

-- JamieShiers - 22 Mar 2006

June 2006

- 12/06 TRIUMF, IN2P3, FZK, PIC and SARA reply confirming disk and tape end-points for ATLAS.
- 07/06 RAL reply confirming disk and tape end-points for ATLAS T0-T1 transfers
- 05/06 Due date for setup for ATLAS T0-T1 tests (19/06 - two weeks)
- 01/06 00:00 UTC The official start date of the WLCG pilot / SC4 Service Phase!

May 2006

- 29/05 12:00 SARA announce service downtime extending from 12:00 today until Wednesday 18:00. According to WLCG procedures, this intervention should have been announced one week in advance (See WLCG Management Board minutes of 16 May. Jamie (Announcement: srm.grid.sara.nl and ant1.grid.sara.nl will be down for maintenance from 12:00 CET today to 18:00 CET wednesday. This is due to a modification to our storage infrastructure.)
- 19/05 02:00 ASGC stable at 60 MB/s, even 70 MB/s for the last 8 hours. CNAF came back at 8 GMT and averaged ~180 MB/s except for a 2-hour dip. NDGF did 30 MB/s until 8 GMT, then averaged 70 MB/s except for a gap of a few hours. PIC stable at 50-60 MB/s except for a 3-hour gap due to a misconfiguration in their stage pools. RAL rose from an average of 70 MB/s to an average of 150 MB/s since 9 GMT, the pattern remaining bumpy, the error rates lower. SARA may have been left on accidentally. Their average rose from 40 to ~70 MB/s at 12 GMT. *Maarten*
- 18/05 01:20 CASTOR came back in the morning and around 12 GMT all channels were set active, even the ones that should have been left off. This was corrected a few hours later, after which a problem with the CASTOR DLF service caused all transfers to fail on SRM GET. This was fixed by Miguel, and the channels were re-enabled after the Champions League final. 😊 A problem was then spotted for one of the gridftp servers at PIC. *Maarten*
- 17/05 13:45 There was a problem in a Surfnet's device at CERN. Since few minutes the connections to RAL, TRIUMF, ASGC and SARA are up again. *Edoardo*

On 05/16/06 15:18, Bly, MJ (Martin) wrote: > *We understand that the OPN link to RAL is down due to a fibre cut somewhere between France and Belgium. _Martin_*

- 17/05 01:30 Business as usual until about 12 GMT, when a CERN-wide power cut stopped everything in its tracks. CASTOR hopefully will be back Wed. morning. All channels have been switched off. The one for DESY will remain off to avoid interference with a big CMS data replication. *Maarten*
- 16/05 01:20 ASGC doing 50 MB/s most of the time, with instabilities due to kernel crashes, possibly related to XFS. The latest CERN kernel will be tried next. CNAF averaging 170 MB/s. DESY averaging 160 MB/s. NDGF had a dip lasting 9 hours due to transfers stuck on a pool node and possibly some other issues. The channel did 80 MB/s for the last 2 hours, maybe to try and make up for the bandwidth lost? RAL came back around 12 GMT and did about 150 MB/s for a few hours, then sunk to 70 MB/s for the last 4 hours, with an error rate that has peaks and valleys. *Maarten*
- 15/05 00:30 ASGC recovered from yesterday's problems and did 50 MB/s or better the last 8 hours, with a very low error rate. CNAF averaged 170 MB/s until 15 GMT and 150 MB/s from then on. DESY doing 160-170 MB/s. NDGF very flat at 60 MB/s with *zero* errors! *Maarten*

- 14/05 04:30 ASGC averaging 50 MB/s, but most requests failing due to the recurring problem that one of the round-robin SRM services crashes and is not automatically restarted. CNAF averaging 170 MB/s. DESY averaging 160 MB/s. NDGF stable at 60 MB/s. FZK got to 100% error rate due to no pool being available. RAL got to 100% error rate due to the gridftp door nodes immediately closing every connection. *Maarten*
- 13/05 00:20 ASGC stable at 50 MB/s. CNAF averaging 170 MB/s over the day, 180 MB/s the last 5 hours. DESY as usual showing the remarkable battlement pattern averaging 170 MB/s. FZK stable at 50 MB/s. NDGF flat at 60 MB/s. RAL rocky until about 12 GMT, then rising to ~140 MB/s with a dip to 100 MB/s around 17 GMT, all with lower error rates. *Maarten*
- 12/05 02:00 ASGC stable at 50-60 MB/s. CNAF averaging 160 MB/s. DESY averaging 160-170 MB/s. FZK came back and has been stable at 50 MB/s since 11 GMT. NDGF stable at 60 MB/s. RAL bouncing between 10 and 100 MB/s with a high error rate. The GridView histograms stopped getting filled around 14 GMT, but by 16 GMT the developers had fixed the problem and the gap then disappeared. *Maarten*
- 11/05 01:00 ASGC doing 50-60 MB/s most of the time, with some instability due to a node that will be replaced (possible HW problem). CNAF came back around 11 GMT and averaged 170 MB/s for the next 11 hours, with an error rate that is not so small, though. DESY had a rough day but still managed to average 130 MB/s. NDGF doing 60 MB/s except for a 2-hour dip to ~40 MB/s. RAL discovered that their rate went down from 150 to 50 MB/s due to the gridftp door nodes having auto-negotiated themselves to 100 Mbps after the network maintenance. Since 16 GMT they have averaged 110 MB/s for the next 6 hours. The GridView developers found that a bug in their code was the cause of yesterday's peculiar gap from 4 to 5 GMT. Investigating a 10% discrepancy between the CNAF network statistics report and the rates shown by GridView we established that GridView takes "MB" to mean $1024 * 1024$ bytes, which is OK, but needs to be taken into account when comparing with other reports (the remaining 5% was due to network protocol overhead). *Maarten*
- 10/05 23:00 Michael Ernst wrote:

