LARGE SYNOPTIC SURVEY TELESCOPE

Large Synoptic Survey Telescope (LSST)

Data Management Releases for Verification/Integration

William O'Mullane, Frossie Economou, Tim Jenness, Andrew Loftus

LDM-564

Latest Revision: 2017-08-18

Abstract

This document describes release management at a high level and specific features for upcoming releases.

Change Record

Version	Date	Description	Owner name
1.0	2017-08-18	Initial version. Approved in RFC-373.	W. O'Mullane

Document source location: `https://github.com/lsst/LDM-564`

1 Introduction

1.1 Scope

This document covers releases of software from the Data Management Subsystem of LSST for verification/integration tests. . It discusses the delineation between the Data Facility as an operational entity and DM producing and testing software. It does not cover the normal releases to the community of the software stack - see `https://developer.lsst.io/` for that.

2 Release Management

This section outlines the current understanding of the release management process. Complete definition is pending the appointment of the DM Release Manager.

2.1 Preparation of Releases

DM develops code in GitHub following its developer guidelines and coding standards [1]. This includes automated testing and continuous integration. Tested releases are tagged by SQuaRE weekly and major releases are made each cycle (six months).

There are specific packages and systems deployed together to form the high level components of DM as depicted in Figure 1. The orchestration of deployments on multiple machines is facilitated by the use of containers and a machine readable configurations. DM prepares Docker containers and Puppet configurations for deploying these systems on Kubernetes enabled cluster. These artifacts are tagged as part of the release.

In addition, specific releases with features required to support the LDM-503 test milestones will be tagged and released in advanced of each verification test. The preliminary feature lists for these releases are defined in Section 3.

