Table of Contents

WLCG Tier1 Service Coordination Minutes - 3 June 2010

Attendance

Jean-Philippe, Patricia, Roberto, Manuel, Harry, Jamie, Maria, Simone, MariaDZ, Dirk, Lola, Eva, Maarten, Nilo, Kate, Julia, Tony, Alexander, Andrea, Tony, Dan, Helge, Tim

Site	Name(s)
CERN	
ASGC	Felix Lee (ASGC),Jhen-Wei Huang
BNL	Carlos Fernando Gamboa
CNAF	cristina vistoli Infn-Cnaf
FNAL	Jon
KIT	
IN2P3	Fabio Hernandez
NDGF	Jon - NDGF
NL-T1	Alexander Verkooijen
PIC	
RAL	Gareth Smith, Carmine Cioffi
TRIUMF	Andrew

Experiment	Name(s)
ALICE	
ATLAS	Pavel Nevski, Stephane Jezequel, Florbella, Gancho, Alexei, Gancho, John DeStefano, Dario Barberis
CMS	
LHCb	Ricardo Graciani

ATLAS DB Patch incident

Eva - On Tuesday 1st June started patching offline DB at 14:00. Observed problem with clusterware which caused all services to go down. Communicated and decided to continue in non-rolling way. Online DB patched day before without any problem. At same time as applying patch received mail from ATLAS saying performance problems on ATLAS online. Started to investigate and saw same problems on ATLAS offline (and LHCb). Patch applied introduced new bug affecting mainly DBs were auditing enabled and COOL used to access DB. On Wednesday when we had complete clear picture notified ATLAS and asked for permission to roll back patch in a rolling way asap. Started rolling back ATLAS online and continued with offline. During rollback on offline one node went down due to corrupted file system - also high load problem. Florbella - since yesterday at 17:00 haven't seen any anomalies. From ATLAS viewpoint 1h of downtime yesterday due to rollback + 1h due to initial problem. A lot of services affected - intermittent problems. Panda etc do not use COOL. Alexei - comment: for us we have 1 day with unstable situation for all applications using offline DB. What is procedure to avoid such situations? Is it tested on testbed? Eva -patch installed first on test DB then on integration DB. Each experiment has 1-2 integration DBs where patch is validated. Maria - any COOL related apps on integration? Florbella - yes, several. On online did see some spikes but not correlated. Only offline experienced service degradation. Carlos - could you disable auditing before rolling back to see if this changes things? Dawid - no, didn't have time. Rolled back on LHCb online. Decided to roll back as similar symptoms on LHCb offline. Also something from RAL today - same problem with patch. Carmine - in process of applying patch - problem while doing patch. Symptoms seemed similar. Opened Service Request. Tried to follow requests from Oracle for more info but decided to roll back. Andrea - for LFC or COOL? Carmine - DB is FTS and LFC for ATLAS. Carlos - my experience was I could do patches in

rolling fashion. Carmine - applied PSU on 3D without problem. (But these use COOL...) Dawid - 2 problems: services went down due to clusterware. 2nd problem is patch introduces bug with auditing problems. DIrk - do we know what security impact of not applying patch is? Maybe sites should leave this one out... Carlos - good idea: also consolidated SR which could be addressed by Oracle. Florbella - is patch still on archive and integration? Left in integration. Maria - please give list of production DBs that you will roll back with list of applications. Andrea - don't say COOL but something more specific. Simone - for Tier1s suggestion is not to apply patch. RAL rolled back yesterday.. Are there other T1s in trouble? DB can be shared with other things. DB can be co-hosting many things. Even if bug didn't manifest itself yet... Eva - if auditing enabled and COOL at anytime suggest to rollback. Simone - access via COOL might not be that often.. Rolled back ATLAS & LHCb online and offline. ATLAS requested also archive DB. Alexei: we want to see this symptom on integration db. Nilo - you will see problems like this in the future. Carlos - at BNL affected by this bug in terms of high CPU usage or connections from T2s or connections from WNs. In my case would lilke to understand more about follow up of SR by Oracle. if service at BNL not affected at same way as at CERN would like to wait until clear from Oracle. At BNL DB is outside security firewall hence patches important. This concerns conditions DB cluster. Florbella - try to reproduce for SR. Eva - need to do so for SR to progress. Nilo - bug acknowledged in bug DB. Can ask for backport to existing release.

Results from test on Alarm Chain

MariaDZ: motivation: to test end-end workflow for GGUS alarms.

- Email notification reception by the experiment experts, members of -operator-alarm@cern.ch
- Email notification reception by the CERN operator on duty and existence of procedures per WLCG Critical Service.
- Quick and correct assignment in CERN Remedy PRMS to the right category.
- Acknowledgment and ticket update by the CERN IT Service manager.

