

# Database Meeting 2016-08-10

## Date

10 Aug 2016

## Attendees

Local: [Brian Van Klaveren](#), [Unknown User \(npease\)](#), [Andy Salnikov](#), [Igor Gaponenko](#), [Vaikunth Thukral](#), [Andy Hanushevsky](#), [Fritz Mueller](#) (for second half)

Remote: [Fabrice Jammes](#)

## Discussion items

### Docker Infrastructure / Webserv

- DM-6376 - Brian still working on it, need image pushed to docker hub and need to check with Fritz about official webserv repo. Make lsst-wide repo or our own for now?
- Friday meeting with NCSA about infrastructure - no discussion of docker, mostly talked about kubernetes
- Talk to KT about a session for system-wide docker infrastructure at AHM - with Josh/Square, us/Qserv, NCSA
- DM-7103 is blocked since it depends on needing the aforementioned docker repo

### Butler

- DM-6988 and DM-7123 composite dataset requirements and design - writeup for design requirements etc is up on confluence for comments
- Some people have commented already, waiting on Gregory's comments to finish DM-6988. Other comments will dictate status of DM-7123
- Working on making policies pretty (DM-7178)

### Qserv cluster and CI

- Setting up Qserv cluster for Igor @ NCSA with OpenStack VMs
- DM-7126 needs reviewed soon since Fabrice will be on vacation second half of August
- CI build was broken - fixed now with updated documentation for review for John and AndyH (DM-7168)
- Docker updated to v1.12. - new ticket for updating docker API which breaks Swarm. Shmux still works but we are moving from it anyway and the new interface is much simpler

### L1DB Proto

- AndyS has mysql-server and postgres-server running on ccqserv124, injecting data to see performance - both are pretty bad
- Attempt changing indexing schema to look for improvements. Baseline scheme indices are not all useful, test changing those
- PK may need changing on old tables.
- Takes a long time to generate data and go through it to measure performance - may need another machine to run at the same time to test both DB's  
(Vaikunth to check with John on how to release one node from test cluster and let Andy know by Thursday)
- Try not writing directly to DB, maybe smart cache or work with memory to avoid disk I/O (current bottleneck) - monitoring also shows that cpu-wait-io is quite large on ccqserv124
- Overall looks like improvements are marginal through these optimizations, probably need SSDs or some better disk options

### Data loader and MySQL dump

- Igor working on DM-7053 mostly, assemble complete catalog of Stripe82 data using backups from in2p3 (4TB) to NCSA, 1 table left
- Dumped some data already to NCSA, next step was to set up DB at lsst-db machine but ran out of space. No solution currently so trying VM now
- After VM testing and fixes, loading up to 1.2Gbps in mariaDB of the dumped data at NCSA. The loading process seems to be capped due to mariadb constraints
- Since data was not dumped in csv but in sql format, this could be cause of slowdowns
- Need to make a ticket to resolve duplicated data issue ([Igor Gaponenko](#))

### X-SWAP

- Made new cpu-wait-on-disk-I/O plots for John. Custom hostname removals/additions adapted
- Mostly same as last week, implementing queryID into pipeline
- Re-setting up tests with uniform sampling frequency
- Vaikunth needs to talk to John about how to release 1 node from cluster setup (to give to AndyS)

#### XrootD

- Finishing migration to 4.4 for XrootD
- AndyH is reviewing Fabrice's documentation
- Consider rpm while using eups tags? Probably not
- Cleaning up orphan tickets - merged branches don't auto-close PRs when the ticket branch is not also updated in github before the merge
- Working on DM-4473, finishing soon but have deployment issues

#### Other

- AndyH on vacation next week
- All hands meeting next week
- Igor needs to talk to Greg Daues @ NCSA about calibrated images
- Fritz sharing T/CAM presentation for NSF/DOE review