

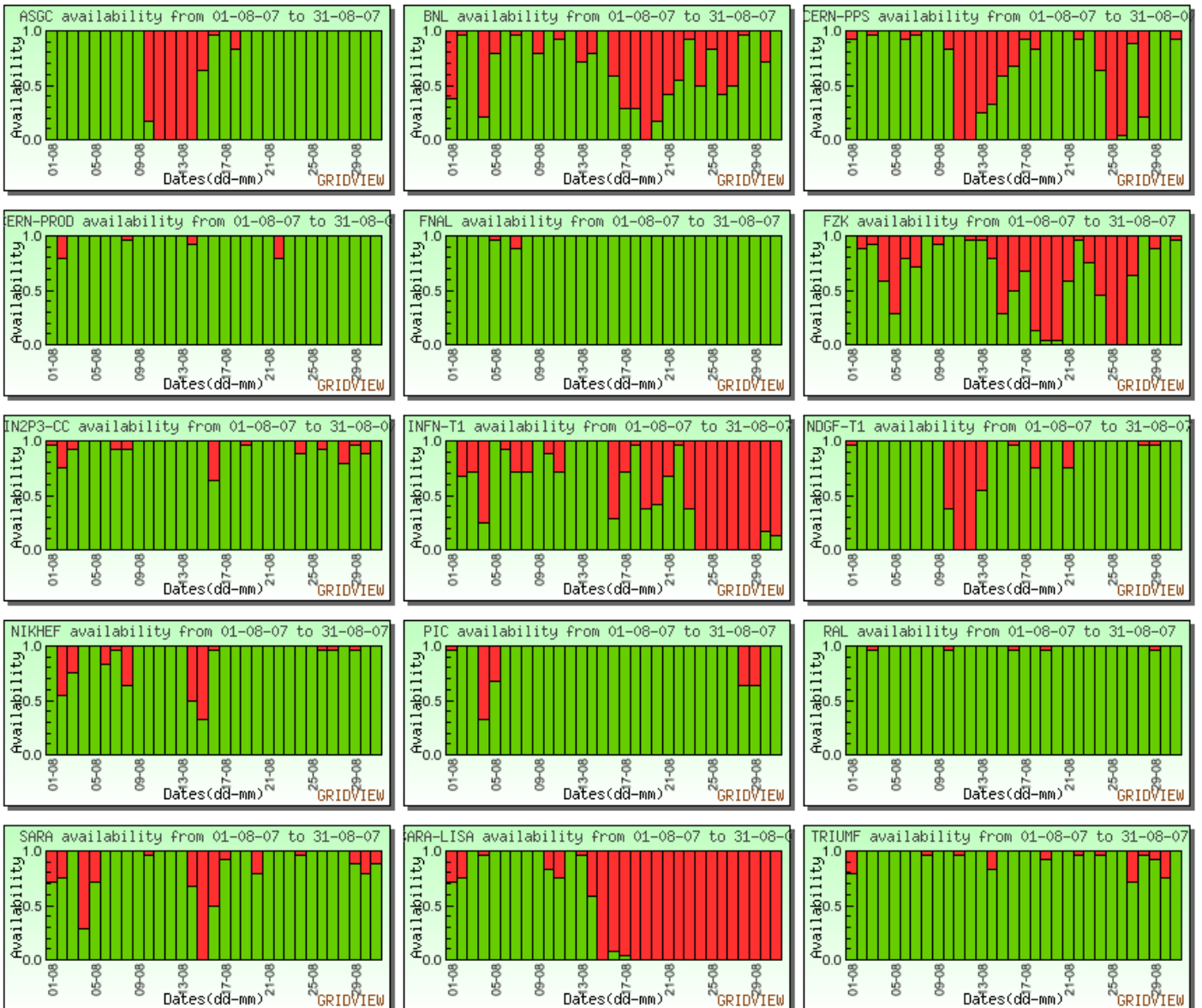
# WLCG Site Reliability Reports August 2007

- Please review and complete the Site Reports below.  
Edit your section and mail the document back to A.Aimar.
- **Deadline Friday 21 September 2007.**
- Some reports need to be completed, no reports from ASGC, CNAF, IN2P3, SARA

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

## Tier-1/0 Site Availability VO:OPS (Daily Report)

(Click on the Graph below to see Availability of Individual Services at the Site)



