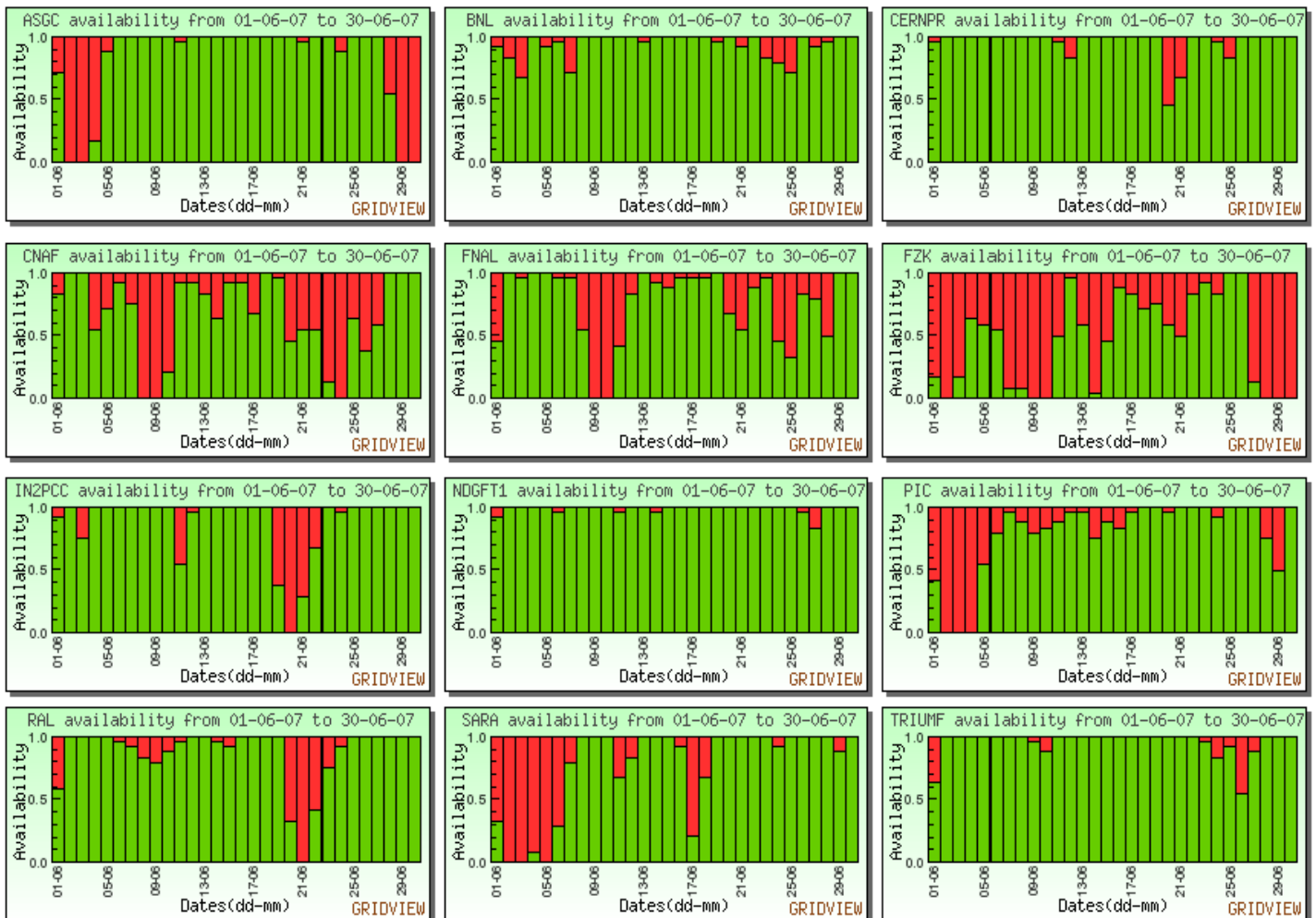


## Site Availability Reports June 2007

- Please review and complete the Site Reports below.  
Deadline Friday 6 July 2007.  
Edit your section and mail the document back to A.Aimar.
- There examples of good reports (BNL, CERN, RAL) and report completely missing at the OPS meeting (ASGC, CNAF, IN2P3) for the whole month.
- Please pass this information to your representative at the Operations meeting so that we do not need to complete them at the end of each month.

