

Post mortem LFC indicent 23-26 May 2009

De WikiPIC, la enciclopedia libre.

Table of contents

- 1 Executive summary
- 2 Follow-up actions
 - 2.1 Final Solution
- 3 Problem detection and reporting
- 4 Problem timeline/follow-up
- 5 Technical Analysis

Executive summary

Starting on Saturday the 23rd. of May about 8:00 UTC the LFC daemon at PIC suffered instabilities. The LFC service crashed when trying to delete some specific LFC entries triggered by central deletion machinery used for ATLAS. Each request to LFC involving this entries (corrupted, wrong ACLs) killed the LFC daemon and produced SAM LFC test to fail as well. The problem was not completely solved until Tuesday 26 of May at 14:00 UTC after deleting manually the corrupted entries in the DDBB.

Follow-up actions

- First intention was to ban the user issuing over the corrupted LFC entries, but the LFC service has no banning option. ATLAS deletion requests and data transfers are triggered centrally and always use the same DN. The workaround was to filter the IP's coming from that DN within ATLAS (iptables).
- Situation reported to LFC developers (corrupted entries should not led to a service crash)
 - Savannah bug report opened: <http://savannah.cern.ch/bugs/?50818>
- Ask LFC experts to procure a manual deletion recipe for this use case:

https://gus.fzk.de/ws/ticket_info.php?ticket=48980 https://gus.fzk.de/ws/ticket_info.php?ticket=49035

Final Solution

- Remove corrupted entries from the ORACLE DDBB for the ATLAS LFC service.
- These entries were related to CCRC08_M5 data (fake) that had no replica on disk, about 23k entries were deleted from the LFC.
 - Directories involved:

```
'ccrc08_m5.0028940.Default.L1TT-b00000110.RAW.ol
'ccrc08_m5.0028940.Default.L1TT-b00000110.daq.RAW.ol
'ccrc08_m5.0028997.Default.L1TT-b00000110.daq.RAW.ol
'ccrc08_m5.0029499.physics_cosmics.daq.RAW.ol
'ccrc08_m5.0029517.Default.L1TT-b00000010.daq.RAW.ol
'ccrc08_m5.0029548.Default.L1TT-b00000010.daq.RAW.ol
'ccrc08_m5.0029563.Default.L1TT-b00000100.daq.RAW.ol
'ccrc08_m5.0029565.Default.L1TT-b00000100.daq.RAW.ol
'ccrc08_m5.0029568.Default.L1TT-b00000100.daq.RAW.ol
'ccrc08_m5.0029576.Default.L1TT-b00000100.daq.RAW.ol
'ccrc08_m5.0029576.physics_cosmics.daq.RAW.ol
'ccrc08_m5.0028940.Default.recon.ESD.ol_r436_tid022728
'ccrc08_m5.0028997.Default.L1TT-b00000110.recon.ESD.ol_r436_tid022488
'ccrc08_m5.0028997.Default.L1TT-b00000110.recon.ESD.ol_r436_tid022667
'ccrc08_m5.0028940.Default.recon.log.ol_r436_tid022728
'ccrc08_m5.0028997.Default.L1TT-b00000110.recon.log.ol_r436_tid022488
'ccrc08_m5.0028997.Default.L1TT-b00000110.recon.log.ol_r436_tid022667
```

- ■ Deletion command at the DDBB:

```
'SQL> delete from CNS_FILE_REPLICA where SFN like '%ccrc08_m5%';
```

- ■ 23304 rows deleted from the DDBB.

Problem detection and reporting

- On the 23rd of May around 11:30 UTC, PIC Manager on Duty noticed that ATLAS LFC was down and this triggered SAM failures. Service manual restarting of the lfc instance didn't help, crashing almost immediately after the restart. The problems starts to appear on both instances of the ATLAS LFC service (we have two instances for the same LFC service load balanced):

```
'[root@lfcAtlas02 run]# /etc/init.d/lfcdaemon status
'lfcdaemon dead but pid file exists [FAILED]
```

