Database Meeting 2016-07-20

Date

20 Jul 2016

Attendees

Fritz Mueller, Vaikunth Thukral, Unknown User (npease), Andy Salnikov, John Gates, Igor Gaponenko, Andy Hanushevsky, Kian-Tat Lim

Discussion items

Qserv

- John has added code to collect task statistics on the qserv workers; now working on code to use these statistics to move scans between scan schedulers. Initially there will just be a single heuristic which moves any slow-running scan to the snail-scan scheduler.
- John has done more testing and profiling on the cluster, looking to identify performance killer factors for scans.

Butler

- Nate finished and closed the X16 Butler and F16 Butler Repository Refactor epics. Checkin caused some minor bustage on ci-hsc, which is not currently in the jenkins autobuild. John Swinbank helped address the bustage, which had a nice side effect of removing some internal Butler dependencies at the same time. Moving forward, it was agreed that a manual trigger of ci-hsc build will be done before Butler merges.
- Nate will be digging in on composite dataset design for Butler next. Design discussion to be help with K-T, Jim Bosch, and John Parejko (look for upcoming RFD).
- Storage and format refactoring work will interleave.

PDAC

- Igor has been doing detective work tracking down the S13 Stripe82 dataset(s). These were apparently processed in parallel in two overlapping parts; part at NCSA and part at IN2P3. Partial ingest of both results was then cross-checked for QA. Igor has tracked down some data at NCSA, but we are still looking for the IN2P3 parts. Efforts in this regard continue...
- Once we locate all the data, we will have to deal with merging and de-duping as well.
- Igor notes that previously mentioned schema incompats between the former processing outputs and the current LSST schema of record should not be an issue because Qserv doesn't care.

L1/AP Database

- AndyS has AP generator script up for review. Performance running this against the current L1/AP database is awful. Andy would like to check this on a bare metal server. We will peel off a worker from the IN2P3 cluster to use for this testing.
- AndyS would like to try with postgres backend as well as SQL. Some minor portability issues need to be resolved.

XRootD

- AndyH has been working on cleaning up loose ends in documentation and fixing some bugs and perf. issues.
- XRootD release (and subsequent xrdssi branch merge) anticipated sometime next week.
- Logging integration needs a few final tweaks.

X-SWAP

- The inconsistencies in memory usage statistics and sample frequency that were formerly observed at IN2P3 have now gone away. Vaikunth still investigating with IN2P3 admins to try and determine the cause.
- Vaikunth and Fritz attended the X-SWAP all-hands at NERSC last week. Beyond hearing status reports from participating projects, the following items were of note:
    - Participating projects would like to set up some common ELK infrastructure at NERSC, and build a set of shared analysis tools on top of this.
    - X-SWAP project is up for renewal at the end of the year.