

== ASGC T1 ==

Castor2

\* 9 disk servers have been brought online for a total of 15 disk servers and a total capacity of 218.92T.

LFC

\* LFC database table has been fixed and now upgraded to 1.6.3

Events:

Power cycle event

\* Date: April 23, 2007. 03:40-3:55 UTC

\* Reason:

\* Human error disconnected power to Grid services.

\* Severity:

{}{}

The following Grid services were restarted due to power cycle.

disk servers of castorsc:

1. wcp006

2. wcp007

3. lcg00007

4. lcg00008

old castor server:

1. castor

nfs servers:

1. exp s/w: nfs

2. pragma005: asgc users" home

grid services:

1. ui (lcg00122)

2. rb (lcg00124)

3. prod bdii (lcg00126)

} }

\* Solution:

\* Power was immediately recovered after incident and services tested.

\* sam js error

\* time stamp: 19-Apr-2007 08:14:18

\* err msg: Cannot plan: BrokerHelper: no compatible resources

\* system cured itself automatically, could arise from high load of giis srever. need evidence from ganglia.

\* fixed gmond info publish err, confirm also from central ganglia monitoring page.

\* unable to reproduce the err we have.

## CERN Tier-0 Site Report (23/4/07)

---

### Castor

---

- \* Two new database corruption appeared in the CMS Stager database this week.
- \* The CMS Stager database was migrated to the new system running Oracle RAC RDBMS 10.2.0.3 64bits on two HP Proliant computers and two IBM Nas Boxes. The migration started earlier than expected due to another corruption in the database affecting the table CASTORFILE. The corrupt block could be recovered and the migration completed successfully.
- \* The CMS DLF database was moved to a new server, running 64bit, single instance. It was upgraded to RDBMS 10.2.0.3
- \* The ITDC Stager database is totally overloaded with the tests run by the Castor team, to the point where it is not possible to so an ssh connection to the machine, and the Lemon agent is not able to extract any info from the machine. The machine needs to be replaced in order to run any Castor meaningful test.
- \* The Castor Atlas Stager database had to be restored after an accidental removal of all the objects in the castor\_stager database user. The recovery was performed to a point in time just before the "incident" took place.
- \* The new Gridview Publisher (fixing a memory leak) failed to restart cleanly on ~20 machines, and had to be restarted by hand. This was done on Tuesday.

### AFS

---

- \* Atlas nightly software builds hampered by performance problems on afs62 (transtec scratch server). Moving volumes off is slow due to high traffic. They will start builds on "q." volumes instead to see if the move to expensive hardware helps. Unloading of afs62 continues at the maximum possible pace as the machine needs the WD firmware update anyway. A discussion with Atlas about their requirements is scheduled for Friday, it is not unlikely that we may be forced to increase the number of servers for scratch space.

### Networking

---

- \* TS group will replace a display module on PDU EOD57 20/04 at 08:00. The OPN routers will be running on reduced power supplies for five minutes

## Cooling

---

- \* A mechanism to automatically shut down and power off batch nodes when they reach a predefined threshold is going to be implemented. It will help in case of Hvac failures. Initial threshold value will be 48deg C for each node

## Physics Database Services

---

- \* The April security patch for Oracle databases was released. Only one vulnerability might affect our databases. As the patching is not rolling we are checking if there is a workaround to remove the vulnerability.
- \* We have partitioned one very big LCG SAM table on LCG production RAC by setting most of the data as read-only. The aim of this exercise is to decrease the necessary backups. Next week some Gridview tables will be partitioned as well, in a similar way.
- \* New scripts were prepared and tested to automatically generate the procedures to perform the split and merge solution in case of one destination site is down.
- \* The second RAC database cluster at PIC for the ATLAS experiment is up and running. Streams throughput tests preparation ongoing. Triumf network frequently broke down during the last days, the causes are under investigation.

## FNAL Tier-1 Site Report (23/4/07)

---

Nothing much to report, quiet week.  
Our local bdii still doesn't work, we don't know why. We continue to use the one at CERN.

## Triumf Tier-1 Site Report (23/4/07)

---

WN disks filled by user job (30GB athena log file). Ban user, mail and clean up.  
Consider ulimit on file size?  
Declare maintenance for breaker replacement  
Power breaker tripped. It has happened before and is not overload - needs replacing.

- IN2P3-CC T1/French ROC Report:

---

\* Regional Top BDII has been put in production at IN2P3-CC. For now, it is only used by IN2P3-CC site but it will be proposed to all french sites as soon as it has been validated by IN2P3-CC site which is checking the load is ok.

\* Accounting normalization is a current concern of the french sites. There is not a common understanding of the information to provide to compute the normalized CPU time. In particular, with the new hardware (multi-cores, hyperthreading, ...), what is the pertinent information for accounting ? What is the official way to compute it ? May be this point should be addressed during a ROC managers meeting

Report for Tier1 GridKA (FZK):

---

[ author : Jos van Wezel]

During the weekend of 14/15 april some of the dcache pool went into an unresponsive state. Machines had to be restarted.

RB machine stability problems. RB no longer monitored until a replacement is installed which is targeted for 23/4.