Sites availability, for all details see here:

[http://lcg.web.cern.ch/LCG/MB/availability/site\\_reliability.pdf](http://lcg.web.cern.ch/LCG/MB/availability/site_reliability.pdf)



	CERN-PROD	FZK-LCG2	IN2P3-CC	INFN-T1	RAL-LCG2	SARA-MATRIX	TRIUMF-LCG2	Taiwan-LCG2	USCMS-FNAL WC1	PIC	BNL-LCG2
1	96%	17%	92%	83%	58%	33%	63%	71%	50%	42%	92%
2	100%	0%	100%	100%	100%	0%	100%	0%	100%	0%	83%
3	100%	17%	75%	100%	100%	0%	100%	0%	96%	0%	67%
4	100%	63%	100%	54%	100%	8%	100%	17%	100%	0%	100%
5	100%	58%	100%	71%	100%	0%	100%	88%	100%	54%	92%
6	100%	54%	100%	92%	96%	29%	100%	100%	96%	79%	96%
7	100%	8%	100%	75%	92%	79%	100%	100%	96%	96%	71%
8	100%	8%	100%	0%	83%	100%	100%	100%	54%	88%	100%
9	100%	0%	100%	0%	79%	100%	96%	100%	0%	79%	100%
10	100%	0%	100%	21%	88%	100%	88%	100%	0%	83%	100%
11	96%	50%	54%	92%	96%	73%	100%	96%	42%	88%	100%
12	83%	96%	96%	92%	100%	83%	100%	100%	83%	96%	100%
13	100%	58%	100%	83%	100%	100%	100%	100%	100%	96%	96%
14	100%	4%	100%	63%	96%	100%	100%	100%	92%	75%	100%
15	100%	46%	100%	92%	92%	100%	100%	100%	88%	88%	100%
16	100%	88%	100%	92%	100%	92%	100%	100%	96%	83%	100%
17	100%	83%	100%	67%	100%	21%	100%	100%	96%	96%	100%
18	100%	71%	100%	100%	100%	67%	100%	100%	96%	100%	100%
19	100%	75%	38%	96%	100%	100%	100%	100%	100%	100%	96%
20	46%	58%	0%	46%	33%	100%	100%	100%	67%	96%	100%
21	67%	50%	29%	54%	0%	100%	100%	96%	54%	100%	92%
22	100%	83%	67%	54%	42%	100%	100%	100%	88%	100%	100%
23	100%	92%	100%	13%	75%	100%	96%	100%	96%	100%	83%
24	96%	83%	96%	0%	92%	92%	83%	88%	46%	92%	79%
25	83%	100%	100%	63%	100%	100%	92%	100%	33%	100%	71%
26	100%	100%	100%	38%	100%	100%	54%	100%	83%	100%	100%
27	100%	31%	100%	58%	100%	100%	88%	100%	79%	100%	92%
28	100%	0%	100%	100%	100%	100%	100%	54%		75%	96%
29	100%	0%	100%	100%	100%	88%	100%	0%	100%	50%	100%
30	100%	0%	100%	100%	100%	100%	100%	0%	100%	100%	100%
ility	96%	46%	88%	67%	87%	75%	95%	80%	77%	79%	94%

---

## ASGC

➤ *Please report on the site unavailability period(s), see page 1*

**Day:**

02-Jun-2007 03:06:48, 02-Jun-2007 07:23:17, 02-Jun-2007 11:08:33 and 02-Jun-2007 13:26:41

**Reason:**

lcg-cr fail with permission error, while manual data management validation works fine from client, first profile could only be done base on generic dteam vo membership, rather than ops monitoring VO that data registration request will map to opssgm. (error messages: lcg\_cr: Permission denied)

**Severity:** all SAM functional testing fail if data registered base on specific user dn that was reject due to the acl settings. Further analyze with ops sgm proxy, we're able to identify the problem, while dpm/dpns log messages wasn't able to provide detail log info that we can figure out the root cause except for practical validation with ops VO.

**Severity:**

All SAM functional testing fail if data management validation performs by wrong DN with respect to the 'generated' directory created by some other DN mapping before.

**Solution:**

applied the patch provided by deployment team (Maarten) sent to rollout list, and this help updating the ACL require for all 'generated' name space dedicated for generic SAM functional testing. The patch has been applied for all VO support on production DPM server.

**Date:** 03-Jun-2007 17:26:26

**Reason:**

SAM replica management testing fail with connection timeout problem, this arises from high CPU load on top level BDII server, that GFAL query timeout frequently. Manual testing fail with exactly the same error we observed in SAM.

**Severity:** intermittent SAM critical error referring to lcg RM, along with DPM permission error arise from ACL settings, site availability drop to zero at Jun 03.

**Solution:** the CPU load problem could only be resolved if bdii release upgrade to 3.8.8-1, and we found exactly the same problem at some other regional resource centers that upgrading the bdii release tag help resolving also the load errors.