[1] `https://developer.lsst.io/`

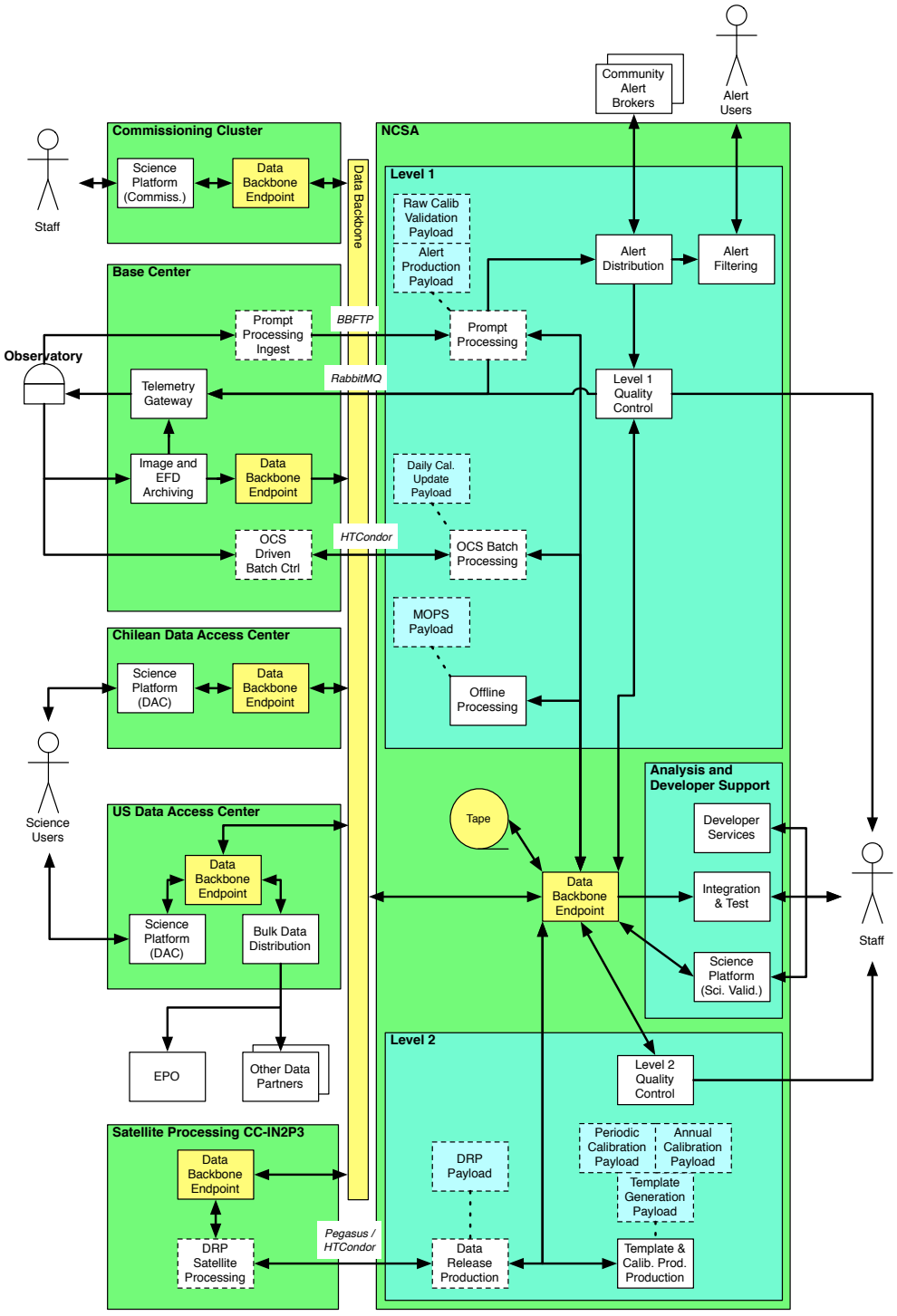

FIGURE 1: DM components as deployed during Operations. Where components are deployed in multiple locations, the connections between them are labeled with the relevant communication protocols. Science payloads are shown in blue. For details, refer to LDM-148.

2.2 Deployment of Releases

Although DM will provide ready-to-install products, these will be further tested before being deployed. Hence, releases will initially be installed on test systems at NCSA and will undergo smoke testing before they are made available in the production environment. This will serve as operational validation of the release.

Once smoke tested, the Docker containers will be made available in the NCSA Docker repository. Using this secure internal repository, operators may deploy containers for specific releases in the operational environment.

2.2.1 Levels of operational validation

Certain containers will be used to provide kernels and supporting libraries for the JupyterLab environment. Multiple versions of these containers can be made available simultaneously — for example, providing a series of minor releases of the software stack — with the user selecting which to deploy for their particular use case. Since they will not be deployed as part of the core operational system, acceptance testing can be relatively minimal.

Some containers will be made available on development systems in support of ongoing development of the code. Again, these should be made available rapidly, with security checking and validation testing kept to a minimum.

Similarly, during Commissioning, availability of containers on the Commissioning Cluster should be on the order of hours (not days). The level of smoke testing and the time to availability of a container may need to be compressed in Commissioning.

Containers to be used for prompt or batch processing on operational systems, on the other hand, must be rigorously validated.

3 Functionality in future DM releases

This is currently not an exhaustive feature list, but rather gives an indication at a high level of the features in each release which will be verified by the corresponding verification test

campaign. As the test plans are written this will become a list of requirements to be tested for that release and thus begin to fill out the verification control database (currently to be in JIRA).

In the feature lists below, the corresponding internal milestone is given in parenthesis.

Each section here is a test milestone from LDM-503 — the same labels are used. The timeline is in the DM schedule using the same labels and depicted in Figure 2