	CERN-PROD	FZK-LCG2	IN2P3-CC	INFN-T1	RAL-LCG2	SARA-MATRIX	TRIUMF-LCG2	Taiwan-LCG2	USCMS-FNAL-WC1	PIC	BNL-LCG2
1	100%	100%	96%	100%	100%	71%	79%	100%	100%	96%	38%
2	79%	88%	75%	67%	100%	75%	100%	100%	100%	100%	96%
3	100%	92%	92%	71%	96%	100%	100%	100%	100%	100%	100%
4	100%	58%	100%	25%	100%	29%	100%	100%	100%	33%	21%
5	100%	29%	100%	100%	100%	71%	100%	100%	96%	67%	79%
6	100%	79%	100%	92%	100%	100%	100%	100%	100%	100%	100%
7	100%	71%	92%	71%	100%	100%	100%	100%	88%	100%	96%
8	96%	100%	92%	71%	100%	100%	96%	100%	100%	100%	100%
9	100%	92%	100%	100%	100%	100%	100%	100%	100%	100%	79%
10	100%	100%	100%	88%	96%	96%	100%	17%	100%	100%	100%
11	100%	100%	100%	71%	100%	100%	96%	0%	100%	100%	92%
12	100%	96%	100%	100%	100%	100%	100%	0%	100%	100%	100%
13	100%	96%	100%	100%	100%	100%	100%	0%	100%	100%	82%
14	92%	79%	100%	100%	100%	67%	83%	0%	100%	100%	91%
15	100%	29%	100%	100%	100%	0%	100%	63%	100%	100%	100%
16	100%	50%	63%	29%	96%	50%	100%	96%	100%	100%	67%
17	100%	67%	100%	71%	100%	92%	100%	100%	100%	100%	29%
18	100%	13%	100%	96%	100%	100%	100%	83%	100%	100%	29%
19	100%	4%	96%	38%	96%	100%	92%	100%	100%	100%	0%
20	100%	4%	100%	42%	100%	79%	100%	100%	100%	100%	17%
21	100%	58%	100%	67%	100%	100%	100%	100%	100%	100%	42%
22	100%	96%	100%	96%	100%	100%	96%	100%	100%	100%	54%
23	100%	75%	100%	100%	100%	100%	100%	100%	100%	100%	92%
24	100%	46%	88%		100%	96%	96%	100%	100%	100%	50%
25	100%	0%	100%		100%	100%	100%	100%	100%	100%	83%
26	100%	0%	92%		100%	100%	100%	100%	100%	100%	42%
27	100%	63%	100%		100%	100%	100%	100%	100%	100%	50%
28	100%	100%	79%		100%	100%	100%	100%	100%	63%	96%
29	100%	88%	96%	0%	96%	88%	92%	100%	100%	63%	100%
30	100%	100%	88%	17%	100%	79%	75%	100%	100%	100%	71%
31	100%	96%	100%	13%	100%	88%	100%	100%	100%	100%	100%
lity	99%	67%	95%	70%	99%	86%	97%	83%	99%	94%	71%

---

## ASGC

### > 2007-08-10

Problem:

SAM Job submission testing failure start from 09-Aug-2007 09:23:11, and end at 09-Aug-2007 16:21:22, at least 8 events have been detected.

Cause:

The root cause arises from reboot of central NFS server, and all clients couldn't cure the connectivity with NFS server and recovery the mount points automatically. Remounting NFS file system manually will fail with resources or dev busy, and accessing NFS scratch fail also with 'Stale NFS file handle', error msg.

Severity:

Due to the inaccessible of network file system, not only the running jobs already submit to backend computing nodes is affects, but also the new arriving jobs couldn't read experiment software from NFS scratch that all SAM job submission testing fail at site CE (at least 11 jobs have been cancel from batch server directly) with backend WN sharing the same mount points from NFS server.

Solution:

To avoid jobs being stuck forever, we have force removing them from queue, and force remounting the file system from NFS, mount option have been tuned further to avoid the same problem in the future.

### > 2007-08-15

Problem:

Replica management testing error detected by SAM at 14-Aug-2007 16:20:37, with error message 'Timeout when executing test CE-sft-lcg-rm after 600 seconds'

Cause:

lcg-rm testing fail with timeout limit, rather than execution failure, or data transferring problem. Severity constrains to wn itself, but we're not able to reproduce also the sanity testing errors again.

Severity:

The severity is limited to the time stamp only, and rests of site services are functional normally.

Solution:

We're not able to reproduce the error encountered, and all lcg-rm testing pass but super testing fail with timeout error. SAM testing event pass and recover itself.

Problem:

SAM job submission error has been detected by another CE, quanta.grid.sinica.edu.tw, starting from '14-Aug-2007 10:56:46', and end at '14-Aug-2007 15:11:39', at least 6 events observed

Cause:

again, we're observing lots of disk space are full in quanta backend computing nodes, seems user cert in globus tmp are referring to same DN observed in w-ce01.grid earlier this morning. (O=GRID-FR, C=FR, O=CNRS, OU=LRI, CN=Alexandru Munteanu).