**Date:** 28-Jun to 30-Jun

**Reason:** due to the migration of GOCDB, that we're facing the duplicate entry of DPM (dpm01.grid.sinica.edu.tw) registered in PPS as well as in production site db node list, say Taiwan-LCG2. However, this wasn't found in old GOCDB interface, and SAM/gstat now showing any particular service failure of functional testing error during the same time (the problem extend to Jul 6, until Judit inform us the drop of site availability due to the SRM service failure in GridView, while information system confirm the exist of DPM end point publish, which is inconsistent with the info carried out from GOCDB adopted by SAM).

**Severity:**

## LHC Computing Grid Project

Site availability drop to zero, and the problem could only be found if you check from grid view, rather than from SAM monitoring page, or gstat we used to for daily service monitoring.

Solution:

While all SAM functional testing passes at Taiwan-LCG2, including also the generic data management validation, but SRM fail to find available DPM node from node list def in GOCDB, that site availability drop to zero since GOCDB interface migrated, we're not able to act promptly since we're not aware before that it's not support conflicts FQDN of node defined in new GOCDB, and this shouldn't be regarded as site service failure as well. We're removing dpm01 from PPS and add it back to Taiwan-LCG2 and add new DPM node into PPS site db info to avoid the confusion of site availability.

---

## BNL

### > Friday: June 01/2007

We continue to observe sluggish Tier 0 data export from CERN to several Tier 1 sites. Our PNFS should high load generated by USATLAS production and Tier 0 data exports. We saw many time-out in data transfer.

### > Saturday, June/02/2007

Problem: SRM server in BNL was off-line between 4:00PM and 7:00PM.

Cause: The SRM server, GridFtp door server, and write pool server certificates expired around 4:00PM.

Severity: the data transfer completely stopped during 4:00PM and 7:00PM. Tier 0 data export to BNL, USATLAS Production and AOD/NTUPLE data replication were impacted.

Solution: We renewed the certificates around 6:00PM.

Problem: USATLAS production has problems to write data into BNL dCache system at 8:48PM. No data can be written into the subdirectories of the dCache root directory: /pnfs/usatlas.bnl.gov/AOD01, RDO01, etc.

Cause:

We changed the ownership of directories to the production account ???usatlas1???. But did not attach storage tag, which means that there were no write pool resources assigned to these sub-directories. The production could not create subdirectories and write files.

Severity: The USATLAS production was affected for two hours.

Solution: attach the storage tag into subdirectories, as we agreed on the morning meetings.

Problem: The problem happened on Wednesday continued. USATLAS production manager reported the performance of the data export from BNL to some Tier 2 sites was sluggish.

Cause: Production requests data files that are only in HPSS, but not in disk areas. It takes long time to stage-in files into disks and transfer them to the remote sites.

Severity: many Tier 2 sites are running out of data files and waiting for data transfer since Tuesday morning. The resource utilization for USATLAS decreased.

Solution:

We reallocated 10 HPSS drives to speed up the data stage-in. The USATLAS production manager gave us a list of files. Our data management expert created a script to stage-in the input files from HPSS to disks.

### > Sunday: June/03/2007.

Problem:

## LHC Computing Grid Project

The Panda monitor (USATLAS production monitoring server) crashed. The kernel crashed from a stack overflow (do\_IRQ: stack overflow: 488), probably because it was not able to keep up with the increasing memory pressure.

Cause: from the message that is still visible on the console and the system logs, it looks like that apache was using too much memory, triggering the kernel's out-of-memory killer several times in the hour prior to its crashing:

```
Jun 3 00:12:33 gridui02 kernel: Out of Memory: Killed process 27405 (httpd).
```

```
Jun 3 00:29:04 gridui02 kernel: Out of Memory: Killed process 28348
```

there could be a memory leak in the panda monitor or the version of apache it is using.

Severity:

USATLAS production dashboard is off air for twelve hours. Production runs blindfolded.

Solution: Reboot the server, and add a memory warning watermark in BNL Nagios monitoring page.

Problem: USATLAS production manager reported the performance of the data export from BNL to some Tier 2 sites was sluggish.

Cause: Production requests data files that are only in HPSS, but not in disk areas. It takes long time to stage-in files into disks and transfer them to the remote sites.

Severity: many Tier 2 sites are running out of data files and waiting for data transfer since Tuesday morning. The resource utilization for USATLAS decreased.

Solution: No Solution was given yet.

problem: data export from Tier 0 and USATLAS production data transfer suddenly failed around 1:00PM.

Cause:

The dCache map file is updated around 1:00PM. Both update scripts and SRM read/write from/to the same dCache grid user map file. SRM reads only a fraction of the file around 1:00PM, and two accounts are missing from the file: USATLAS1 and USATLAS4

Severity: Both USATLAS production and Tier 0 data export have been off-line for twenty minutes.

