

Table of Contents

Raw collection of issues.....	1
June 2.....	1
ALICE.....	1
ATLAS.....	1
CMS.....	1
LHCb.....	1
Sites.....	1
June 3.....	1
ALICE.....	1
ATLAS.....	1
CMS.....	2
LHCb.....	2
Sites.....	2
June 4.....	2
ALICE.....	2
ATLAS.....	2
CMS.....	2
LHCb.....	3
Sites.....	3
June 5.....	3
ALICE.....	3
ATLAS.....	3
CMS.....	3
LHCb.....	3
Sites.....	3
June 6.....	4
CMS.....	4
June 7.....	4
CMS.....	4
June 8.....	4
ALICE.....	4
ATLAS.....	4
CMS.....	4
LHCb.....	4
Sites.....	5
June 9.....	5
ALICE.....	5
ATLAS.....	5
CMS.....	5
LHCb.....	5
Sites.....	5
June 10.....	6
ALICE.....	6
ATLAS.....	6
CMS.....	6
LHCb.....	6
Sites.....	6
June 11.....	7
ALICE.....	7
ATLAS.....	7
CMS.....	7
LHCb.....	7
Sites.....	7
June 12.....	8

Table of Contents

Raw collection of issues

ALICE.....	8
ATLAS.....	8
CMS.....	8
LHCb.....	8
Sites.....	8

Raw collection of issues

June 2

ALICE

- Will run FTS transfers only during 2nd week.

ATLAS

- Tests are starting, including pseudo-reprocessing jobs reading tapes.
- Transfer failures from IN2P3 to ASGC *[why?]*.
- Low transfers from TRIUMF to CNAF *[why?]*.
- Also slow from TRIUMF to 2 Canadian T2 and from SARA to T2.
- Failures to Coimbra and Lisbon.
- Bug in a GFAL version included with DQ2, had to downgrade GFAL.

CMS

- FZK and IN2P3 will not participate to tests involving tape because of serious performance problems due to network at FZK, and of an HPSS upgrade at IN2P3. However, FZK was able to migrate to tape incoming AODSIM data written to the tape write pool, and copied to the tape read pool to be available for reprocessing jobs.
- FNAL and RAL have set tape families to import the AODSIM data, while ASGC, CNAF, FZK, IN2P3, PIC plan to import in a disk-only area.
- ASGC, RAL, PIC will use PhEDEx pre-staging agent; CNAF will use the SRM staging script executed by Andrea Sciabà; FNAL will do manual staging; IN2P3 will use manual prestage after HPSS downtime.
- ASGC cannot yet prestage due to a tape backlog.
- At RAL, prestaging happened on the wrong pool due to a mistake in the PhEDEx configuration.
- [DONE]

LHCb

- They start only in the second week. In the first week, preparation, including deletion from disk of files to prestage.

Sites

- NIKHEF (or SARA?) had LHCb SRM tests failing due to the dcache disk pool 'cost' function not load balancing correctly across full disk pools.

June 3

ALICE

ATLAS

- Data distribution going well.
- Soon will test xrootd analysis at CERN.

- Prestaging going well apart from NDGF where prestage request do not happen (see below for explanation). Even at FZK it went not bad, with a degraded tape system.
- Some backlog in writing to tape at NDGF.
- Want to check that at CERN data committed to tape is going to tape (tapes being recycled).

CMS

- PIC sees a prestage rate of 60 MB/s, the maximum possible with only two tape drives.
- RAL has a rate of 250 MB/s.
- CNAF has a rate of 400 MB/s.
- Low quality seen in T1-T1 transfer tests.
- At PIC, stager agent was set up to make more efficient WAN outgoing transfers for data on tape.
- [DONE]

LHCb

Sites

- RAL had some low level disk server failures.
- ASGC has now 5 working tape drivers, and they should be 7 next week. One tape drive has a hw issue.
- ASGC will upgrade CASTOR to 2.1.7-18 to fix the "big ID" issue.
- BNL had a site services problem affecting transfer performance due to a gfal library compatibility issue. After fixing this they observed up to 20000 completed srm transfers per hour, a very good rate. 16 drives were used for staging and 4 were reserved for migration. Data migration rates of 800 GB/hour from hpss to tape were observed at the same time as running merge jobs.
- NDGF had to increase FTS timeouts.

June 4

ALICE

ATLAS

- NDGF is slow getting data from CERN. It turned out to be a too low number of concurrent transfer jobs on FTS. Changed from 20 to 40.
- NDGF changed also other parameters to accommodate for long running transfers (see above).
- Problems transferring data to FZK due to a too low limit on number of file handles on a storage node.
- Some T2 sites had storage problems.
- Also log files from reprocessing jobs end up in tapes.

