Table of Contents

Week of 080825

Open Actions from last week:

Daily WLCG Operations Call details

To join the call, at 15.00 CE(S)T Monday to Friday inclusive (in CERN 513 R-068) do one of the following:

1. Dial +41227676000 (Main) and enter access code 0119168, or
2. To have the system call you, click here⊞

General Information

- Web archive: https://mmm.cern.ch/public/archive-list/w/wlcg-operations⊞
- CERN IT status board: http://it-support-servicestatus.web.cern.ch/it-support-servicestatus/⊞
- elogs: https://prod-grid-logger.cern.ch/elog/CCRC'08+Logbook/⊞

See the weekly joint operations meeting minutes

Additional Material:

Monday:

Attendance: local(Julia, Daniele, Andrea, Maria, Jamie, Harry, Nick, Roberto, Patricia, Miguel, Gavin);remote(Jeff).

elog review: See under CMS Tier 0 Workflows below.

Experiments round table:

CMS (DB):

General: CRUZET-4 over at ~8am in the morning, ~38 ml evts collected during the exercise, most interesting part from Thursday on, >25 only in last weekend. Plenty of precious info and feedback on a real-life exercise. CRUZET Jamboree on Wednesday afternoon. We will reconvene globally again with magnetic field at the end of the week.

Tier-0 workflows: SLS reports that the "CMS Online databases" have 0% availability. As the overall DB tests are designed atm, it means that it's not possible to connect to some instance of the cluster (see [*1] and [*2]). It was related to a CMS Online intervention, but it was notified to the CMS Databases HN, and not to the Computing Announcements HN, so since there was no scheduled intervention in the IT Service Board [*3], and no message in the general Computing HN, we opted for pinging support lists and open tickets, e.g. GGUS [*4] and mail to PhyDB.Support@cernNOSPAMPLEASE.ch (mail address as in the SLS pages), plus ELOG entry in the WLCG central ELOG system [*5]. An immediate reply to the GGUS was received, explaining us (thanks Jacek). We are now working to improve this CMS-internal communication flow among Online and Offline/Computing.

[*1] http://sls.cern.ch/sls/service.php?id=phydb-CMS⊞ [*2] http://sls.cern.ch/sls/service.php?id=phydb_CMSONR⊞ [*3] http://it-support-servicestatus.web.cern.ch/it-support-servicestatus/⊞ [*4] https://gus.fzk.de/pages/ticket_details.php?ticket=40081⊞ (now closed and verified) [*5] https://prod-grid-logger.cern.ch/elog/CCRC'08+Logbook/444⊞

Distributed Data Transfers: We see 1) issues with the stager agent (experts aware and investigating) + 2) some Castor issues causing problems to the CAF (2 tickets to CERN-IT still pending over the weekend, see [$1] and [$2]) + 3) issue with download agents in at least 2 T1 sites. This overall causes PhEDEx service to be labelled as 'degraded' in SLS. CMS needs faster reaction from CERN-IT on tickets opened over the weekend, or suggestions on how to improve on better notifications. Jamie emphasised the need for such good communications during beam-on time.

[$1]
http://remedy01.cern.ch/cgi-bin/consult.cgi?caseid=CT0000000546182&email=stephen.gowdy@cern.ch
[$2] http://remedy01.cern.ch/cgi-bin/consult.cgi?caseid=CT0000000546181&email=peter.kreuzer@cern.ch

LHCb (RS): 1) could not access SAM this morning - Remedy ticket opened. 2) The CNAF 'operator emergency' ticket sent last week did not arrive to CNAF. To be followed up with GGUS. 3) Not much activity from LHCb at the moment as we are working on our data processing/jobs model.

ALICE (PM): We have not yet deployed the next version of aliroot. We have minor issues with site voboxes. As regards deployment of FTS services under SLC4 we are neutral as long as there is no disruption to the service. Harry said this is a hot topic being handled by the EGEE operations meeting but needed a rapid decision and actions.