Solution:

Our dCache administrators redirected the output of the dCache grid map file updating script to a temporary file. After the update is finished and validated, then the script swaps the new grid map file with the existing dCache grid map file.

### > Thursday June/07/2007

Problem: SRM data transfer timed-out.

Cause: the dCache write pool caches are filled up because the majority of tape drives were allocated to file stage-in. New files have to wait for the available caching space.

Severity: dCache write pools performance is slow, and large number of SRM requests are queued.

Solution: after the input files are completely staged-in from tapes, we will return the tape drives allocated for write pools to speed up the data migration from the write pools to tapes.

Maintenance:

We reallocated 10 HPSS drives to speed up the data stage-in to fix the sluggish production problem reported on Wednesday. The USATLAS production manager gave us a list of files. Our data management administrator created a script to stage-in the input files from HPSS to disks.

### > Monday: June/11/2007

Problem:

## LHC Computing Grid Project

One of GridFtp servers node (dcdoor02) crashed on 6:00AM Monday morning.

Cause: an IRQ kernel error.

Severity: the GridFtp server was off-line for four hours and 14% connectivity was lost due to this problem.

Solution:

Our system administrator updated the server since it hadn't been updated in many months. Then, after the update, we rebooted it a second time so it would be running the updated kernel.

Problem:

Two OSG gatekeepers were reported critical at 04:54:56 EDT 2007

Cause:

GUMS server went off-line (but log output had stopped. we restarted Tomcat and it appears to be functioning).

The new host certificate generated for the GUMS server was missing an attribute needed for a server. The old one had

X509v3 extensions:

Netscape Cert Type:

SSL Client, SSL Server

while the new one had:

X509v3 extensions:

Netscape Cert Type:

SSL Client, S/MIME

Severity:

Two OSG gatekeepers were impacted for three hours before our administrator intervention.

Solution:

We disabled the GUMS on two OSG gatekeepers and used the static Grid map files to allow BNL CE to be accessible right away.

In the mean time, we obtained a new host certificate with a proper server attribute for our GUMS server.

### > Tuesday: June/12/2007

Maintenance:

Kernel update was performed on all seven GridFtp server nodes one by one.

Each GridFtp server has less than an hour downtime. The update is transparent to users. No need to make announcement.

### > Wednesday: June/13/2007

Problem:

A fraction of data transfers from BNL to other ATLAS Tier 1 sites failed with certificate mismatch errors.

Cause:

A fraction of our dCache read pool nodes have bad certificates that their DNS do not match with their hostnames.

The error message is shown as follows:

-----  
06/07 12:01:04 Cell(SRM-dcsrcm@srm-dcsrcmDomain) : Authentication failed.

Caused by GSSEException: Operation unauthorized (Mechanism level:

[JGLOBUS-56] Authorization failed. Expected

"/CN=host/acas0203.usatlas.bnl.gov" target but received

"/DC=org/DC=doegrids/OU=Services/CN=acas0399.usatlas.bnl.gov")

Severity: This problem affects the data transfer directly between the remote party and these affected dCache read pool nodes

while the data transfer via (GridFtp servers) was not affected. BNL to other Tier 1 data transfer does not use GridFtp

server nodes; all data transfer from BNL affected nodes to Tier 1 sites experienced transfer failures.

USATLAS production is NOT affected by this problem.

Solution:

## LHC Computing Grid Project

We replaced these bad certificates on Wednesday. We notified the ATLAS data operation team to confirm whether the lower performance problem with the data transfers from other Tier 1 sites to BNL. We will add scripts to validate the host certificates.

### > Friday: (June/22/2007).

Problem: a user repetitively copies a file from dCache into the local area on worker nodes. The dCache name space server (PNFS) load is about 5.

Cause: One user tries to run the reconstruction software over some RDO files repetitively (hundreds of times).

Severity: SAM tests experienced large number of time-out on Thursday (June/21/2007). BNL DQ2 0.3 site service validation failed in the middle because of the PNFS high load.

Solution: Our site administrators killed all local batch jobs generated by this user. The dCache file copy activity was stopped, and the load on PNFS was back to normal.

### > Saturday ~ Monday: June/23/2007~June/25/2007.

Problems: BNL write servers show large number of time-out errors since June/13/2007.

Cause: A user uses SRMCP command to copy data files from the remote ATLAS Tier 1 sites directly into our internal write pool servers which do not have public DNS entries and necessary firewall conduits at BNL perimeter firewall. All of data transfer stuck, and eventually failed.

Severity: these data transfer are hanging for one hour before they timed out. It hold the limited write pool resources and affected data transfer requests into these write pools during the weekends and the following Monday.

