

Meeting Notes 2019-09-19

Attendees:

Michelle Gower Jim Bosch Tim Jenness Andy Salnikov Unknown User (cs2018) Unknown User (cpond) Unknown User (mbutler)

Status

- [Michelle Gower](#) Other than misc conversations and planning, nothing Gen3 to report.
- [Unknown User \(cs2018\)](#)
- [Jim Bosch](#): DM-17023 is almost done. Will probably land early next week, after addressing review comments and rebasing one more time. After that, I need to write an SST technote. And after that, I'll dive into the next set of Registry changes unless we want to do any restructuring or renaming first.
- [Tim Jenness](#)
- [Andy Salnikov](#)
- [Fritz Mueller](#)

Misc items

- Upcoming maintenance at NCSA Sept 28th. [LSST Service Status page](#)
- [Fritz Mueller](#) Are there updates on the requested Provenance presentation/demo?

Short Term Planning

1. Run ci_hsc with large set of Gen3 changes.

a. Waiting for



DM-17023 - Jira project doesn't exist or you don't have permission to view it.

to merge to master

2. Run RC2 patch through as much of DRP as applicable
 - a. needs RC2 bootstrap
3. Run RC2
 - a. DB folks will now have larger set of data flow information for analysis (along with updated schemas)

Planning

- When OODS includes Gen3 (Tucson and Mountain)
 - Raw Ingest (AuxTel and ComCam)
 - 2019-09-19 obs_lsst package ready, ingest API ready (see ci_hsc_gen3 for example code). Request to not have to specify instrument class with every ingest. Discussed saving class in registry when registering instrument.
 - Purge (to keep size of registry small)
 - 2019-09-19 Butler delete functionality ready (but not duration related). Perhaps could code a Datastore with expiration feature instead of an external driver.
 - Multi-registry
 - Staff schemas (db admin procedures, etc)
 - Unique Site IDs?
 - 2019-09-19 could be implemented in Jim's next big ticket.
 - Get calibs from any production runs
 - 2019-09-19 Have export and import when current big ticket merges to master. Not clear if import is enough or will need a merge.
- When NCSA DBB endpoint includes Gen3
 - Raw Ingest (AuxTel and ComCam)
 - Ingestion of certain LFA files
 - 2019-09-19 Wavefronts should be treated like raw, calibration spectrograph images and CBP images should be new instruments. **Need example files to determine whether any work needed.**
 - Multi-registry
 - Staff schemas (db admin procedures, etc)
 - Unique Site IDs
- EFD ETL (Quick Look & DBB copy)
 - No additional Gen3 work because joins would happen on EXPID?
 - 2019-09-13
 - **Need to figure out how this is going to work**
 - Will Butler.get read header information from the EFD tables instead of from disk?
 - (If yes, then a job scratch repo/subset will need to also copy these EFD entries)
 - A third source of exposure information? CAOM-compliant exposure table
 - Will EFD ETL values be used in QuantumGraph generation?

- Gen3 "production" pipeline runs
 - HSC RC2 DRP running in Gen3 at some cadence (like we currently do with Gen2)
 - AuxTel pipelines (calibration? drp?)
 - BPS
 - HTCondor version of BPS
 - BPS job activator
- Next OPS rehearsal (Feb-ish)
 - Not clear that it will use Gen3, but can assume that is the hope.
 - Could include everything from image generation through accessing it via LSP (multi-registry)

Technical Discussions

- Discussion of Tim's review of ticket



[DM-17023](#) - Jira project doesn't exist or you don't have permission to view it.

- Repo import/export work done to make ci_hsc_gen3 less fragile.
- Export is pretty much the subset functionality
- Import does not replace merge. Data is inserted into repo as is (with no id modification).
- Support re-run of pipetask on the same output collection (



[DM-20845](#) - Jira project doesn't exist or you don't have permission to view it.

- Discussion of how using collection in filenames and using collections to look up run. Run ids are not human friendly. And collections can be more than 1 run. Instead switch to using a run name in the filename and to identify runs. [Jim Boschto](#) write up in ticket.

- obs_lsst gen3 support now merged (



[DM-16297](#) - Jira project doesn't exist or you don't have permission to view it.

- Need to sort out some things in data model:
 - discussion started because needing to use OBSID in the OCS Batch Service
 - e.g. OBSID vs EXPID, sequence number, etc.
 - [Tim Jenness](#) to add as needed.
- Provenance
 - Near-term work Butler could start adding quantum records at execution (i.e., saving provenance tree).
 - Longer-term understanding what pieces are responsible for saving information that can be duplicated information (e.g., environment when running a quantum)