Severity:

## LHC Computing Grid Project

Due to the root are full, no further job execution are able to sending into WN allocated by batch system, that all SAM job submission testing fail at quanta CE.

Solution:

Have deliver simple script to check the account usage, and remove forcibly all globus-tmp if larger than 1GB. By scanning all WN, we have detected 6 WN are suffering from disk space outage error, that we have scan all pool accounts as well as all CPU farm and make sure the other WNs wont have same problem. After forcibly removing globus tmp, seem all are back to normal including also SAM functional testing. Validations from batch server, with at least 25 testing jobs, are all able to carry out final results from backend WN.

Problem:

SAM job submission error observed at site CE, lcg00125, starting from 14-Aug-2007 01:11:44, and end at 14-Aug-2007 03:16:41, total 4 events

Cause:

The root cause of job submission failure is due to some of the WN disk space has been full due to the Biomed jobs. Some of them have been occupied around 30GB, and still some occupied more than 60GB. Jobs sending to these WN will fail eventually, due to the lack of free space for cache export generating by job submission via globus gatekeeper.

Severity:

Due to the root of affected WNs are full, w-ce01 have fail with 4 job submission testing observing from SAM. The impact are limited since the responding between job failure and dispatching might not coherent and could have sort of delay that bulk job submission from experiment production could still allocated to WNs that are healthy.

Solution:

Identify the WN with simple host base authentication testing, and extract also from batch server for those job submitted by specific VO, have forcibly offline them immediately from batch pool and remove job id to avoid confusion with maui scheduler. Re-online these WNs and add them into batch again after disk space outage problem being cured.

### > 2007-08-18

Problem:

Replica management testing failure has been observed at '18-Aug-2007 05:20:50' and '18-Aug-2007 06:09:40', with error messages 'No information found for SE: dpm01.grid.sinica.edu.tw'. Besides to this, we also have job submission error found at '18-Aug-2007 10:23:43, 18-Aug-2007 14:12:24, 18-Aug-2007 12:12:02, 18-Aug-2007 15:11:48, and 18-Aug-2007 16:12:27' due to the same domain name lookup failure.

Cause:

The root cause of the lcg-rm testing failure is arising from connectivity of Sinica root DNS that was inaccessible during the maintenance window of computing center. Where top BDII fail to query GIIS of all regional resource centers including also ASGC info sys, the problem was discovered late in the evening, which is 9 hours later.

Severity:

The severity have covering also the impact to the GFAL access from all resource centers in AP region, where wrong DNS entries are adopted in resolve config file. This includes also the impact for the only DPM server accepting OPS rm testing.

Solution:

## LHC Computing Grid Project

By switching to ASGC central domain name server had resolved all the problem detected by SAM.

---

### BNL

#### > 2007-08-01

Problem:

DCache information provider does not report dCache information.

Cause:

The host name of the dCache information provider is changed. Detailed please see Wednesday report.

Severity:

FTS data transfers relying on BDII information fail.

Solution:

See Wednesday report.

#### > 2007-08-02

Problem:

Starting Monday and getting worse on Tuesday, load on the primary OSG gatekeeper has been increasing, and the swap usage has been become critical. On Wednesday, load is 600, 2GB swap. System spends little time in user--mainly does swap. 8-minute log in time is observed. Nagios gatekeeper timeouts (presumably load related) warning was sent out to our OSG administrator.

Cause:

We did initial assessment, observed:

-able to log in

-java memory and cpu somewhat heavy (CEMon and GT4 globus-ws)

-close to 3000 globus-job-manager processes. 1500 from Neng Xu in usatlas, and most of the rest from 3 Geant4 users.

-our OSG administrator concurs this is an unavoidable outcome of the inability to throttle globus job submission, and the load is probably caused by the perl jobmanagers (one per submitted job).

-System is software RAID5, so once swapping begins load spikes quickly.

-our system administrator noted that processes doing heavy IO, even though they are in a wait state, are still counted as contributing to load. This is supported by the fact that although the load is 400-600, and the system is slow, it is still functional (all commands get executed after a couple minutes).

Severity: the OSG gatekeeper is sluggish, cause the Nagios probe time out. RT tickets were opened to track this problem.

Solutions:

-collected all sorts of ps, lsof, df,logs, etc. information and copied to other host.

-restarted CEMon

-disabled GT4 WS

-killed 1000 Neng Xu jobs

-killed all geant4 jobs

-disabled geant4 in GUMS

Current Status:

-load down to 400, then down to 80

-reniced root login responsive.

