

Vera C. Rubin Observatory

Data Management Monthly Report

January 2020

High-level Summary

Community Interactions, Meetings and Workshops

DM Project Manager Wil O'Mullane attended the [235th meeting of the American Astronomical Society](#) in Hawai'i, at which the renaming of the project to the Vera C. Rubin Observatory was announced. This useful meeting provided the opportunity for a number of useful interactions with the scientific community and the project's funding agencies.

Several members of the DM team attended the Dark Energy Science Collaboration winter meeting in Tucson. Leanne Guy and Robert Lupton were there as official liaisons for the DM System Science Team, while Wil O'Mullane, Chris Morrison, Melissa Graham and Simon Krughoff also participated in the meeting.

Much of the DM team is now focused on preparations for the [Rubin Observatory Algorithms Workshop](#), a key part of our strategy to engage with the scientific community, which will take place in Princeton in March.

Technical Progress

This month saw the DM Middleware Team transition to new leadership under Middleware Manager Tim Jenness (AURA) and Product Owner Robert Gruendl (NCSA). We thank Fritz Mueller (SLAC) for his leadership of this effort over the last several years. This new team got rapidly to work, enabling ingest of precursor data from DECam into the "Generation 3" Data Butler system.

The Architecture (primarily Lim and Jenness), Data Release Production (Fisher-Levine and Plazas), System Science (Lupton), and Data Facility (Pietrowicz, Morganson, Menanteau, Win, and others) teams spent much of this month directly supporting System Integration, Test, and Commissioning for LATISS¹ and the Commissioning Camera.

¹ The LSST Atmospheric Transmission Imager and Slitless Spectrograph; the primary instrument on

The Architecture team also documented the impacts of LCR-1923 — descoping the Camera Data Acquisition System crosstalk correction capabilities — on the DM requirements and system design.

The Pipelines team made a number of algorithmic improvements, including releasing a prototype implementation of the HeliLinC algorithm ([Holman et al, 2018](#)) for linking solar system objects; substantial improvements to photometric calibration through improved aperture corrections; and better brighter-fatter effect correction kernel generation ([Antilogus et al, 2014](#)), with many thanks to Craig Lage (UC Davis) for help with the last of those. We have also tested the new [SCARLET](#) deblender release on precursor data.

The first acceptance test campaign for a Science Pipelines release is currently being run by the DM System Science team. Many performance requirements and other metrics have been verified; the results are reported in [DMTR-201](#).

A Python helper package that wraps some complexity encountered by users of our Engineering and Facility Database (EFD) has been demonstrated to stakeholder and released as a prototype.

New hardware has been deployed at the Data Facility and is ready for developers to use. This includes a number of new Kubernetes nodes, 3 PB of additional data storage, and extra test machines for the Qserv database system. Furthermore, the NCSA camera test stand is now configured with Kubernetes.

This new hardware is already being for evaluating [Apache Cassandra](#) as a platform for implementing the Alert Production Database (APDB), which has extreme performance demands.

Puppet automated deployments are now stable in Chile. All network links remain operational and we are preparing to migrate LATISS traffic off the general AURA network to the LSST data network in February.

Risk Management

The DM Risk Register was reviewed in the monthly process. No new risks were added and no significant changes to existing risk exposure were made.

Milestone Summary

Milestones Completed

DM-NET-4: Base LAN installed

This includes the Base Campus (User) Network for the data center and offices as well as the VOIP telephone system connected to the general phone system in Chile. (The Base Core/Control Network is a separate milestone IT-725-M below which is still in work).

Completion date (as reported in PMCS): 1/31/20

Documentation is captured in JIRA in issues in the following components:

Base Data Center Infrastructure (37 issues)

Base Campus (User) Network (33 issues)

Base VOIP System (15 issues)

Milestones Delayed

Data Release Production

DM-DRP-8: Calibration product generation for the Auxiliary Telescope

Due 2020-01-21

The DM team is currently supporting initial observations with the Auxiliary Telescope, including the generation of calibration products. This milestone will be completed when that system achieves a relatively mature capability.

DM-DRP-29: Moving point source model fitting now available

Due 2018-11-30

This milestone was delayed due to staff shortages; it will be reprioritized and addressed during the current calendar year.

DM-DRP-37: Artifact rejection and background matching during coadd construction

Due 2018-11-30

Artifact rejection steps are complete and well tested. Background matching is regarded as a scientifically lower priority, and will be addressed during calendar year 2020.

Science User Interface and Tools

DM-SUIT-5: Search and display processed HSC data

Due 2019-02-28

This is still awaiting the HSC data to be in OBS-CORE, which is underway.

DM-SUIT-8: SUIT portal integrated with workspace

Due 2018-11-30

Work is ongoing on [DM-22559](#), with completion expected in Feb 2020.

DM-SUIT-10: SUIT deployment procedure

Due 2019-05-31

Still awaiting formal documentation, in the shape of [DMTN-136](#).

Science Data Archive & Application Services

DLP-802: Alert Production Database (APDB) Design

2016-11-30

The combination of aggressive performance and science requirements for this database have made a working design at scale more difficult than was originally anticipated. A functional prototype at reduced scale has been produced, and was evolved with feedback from the AP team. Work in search of a full-scale solution continues in S20 with scheduled experiments involving different back-end database technologies.

DM-DAX-5: Database ingest in support of HSC reprocessing (i.e., large catalog ingest)

2019-02-28

This milestone is pending completion of DM-DRP-11 (Support for database ingestion of results). The dependency between these milestones was not previously correctly captured in the project management system; it will be corrected by an LCR.

LSST Data Facility

DM-NCSA-11: Verified acquisition of raw and crosstalk-corrected exposures at raft scale, incl. correct metadata

2019-07-29

Following LCR-1923, we no longer expect to acquire crosstalk-corrected exposures, and the scope of this milestone is modified accordingly. The milestone will be met after ComCam is available on the Tucson test stand and is capable of providing raft-scale raw images, currently expected in March 2020.

