



LARGE SYNOPTIC SURVEY TELESCOPE

Large Synoptic Survey Telescope (LSST) Data Management Organization and Management

William O'Mullane, John Swinbank, Mario Juric and DMLT

LDM-294

Latest Revision: 2017-06-23

Draft Revision NOT YET Approved - This LSST document has been approved as a Content-Controlled Document by the LSST DM Technical Control Team. If this document is changed or superseded, the new document will retain the Handle designation shown above. The control is on the most recent digital document with this Handle in the LSST digital archive and not printed versions. Additional information may be found in the corresponding DM RFC. - **Draft Revision NOT YET Approved**

Abstract

This is the DM plan updated from the v2 of 2015. It covers the organisation and management of DM for LSST.

Change Record

Version	Date	Description	Owner name
1	2004-06-23	Initial version	JK
2	2015-03-11	Updated with new RFC process, realignment of TCT, SAT, DMLT - other versions in between	JK
3	2017-06-29	Update in TeX and complete overhaul	WOM,MJ

Draft

Contents

1	Introduction	1
1.1	Purpose	1
1.2	Mission statement	1
1.3	Goals And Objectives	1
2	Data Management Conceptual Architecture	3
2.1	External Interfaces	3
2.1.1	Auxiliary data in DM	7
3	Data Management Organization Structure	7
3.1	Meetings	8
3.2	Document Management	9
3.2.1	Documentation of Cross-Cutting Aspects for services	11
3.2.2	Documentation of Service Layers	12
3.2.3	Draft Documents	13
3.3	Configuration Control	13
3.3.1	Software Configuration Control	13
3.3.2	Hardware Configuration Control	14
3.4	Risk Management	14
3.5	Quality Assurance	14
3.6	Action item control	15
3.7	Verification and Validation	15



4 Project Controls 16

4.1 Schedule 16

4.2 Work Breakdown Structure 17

5 Products 17

6 Roles in Data Management 18

6.1 DM Project Manager (DMPM) 18

6.2 DM Deputy Project Manager (DDMPM) 20

6.3 DM Subsystem Scientist (DMSS) 20

6.4 Project Controller/Scheduler 22

6.5 Product Owner 22

6.6 Pipeline Scientist 22

6.7 Systems Engineer 23

6.8 DM Interface Scientist (DMIS) 24

6.9 Software Architect 24

6.10 Operations Architect 24

6.11 Configuration Manager (CM) 25

 6.11.1 Configuration Item List 26

 6.11.2 Release management 26

 6.11.3 Configuration Baseline 27

6.12 Lead Institution Senior Positions 27

 6.12.1 Technical/Control Account Manager (T/CAM) 27

6.12.2	Institutional Science/Engineering Lead	28
7	Data Management Groups/Bodies	28
7.1	System Science Team	29
7.1.1	Organization and Goals	29
7.1.2	Communications	30
7.2	DM Systems Engineering Team	30
7.2.1	Communications	31
7.3	DM Leadership Team	32
7.3.1	Communications	32
7.4	DM Change Control Board	33
8	Lead institutions in DM	34
8.1	LSST Tucson	34
8.1.1	Science Quality and Reliability Engineering	34
8.1.2	LSST International Communications and Base Site	35
8.2	Princeton University	36
8.2.1	Data Release Production	36
8.3	The University of Washington	37
8.3.1	Alert Production	37
8.4	California Institute of Technology/IPAC	37
8.4.1	Science User Interface and Tools	37
8.5	SLAC	38