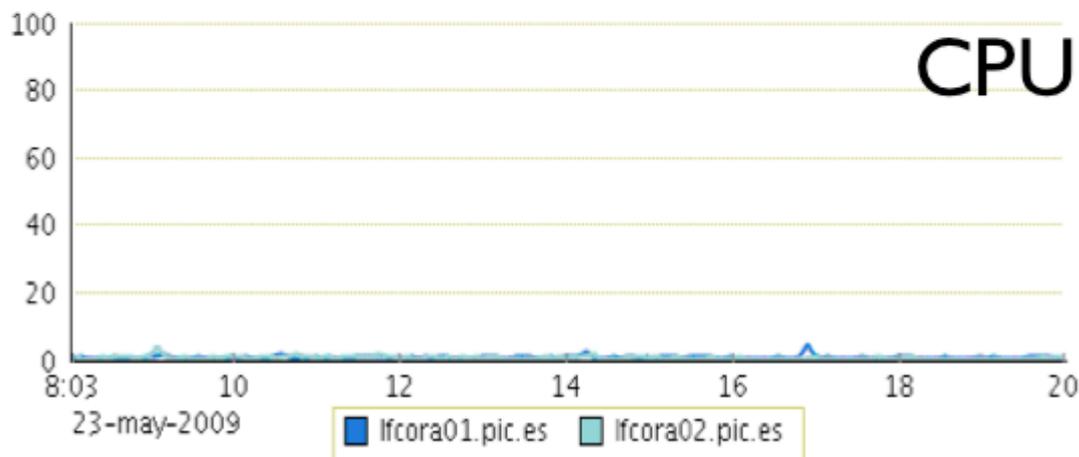
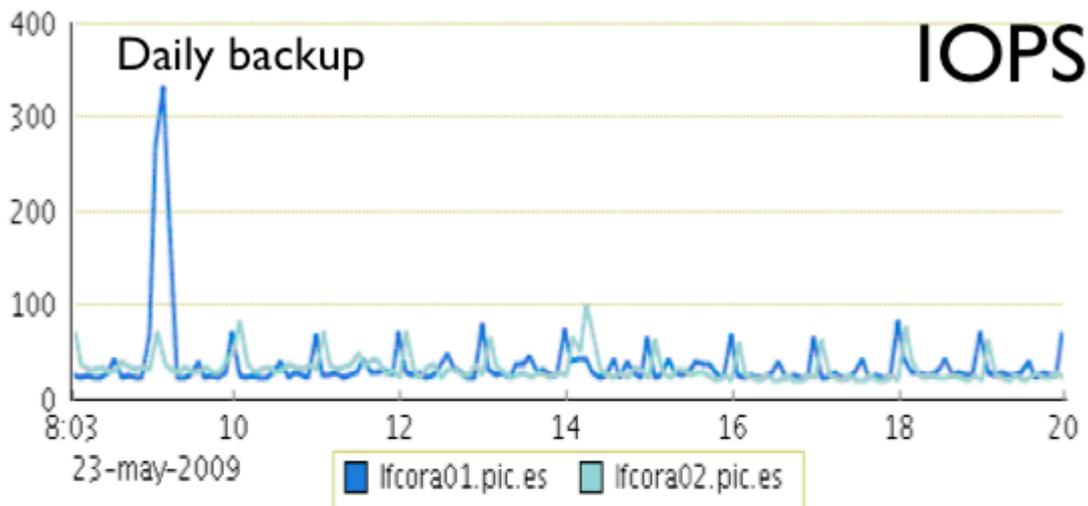
- ■ After trying it several times, he decided to reboot the machine. No success. Deletion/transfers request were continuously issuing the LFC. PIC MoD send an email to the person using the DN for deletion and the ATLAS contact at PIC. And in parallel filtered the IP's causing the problem (banning workaround)
- On the 24th of May:
 - ATLAS expert on call stopped deletion and transfer machinery waiting for the service to be recovered.
- On the 25th of May:
 - ATLAS expert on call identified the path holding the wrong ACLs (all of them issued by the deletion machinery), so the file transfer service was restarted as this activity was not trying to access the LFC corrupted entries.
- On the 26th of May:
 - The ATLAS contact at PIC (on vacation during the day before) together with the DBA expert at PIC removed manually the corrupted entries. After it, deletion machinery was started successfully and the deletion activity was happily progressing.

Problem timeline/follow-up

1. 23 May, 11:00 UTC: PIC manager on duty opened an internal PIC ticket after spotting the first failures
2. 23 May, 14:00 UTC: ATLAS ADCoS shifter opened a GGUS ticket , that was routed to LCG support at PIC.
3. 23 may, 15:00 UTC: ATLAS Expert on call jumped in the GGUS ticket
4. 25 May, 10:30 UTC: LFC admin at PIC sent two emails to LCG-rollout asking for support for banning a user (23 May, 15:00 UTC) and how to perform a cleanup in the LFC directly into the DDBB .
5. 25 may, 11:00 UTC: LFC admin opened a Savannah bug report to LFC developers , entitled: operation with entries with wrong/inexistent uid crashed lfcd daemon. Debugging is ongoing, we offer to provide a complete DDBB copy and the list of affected paths for a proper post-mortem (ONGOING)
6. 26 may, 08:30 UTC: LFC admin at PIC open a new GGUS: https://gus.fzk.de/ws/ticket_info.php?ticket=49035
7. 26 May, 13:00 UTC: The corrupted LFC entries were deleted manually form the DDBB and ATLAS activities were working smoothly.

Technical Analysis

One f the first action was to perform forensics at DB level, to discard potential DDBB overloads or high activity peaks, it was clearly seen that there were no problems due to the LFC ORACLE DDBB backend. This can be seen in DDBB activity plots during the 23rd of May.



So the problem was related only with the corrupted entries in the LFC with wrong ACLs:

```
lfc-getacl /grid/atlas/dq2/ccrc08_m5/log/ccrc08_m5.0028940.Default.recon.log.o1_r436_tid022728 /log.0
# file: /grid/atlas/dq2/ccrc08_m5/log/ccrc08_m5.0028940.Default.recon.log.o1_r436_tid022728/log.02272
No such uid
# owner: 31801
user::rw-
```

The entries couldn't be deleted using lcg-uf (unregister file), so those had to be deleted manually in the LFC DDBB using:

```
SQL> delete from CNS_FILE_REPLICA where SFN like '%ccrc08_m5%';
```

As mentioned the problem disappeared after the manual deletion of the corrupted entries. Although there is a couple of issues opened:

- Official procedure for removing corrupted entries in the LFC DDBBs
- Correction of what we think is a bug: LFC service crashing when trying to perform operations over a corrupted entry.

Obtenido de "https://wiki.pic.es:443/index.php/Post_mortem_LFC_indicent_23-26_May_2009"

- Esta página fue modificada por última vez el 14:09 27 may, 2009.