Table of Contents

Week of 080602

Open Actions from last week:

Daily CCRC'08 Call details

To join the call, at 15.00 CET Tuesday to Friday inclusive (usually in CERN bat 28-R-006), 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-ccrc08/⬈
- 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(G.Mccance, S.Campana, R.Santinelli, H.Renshall, J.Andreeva);remote(D.Ross - RAL, M.Ernst - BNL).

elog review: 5 reports from LHCb all under investigation (no permission to write to a pnfs directory at SARA, CNAF has failures to write in a space token space and also in a shared storage area, LFC access is failing at FZK and PIC and finally gfal_ls fails intermittently at IN2P3). One report from ATLAS on Saturday of transfers out of CERN having stopped - see below.

Experiments round table:

ATLAS (SC): the transfer failures out of CERN on Saturday were due to two incidents. Firstly there was an internal ATLAS problem of a clash between the data export pool and movement of M7 data to another pool in CERN. The second cause was that the CERN power failure on Friday took down the M7 migration disk pool and it did not restart leaving entries in Castor and ddm that were catalogued but neither on disk nor tape. The ddm parsing does not proprly allow for this and it has being repeatedly retrying and generating many srm failures. The solution will be to delete the catalog entries pointing to the lost data. The current planning is finish transferring M7 data then, about midday tomorrow, switch to the FDR2 exercise and start transferring the 6 hours worth of generated MC data (about 17 TB).

LHCb (RS): A preliminary analysis of the May run issues has been entered in the observations log book. It is thought many of the dcache problems at IN2P3 and SARA are due to old versions of dcache and upgrades are being planned.

CMS (JA): have also been seeing the myproxy problems and are experiencing lower FTS transfer rates.

Sites round table: BNL 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.

Core services (CERN) report:

DB services (CERN) report:

Monitoring / dashboard report:

Release update:

AOB:

Tuesday:

Attendance: local(H.Renshall, J.Andreeva);remote(M.Ernst - BNL, J.Templon - NIKHEF, M.Jouvin - GRIF).

elog review: LHCb have closed three of their entries from yesterday (SARA SRM problem was an unscheduled downtime, CNAF failures to write fixed by SRM restart but in longer term need an SRM upgrade and IN2P3 gfal_ls problems expected to be fixed by dcache upgrade tomorrow) leaving only the LFC access problems at PIC and FZK open.

Experiments round table:

ATLAS: From K.Bos - Yesterday we switched to FDR2 and the data that was prepared successfully at BNL was uploaded into the SFO's. Let me remind you, this data was put together from Monte Carlo data into a cocktail that looks very much like data we expect from the detector in real life, with beams, with the trigger and with a certain luminosity in the machine and even with imperfections in the detectors. This data will be streamed from the SFO's into Castor at CERN and from here on we play the game as if these were real data.

In the T0 byte stream data from the various streams will be merged (5 to 1) into bigger files and then be written to tape at CERN and exported to tape in the T1's. This goes by data set (run) and stream and so again, initially there won't be enough to be sending data to everybody all the time so the transfer pattern will be spiky again. When the processing of those data in the T0 has been finished the derived data will also be exported: the ESD will go to the sites where the corresponding RAW went and the AOD and DPD will go to all T1's and then the T2's will get from the T1's what they requested. There is a delay between the RAW and the derived data because of the procedure for validation.

Addition from S.Campana - FDR raw data are being produced (not registered yet), while yesterday there was still a problem with the software to produce AODs. Therefore AODs and ESDs (and DPDs and TAGS and HISTSs) will come only late today/night.

LHCb: from R.Santinelli - This week is DIRAC3 week and we are draining the systems (we stop all transfers). It was agreed that priorities were to implement fixes in DIRAC as identified during CCRC and the migration to new h/w for central dirac services. Next week we will restart with all these changes in place (like the use of thread-safe gfal library, all LFC at T1s in place, download of input data and local in the WN process... Also discussions and preliminary investigations on xrootd are happening.

Sites round table:

NIKHEF: J.Templon asked why there were no LHCb jobs being scheduled at NIKHEF - the answer from Roberto is: Your question is not answered from the report I sent to Harry (as he assumed). Those are plans for this week and not a report of last week activity. Your question is rather answered by the (now famous) NEARLINE bug of dCache 1.8.0-15p4 that prevents the DIRAC stager service (issuing gfal_ls against input data to be processed by jobs in the WMS belly) to know whether the file is on the disk/cache or not. Having not this confirmation the payload for processing this data is not dispatched to pilots at NIKHEF and then not

CPU is consumed.

BNL: 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.

Core services (CERN) report:

DB services (CERN) report:

Monitoring / dashboard report:

Release update:

AOB:

Wednesday

Attendance: local(H.Renshall, S.Campana, J.Andreeva, G.Mccance, P.Mendez, R.Santinelli);remote(M.Ernst - BNL, D.Bonacorsi - CMS, D.Ross - RAL, J.Coles - RAL/Tier2, J.Templon - NIKHEF).

elog review: Nothing new.