-load just really dropped when I killed grid7790''''s jobs (Patricia Lorenzo -- geant4).

## LHC Computing Grid Project

```
[root@gridgk01 post-install]# ps aux | awk '{ print $1}' | sort |
uniq -c | sort -rn
  218 root
  211 agrd0093
  108 grid7915
   47 grid7941
   38 condor
   27 grid0040
   26 grid7939
   20 grid7916
   15 grid7790
   10 usatlas1
```

### Further Solutions:

- Our OSG administrator has troubleshooting info (logs, ps output) and will continue to investigate.
- Our OSG administrator will e-mail geant4 users to ask about their activities and to reduce job submissions.
- We will both keep watching the server for load and activity.

Problem: dCache does not publish information to LCG BDII.

Cause: the host name (dcsrcm02.usatlas.bnl.gov) for dCache information provider needs to be recycled and reassigned to a new node in preparation for the dCache upgrade. The new hostname for dCache information provider is dcinfor.usatlas.bnl.gov. Before LCG BDII is reconfigured to point to the new hostname, dCache information is not published into BDII.

Severity: the data transfer based lcg-util failed because of losing dCache information. The SAMS tests failed between the time when the host name for the dCache information provider was changed and the time when the new host name was granted by ITD networking group and the dCache information provider was reconfigured with the new host name.

Solution: Reconfigure LCG BDII to use the new host name/IP address for dCache information provider.

### > 2007-08-03

#### Maintenance:

One of GridFtp servers was taken out production and added into the dCache testbed.

### > 2007-08-04

Friday August/03/2007

#### Problem:

Data transfer from Canada to BNL failed. The file size in BNL dCache is 0.

#### Cause:

ATLAS DDM software had a bug, and registered the failed data transfer as success.

Severity: For those data files produced by Panda-based Canada WLCG sites, the LRC data catalog entries might be inconsistent with the BNL dCache.

#### Solution:

ATLAS production team fixed the software bug in ATLAS DDM. It only registers successful data transfer.

### > 2007-08-05

Friday August/03/2007

## LHC Computing Grid Project

**Problem:**

Data transfer from Canada to BNL failed. The file size in BNL dCache is 0.

**Cause:**

ATLAS DDM software had a bug, and registered the failed data transfer as success.

**Severity:** For those data files produced by Panda-based Canada WLCG sites, the LRC data catalog entries might be inconsistent with the BNL dCache.

**Solution:**

ATLAS production team fixed the software bug in ATLAS DDM. It only registers successful data transfer.

**> 2007-08-06**

Problem: The ATLAS DDM problem continues.

**> 2007-08-07**

None

**> 2007-08-08**

Problem: ATLAS DDM server managing BNLDISK/BNLTAPE storage for Tier 0 data export and AOD data distributed does not respond to data transfer requests.

**Cause:**

The transfer database of the ATLAS DDM software stack used large fraction of local disk space. Further ATLAS DDM operation requesting more disk space failed.

**Severity:**

The AOD data distribution performance was impacted since August/04/2007.

**Solution:**

The transfer database is in the process of migrating into a more reliable hardware. Therefore, no major fix is planned. ATLAS DDM administrator will keep track of the server status, and clean up disk space when it approaches disk limits. New hardware replacement will be added in two weeks. Disk monitoring tools will be in place in the new hardware to report disk full errors in future.

**> 2007-08-09**

**Problem:**

The network problems on one of core RHIC switch hosting HPSS services returned around 3:00PM afternoon. Users experienced flaky network connectivity.

**Cause:** another legacy CISCO line card was found in the switch chassis, it brought down the backplane speed and caused ip packets drops.

**Severity:** HPSS service was affected, and provided unreliable data service. The unstability problem with HPSS propagated into dCache services, and ended up failed data into/from dCache. Nagios server sent out large number of false alarms.

**Solution:**

Measures have been taken to boost performance but the benefits may be short-lived. BNL Network engineers planned to migrate the servers on the legacy line card to a newer one. The work will be finished on Thursday afternoon. Some fiber cable needs to be replaced under floor to connect new gbics next week.

**Problem:**

## LHC Computing Grid Project

HPSS needed an emergency maintenance due to its core service deterioration.

Cause:

HPSS had five stuck LTO 3 tape drives. The tape drive problem correlated with the under-floor maintenance work which might nudge the HPSS core network cable.

Severity:

HPSS lost data service between 7:00AM and 10:00AM.

Solution:

The HPSS system has to be restarted to recover the service.

### > 2007-08-10

Problem:

HPSS has flaky services on Thursday.

