Table of Contents

Week of 090511

WLCG Baseline Versions

WLCG Service Incident Reports

- This section lists WLCG Service Incident Reports from the previous weeks (new or updated only).

GGUS Team / Alarm Tickets during last week

- Tickets: see https://gus.fzk.de/ws/ticket_search.php⬚ and select "VO" - this will give all tickets, including team & alarm - or use the following link:
  - ♦ In the GGUS Escalation reports every Monday⬚
- Please observe the procedure to become a LHC Experiment VO TEAM member⬚ and other most recent GGUS FAQs⬚.

Weekly VO Summaries of Site Availability

- ALICE⬚
- ATLAS⬚
- CMS⬚
- LHCb⬚

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/⬚
- M/w PPS Coordination Work Log: PPSCoordinationWorkLog

See the weekly joint operations meeting minutes

Additional Material:

Monday:

Attendance: local(Jamie, Miguel, Harry, Nick, Jean-Philippe, Markus, Jeff, Olof, Maria, Simone, Diana);remote(Angela, Jeremy, Daniele, Gareth, Michel, Gang).

Experiments round table:

- ATLAS - (Simone): Dispatching of Cosmic data has stopped. The exercise ran smooth over the Weekend. No problems to report. One problem at SARA with memory in dCache but it was resolved

within a couple of hours. BDII lookups from WNs at NIKHEF: not understood but Graeme and Hurng are investigating this issue. Gareth: *in standard ATLAS production jobs, when copying files back to the storage element, do you specify the space token?* Simone: yes but it depends on whether uploading or downloading the files. It might be that for some sites the space token is not passed on a 'get'.

- CMS reports - (Daniele) Brief report. mostly following up a number of tickets for transfer problems. Some timeout problems for big files (>50GB?!!!). Daniele has closed ticket for CNAF about Custodial data slow to show up on T1_IT_CNAF_MSS - issue solved. Another issue with 11 unavailable files at IN2P3. Daniele mentioned that there are a number of announced interventions affecting CMS: CASTOR upgrade at CERN - time slot will be discussed at the CMS ops meeting. Daniele also wondered about the CMSONR and CMSR DBs rolling upgrades foreseen for 14th of May (http://it-support-servicestatus.web.cern.ch/it-support-servicestatus/ScheduledInterventionsArchive/CMSONR): were they announced to CMS? what's the impact? Daniele will follow up with DBAs.

- ALICE -

- LHCb reports -

Sites / Services round table:

- NIKHEF (Jeff): tough weekend with Storage. One of the poolnodes didn't come back after some intervention last week. dCache is running out of memory. The memory was increased by a factor 4. Simone: *do you know when the memory gets exhausted?* no? Jeff mentioned that a similar problem might have been seen by LHCb on dCache...
- FZK (Angela) issues: one problem with a FTS channel, which was down over the Weekend. Unknown why, and the developers don't understand. Will try to recreate the channel today. Planned downtime for part (1/2) of tape system, probably on Thursday.
- RAL (Gareth) things coming up: tomorrow network intervention, 1 hours site down in GOCDB. The week after: 18-19/5 the CASTOR dbs are migrated to new hardware - two days outage for CASTOR.
- GRIF (Michel) NTR
- ASGC (Gang): timeouts from LCG commands. Problem is likely to be due to the BigId issue.
- CERN (Miguel)
  - ♦ today there was MyProxy upgrade. Mail px.support@cernNOSPAMPLEASE.ch in case of problems.
  - ♦ Change on CASTORCMS for starting stager daemons in loadbalanced manner.
  - ♦ Transparent intervention tomorrow for SRM upgrade (http://it-support-servicestatus.web.cern.ch/it-support-servicestatus/ScheduledInterventionsArchive/09) . One question: the intervention was registered in GOCDB last week but still flagged as 'unscheduled' in GOCDB, why? Nick recommends to submit a GGUS ticket.
  - ♦ Linux Scheduled upgrade today (http://it-support-servicestatus.web.cern.ch/it-support-servicestatus/ScheduledInterventionsArchive/09)

AOB:

- Maria wonders about some CMS member with DoE certificate. Daniele is following up. Another issue with Coreen person requesting a certificate: Maria will mail Gang

Tuesday:

Attendance: local(Jean-Philippe, Olof, Simone, Roberto, Miguel);remote(Daniele, Michael, Brian, Gareth, Angela).