>
> *Hi Maarten,*
> *it's been a rough day and I expect you have seen a good*
> *number of failed transfers over the course of the last*
> *12 to 15 hours. Instabilites were caused for the following*
> *reasons*
>
> *1) Our NREN provider applied maintenance to the border*
> *router which has caused loss of connectivity between*
> *9PM and 10PM*
>
> *2) The loss of external connectivity has caused our Firewall*
> *to crash ('was told Cisco knows about this bug) which*
> *extended the outage. Though the path through the FW is*
> *not used for data transfer the management related*
> *communication is affected.*
>
> *3) In addition we have had LAN problems in the morning, also*
> *affecting communication between the admin and the pool*
> *nodes*
>
> *We have never dropped out completely, but have seen a significant*
> *reduction in the throughput compared to the previous days.*

>

> *Just wanted to let you know,*

>

> *Michael*

- 10/05 02:00 ASGC recovered during the night and ran at a stable 60 MB/s. CNAF ran at 100 MB/s until 4 GMT, then falling to zero by 5 GMT. DESY had a glitch or two, but still averaged 150 MB/s. IN2P3 was averaging 230 MB/s, then set inactive around 13 GMT. NDGF oscillating around 50 MB/s average. RAL off during the night, then came back around 9 GMT, but hardly reached 50 MB/s with the same parameters as before, now with a very low error rate, so there is a new bottleneck somewhere. GridView displays a gap of zero activity between 4 and 5 GMT, but the FTS logs show that transfer rates were very similar to those during the hours right before and right after the gap. The logging to the MON box shows no problems and on the latter node I did not find any anomalies either. I asked the GridView developers to look through their logs and inspect their DB. *Maarten*
- 09/05 02:20 ASGC doing a stable 60 MB/s until about 19:30 GMT, when SRM put timeouts brought the rate down to zero, prompting further work on their cluster. CNAF came back around 14 GMT, rose to ~150 MB/s, then fell back to a stable 100 MB/s, but with a high error rate. DESY averaging ~170 MB/s as usual. FZK had most transfers failing and dropped into GridView's "others" category. IN2P3 averaging 240 MB/s. NDGF experiencing some hiccups that brought their rate down from 60 to 40 MB/s. RAL switched back to disk at 15 GMT and averaged 150 MB/s until the channel was set inactive in anticipation of network maintenance. *Maarten*
- 08/05 02:30 ASGC doing a constant 60 MB/s until about 14 GMT, when a stager had become overloaded with processes cleaning up the SC4 data. A fix was applied and from 20 GMT onward the rate was 60 MB/s again. DESY nicely averaging 170 MB/s. FZK doing 60 MB/s during the night, then suffering instabilities throughout the day. IN2P3 very nicely averaging 240 MB/s. NDGF had a dip for a few hours, doing 60 MB/s for most of the day. RAL had a more pronounced dip for a few hours, still averaging 40 MB/s, with significant error rates. *Maarten*
- 07/05 02:20 ASGC doing a very stable 60 MB/s most of the day. CNAF at 200 MB/s until 17 GMT, when the error rate started increasing sharply, causing all transfers to fail (EBUSY) from 21 GMT onward. DESY stable at an average of 170 MB/s. FZK doing 50 MB/s or more most of the time, but with various gaps; the error rate was not high, so requests must have been held up at the destination. IN2P3 came back at 13 GMT and averaged 250 MB/s, except between 21 and 23 GMT as usual, when many SRM put requests timed out due to the daily PNFS backup. NDGF OK at 60 MB/s. RAL averaging 40 MB/s with alternating periods of low and high error rates. *Maarten*
- 06/05 02:20 ASGC was doing 100 MB/s for many hours, but with many errors. Around 12 GMT the number of files was halved to 30, to make room for CMS PheDEx transfer tests. The SC4 rate now is a stable 70-80 MB/s and the error rate became negligible when an upgraded gridftp server was removed from their CASTOR cluster. CNAF got to 100% error rate (name server DB) just after midnight and the channel was off until 12 GMT, when it quickly rose to a fairly constant ~190 MB/s average for at least 11 hours in a row. DESY fairly stable at ~170 MB/s. FZK fairly stable at ~60 MB/s (tape) throughout the day, then got to 100% error rate (pool request timeouts) early in the night. IN2P3 was doing about 200 MB/s, then set the channel inactive to clean up the dCache DB and tapes. NDGF suffered disk faults on one pool node, which caused a slight dip for a few hours, otherwise stable at 60 MB/s. RAL was unstable until 12 GMT and at a flat 50 MB/s from then on, with a much lower error rate. TRIUMF was fairly stable at 50 MB/s until 12 GMT, when an end was put to their disk-tape transfers. Yesterday's dip between 4 and 6 GMT turned out to have been caused by a problem in the LCG DB server cluster, slowing down the FTS. *Maarten*
- 05/05 00:40 ASGC averaged 100 MB/s, but less constant than before, and with a high error rate that is not yet understood. CNAF have really started the tests of their upgraded CASTOR-2 system and

so-far ran at ~180 MB/s (disk) for 6 hours in a row. DESY averaging 160 MB/s, fairly stable. FZK were averaging 150 MB/s to disk, then switched to a single tape drive for now, which should handle ~30 MB/s. So-far the transfer rate has averaged to 50 MB/s, but maybe the data is steadily accumulating on the buffer disks. IN2P3 averaging 200 MB/s, with a 1-hour dip right after midnight due to many SRM PUT timeouts. NDGF stable at 60 MB/s. RAL averaging 40 MB/s with a high error rate. TRIUMF stable at 50 MB/s. Between 4 and 6 GMT there was a significant dip affecting all sites. There was no increase in the error rates, so it seems most transfers were hanging in CASTOR, e.g. due to LSF or some DB operation being stuck. We have seen such a dip on other days at the same times, but not every day. *Maarten*