Cause:

The Network problem continued to disrupt the HPSS core services

Severity:

Data read and write operations with HPSS were suspected between 2:00PM and 4:00PM.

Solution:

HPSS group recycled HPSS and brought it back on-line on 4:00PM.

### > 2007-08-18

Friday: August/17/2007

One user reported that he could not transfer USATLAS production files around 4:00PM.

Cause:

The problem is down due to the expired CRL files on the myproxy server at BNL. It looks like the CRL files have not been updated since July 18th on the machine. The CRL cron process appears running, but does nothing.

Severity:

The new data transfer requests sent by FTS failed because of the authentication error with the MyProxy server. The downtime is within twenty minutes.

Solution:

The CRL cron process was restarted, and the CRL file was updated. Then we triggered an update of the both the CA file RPMs and subsequently a CRL update on 5:00PM. User confirmed that the problem was fixed about. We've initiated the process to get a monitoring probe for our local Nagios system to check the age of CRLs (in addition to host cert checking which it does already).

We will the ticket open a while in case there are any follow up comments.

Friday August/17

A user couldn't access SRM

Cause: kpwd not updated because GUMS wasn't installed as it wasn't in

the satellite for 64 bit machines

Severity: few users affected

Remediation: install gums by hand

Friday 17 - Sunday 19

This problem happened twice, Friday evening and Sunday at noon. Some requests would hang up.

Cause: the first request to a file would generate a suspended dcache requests for HPSS transfer. The subsequent requests for the same file would hang.



## LHC Computing Grid Project

Severity: problems on some files in USATLAS production

Remediation: force retry prestage request from the admin console for those files

### > 2007-08-19

Between midnight to noon, SRM was down

Cause: PNFS was out of disk space due to high level of log. Also, the default value for a logging parameter of the new version of Postgres didn't work as expected

Severity: dCache down, USATLAS production down

Remediation: fix log

### > 2007-08-20

From 19:20 Saturday night to Monday 20 morning, SRM wasn't working properly.

Cause: too little memory was assigned to SRM

Severity: SRM down, USATLAS production down

Remediation: restarted and reconfigured SRM

### > 2007-08-21

From 19:20 Saturday night to Monday 20 morning, SRM wasn't working properly.

Cause: too little memory was assigned to SRM

Severity: SRM down, USATLAS production down

Remediation: restarted and reconfigured SRM

Problem:

dCache is completely down.

Cause:

Log area on PNFS is full.

Severity: dCache is out of service for a few hours.

Solution:

BNL dCache is back after moving the location of the logs to a big disk areas. Log level will be adjusted to avoid overflow the disk. The monitoring script is going to be put into place to raise alarms.

Problems:

Partial loss of cooling at 21:02:03. It cannot be resolved at this time unfortunately. The right equipment and technical expertise is not available to resolve this problem until tomorrow morning (Tuesday, August 21).

Cause: N/A until the technical expertise is available to diagnose the problem.

Impact:

Have to shut part of Linux farm and distributed storage to reduce the heat load.

Solution:

To minimize the risk of damage to high-value computer equipment due to elevated temperatures, it was decided that a partial shutdown of the Tier 1 Linux Farm at Brookhaven (on the order of 31% of the total computing capacity and 19% of the distributed storage capacity) is necessary for the remainder of this evening (Monday, August 20).

## LHC Computing Grid Project

### > 2007-08-22

Problem:

Partial loss of A/C on Monday evening.

Cause:

Plant Engineering technicians have found two small leaks in the cooling hoses, and they believe this is responsible for the partial cooling failure last night.

Impact:

We have to shut down a fraction of Tier 1 Linux farm to reduce load. Some distributed dCache storage was stopped, and users will experience time out if their requested data files happened to reside on the affected storage system.

Solution:

The plant Engineers repaired the leaks and restored the cooling system around 15:45PM.

We have restarted about 40% of the equipment powered down last night, since the

Temperature in the computing facility has been stable and low overnight in the Tuesday morning. . The remaining systems were restarted after the cooling system was repaired by Plant Engineering last afternoon.

Problem:

Some user files are not accessible on the disk areas.

Cause:

dCache was affected by the cool problems.

Severity: 19% percentage of dCache disk storage is not accessible.

Solution:

A/C was fixed late afternoon. The entire dCache storage was resumed.

Problem:

The panda monitor apache started getting errors about libcurl incompatibility with python. A simple "df" command hangs up.

Cause:

dCache PNFS server hangs due to the log space overflow. The panda monitoring server mounts PNFS directories. The panda server happened to do file accesses to the PNFS directories during PNFS server outage. The PNFS directories can not be recovered after PNFS was brought back to service.