Solution: No solution was provided until Tuesday.

### > Tuesday: June/26/2007

Problems: large number of stuck SRM data transfers were found at our SRM log file since June/22/2007.

Cause: A user uses SRMCP command to copy data files from the remote ATLAS Tier 1 sites directly into our internal write pool servers which do not have public DNS entries and necessary firewall conduits at BNL perimeter firewall. All of data transfer stuck, and eventually failed.

Severity:

It caused some time-out errors to SAM tests, and increased the SRM database load.

Solution:

We restrict such direct pool to pool data transfers to only use the dCache write servers with firewall conduits, and reserve the internal write pools for data transfer via GridFtp servers.

---

## CERN-PROD

### > Remark(s) on 2007-06-01

01/06: OK, only 1 h unavailability

### > - 12/06/2007, 4 h (7-10 am):

problem: replica management (CE-sft-lcg-rm ) test failure  
cause: CASTORPublic Unavailable

## LHC Computing Grid Project

broadcast sent in advance: The public CASTOR2 service at CERN will be unavailable from 07.00 to 16.00 UTC (09.00 to 18.00 CET) on Tuesday 12 June TODAY while a major upgrade is performed.

As it is seen from the SAM monitoring, the intervention only caused a 4 h interruption.

There was an scheduled downtime on the SE for this interruption:

Upgrade of the public CASTOR2 stager at CERN 12nd June 2007 - 09:00 12nd June 2007 - 14:30

But as the failure was on the CE (replica management test) it was not taken into account by SAM.

### > Remark(s) on 2007-06-20

Problem: an attempt to fix the configuration for srm-durable-<VO>.cern.ch reported by LHCB was attempted at 11am. Unfortunately the fix introduced another misconfiguration for the non-durable SE srm.cern.ch.

Solution: the configuration for srm.cern.ch was fixed in the evening.

Problem: SAM tests not running since 18:00 and site availability continued to be flagged red for the whole night even if the SE problems were fixed in the evening

Solution: SAM team contacted in the morning and they admitted the problem and restarted the service

### > Remark(s) on 2007-06-21

Problem: SAM tests not running since 18:00 the day before and site availability continued to be flagged red for the whole night even if the SE problems were fixed in the evening

Solution: SAM team contacted in the morning and they admitted the problem and restarted the service

### > Remark[s] on 2007-06-26

Problem: wrong SE records published, breaking lcg-utils, during 3 hours in the afternoon.

Solution: roll back to previous version of Castor SRM information provider

---

## CNAF

➤ *Please report on the site unavailability period(s), see page 1*

Day: Jun 4,5,13,14,17,20,21,22,23,24,25

\* Reason: CASTOR unavailability (several meltdown of several CASTOR servers, e.g. lsf, rmmaster) The meltdown of these services was typically due to a excessive load on the stager (e.g. too many lsf processes due to the Phedex probe, continuous write access to CASTOR D1T0 storage class with no available disk space)

\* Severity: CASTOR storage system not available

\* Solution: restart of CASTOR services Day: Jun 6-11

Day: Jun 6-11

\* Reason: power cut

\* Severity: all services affected (including core switches)

\* Solution: farm services were restored on June 7 evening. Castor services restored on June 11.

Day: Jun 26-27



## LHC Computing Grid Project

- \* Reason: power cut on a rack due to a shortcut on a server
- \* Severity: mainly all tape servers failed due to the power cut, affecting CASTOR services.
- \* Solution: the faulty server was isolated and the service restored

---

### FNAL

➤ *Please report on the site unavailability period(s), see page 1*

#### > Remark(s) on 2007-06-01

unscheduled cooling outage

#### > Remark(s) on 2007-06-03

fully operational, test defect

#### > Remark(s) on 2007-06-06 to 07

fully operational, test defect

#### > Remark(s) on 2007-06-08 to 11

Failure of DNS at CERN

#### > Remark(s) on 2007-06-12 to 24

USCMS was fully operational, test defect

#### > Remark[s] on 2007-06-25

USCMS was operational defective tests for single bars

Disk failure on CRL squid cache led to authentication problems.  
Squid cache now a critical service so we will be alerted of problem.

#### > Remark[s] on 2007-06-26

Disk failure on CRL squid cache led to authentication problems.  
Squid cache now a critical service so we will be alerted of problem.

#### > Remark[s] on 2007-06-27

SRM database failure, dropped all SRM tables to recover

#### > Remark[s] on 2007-06-28

USCMS was operational, defective tests

#### > Remark[s] on 2007-06-29

Scheduled Downtime, should not be red.