Sites round table:

CERN (MCS): 1) The CASTOR problems seen by CMS and LHCb from Friday and over the weekend are due to intermittent network router problems which do not trigger any operational alarms. To be followed up with CS group. 2) A question to CMS - we observe CMS doing repack operations on the CASTOR T0EXPORT pool at 500 MB/sec which seem to include writing ? Daniele said there should be no rewriting into this pool - MCS will check.

Core services (CERN) report:

DB services (CERN) report (MG): We may need an upgrade of our RHEL4 linux kernel to fix some Oracle crashes that we see from time to time. This has been tested on our test system and we intend to deploy it on our integration cluster for a few days. For some time we have had a high load on some CMS Oracle servers and this has now been traced to a heavy and frequent query that was, in fact, unecessary and has now been removed.

Monitoring / dashboard report (JA): There has been a change in what is published in the xml of the Imperial College Grid jobs real time monitor and which we use to build our dashboard displays. The field that should contain the jobid now contains the WMS host name. We have contacted the developer directly.

Release update:

AOB:

Tuesday:

Attendance: local(Miguel, Roberto, Julia, Andrea, Harry, Jean-Philippe, Maria, Gavin);remote(Derek, Jeremy, Jeff, Xavier).

elog review:

Experiments round table:

ATLAS (by mail from SC): After few discussion, turns out FDR-2c scope will be relatively modest for what data exports are concerned. According to Kors data volume will be small and data won't be available before Sat Aug 30th. Therefore, data will not be exported before saturday and T1s will receive only AODs and DPDs (the latest are the interesting part for clouds, since they are produced with the latest software release and therefore are an enhancement in respect of DPDs of FDR2-a. Therefore, no RAW data distribution, no ESD distribution, no export shifters at P1. Obviously, the all T0 machinery will work as during data taking, i.e. reconstruction of the express stream, daily meeting to discuss data quality plots, bulk processing only in case of green light from data quality.

Cosmic Data are daily collected and exported, no issues to report. Some data were not available on the weekend on CASTOR (disk server failed before data could go to tape and vendor has been consulted). I would like to ask for the status of this. Answer from FIO group is: The disk server (lxfsrd4101) failure in ATLAS is not related to the weekend network problem. There is a multiple hardware failure on the box, the vendor call is open but the vendor has not provided a complete solution. This morning we contacted the vendor again.

I would like to ask for the status of the ATLASUSERDISK space token at CERN (100TB). Requested some time ago. Answer from FIO group is: Regarding ATLASUSERDISK, the creation process is mostly finished. Some last details regarding SRM and the publishing on the space token are being looked at. Simone Campana should be contacted to test soon.

CMS (DB):

Distributed Data Transfers: Greenish quality in CERN-T1 routes. After an injection operational mistake (over the busy weekend) was sorted out [1*], transfers from CERN and T1 sites to the CAF also restarted to flow. A suspect of a dataset info inconsistency in PhEDEx+DBS popped out with the dataset CRUZET4*TkAl*ALCARECO: to be followed up [2*]. [1*] data access problems on the CAF (related to Castor and the nwtwork problems over the weekend) -- https://savannah.cern.ch/support/?105337 ---> CLOSED [2*] suspect of a dataset info inconsistency in PhEDEx+DBS -- https://savannah.cern.ch/support/index.php?105364

Tier-1 workflows: Not much happens apart from transfers atm. --- Castor at CNAF T1 did not come up in a good state after the scheduled downtime (19-20/8) for core services update. Migration problems arose, probably caused by the change of the nameserver, stopping Buffer->MSS transfers for all VOs. Experts are working on it in close contact with the Castor devs @CERN. They have a clue already, and a script to solve the situation may be available by Castor devs already within the day. To be followed up. Effect on CMS atm: approved subscriptions flew, files got transferred, did not go to tapes yet, 10 TB high-prio threshold was hit, no more files flow to Buffer atm until some migration is demonstrated to work.