DM-NCSA-20: ComCam Archiving Service

Due 2019-09-25

This milestone will be met after ComCam is available on the Tucson test stand, which is currently expected in March 2020.

International Communications & Base Site

IT-725-M: Base Data Center Network Complete

2019-07-29

This milestone has been delayed due to late acceptance of the Base Facility (including the Data Center) from GHG, the construction contractor (a Telescope & Site team deliverable). The Data Center and Office Building LANs are installed sufficient to handle construction, Telescope AIV, Auxiliary Telescope Commissioning, and Commissioning Camera bench testing, but a few non-critical elements remain to be installed. Currently projected date is the end of April 2020. The critical path is not impacted.

Detailed Project Progress

1.02C.01: System Management

Current accomplishments

The DM Project Manager:

- Attended the [235th meeting of the American Astronomical Society](#) in Hawai'i, where he:
 - He supported the Rubin Observatory booth;
 - Spoke to many attendees about the Observatory;
 - Gave several virtual tours.
- Participated in the Joint Quarterly Status Meeting where there were discussions about the operations-era Data Facility.
- Supported the Rubin Observatory Project Manager at the AURA Management Committee for LSST meeting in Tucson.
- Worked on updates to baselined documentation.

Planned activities

The DM Project Manager will work on the Rubin Observatory Operations Proposal.

1.02C.02.01: Data Management Science

Current accomplishments

The DM Subsystem Scientist:

- Attended the Dark Energy Science Collaboration (DESC) winter collaboration meeting as DESC liaison.
- Attended the Project Science Team (PST) face-to-face meeting in Tucson, reporting on DM status and progress.
- Continued preparations for the Algorithms Workshop in March.
- Prepared the Rubin Observatory Operations proposal, and staffing and activity plans for the System Performance operations team.

Other DM System Science Team (DM-SST) activities include:

- Robert Lupton (DESC liaison), Melissa Graham and Simon Krughoff attended the DESC winter science collaboration meeting.
- Several new performance metrics were implemented and added to the `validate_drp` package
- Jeff Carlin carried out DM acceptance testing as outlined in the acceptance test plan, [DMTR-201](#).
- Jeff Carlin and Simon Krughoff reviewed the architecture of the `validate_drp` package with a view to redesigning it to scale to processing "LARGE" size datasets.
- Eric Bellm reported on recent progress on time series features to compute.
- Hsin-Fang Chiang defined and created a new validation dataset based on the DESC Run 2.2i dataset, which corrects errors in the simulations with active galactic nuclei that had too-bright magnitudes.
- Hsin-Fang Chiang published [DMTN-137](#), summarizing the Proof of Concept of cloud data processing on the Amazon Web Services platform.
- Hsin-Fang Chiang worked on the post-processing pipeline `TransformObjectCatalogTask` to make output object tables independent of the input data and improve test coverage.

Planned activities

The DM Subsystem Scientist will:

- Continue to focus on preparations for the Algorithms Workshop in March, including a review and update of the Data Products Definition Document ([LSE-163](#)).
- Continue to work on the Rubin Observatory Operations Proposal.
- Review draft outlines of construction papers as coordinating editor for DM.

Other DM System Science Team (DM-SST) activities will include:

- The team will be mostly focused on preparations for the Algorithms Workshop.
- Hsin-Fang Chiang will begin reprocessing the HSC PDR 2 data in preparation for the Algorithms Workshop.
- Melissa Graham will contribute to the review of the operations proposal.
- Jeff Carlin will visit Princeton for a week to work with the team there on analyzing precursor data in preparation for the Algorithms Workshop.

1.02C.02.02: DM System Architecture

Middleware Team activities, which are managed by the Architecture Team but involve members from across DM, are now reported in this section.

Architecture accomplishments

In release, verification, and standards activities, the Architecture Team revised the public communication of the official release process. A team member improved the docsteady document generation tool (verification element baseline function, recipe for building the tool, automated document build system) and completed [LDM-732](#), the Network Verification Elements Document. The Python style guide was updated in accordance with [RFC-650](#) to match current community practice. The Qserv-related products in the product tree were reviewed.

In requirements and design activities, specific requirement impacts of LCR-1923 (descope Camera DAQ crosstalk correction) were derived, and a proposal was written for simplification of pixel data retrieval enabled by this. A (near-) final version of [DMTN-133](#) designing an automated observing-script-controlled processing system was issued.

Key coding activities supported System Integration, Test, and Commissioning, including fixing header metadata problems with LATISS on-sky data (OBJECT, right ascension, declination, airmass, and filter names) and changing Generation 2 Data Butler ingest to not require reconnection to the database to see new datasets. The Team enabled the metadata translator to deal with groups of related exposures, necessary for LSST standard science visits and useful for calibration modes. Relatedly, the Gen2 Butler configuration in the obs_lsst package was updated to provide and use the numeric exposure id and string observation id in place of the visit id (especially for raw images). Incompatibilities of Science Pipelines code with newer versions of Numpy, Astropy, and matplotlib were fixed. Large, unnecessary test outputs were removed from the installed pipe_tasks package. Deprecation warnings about configuration items were improved. The requests package

from conda is being used instead of the eups-packaged version. The `translate_header` utility in `astro_metadata_translator` can now output the fixed header in addition to the original one, and the tool was given a summary tabular output mode.

Middleware accomplishments

During January, management of the Middleware Team transitioned from the Data Access team to Architecture. This is reflected in updates to [LDM-294](#) being made on [RFC-666](#). The registry access code was refactored to use the new abstract database interface, resulting in a more consistent interface to different database backends. Enhancements were made to the ingestion code to support vectorized dataset inserts and support was added to allow DECam data to be ingested where multiple datasets are associated with a single file. Improvements were made to the handling of read-only butlers and work was started to clarify how collections relate to runs. We continue to verify that we can ingest Hyper Suprime-Cam data into an Oracle-based registry.

