



Barbara Martelli

REVIEW OF TIER1 DB INTERVENTIONS AND SERVICE PROCEDURES

Sources of Information

Information summarized from the following sources about 2008 activities:

- 3D mailing list
- Castor external mailing list
- 3D phone conference minutes and workshops
- Castor F2F meeting
- Castor External operation phone conference minutes
- Additional information from sites
- WLCG workshops and documentation

Open Discussion Session



DB Service Intervention Procedure

- Declare the downtime on the 3D wiki service interventions page: an email is sent to the grid-service-databases@cern.ch.
- If LCG services are impacted, declare the Downtime on GOC DB too: an EGEE broadcast will be sent to all relevant mailing lists.
- Downtime notification rules:
 - Short interventions (less than 4 hours) → one working day in advance
 - Medium interventions (from 4 to 12 hours) → announce to WLCG weekly operations meeting plus EGEE broadcast reminder
 - More than 12 hours → announce at list 1 week in advance via EGEE broadcast, announce to the WLCG operations meeting
- Notification rules on CIC portal:
<https://cic.gridops.org/index.php?section=rc&page=SDprocedure>

Ticketing System

- We use the WLCG GGUS portal
 - <https://gus.fzk.de/pages/home.php>
- A 3D database problem should be notified specifying "3D/database" as type of problem
- But...

GGUS ticket search engine

show columns in search result Ticket-ID Type Virt. Org. Resp. Unit Status Date Last Update Info

Ticket ID: Status: ?

Support Unit: Priority: Order tickets by

VO: Type of problem:

User: MoU Area:

Keyword:

Involved supporter:

Assigned to person:

Affected SITE:

Special attributes (TPM/ENOC/WIKI)

Timeframe

All tickets from

to

Untouched since

new search with default values

show/save search result as: CSV | HTML | XML | PDF

3 Tickets found

Ticket-ID	Type	Virt. Org.	Resp. Unit	Status	Date	Last Update	Info
35929		cms	ROC_DECH	verified	2008-04-30	2008-11-03 00:45	Request to clear CMS Frontier/Squid cache at T1 Gr...
31843		lhcb	VOSupport	solved	2008-01-29	2008-07-18 23:04	LFC failure on lfc-lhcb-ro.cr.cnaf.infn.it (INFN-T...
31223		cms	ROC_DECH	solved	2008-01-11	2008-07-18 23:04	Upgrade of Squid needed.

Big Computing Centers Infrastructural Works

- CNAF down for 2 weeks in March/April
 - A second downtime in summer due to infrastructural works/problems
- ASGC :
 - facilities relocation at data center (computing nodes and also disk/tape servers) in January
 - Schedule Power Maintenance and also for new power generator installation in March.
 - annual power maintenance last October.
- GridKa end of October: Complete GridKa downtime necessary for infrastructure work.
- SARA end of October: suffered a power failure. "We've restarted the database and the streams."

Hardware Upgrades and failures

- PIC May: new setup for LHCb. ATLAS migration proposed to first week of June, cluster database is being prepared. LFC for ATLAS will be migrated 2nd week of June
- RAL May: testing new hw which has arrived
- IN2P3 May: 2 new RAC clusters dedicated to LHCb and ATLAS and AMI databases: quad-core boxes and Linux 64bit. 15G RAM.
 - Databases migration – using transportable tablespaces for ATLAS, datapump import and export for LHCb and [LFC](#).
 - AMI very critical, requesting high availability. Streams to be implemented.
- CERN: IN2P3 databases have been migrated to the new hardware and Streams replication has been reconfigured.
- Triumpf Aug: [FTS](#) migrated to RAC environment.
- Sep NGDF: New cluster being approved. Not final date yet.
- Sep GridKa: Storage intervention planned for next month
- Oct GridKa: end of October (28 evening and 29) full downtime - physical hardware (storage) move
- Oct SARA: network failure, database unavailable during 40 min.