8.5.1 Science Data Archive and Data Access Services 38

8.6 NCSA 39

9 Development Process 41

9.1 Communications 41

9.2 Conventions 41

9.3 Reviews 41

10 Data Management Problem/Conflict Resolution 42

A DM Product List 42

B Proposed FY2018 WBS: Data Management Construction 47

B.1 System Management 52

 B.1.1 Project Management 53

B.2 Data Management Science 53

 B.2.1 1.02C.02.02: DM System Architecture 53

B.3 1.02C.03: Alert Production 55

 B.3.1 1.02C.03.00: Management & Leadership 55

 B.3.2 1.02C.03.01: Single Frame Processing 55

 B.3.3 1.02C.03.02: Catalog Association for Alert Production 56

 B.3.4 1.02C.03.03: Alert Distribution System 56

 B.3.5 1.02C.03.04: Alert Generation Pipeline 57

 B.3.6 1.02C.03.05: Tools for Science Pipelines 58



B.3.7 1.02C.03.06: Moving Object Processing System (MOPS) 59

B.3.8 1.02C.03.07: Transform fitting on stacks of images 59

B.3.9 1.02C.03.08: Integration 60

B.4 1.02C.04: Data Release Production 60

 B.4.1 1.02C.04.00: Management & Leadership 61

 B.4.2 1.02C.04.01: Software Primitives 61

 B.4.3 1.02C.04.02: Calibration Products 62

 B.4.4 1.02C.04.03: Image Characterization 62

 B.4.5 1.02C.04.04: Coaddition 63

 B.4.6 1.02C.04.05: Detection & Deblending 63

 B.4.7 1.02C.04.06: Characterization & Measurement 64

 B.4.8 1.02C.04.07: Maintenance & Documentation 65

B.5 1.02C.05: Science User Interface and Tools 65

 B.5.1 1.02C.05.00: Management, Leadership, & Other LOE 65

 B.5.2 1.02C.05.06: Client-server Query & Visualization Framework 65

 B.5.3 1.02C.05.07: LSST Science Platform Interfaces 66

 B.5.4 1.02C.05.08: Applications 67

B.6 1.02C.06: Science Data Archive and Application Services 68

 B.6.1 1.02C.05.09: Alert Interfaces 68

 B.6.2 1.02C.05.10: Integration & Test 69

 B.6.3 1.02C.06.00: Management, Leadership, and Other LOE 69

 B.6.4 1.02C.06.01: Science Data Archive 69

B.6.5	1.02C.06.02: Data Access Services	71
B.6.6	1.02C.06.03: Task Framework	73
B.6.7	1.02C.06.04: Middleware, Infrastructure, and Toolkits	74
C	DM Discussion and Decision Making Process	75
C.1	Empowerment	75
C.2	RFC Process	76
C.3	Exceptions and Appeals	77
C.4	Formalities	77
D	Traceability matrix of DMSR requirements to OSS Requirements	79
E	Traceability matrix of OSS requirements to DMSR requirements	84
F	References	84
G	Acronyms	87

Data Management Organization and Management

1 Introduction

1.1 Purpose

This document defines the mission, goals and objectives, organization and responsibilities of the LSST Data Management (DM). The document is currently scoped to define these elements for the LSST Design and Development, Construction, and Commissioning phases. It does not address any ongoing mission for the DM during LSST operations.

1.2 Mission statement

Stand up operable, maintainable, quality services to deliver high-quality LSST data products for science and education, all on time and within reasonable cost.

1.3 Goals And Objectives

LSST Data Management will:

- Define the data products, data access mechanisms, and data management and curation requirements for the LSST
- Assess current and LSST-time frame technologies for use in providing engineered solutions to the requirements
- Define a secure computing, communications, and storage infrastructure and services architecture underlying LSST data management
- Select, implement, construct, test, document, and deploy the LSST data management infrastructure, middleware, applications, and external interfaces
- Adopt appropriate cybersecurity measures throughout data management and especially on external facing services.
- Document the operational procedures associated with using and maintaining the LSST data management capabilities

- Evaluate, select, recruit, hire/contract and direct permanent staff, contract, and in-kind resources in LSST and from partner organizations participating in LSST Data Management initiatives.

The DM goals in selecting and, where necessary, developing LSST software solutions are:

- Acquire and/or develop solutions: To achieve its mission, LSST DM subsystem prefers to acquire and configure existing, off-the-shelf, solutions. Where no satisfactory off-the-shelf solutions are available, DM develops the software and hardware systems necessary to:
 - Enable the generation of LSST data products at the LSST Archive and Satellite processing center, and
 - Enable the the serving of LSST data products from the two LSST DACs (one in the U.S., and one in Chile).
- Maintain coherent architecture: DM software architecture is actively managed at the subsystem level. A well engineered, and cleanly designed codebase is less buggy, more maintainable, and makes developers who work on it more productive. Where there is no significant impact on capabilities, budget, or schedule, LSST DM prefers to acquire and/or develop reusable, open source, solutions.
- Support reproducibility and insight into algorithms: Other than when prohibited by licensing, security, or other similar considerations, DM makes all newly developed source code public, especially the Science Pipelines code. Our primary goal in publicizing the code is to simplify reproducibility of LSST data products, and provide insight into algorithms used. The software is to be documented to achieve those goals.
- Opportunities beyond LSST: LSST DM codes may be of interest and (re)used beyond the LSST project (e.g., by other survey projects, or individual LSST end-users). While enabling or supporting such applications goes beyond LSST's construction requirements, cost and schedule-neutral technical and programmatic options that do not preclude them and allow for future generalization should be strongly preferred.

Background decision material on choices made in DM will be documented in technical notes (DMTN) which will be lodged in DocuShare (see Section 3.2).

2 Data Management Conceptual Architecture

The DM Subsystem Architecture is detailed in LDM-148. A few of the higher level diagrams are reproduced here to orientate the reader within DM.

During Operations, components of the DM Subsystem will be installed and run in multiple locations. These include:

- The Commissioning Cluster, which may be physically at NCSA in Urbana-Champaign
- The main center in NCSA enclave in Urbana-Champaign
- The US Data Access Center (DAC), also at NCSA in Urbana-Champaign
- The Chilean DAC in the Base Facility in La Serena Chile
- The Satellite Processing Center at CC-IN2P3 in Lyon, France

Figure 1 shows the various DM components which will be used in operations and the physical compute environments in which they will be deployed. Bulk data storage and transport between components is provided by the Data Backbone. This complex piece of infrastructure is displayed in Figure 3.

