmortem for this incident https://twiki.cern.ch/twiki/bin/view/FIOgroup/PostMortem20081008
    - For the affected files the checksum in DBS does not match the files when you retrieve it from castor and ROOT cannot open the file
    - Traced to a defective disk on a fileserver that was not clearly reporting errors.
    - Using the checksums provided by CMS FIO found that 29 files had a corrupted copy on tape. All 29 files could be repaired because the original diskcopy was still available
    - Another 2 files, for which they did not have the CMS checksum, could also be repaired by finding their original copy using the CASTOR log files.
    - 4 user files are unrepairable because they were originally created in a Disk0Tape1 service class (DEFAULT) and then recreated in CMSCAF, which is Disk1Tape0. This is a forbidden transition but unfortunately a CASTOR bug allows for it to happen (a fix will be deployed on the 14/10). The original tape copies are still available and users will be informed.

Experiment reports (3/3)

- ATLAS:
  - PreStageTests

    https://twiki.cern.ch/twiki/bin/view/Atlas/PreStageTests
    - Pre-staging is functional on all CASTOR sites

      CNAF, RAL, CERN, ASGC
    - dCache sites are having more problems
    - BNL, TRIUMF, NDGF should be OK
    - Realistic reprocessing rates from tape look *hard*
    - Phase II will start at T1/T2 clouds which can pre-stage

  - PANDA instance at CERN (the high level workload and pilot jobs manager) planned to become the sole one by the end of 2008.
    - Will need existing UI to be expanded from 1 to 3 load balanced servers
    - Existing core of 3 mysql servers + 3 hot spares to be migrated to Oracle
    - Other servers to be hardened

    - Main developer seconded to CERN from BNL from January

Conclusions

From the September F2F:

(Too) many software upgrades and we do not see this slowing down. (It is not)

Many miscellaneous failures which will also continue. (They are continuing)

To maintain good quality LCG services over the next months is going to require constant vigilance and be labour intensive. (It is labour intensive)