Monday:

Experiments round table:

- ATLAS - (Simone) Nothing special to report other than that some of the ATLAS DM services at CERN will be redeployed in the coming days. There might be intermittent problems. Miguel (CERN): *timeouts reported by Armin yesterday - the network switch turned out to not have a standard fiber for the router uplink. The diskservers on that switch have been removed from production until the link problem is fixed. Short incident on CASTORATLAS this morning where the execution plan had changed for a particular stager query.* Brian (RAL): *test transfers with ATLAS production role between CERN - RAL after network change but didn't manage to transfer files across, is it a known problem?* Simone hasn't seen any transfer issues but suggested that Brian to try going through the whole chain starting from lcg-cp with a spacetoken. If that works, try with FTS. Jean-Philippe: instead of using lcg-cp, you may try lcg-gt (for a file at RAL), which will pinpoint whether it is a SRM or FTS problem.

- CMS reports - (Daniele) Working on tickets related to site problems, especially for transfers. Have been seeing an impressive cycle of ticket resolution in the recent past, which proves that many Tier-2s are able now to reach the 80% availability. CASTORCMS upgrade to 2.1.8-7: a good date would be May 25th? Miguel agrees. The rolling CMSONR and CMSR DB interventions are both supposed to be transparent so they can go ahead.

- ALICE -

- LHCb reports - (Roberto) Another FEST week, which means that by tomorrow usual transfers between CERN and tier-1s should be expected. Several problems with WMS testing at RAL, SARA and PIC - GGUS tickets opened to relevant people. A problem at GridKA SRM not returning TURLs has been understood: DIRAC doesn't always set 'Done' state for 'get' request which caused the system to be unstable.

Sites / Services round table:

- RAL (Gareth) scheduled outage this morning for network intervention. Unfortunately the intervention failed and had to be rolled back and repeated later.
- BNL (Michael) scheduled intervention this morning (local time): move pnfs PostGres from current 32bit to 64bit database system with large caching (up to 48GB). This increased cache will hopefully improve the overall performance significantly.
- FZK (Angela) NTR
- CERN (Miguel) scheduled SRM intervention this morning went fine. The scheduled linux upgrade, mentioned yesterday, was only scheduled today due to an urgent kernel fix. Unfortunately the upgrade caused problems for the LSF batch service (lost jobs). A post-mortem at https://twiki.cern.ch/twiki/bin/view/FIOgroup/PostMortem20090512

AOB:

Wednesday

Attendance: local(Olof, Patricia, Gavin, Simone, Roberto, Jamie, Jean-Philippe, Jamie, Wayne Salter, Miguel);remote(John Kelly, Michael).

Experiments round table:

- ATLAS - (Simone) One relevant news: after discussions with ASGC people the site is considered operational now. All storage related functional tests work ok (SRM, LFC). However, when trying FTS subscription the transfer went into timeout in gridftp phase. Investigations show that all transfers to/from ASCG ends up with ~400-500kb/s with CERN and other Tier-1s. Is the routing via the OPN

or not? If not, this is a blocking issue for the STEP09. Jamie: *do we know the hostnames?* yes, can be provided. Wayne Salter (OPN responsible at CERN) will follow up once the hostnames have been provided to him.

- CMS reports - Daniele busy at GDB today.

- ALICE - (Patricia) back from holidays... Production is running quite smooth. Checking WMS submission issues with CREAM CE at CNAF. Patricia is following this up directly with the site.