Science users will access the data products produced by LSST through the Science Platform, as shown in Figure 2.

Figure 4 shows the common infrastructure and services layer which underlies the compute environments. This does not list specific technologies for management/monitoring, provisioning/deployment, or workload/workflow — these are still under development — but consider industry-standard tools such as Nagios, Puppet/vSphere/OpenStack/Kubernetes, and Pegasus.

2.1 External Interfaces

The DM external interfaces are controlled by the ICDs listed in Table 2.

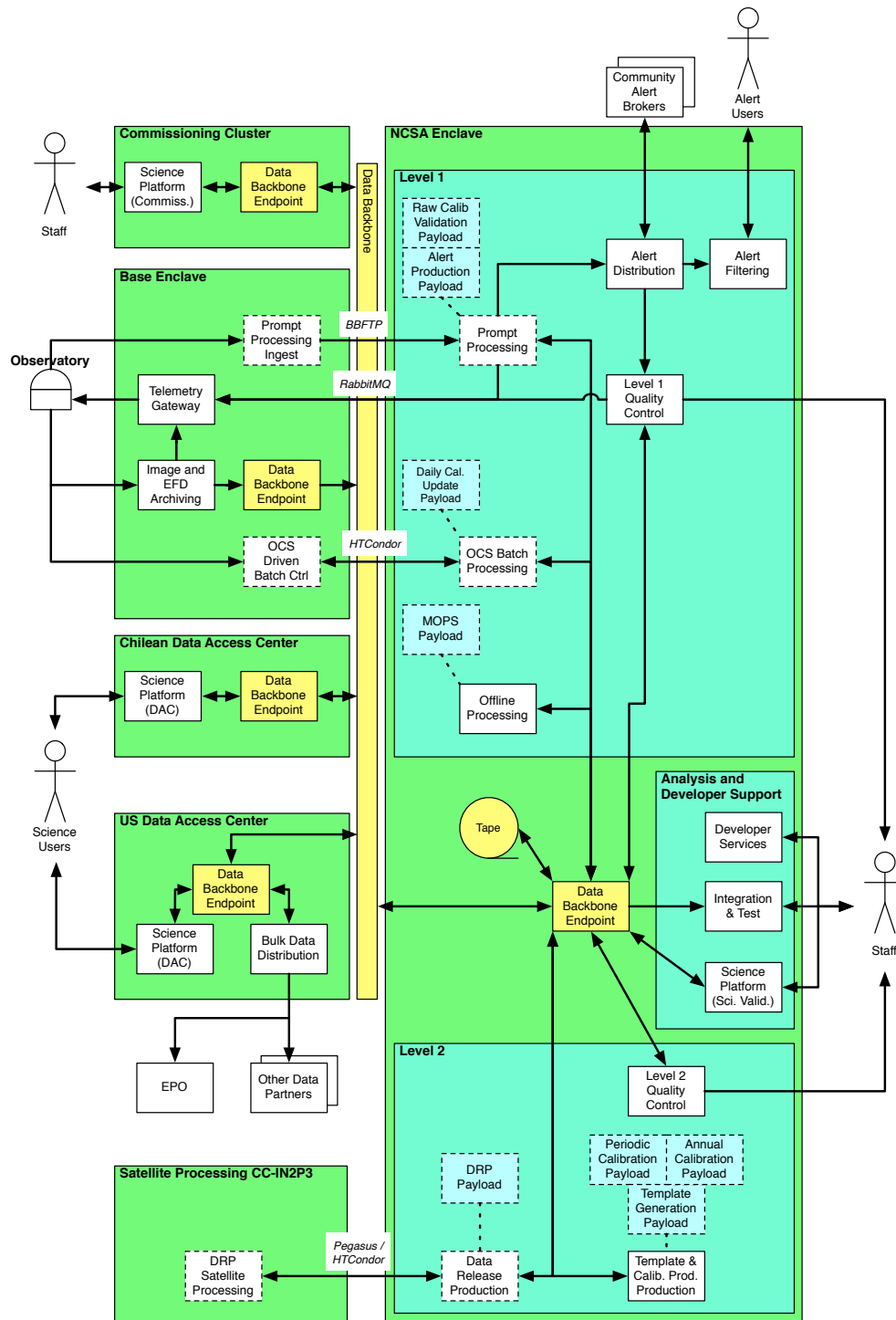


FIGURE 1: DM components as deployed during Operations. Where components are deployed in multiple locations, the connections between them are labeled with the relevant communication protocols. Science payloads are shown in blue.