CMS

- ASGC did 190 Mb/s in prestaging. Some day00 files were requested by jobs before being prestaged, but this did not cause problems.
- RAL saw excessive network traffic due to not using by mistake LazyDownload. The 10 Gb/s link to the storage was saturated at 3 GB/s. Jobs had to be killed and recreated.
- Also at PIC LazyDownload was not active.
- ASGC prestage at 140 MB/s (day01)
- CNAF prestage at 380 MB/s. Difficult to monitor tape usage. StoRM had problems for transfers.
- At FNAL, concurrent prestaging and transfers did not cause problems.
- FZK could write (not read) with one tape drive. No estimate for fix of tape system.

- At IN2P3 HPSS progresses nicely. Even if downtime competes before the weekend, no tape activity is desired. Confirming that they will use manual prestaging.
- At RAL CMS is competing now with ATLAS. Tape monitoring available [?](#).
- T1-T1 transfers with CNAF bad due to PhEDEx not working well on sites with two SE endpoints (CASTOR and StoRM).
- [DONE]

LHCb

- About 7000 files were lost at CERN in the *lhcbdata* service class due to a CASTOR upgrade which turned on by accident the garbage collector (post-mortem).

Sites

- TRIUMF seeing FTS timeouts on transfers from RAL and ASGC and to CNAF especially for the large (3-5 GB) merged AOD files. Have hence increased their timeouts from 30 to 90 minutes.

June 5

ALICE

ATLAS

- Serious SRM problem at RAL for 3 hours [*which problem?*].
- Realized that due to larger data files, one needs to set longer FTS timeouts.
- At SARA prestaging stopped at midnight, due to problems with tape backend solved with a DMF reboot.
- It was discovered that NDGF did not have prestige requests due to a difference in the ARC architecture where there is an extra state in the workflow.

CMS

LHCb

- Cannot read non-root files via gsidcap at SARA and IN2P3 due to some dCache parameters controlling read-ahead which appear to be set to bad values, and not by LHCb software.
- GGUS tickets will be sent to T1 sites with the lists of files to delete from disk buffers.

Sites

- PIC is going to install four new LT04 drives on the STK robot, which will multiply by 3 the stage rate.
- In2P3 will start prestaging on June 8 for CMS and June 9 for ATLAS. A new script to optimize tape access is now in place.
- RAL increased FTS timeouts which eliminated transfer errors with BNL.
- TRIUMF increased timeouts, too.
- CNAF fixed some mispublishing of Storm disk pool information which was causing slow transfers TRIUMF-CNAF [*how is that possibly related?*]

June 6

CMS

- Prestaging at RAL ran with CMS and ATLAS together with no problem; CMS have 4 drives dedicated, ATLAS have 4 drives dedicated, LHCb have 1 drive dedicated.
- FNAL started to recover from tape problems. Traced back to relatively small queue depth in enstore's library manager before it overloads and the very large amount of transfers at FNAL. This increased the rate of "seeks" on tapes dramatically in combination with a 1 minute delay in reads which is 'normal' for non adjacent files on LTO4.

June 7

CMS

- Hardware failure on one CASTOR disk server at CNAF.
- FNAL is recovering tape performance.
- At RAL, still seeing more import rate via PhEDEx than tape writing rate although more tapes have been added. It seems that Castor is only writing to a tape then stopping. James found the problem, backfill was pointed to a tape family only attached to the farm service class, T0 data importing to the Import service class was never going to migrate, fixed now. Stages rolling in.

June 8

ALICE

- Some unspecified problem with FTD preventing the startup of transfers foreseen for this week.

ATLAS

- Over the weekend, reprocessing was smooth at PIC, RAL, BNL and TRIUMF.
- CNAF had some CASTOR problems meaning that it was slow but prestaging worked fine.
- SARA tuned DMF [*what is it?*] and storage R/W is better balanced.
- Last week's problem at NGDF consisting in prestage not being triggered looks like a configuration problem in Panda. Prod system doesn't realise files on tapes and hence doesn't trigger prestage.
- FZK still very slow.
- ASGC not usable [*for what?*]. Eva suggested 2nd listener for Oracle on old port.
- IN2P3 restarted tapes for writing, not yet for reading.

CMS

- FZK: tape writing is working, but there are still issues with tape reading.
- CNAF: stageout failures due to wrong permissions.
- FNAL: no problems in tape writing, up to 1.4 GB/s; problems in reading, being below what was achieved last week (and previous weeks): being investigated over the weekend. Recovered the needed tape read rate on Sunday but had a large backlog since then.

LHCb

- STEP09 starts tomorrow for LHCb.
- 6500 files lost from CASTOR.

- Investigating problem at NIKHEF preventing from reading non-root files via gsidcap.