---

### FZK-LCG2

➤ *Please report on the site unavailability period(s), see page 1*

## LHC Computing Grid Project

### > Remark(s) on 2007-06-21

Almost all difficulties this (and last week) stem from stability problems on the CE\''s. More specific, the info provider system (gris) sometimes returns erroneous data (i.e. no data). Consequently the job requests fail. We are investigating and have setup more extensive monitoring of all relevant activity on the CE. However for the time being the situation remains unsatisfiable.

### > Remark[s] on 2007-06-23

instabilities of CE. (infosystem) Under investigation

### > Remark[s] on 2007-06-24

instabilities of CE. (infosystem) Under investigation

### > Remark[s] on 2007-06-25

instabilities of CE. (infosystem) Under investigation

### > Remark[s] on 2007-06-28

Scheduled downtime

### > Remark[s] on 2007-06-29

Scheduled downtime

---

## IN2PCC

### ➤ *Please report on the site unavailability period(s), see page 1*

Day: 2005-06-01

Reason:

- 1) Classical unexplained CE-sft-lcg-rm failures. Only one CE was impacted because the 'ops' LFC catalog (operated by CERN) was not responding.
- 2) SRM SEs tests failures. That is certainly the reason of unavailability, but I don't know why those failures.

Severity: None

Solution: Ask Lionel Schwarz

Day: 2005-06-03

Reason: Classical unexplained CE-sft-lcg-rm failures

At the same time, our 3 CEs had different failures: SE not found in the IS, no file registration possible into the Central LFC, and lcg-cp timeout.

Severity: Unknown and temporary problem

Solution: Nothing was done

Day: 2005-06-11

Reason: Scheduled downtime was set

Severity: None

Solution: We didn't mind that SAM was going on the monitoring in spite of the SD

Day: 2005-06-19/20

Reason: Scheduled downtime was set

Severity: None

Solution: SAM team was informed of that anomaly (see SAM section)

Day: 2005-06-21/22

Reason: CERN CA CRL was out of date

Severity: Failure for jobs submitted with CERN proxy (in particular, SAM tests)

Solution: Manual update of CRLs was done and cron service was restarted

## LHC Computing Grid Project

Day: 2005-06-24

Reason: "CGSI-gSOAP: Error reading token data: Connection reset by peer"

Severity: Unknown and temporary problem

Solution: Nothing was done

---

## NGDF

Availability report : NDGF-T1

### > Remark(s) on 2007-06-01

Problem: Sam test issue, most likely

Solution: ignore it and it fixed itself

### > Remark(s) on 2007-06-06

Problem: SRM door on srm.ndgf.org had hung.

Solution: Restarted shortly afterwards and worked since.

---

## PIC

➤ *Please report on the site unavailability period(s), see page 1*

Date: From 01/06/2007 at 10:00 until 05/06/2007 at 8:00

Problem: Some of the CE SAM tests were launched with the certificate of Judit Novak, with VOMS role SGM. We had a misconfiguration in the YAIM users and groups files for the OPS VO and SGM role, which only appeared when OPS jobs with this VOMS role were submitted.

Severity: Low. This only showed up with one specific user of the OPS vo. The CE service for the LHC experiments was up and running during that period.

Solution: The misconfiguration for the OPS VO with SGM role was corrected.

--

Day: From 5 until 11 Jun 2007

Reason: Intermittent failures in the SRM-disk (dcache) service. Most of them timeouts. This might be correlated with the fact that the SRM-disk service was quite highly loaded those days. We are about to upgrade to dcache-1.7 in few days, and expect this will help to solve these instability sporadic problems.

Severity: Medium. The SRM-disk service is critical, but the failures we see are timeouts and affect typically to a small fraction of transfers. We believe the retries work at the end.

Solution: For the moment, no definitive solution has been found. Waiting to upgrade the whole SRM-disk service to dcache-1.7

--

Day: 14 Jun 2007, from 15:00 till 21:00

Reason: site-bdii service failures detected by SAM and GSTAT with error "GIIS query failed". We did not see the alarm in our local monitoring system. On top of that, the SAM tests like the SRM ones, that make use of the information published by PIC from the site-bdii, continued working OK during the period in which the site-bdii was detected failing. This could have been a false positive failure of the site-bdii test.

Severity: Medium. The service is a critical one, but the other tests that depend on the info published by site-bdii kept working. Seems that the CE or SE services were not affected.

Solution: No action was taken. After 5h of failures, the SAM tests for site-bdii started to be OK.

--

Day: 15 and 16 June 2007

## LHC Computing Grid Project

Reason: Same problems as those reported for 8-11 June. Intermittent failures in the SRM-disk (dcache) service. Most of them timeouts. This might be correlated with the fact that the SRM-disk service was quite highly loaded those days. We are about to upgrade to dcache-1.7 in few days, and expect this will help to solve these unstability sporadic problems.