OPENED tickets: - low quality transfers to ASGC: it was an expired proxy, fixed by site admins -- https://savannah.cern.ch/support/?105338 ---> CLOSED

T2 workflows: See https://prod-grid-logger.cern.ch/elog/CCRC'08+Logbook/446

LHCb (RS): Have resumed full production running for simulation jobs (only at Tier 2 sites in the LHCb computing model). Preparing to distribute a new version of their Brunel application to Tier 1 today after which production reconstruction and stripping jobs will resume with 2000 files ready to be processed (should be later today). LHCb have requested interactive access to a CERN PPS worker node to carry out gLexec tests.

Sites round table:

RAL (DR): The ATLAS CASTOR instance now reproduces the bulkInsert constraint violation so will be left down for intensive debugging with the CASTOR developers. The LHCb instance will, however, be restarted

Tuesday: 3

but carefully watched in case the problem reoccurs there. If so it will be cleared by a stager restart (only the request queue is affected).

NL-T1 (JT): reported no jobs running there from ATLAS or LHCb and asked why. Roberto reported for LHCb that they do not run Monte Carlo work at the Tier 1 but reconstruction and stripping should start again this afternoon so work should start to flow. The question will be passed to ATLAS.

Core services (CERN) report: Starting at 10 am on Friday 22 August a switch in the computer centre caused an overflow putting a network router in a bad state. This was thought to have been resolved at 11.30 but there was a residual problem causing intermittent access failure to some castor disk pools (mainly of LHCb and ATLAS) leading to some SAM test and experiment job failures for the next two days (they would usually work on retry). The problem was finally understood and corrected by 10 am on Monday as being in one 10 Gigabit module on a router, providing one of the four 10G uplinks to the core, and although everything looked correct on this link the traffic was not forwarded in one direction. CERN CS group are now in contact with the vendor to understand why the router did not show any error message. A post-mortem with a detailed time line, and recommending that in future such degradation in CASTOR services should be promptly logged in the CERN service level status board, has been prepared by FIO group at https://twiki.cern.ch/twiki/bin/view/FIOgroup/NetworkPostMortemAug25

Following up on yesterdays report of high i/o rates on the CMS T0EXPORT pool Miguel CS reported they had suspected a CMS application was reading 10K bytes from 128K blocks so provoking 10 times as many block I/Os as needed but they no longer think this is the case. There was no more news from the CMS side on this issue.

DB services (CERN) report:

Monitoring / dashboard report: The issue reported yesterday with the Imperial College rtm application xml publishing changing is because they have switched from using RB instances to WMS instances and introduced a minor bug. It should be fixed quickly.

Release update:

AOB:

Wednesday

Attendance: local(Julia, Oliver, Roberto, Andrea, Harry, Patricia, Maria);remote(Jeff, Derek, Xavier).

elog review:

Experiments round table:

CMS:

Distributed Data Transfers: CERN export quality started to get worse this night from ~23h00 UTC on, mostly affecting FZK, RAL, FNAL. Some identified errors: - CERN->FZK "an end-of-file was reached (possibly the destination disk is full)" errors. - CERN->RAL "stage_prepareToPut: Internal Error" errors, plus CMS SAM lcg-cp tests failures. - CERN->FNAL "Could not submit to FTS" errors in last hrs.

T1 workflows: News from last shift: CNAF restarted to migrate to tape at good speed. Need to balance among VO activities, so after some hrs of >200 MB/s rates, it was stopped then restarted, from this morning on it should be on and more stable.

T2 workflows: Following up on discussions related to other open tickets, as from previous more detailed shift reports.

ATLAS (SC): 1) ATLAS got the ATLASUSERDISK pool. I tested the basics, works OK for me. Pool dumps are provided. Action closed. 2) I have asked for a new pool after yesterday ATLAS CREM meeting. 100 TB for groups at CERN. Needed in order to phase out "default" pool for users. 3) Graeme had exchange with RAL. Seems they are back online. We are about to restart functional tests also for them. The site should show reliability for > 48h to pass the quarantine and start receiving detector data and production jobs. Every other site shows > 99% efficiency at the moment. so today I basically have only good news.