HW/infrastructure maintenance

- SARA june: maintenance OPCORN network
- Aug RAL: 3rd node to be added to ATLAS RAC, foreseen by the end of the month. No storage upgrade is planned.
- Aug BNL: [FTS](#) to be patched (CPU patch) and downtime to upgrade storage firmware + update asmlib packages.
- Sep BNL: Firmware on the storage upgraded. OS patches to be applied (Luca sent an email with the information about the bug).
- Lyon Oct: IN2P3[?]: network failure last Monday linked to a power cut. CPU patch was applied, problem found but cause unknown (local [LFC](#) database). For LHCB and ATLAS databases, there is not plan yet.
- Gridka Oct: Gridka: Full downtime for atlas and lhcb databases, physical storage moved successfully.

Sniped Sessions Killing

- we have observed at TRIUMF and BNL some problems with db clusters getting close to the maximum number of configured sessions. We therefore ask all sites to check their current number of configured and actual concurrent session and will suggest to implement also session sniping in a consistent way to avoid accumulating hanging/idle sessions. We will discuss the problem and the suggested solution(s) with all sites in more detail at the next 3D meeting this Thursday and will agree on the schedule for implementing that at all sites.
- Oracle client sessions and processes: Reconstruction test ongoing on the conditions database, running jobs through Grid. Number of connection was increasing, Carlos increase the number of sessions.
 - Replication affected due to the maximum number of sessions exceeded.
 - Profile used number of session per user higher than the number of processes in the database.
 - After fixing it, replication went well.
 - Also, idle connections observed in the database.
 - Snipe script implemented. Maximum idle time=1 hour.
- Dawid has shared the script for automatic kill of sniped sessions
 - <https://twiki.cern.ch/twiki/bin/view/PSSGroup/KillingSnipedSessions>

Streams (1/2)

- Gridka 08 May: Replication problem from LHCb to Gridka, propagation is aborted with connection lost contact error. SR opened with Oracle.
- SARA May : apply process aborted two weeks ago: no data found. The accounts were not locked at the destination. Rows being updated at the same time at source and destination. Fixed
- CERN june: parameter related to Streams to be included during Tier1s interventions in order to upgrade to 10.2.0.4
- CERN-PIC June: Migration of the ATLAS database at PIC using transportable tablespaces for the Streams resynchronization
- Sep GridKa: [LFC](#) replication problem to Gridka. Propagation job disabled due to error: connection lost contact. Working on this problem together with Oracle support. No solution yet. Workaround: recreate Gridka configuration split from the main Streams setup.

Streams (2/2)

- Oct Triumpf : FTR2 tests - tests moved to a node running apply - caused bottleneck on streams, snapshot too old problems spotted as well
- Oct BNL : apply stuck last week, nothing in the logs, SR open, after moving backups to other node, no problems observed since then
- Oct Gridka: LHCB [LFC](#) propagation problem, diagnostic patch will be applied on Monday
 - Oct Oracle support is working on the problem observed with the propagation at Gridka. They have produced a diagnostic patch which must be installed in the next 7 days. The aim is to produce more trace information next time this issue happens.
- CERN Oct: Problem with the Streams replication for ATLAS problem during unavailability of NDGF. NDGF database was in unresponsive state and the propagation job could not report any problem but LCRs were not consumed. Queue was filled up during the weekend, causing the whole replication system getting stuck one day later due to the lack of memory.
 - Running memory tests in order to identify how the streams memory is used by the spilled LCRs, consumption is increased linearly with the amount of spilled LCRs in the queue.
 - We have allocated a new node to be added to the downstream cluster. With 4 nodes, we will move the downstream databases to run in separate nodes and this will allow us to add more memory to the streams pool.
 - We have received a new patch in order to fix the ORA-600 error when dropping the propagation job and it has been tested successfully. We will apply it on the production databases within the next interventions.

Streams/LFC/SW interventions

- Oct :[LFC](#) migration out of SRM v1 endpoint intervention done on Tuesday, 7th October. Streams replication was stopped during the intervention and the updates where performed at CERN first and then at all the 6 Tier1 sites. Everything went fine and the intervention finished in time except for RAL where the update was not able to finish in more than one day. The data was finally exported from CERN and imported at RAL.