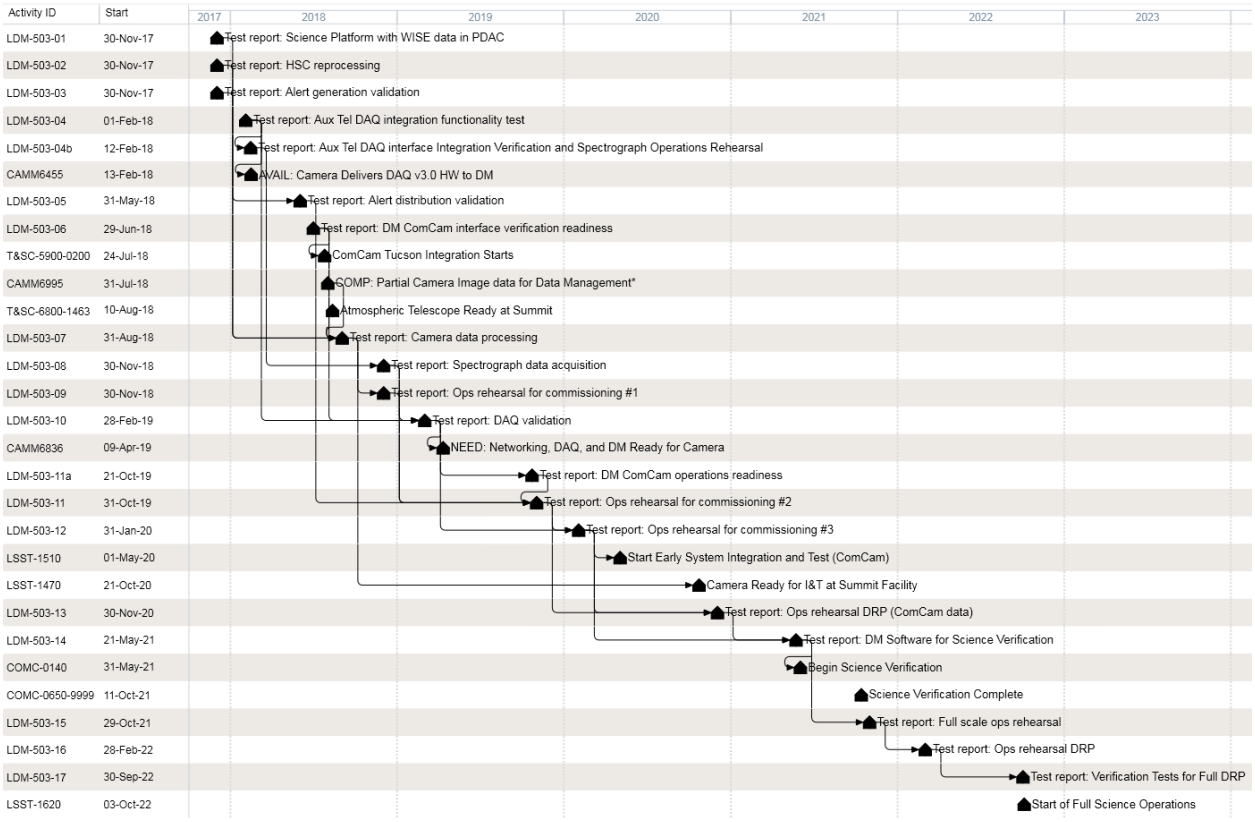

FIGURE 2: DM level 2 milestones (LDM-503-x) in the LSST schedule.

3.1 LDM-503-1: Science Platform with WISE data in PDAC

- WISE data ingest to PDAC. (DM-DAX-1)
- Query service suporting IVOA TAP protocol, with support for asynchronous queries. (DM-DAX-2)

- Image cutout service supporting IVOA SODA protocol. (DM-DAX-3)
- Metadata service suporting IVOA SIAv2 protocol. (DM-DAX-4)
- Search and display WISE sources (objects) in PDAC. (DM-SUIT-1)
- Search WISE coaded data single exposure images in PDAC (the images are from IRSA at IPAC, not NCSA). (DM-SUIT-2)
- Time series analysis tool for WISE data. (DM-SUIT-3)
- Multiple data traces in chart space. (DM-SUIT-4)
- Project internal Jupyter notebook service. (DM-SQRE-1)

3.2 LDM-503-2: HSC Reprocessing

- Basic single frame measurement pipeline. (DM-AP-1)
- Database ingest in support of HSC reprocessing (automatable, large catalog ingest). (DM-DAX-5)
- HSC merger complete: all functionality deployed for the most recent HSC data release processing is now available within the LSST stack. (DM-DRP-1)
- Basic visualization and quality assessment tools operational on HSC-scale data volumes. (DM-DRP-2)
- Regular reprocessing service for HSC data available. (DM-NCSA-1)
- Access to results of regular reprocessing available (NB the form this takes depends upon available DAX functionality). (DM-NCSA-2)
- Provide database for metadata, provenance, location and demonstrate ingest at small scale. (DM-NCSA-3)
- Search and display processed HSC data. (DM-SUIT-5)

3.3 LDM-503-3: Alert Generation Validation

- Basic single frame measurement pipeline. (DM-AP-1)

- Alard & Lupton-style image differencing. (DM-AP-2)
- Point source & dipole measurement on difference images. (DM-AP-3)
- DIASource association. (DM-AP-4)
- DIAObject generation. (DM-AP-5)
- Prototype L1 / Alert Production database (relaxed scale requirements). (DM-DAX-6)

3.4 LDM-503-4: AuxTel DAQ integration functionality test

- Butler interface to retrieve images from data backbone. (DM-DAX-7)
- Minimal support for the small operational schema including file metadata and provenance for every file, and record of ingest. (DM-NCSA-4)

3.5 LDM-503-4b: Aux Tel DAQ Interface Integration Verification and Spectrograph Operations Rehearsal

- Level 1 archiving system able to acquire pixel data from the Aux Tel DAQ, header metadata via OCS, assemble FITS image, deposit in cache. (DM-NCSA-5)
- Ability to transfer files originating from Tucson to NCSA and ingest files at NCSA, including metadata and provenance. (DM-NCSA-6)
- Capability to paint displays for Tucson and NCSA. (DM-NCSA-7)
- Deliver header service code. (DM-NCSA-27)