Sites

- CERN: Post mortem on LHCb data loss (link1 link2).
- FZK: working on tape problem. Fixed config errors. Newer version of TSM. Consolidated staging - saw improvement over w/e. Still looking to find bottleneck. Some of gridftp doors ran into memory problems - investigating it.
- IN2P3: reading from tape works fine since the end of the HPSS intervention, but also ATLAS was hit by the srm-ls locality problem (Fabio's email).

June 9

ALICE

- Began some transfer tests using a new FTD instance.

ATLAS

- Prestaging works fine at NDGF after fix in Panda.
- FZK healthy. Site reports that analysis activity, using excessive lcg-cps to stage data, was loading the gridftp doors. After this was stopped the SE has performed well.
- Prestage working fine at IN2P3 after fixing srm-ls bug.
- ASGC: good news is that late yesterday listener on old port working. Prestaging is slow, though.
- BNL has an increasing transfer backlog. They had to reboot their SRM early this morning.
- RAL reported that they are not writing to tape.
- T2s: quite a few with backlog (~12 sites >24h). Check for lack of FTS slots, slow transfers hogging slots, overload from analysis.

CMS

- A file corrupted after transfer to ASGC.
- CNAF: prestaging slow, and first files started to arrive only after 2.5 hours, due to other concurrent activities. One file never came online according to statusOfBringOnline while it was for Ls. At risk tomorrow due to late delivery of cleaning tapes.
- FZK ready to start tape activities.
- IN2P3 has issues at the pool level due to a bug. The workaround is to reboot the pools. A possible fix is to upgrade to a more recent version of dCache. *[find out more if possible]*
- At RAL, fixed issue with CASTOR migrator. *[see Facops HN for details]*

LHCb

- FZK: many users complain and production activities report problems accessing data through dcap.
- NL-T1: problem accessing raw files still being investigated.
- IN2P3 also had dcap problem. Discovered some limitation in dcache client library. New unreleased version recommended.

Sites

- BNL discovered that 3 of 4 write buffer servers are idle. Under investigation. Result is BNL not keeping up.

- IN2P3 introduced a new component scheduling stager requests before sending to HPSS. Started yesterday for CMS - very happy with initial results, but must analyse further. Reduction of number of tape mounts is the purpose of this new component. Being activated for ATLAS.
- ASGC: streams replication is working fine now.

June 10

ALICE

- NDGF: problem was with destination path in FTD. After changing was ok.
- RAL: site admins needed to fix write permissions.

ATLAS

- IN2P3 dCache went down at 0100UT. We sent a team (49389) then alarm (49392) ticket and the situation was resolved at about 0800UT. IN2P3 pre-staged really nicely, then were killed by their dCache issues (stage-out failed). Now recovering.
- SARA_MCDISK almost ran out of space.
- RAL reported severe tape robot problems this morning. Looks like it's fixed now.
- CNAF have reported drives disabled awaiting fresh cleaning cartridges, so degraded tape performance.
- Looks like global rate limitation in ATLAS pool used for export from CERN: 4GB/s: Simone - anything above 3GB/s of total I/O traffic "challenging". Miguel: pool configured to much lower rate than now delivering - don't expect to give > 3GB/s unless we change h/w.

CMS

- Only two tape drives available at RAL because of a stuck cartridge. Rebooting the robot recovered all drives apart from the stuck one. Performance was enough to absorb the backlog.
- Could not transfer AODSIM to FNAL because their FileDownloadVerify agent failed after the PFNs were changed to avoid putting too many files in a single directory. Solved by reverting to the standard PFNs.
- IN2P3: Since 3:00 AM this morning all gridFTP transfers were aborting. This was due to a failure with the component in charge of routing requests to gridFTP servers. The corresponding service has been restarted and Dcache is full available since 11:00 AM.

LHCb

- At CNAF, stage attempts failed due to the problem with the cleaning cartridges.
- Investigation with Jeff has discovered a middleware issue with dcap plugin for root. Enabled in DIRAC the configuration to read data after a download into the WN (ATLAS approach). Running smoothly now.
- LFC issue: due again to Persistency interface.
- MC_DST and DST space tokens at CNAF. It seems that the space token depends on the path while StoRM should guarantee the independence between space token and namespace path.

Sites

- FZK made available a post-mortem for their tape system problems. The main problem was an unstable SAN connectivity, but also memory problems in the disk servers. Transfers seem better now.
- IN2P3: a problem with dCache occurred this morning at IN2P3-CC. Since 3:00 AM this morning all gridFTP transfers were aborting. This was due to a failure with the component in charge of routing requests to gridFTP servers. The corresponding service has been restarted at 10:30 AM and dCache is

fully available since 11:00 AM. (Probably due to some load - many pending requests).

June 11

ALICE

- Having problems with the `-o` option of *glite-transfer-submit*, as the "File exists" error message is received.

ATLAS

- Geneva-Frankfurt fibre cut. Due to rerouting, ASGC have a serious problem with `gridftp_copy_wait`: Connection timed out from other T1s and from CERN.

