

Database Meeting 2016-06-15

Date

15 Jun 2016

Attendees

Local: [Brian Van Klaveren](#), [John Gates](#), [Unknown User \(nplease\)](#), [Vaikunth Thukral](#)

Remote: [Andy Salnikov](#), [Fabrice Jammes](#), [Unknown User \(oachbal\)](#)

Discussion items

Swarm

- [Unknown User \(oachbal\)](#) presenting Swarm demo today with [Fabrice Jammes](#) on Shared Screen
- Tested on Qserv OpenStack-cluster@NCSA
- Good replacement for shmux@IN2P3, working on documentation to propose doing that on Qserv-cluster@IN2P3
- Link to talk: <https://drive.google.com/file/d/0B-VJpOQeezDjbMxdEthWmhTkE/view?usp=sharing>
- Cleaning up Swarm scripts - ETA next week
- DC2013 data-loading attempted on Swarm OpenStack test cluster - queries launched on it

Documentation

- Documentation on improving container infrastructure underway
- DM-5039: Jonathan Sick is back so Brian needs to talk to him. Don't want to add too much to LDM-463 but refer/link to it for this ticket
- Discussion - would it be better to have some doc inside the code rather than update it in LDM-463 for frequent changes (like API parameters etc.)

DAX

- Brian started work on building docker container for DAX
- Deploying Qserv@newOpenStack@SLAC - issues with SLAC admin stuff

Butler

- DM-6459 ticket under epic DM-6224: RFC done, work from that RFC under review (KT)
- Work continues on DM-5548 - will likely need RFC too, have a draft but need a discussion session first with KT when he's around

Shared Scans

- Mem-man mlock function is now locking memory on the cluster with Fabrice's changes - all queries in scan-scheduler running ~twice as fast
- mlock call takes longer than expected time (~5sec) - made new ticket for this (DM-6518), under review
- Perhaps investigate further issues in a new ticket

XSWAP

- Elasticsearch JSON-output parsing finished using R. Organizing output in CSV files to work with for analysis for both us and NERSC
- Planning on optimizing (and maybe automating?) this parsing for as long as ES doesn't provide native CSV output support
- Working on having basic plots rest of this month, but complete QueryID implementation missing so results not useful for decisions yet
- Thinking about potential pitfalls of repeatable vs randomized queries in test-matrix
- Ticket for cleaning up and adding QueryID at new locations is created, John planning to do on next sprint

Misc

- AndyS back from vacation today, still @CERN and continuing work on assigned tickets
- Fabrice reviewing John's tickets
- John's other tickets under review
- DM-6160 - very hard to figure out what's going on, problem seems to be within log4cxx, help from Fabrice?
- AndyS' large-results testing underway
- Mike is @ Director's Review today
- AndyH is @CERN, back on 6/20