- 04/05 01:50 ASGC back to 120 MB/s, but many requests immediately fail as their connections are reset on the SRM put. DESY now at 150 MB/s (!), still using the same FTS parameters, but their link got upgraded from 1 to 10 Gbps! They have also done tests with FZK and reached ~300 MB/s. FZK still bouncing between ~150 and 200+ MB/s. IN2P3 bouncing between 150 and 240 MB/s. NDGF stable at 60 MB/s. PIC channel switched off after successful demonstration of target rates still using CASTOR-1. They will try and boost the upgrade to CASTOR-2 now. RAL at 50 MB/s, with fewer errors, but still too many. TRIUMF stable at 50 MB/s. During the day there were many SRM get failures due to modifications in the CASTOR SC4 cluster, which is being reduced to half the original size, allowing CASTOR instances for experiments to be enlarged, while still providing sufficient disk servers for the SC4 "background" traffic, which can amount to 1 GB/s or more continuously. There also were many transfer failures due to a stuck tape recall. The third cause for the big gap in the middle of the day was a rare problem in the FTS: a single stuck transfer managed to block four channels in the DB. We investigated the matter with the debugger, but could not understand how the process could have got itself into the situation observed. In the end we saved the stack traces, logfiles etc. and killed the process, after which normal operation resumed. The GridView developers introduced requested enhancements in the display options, with more to come. *Maarten*
- 03/05 00:30 ASGC having many instabilities due to multiple nodes having crashed, with subsequent investigations and cleanup. FZK at ~200 MB/s for the first half of the day, then sinking to below 150 MB/s. In the last few days a periodicity has been observed with a dip about every 6 hours, not yet understood. It does not seem likely that the matter is related to the 6-hour update cycle for the gridmap or the CRLs. IN2P3 doing 200 MB/s or better during the first half of the day, then falling below 150 MB/s for a few hours. Both for FZK and IN2P3 the cause may well have been in the CASTOR setup, because there have been many more SRM GET timeouts and suspicious transfer timeouts than usual. NDGF fell to 20 MB/s during the night, then climbed back to 50-60 MB/s. INFN have started testing their upgraded CASTOR-2 system at low rate. PIC fixed their SRM timeout problem and climbed to 50 MB/s, then got hit by a full file system causing most transfers to fail. RAL became reachable again via the OPN, but their error rate remains high. TRIUMF stable at 50 MB/s. *Maarten*
- 02/05 07:30 IN2P3-CC switched off one disk server to see the impact on the rate. Still 3 disk servers and 40 files in the channel. *Lionel*
- 02/05 00:30 ASGC had a 1-hour dip with many SRM timeouts, otherwise doing 100 MB/s or better. BNL were doing 90 MB/s, then ran out of tape and decided to switch the channel off for the time being, given that the first disk and tape phases of SC4 have ended. FZK had a 1-hour dip to 120 MB/s during the night, a few dips to about 200 MB/s, running at about 240 MB/s most of the time. IN2P3 doing 250 MB/s or better most of the time. NDGF dropped to zero during the night due to no write pool being available, then came back to a steady 60 MB/s. PIC still at 20 MB/s with many SRM timeouts. TRIUMF stable at 50 MB/s. *Maarten*

April 2006

- 01/05 02:20 ASGC OK at 120 MB/s. BNL stable at 90 MB/s. DESY at 70 MB/s, then set inactive at 19 GMT in preparation of high-speed transfer tests with FZK. FZK/GridKa averaging 230 MB/s, doing 240 MB/s or better most of the time, falling to 200 MB/s a few times per day. IN2P3 averaging about 250 MB/s, with a drop to 200 MB/s between 21 and 23 GMT, just like yesterday, possibly due to a daily backup or so. NDGF OK at 60 MB/s. PIC still at 20 MB/s due to many SRM timeouts. TRIUMF OK at 50 MB/s. *Maarten*
- 30/04 02:20 ASGC stable at 120 MB/s. BNL doing 90+ MB/s. DESY 70 MB/s. GridKa doing 250 MB/s or better most of the time, but occasionally falls slightly below 200 MB/s. IN2P3 slightly above 250 MB/s most of the time, but occasionally dropping to about 200 MB/s. NDGF stable at 50 MB/s. PIC at one third of their usual rate due to many SRM timeouts. RAL dropped to zero around 9 GMT due to a problem with the OPN. TRIUMF stable at 50 MB/s. *Maarten*
- 29/04 02:30 ASGC essentially stable at 120 MB/s the whole day! Further improvements to their CASTOR-1 setup are planned to reduce the memory consumption. The upgrade to CASTOR-2 will be prepared in the coming months. BNL averaged 100 MB/s over the full day. DESY at 70 MB/s as always. GridKa exceeded 200 MB/s for 10 hours in a row, with 4 hours at 250 MB/s or higher! This became possible after various changes in their setup, most notably a fix of the disk striping configuration. Occasionally some transfers get stuck (also seen by IN2P3) and the pool node selection is not always balanced; both issues are looked into by the dCache developers. IN2P3 have been at 200 MB/s or higher essentially the whole day, peaking at 270 MB/s! This became possible after another disk server got added and the number of files was increased. We want to demonstrate stable running at about 250 MB/s, comfortably above the 200 MB/s nominal rate. NDGF stable at 50-60 MB/s. PIC stable at 70 MB/s, with a high number of errors all due to a single misconfigured gridftp server. RAL has climbed to 50 MB/s, still with a high error rate, though. TRIUMF stable at 50 MB/s. *Maarten*
- 28/04 00:30 ASGC did further work on their system and have been at 100 MB/s or higher for 14 hours in a row already, most of the time at 120 MB/s! BNL back at 100 MB/s after maintenance, but with instabilities. CNAF at 100 MB/s for many hours, then gradually crumbling due to the usual problem with their old CASTOR-2 version. DESY outright flat at 70 MB/s. GridKa at 200 MB/s or higher for 5 hours in a row, then falling back to 170 MB/s for a while, then up to 190 MB/s, still only using 20 files with 1 stream each. IN2P3 switched back to disk and ran at 200 MB/s or higher for 7 hours in a row, then fell back to 180 MB/s; they are using 25 files with 5 streams each. NDGF back doing 50 MB/s or higher. PIC fairly stable at 70 MB/s. RAL doing 30 MB/s with a high error rate. SARA working on their system, doing occasional tests. TRIUMF stable at about 50 MB/s. *Maarten*
- 27/04 14:00 IN2P3-CC switch back to disk-disk transfers with 25 files, target=200MB/s *Lionel*
- 27/04 00:30 ASGC climbed to 70 MB/s after downgrading gridftp server rpm, upgrading one of the pool nodes to the 2.6 kernel again, and tuning the TCP window size. BNL averaging 80 MB/s, with some instability and HPSS maintenance work. CNAF climbed back to 100 MB/s, but are ready to start the upgrade to the latest CASTOR-2 version, which should make their setup stable without frequent admin interventions. DESY at 70 MB/s as usual. FNAL channel set inactive for now after end of first SC4 tape test phase. GridKa analyzing and reconfiguring their setup to try and get to 200 MB/s and higher; the first few hours were again above 150 MB/s with a peak of 180 MB/s, but then the rate tumbled again to 90 MB/s for unknown reasons; we only have 20 files with 1 stream each in the channel at the moment, so we have some room to get comfortably above 200 MB/s. IN2P3 averaging 90 MB/s. NDGF back at 50 MB/s. PIC stable at 70 MB/s, RAL at 30 MB/s and TRIUMF at about 50 MB/s. New DB cleanup script deployed in FTS, which probably caused the temporary dip between 15 and 16 GMT. *Maarten*