Outcome:

- The exercise remained unclear for some experiment members till the end despite the documented steps-to-follow and the agreed CriticalServices twiki.
- Test ALARM tickets for the network service reclassified by ROC_CERN (now IT/PES) to IT Services-Network-Netcom-All in PRMS still remain in progress with NO update from the service.
    - ♦ Maybe a problem with choice of Remedy category - to be checked. Action: IT-PES to find out which is right category.
- Test ALARM tickets for the VOboxes service showed the lack of operators procedures.
    - ♦ Manuel - procedures exist. MariaDZ - someone should close ticket then.
    - ♦ Roberto - was agreed that LHCb VOboxes were not in scope of this. Maria - if wanted experiments could send basic VO box alarm.

Conclusions:

- IT services participating in the WLCG daily meeting were aware of the exercise and alert to respond and close tickets as solved .
- It is worth to check that full procedures do exist for all services and for real cases.
- grid-cern-prod-admins was added in the 4 -operator-alarm e-groups for faster notification of the service managers.

Interventions for next technical stop

- Conclusion from earlier is that no more than 3 sites (serving a given VO) nor 30% of total resources (again for a given VO).

- RAL (Gareth) - for next technical stop (28-30 June) have some transformer work - site "AT RISK" planned.

- FNAL(Jon) - nothing planned

- CNAF(Cristina) - nothing planned for the moment.

Hammercloud

Presentation by Dan van der Ster.on potential use of Hammercloud by LHCb.

Summary:

- HammerCloud is a DA functional and stress testing system used widely by ATLAS and coming soon for CMS
- Two basic use-cases:
    - ♦ Continuous stream of test jobs to measure site availability
    - ♦ Enable central managers to define standardized (stress) tests, and empower site managers to invoke those tests on-demand.
- An HC-LHCb plugin would leverage the existing GangaLHCb work
    - ♦ A prototype plugin would not take significant effort

- Andrej TSAREGORODTSEV - have to compare with LHCb needs, in particular our computing model with analysis at a small number of sites (T1s).

- Roberto - some sites, e.g. PIC & Lyon, asked for a way of testing sites. This would be a useful tool for sites.

- Ricardo - have some production testing jobs. Including some analysis jobs would be trivial and could also be trivially extended to allow others. But could still look at HC as a tool.

- Andrej - strong point of HC Is presentation.

Data Management & Other Tier1 Service Issues

glexec

- LCG-CE test results☑
- CREAM test results☑

Site	"/ops/Role=pilot" job + glexec test	glexec capability in BDII
ASGC	CE/batch authZ error (5 CE) or glexec absent (1 LCG-CE)	-
BNL		
CERN	lcg-CE OK, CREAM fails	absent for 1 CREAM
CNAF	home directory absent	-
FNAL	OK for CMS	OK
IN2P3CC	CE authZ error or glexec absent	-
KIT	OK	OK
NDGF	n/a	n/a
NIKHEF	lcg-CE OK, CREAM authZ error	OK
PIC	OK	OK
RAL	OK	OK

SARA	argus installed and configured, glexec installation nearly completed	-
TRIUMF	OK	-

Storage systems: status, recent and planned changes (please update)

Site	Status	Recent changes	Planned changes
CERN	CASTOR 2.1.9-5 (All) SRM 2.9-3 (all)	None	A hotfix will be required for the problem seen with CMS on 1/6/2010. Code under development.
ASGC	CASTOR 2.1.7-19 (stager, nameserver) CASTOR 2.1.8-14 (tapeserver) SRM 2.8-2	None	None
BNL	dCache 1.9.4-3	?	?
CNAF	CASTOR 2.1.7-27 (ALICE) SRM 2.8-5 (ALICE) StoRM 1.5.1-3 (ATLAS, CMS, LHCb,ALICE)	None	waiting for green light from ALICE to dismiss CASTOR (to be discussed at today's TF meeting); StoRM endpoints to be upgraded soon with bugfix release, no dates agreed yet; configuring the new disk to meet the pledges, ready for users hopefully next week
FNAL	dCache 1.9.5-10 (admin nodes) dCache 1.9.5-12 (pool nodes)	?	Adding 2010 disk in June (~2 PB), retiring 0.7 PB in July, no downtimes foreseen
IN2P3	dCache 1.9.5-11 with Chimera	nothing to report	nothing to report
KIT	dCache 1.9.5-15 (admin nodes) dCache 1.9.5-5 - 1.9.5-15 (pool nodes)	none	Change of authentication method on ATLAS dCache instance planned for 8th of June. A restart of SRM and gPlazma is needed.
NDGF	dCache 1.9.7 (head nodes) dCache 1.9.5, 1.9.6 (pool nodes)	?	?
NL-T1	dCache 1.9.5-16 with chimera (SARA), DPM 1.7.3 (NIKHEF)	?	?
PIC	dCache 1.9.5-20rc1	Installed on 1-June-2010, to test the tape protection patch	As soon as the version 1.9.5-20 containing the tape protection patch in production we will apply it
RAL	CASTOR 2.1.7-27 (stagers) CASTOR 2.1.8-3 (nameserver central node) CASTOR 2.1.8-17 (nameserver local node on SRM machines) CASTOR 2.1.8-8, 2.1.8-14 and 2.1.9-1 (tape servers) SRM 2.8-2	?	?
TRIUMF	dCache 1.9.5-17 with Chimera namespace	SRM service at 1.9.5-19 (CRL fix)	None

