

Database Meeting 2016-06-01

Date

01 Jun 2016

Attendees

Fritz Mueller, John Gates, Brian Van Klaveren, Andy Hanushevsky, Vaikunth Thukral, Andy Salnikov, Fabrice Jammes, Unknown User (npease)

Discussion items

AWS

- Some interesting discussion was had with some of the Amazon team during the XLDB conference. There is interest from both sides in exploring use of AWS by the project. Exploratory at this point; some follow-up discussions including NCSA to occur soon.

Documentation

- Brian is working on DAX services docs; would like API docs to be split out separately from LDM-463 and referenced/linked from within so APIs can be updated without TCT involvement. Working with Jonathon to work out the details of this.
- Nate has updated LDM-463 for Butler
- Serge has updated the sphgeom README

IVOA

- Brian attended the IVOA interop conference in Cape Town. After discussions with many members of the IVOA community, has now refined the recommend approach for which IVOA standards might best be adopted by the DM team for which purposes and when. In summary (see [DM-5317](#) for detailed write-up):
 - We definitely need to support TAP (with some limitations on ADQL) as a public interface
 - SCS (simple cone search) is the most popular IVOA interface so we should support it as a public interface as well
 - TAP is not well suited for high-volume within-the-datacenter service-to-service interfacing; we should do something else for those use cases
 - We should ignore SIAP for now; existing standard is being retired and the replacement is still very immature
 - Brian did not have a chance to evaluate CAOM; we will be looking further into this moving forward
- There is a chance that we could inspect TAP query logs of some other TAP providers, as a way of assessing which APIs and entry points are the most used in the field

L1/AP Database

- Andy has published [DMTN-018](#), which reflects the database design revisit to date
- Andy and Fritz to coordinate a meeting with Simulations group to see if they can help out with the upcoming input simulator task

Butler

- [RFC-184](#) went up, and is making progress
- Nate had a meeting with Jim Bosch to work out desiderata of repository tagging feature

Shared Scans

- John is working on an enhancement/cleanup to replace the heap data structures within the ScanScheduler with a list of buckets; this will make it easier to add Tasks to actively running chunks

XRootD

- AndyH has made progress on the logging consolidation task, but needs some additional input from the team to finalize (how should messages be classified, how should we plumb through the correct thread id, etc.)
- Working with John on response pacer feature to avoid thread-lockup on the czar with large results
- Working on XRDSSI single session feature which will simplify code in the czar

Docker

- Oualid has a single Python script that provisions a cluster on the NCSA OpenStack, then deploys and runs the Qserv multi-node integration test
- Fabrice and Oualid have had initially good experiences using docker Swarm for cluster orchestration (instead of the previously scripted shmux approach)
- Docker repository cache is now up and running at IN2P3
- Fabrice has heard good things about the Rocket project (a CoreOS alternative to Docker for containers) and recommends we keep an eye on it

XSWAP

- Vaikunth now has some raw data to work with, and is looking at data transformations

- Next steps will need to make use of queryids (already in logs)
- XSWAP folks wonder if ML techniques might be used to ferret out memory contentions between query types?

Misc

- A pre-proposal was submitted to NSF in answer to a call re. HPC and machine learning. Subsequent proposal submission for this grant was discourage by NSF 😞