CASTOR

- Taiwan May: deadlock problem with asm instance. Dirk suggested to contact to Castor support because the problem was observed (castor deployment list).
- CNAF Sep:
 - After the upgrade to 10.2.0.4 the agents were postponed and this caused a problem on the database. Agents are upgraded now.
 - Castor middleware upgrade: several sites reported problems after the upgrade when running more than one VOs per cluster.
 - Workaround proposed by Nilo: set init parameter "_kks_use_mutex_pin" to false. Sites observing the problem should open a SR reporting this problem so Oracle support can track the problem and produce the correct fix patch to be included on 10.2.0.4 and 10.2.0.5.
 - Carmine:
 - Getting ORA-600 errors. Oracle support suggested to upgrade to 10.2.0.4. CNAF is already running on 10.2.0.4.
 - The patches identified and circulated by Nilo (for Oracle version 10.2.0.3) must be also identified for version 10.2.0.4.
 - Eva will collect a summary of the configurations at the Tier1 sites.
 - [cnaflastor.rtf](#): Database configuration for castor at CNAF
 - [ralcastor.rtf](#): Database configuration for castor at RAL
- Oct CNAF:Castor: stager and DLF job failing every 1 hour, application and users not affected, looking after this.
- Oct RAL:Castor DB upgrade to 10.2.0.4. Many problems have been fixed. Problem with a process trying to insert data. This behavior only applies to one of the sessions running. Reset the session fixes the problem. RAL is the unique site observing this problem. CERN has tried to reproduce this, without success.

Castor @ RAL

- By the end of July we have moved the Castor databases over to two 5 nodes RACs version 10.2.0.3.
- Soon after we started to have many problems including the
 - ora:600 and between these we saw the
 - deletion of 14000 files due probably to the "cross talk" problem and the insertion of a big number.
- We changes few parameters like the
 - cursor_sharing from force to exact and we added the
 - hidden parameter _kks_use_mutex_pin=false
 - and we adjusted few other parameters but with no much success.
 - unfortunately at the same time we upgraded the Castor schema to 2.1.7.* so we are at the stage where we don't know how much these problems are Oracle or application related.
- Eventually under Oracle support suggestion we have upgraded Oracle to 10.2.0.4. Many ORA errors disappears but we are still having problems like the insertion of big number.

Castor @ ASGC

- April: emergent stager db service intervention due to oracle error code 600, and high CPU arise from from index contention error. all transfer request couldn't insert properly into database
- Oct: stager db recovery - following previous intervention (unscheduled)
- Oct: SRM transfer efficiency degraded due to poorly performing castor instance with long wait times for stage_filequery and/or stage_putdone.
- Oct: castor release upgrade, from 2.1.7-16 to 2.1.7-19-2. in parallel the backend DB SGA size is increasing within the same intervention.
- Sep: castor upgrade from 2.1.6 to 2.1.7 and also move all core services to SLC4/64bit.
- May: castor 2.1.6-12-1 will be patched with 2.1.6-12-2 hotfix and SRM will be upgraded to 2.1.3-21
- March: castor upgrade from 2.1.4-10 to 2.1.6-10 and hardware migration for Oracle RAC.

Sw upgrades 10.2.0.4

- CERN june: Downstream databases upgraded to 10.2.0.4 Mail to the 3d list – parameter related to Streams to be included during Tier1s interventions in order to upgrade to 10.2.0.4 . d3r RAC database being reinstalled with 10.2.0.4 64 bit
- RAL june: Upgrade to 10.2.0.4 in the next couple of weeks (to be scheduled)
- RAL july: RAL: 10.2.0.4 upgrade is not done. Problem with clusterware. Dbs are online but not upgraded. Report will be sent. Try again in few weeks. Adding 3rd node for ATLAS RAC planned to the end of the month.
- Aug RAL: upgrade to 10.2.0.4 + July CPU patch applied
- July Triumpf: 10.2.0.4 upgrade done
- July PIC: Upgrade to 10.2.0.4 done
- July Taiwan: upgraded 3D Oracle RAC and the agent to 10.2.0.4
- Sep PIC: Databases upgraded to 10.2.0.4 and July cpu patch applied.
- Sep BNL : Upgrade to 10.2.0.4 already scheduled for 23rd June NDGF (Olli): Upgrade to 10.2.0.4 in the next 2 weeks (before Olli goes on holidays) Gridka (Andrea): Upgrade to 10.2.0.4 scheduled for July 19th.
- NDGF (Olli): Database upgrade to 10.2.0.4 on 2nd July. Parameter aq_tm_processes parameter setup during upgrade using dbua assistant
 - **This parameter must not be set explicitly. Please check and remove it from the spfile/pfile**
 - Review metalink Note:428441.1 <https://twiki.cern.ch/twiki/bin/view/PSSGroup/StreamsRecommendations>
- July: **Please remember to setup the parameter "_buffered_publisher_flow_control_threshold" and re-enable the propagation job after your interventions** <https://twiki.cern.ch/twiki/bin/view/PSSGroup/StreamsRecommendations> or metalink note 551516.1
- Sep CNAF:Grid cluster upgraded to 10.2.0.4. Internal problem with backups on ATLAS and LHCb clusters. Upgrade to 10.2.0.4 postponed till October
- Aug gGridka: some problems after upgrading ATLAS and LHCb databases to 10.2.0.4 – instances being rebooted. Cluster interconnect has been changed (Oracle support does not support direct cluster interconnect). CPU patch to be scheduled.
- Sep BNL:Plan to upgrade the agents to 10.2.0.4. Databases on version 10.2.0.4 require agents on version 10.2.0.4.