Severity:

USATLAS Production run without the monitoring pages for Tuesday.

Solution:

Unmounting and remounting the PNFS directories fixed the problem. We tested it with lcg-util commands, and confirmed that system/service was recovered on Tuesday night 9:30PM.

SRM down

Cause: ran out of disk space due to log. Only one of the logs was changed to the log partition

Severity: SRM down for a few hours

Remediation: log reconfigured

### > 2007-08-23

Client would hang while accessing some files

## LHC Computing Grid Project

Cause: the cooling failed on some nodes on the farm, so part of the read pool disappeared

Severity: some files were slow to access as they had to be taken from HPSS

Remediation: None. dCache restored the files given some time

Problem:

One of Panda Database service becomes un-accessible around Tuesday afternoon around 5:30PM.

Cause:

The power supply on this server failed due to loss of A/C from Monday night to Tuesday afternoon.

Severity:

USATLAS production continues with the existing jobs running, but no new jobs can be submitted to prevent job metadata from being corrupted.

There was no new jobs dispatched between 5:30PM Tuesday afternoon and Wednesday 11:30AM.

Solution:

Replaced the faulty hardware, and restarted the database server, and performed sanity check. All of these operations were completed before 11:30AM Wednesday morning. A long term solution to replace the aging hardware is being discussed. We will add redundancy to the critical database servers.

### > 2007-08-24

SRM transfer to the ITB would fail because of "Unknown CA" from Thursday afternoon till Friday at 16:00rCause: certificate directory was not setup correctly. When fixed, the server didn't reload the new CA files because didn't really restart due to a Solaris/Linux incompatibility

Severity: SRM down in ITB

Remediation: fixed the startup script and the CA file, then restarted the server

FTS/DQ2

Thursday 23

DQ2 down for BNLDISK and BNLLTAPE Cause: DQ2 was upgraded, but the new version had a problem. Needed a fix from the developers

Severity: DQ2 down for BNLDISK and BNLTAPE for a few hours

Remediation: once the patch arrived from the developers it was applied

### > 2007-08-25

SRM and SE down in SAM. BNLTAPE occasionally failed for Miguel Branco. Cause: kpwd files were stale and were not properly generated. The script that was supposed to run as cron job was failing as the wrong shell was used

Severity: only users with a newly registered certificate were affected.

In particular, certificate added between the 16 and the 19 where failing only on one door out of eight, making the behavior chaotic. USATLAS production was not affected.

Remediation: refreshed the kpwd files and fixed the script

### > 2007-08-26

SRM and SE down in SAM. BNLTAPE occasionally failed for Miguel Branco. Cause: kpwd files were stale and were not properly generated. The script that was supposed to run as cron job was failing as the wrong shell was used

## LHC Computing Grid Project

Severity: only users with a newly registered certificate were affected.  
In particular, certificate added between the 16 and the 19 where failing only on one door out of eight, making the behavior chaotic. USATLAS production was not affected.

Remediation: refreshed the kpwd files and fixed the script

### > 2007-08-27

SRM and SE down in SAM. BNLTAPE occasionally failed for Miguel Branco.  
Cause: kpwd files were stale and were not properly generated. The script that was supposed to run as cron job was failing as the wrong shell was used

Severity: only users with a newly registered certificate were affected.  
In particular, certificate added between the 16 and the 19 where failing only on one door out of eight, making the behavior chaotic. USATLAS production was not affected.

Remediation: refreshed the kpwd files and fixed the script

### > 2007-08-28

SRM and SE down in SAM. BNLTAPE occasionally failed for Miguel Branco.  
Cause: kpwd files were stale and were not properly generated. The script that was supposed to run as cron job was failing as the wrong shell was used

Severity: only users with a newly registered certificate were affected.  
In particular, certificate added between the 16 and the 19 where failing only on one door out of eight, making the behavior chaotic. USATLAS production was not affected.

Remediation: refreshed the kpwd files and fixed the script

### > 2007-08-29

Justin Ross couldn't transfer files from BNL

Cause: client dual nic was sending wrong IP to the server, which couldn't connect back

Severity: one user affected

Remediation: helped used configure his client

SRM was down and was automatically restarted

Cause: out of memory error

Severity: downtime of less than 20 minutes

Remediation: increasing memory setting

### > 2007-08-30

Wednesday 29 - Thursday 30

Starting at 6pm, high PNFS load. Thursday morning performance of SRM and gridFTP deteriorated (5 seconds to connect), and caused timeout errors.

