Database Hangout 2015-04-21

Date

21 Apr 2015

Attendees

- Fabrice Jammes, Serge Monkewitz, Vaikunth Thukral, Fritz Mueller, Unknown User (danielw), Andy Salnikov, John Gates, Jacek Becla
- Brian Van Klaveren for sqlalchemy discussion

Discussion items

Large scale tests

- have qserv running, on 25 nodes @in2p3
- can easily change qserv versions
- caveat: qserv still shows cmsd bug (DM-2547)
    - made progress with debugging during the meeting
        - issues with eups: .so is not copied to run-dir
        - there are more issues with cmsd, Daniel looking into it
- Serge is in the process of loading data to shared file system now
    - transferred Object dump
    - doing ForcedSource
    - flux cutout for Source? Per official LSST key numbers in DR1: 18 billion Objects, 350 billion Sources. So try 10% of that
    - will start duplicating Object table now (copy all onto each node, and duplicate)
- would be good to have --keepdata option (DM-1429), can't afford to destroy the 100+ TB data set
    - quick and dirty solution for now: put data outside of run-dir, symlink it, "qserv-configure --all" will not destroy data, it will only destroy symlink
        - Fabrice will implement that

Parallel ssh vs off-the-shelf

- Serge discovered saltstack.org, ansible.com, both look good. Will try
    - saltstack is daemon-based
    - ansible supports kerberos (required by in2p3)
    - both can run adhoc shell commands
- But Fabrice just finished implementing parallel ssh (DM-2341)
    - Put deploying our custom parallel ssh on hold, decide in 1-2 weeks
    - Lesson learned: don't jump to quickly into implementing new feature, do more research and discuss more first

Mysql-proxy

- writing 1 pager for mysql team, made progress at the meeting
- would be nice to get their protocol plugin nicely packaged, and add it to czar and get rid of mysql-proxy and python layer

SqlAlchemy

- see DM-2299
- consensus: don't want to go to much towards ORM: it is too easy to write **very** inefficient code

Improving communication within group

- do post status in hipchat, don't be shy if you are reporting you are late
- be on-call during meeting times (e.g., people working on non-qserv and not attending qserv meeting should be on call in case we want to pull them in on very short notice, same with people working on data access not attending data access meetings)
- also, be on-call during ~11-3pm so that we can quickly summon ad-hoc meetings as needed (as today)

Daniel

- Daniel's last day is end of April