OS upgrade

- Aug Triumph question: Upgrade to RedHat 5 not planned at any Tier1 site neither at CERN for the moment..
- Actually at CANF we have...

Parameter settings

- July BNL: Prepare space and account for TAGS tests. Reconfigure memory parameters current SGA 4G, db_keep_cache_size to 1200M node 2 (this parameter is not needed anymore – Florbela). Stream pool was not modified on this node.
- Oct RAL: set _buffered_publisher_flow_control_threshold parameter (rolling intervention)
 - memory tuning problems, identified, buffer cache adjusted
- Oct RAL: 2 interventions: first to set up the memory parameter related to Streams, second to increase the sessions number (150 to 300).
- [FTS](#) and [LFC](#) database for ATLAS, Oracle recommended to install one patch – Carmine will sent

Bug/problems

- July BNL: Agents state in OEM? State changes to unknown without any reason. SR opened by Carlos
- Sep TRIUMF: 'Out of Sessions' errors during FDR2 testing. The problems is that they are submitting too many sessions overloading the database and impacted the Streams performance
- Sep NGDF: During ATLAS stress tests, single instance database was overloaded (I/O limitation) causing replication rate being decreased. Olli tried to implement consumer groups causing Streams replication getting stuck. Now the number of ATLAS_COOL_READER sessions is limited to 16 active sessions. New cluster being approved
- Sep PIC: Problem with raw devices at the ATLAS cluster, asm disk group metadata was not updated correctly. Fixed with help of Oracle support. Eric proposed to ask support for the cause/bug of this problem to communicate to other sites.
- Sep GridKa: During ATLAS high load, several problems observed related to COOL queries. Fixed by increasing the undo retention.

Bug/problems 2

- Sep ATLAS T1s, To: Tier1 sites problems during FDR ATLAS stress tests: Tests are run without notifying: nor ATLAS dbas, nor Tier1 sites dbas, nor Eva.
 - Tier1 databases are overloaded with ATLAS_COOL_READER sessions and Streams performance is impacted.
 - From the first checks, systems are I/O limited.
 - Requirements to Tier1 sites were only on volume. Stress tests are showing other limitations, new requirements to be added?
- Sep : Must check the OEM configuration for their targets. Several agents are inaccessible and database are not configured. OEM monitoring is not useful in this way. She has already sent an email.
- Oct Triumf : FTR2 tests - tests moved to a node running apply - caused bottleneck on streams, snapshot too old problems spotted as well
- Oct SARA : * reboots, agent problems
- Oct General: 3d OEM agents/targets still not well configured at IN2P3² (all databases), SARA and CNAF (ATLAS database). observed different undo retention on different sites - recommended at least 1 hour, 2 hours preferred
- Oct BNL: one of the nodes crashed last week, asm instance problem, impossible to re-start it, node was rebooted to fix the problem. Identified as a bug by Oracle, but there is not patch for BNL architecture yet (assigned to development). No news from SR on apply problem (apply process getting stuck).