Planned Architecture activities

A meeting with Google Cloud representatives to discuss potential usage of cloud resources will be held. The Team will prepare for and attend the DMLT virtual face-to-face meeting. The Team will begin transitioning authentication and authorization code to the new SQuaRE security architect.

A Team member will develop an outline for an end-of-Construction paper describing the overall DM system ([PSTN-017](#)).

The SUI-related products in the product tree will be reviewed, and obsolete products will be removed. The Team will start to review the requirements-to-products traceability matrix.

The Team will fix various `obs_Isst` Butler Gen2 configuration problems and add metadata translator code to support LATISS observing and ComCam testing (calibration and defects lookup, CCD identification, filter/grating naming, right ascension/declination handling, image types). Problems in the handling of multiple simultaneous conda environments by the Jenkins continuous integration system will be fixed, and maintenance of the `Isstsw` Science Pipelines build tool will be performed. A mechanism for marking LaTeX documents as obsolete will be provided. A feature will be added to the `pex_config` configuration package to simplify loading of configuration files that are located in the same or nearby directories as others. The base conda environment will be updated to newer package versions, with any resulting incompatibilities being fixed.

Planned Middleware activities

For February, we intend to further improve registry performance and also release a more general tool for converting Gen2 repositories to Generation 3. It should also be possible to run cp_pipe on LATISS data. Work will begin on changing the way that ingest works in order to separate visit definitions from exposure definitions and to support incremental ingest.

1.02C.03: Alert Production

Current accomplishments

02C.03.00 – Management and Leadership

- Eric Bellm and John Swinbank visited NSF's OIR Lab in Tucson for a productive discussion with Carl Stubens, developer of the [ANTARES](#) alert broker, about the future of alert distribution and filtering.
- Chris Morrison attended the Dark Energy Science Collaboration meeting in Tucson, AZ.
- We continue to collaborate with a team from the [Astronomical Observatory of Belgrade](#) on development of an alert stream simulation package.
- The AP team is focused on preparations for the [Rubin Observatory Algorithms Workshop](#), which will take place in mid-March in Princeton, NJ.
- John Swinbank joined the team planning for Rubin Observatory operations; over the next several months, he will spend less time on DM construction, and will instead work on preparing for the operational era.
- Planning has begun for the DM Leadership Team (virtual) face-to-face meeting, which will take place in mid-February.

02C.03.01 – Single Frame Processing

- No work was completed in this WBS element.

02C.03.02 – Catalog Association for Alert Production

- Numerical warnings generated by some DIAObject characterization plugins have been eliminated. [[DM-22727](#)]
- The algorithm used to select DIASources for forced photometry has been improved, eliminating a class of misleading results which were being stored to the AP database. [[DM-22777](#)]
- The association system has been restructured in preparation for integration with

the new “Generation 3” DM middleware. [\[DM-22741\]](#)

02C.03.03 – Alert Distribution System

- No work was completed in this WBS element.

02C.03.04 – Alert Generation Pipeline

- Differential Chromatic Refraction (DCR) mitigation [\[DM-21438\]](#):
 - An error in image differencing with DCR-corrected templates was traced to an attempt to use template data which did not contain enough pixels overlapping the area of the science image to be worthwhile. This has been resolved by automatically skipping those patches. [\[DM-21237\]](#)
 - This completes activity [DM-21438](#). Further work on mitigating differential chromatic refraction will occur as part of ongoing pipeline QA and algorithmic refinement on [DM-22634](#).

02C.03.05 – Tools for Science Pipelines

- Emergent work [\[DM-22484\]](#):
 - A total of nine tickets were resolved, fixing bugs, addressing emergent feature requests, and improving documentation across the codebase.
 - Highlights included:
 - Clarifications to the Task framework documentation. [\[DM-11097\]](#)
 - The Jointcal calibration tool can now write its initial model to disk to facilitate debugging. [\[DM-21004\]](#)
 - An astrometric reference catalog based on [Gaia](#) Data Release 2 is now available for the ap_verify_hits2015 test dataset regularly used by the AP team. [\[DM-22676\]](#)
 - The [DM glossary](#) has been updated to clarify terminology around “prompt processing”. [\[DM-23025\]](#)
 - A number of spurious and confusing warning messages being emitted by DM code have been eliminated. [\[DM-23056\]](#)
- Pipeline infrastructure updates [\[DM-22633\]](#):
 - The team continues to work on integrating the AP pipeline with the new, “Generation 3”, middleware. This has involved integrating AP-specific data structures (like the DcrModel and its attendant subfilters) with the new Data Butler, and engaging with the Middleware Team to resolve issues which are unique to the DECam data used for testing the AP system. [\[DM-22648, DM-22708, DM-23008, DM-23249\]](#)

- Building on the above work, the team has given some thought to handling of calibration products in the operational era, and have proposed this topic for discussion by the DM Leadership Team at their face-to-face meeting in February. [[DM-22482](#)]

02C.03.06 – Moving Objects Processing System (MOPS)

- A prototype Rubin Observatory implementation of the HeliLinC algorithm ([Holman et al, 2018](#)) has now been completed and is being tested. [[DM-19725](#)]
- Rubin Observatory's interface to the [NASA Jet Propulsion Laboratory Small Body Database](#), pysbdb, now supports the [Astroquery](#) system. [[DM-22374](#)]
- Orbit covariances can now be converted to Cartesian states and properly propagated. [[DM-23209](#), [DM-23210](#)]

02C.03.07 – Transform Fitting on Stacks of Images

- No work was completed in this WBS element.