Severity: Medium. The SRM-disk service is critical, but the failures we see are timeouts and affect typically to a small fraction of transfers. We believe the retries work at the end.

Solution: For the moment, no definitive solution has been found. Waiting to upgrade the whole SRM-disk service to dcache-1.7

--

Date: From 22/06/2007 evening until 27/06/2007 evening.

Problem: The FTS agents failed on 22nd June, for unknown reasons. Trying to solve them, an update of the system was issued, which ended up in installation of the FTS-2.0. We understand this was a mistake of the FTS deployment, since the FTS-2.0 rpms should not have been put in the production repository.

Severity: High. The PIC FTS service was down during all the period. Luckily, the experiments were not transferring files between T1 and T2s during those days. The only activity was CERN-PIC (from CMS, and not very intense) which is driven by CERN FTS.

Solution: The downgrade was extremely painful due to the way the meta-rpms are built. With the help of fts-support@cern, the downgrade could be finally done manually and the service restored with the original version.

--

Date: 24/06/2007

Problem: Two CE-sft-lcg-rm spurious errors at 00h and 08h due to one misconfigured WN pointing to a default\_SE recently migrated, with the Information Provider still to be configured.

Severity: Low. The problem appeared in only one of the WNs of the farm.

Solution: The bad WN was detected and correctly configured on Monday 25 June at 11h.

--

Date: From 28/6/2007 at 19:00 until 29/06/2007 at 12:00

Problem: The CE service fails Job Submission due to a human error when changing the batch scheduler configuration for testing one WN with SL4. The exact problem was with a wrong maui reservation that was aimed to assign a test WN to a test queue. Adding a GROUPLIST tag to this maui reservation, resulted in a wrong behaviour.

Severity: Low. This configuration problem was only affecting the OPS and DTEAM vos, since it was a test. Job submission for ATLAS, CMS and LHCb continued with no problems.

Solution: The problem was detected and solved by deleting the reservation, and substituting it by associating nodes to queues using a Torque directive "resources\_default.neednodes".

--

Date: From 28/6/2007 at 19:00 until 29/06/2007 at 12:00

Problem: SRM-disk service fails SAM tests because the reconfig of the info provider publishes the hostname (dcstrm02) by default, and not the alias of the service.

---

Severity: Medium. We see the FTS transfers CMS and ATLAS continued, since these transfers seem not to depend on the info published.

Solution: Solved by hardcoding the alias in the static info published by the dCache Information Provider.

## RAL-LCG2

Availability report : RAL-LCG2

## LHC Computing Grid Project

### > Remark(s) on 2007-06-01

Problem: high load on CE

Solution: problem eventually resolved itself after rebooting

### > Remark(s) on 2007-06-06

Problem: top-level BDII timed out

Solution: none (load on BDII not generally a problem)

### > Remark(s) on 2007-06-07

Problem: Scheduled maintenance + during reconfiguration a change was introduced that meant the OPS tests were mapped to accounts which could not run jobs successfully.

Solution: The reconfiguration was backed out, only OPS jobs would have been affected

### > Remark(s) on 2007-06-08

Time: 1300 GMT

Problem: Network Problem caused by switches going into odd state

Solution: Switches were reset.

Time: 2000 GMT until 0000 GMT

Problem: After reconfiguration on previous day, a local change to the job manager was overwritten causing the CE to remove jobs it found in a Waiting State, many jobs at RAL-LCG2 enter this state but then go on to run successfully, so the local change to the job manager to ignore waiting jobs was reintroduced on Monday 11th.

### > Remark(s) on 2007-06-09

Time: 0100 GMT until 0900 GMT

Problem: After reconfiguration on previous day, a local change to the job manager was overwritten causing the CE to remove jobs it found in a Waiting State, many jobs at RAL-LCG2 enter this state but then go on to run successfully, so the local change to the job manager to ignore waiting jobs was reintroduced on Monday 11th.

### > Remark(s) on 2007-06-10

Time: 0100 GMT until 2100 GMT

Problem: After reconfiguration on previous day, a local change to the job manager was overwritten causing the CE to remove jobs it found in a Waiting State, many jobs at RAL-LCG2 enter this state but then go on to run successfully, so the local change to the job manager to ignore waiting jobs was reintroduced on Monday 11th.

### > Remark(s) on 2007-06-11

Time: 0300 GMT

Problem: After reconfiguration on previous day, a local change to the job manager was overwritten causing the CE to remove jobs it found in a Waiting State, many jobs at RAL-LCG2 enter this state but then go on to run successfully, so the local change to the job manager to ignore waiting jobs was reintroduced on Monday 11th.