- 26/04 01:40 ASGC had an instability for a few hours, then got back to a stable 50 MB/s. BNL had many gridftp timeouts causing the rate to plunge. DESY did some configuration changes, then got back to their usual 70 MB/s. FNAL stable at 80 MB/s. GridKa decreased from ~140 MB/s to ~90 MB/s for unknown reasons. IN2P3 did some changes and climbed back to 100 MB/s. PIC problems were due to the wrong endpoint being used, viz. a slow node that could not really handle the load; now they are at 70 MB/s with a low error rate! RAL did some changes and are at a stable 40 MB/s now. SARA working on their system, reached more than 50 MB/s for a few hours. TRIUMF very stable at 50 MB/s. CASTOR gave many SRM and gridftp timeouts between 14:00 and 16:00 GMT. FTS srmCopy agent file descriptor leak was due to unintentional rpm downgrade. *Maarten*
- 25/04 09:00 CERN-IN2P3 channel closed between 7 GMT and 9GMT for maintenance *Lionel*
- 25/04 00:30 ASGC stable at 50 MB/s throughout the day. BNL at 90 MB/s. CNAF recovered around 13 GMT and averaged 90 MB/s so far. FNAL active again since 9 GMT and immediately shot to 200 MB/s (to cache) with the parameters that gave 100 MB/s to tape before the weekend, so it seems that something has changed in the dCache configuration at FNAL. Even with 30 files the rate still remained 140 MB/s. Between 17 and 18 GMT the rate dropped to zero due a file descriptor leak in the FTS srmCopy agent: it could be the same leak we experienced before, but then the FTS rpms must have got downgraded somehow; the developers have been alerted. PIC investigating the problems affecting their channel. RAL back at a stable 30 MB/s. SARA back around 40 MB/s. TRIUMF reached 50 MB/s after increasing the number of files. IN2P3 a bit bumpy, averaging 90 MB/s. *Maarten*
- 24/04 16:00 Increased the number of files from 10 to 20 for IN2P3 and increased the number of movers from 5 to 10. LHCb transfers are using the same pools as dteam and their files are smaller (700MB), that could explain the drop at 13:00 GMT *Lionel*
- 24/04 02:00 ASGC very stable at 50 MB/s with low error rate! RAL suffered the same problem recently experienced by NDGF: today's automatically created directory for SC4 files was owned by root and hence unwritable, causing all xfers to fail when yesterday's backlog had been fully processed. SARA suffering from problems with the network, dCache or yet something else, causing all xfers to get aborted by the CASTOR gridftpd after 20 minutes of inactivity. TRIUMF stable at 40 MB/s now with a somewhat lower error rate. BNL stable around 90 MB/s. IN2P3 fairly stable around 100 MB/s. PIC climbing to 60 MB/s. NDGF channel set inactive because of problem in their dCache setup. DESY very stable at 70 MB/s, GridKa at 140 MB/s. *Maarten*
- 23/04 04:00 On Sat. I tried various settings for the ASCC and GridKa channels, varying the number of files and/or the number of streams in both directions. For ASCC I did not manage to get a significant performance increase and in the end I left the channel at 60 files and 20 streams, giving a stable 50 MB/s. I tried to vary the TCP buffer size, whose default setting is 2 MB, but the CASTOR gridftp server always logs 2 MB for any transfer, as if the request for a different size is ignored, even though the response is affirmative. In the code I did not yet see how this situation can arise. For GridKa I tried many settings, but never got more than 150 MB/s, which is already achieved with 20 files at a single stream each. With 10 files the rate dropped to 100 MB/s. A second stream did not make a difference. We will have to check if the dCache configuration has a limit on the number of active transfers. If not, then it really looks as if Geant/DFN/... has a cap of about 150 MB/s for the channel. NDGF has developed many timeouts and other errors during the day. BNL at a stable 90 MB/s. CNAF at 100% error rate. IN2P3 has a low error rate, yet has highs and lows; in particular the drop at 12 GMT is not understood. PIC has recovered and reached 40 MB/s. SARA was able to take up to 50 MB/s. TRIUMF at about 40 MB/s, but with a high timeout rate. RAL averaging about 30 MB/s. DESY at a solid 70 MB/s. *Maarten*
Note from IN2P3-CC: the drop at 12 GMT is understood: the night before, the migration process was stuck on both pools on one of the disk servers ("ccxfer10"), so the buffer was half full in the morning. To allow the disk to be emptied, I have disable access to "ccxfer10" for incoming transfers. For 3

hours, all transfers went to one single host ("ccxfer09") that is why the rate dropped. However, why the migration process stopped is not understood. *Lionel*

- 22/04 01:00 During the night there was again a dip between 1 and 3 GMT. Then between 7 and 9 GMT there were many CASTOR errors due to an unavailable file system and at the same time tape recall being stuck after an intervention on the IBM robotics two days ago. Between 11 and 14 GMT the FTS was out of service after an unexpected hiccup in the deployment of the regular DB cleanup procedure. Between 15 and 17 GMT the rates fell as the load generator had not yet been restarted. SARA found a bottleneck in their SAN system limiting them to 30 MB/s to tape for the time being, with short- and long-term solutions being investigated. FNAL channel set inactive around 20 GMT because of maintenance work on the tape system over the weekend. NDGF very stable at 50 MB/s. ASCC becoming more stable at about 50 MB/s (disk), but with a large number of concurrent transfers; we will try out various settings to identify the bottlenecks and find an optimal working point for the current setup. We will also do that for GridKa, which seems to be limited to 150 MB/s (disk) while we easily reached > 200 in the past. PIC has many gridftp transfers timing out after half an hour because of very low data rates. CNAF still suffering SRM timeouts and EBUSY. RAL was doing about 50 MB/s, then set the number of streams to just one, giving some 25 MB/s. BNL rates recovering after dCache upgrade. IN2P3 averaging about 80 MB/s. DESY very stable at 70 MB/s (disk). TRIUMF just added new tape drives, which should double their rate to about 50 MB/s.