PIC Tier-1 Report:

---

The only relevant incident was a network problem that affected PIC from 18-Apr at about 22:45 until 19-Apr at about 9:00. The catalan network provider informed us that the problem was in a faulty rectifier, which was substituted on 19-Apr in the morning.

lcg-CE, SRM-disk and SRM-tape: Ok. No major incidences.

top-BDII: We have installed a new top-bdii with 4 GB of memory. We were getting too much timeout errors from the CE replica manager tests. The alias bdii.pic.es now points to the new machine, so the tier2s which are pointing to our regional top-bdii won't have to reconfigure the WN's environment.

FTS: Still have problems with the fts server when trying to send transfer jobs to/from the dcache system. We have changed the hardware to a new machine just in case it was a problem with a high load in the machine. After the upgrade of the machine we noticed we were still getting the same errors. We are still investigating the problem.

BNL

---

Friday,

Panda shifter reported that some BNL jobs were put on hold -- reasons unknown.

Saturday:

The problem continued to develop. Production showed error message “Pool is disabled (errno 104).”. 1292 jobs at BNL were put on-hold. dCache system administrators found that one write pool was not restarted properly, and leftover Java pool processes were not cleaned up properly.

dcdoor02 went off-line around 7:00PM. Nagios sent out a warning notification. A dCache administrator restarted the server around 9:00AM Sunday morning. 14% connectivity was lost during downtime. The door logs did not provide enough information to understand the cause -- further investigate is needed.

Sunday:

Two GridFtp doors were restarted at 9:00AM and 7:00PM.

Monday:

Five doors and two servers attached to LHCOPN dropped a large number of received packets. Our network administrator investigated the problem and tried to make sure that there was no network problem.

Three write-pool servers were restarted early in the morning.

BNL’s Oracle 3D instance was setup in active log mode and RMAN (recover managers) configured per Eva’s request.

Tuesday:

GridFtp Door 01 and 02 were brought off-line to diagnose the ongoing network problem and packet loss.

From 13:10-16:30 CERN time, all data transfers of ATLAS Tier 0 exercise failed. Hiro investigated the problem. It turned out that door had the maximum allowed number of streams per file set to 10. The ATLAS Tier 0 exercise recently increased the number of streams per file to 40 based on Miguel’s recent investigation. Hiro did invention and increased the value of gsiftpMaxStreamsPerClient from 10 to 50.

Some users deleted files right away after they copied into dCache before the files were archived. The PNFS entries were removed, and the data was still there. By default, dCache does not delete precious files. We changed the configuration (removeUnexistingEntriesOnFlush=true) and restarted all write pools.

Wednesday:

Three dCache doors went offline at 3:00AM, 7:00AM and 8:30AM. The memory consumption pattern was like SAW tooth. The console showed: “Kernel Panic: not

synching, out of memory, no killable processes". These three doors were brought back on-line around 9:30AM.

Three write pools were restarted again. Our dCache administrators removed the home-grown write pool restart script, and let dCache itself clean up the stuck data transfer.

The maximum allow TCP stream per file on GridFtp doors were adjusted again to enhance the stability of dCache.

MyProxy was offline from 5:00PM to 9:30PM when we upgraded some Glite packages. The ATLAS DDM operator noticed it at 8:00PM and fixed it around 9:30PM. There was a four-hour downtime for USATLAS Production data transfer.

The network problems discovered on Tuesday turned out to be caused by a faulty network card.

Thursday:

GridFtp doors were stabilized after we decreased the number of streams per file allowed at GridFtp doors. There was a two-hour network outage while we replaced our faulty CISCO chassis. The morning network outage caused some left over Java processes in the SRM server. SRM was off-line for two hours.

#### NDGF-T1

---

Migration of data to alternative site scheduled to be completed tomorrow, some data will not be recoverable.

Data pools at one site inaccessible due to errors in underlying file system. Migration to other site underway, recovery of data in a damaged pool started.

#### SARA

---

gris on ce.gina.sara.nl is down resulting in a job list match failed.

#### NIKHEF

---

We installed a new RB+BDII machine and found that the networkserver of the RB frequently died on new connections.

On further inspection, it turns out that also two other RBs suffer from this problem since Tuesday, April 17.

The cause is currently under investigation.

We discovered that one of the enclosures (containing harddisks) of a DPM disk server was bad. The result was that many files, particularly those larger than ~40 MB, were corrupted due to random bit flips.

The cause is probably in the (cache) memory of the enclosure, which explains why there were no file system errors, no errors in the transfer logs or any other logs. This problem

was detected by Atlas production, comparing the md5sum of the files that were downloaded to the md5sums stored in the LFC.

This problem potentially affected over 170000 files (5.5 TB).

## RAL

---

Downtime announcement : The RAL-LCG2 Castor will be down for upgrade on the 30th April and the 1st May, this will affect the 6 SRMs ralsrm[a-f].rl.ac.uk.

## Availability reports

Date: 14/04 - 16/04

Reason: Castor Stager hung and two srm hosts locked up

Severity: Critical

Solution: Stager was restarted and the srm hosts were rebooted

Date: 17/04

Reason: Continuing instabilities with Castor

Severity: Critical

Solution: The default SE for ops was switched to an non-Castor SE

(dcache.gridpp.rl.ac.uk) while the Castor team attempted to fix the problem with the system up, this failed to resolve the problem and all Castor SRMs were taken out of service for an intervention on the Oracle Database.