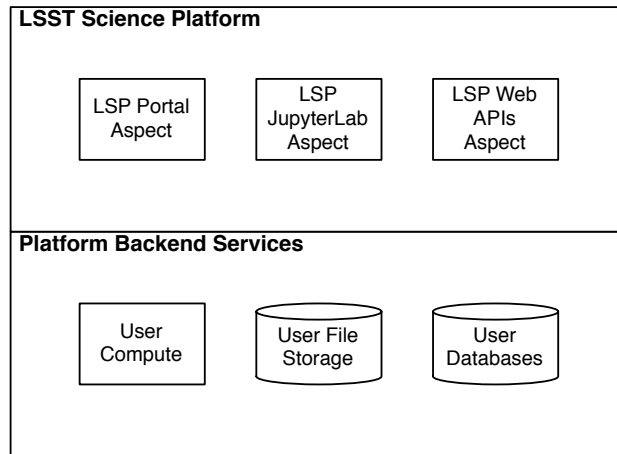


FIGURE 2: The sub-components of the Science Platform.

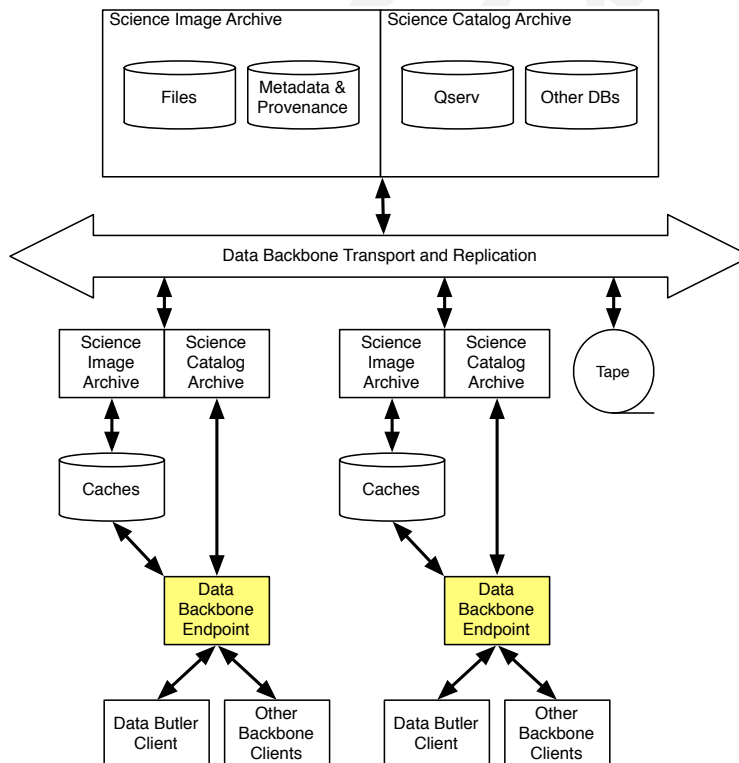


FIGURE 3: The Data Backbone links all the physical components of DM.