02C.03.08 – Integration

- Pipeline testing and integration [[DM-22634](#)]:
 - Template generation for AP test dataset reprocessing has now been automated. [[DM-18828](#)]
 - Updated templates from the [HiTS](#) dataset, correcting known issues with the previous template set, have been generated and are now in use regular reprocessing and test campaigns. [[DM-21330](#)]
 - Regular monthly reprocessing and [analysis](#) of the test datasets has been completed. This campaign used improved settings for template generation, only including images with good point spread functions. This was reflected in the outputs: excess sources which were seen in previous test campaigns have been eliminated. [[DM-23069](#)]

Planned activities

02C.03.00 – Management and Leadership

- Prepare for and participate in the DM Leadership Team (virtual) face-to-face meeting.
- Continue preparations for the Rubin Observatory Algorithms Workshop.
- Contribute to developing the Rubin Observatory Operations Proposal.

02C.03.01 – Single Frame Processing

- No work is scheduled in this WBS element.

02C.03.02 – Catalog Association for Alert Production

- Continue working on Science Data Model support within the Alert Production system.

02C.03.03 – Alert Distribution System

- No work is scheduled in this WBS element.

02C.03.04 – Alert Generation Pipeline

- Prepare algorithmic material for discussion at the Rubin Observatory Algorithms Workshop next month.
- Debug issues identified with the decorrelation “afterburner” used in image differencing.

02C.03.05 – Tools for Science Pipelines

- Continue efforts to update the Alert Production system to integrate with the new “Generation 3” middleware which is being rolled out across the Data Management subsystem.
- Research alternative data persistence frameworks.

02C.03.06 – Moving Objects Processing System (MOPS)

- Shepherd [RFC-620](#) to a conclusion.
- Complete end-to-end processing of an example solar system dataset with a prototype of the Rubin Observatory solar system processing pipeline.

02C.03.07 – Transform Fitting on Stacks of Images

- Work with the DRP team to identify performance bottlenecks in Jointcal and plan a path forwards.

02C.03.08 – Integration

- Regular pipeline reprocessing and quality analysis of the results.
- Develop a system for tracking simulated (“fake”) sources through alert production processing.

Staffing update

- Spencer Nelson has accepted a position as an alert stream developer at the University of Washington. He will join the DIRAC Institute at the University of Washington; and will be co-funded by Rubin Observatory and by other, related, projects underway at DIRAC. He will start in mid-March.
- Eric Bellm, Meredith Rawls, and John Swinbank conducted interviews for a software developer to join the Rubin Observatory team at the University of Washington. At time of writing, an offer has been made to one of the applicants, and negotiations are ongoing.

1.02C.04: Data Release Production

Current accomplishments

02C.04.00 – Management and Leadership

- Jim Bosch, DRP Science Lead, is serving on the organizing committee for the [Rubin Observatory Algorithms Workshop](#). Meanwhile, the rest of the DRP team is focused on preparations for this meeting.
- Yusra AlSayyad, Image Display Working Group chair, completed a first draft of the working group's report. A complete version will be presented to the DM Leadership Team in February.

02C.04.01 – Software Primitives

- “Generation 3” middleware development [[DM-21254](#), [DM-22586](#)]:
 - Throughout this month, members of the DRP team continued to participate in development of the “Generation 3” middleware (task execution framework and “Butler” data access abstraction) and to convert existing algorithmic code to the new framework. This work is carried out under the overall coordination of the Architecture team. [[DM-21786](#), [DM-22609](#), [DM-22771](#), [DM-23079](#), [DM-23180](#)]
- Emergent work [[DM-22588](#)]:
 - A total of seven tickets were resolved, fixing bugs, addressing emergent feature requests, and improving documentation across the codebase.
 - A particular highlight was that improvements to improved aperture corrections, based only on bright stars, have substantially improved photometric calibration in single-frame processing of precursor (Hyper

Suprime-Cam) data. [[DM-23071](#)]

02C.04.02 – Calibration Products

- Detector characterization [[DM-21282](#)]:
 - A major upgrade to the way that correction kernels for the brighter-fatter effect ([Antilogus et al, 2014](#)) has been completed. Thanks to Craig Lage (UC Davis) for contributing the algorithmic thinking behind this approach. [[DM-18683](#), [DM-22659](#)]
 - It is now possible to run instrument signature removal on detectors which have no linearizer. [[DM-22680](#)]
 - Masked regions are now correctly grown to account for the effects of convolution during brighter-fatter effect mitigation. [[DM-23083](#)]
 - This marks the completion of activity [DM-21282](#); detector characterization efforts will continue on [DM-22594](#).
- Auxiliary Telescope development [[DM-22593](#)]:
 - Rubin Observatory's version of the [Spectractor](#) system used for analyzing spectroscopic data from the Auxiliary Telescope has been updated to incorporate the latest upstream changes.
 - The DM team has been working to support early on-sky results from the Auxiliary Telescope.

02C.04.03 – Image Characterization

- Background estimation [[DM-22597](#)]:
 - Exploratory work into understanding amp-to-amp offsets in Hyper Suprime-Cam data is now underway. Preliminary results are presented in [these slides](#); analysis will continue next month.

02C.04.04 – Coaddition

- No work was completed in this WBS element.

02C.04.05 – Detection and Deblending

- Deblender development [[DM-22587](#)]:
 - The new proximal adaptive moment estimation code in [SCARLET](#) 1.0 was tested on precursor (Hyper Suprime-Cam) data. It was found to be much more robust and produced lower residuals. However, it introduced a bias in galaxy photometry; this was traced to using a high detection threshold, and solved. [[DM-22137](#), [DM-23035](#)]

- A framework of plots and catalogs which can be used for comparing deblending algorithms has been developed. [[DM-22536](#)]
- Improvements to the weighting scheme used in `meas_extensions_scarlet` have substantially reduced the number of iterations required to reach convergence. [[DM-23192](#)]

02C.04.06 – Characterization and Measurement

- Galaxy photometry metrics [[DM-22592](#)]:
 - An extensive comparison of [MultiProFit](#) processing of Hyper Suprime-Cam observations of the COSMOS field with observations of the field from the Hubble Space Telescope has been completed. The results are summarized in this (lengthy) [Jupyter notebook](#). [[DM-21997](#)]