CMS

- OPN cut impact on CMS, mainly links to/from ASGC.
- At IN2P3 the Tier-2 was closed because the analysis jobs caused files to be staged from HPSS and thus interfering with Tier-1 operations. This was a known issue and some actions were already identified for strictly separating the storage spaces of the tier-2 from those of the tier-1 and to modify the configuration of the tier-2 to make it a disk-only site. Unfortunately, this could not be finished before the beginning of STEP'09.

LHCb

- STEP09 Reprocessing and staging confirmed that Persistency interface to LFC needs urgently to be updated. A lot of jobs failing to contact LFC server because brought down by the inefficient way Persistency queries it. A Savannah bug open (ref. <https://savannah.cern.ch/bugs/?51663>). Like for dCache client issue reported days before, this is another important middleware issue that prevents to continue the exercise STEP09. We cannot process data w/o ConditionDB and then we will kill currently running jobs. LHCb will ask again sites to wipe data from the disk cache and rerun the reprocessing using SQLite slices of the Condition DB (instead of using it directly) in order to compare the two approaches.
- Staging at SARA. All files to be reprocessed failed after 4 attempts
- File access problem on CASTOR at CNAF.
- Issue transferring to MC_DST at Lyon. Disk really full.

Sites

- RAL: Still have 78k files refusing to budge. After running last night, the migrator process dealing with these started impacting the server, taking >90% of memory and still using 20%+ CPU time with no sign of a migration. I have restarted the mighunters again just to free up resources on the machine, and will continue to monitor (see here).
- FZK: stage throughput cut in half for CMS due to a stager down.
- FZK: Failing stageout for ATLAS: one gridftp 'door' (dCache) had to be restarted because out of memory. One library down for 2 hours. Fixed. Currently some drives used for writing are not accessible because of SAN zoning problem (we've hit a FW bug).
- BNL: SRM server issue: overloaded due to analysis jobs staging out output files via `lcg-cp` so to write to the correct space. Need a version of `dccp` capable of that. Developers say dCache 1.9.2 but becoming an immediate and urgent issue.
- BNL: Networking: big fibre cut in Frankfurt area affecting also US. For BNL bandwidth reduced from 10 to 4.2 Gbps.
- Network: Two independent fibre cuts .

- ASGC: a problematic LTO4 drive has been fixed and all new tapes are now functional.

June 12

ALICE

- The problem with FTS reported yesterday was due to not using fully qualified URLs to force SRM 2.2 (SRM 1 does not have an overwrite option). After using proper URLs the problem was solved.

ATLAS

- Many jobs died at FZK due to a gridftp door problem.
- Problem with storm at CNAF overnight - fixed this morning.

CMS

- FZK tapes fine now, but the number of recalled files marked as being on disk was still quite low. This was due to two reasons:
 1. One of the 2 CMS stager nodes had limited throughput due to copying data to itself (stager runs on same as for some other pools).
 2. Config error on the stager pools. Majority of the data staged back was copied to the tape write pools instead of the tape read pools. (Some behaviour which seems to show up under heavy load of the system).

LHCb

- During the weekend, staging and reprocessing data will be done w/o the remote ConditionDB? access (that currently would require a major re-engineering of the LHCb application). but via available SQLite CondDB slices.
- Staging at SARA issue seems related to the fact that all directories to be staged were also removed from tape.

Sites

- GRIF: observed with ATLAS lcg_cp timeouts at all Fr sites (Hammercloud tests). Suspect not site nor load - maybe some race condition? Similar problem reported also with STORM during pre-GDB. Investigating this.
- FZK: SAN switch reset after contact was lost to 8 LTO3 drives. Datasets from LHCb were unavailable. Some data from CMS. Restarted movers that stopped because of this outage. Restarted a mover for ATLAS which was stuck for more than 5 hours. Trigger went unnoticed because of intense other activities. Rates picked up to >100MB/s after restart. Network problem on a CMS stager node limits throughput. Possibly the node is copying data to itself because the stager runs on the same machine as some other pools. Config error on the stager pools. Files staged were copied to tape write pools. Exact life of a file to be determined. Also a CMS pool was recalling data but then never showed them when they arrived. Restart fixed this. 3 of 4 gftp proxies (doors) locked up. Cause unknown. Restarted. We have now 2 more stager nodes ready that contain no other pools. Experiments please keep the load. GridKa will fight to keep up the traffic and fix problems.

-- AndreaSciaba - 05 Jun 2009

This topic: Sandbox > Step09ExpIssues

Topic revision: r13 - 2010-06-11 - PeterJones



Copyright &© 2008-2021 by the contributing authors. All material on this collaboration platform is the property of the contributing authors.
or Ideas, requests, problems regarding TWiki? use [Discourse](#) or [Send feedback](#)