- LHCb reports - See Recap on LHCb issues with dCache at FZK⧉ for a discussion of the issues seen and possible ways forward. (Roberto) Main points:
    - ♦ issues with unstable WMS @ Tier-1s - the service is not really usable for daily activity.
    - ♦ After SRM upgrade at CERN yesterday, LHCb started to see massive error rates. Gavin: *seems to be a problem with 2.7-17 release, which has a specific issue with multi-file 'get' requests and affects at least LHCb and CMS. A bug report has been submitted to CASTOR/SRM developers. All production end-points (except PPS) have been downgraded to 2.7-15. A post-mortem will be produced*

Sites / Services round table:

- RAL (John Kelly): nothing specific to report.
- BNL (Michael): yesterday's intervention was completed in time. New PNFS server is running on improved hardware. Stable operation during night and load has come down.
- ASGC
    - ♦ ASGC T1 is brought back to ATLAS DDM Functional Test this morning. Transfer from INFN,NDGF and TRUIMF to ASGC are fine with 100% efficiency, but transfer from CERN to ASGC did not succeed due to low transfer speed.
- CERN (Miguel): besides the SRM issue mentioned above, tomorrow the CASTOR Oracle db services will be patched with the latest security patch. Intervention is transparent and will start at 09:00 (CEST). The name server database will be patched at 13:00.

AOB:

Thursday

Attendance: local(Jamie, Maria, Miguel, Harry, Jean-Philippe, Nick);remote(Gang, Roberto, Michael, Gareth, Angela).

Experiments round table:

- ATLAS (Simone) - exchange of info with ASGC this morning. Jason started looking at problem reported yesterday (slow transfers) - about to contact CERN for iperf tests CERN-ASGC to test performance between 2 disk servers (castor here & there). JPB - traceroute? Simone - path goes through OPN. Some news tomorrow... Miguel - should setup iperf server & give host and port name and we'll trigger test. Simone - STEP09: exchange with CMS - 1 scenario not thought of but important to test which is concurrent writing of data into CASTOR@CERN. ATLAS test suite uses recycle class - data go to tape but always same one(s). Can change class - run test for 48h. This would burn 60TB of tape. Miguel - from tape side is same if recycle pool or non-recycle pool. Same drives & diskservers... Have to agree 48h - could be extended from ATLAS side. Distribution of data for AODS and DPDs stopped in 6/10 clouds in Feb - disk space in Clouds. If in STEP09 want realistic situation have to distribute copies according to computing model. To be discussed today... Data movement will be triggered next week.

- CMS reports - Report by Daniele: The WLCG Ops call today clashes with one of CMS-internal STEP planning meeting, so I cannot attend, apologies. One thing to report is that an operational mistake in a removal cycle of the normal CMSSW deployment activity on EGEE sites has caused swinst CMS-specific SAM tests to fail at all sites. It's a tool/infrastructure issue, and not a site(s) issue.

- ALICE -

- LHCb reports (Roberto) - SRM issue with upgrade reported yesterday fixed after rollback. This is fest week for LHCb. FEST stuck somewhere between online and castor - lhcb bookkeeping service?? Problems from previos days: WMS stability issue at several T1s, SARA isolated - user(s) overloading WMS. Banned DN of users & restarted - looks much better. RAL & PIC still same - GGUS. Network intervention at RAL over last days - many user jobs stalled & returned to DIRAC central services. Gareth - had declared a 1h outage followed by ~2h at risk.. Around problem when network worked on paused batch jobs and then recontinued. Maybe a side effect of that? Did not drain all qqqq. Intervention on CASTOR next Mon/Tue.. Miguel - Gavin will add postmortem of SRM problem to wiki page. Problem case is actually tested but was not caught in tests.. (2 TURLs in same command ..) Postmortem link: https://twiki.cern.ch/twiki/bin/view/FIOgroup/PostMortem20090513

Sites / Services round table:

- RAL (Gareth) - We have the plans in place for the move of the RAL Tier1 to the new computer building. Our blog entry that details the timetable for this can be found here ⬚. Significant outage during this move! 22nd June - 3rd July

- DB (Maria) finished applying Oracle April CPU. Last cluster is cms offline - being down now. Miguel - Castor DB also being patched.