Maarten

- 21/04 00:30 Between 06:00 and 08:00 GMT there was a 250 MB/s dip that is not understood, affecting all channels; the CASTOR team did not yet say if their logs pointed to any smoking gun. Around 10:15 GMT TRIUMF became unreachable via the OPN, and the same fate befell RAL around 14:43 GMT. Edoardo found one of the two Geneva-Amsterdam links down, affecting all the dedicated 1-Gbps channels to ASGC, RAL and TRIUMF. Remarkably the rate to ASGC was not affected, presumably due to an alternative fail-over route, RAL was only unreachable for some 3 hours, while TRIUMF is still unreachable as I write this. The majority of the sites are fairly stable around their target rates; a few sites did some maintenance or reconfiguration to improve running over the next days. *Maarten*

- 20/04 12:00 The IN2P3-CC rate dropped twice this morning:

- ◆ 4:00 GMT: due to a bad transfer, a mover hanged last night at 19:00 GMT and when the other pools on the machine ("ccxfer09") got full (at 4:00 GMT), no transfers did happen anymore on this machine because of dCache load balancing. The rate fell because all transfers went a single server ("ccxfer10"). When the stuck mover was killed transfers could resume on ccxfer09
- ◆ 6:00 GMT: The number of files received on the channel started to be very low, even 0 sometimes. The situation got better at 8:30 GMT without any intervention at IN2P3-CC

Lionel

- 20/04 Failures detected by Maarten at PIC seem to be "false" failures. We are able to reproduce the error, but can't see the reason. Anyway, files are being successfully transferred and migrated to tape through the srm. Someone (Maarten?) has upped the number of files in our channel to 60, so we are seeing relatively good performance numbers at PIC again. As a coarse approximation to our tape performance, we use <http://ganglia.pic.es/?r=day&c=PIC-GENERAL&h=stg005.pic.es> and check the dteam_migr_files graph. *Paco*
- 20/04 09:40 Based on GridView plots, rate to NDGF was 54MB/s averaged over 19/04 - with a target of 50MB/s - congratulations! Let's see how this keeps up over several days... *Jamie*
- 20/04 00:50 Bernd's T0 interference tests were successful according to the Lemon plots: the output rate was not affected by the input rate. As the tape migration was intensified, however, there was a noticeable fraction of gridftp timeouts bringing the SC4 rates down. The GridView plot shows two

declines and a valley that are not understood. The rate finally recovered after 13:00 GMT when the disk servers only served SC4 output data, still going to disk at the partner sites. Around 14:15 GMT 7 sites were switched to tape: BNL, FNAL, IN2P3, INFN, RAL, SARA, TRIUMF. NDGF have been writing to tape all the time since they started in SC4. FNAL met their target rate immediately, INFN and IN2P3 are very close. ASCC downgraded the kernel on their disk servers and have exceeded 50 MB/s which may be further improved by tuning various parameters. *Maarten*

- 19/04 12:00 Executive summary as reported to LCG Management Board of 18/04 and posted to service-challenge-tech:
 - ◆ We did not sustain a daily average of 1.6MB/s out of CERN nor indeed the full nominal rates to all Tier1s concurrently for the target period (April 3 - April 10).
 - ◆ We were nevertheless just under 80% of target in week 2.
 - ◆ Things clearly improved --- both since SC3 and during SC4:
 - ◇ Some sites meeting (includes exceeding) the targets!
 - ◇ Some sites within spitting distance optimisations? Bug-fixes?
 - See blog for CNAF castor2 issues for example
 - ◇ Some sites still with a way to go
 - ◇ Some sites (FNAL, RAL, others?) had on-going production activity in parallel with these transfers
 - ◆ The Operations of Service Challenges are still very heavy
 - ◇ Special thanks to Maarten Litmaath for working > double shifts
 - ◆ Need more rigour in announcing / handling problems, site reports, convergence with standard operations etc.
 - ◆ As a positive "postscript" to this exercise, the write streams that Bernd added last night did not perturb the data rate out of CERN (see Lemon monitoring) significantly - as had been feared could be the case.
 - ◆ There will be an update on SC4 at the next meeting with the LHCC referees (May 9th), for which I think we need to prepare:
 - ◇ A detailed post-mortem of the exercise (including the disk-tape transfers just starting now), summarising the key issues and lessons learnt;
 - ◇ A clear programme for bringing remaining sites up to (and beyond) their target nominal rates --- obviously whilst maintaining nominal+ rates at those sites that achieved this during the past couple of weeks.
 - ◆ Once again, I strongly encourage sites (which as always includes CERN) to include a summary of all key events and issues as input to the weekly operations meeting.
 - ◆ ***We really need to get things going in this area, so all sites, all VOs, as well as network operations are also systematically reported to these meetings.***
 - ◆ As Les stressed in his talk to HEPiX in Rome the week before last, SC4 production phase (1st June on) == the pilot WLCG service.
 - ◆ Thanks once again to everyone who continues to make this possible.
- 19/04 01:20 A bumpy day. Steady decline during the night, ending in a crevice between 04:00 and 05:00 GMT, not understood: no significant error rates reported by FT agents or CASTOR. The system recovered to 1.5 GB/s and then plunged into a crevasse between 08:00 and 09:00 GMT due to an Oracle deadlock plus LSF bug. The fall-out was cleaned up and the system soared back to 1.4 GB/s for two hours, after which the output rate seemed to be tumbling to 0.8 GB/s over the next few hours, according to GridView. From the Lemon plots, though, it was clear that the output rate actually never got below ~1.3 GB/s. The mismatch seems to have been caused by a problem with Tomcat on the Mon box: it ran out of memory, got restarted automatically, but apparently left the lcg-mon-gridftp clients on the CASTOR disk servers in a messy state, whereby somehow only about half of the gridftp records were making it to GridView. At 16:23 GMT I restarted all of them and the plot for that hour looks a lot better. CNAF is back at about 100 MB/s for now. Shortly after 16:00 GMT Bernd started a T0 interference test: while SC4 data is being sent out as usual, a steadily increasing input rate (simulating DAQ streams) is sent to the same disk pool, to see how the output rate gets affected.