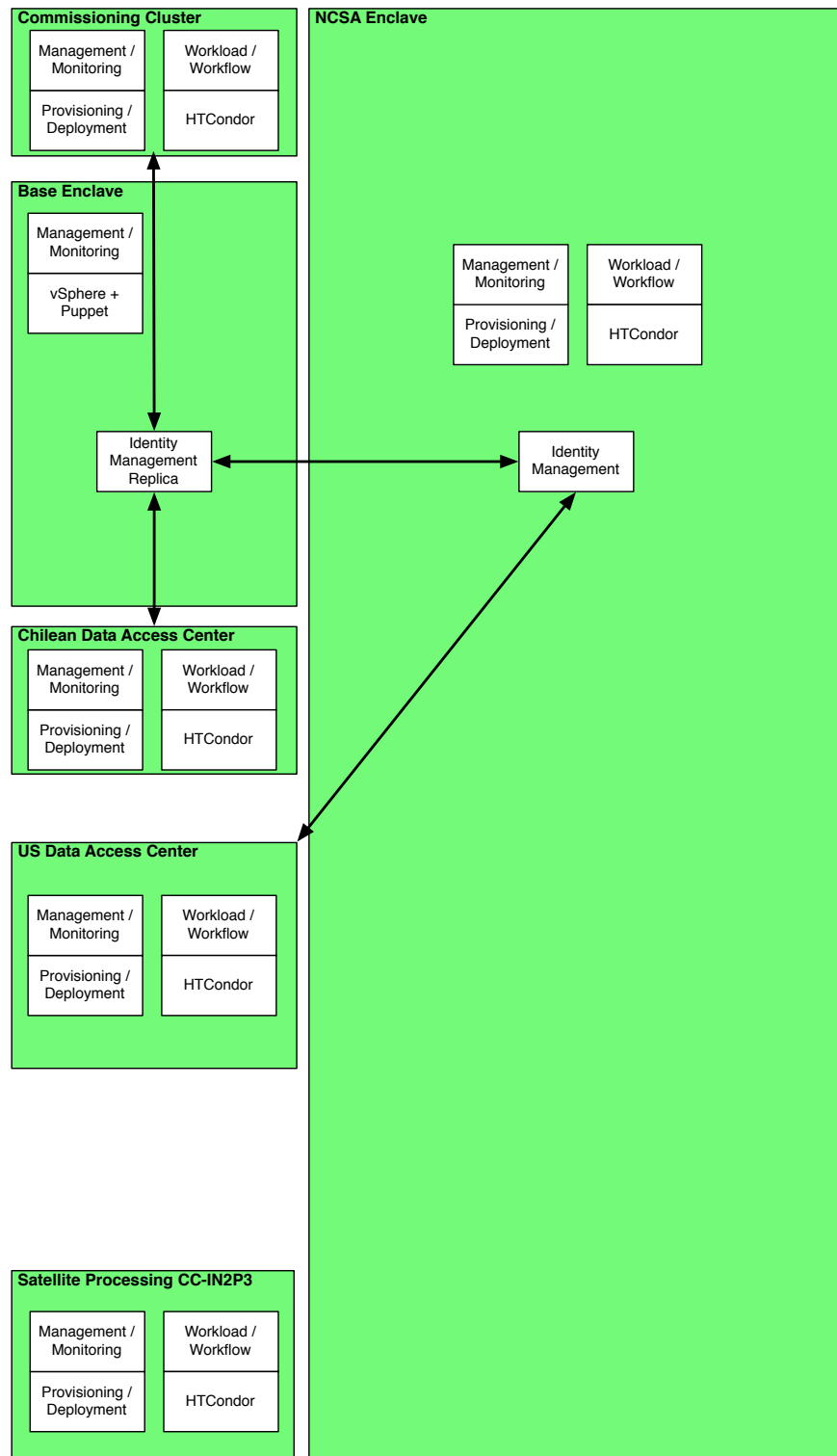


FIGURE 4: Common infrastructure services available at each DM location.

TABLE 2: DM Interface Control Documents

LSE-68	Data Acquisition Interface between Data Management and Camera
LSE-69	Interface between the Camera and Data Management
LSE-72	OCS Command Dictionary for Data Management
LSE-75	Control System Interfaces between the Telescope and Data Management
LSE-76	Infrastructure Interfaces between Summit Facility and Data Management
LSE-77	Infrastructure Interfaces between Base Facility and Data Management
LSE-130	List of Data Items to be Exchanged Between the Camera and Data Management
LSE-131	Data Management Interface Requirements to Support Education and Public Outreach
LSE-140	Auxiliary Instrumentation Interface between Data Management and Telescope

2.1.1 Auxiliary data in DM

Certain tasks in DM rely on external catalogs and other information. Currently we believe we need:

1. Gaia catalogue (Release 2) as a photometry baseline.

3 Data Management Organization Structure

This section defines the organization structure for the period in which the DM System is developed and commissioned, up to the start of LSST Observatory operations.

The DM Project Manager (William O'Mullane), Deputy Project Manager (John Swinbank) and DM Project Scientist (Mario Jurić), who are known collectively as DM Management, lead the DM Subsystem. The Project Manager has direct responsibility for coordination with the overall LSST Project Office, the LSST Change Control Board, the LSST Corporation, and LSST partner organizations on all budgetary, schedule, and resource matters. The Project Scientist has primary scientific and technical responsibility in the DM and responsibility for ensuring that the scientific requirements of the LSST are supported, and is a member on the LSST Project

Science Team (PST).

As shown in Figure 5, the organization now features major products each with a product owner relating to a major element of the DM Subsystem (Level 2 Work Breakdown Structure elements).