AOB:

- Nick - sites might want to get glite 3.2 UI. Going into cert now. Expect to see next week... Has fix for lcg_cp.

Friday

Attendance: local(Jamie, Jean-Philippe, Harry, Miguel, Olof, Nick, Gavin, Patricia, Simone);remote(Gonzalo, Gang, Michael, Gareth, Jeff).

Experiments round table:

- ATLAS - (Simone) Two points:
    - BNL report high load on dCache yesterday evening. This is because ATLAS site services, especially in Germany, were attempting to do large transfers of data residing on tape at BNL. This caused a high 'poll' load on the dCache service at BNL. Some actions were taken on the ATLAS side to lower the load. Simone sent a proposal how to proceed with this. Being discussed
    - ASGC: it seems that the bottleneck has been identified. Gang: distance between switch and routers is more than 10km. Had therefore introduced a component, ER, to compensate for this but it seems that it was not behaving as expected. (More details see below)

- CMS reports -

- ALICE - (Patricia) Smooth production. 8000 concurrent jobs running. Small issue observed two days ago at CNAF CREAM CE has been solved. Minor issues with some ALICE VOBoxes service in some French Tier-2s - a restart of the service was sufficient, no intervention required from the site.

Jeff: do know why your jobs stop at NIKHEF (last time a job was seen last Tuesday)? see jobs submitted by the gatekeeper and stay in the queue.

- LHCb reports -

Sites / Services round table:

- PIC (Gonzalo) Incident yesterday: site down from 14:00 - 17:00 due to an accidental power-off of the cooling system. Centre was gracefully shutdown but fast (jobs were killed). Afterward a couple WNs didn't come back, which Torque didn't cope well with and the batch system was degraded until ~midnight. A SIR is in preparation.
- ASGC(Gang) Slow transfer from CERN: before the fire, core switch and router were put in the same data center. After the fire they were separated, core switch was moved to new data center, and router was placed near to the old data center. The distance between them is more than 10 KM. This required to use a small component 'ER interface' which we are not familiar with. Actually we won't use it when we moved machines back to our previous data center. Today during the iperf tests our local admin found that this 'ER interface' does not work properly at certain frequencies. We have contacted the service provider for replacement.
- BNL (Michael) only incident we had was the one reported by Simone above. To recap: the issue is massive srmBringOnline followed by polling with srmLs, which causes high load on the namespace component. The polling is on a file-by-file basis, with a ~minute between polls. As temporary workaround the associated space (BNLPANDA) has been configured as no-tape storage class. Simone: the proposal is to extend the poll time.
- RAL (Gareth) at the moment a emergency intervention is ongoing at the RGMA service. CASTOR will be unavailable Monday - Tuesday next week while the backend db is moved to new hardware
- NIKHEF (Jeff) incident: user-analysis for people using GANGA 5.1.10 there is a bug if a site supports more than one type of SEs causes a high load on BDII and SRM. It seems that the client gets confused which SE to use. The solution is to upgrade to newer GANGA release. Simone hasn't heard about this problem but it seems specific to sites with multiple SE for same storage type, which is only NIKHEF/SARA.
- GRIF (Michel) following the success with WMS 3.2 all WMSs are now being upgraded.

AOB:

- Olof: In GDB LHCb presented file open failure rate of 21% at CERN but the members of CASTOR support team were not aware of this problem and would appreciate a GGUS ticket with few examples that could be investigated. Jeff said that the same was true for NIKHEF.

-- JamieShiers - 06 May 2009

This topic: LCG > WLCGDailyMeetingsWeek090511
Topic revision: r15 - 2009-05-15 - OlofBarring