Unfortunately, for most of the evening the BNL channel has been down, which already reduced the output rate by 250 MB/s. *Maarten*

- 18/04 (To CNAF): As far as I understand, the CASTOR2 LSF plugin and rmmaster version you are running is leaking jobs. We had the same problem at CERN during the SC3 re-run in January. The symptom is that LSF dispatches some fraction of the jobs without the message box 4, which is posted from the plugin. It is briefly described on slide 4 in the CASTOR2 training session 5: <http://indico.cern.ch/getFile.py/access?resId=33&materialId=0&confId=a058153>. Unfortunately it only describes a workaround for 'get' requests, while you are seeing the problem for 'put'. The procedure should be similar but you have to decide the filesystem yourself (for 'get' it is much simpler since the list of target filesystems is already posted in msg box 1). Alternatively you may just post an empty or non-existing filesystem, which will cause the job to fail. Simply killing the job with bkill would also work but result in an accumulation of rows with STATUS=6 in the SUBREQUESST table. The problem has been fixed (a workaround for a timewindow in LSF) in 2.0.3 so the best would be to upgrade. *Olof*
- 18/04 01:30 Easter Monday has been above 1.4 GB/s for most of the day, averaging about 1.5 GB/s, peaking at 1.6 GB/s right after the CNAF channel was switched on again. The problems of the day were with the CNAF channel, which was off except for a 6-hour spell, and with the BNL channel, which experienced many errors that were not very much reflected in the rate until 22:00 GMT. NDGF experimenting with dCache parameters. *Maarten*
- 17/04 02:30 Easter Sunday was the first full day averaging 1.6 GB/s! All channels were stable except for CNAF, whose LSF queue got full with stuck jobs a few times, requiring admin interventions. *Maarten*
- 16/04 03:50 A rocky night with BNL going up and down, and a peculiar dip between 02:00 and 04:00 GMT that could have been due to the large number of BNL errors during that period. On various other days there has been a dip between 01:00 and 03:00 GMT, so one is inclined to suspect interference by some backup or so. To be continued. At 12:00 GMT I scaled FNAL down from 100 to 70 files, which should still allow them to meet their nominal rate (and they do), while giving some room to BNL and other channels. The shared transatlantic network link apparently has independent light paths for BNL and FNAL, so the rate on one channel should not influence the other. At CERN, though, there is contention for the FTS and CASTOR, and as the FNAL channel uses 20 streams per file transfer while others use at most 10, a reduction there could make a lot of room for other channels. It turned out, however, that the problem at BNL was due to one or more stuck transfers, that were cleaned up in the morning, after which the rate quickly rose to just under 200 MB/s, and further increases in the number of files did not help. In the meantime the problem at INFN got fixed: their LSF queue was full of stuck jobs. I restarted their channel at 15 files and added 5 every hour, very visible in the GridView plot. At 55 it almost reached 200 MB/s. Now it is at 60 files and we can try even more, though it is disconcerting that we need so many files on a very fast and otherwise empty link between CERN and CNAF. The total rate was inching toward 1.4 GB/s, which suggested that we were again getting limited by the FTS DB, which has only been filling up after the reset a few days ago. A continuous cleanup procedure is the highest priority item for next week. But then suddenly after 22:00 GMT the rate jumped to almost 1.6 GB/s, for unknown reasons. There had not been any significant errors for hours on any of the channels, and there were no parameter changes that could explain the increase. It is suspicious that it happened right after midnight CEST. The rate crossed 1.6 GB/s the next hour. Meanwhile the BNL channel had started to develop a significant error rate that did not lower its throughput yet, but led us to lower the number of files from 100 to 80, after which the error rate dropped to insignificance and the throughput rose a bit further to 270 MB/s. *Maarten*
- 15/04 10:30 IN2P3-CC switched off 1 dCache server, that explains the drop from 7:00 GMT *Lionel*
- 15/04 02:15 Interesting times. Fairly stable running at 1.4 GB/s until 14:00 GMT, when an unknown

user or application started doing recursive listings of the dCache PNFS tree at FNAL, causing more and more srmCopy requests to hang and the rate to go to zero. A big service restart got FNAL back in business around 21:00 GMT. At BNL a lot of effort was spent to try and get the channel working again at a reasonable level, even though there still is the idea that there may have been something wrong with the network path at CERN yesterday. Also at BNL there was an abuse of the PNFS system, causing many timeouts, but the incidents are deemed unrelated. The INFN channel started failing 100% just before 17:00 GMT, every request hitting a CASTOR stager EBUSY. After lowering the number of files I switched the channel off a few hours later. NDGF have been running at a fairly stable 40 MB/s for two days in a row. *Maarten*

