

WLCG DB Service Reviews

Maria.Girone, IT-DM

~ ~ ~

Oracle Operational Review Meeting, 6th
April 2009

- Motivation for such reviews
- Some concrete examples with technical details & service impact
- Status of the discussions with Oracle
- Outlook
- Summary

- As highlighted by the experiments' list of Critical Services and as confirmed by regular WLCG service reports, databases play a key role
 - Used for online and offline applications, at Tier0 as well as Tier1s, for experiment applications, Grid middleware services, monitoring, data management, ...
- Services designed and deployed with resilience in view but we need to be mindful of what can go wrong and the possible service impact
- **Cannot expect Oracle support engineers looking at a specific problem report to understand the overall WLCG scope or impact**
- Need to be both pro-active as well as re-active



- Atlas, Streams and LogMiner crash
- CMS Frontier, change notification and Streams incompatibility
- Castor, BigID issue
- Castor, ORA-600 crash
- Castor, crosstalk and wrong SQL executed
- Oracle clients on SLC5, connectivity problem
- Restore performance for VLDB



- Atlas streams replication aborted with a **critical error** on 12-12-2008
 - **Full re-instantiation** of streams was necessary
 - PVSS replication was **down** for **4 days**
 - Conditions replication was **down** for **3 days**
 - Replication to Tier1s was severely affected, part of the replication was down during the CERN annual shutdown period
 - Open **SR 7239707.994** with Oracle

- **Streams issue:**
 - Generated by **LogMiner** aborting and not being able skip 'wrong redo' and/or bypass single transactions
 - Issue could not be reproduced exactly, **but likely to come up any time LogMiner will run in a critical error** while reading a redolog/archivelog file
 - In a test environment we could see similar behaviour by manually corrupting redo files
- **Workaround proposed by Oracle support on 10g:**
 - skip the log file which contains the transaction causing the problem
 - **Not acceptable in production environment**
 - Will cause logical corruption, i.e. **data loss at the target database**

- Streams capture process aborts with errors ORA-16146 and ORA-07445
 - ORA-16146 reported intermittently and capture aborts (SR 20080444.6) and
 - ORA-07445: [kghufree()+485] repeatedly in a Streams environment (**SR 7280176.993**)
 - **Streams and change notification have incompatibilities**
 - The issue also causes high load on the server (**performance issue**)
 - Forces **manual capture restart**
 - Oracle support recommendations and actions
 - Proposed patch 5868257, did not solve the issue
 - Oracle support will try to reproduce in house
 - Alternative solution for CMS: replace the use of change notification with systems info of last modified tables.

- Wrong bind value being inserted into castor tables
 - i.e. ‘Big ID’ problem in CASTOR @RAL, ASGC & CNAF
 - OCCI NUMBER TYPE OVERFLOW
 - Oracle support [SR 7299845.994](#)
 - OCCI 10.2.0.3 application build with GCC 3.4.6 on Linux x86, Oracle 10.2.0.4 database
 - Randomly the application inserts huge values (10^{17} instead of 10, for example)
 - **Difficult to trace OCI or OCCI in multi-threaded environment**
 - RAL working on a test case to reproduce the problem for Oracle support

- SQL executed in wrong schema when fully qualified names are not used
 - Seen in 10.2.0.2 (LFC)
 - Similar bugs are reported on 10.2.0.3
 - Supposed fixed in 10.2.0.4
 - Although Castor @ RAL has reported one occurrence of ‘crosstalk’ which caused loss of 14K files
 - [SR 7103432.994](#) (September 2008)
 - Issue not reproduced since
 - Problem seen also on VOMS (average of once/day). To follow-up

Invalid Number on Bind Variable for Castor

- Invalid number found when updating with bind variable value of type OCCIBDOUBLE
 - Oracle support [SR 7420155.994](#)
 - Similar to the BigId issue ([SR 7299845.994](#))
 - The effect is a wrong bind variable being detected by Oracle. This generates 'data corruption'.
 - Sporadic issue, difficult to reproduce (although when present is consistently seen)
 - [Work in progress](#) with Oracle analyst to understand the source