ALICE (PM): minor issues with Tier 2 sites.

LHCb (RS): After having installed the new version of Brunel everywhere and changed the workflow we started testing CCRC-like activity (production 3008) across all T1s (activity going in parallel with normal MC simulation activity across all other centers). All jobs were crashing at NIKHEF and GRIDKA. It looks like the process is aborted as soon as the CondDB access is attempted. The jobs went normally through reading the input file. We also started to have some problems at NIKHEF retrieving the tURL with a error message we saw in the past during May phase of CCRC at IN2P3. The solution at that time suggested was to upgrade the clients. It seems the problem is still there however. This is rather symptomatic of a SRM problem at SARA (timeout) wrongly interpreted.

2008-08-27 10:31:09 UTC Wrapper_279 DEBUG: SRM2Storage.__gfal_turlsfromsurls: Performing gfal_turlsfromsurls
2008-08-27 10:34:35 UTC Wrapper_279 ERROR: SRM2Storage.__gfal_turlsfromsurls: Failed to perform gfal_turlsfromsurls: [SE][StatusOfGetRequest] httpg://srm.grid.sara.nl:8443/srm/ managerv2: org.xml.sax.SAXParseException: An invalid XML character (Unicode: 0x10) was found in the element content of the document. Unknown error 4294967295. Opened an ALARM GGUS ticket to SARA for this error.

Sites round table:

NL-T1: Jeff reported on the LHCb turl problem - not simple as some turls are returned correctly while others give strange results. All worker nodes have been upgraded to the latest kernel.

RAL (DR): Last night the CMS CASTOR instance showed the constraint violation problem so has been taken offline till the end of today. They will restart services and simply restart if they hit the bug out of hours otherwise resume debugging. The current suspicion is an Oracle bug connected to having two experiments share a RAC - they may try a single RAC test. He confirmed later they are running Oracle 10.2.0.3 which is where CERN saw the cache problem when not using fully qualified names for name spaces (RAL do not think this is their case).

Core services (CERN) report:

DB services (CERN) report: RAL replicated databases have received the latest security patch.

Monitoring / dashboard report:

Release update: Roberto asked the status of the WMS problem of wrong proxies when two different roles are used for the same user. Oliver thought the fix for this was in a major release of the WMS (a 6 months accumulation) that should go into certification in about a weeks time.

AOB:

Thursday

Attendance: local(Andrea, Laurence, Daniele, Jan, Gavin, Julia); remote(Derek, Jeff).

elog review: 448

Experiments round table:

- CMS (Daniele):
  - ♦ problems at Tier-0 with the Data Discovery service, needed to roll back to a previous version.
  - ♦ There is a pending hardware request (see eLog).
  - ♦ Since yesterday night in PhEDEx the quality of data transfers from CERN was very bad due to a problem with the CMS SRM instance. Jan confirmed that there was a problem causing a 95% degradation of the service, fixed this morning after a six hours downtime. Daniele reported that now transfers seem to be working but only to ASGC and CNAF.
  - ♦ At RAL CMS is having the same CASTOR database corruption issues as ATLAS and LHCb, developers are looking into them.
  - ♦ Tape migration agents at CNAF are now running fine, backlog is being digested.

Sites round table:

- RAL (Derek): unscheduled downtime of the ATLAS SRM instance, seems a new problem and it is being debugged. Not affecting the other VOs. According to Derek the CMS SRM should be working, but Daniele says it is not the case; Derek will check.
- NIKHEF (Jeff): the alarm ticket submitted yesterday by LHCb has been solved, but the internal response was not satisfactory and a post-mortem will be performed.

Core services (CERN) report (Jan): clarified that only CMS had problems with SRM, the other VOs were fine.

DB services (CERN) report: nothing to report

Monitoring / dashboard report: nothing to report

Release update: nothing to report

AOB: Daniele announced that from now on, on Fridays he will announce the CMS plans for the weekend at this meeting.