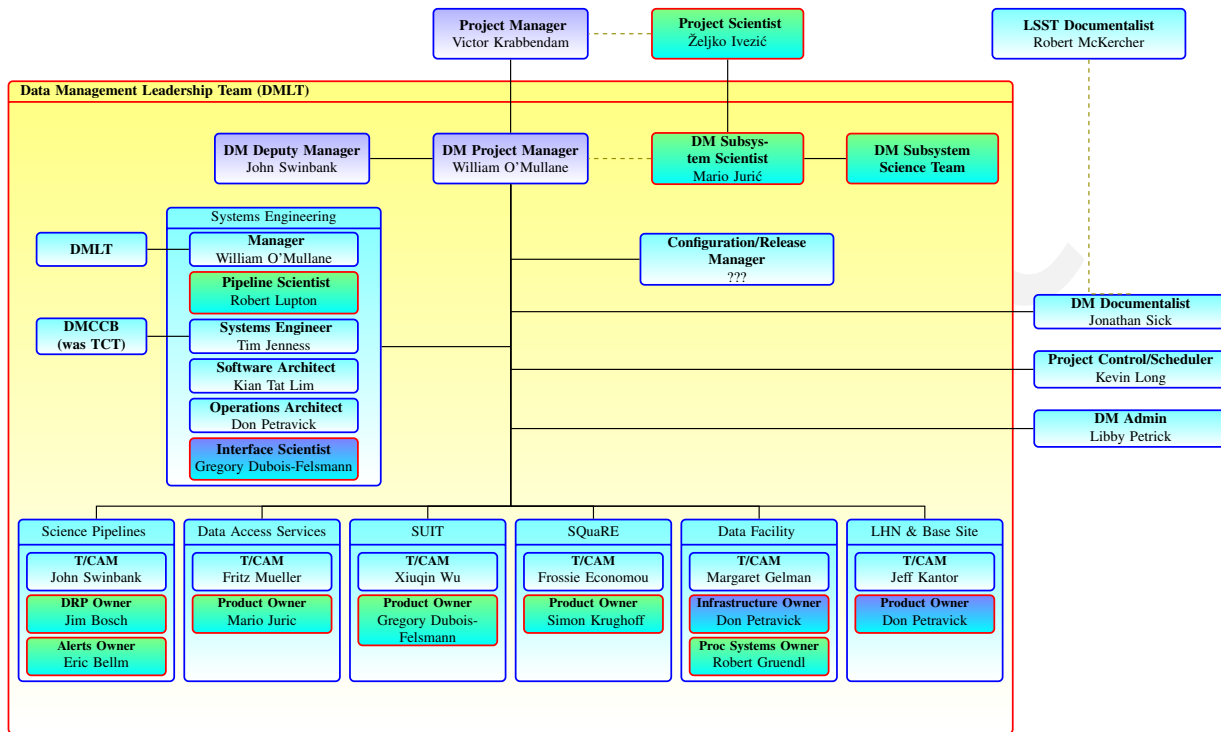


FIGURE 5: DM organization with Scientists in Green.

3.1 Meetings

As a diverse and distributed organization DM staff will participate in a considerable number of meetings. NSF and Aura have many rules on meeting attendance and LSST keep policies updated accordingly on the project travel web page <https://project.lsst.org/travel/policies-procedures>. This include the travel summary report template <https://project.lsst.org/sites/default/files/Travel/Travel%20Summary%20Report.pdf> every traveler must fill after attending a meeting. More details are in Document-13760.

A detailed debrief note or presentation may be asked of travelers to specific meetings of interest by the DMLT.

3.2 Document Management

DM documents will follow the System Engineering Guidelines of LSST. PDF versions of released documents shall be put in DocuShare in accordance with the Project’s Document Management Plan [LPM-51]. LPM level documents are released on agreement of the DMCCB (Section 7.4), uncontrolled documents such as technotes may be released when the author decides it is appropriate or they are asked to release it by the Project Manager.

The Document Tree for DM is shown in Figure 6, it is not exhaustive but gives a high level orientation for the main documents in DM and how they relate to each other. Some documents shown in red are not yet written.

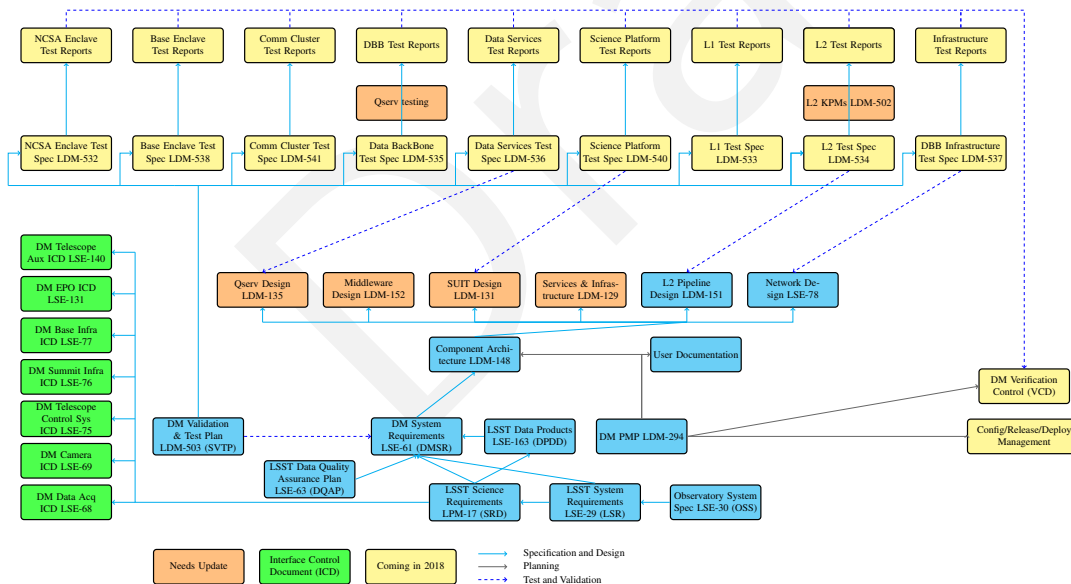


FIGURE 6: Outline of the documentation tree for DM software relating the high level documents to each other.