- Oracle crash due to ORA-600 on 14-3-2009
 - Castor stager for Atlas down for about 10 hours
 - Because of bug 6647480
 - Instance crashed and crash recovery failed
 - **SR 7398656.993** opened with Oracle support
 - Issue resolved with patch
 - CERN also requested merge for 7046187 and 6647480
 - Note: bug 6647480 was available since October 2008, but not listed as critical in Oracle main support doc “10.2.0.4 Patch Set - Availability and Known Issues”
 - **See**
<https://twiki.cern.ch/twiki/bin/view/FIOgroup/PostMortem20090314>

- **Clients cannot connect**
 - It's a **blocking issue** for Oracle clients on SLC5 (and RHEL5)
 - Based on Bug 6140224, "SQLPLUS FAILS TO LOAD LIBNNZ11.SO WITH SELINUX ENABLED ON EL5/RHEL5"
 - The root cause is explained in **Metalink Note 454196.1**
 - The first proposed workaround by support (use of SELinux in 'permissive' mode) is not acceptable in our environment
 - A second workaround proposed from support, involves repackaging of the Oracle client as rpms installed on localdisks (instead of AFS)
 - requires significant changes in CERN and Tier 1 deployment model
 - From oracle support, the issue should be fixed in 11gR2 Beta 2
 - We requested a backport to 10.2

- Backup **restore performance issue**
 - Affects significantly time to restore for very large databases (VLDB)
 - Restores are limited by gigabit Ethernet speed (100 MBPS)
 - The issue effectively disallows to use multiple RAC nodes/NICS to restore a VLDB
 - Example: 10TB DB, the use of a single node/channel limits **restore time to 1 day or more**
- Status:
 - Opened as **SR 7241741.994**
 - Bug 8269674 opened by Oracle support for this issue

Issue	Services Involved	Service Request	Impact
Logminer	ATLAS Streams	SR 7239707.994	PVSS, conditions service interruptions (2/3 days); Replication to T1s severely affected
Logminer	Streams		Proposed 'workaround' not acceptable -> data loss!
Streams/ Change notification	CMS Frontier	SR 7280176.993	Incompatibility with Streams
BigID	CASTOR	SR 7299845.994	Service interruption
"Crosstalk"	CASTOR + ??	SR 7103432.994	Logical data corruption
ORA-600	CASTOR	SR 7398656.993	Service interruption
	Clients on SL5		Clients can't connect!

- After some initial contact and e-mail discussions, a first preparatory meeting is scheduled for April 6th at Oracle's offices in Geneva
- There is already a draft agenda for this meeting
 - Introductions & Welcome
 - WLCG Service & Operations – Issues and Concerns
 - WLCG Distributed Database Operations: Summary of Current & Recent Operational Problems
 - Oracle Support – Reporting and Escalating Problems for Distributed Applications (Customers)
 - Discussion on Frequency, Attendance, Location, ...
 - Wrap-up & Actions
 - <http://indico.cern.ch/conferenceDisplay.py?confId=53482>
- A report on this meeting will be given at the Distributed DB Operations meeting to be held in PIC later in April
 - Agenda:
<http://indico.cern.ch/conferenceOtherViews.py?view=standard&confId=54037>

- Depending on the outcome of the discussions with Oracle, we would expect to hold such “WLCG Oracle Operational” (W) reviews roughly quarterly
- RAL have already offered to host such a meeting
- The proposed initial scope is to:
 - Follow-up on outstanding service problems;
 - Be made aware of critical bugs in production releases, together with associated patches and / or work-arounds.

- Following proposals to the WLCG MB and elsewhere WLCG Oracle Operational review meetings are being established
- Such meetings will initially take place quarterly
- Input from regular WLCG Distributed Database Operations (ex-3D) con-calls and workshops will be essential in preparing and prioritizing lists of issues, as well as to exchange knowledge on workarounds and solutions

Backup Slides

- CERN: 6 instances (4x LHC, Public, cernT3)
 - Each instance: stager + DLF + SRM DB schemas, each on a different cluster (2 nodes)
 - Name server on yet another cluster
 - 38 machines!
- RAL: 4 different stagers + DLF + SRM, one Name server, two clusters
- CNAF and ASGC: one castor instance (Stager + DLF + SRM + Name server), one cluster