02C.04.07 – Maintenance, Quality and Documentation

- Forward Global Calibration Method (FGCM; [Burke et al, 2018](#)) development [[DM-22591](#)]:
 - `fgcmcal` can now distinguish between the WCS-based Jacobian and the illumination correction. [[DM-20163](#)]
 - The coaddition system can now take account of calibrations from different sources (FGCM, FGCM-tract, Jointcal) when building coadds. [[DM-21308](#)]
 - Assorted minor modifications and enhancements have been made to the `fgcmcal` codebase. [[DM-22470](#), [DM-22798](#)]
 - `fgcmcal` can now perform a local background subtraction on aperture fluxes. This has been shown to substantially improve photometric repeatability in all bands. [[DM-23036](#)]
- Quality assurance [[DM-21253](#)]:
 - An issue in which some diagnostic data was not being properly written when running with simulated sources has been resolved. [[DM-23026](#)]

Planned activities

02C.04.00 – Management and Leadership

- Focus on preparing for the [Rubin Observatory Algorithms Workshop](#). In particular, a preparatory data analysis sprint will take place during the weeks of 17 and 24 February. Jeffrey Carlin will visit Princeton to participate in this sprint.
- DRP team leadership will prepare for and participate in the [DM System Science](#)

[Team](#) and [DM Leadership Team](#) (virtual) face-to-face meetings in mid-February.

02C.04.01 – Software Primitives

- Continue development of the Generation 3 middleware.
- Service emergent feature requests and bug reports.

02C.04.02 – Calibration Products

- Act in support of ongoing Auxiliary Telescope observations.
- Work towards a continuous integration system for calibration products production.
- Simplify linearity correction infrastructure.

02C.04.03 – Image Characterization

- Continue investigating amp-to-amp background offsets.

02C.04.04 – Coaddition

- No work is planned in this WBS element.

02C.04.05 – Detection and Deblending

- Present the current status of deblending to the DM System Science Team.
- Investigate and eliminate causes of poor deblending performance with SCARLET.

02C.04.06 – Characterization and Measurement

- Continue exploring the galaxy-model-fitting parameter space using MultiProFit.

02C.04.07 – Maintenance, Quality and Documentation

- Analysis of precursor and simulated datasets in support of Algorithms Workshop preparation.

Staffing update

- Morgan Schmitz joined the DRP team in early January as a Science Pipelines developer. Morgan obtained a PhD from CEA Saclay in 2019.

1.02C.05: Science User Interface & Tools

This WBS supports the activities of the Science Platform Scientist Gregory Dubois-Felsman, which are reported under the Architecture (1.02C.02.02) and DM Science (1.02C.02.01)

teams. An additional fractional FTE is used to support bug fixes on Firefly which are not reported here in detail.

1.02C.06: Science Data Archive & Application Services

Current accomplishments

02C.06.00 Management & Leadership

- SLAC had winter closure from December 23 to January 3.

02C.06.01.01 Catalogs, Alerts and Metadata

- Salnikov updated his Cassandra APDB (Alert Production Database) development branch [[DM-23214](#)], and Cassandra evaluations were begun using new hardware deployed at NCSA.

02C.06.01.02 Image and File Archive

- No work was carried out in this WBS element this month.

02C.06.02.01 Data Access Client Framework

- Management of work for this WBS element has been transferred to ARCH, and corresponding progress for this and future months will be reported in that section (02C.02.02) of this report.

02C.06.02.02 Web Services

- Work and staffing for this WBS element have been transferred to SQuaRE, and corresponding progress for this and future months will be reported in that section (02C.10) of this report.

02C.06.02.03 Query Services

- Gates fixed an issue with handling large results in Qserv [[DM-22220](#)].
- Gates tested prototyp distributed secondary index code on a full Qserv deployment [[DM-17656](#)].
- Gates fixed an issue where Qserv could silently fail if two worker instances tried to use the same MySQL instance.
- Gaponenko added a batch mode to the Qserv ingest system for allocating chunks during catalog ingest [[DM-23087](#)].
- Gaponenko added an identity mechanism into the Qserv Replication system to

prevent incorrect communication between Qserv instances [[DM-23191](#)].

- Mueller fixed Qserv container builds by removing explicit eups install of pytest, which is now installed by Conda. [[DM-23084](#)].
- Mueller fixed the LSP (LSST Science Platform) logging config to reflect the recently changed Qserv logging settings [[DM-23006](#)].
- Jammes changed the Qserv docker container build scripts to use fixed tags including the build the date [[DM-20332](#)].
- Jammes modified the Qserv Kubernetes Operator to run at NCSA [[DM-21824](#)].
- Jammes and Gaponenko to duplicated data on the NCSA Qserv nodes, permitting running stable deployments in support of LSP users and experimental Kubernetes-based deployments simultaneously on the same set of hardware at NCSA.

02C.06.02.04 Image Services

- Lo continued work on porting SODA image service to Gen3 Butler, and on creating a Postgresql ObsTAP test dataset.

02C.06.02.05 Catalog Services

- No work was carried out in this WBS element this month.

02C.06.03 Task Framework

- Salnikov implemented printing of Gen3 task configs in a manner that is easily diff'd [[DM-22301](#)].
- Salnikov removed extra newlines from output of pipeline "--show=config" [[DM-22851](#)].

Planned activities

02C.06.00 Management & Leadership

- Mueller and Pease to attend DMLT Virtual Face-to-Face meeting.
- Pease and Lo have vacations planned for this month.

02C.06.01.01 Catalogs, Alerts and Metadata

- Salnikov and Hanushevsky to continue APDB Cassandra implementation prototype and performance characterization work.

02C.06.01.02 Image and File Archive

- No work is planned for this WBS element this month.