SRM and SE down in SAM for a few hours

Cause: USATLAS tests on dq2\_cr jobs kept retrying transfers that were failing, causing to add and remove entries in PNFS. 188 such jobs were submitted. The high PNFS load caused all of dcache performance to deteriorate

Severity: Production at BNL was affected: 90% failure for local production as dccp would fail. Overall dCache performance was affected. DQ2 transfer also were affected by timeouts for both SRM and gridFTP

Remediation: contacted the developers, and killed the dq2\_cr jobs

## LHC Computing Grid Project

---

### CERN-PROD

#### > 2007-08-03

Problem: All CEs were failing job submission test for the whole morning. The reason was that the LSF batch system was overloaded due to one local user that submitted ~80000 jobs on Wednesday evening.  
Solution: The CE managers had to kill the jobs and drain the queues.

#### > 2007-08-15

There was a scheduled network intervention between 8am and 8:30 (CEST) for which the batch system was paused. This intervention was announced but it was entered into GOCDB too late for the SAM tests which reported a failure for some time after. The site was fully functional again at 9:00 CEST. See also scheduled interventions.

#### > 2007-08-23

Planned intervention: router upgrade (batch/CE paused) + castorpublic/SE upgrade

---

### CNAF

➤ *Please report on the site unavailability period(s), see page 2*

#### > 2007-08-02, 2007-08-03, 2007-08-04, 2007-08-07/08, 2007-08-10/11

CASTOR problems

#### > 2007-08-16/17, 2007-08-19/21

SAM tests for Site GIIS failed due to a problem with the sensor. The issue has been discussed on LCG ROLLOUT mail list.

#### > 2007-08-24

Downtime scheduled for general site shutdown, due to AC power upgrade. Site will come back next week, hopefully wed the 29th.

#### > 2007-08-30/31

Downtime extended due to network problems

---

### FNAL

#### > 2007-08-06

USCMS was operational, trouble with test.

#### > 2007-08-08

USCMS was operational, trouble with test.

## FZK-LCG2

➤ *Please report on the site unavailability period(s), see page 2*

> **2007-08-03**

See above (gridftp).

> **2007-08-05**

2007-08-04 - 2007-08-05:

Problem: sBDII SAM test failed only once (on 2007-08-04 17:33:26), no new SAM test was done before 2007-08-10

Solution: Ticket was already opened last week

> **2007-08-06**

2007-08-05: See above.

> **2007-08-07**

2007-08-06:

Problem: SAM tests failing because of BDII timeout

Solution: -

> **2007-08-08**

2007-08-07: See above

> **2007-08-15**

Date: 14-08:

Problem: SRM unstable

Solution: restarting SRM several times

> **2007-08-16**

Date: 15-08:

Problem: SRM unstable

Solution: restarting SRM, dropped SRM DB, restarted headnode, informed developers.

> **2007-08-17**

Date: 16-08:

Problem: SRM unstable, even restarting SRM not working any more

Solution: rebooting SRM machine

> **2007-08-18/21**

Date: 18-08 to 21-08

Problem: centralSE @ CERN (lxdpm1001.cern.ch) not included in the FZK BDII after migration of BDIIs to new machines. As a result, command lcg-rep to the CERN machine fails.

Solution: Configuration fixed.

Severity: Zero. Only affects SAM tests.

## LHC Computing Grid Project

### > 2007-08-23/27

Date: 23-08 to 27-08

Problem: SAM tests need an rpm to check "lcg-version" which is missing in the WN distribution 3.0.24-1 after upgrading.

Solution: Ticket opened to developers, extraction of the missing rpm from WN package 3.0.22 and distribution on the production cluster. Configuration fixed.

Severity: Zero. Only affects SAM tests.

### > 2007-08-24

All major (>1 hour) this week stem from the a problematic host change of the BDII service. The actual cause is still unknown and there will be no followup.

Severity of this impact was almost NIL because the the SFT failure was a replication test (to the CERN SE) which is not used by the VO's

anyway. The latter was confirmed in the operations meeting last Monday.

### > 2007-08-28

24-08-07 / 27-08-07 availability < 50% though VOs not affected. Caused by a

missing rpm in the wn glite package. The trouble ticket is here:

<https://gus.fzk.de/ws/overview.php?ticket=26248>. According to the bug report, the problem is fixed in the next update. GridKa will worked around the version info manually.