3.6 LDM-503-5: Alert Distribution Validation

- Alert format defined & queue system available. (DM-AP-6)
- Test instance of feeds to LSST mini broker in online (live stream) and offline (replaying from files) modes. (DM-NCSA-8)
- Test instance of alert distribution hosting service and L1 database in Development & Integration Enclave. (DM-NCSA-9)

3.7 LDM-503-6: Limited ComCam Interface Verificaton using Spectrograph

- Sustained archiving service that is Observatory Control System (OCS) commandable. (DM-NCSA-10)
- Verified acquisition of raw and crosstalk-corrected exposures at raft scale, incl. correct metadata. (DM-NCSA-11)
- Interface Verification: Single Visit (LSST-1200)

3.8 LDM-503-7: Camera Data Processing

- Camera package supporting the LSST Camera. (DM-DRP-5)
- LSSTCam data display and visualization. (DM-SUIT-6)
- Mapping between SUIT systems & NCSA auth system. (DM-SUIT-7)
- SUIT portal integrated with workspace. (DM-SUIT-8)
- Basic instrument signature removal (ISR) capability. (DM-AP-7)
- Camera package supporting the Commissioning Camera. (DM-DRP-38)
- Calibration product generation in support of basic ISR. (DM-DRP-4)

3.9 LDM-503-8: Spectrograph Data Acquisition

- Basic instrument signature removal (ISR) capability. (DM-AP-7)
- Camera package supporting the Auxiliary Telescope. (DM-DRP-6)
- Calibration product generation for the Auxiliary Telescope. (DM-DRP-8)
- Data reduction pipeline for the Auxiliary Telescope. (DM-DRP-9)
- EFD ETL Service. (DM-NCSA-12)
- Spectrograph Archiving Service. (DM-NCSA-13)
- Data Backbone endpoints in Chile for ingestion and access, distribution over WAN, ingest at NCSA into custodial file store. (DM-NCSA-14)

- Batch Processing Service for offline spectrograph data processing. (DM-NCSA-15)
- Mountain - Base Network Functional 2 x 100 Gbps. (DM-NET-2)
- Initial Network Ready (Summit). (DM-NET-3)
- Base LAN installed. (DM-NET-4)
- Spectral data display. (DM-SUIT-9)

3.10 LDM-503-9: Verification Tests in Advance of Pre-Ops Rehearsal for Commissioning

- Advanced ISR, including ghost and linear feature masking, correction for the Brighter-Fatter effect and compensation for pixel response non-uniformity. (DM-AP-8)
- Jointcal at a functional (but not necessarily algorithmically complete) level. (DM-AP-9)
- Calibration products include an optical ghost model. (DM-DRP-10)
- Pipelines code provides supports for database ingestion of results. (DM-DRP-11)
- Background estimation over the full visit. (DM-DRP-12)
- PSF estimation over the full visit. (DM-DRP-13)
- Insertion of simulated sources into the data stream to check pipeline performance. (DM-DRP-14)
- All varieties of coadd required for object detection and characterization are now produced during normal pipeline operation (although not necessarily at the ultimately required level of fidelity). (DM-DRP-15)
- Global photometric fitting (e.g. Burke et al. Forward Global Calibration Method). (DM-DRP-16)
- Simultaneous photometric and astrometric fitting to multiple exposures. (DM-DRP-17)
- Initial multi-band deblending algorithm available. (DM-DRP-18)
- QA metrics are generated during pipeline execution. (DM-DRP-19)
- Perform ISR processing on ComCam-scale data. (DM-NCSA-16)

- QA on WCS, PSF, etc returned to Observatory using JupyterLab. (DM-NCSA-17)
- Validated disaster repsonse recovery for data and calibration products. (DM-NCSA-18)
- 8x7 incident response system available. (DM-NCSA-19)
- Supertask-based system capable of efficient processing across a full focal plane. (DM-DAX-8)
- Provenance system, details TBD. (DM-DAX-9)
- Commissioning notebooks running on the commissioning cluster. (DM-SQRE-2)
- Artefact rejection and background matching during coadd construction. (DM-DRP-37)
- Coordinate transformation tool provided for use with the Collimated Beam Projector. (DM-DRP-7)

3.11 LDM-503-11a: ComCam Operations Readiness.

- Advanced single frame measurement pipeline for Alert Production. (DM-AP-10)
- Alard & Lupton-style image differencing. (DM-AP-2)
- Point source & dipole measurement on difference images. (DM-AP-3)
- Alert format defined & queue system available. (DM-AP-6)
- Base - Archive Network Functional 100 Gbps. (DM-NET-5)
- ComCam Archiving Service. (DM-NCSA-20)
- L1 Offline Processing Service, Raft Scale, ComCam. (DM-NCSA-21)
- Information in consolidated database available to QA portal. (DM-NCSA-22)
- SUIT deployment procedure. (DM-SUIT-10)