02C.06.02.03 Query Services

- Gaponenko and Gates to profile and characterize Qserv query dispatch performance in preparation for 50% DR1 large-scale testing.
- Mueller to begin rework of Qserv build procedures decrease build time, decrease container sizes, and simplify dependency management.
- Pease and Jammes to work integrating Redis into Kubernetes-operator deployed Qserv instance and NCSA, to facilitate evaluation of Redis as a potential technology for hosting Qserv global spatial indices.

02C.06.02.04 Image Services

- Lo continued work on porting SODA image service to Gen3 Butler, and on creating a Postgresql ObsTAP test dataset.

02C.06.02.05 Catalog Services

- No work is planned for this WBS element this month.

02C.06.03 Task Framework

- Salnikov to continue with enhancements/fixes to task framework in support of ongoing Gen3 middleware development.

Staffing update

- Nothing to report.

1.02C.07: LSST Data Facility

Current accomplishments

02C.07.05 LSST Data Facility Management, Service Architecture, and Project Controls

The Management Team at NCSA:

- Continued initial planning for Operations Rehearsal #2 ([LDM-643](#)).
- Continued regular steering meetings for internal oversight over all technical areas of the LSST Data Facility.
- Continued working toward a working NCSA test stand for all CSCs involved in the AuxTel and ComCam environments in Tucson and at the summit.
- Begin implementing a 7x24 hour staffing cycle at NCSA to monitor the IHS ticket

system and assign tickets to the NCSA staff as needed along with keeping close watch on the #dm-infrastructure Slack channel for possible problem reporting after normal working hours.

- Continued participation with System Science Team, LDF Infrastructure meetings, Data Management Leadership Team, Commissioning meetings, ComCam meetings, AuxTel meetings, IT Support Committee, and operations proposal writing team.

02C.07.06 LDF Production Services

- All hardware ordered in the 2020 budget year for NCSA has been racked and stacked with OS's loaded and are ready for service for the development staff.
- Continued monthly reprocessing of each the RC2 and DC2 datasets to support software stack testing and pipeline code development.

02C.07.07 Data and Compute Services

- We have completed the build for the APDB cassandra testing to be done in Feb for the alert processing database.
- We continuously improve the Kubernetes cluster infrastructure.
- We built a K8 infrastructure in the NCSA test stand for the Telescope and Site to upload software to for testing before deploying on the summit for AuxTel.

02C.07.08 LDF Service Software

- For Prompt Service software, we continued incremental development and improvements for the Archiver, Forwarder, Observatory Operations Data Service (OODS) and Header Service software to support LATISS on the mountain and helping to build that system during its initial deployment.
- For Batch Production Services, work continued to integrate the HTCondor workload management system, as well as responding to emerging changes in the Gen3 middleware code base.
- For the Data Backbone, we continued testing file transfer elements to evolve the mechanism to a more robust solution that would search for new files and send them as soon as they were found.
- We continued supporting data transfer from test stands at SLAC, monitoring file transfer and verifying data ingestion.
- We continued work adding features to the new Disaster Recovery service, aimed at hardening and improvements to operationally support on-sky LATISS data.

02C.07.09 ITC and Facilities

- The new Qserve nodes, K8 nodes extensions at NCSA and BDC, storage disk expansions for both Chile and NCSA and a development AMD ROME node all arrived were deployed.
- We continued work with Configurable SAL Components (CSCs) on the NCSA test stand, installing Kubernetes with 6 nodes, moved the EFD environment to the K8 environment while working with the developers on the deployment.
- We continued responding to incidents and requests concerning as-is development systems and services, as well as addressing emerging hardware and system needs, and reviewing vulnerability reports and performing security reviews for system changes.

Planned activities

02C.07.05 LSST Data Facility Management, Service Architecture, and Project Controls

The management team at NCSA plans to:

- Continue regular steering meetings for internal oversight over all technical areas of the LSST Data Facility.
- Continue developing and documenting test plans for Data Facility components and services.
- Complete and upload to P6 planning activities for the next development cycle.
- Continue regular meetings with CC-IN2P3 to coordinate near-term work and discuss operations.
- Continue participation with the Summit/Base Tiger Team, Science System Team, LDF Infrastructure meetings, Data Management Leadership Team, IT North/South Team, Data Management Change Control Board, ITS committee, Comcam/ATS meetings, and Commissioning Team.
- Finish the initial deployment plan with the 7x24 staff for watching the tickets and #dm-infrastructure Slack channel for problems, while getting people involved in off hours if needed.

02C.07.06 LDF Production Services

- We will continue periodic reprocessing of datasets in support of stack testing and pipeline development. In February, we are working with the pipeline developers to run the PDR2 software run before the Algorithms Workshop in March.
- We will begin data ingest from the ComCam test stand in Tucson continue with

LATISS observing and continue ingest from other SLAC data, enabling Science Platform (LSP) viewing.

- Configuration work toward providing more nodes for the LSP for use by the commissioning team. We also plan to add more storage as LATISS data will be coming into the LDF.
- Build a larger compute cluster by moving Qserv nodes from stable, K8 nodes from stable, and putting the condor testing on hold and add those nodes back in for the other processes needed for PDR2.

02C.07.07 Data and Compute Services

- We will continue work implementing the current phase of file transfer, service endpoint, and file ingestion into the Data Backbone for early commissioning data generated by the LATISS instrument on the summit, the ComCam test stand, and test stands at SLAC.
- We will continue working with developers on improvements to the Kubernetes service as well as Kafka installations to support development for the EFD and alert distribution.

02C.07.08 LDF Service Software

- We will continue work on Prompt Service software with new libraries and Xml needed the LATISS needs as they arise, as well as upcoming support for ComCam testing and integration on the test stand in Tucson.
- We will continue modifications to Forwarder, Archiver, OODS, and Header Service components to support the new DAQ hardware and software on the NCSA Test Stand.
- We will continue to work with Gen3 middleware, using emergent versions of the middleware software to test pipeline execution using Gen3 components with workflow management and database systems, as well as responding to changing needs for database support.
- For the Data Backbone, we will continue facilitating and monitoring automated file transfer and ingestion in support of test instruments at SLAC and in Tucson as well as continue to prototype elements needed for future versions of the data backbone.
- We will continue development work on version 2.0 of the Disaster Recovery service.