Other Tier-0/1 issues

CASTOR news

dCache news

StoRM news

- bugfix release expected quite soon

DPM news

- DPM 1.7.4 awaiting staged roll-out

LFC news

- LFC 1.7.4 awaiting staged roll-out

- (JPB) LFC 1.7.4-6 certified, 1.7.4-7 mainly bug fixes for Python 2.5 i/f - fixed in new release. Entering certification

FTS

- FTS 2.2.4 patch certified for gLite, in use at TRIUMF, tested at CERN by ATLAS, who concluded the following:
    - ♦ ATLAS would like to ask T1s to upgrade to FTS 2.2.4 since it seems to cure the problems which have been reported and did not show so far any new issue.
    - ♦ In addition, the upgrade seems very simple to roll back, so the risk is minimal.
    - ♦ The most urgent sites for the upgrade are RAL (checksum case sensitivity problem fro files from tape), NDGF and CNAF (affected by the source file deletion in case of failed transfer).
    - ♦ Jon : FTS rolling upgrades?
    - ♦ JPB : never tested, could be enabled in FTS 2.2.5
    - ♦ Maarten : agreed that this is a desirable feature. Noted and will report on progress in coming meetings.
    - ♦ RAL - can't comment on schedule - will inform Mat.
    - ♦ CNAF - will inform Paolo can probably apply patch from next week

WLCG Baseline Versions

- Release planning for WLCG
- Release report: deployment status wiki page
- WLCG Baseline versions: table

Conditions data access and related services

- ATLAS weekly FroNTier meetings

Database services

- Experiment reports:
    - ♦ ALICE: Database patched with latest Oracle security and recommended updates on 31st May.
    - ♦ ATLAS:

◊ Databases patched with latest Oracle security and recommended updates. While patching we experienced some cluster-wide issues that affect connectivity and usage of services on the Atlas offline database so intervention was not rolling as expected.

◊ Since PSU APR10 was applied, large spikes of load (mainly reported in Oracle as 'wait events' related to commit time) every few hours were observed in production databases for a duration of 5-10 minutes. During such high load spikes database access and quality of DB services are compromised. Patch was rolled back.

♦ CMS:

◊ 3rd node of the CMS offline cluster down due to hw failure for one week (vendor could not replace failed hw in the negotiated time).

◊ Databases patched with latest Oracle security and recommended updates. While patching CMS online database it was discovered a problem with the quattor certificates renewal (being followed up). Also, some cluster-wide issues affected connectivity and usage of services so intervention was not rolling as expected.

♦ LHCB:

◊ Databases patched with latest Oracle security and recommended updates.

◊ Same problem with PSU APR10 as ATLAS. Patch rolled back.

◊ Marco: memory problems observed with the LHCb database at Lyon. Is this related to the security patch? Eva: No, IN2P3 has not applied it yet. Osman: problem is caused by limited sga memory. We will increase as soon as possible.

• Site reports:

Site	Status, recent changes, incidents, ...	Planned interventions
CERN	Downstream databases patched with latest Oracle security and recommended updates on 31st May. Also patch to fix the queue monitor processes memory consumption was applied.	
ASGC	ntr	
BNL	COND Database cluster, Tags Test database cluster, BNL FTS and LFC and Tier 3 LFC database cluster were successfully patched (9294403, 9352164). Network service disruption to database services due to a site downtime maintenance. Database service (network) was fully operational after 1:00PM EDT.	
CNAF		
KIT	ntr	At risk on June 29 for FTS, LFC, ATLAS 3D, and LHCb 3D services due to application of the April 2010 Oracle security patches
IN2P3	ntr	April PSU on 8th June on atlas,ami and lhcb databases **cancelled**.
NDGF	20th May: Firmware upgrade on ATLAS DB storage controllers (transparent). 26th May: Install April PSU on ATLAS conditions DB (transparent).	
PIC	HW problem (a LAN card) on a FTS node. Running on one node. Hw changed yesterday but did not fix the problem. Investigating if the cause is the switch. On 01/06 we reorganized bonding on all the serves to grant a major HA.	
RAL	April PSU applied on the 3D databases last week. Problems observed during patching of the FTS/LFC RAC, Oracle SR open, patch rolled back.	
SARA	ntr	April PSU+CPU path on July 1th between 7:00 UTC and 10:00

		UTC
TRIUMF	Applied APRIL2010 PSU	

- Database weekly reports:
  - Pending: ASGC, CNAF and NDGF.
  - NDGF will deploy them next week.

- Recommendation from CERN to Tier1 sites regarding the April PSU patch: for those sites which have applied the April PSU patch on their databases and where auditing is enabled and COOL or similar (multiple sessions connected to one server process) is used to access the database, it is advised to rollback the patch.

AOB

-- JamieShiers - 31-May-2010

This topic: LCG > WLCGTier1ServiceCoordinationMinutes100603
Topic revision: r18 - 2010-06-19 - DiQing