Experiments round table:

ATLAS (SC): An ATLAS application has uncovered a bug that has appeared after a new linux kernel security patch was installed on many machines automatically after the Friday power cut. A rogue python process takes over the cpu and the only way out is to power cycle the machine. They are currently doing the 5 to 1 merge of FDR2 byte stream data and after finding small differences logged between the original and merged files are taking time to validate the merged files. They request CERN to set up a Tier0 to Tier0 FTS channel as they wish to make independent disk and tape copies of some files as needed by their ddm site services. Currently such copies are going via remote Tier-2 sites as the fallback channel. G.Mccance agreed to set this up.

ALICE (PM): Are having problems setting up the correct Alien environment on some sites. The detector commissioning exercise is nearly ready to start.

LHCb (RS): Are testing downloading of data to local worker node disk.

CMS (DB): Are finishing the last reconstruction activities of CSA08 and will be starting physics analysis. They are trying to increase the number of skim jobs a site can concurrently support and are preparing analyses of CCRC'08 CMS site experiences for next weeks post-mortem workshop.

Sites round table:

NIKHEF (JT): 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.

Core services (CERN) report (GM) : They are preparing data to allow users to analyse FTS performance during the CCRC May run. This involves extracting 1.6 million log files from the FTS database.

DB services (CERN) report:

Tuesday: 3

Monitoring / dashboard report (JA): 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.

Release update:

AOB:

Thursday

Attendance: local(H.Renshall, J.Andreeva, G.Mccance, R.Santinelli, J.van.Eldik, S.Campana, A.Sciaba);remote(M.Ernst - BNL, D.Bonacorsi - CMS, J.Templon - NIKHEF).

elog review: Two new entries - CERN made a supposedly transparent change to their production SRM entry points which stopped transfers for ALICE, ATLAS and CMS from 10.30 to 14.30 - CNAF have understood that the LHCb failures to write into STORM are due to hardware and will be moving a GPFS server to new hardware.

Experiments round table:

ATLAS (SC): At their daily 16.00 data quality meeting yesterday ATLAS decided to delay again releasing their fake raw data for reconstruction at CERN due to problems in the liquid Argin calorimeter part. They need to recalculate constants. The merge procedure (5 files into one) has, however, now been validated so the accumulated 3-days worth of raw data is now being registered in DDM and should start going out to the Tier-1 from about 16.00 CEST. The muon calibration streams (4 per day) are already going out to the concerned Tier-2 sites. If the data is released for reconstruction then AOD and ESD will start to be distributed from about midnight.

CMS (DB): Are analysing their CCRC'08 performance preparing for next weeks post-mortem workshop. They are having discussions on how to use VOMS roles in Castor at CERN. The CERN SRM failure today was unfortunate as they were trying to reach a long running metric on data transfers from CERN to FZK.

LHCb (RS): There is a very sustained transfer activity out of CERN (80-90MB/s) in the last 24 hours due to replication of DST stripped at CERN to Tier 1s.

Sites round table:

NIKHEF (JT): 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.Fuhrman.

BNL: 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.

Core services (CERN) report: J.v.E reported that the SRM upgrade went in at 09.30 and that FTS export traffic stopped about 10.30. The software was rolled back and the service resumed at about 14.30. The upgrade had been well tested but not all the features and it is thought the problem lay in tape recall - more tomorrow.

DB services (CERN) report:

Monitoring / dashboard report:

Wednesday   4

Release update:

AOB: RS, for LHCb, asked why they needed two separate (lightly loaded) VOBoxes to deploy the single xrootd based Castor service class. JvE replied there are two xrootd services - peer and manager - that need to be on separate boxes today though they are planning to remove this restriction.

Friday

Attendance: local(H.Renshall - WLCG, R.Santinelli - LHCb);remote(M.Ernst - BNL, S.Gowdy - CMS).

elog review:

Experiments round table:

ATLAS : FDR2 report from K.Bos - Last night we again had problems with the reconstruction. To not further overload the system (and the people) we did not simulate another fill today. This morning the reconstruction was fixed and now the Tuesday express-line data is being processed. So this is backlog. On schedule is the Thursday express line reconstruction which is also ongoing right now.

When the Tuesday express line reconstruction has finished, I assume we will go to bulk reconstruction of the Tuesday data. Officially we should wait until the 4 pm meeting but because we have discussed these results already 2 times and because it is Friday afternoon we will probably jump the gun and go ahead. We can still sign it off officially at 4:00. But this is of course up to the FDR coordinators to decide. This would mean we could expect the first ESD and AOD files later today.

The RAW of all 3 days were transfered yesterday, well within 20 hours and remind you this was 18 hours of data taking. All these files were big and that makes it easier but still this is a remarkable success. The efficiency to all T1's was in the high 90's and to Triumf, PIC and FZK not a single transfer had to be attempted twice. The data got distributed according to the shares defined in the plug-in Santa Claus and indeed the number of data sets in each T1 are what could be expected from this.

LHCb (RS): They are analysing performance of the May CCRC'08 for presentation at next weeks post-mortem workshop. On Monday they will restart the full LHCb chain using as many sites as possible. One objective is to compare the performance of data access on local worker node disks to that seen using the previous remote input/output.

Sites round table:

Core services (CERN) report:

DB services (CERN) report:

Monitoring / dashboard report:

Release update:

AOB: FIO has completed a post-mortem of yesterdays CERN srm failure and posted it on the Data Operations elog 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

-- JamieShiers - 29 May 2008

This topic: LCG > CCRC08DailyMeetingsWeek080602
Topic revision: r9 - 2008-06-06 - JamieShiers