02C.07.09 ITC and Facilities

- Activities for February will include configuration for systems to support ComCam on the Tucson test stand, as well as systems that will support LATISS on the summit.

- We will complete planning for and perform scheduled preventive maintenance. (Feb 27, 2020)
- We will continue responding to incidents and requests concerning as-is development systems and services, as well as addressing emerging hardware and system needs.

Staffing update

- We added Kevin Deptula, an extra help project manager, to help with the project management load after Joel Plutchak's retirement.

1.02C.08: International Communications & Base Site

Current accomplishments

02C.08.00 International Communications and Base Site

- Nothing to report.

02C.08.01 – Base Center

- We continued working on the campus and control networks, providing support to different areas at the Base and Summit facilities.
 - Amor and Andes clusters were connected to the Control network
 - Out of band access was configured in IPMI switches.
 - PTP configuration has been completed and it is available for general deployment
 - Puppet service has entered a stable state
 - IPA service has entered a stable state
 - Continued deployment of network connections, servers, and services

02C.08.02 – Chilean Data Access Center

- Nothing to report.

02C.08.03 – Long-Haul Networks

02C.08.03.01 – Chile National WAN

- Summit - AURA Gatehouse:
 - The link remains operational and working properly.

- Dense Wave Division Multiplex (DWDM) Equipment:
 - The DWDM system over the Long haul path remains operational and working properly. Migration of the LSST client card completed. The new card has 4x100Gbps interfaces. A test performed on one of these interfaces using a traffic generator yielded a result of 99.97Gbps.
- Santiago-La Serena:
 - The Santiago ring and La Serena – Santiago primary and backup links remain operational and working properly.
 - Fiber cuts occur fairly frequently on this segment, but the backup links are working correctly and no traffic has been lost. The backup link is currently 4 Gbps, but will be upgraded to 40 Gbps by June, 2020.
- La Serena - AURA Gatehouse:
 - The link remains operational and working properly

02C.08.03.02 – International Chile–US WAN

- *100 Gbps Managed Ring:*
 - The links remain operational and working properly.
- *Management and Coordination Contract:*
 - The LSST NET meeting was held on January 9, 2020. The meeting had the following topics:
 - NET Action Items Status
 - Continental US Networks (CONUS) update
 - Chilean/International Networks update
 - Transatlantic Networks (NCSA - CC-IN2P) update
 - LSST Operations Network Service-Level Agreements (SLAs) and Virtual Network Operations Center (VNOc)
 - LSST Network Verification and Validation
 - Network Design and End-to-End Test Plan and QoS planning
 - Conferences and Workshops
 - Dr. Julio Ibarra attended the Trans-Pacific Research and Education (TPRE) Workshop – January 17-18, 2020 to participate in the planning and operational discussions involving future network connectivity to support science applications.
 - Italo Valci attended the 4th SIG-NGN meeting - January 15-16, 2020 which took place at CERN and is co-located with a two-day workshop on networking for LHC and High-Energy Physics
 - The South American Astronomy Coordination Committee (SAACC) &

LSST NET face-to-face meetings are scheduled for April 22-23, 2020, in Miami, FL. The [meeting web page](#) includes the logistics and draft agenda.

- *Miami - Boca Raton - Atlanta:*
 - No activity, the 100G LSST wave is considered production.
- *São Paulo – Miami Spectrum:*
 - No activity, the 100G LSST wave is considered production.
- *São Paulo – Santiago Spectrum:*
 - CLARA's bid process is finally over and two companies were selected, Century Link for the path Porto Alegre - Buenos Aires and Silica Network for the path Buenos Aires - Santiago. The spectrum from Santiago to Porto Alegre/Brazil is planned to be available by the end of Q2 2020 with full installation by Q3 2020. From Porto Alegre to Sao Paulo, RNP is working with two power companies to use their OPGW fibers: Furnas and Eletrosul. The DWDM installation is in progress and the 100G link from Porto Alegre to Sao Paulo is planned to be active by the end of Q2 2020.
- *US National WAN:*
 - The Continental United States (CONUS) Network Implementation Team (CNIT) led by Paul Wefel of ESnet completed testing the implementation of MIA-ATL-CHI. There is no operational traffic as of yet, but tests are positive. Note that these paths use shared leased paths from Miami - Atlanta currently. Dedicated paths will be implemented during operations.

Planned activities

02C.08.00 International Communications and Base Site

02C.08.01 – Base Center

- Deployment of Core Cluster at the Base
- Integration of M2 Servers to Control Network
- Integration of M1M3 Servers to Control Network
- Deployment of several fibers for DAQ/CCS and other systems.

02C.08.02 – Chilean Data Access Center

- Nothing planned.

02C.08.03 – Long-Haul Networks

02C.08.03.01 – Chile National WAN

- Summits - AURA Gatehouse Network:
 - Enable transfer of LATISS traffic over this segment.
- La Serena - AURA Gatehouse:
 - Enable transfer of LATISS traffic over this segment.
- Santiago-La Serena:
 - Enable transfer of LATISS traffic over this segment.
- DWDM:
 - No planned activity.
 - Install perfSonar servers on summit and base connected to LSST DWDM.

02C.08.03.02 – International Chile-US WAN

- *São Paulo – Miami Spectrum:*
 - Enable transfer of LATISS traffic over this segment.
- *São Paulo – Santiago Spectrum:*
 - Monitor work on all phases
- *US National WAN:*
 - Enable transfer of LATISS traffic over this segment.

Staffing update

- Nothing to report.