Figure 6 has one box for End User documentation, this is a major set of documentation for DM which will be web based as described in LDM-493. Figure 7 shows the intended web hierarchy for the end user documentation.

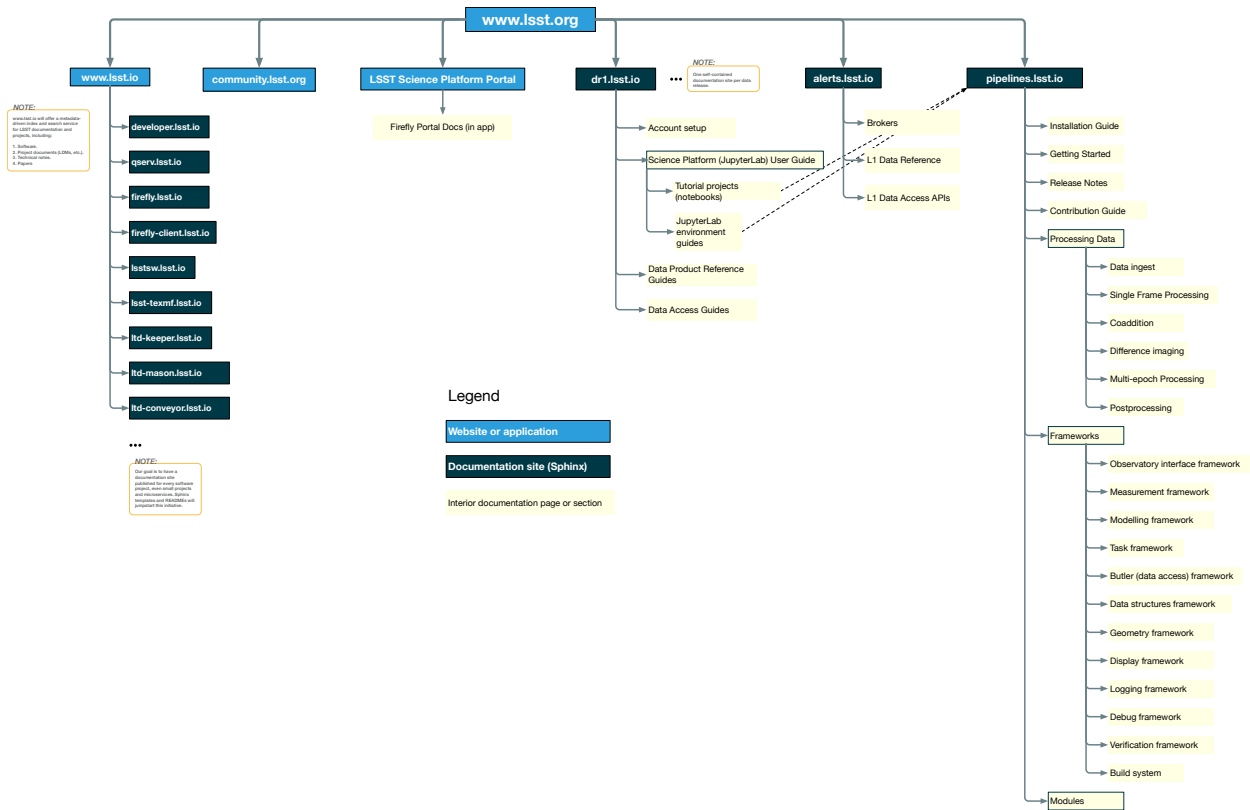


FIGURE 7: Outline of the web hierarchy for the DM end user documentation.

Service-level documentation follows the layered service architecture of the LSST Data Facility (see Figure 8).

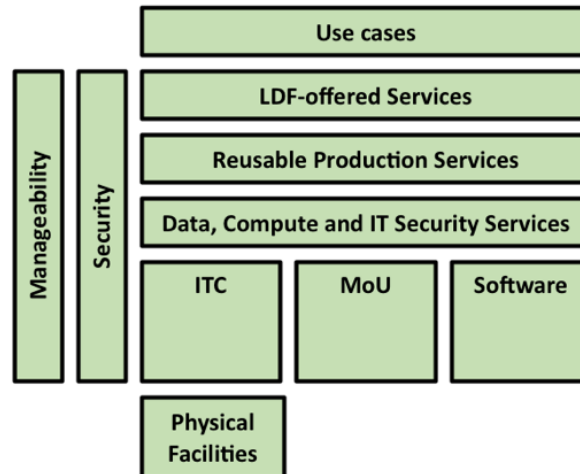


FIGURE 8: Outline of layered service architecture of the Data Facility.

