

Data Backbone and Consolidated DB: Features and SLA



Data Backbone: What Is It?



Transfers data from Enclave to Enclave

Raw images from Chile to NCSA (offline); data products from NCSA to Chile

Stores and manages archived data permanently (with replication, checksums)

- Prompt, Calibration, Data Release, Special Programs products (files, DBs, MOPS DB, Alert DB)
- Transformed EFD
- Handles deep storage and disaster recovery
- Handles certain limited-lifetime caches
- Records metadata and provenance (shared with Butler Registry)

Delivers archived files via filesystem (Butler) and possibly web service interfaces

Consolidated Database: What Is It?



Holds most non-Oserv database instances

- Butler Registries and any other image metadata tables
- MOPS DB (assuming relational)
- Data Backbone tracking/metadata/provenance tables
- User Database Workspaces
- Workflow state

Not expected to hold "live" PPDB, but could hold user replicas

Projected uses for DRP intermediates appear to be replaced by Butler Gen3 functionality

Will nevertheless need spatial query capability for TAP/ADQL

Ideally via internal functions, but could be via pixel index

Consolidated to ease deployment and management

SLA Issues



Hardware maintenance

Software maintenance

Schema updates

- Column additions, deprecations
- Reorganizations
- Sharing tables between DBB and Butler; between Prompt and Data Release

Speed and latency and synchronization of file and database transfer

Daily calibration is a limiting use case

Identifier generation

Data Backbone Timeline



Initial DBB Storage already acquired

Initial (minimal) DBB Metadata deployed

Almost ready to start ingesting test stand data into permanent archive (small test)

Ingestion tools need work

- Awaiting refinement of dataset metadata (Butler Gen3 Registry and more?)
- Awaiting astro_metadata_translator
- Awaiting "new dataset appearance" trigger

Retrieval interfaces other than filesystem mount not yet defined