3.12 LDM-503-11: Verification Tests in Advance of Pre-Ops Rehearsal for Commissioning #2

- Integrated image characterization pipeline for Data Release Production. (DM-DRP-21)

- Template generation integrated with Data Release Production pipelines. (DM-DRP-22)
- Atmospheric characterization based on data from the Auxiliary Telescope now available. (DM-DRP-23)
- Overlap resolution at tract & patch boundaries. (DM-DRP-26)
- Middleware support for MultiFit. (DM-DAX-10)
- ComCam data search, display, and visualization. (DM-SUIT-11)
- Integrated SUIT portal with the notebook aspect of the Science Platform. (DM-SUIT-12)
- Alert subscription (beta). (DM-SUIT-13)
- ”All-sky” visualization, including the single focal plane image visualization. (DM-SUIT-14)
- Difference imaging includes noise decorrelation and correction for differential chromatic refraction. (DM-AP-11)

3.13 LDM-503-12: Verification Tests in Advance of Pre-Ops Rehearsal for Commissioning #3

- Object generation: association and assembly of (DIA, coadd, etc) sources to form objects. (DM-DRP-27)
- Difference images are now a first-class data product during data release processing. (DM-DRP-28)
- Moving point source model fitting now available. (DM-DRP-29)
- Forced photometry is now performed on individual processed visit images during data releases. (DM-DRP-30)
- Hardened Jupyter deployment on Commissioning Cluster. (DM-SQRE-3)
- Physically motivated PSF model, including separate characterization of contributions from the atmosphere and the telescope system. (DM-DRP-24)
- Refined set of LSST calibration products. (DM-DRP-20)

3.14 LDM-503-13: Ops Rehearsal for Data Release Processing with ComCam Data

- Difference imaging is now agnostic to the PSF of the template image. (DM-AP-12)
- Trailed source measurement on difference images. (DM-AP-13)
- Alert filtering system available. (DM-AP-14)
- Object classification system available. (DM-DRP-32)
- Generation of coadded images suitable for use in EPO activities. (DM-DRP-33)
- Selection maps are generated during data releases. (DM-DRP-34)
- Simultaneous measurement across a suite of coadds representing different bandpasses, epocs, and flavors. (DM-DRP-35)
- Operational processes for preparing for & producing a data release defined and tested. (DM-NCSA-23)
- Alert subscription. (DM-SUIT-15)
- Commissioning DAC. (DM-SUIT-16)
- Prototype multi-epoch fitting system available. (DM-DRP-25)

3.15 LDM-503-14: DM Software for Science Verification

- Alert distribution system fully integrated. (DM-AP-15)
- Commissioning DAC. (DM-SUIT-16)
- Notebook service ready for verification & validation. (DM-SQRE-4)
- Generation of coadded images suitable for use in EPO activities. (DM-DRP-33)

3.16 LDM-503-15: Verification tests in advance of full scale Ops Rehearsal

- Full integration of the Alert Production system within the operational environment. (DM-AP-16)

- Moving object processing system (MOPS) available. (DM-AP-17)
- MOPS integrated to data release processing. (DM-DRP-36)
- Production batch service for data release processing. (DM-NCSA-24)
- Demonstrate operational coordination with and processing at CC-IN2P3 satellite computing facility. (DM-NCSA-25)
- Notebook service ready for general science users. (DM-SQRE-5)
- A photometric redshift is now provided for each object. (DM-DRP-31)

3.17 LDM-503-16: Ops rehearsal DRP

- Demonstrate full delivery of Data Facility ConOps. (DM-NCSA-26)

4 References

[1] **[LDM-148]**, Lim, K.T., Bosch, J., Dubois-Felsmann, G., et al., 2017, *Data Management System Design*, LDM-148, URL `https://ls.st/LDM-148`

[2] **[LDM-503]**, O'Mullane, W., Jurić, M., Economou, F., 2017, *Data Management Test Plan*, LDM-503, URL `https://ls.st/LDM-503`

5 Acronyms

Acronym	Description
DAQ	Data AcQuisition (system)
DM	Data Management
DRP	Data Release Production
LSST	Large Synoptic Survey Telescope
NCSA	National Center for Supercomputing Applications

PDAC	Prototype Data Access Center
SQuaRE	Science Quality and Reliability Engineering
WISE	Wide-field Survey Explorer