1.02C.10: Science Quality and Reliability Engineering

Current accomplishments

- From our project to bring all our Kubernetes-deployed services under the ArgoCD platform in order to bring more uniformity and configuration control:
 - We are now able to deploy our elasticsearch-fluentd-kibana logging service stack with ArgoCD.
 - We are now able to authenticate users with Github OAuth2 tokens in Chronograf services deployed under ArgoCD.
 - The Tucson teststand environment has a small nublado deployment which now is also deployed with ArgoCD.

- We performed extensive work with Documenteer in evaluating ways of supporting the C++ codebase, including prototype sprints with Breathe and Exhale; ultimately after feedback from Science Pipelines developers we have converged on using Doxylink to allow us to link from the Sphinx pages to Doxygen generated ones.
- We demonstrated and collected feedback for EFD helper class whose purpose is to wrap obscure and non-standard queries around a friendly interface available from the notebook environment.
- We deployed ArgoCD and on that an EFD service at the NCSA teststand.
- We put together a TAP service with postgres/pgsphere as preparation for upcoming ObsTAP work.
- We fixed an issue with sensor definitions in obs_lsst (see image)
- Notable external interactions:
 - Frossie Economou gave an invited talk at a Science Cloud meeting in Germany.
 - Simon Krughoff gave a presentation at the DESC Collaboration meeting.

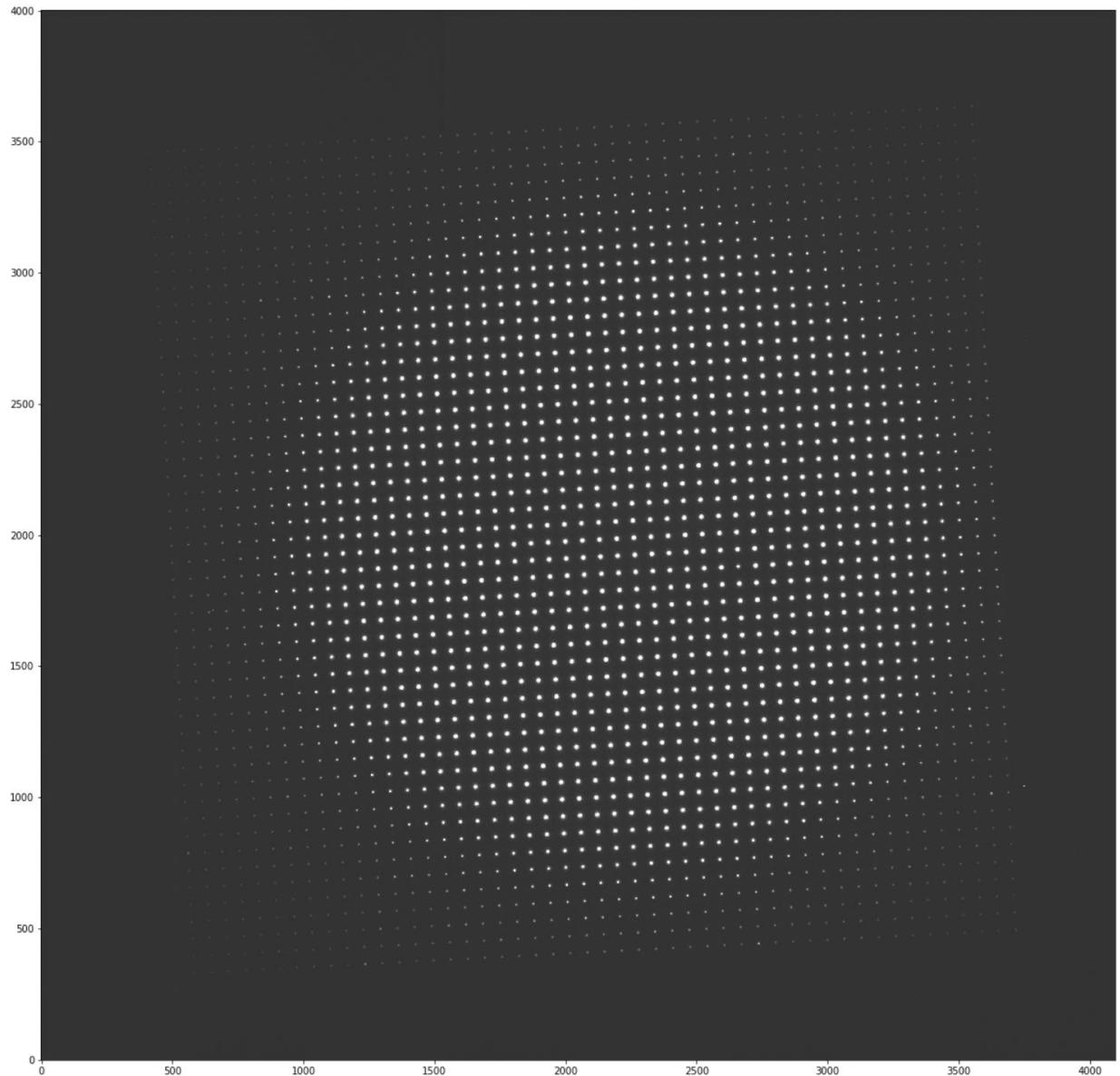


Image: an assembled E2V sensor with a projected spot pattern, used to verify that the configuration in obs_lsst now has the correct sensor definitions so as to result in their right assembly

Planned activities

- Import WISE data to back prototype OBS-TAP service
- Bring Vault under ArgoCD / roundtable
- Test prototype for notebook “headless” execution
- Draft document identifying security gaps in SQuaRE services

- Deploy the EFD at the NCSA test stand
- Deploy logging infrastructure to roundtable
- Move ghsacker to roundtable
- Implement helper class for EFD queries from notebooks

Staffing update

- Our newest team member, Russ Allbery, on boarded on Jan 28th as the Science Platform Security Architect.