

Meeting Notes 2019-09-26

Attendees

Michelle Gower Jim Bosch Andy Salnikov Tim Jenness Michelle Butler Christopher Stephens Christopher Pond

Status

- [Michelle Gower](#) : Running ci_hsc_gen3 with Oracle using weekly and daily 2019-09-26 which has big ticket changes reporting any problems.
- [Christopher Stephens](#) :
- [Jim Bosch](#) : Merged  [DM-17023](#) - Refactor the Dimensions and query system  . Threw together some designs for feedback for the next step in the refactoring ( [DM-21448](#) - Clean up DatasetRef comparisons and immutability  ,  [DM-21454](#) - Remove DatabaseDict and vectorize Datastore/Butler ingest APIs ), started coding up the latter. Next up will fix the foreign key name conflict issue Michelle discovered and help get DM-18610 merged (not middleware related, other than it's what is keeping Chris Waters from working on CPP in Gen3).
- [Tim Jenness](#) : Working on tech note DMTN-133.  [DM-21384](#) - Write a tech note on OCS triggered data processing 
- [Andy Salnikov](#) : Started working on  [DM-21201](#) - Research cross-database approach to inserts with custom conflict resolution  last week, but was distracted by many other things, not much progress.

Misc items

- Upcoming maintenance at NCSA Sept 28th. [LSST Service Status page](#)
- [Fritz Mueller](#) (copy from last week waiting for Fritz to return from vacation) Are there updates on the requested Provenance presentation/demo?

Short Term Planning

1. Run ci_hsc with large set of Gen3 changes.
 - a. Running daily that has for  [DM-17023](#) - Refactor the Dimensions and query system  to merge to master
 - i. If set DAF_BUTLER_CONFIG_PATH to a location with a registry.yaml file containing Oracle connection info. Much, much better than having to use the previously required hacky shell script. ★★ (still have to manually run a clobber sql manually to clean Oracle schema)
 - ii. Ran into identifier length issues, etc. Discussions on dm-middleware about moving to newer version of sqlalchemy and applying patch Tim found for identifier length.
 1. <https://github.com/sqlalchemy/sqlalchemy/issues/4857>
 2. [Tim Jenness](#) will create eups version of new sqlalchemy (no patches):
 [DM-18601](#) - Investigate newer SQLAlchemy versions 
 3. [Michelle Gower](#) will apply patch for determining max identifier length to the oracleRegistry create_engine code until sqlalchemy version allows longer identifier lengths. (will wait to see if any other Oracle-specific changes are needed to run ci_hsc_gen3 before pushing/merging).
 - iii. Foreign key constraint errors  [DM-21454](#) - Foreign key error when running makeButlerRepo.py against Oracle 
2. Run RC2 patch through as much of DRP as applicable
 - a. needs RC2 bootstrap - approximately 2 or 3 weeks out
3. Run RC2
 - a. DB folks will now have larger set of data flow information for analysis (along with updated schemas)

End-of-October DMLT meeting

- Middleware development status 30 minutes
 - How close are we to "One Butler for Christmas"?
 - [Fritz Mueller](#) said he has a pretty good idea of how to track this through the month and report
- Middleware staffing transition - 30 minutes
 - Who will be working on middleware during 2020? How can we free up folks — in particular Jim Bosch — to focus on other tasks? Fritz Mueller has agreed to come up with a plan which he will discuss at this meeting.
 - (This is downstream of a meeting with Wil and John and Michelle Butler; basically, how and when can we transition Jim back to DRP? How do we distribute the torch? If NCSA is going to be taking more lead as we work through production and workflow issues, what are deliverables and acceptance criteria for a cross-fade?)
 - Decided to wait to have discussions when Fritz is available

Technical Discussions

- See Short Term Planning for discussion about running with Oracle
- DMTN-133 coming with discussions about how to trigger pipelines in chained processing.
- What processing needs to happen at summit vs what processing can happen at Base + OODS vs what processing happens in Prompt Processing (which includes AP, but wasn't limited to AP)
- Discussion about needing to get data out of Butler at summit into other Butlers (OODS? User Schemas near OODS? DBB? User Schemas near DBB?)
- Small discussion about the single-node activator restarts and that some current Butler work will help with some of the issues with restarts.
 - Continue running - (new smaller QuantumGraph)
 - Skip existing - (same QuantumGraph) skips running Quantum if output exists