Severity:none

---

## IN2PCC

➤ *Please report on the site unavailability period(s), see page 2*

### > 2007-08-02

### > 2007-08-16

### > 2007-08-24

### > 2007-08-28

### > 2007-08-30

---

## NGDF

### > 2007-08-11

Problem: srm2 was not configured to allow the OPS vo

## LHC Computing Grid Project

Solution: srm2 was configured to allow the OPS vo

### > 2007-08-12

Problem: srm2 was not configured to allow the OPS vo

Solution: srm2 was configured to allow the OPS vo

### > 2007-08-13

Problem: srm2 was not configured to allow the OPS vo

Solution: srm2 was configured to allow the OPS vo

### > 2007-08-14

Problem: srm2 was not configured to allow the OPS vo

Solution: srm2 was configured to allow the OPS vo

---

## PIC

### > 2007-08-01

Date: From 30/07/2007 at 17:00 UTC until 31/07/2007 08:00 UTC

Problem: The SRM-disk service becomes unavailable. The Slony application for backing up the PNFS DB was started on 30-Jul in the evening, and it overloaded the DB resulting in an unavailability of the PNFS and hence the whole SRM-disk service.

Solution: Killing the Slony DB replication processes and restarting the PNFS immediately solved the problem. We are currently studying if there is a way to restart the Slony in a non-destructive way, or if we need to schedule a downtime in the future to do this.

Severity: High. The SRM-disk service was completely unavailable during the affected period. All the transfers from/to PIC for all VOs failed.

### > 2007-08-02

Date: 01/08/2007 5 UTC (only 1 hour)

Problem: We do not understand this one hour unavailability. The detailed SAM results for the CE ce06.pic.es for instance do not show an ERROR at that hour.

### > 2007-08-05

Date: from 04/08/2007 at 08:00 until 05/08/2007 at 08:00

Problem: We believe this is a SAM problem. Seems SAM detected problems in our site-bdii, but this was not the case since we could check that the service was running ok.

Solution: We did not do anything. We believe this was a SAM failure, then we opened a GGUS ticket on SAM, GGUS-25458.

Severity: None. This was not a real problem. Just a false negative from SAM.

### > 2007-08-06

Date: from 04/08/2007 at 08:00 until 05/08/2007 at 08:00

Problem: We believe this is a SAM problem. Seems SAM detected problems in our site-bdii, but this was not the case since we could check that the service was running ok.



## LHC Computing Grid Project

Solution: We did not do anything. We believe this was a SAM failure, then we opened a GGUS ticket on SAM, GGUS-25458.

Severity: None. This was not a real problem. Just a false negative from SAM.

### > 2007-08-29

problems with dcache information system. The ldap server hanged and we were not publishing correctly the information about srm-disk.pic.es. Problem solved.

### > 2007-08-30

problems with dcache information system. The ldap server hanged and we were not publishing correctly the information about srm-disk.pic.es. Problem solved.p

---

## RAL-LCG2

### > 2007-08-30

Unavailability due to local networking (firewall) problem.

---

## SARA – MATRIX

➤ *Please report on the site unavailability period(s), see page 2*

### > 2007-08-1/2

Problem with a dcache pool node.

### > 2007-08-04/5

Moved old FTS to a new server. Was unable to remove old FTS server from GOC database while SAM tests were still directed towards the old FTS server.

### > 2007-08-06

Problem: The sitebdii had allegedly problem.

Solution: We are not aware of any problems and it works just fine as far as I can see.

### > 2007-08-14/16

Globus-mds has crashed on CE. Postgres problem with the dCache pnfs database. Solved by doing a manual vacuum.

### > 2007-08-20

Gfail\_read test was hanging due to a firewall problem which caused the SAM job to run past the wallclock limit.

## LHC Computing Grid Project

### > 2007-08-29

Full root file system on gridftp door nodes caused by excessive logging of dCache gridftp door. This problem has been submitted to the dCache developers which passed it on to FNAL.

### > 2007-08-30

Gridftpdoor has crashed.

### > 2007-08-31

Unknown dCache-related error that went away by itself.

---

## TRIUMF-LCG2

➤ *Please report on the site unavailability period(s), see page 2*

### > 2007-08-01

Replicate to lxdpm101.cern.ch failed.  
d-cache head node has been upgraded to 1.7.0.39 - caused short outages.

### > 2007-08-02

Dcache pool nodes have been upgraded to 1.7.0.39 - caused short outages.

### > 2007-08-09

```
+ lcg-cr -v --vo ops file:/home/samops/.same/SRM/testFile.txt -l lfn:SRM-put-srm.triumf.ca-1186555082 -d srm.triumf.ca
globus_ftp_control_connect: globus_libc_gethostbyname_r failed
apparent DNS problem
```

### > 2007-08-12

Job submission failure - reason unknown

### > 2007-08-15

dache srm port 8443 blocked twice today. Restart fixes it but reason unknown.  
Hardware arrives and physical installation begins.

### > 2007-08-30

---

## SAM unavailability

None reported.