- 14/04 02:22 Fairly stable night around 1.6 GB/s, with a dip to 1.5 GB/s lasting for two hours, not understood. Then at 08:00 GMT something caused the rate to BNL to get almost halved and getting worse throughout the day. A lot of effort was spent at BNL to try and find anything amiss there, but then Dantong ran an iperf from CERN to BNL and found lots of packets getting dropped. The shared transatlantic link to Michigan works fine for FNAL, and the link from BNL to Michigan works as well, so the suspicion now is that something is wrong at CERN, where BNL and FNAL do not have exactly the same route, it seems. Meanwhile FNAL tested transfers using their SRMv2 endpoint: many/most of them actually succeeded, but the FTS considered all of them failed because of some problem in the srmCopy handshake, to be investigated by the developers. Around midnight CEST they switched back to SRMv1. At the same time we increased their number of files from 80 to 100 to allow for some of the performance losses to be recovered and to get us closer to 1.6 GB/s again. *Maarten*
- 13/04 09:40 Dip in transfer rates to IN2P3 around midnight (many SRM timeouts) understood - this happens every night because of the load on the head node due to the backup of the databases. We need to find another solution. We also plan to add some memory in it. The situation was back to normal shortly after then end of the backup. *Lionel*
- 13/04 00:15 Today we broke the 1.6 GB/s barrier! The preceding evening we had started a simple cleanup procedure on the FTS DB, instructing Oracle to delete the few million SC3 rerun records. In the morning we found it still busy constructing the roll-back log and we decided to abort it and go for a faster method instead, deleting the entries one by one with an immediate commit: this brought the FTS to a grinding halt and caused all channels to dry up around 13:00 GMT. We then decided to stop bothering and simply drop the tables completely. At 14:00 GMT all was operational again and we immediately reached the highest throughput so far, exceeding the nominal SC4 rate of 1.6 GB/s, which was sustained for 4 hours in a row, after which problems at BNL caused a dip lasting a few hours. In the meantime we found that one of the FTS agent nodes has become overloaded due to the very much improved FTS response time; we will move a few more channels to the other node. PIC is back at about 70 MB/s. ASGC still debugging their setup. RAL traffic competing with other users of their production endpoint, which may explain the lower transfer rates we see for them in SC4. CNAF still at 150 MB/s. FNAL nicely at 250 MB/s. FZK rose to almost 200 MB/s, but still not quite. IN2P3 at 200+ and SARA at 200 MB/s. DESY stable at 70, TRIUMF at 60. *Maarten*
- 12/04 01:20 Most of the day above 1.3 GB/s, reaching 1.5 for one hour. Various normally constant sites showed steps up and/or down that are not understood, though some appear correlated. We did the first (largest) part of the DB cleanup, to be finished on Wed. Certain FTS queries were taking e.g. one minute and a half instead of a few seconds, so we may very well see a dramatic difference in SC4 performance and stability afterwards. We tested the SRMv2 endpoint at FNAL in SRMv1 compatibility mode and managed to write some files and even read one back (not needed for now). *Maarten*
- 11/04 PIC will be in Scheduled Downtime from 10th till 12th April. This is because the yearly electrical maintenance works in our building are happening those days. *Gonzalo*

- 11/04 01:20 A rocky day. During the night the srmCopy transfers to FNAL recovered and were >~ 350 MB/s for 7 hours in a row with a peak of ~450 MB/s. That rate is twice the nominal rate for FNAL at this time, but may serve to debug srmCopy and to find possible rate limitations at the T0. At 08:00 GMT we switched the FNAL channel off to test a theory that other channels might then pick up part of the slack. The rate dropped 350 MB/s and then started to climb slowly with the GridKa channel that had just been recovered. Before we could conclude the test we were hit by a DNS FUBAR causing the alias for the CASTOR name service to disappear, whereby the SC4 rate abruptly dropped to zero... This took 1.5 hours to recover from. Meanwhile PIC had switched themselves off in preparation for a power cut that will keep them offline until Wed. morning. Between 22:00 and 23:00 GMT on Sunday evening the rate to IN2P3 mysteriously got halved and has not recovered; Lionel noticed a coincidence with a DB backup time; the network experts at both ends see nothing wrong with the link. GridKa ran at 200 MB/s for a few hours, then mysteriously dropped back to 150 around 20:00 GMT, then recovered 2 hours later. RAL did some network configuration changes that appear to have had a small positive effect, allowing them to stay at 150 MB/s more easily. It is not understood why we cannot get more than 100 MB/s into CNAF; James and Tiziana did some iperf tests that were faster by almost an order of magnitude. In the evening I changed the FNAL srmCopy parameters to try and get a smoother sequence of transfer requests. ASGC are stable now, but I did not yet manage to get them to exceed 40 MB/s. *Maarten*
- 10/04 00:30 And another one. The srmCopy channel agent turns out to have a file descriptor leak, so I installed a cron job to restart it every hour. On top of that the FNAL agent has got itself into a persistent bad state causing most transfers to fail with some vague error message, to be investigated by the developers. GridKa failing 100% since about 09:00 Sunday morning. ASGC bouncing up and down due to instabilities at their end that are not yet understood. During the night the total rate was around 1.4 GB/s with a one-hour peak of almost 1.5 GB/s, while since 13:00 GMT the rate is about 1.1 GB/s. BNL has slightly taken up the slack, while other channels are not profiting, which reinforces the idea that the FTS is the limiting factor, since it still has to process all the failed requests for FNAL. *Maarten*
- 09/04 02:30 Another rough day. Twice the rates dropped by a few hundred MB/s for a few hours. For each period Olof found one of the older disk servers having become a hot spot, queuing up a lot of requests, leaving fewer LSF slots for the other servers. There exists a negative feedback loop that should prevent overload on any disk server, but it must be tuned a bit more. Meanwhile we decided to take out the 8 old servers after migrating their files to the 43 new servers, which should be more than sufficient for SC4. After doubling the data set to 8k files I found that we had been using only 900 files instead of the 4k available, which explains Bernd's observations. Now we are using all 8k files, such that the collision rate should be down by an order of magnitude. We hit two FTS bugs affecting srmCopy transfers to FNAL: occasionally a failed request is not cleaned up, whereby it keeps occupying a slot, and occasionally the channel agent gets into a bad state causing all requests to fail immediately with a CGSI-gSOAP error. Two cron jobs have been installed to ensure the cleanup and to restart the agent as needed. BNL repeatedly expressed their worries about well-configured sites unjustifiably suffering from issues with other channels. BNL have an Atlas milestone to demonstrate 7 days of stable running at ≥ 200 MB/s, so they are not very happy with the recent T0 instabilities. ASGC are back online, but the individual file rates are much lower than before, which we can compensate to some extent by increasing the number of files and streams, but there must be something wrong on their channel. NDGF has been doing low-rate tests throughout the day. *Maarten*
- 08/04 02:12 It seems that something is limiting the total output rate at the source: whenever some site gets a big increase (typically BNL or FNAL), a few other sites suffer decreases without correlated changes in their error rates. Olof sees no CASTOR services maxing out somewhere (SRM, LSF, DB); Bernd notices the disk I/O rates are a lot lower than they should be, due to inadequate load-balancing: we must double the set of source files such that we are less susceptible to statistical fluctuations. Another possible bottleneck is the FTS, whose DB has been accumulating the history of all requests, which may slow down certain queries; to be investigated next week. GridKa investigating if they are