Friday

Attendance: local();remote().

elog review:

Experiments round table:

ATLAS (by email from SC): 1) FDR bulk reprocessing is stuck. Problem seem to be connected to CondDB. I quote message from Armin: 'FDR2c bulk processing over night has failed miserably because of connection problems to the Oracle CondDB'. Due to the delays we are in competition with cosmics processing, and although I've throttled both to a combined job submission rate of 10-12 per minute (which is less than the usual one for cosmics alone), the problems persisted during the whole night and basically not a single job succeeded after 11pm yesterday. Error message is 'Oracle error ORA-02391: exceeded simultaneous

SESSIONS_PER_USER limit'

2) RAL seems to be back online. Efficiency is approx 90% for the current 5% test. RAL people asked to increase load, I am evaluating how to do this without complicating T0 activities.

3) There are currently errors for transfers of files on tape. Those are FDR2a data (June 2008) being replicated to every T1 for prestaging test next week. Inefficiencies are therefore expected.

CMS (from elog by DB):

General : plan to play with cosmic runs with magnetic field, may start this evening. Plan another (last one) mid-week GR exercise for mid next week. --- Patricia is still waiting for a reply on a hw request [&1].

Distributed Data Transfers : Thanks Gavin/Jan/all for the executive summary of the Aug28 incident [*1]. Overall bad quality in **all** CERN->T1 routes since few days persists [*2], though: it may be a superimposition of some different problems (e.g. we know at RAL for sure), the overall situation seems drastically worse than before, still. Investigating. If you expect some CERN export problems may still be seen, please point to its precise signature.

T1 workflows : mostly preparing tape families and name space in connectiont with T1 contacts, for the next exercises.

T2 workflows : Still following up on tickets.. (slow replies in many cases)

1st PRIORITY OF THE DAY for CMS to this WLCG forum: * CERN outbound transfers to T1 sites * NOTE: we have this same priority since 2 days now.

Plans for the weekend: Preparing for the cosmic exercise with magnetic field, may start this night. We would need a high presence and coverage in case of CMS notification of problems on the CERN instractructure on Saturday and Sunday, August 30th-31st.

Links: [&1] http://cern.ch/helpdesk/problem/CT546412&email=pbittenc@cern.ch [*1] https://twiki.cern.ch/twiki/bin/view/FIOgroup/FtsPostMortemSrmAug28 [*2] http://cmsweb.cern.ch/phedex/prod/Activity::QualityPlots? graph=quality_all&entity=dest&src_filter=CERN&dest_filter=T1&no_mss=true&period=l96h&upto=&.submit=Upda

Sites round table:

RAL (DR): They found a problematic disk server that had been affecting CASTOR stage_put operations and have removed it. All three CASTOR instances are running at the moment but the constraint violation bug must still be in the system. They have looked at CMS CERN to RAL transfer failures yesterday and suspect it is because file sizes have increased (from 2 to 4 GB) leading to timeouts (DB said these would be CRUZET 3 files) and requested CERN to increase the channel timeout from its current 10 minutes.

Core services (CERN) report: The SRM service for CMS (srm-cms.cern.ch) was severely degraded (90-95% down) between the hours of 03:00 and 09:00 UTC on 28 August 2008 due to a tablespace reaching its quota limit. A post-mortem report has been filed at https://twiki.cern.ch/twiki/bin/view/FIOgroup/FtsPostMortemSrmAug28

CERN will upgrade it's linux system software on Monday and this includes a new RFIOD. We have also scheduled an upgrade of all SRM services next Wednesday morning - probably 10.00 to 11.00 but to be confirmed.

DB services (CERN) report:

Monitoring / dashboard report: The Imperial College RTM team has implemented modifications to publish not only RB name, but also LB. Needed changes are also implemented on the Dashboard side. CMS XML collector is now up and running. Due to wider coverage of job status changes information from IC XMLs the reliability in all Dashboard instances (all 4 LHC experiments) for job monitoring will increase.

Release update:

AOB:

This topic: LCG > WLCGDailyMeetingsWeek080825
Topic revision: r11 - 2008-09-01 - HarryRenshall