## LHC Computing Grid Project

### > Remark(s) on 2007-06-12

n/a

### > Remark(s) on 2007-06-13

n/a

### > Remark(s) on 2007-06-14

Problem: OPS tests were not being scheduled with sufficient priority to run soon after submission

Solution: Ops tests were previously running using dteam pool accounts and queues but have now been switched to a separate set of pool accounts and queue and the priority for this group has been increased.

### > Remark(s) on 2007-06-20

Problem: OPN link to CERN went down, causing CE replication and SE and SRM tests to fail

Solution: Link was repaired 22nd June approximately 14:00 (BST)

### > Remark(s) on 2007-06-21

Problem: OPN link to CERN went down, causing CE replication and SE and SRM tests to fail.

Solution: Link was repaired 22nd June approximately 14:00 (BST)

### > Remark[s] on 2007-06-23

Time : Midnight-13:00

Problem : OPN connection to CERN was down

Solution : OPN connectivity was restored at 13:00

### > Remark[s] on 2007-06-24

Time : 09:00 - 13:00 & 15:00 - 17:00

Problem : Timeouts in contacting the top-level BDII

Solution : BDIIs will be updated with new release to reduce load

---

## SARA - MATRIX

### > Remark(s) on 2007-06-02 to 07

Problem: problems with sgm mappings and change in SAM test to run as ops sgm

Solution: fixed now.

### > Remark(s) on 2007-06-11

Problem: Problems with site bdii due to a misconfiguration.

Solution: misconfiguration has been corrected

### > Remark(s) on 2007-06-12

Problem: Srm problems.

Solution: Service which was very slow has been restarted.

Problem: Problems with the oracle LFC which was very slow.

## LHC Computing Grid Project

Solution: Restarted the service with an increased number of threads which fixed the problem.

### > Remark(s) on 2007-06-17

Problem: We have had problems with dcache pools running out of disk space on their root file systems. This problem was caused by idle gridftp doors generating lots messages stating that it has nothing to do. This generated huge log files.

Solution: Removed the gridftp logs and restarted the gridftp door and pools. In addition we have tightened the logrotate rules so that the dcache log files are not only rotated and compressed each day but also when they exceed a 2 GB limit.

### > Remark(s) on 2007-06-18

See above.

Availability report : SARA-LISA

### > Remark(s) on 2007-06-20

problem: fileserver/storage crash

solution: reboot of cluster

### > Remark(s) on 2007-06-21

problem: nat to internet from WNs broke

solution: network department fixed the issue

Availability report : SARA-MATRIX

### > Remark[s] on 2007-06-25

Problem: Could not replicate file to CERN

Solution: Problem dissappeared by itself.

---

## TRIUMF-LCG2

### ➤ *Please report on the site unavailability period(s), see page 1*

Day: Network unstable during July 25 - July 27

Reason: various network interruptions and combinations of effects. It all started by a fiber cut on the Telus lines along with scheduled maintenances on our 5 GigE and 1 GigE lightpaths. These links act as failover to each other.

Severity: Network unstable. TRIUMF was still accessible through the LHC OPN for most of the time but sites on the normal research network could not see us.

Solution: network topology is being revisited for fail overs to the routed research network so traffic from LHC OPN nodes is properly routed onto the research network.

### 1. Remark(s) on 2007-06-01

Problem: GC CA crl was not updated due the CA operator error.

Solution: CA authority updated the CRL on web site

## LHC Computing Grid Project

### 2. Remark(s) on 2007-06-09

Problem: SRM directory creation failed. Unknown cause but likely due to a permission problem.

Solution: this is a known dCache issue. It happens randomly. We have a cron job that reset the proper directory permission.

### 3. Remark(s) on 2007-06-10

CE, SE, BDII  
delete failed

2000 jobs in queue due to ATLAS job priority problem - possible load issue.

---

## SAM unavailability

- 11 June, 16:00-17:00 Piotr testing firewall script
- 20 June 20:00 - 21 June 10:00 tomcat down
  - reason: memory leak in the software
  - workaround: cron job
  - resolved: when Piotr is back
- 20 June: rb118 WMS down during night (rb108 was working)
- 15 June - 20 June: GOCDB synchronization was stopped (mistake) => Downtimes problem
- 25 June: ~14:00-15:30 due to a failure on the SAM SE the tests were using srm.cern.ch, which also showed failures. This way some sites got fake Replica Management test alarms.
- 26 June: ~8:00-~10:00 tomcat down, no test results could be published
- 28 June: ~16:00-~18:00 GOCDB3 migration => bug in the synchronization script => test submission disabled for the time of debugging