getting limited by Geant or their national network. PIC has improved a lot, running 70 MB/s with only a small error rate (but with a lot of concurrent files). DESY solid as a rock at 70 MB/s, TRIUMF at 50 MB/s. ASGC tried to improve some network parameters, but all transfers fail since that intervention. Meanwhile FNAL has set a new record: 350 MB/s for 3 hours in a row. They are running at a maximum of 160 concurrent transfers, though the actual numbers fluctuate a lot due to variable srmCopy request bunching that is not yet optimized. The total SC4 rate has been over 1.4 GB/s for 3 hours in a row plus another 2 hours. *Maarten*

- 07/04 11:30 Two sites (BNL, TRIUMF) now above nominal rates for past day and three more sites (RAL, SARA and IN2P3) within 15% of nominal. (TRIUMF's average daily rate so far is above nominal too!) *Jamie*
- 07/04 00:30 FNAL bumped to 80 concurrent srmCopy transfers leading to almost 200 MB/s. BNL bumped to 60 (3rd party) transfers to get rid of idle time, whereby their easily reached 250 MB/s may be sustained. GridKa bumped from 40 to 60 with zero effect, so we are getting limited by something we do not understand yet; the same seems to happen for RAL. SARA slowly oscillating between 100 and 200 MB/s for unknown reasons (modulo a temporary SRM downtime). INFN investigating configuration issues in their mixed CASTOR-1/CASTOR-2 setup. We upgraded the FTS to gLite 3.0 RC2 to enlarge the SRM get- and putDone timeouts from 40 to 180 seconds, and the GridFTP progress marker timeout from 180 seconds to infinity; both issues caused significant failure rates. For one hour we exceeded 1.3 GB/s. *Maarten*
- 06/04 13:15 BNL now running at over 200MB/s for the past 24 hour period. Congratulations! *Jamie*
- 06/04 09:00 TRIUMF exceeded their nominal data rate of 50MB/s yesterday, despite the comments below. Congratulations! *Jamie*
- 05/04 23:59 A rough day with problems that are not yet understood (see the [tech](#) list), but we also reached the highest rate ever (almost 1.3 GB/s) and we got FNAL running with srmcopy. Most sites are below their nominal rates, and at that they need too many concurrent transfers to achieve those rates, so we still have some debugging ahead of us. CASTOR has been giving us timeouts on SRM get requests and Olof had to clean up the request database. To be continued... *Maarten*
- 05/04 16:30 The Lemon monitoring plots show that almost exactly at noon the output of the SC4 WAN cluster dropped to zero. It looks like the problem was due to an error in the load generator, which might also explain the bumpy transfers BNL saw. *Maarten*
- 05/04 11:02 Maintenance on USLHCNET routers completed. (During the upgrade of the Chicago router, the traffic was rerouted through GEANT). *Dan*
- 05/04 11:06 Database upgrade completed by 10am.DLF database was recreated from scratch. Backup scripts activated. DB Compatibility moved to release 10.2.0.2, automatic startup/shutdown of the database tested. *Nilo*
- 05/04 10:50 DB upgrade is finished and CASTOR services have restarted. SC4 activity can resume. *Miguel*
- 05/04 09:32 SC4 CASTOR services stopped. *Miguel*
- 05/04 09:30 Stopped all channels to allow for upgrade of Oracle DB backend to more powerful node in CASTOR. *James*
- 04/04 IN2P3 meet their target nominal data rate for the past 24 hours (200MB/s). Congratulations! *Jamie*
- 01/04 Problems with LCGR Oracle cluster (RAC) - As for the details of what happened to the service, the problem was most likely caused by a bug in Oracle RAC software. As the result of this bug, the RAC distributed cache froze, which caused in turn the unavailability of the whole service (all 4 nodes of the cluster have been affected). The service was not available from Sat-01-Apr-06 06:55 to ~Sat-01-Apr-06 8:20, affecting all of LCG_FTS, LCG_LFC, LCG_GRIDVIEW and LCG_SHIVA. *Andrea & Luca*

March 2006

- 30/03 Castor@cern: a rough day yesterday... First, we discovered a misconfigured SRM node, sending 10% of the requests to the wrong stager... Then, we found that the new vixie-cron did not run the job that creates gridmap file (wrong mode bits!)... Once all that was found 'n fixed, Maarten reported transfers timing out (fixed this morning).
On the brighter side: we are moving more and more LHC experiment users to their Castor-2 instances. *Jan*
 - 23/03 Castor@cern: 50 disk servers froze at 17:30, and needed to be rebooted. No service until 20:30. Bad & Ugly... *Jan*
 - 23/03 ccsrm.in2p3.fr is back. One server still missing (LHCb T1-T1 transfers) *Lionel*
 - 22/03 Castor@cern: ~30 disk servers are being prepared. Should be ready tomorrow... *Jan*
 - 22/03 srm.grid.sara.nl will not be available tomorrow due to maintenance. This will last from 10:00-12:00 CET. *Ron*
 - 22/03 As previously announced, all services at CCIN2P3 are down for the day. dCache/SRM will be restarted tomorrow. *Lionel*
 - 22/03 Notes (reverse chronological order) of events - good, bad and ugly - during the SC4 preparation and throughput phase. *_Jamie_*
-

This topic: LCG > ServiceChallengeFourBlog

Topic revision: r87 - 2007-02-19 - AlbertoAimar



Copyright &© 2008-2020 by the contributing authors. All material on this collaboration platform is the property of the contributing authors.

Ideas, requests, problems regarding TWiki? Send feedback