



Arch S22B Status & Plans

Kian-Tat Lim and Tim Jenness



U.S. DEPARTMENT OF
ENERGY

SLAC

CHARLES AND LISA SIMONYI FUND
••• FOR ARTS AND SCIENCES •••



Status for S22B

Key:

Added since last vF2F

Completed, In Progress, Not Started

Requirements and Design

- Consolidated Database design document
- Produce drafts for DMTN-204 (Dataset annotations, punted to Campaign Management) and DMTN-205 (provenance, punted to Jim Bosch)
- Describe compute environments for rapid processing
- DP0.2
 - IVOA integration design
 - ObsCore Felis
 - HiPS design
- Define Large File Annex integration with the Butler

Oversight and Release

- Build engineering
 - `lsstinstall` release and incorporation into `lsstsw`
 - ~~`rubin-env: Python 3.10`~~ and ~~`rubin-env-extra`~~, `conda-forge cfitsio/compiler updates`
 - Keep working with MatthiasW (`changelog` and `stack containers` and `reusable workflows`)
- `Release 23.0 support (and rubin-env 0.8.*)`

Investigation and Coding — Middleware

- Release documentation on how to write pipeline tasks
- Possible schema migration with new abilities to support summit observing
- Provenance
- Dataset annotations (if part of butler or butler adjacent — DMTN-204)
- Resolve handling of metrics (is butler involved once file is written?)
- Graph-backed registry for batch production
- Support Rucio work
- Support prompt processing
- More flexible “runner” for PanDA
- pipetask Python API

Investigation and Coding – Other

- Productionize (but not optimize) Prompt Processing
 - Complete middleware interface
 - Build Kafka and MinIO into prototype
 - Move to SLAC K8s if available
- Consolidated Database prototype
 - Per-visit EFD summary
- OCPS
 - Update to salobj 7
- DMTN-143
 - Raw image upload service

Pre-Ops

- Integration and standing up USDF
- Campaign Management
- Security
- Hybrid model
- Rucio/Butler/PanDA integration

Plans for F22A

se design document (define data model)

Oversight and Release

- Build engineering
 - lsstinstall release and incorporation into lsstsw
 - Jenkins move to USDF
 - MatthiasW (PyPI, legacy packages, stack containers, Jenkins move)

Investigation and Coding — Middleware

- Provenance
- Dataset annotations (if part of butler or butler adjacent — DMTN-204)
- Resolve handling of metrics (Datastore backends?)
- Graph-backed registry for batch production
- Support Rucio work
- Support prompt processing
- Support Campaign Management
- pipetask Python API
- Butler client/server

Investigation and Coding — Other

- Productionize (but not optimize) Prompt Processing
 - Build next_visit Kafka into prototype
 - Transfer outputs to permanent Butler repo
 - Post metric outputs to SQuaSH/Sasquatch/EFD
 - Move to SLAC K8s if available
- Consolidated Database prototype
 - Per-visit EFD summary
- DMTN-143
 - Raw image upload service if needed

Pre-Ops

- Integration and standing up USDF
- Campaign Management
- Security
- Hybrid model
- Rucio/Butler/PanDA integration