

Oracle Access at SARA during ATLAS Reprocessing Campaign

Alexandre Vaniachine
WLCG Tier-1 Service Coordination Meeting
March 25, 2010
CERN, Geneva, Switzerland

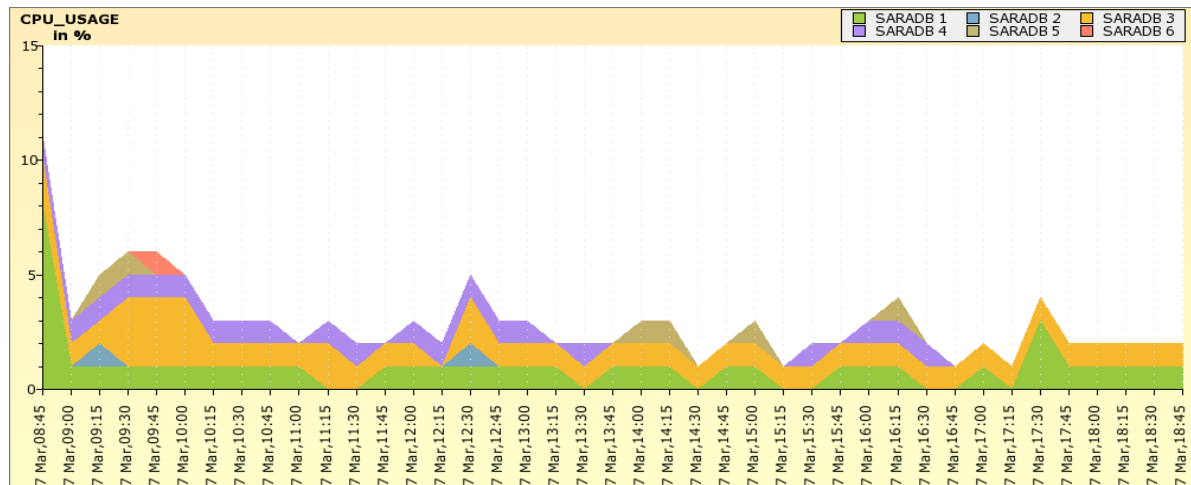
Oracle Access in ATLAS Reprocessing

- The recent ATLAS reprocessing campaign was delayed due to problems that exposed unexpected weak points in Oracle database access
- As a last step of the reprocessing campaign, TAG production tasks were submitted on March 16 to all Tier-1 sites
 - Last TAG production jobs started at SARA on March 17
 - Four of these jobs stuck for about a day and required manual intervention to finish the reprocessing campaign
- The cause of the incident was Oracle access used by TAG production jobs
 - Each successful TAG recovery job made 12 queries in two Oracle sessions to read 232 bytes of data in less than a second, e.g.:
 - INFO Connection COOLONL_GLOBAL/COMP200 : nConnect: 2 nFolders: 1 ReadTime: 0.34 s
- The last four TAG recovery jobs at SARA were not able to connect to local Oracle due to repeated errors:
 - Error ORA-12170: TNS:Connect timeout occurred (attaching a server)
- Seven finished jobs from the same TAG recovery task 121216 (these were able to connect to Oracle at SARA) started
 - from Wed Mar 17 12:18:55 2010 to Wed Mar 17 12:20:09 2010
- All four stuck jobs started at about the same time:
 - from Wed Mar 17 12:19:50 2010 to Wed Mar 17 12:20:27 2010



Oracle Load at SARA at the Time of the Incident

- Each failed job made five CORAL connection attempts to Oracle at SARA waiting 300 seconds between connection retries
- Each connection attempt failed with
 - Error ORA-12170: TNS:Connect timeout occurred (attaching a server)
- As expected, there was no significant load on Oracle at SARA at that time
- After five connection attempts each of four jobs connected to the failover Oracle at FZK:



```
CORAL/Services/ConnectionService Info Connection to service "FZK" established.
Id=27101586-31ca-11df-a920-0026b93b5044
```

```
CORAL/Services/ConnectionService Info New session on connection to service "FZK"
started for user "ATLAS_COOL_READER_U". Connection Id=27101586-31ca-11df-a920-
0026b93b5044
```

```
RalSessionMgr Info Start a read-only transaction active for the duration of the
database connection
```

```
RelationalDatabase Info Instantiate a R/O RalDatabase for 'COOLONL_GLOBAL/COMP200'
```

Follow-up

- Three jobs out of four crashed about two hours later, with no more records in the log file

```
Py:pmon_dso          INFO [2414] /dso/libnuma.so      347.211 Mb
```

```
-----  
Core dump from CoreDumpSvc on am94-15.gina.sara.nl at Wed Mar 17  
14:36:55 2010
```

```
Caught signal 7(Bus error). Details:
```

- One job out of four crashed at about the same time, but with more records in the log file that show that both Oracle sessions were successful before the crash
- Following bus error crashes, all jobs remain stuck for about a day
 - which caused a delay in the reprocessing campaign
 - ATLAS A-team is notified of this problem
- After thorough investigations by the SARA and ATLAS teams it was found that the firewalls on each of the seven Oracle nodes were dropping traffic from the compute nodes
 - The problem is already fixed
 - A signature of the firewall problem is that within the one hour limit for retrials there was only five connection retries (instead of the eleven retries expected)

Conclusions

- Failover to Oracle servers (using CORAL 2.3.5) resulted in job crashes that delayed reprocessing campaign by one day
- Fortunately, reprocessing at one Tier-1 should not be using failover to Oracle at another Tier-1
- SARA Tier-1 and ATLAS database production support teams took measures to prevent the reoccurrence of that problem in the next reprocessing campaign