3.2.1 Documentation of Cross-Cutting Aspects for services

The cross-cutting aspects of the LSST Data Facility, Security and Operational Manageability, are represented by the vertical boxes. Documentation of these aspects describes policies, procedures, and supporting management frameworks, including:

1. LDF service management framework: service catalog, service-level agreements (SLAs), configuration management database (CMDB), service monitoring.
2. LDF service management processes and context in the overall project: incident response, request response, issue tracking, problem management and the problem management database, change management and change control authority, release management.
3. Overview of the security enclave structure
4. Security controls and incident response procedures
5. Disaster recovery and continuity policies

3.2.2 Documentation of Service Layers

The box at the top of the figure, Use Cases, represents subsystem-level and project-level operational use cases. The next layer, LDF-offered Services, represents specific services offered by the Data Facility which satisfy those use cases. Documentation of this layer includes:

1. For each service, a Concept of Operations (ConOps) which summarizes how a service operates to satisfy a use case. The ConOps describes the operational characteristics of the production system, context within overall LSST operations, and representative scenarios.
2. For each service, a Theory of Operations, which provides a mental model of a constructed system. The Theory of Operations explains how the constructed service both fulfills the ConOps and integrates with the cross-cutting aspects of the facility. The document describes the overall architecture of the service and dependency on supporting service layers; integration into aspects of computer security, information security and business continuity; and integration into incident reporting and response, availability and capacity management, and change management.

The next two layers, Reusable Production Services and Data, Compute, and IT Security Services, represent tiers of supporting service. Documentation of these layers includes a Theory of Operations, as described above, explaining the dependencies on supporting service and ITC layers, and integration with cross-cutting aspects of the facility.

The ITC box represents hardware components supporting all LDF services. Documentation of ITC describes the system elements at all facility sites, administration within each security enclave and integration with security operations, the overall provisioning plan, ITC system monitoring and integration into the service monitoring framework, and integration into service management processes including configuration management and change management.

The Software box represents service software components being developed by the LSST Data Facility. Documentation of software elements follows the standards of the LSST software stack.

Documents are managed as configuration items in the LSST Data Facility CMDB.

3.2.3 Draft Documents

Draft DM documents will be kept in GitHub. A single repository per document will be maintained with the head revision containing the *released* version which should match the version on docushare. Each repository will be included as a *submodule* of a single git repository located at <https://github.com/lstt-dm/dm-docs>.

Use of Google Docs or confluence is tolerated but final delivered documents must conform to the standard LSST format, and hence either produced with LaTeX, using the `lsst-texmf` package¹, or Word, using the appropriate LSST template [Document-9224, Document-11920]. The precursor document should then be erased with a pointer to the baseline document, stored in GitHub.

End user documentation will most likely and appropriately be web based and the scheme for that is described in LDM-493.

3.3 Configuration Control

Configuration control of documents is dealt with in Section 3.2. Here we consider more the operational systems and software configuration control.

3.3.1 Software Configuration Control

DM follows a git based versioning system based on public git repositories and the approach is covered in the developer guide <https://developer.lsst.io/>. The master branch is the stable code with development done in *ticket* branches (named with the id of the corresponding Jira Ticket describing the work. Once reviewed a branch is merged to master.²

As we approach commissioning and operations DM will have a much stricter configuration control. At this point there will be a version of the software which may need urgent patching, a next candidate release version of the software, and the master. A patch to the operational version will require the same fix to be made in the two other versions. The role of the DM Change Control Board (DMCCB; Section 7.4) becomes very important at this point to ensure

¹<https://lsst-texmf.lsst.io>

²LSE-14 seem out of date and should be updated or revoked - titled a guideline it seems inappropriate as an LSE.

only essential fixes make it to the live system as patches and that required features are included in planned releases.

We cannot escape the fact that we will have multiple code branches to maintain in operations which will lead to an increase in work load. Hence one should consider that perhaps more manpower may be needed in commissioning to cope with urgent software fixes while continuing development. The other consideration would be that features to be developed post commissioning will probably be delayed more than one may think, as maintenance will take priority.³

3.3.2 Hardware Configuration Control

On the hardware side we have multiple configurable items, we need to control which versions of software are on which machines. These days tooling like Puppet make this reasonably painless. Still the configuration must be carefully controlled to ensure reproducible deployments providing correct and reproducible results. The exact set of released software and other tools on each system should be held in a configuration item list. Changes to the configuration should be endorsed by the DMCCB.

The sizing model for compute hardware purchasing is detailed in LDM-144 LDM-141 and LDM-138.

3.4 Risk Management

Risks will be dealt with within the LSST Project framework as defined in LPM-20. Risks in DM may be sent to the DM project manger or Deputy project manager at any time for consideration to be included in the formal risk register (appropriate costed and weighted). All risks are reviewed regularly by the DM Project manager and System engineer (minimum each 3 months).

3.5 Quality Assurance

In accordance with the project QA plan LPM-55 we will perform QA on the software products. This work will mainly be carried out by SQuaRE (Section 8.1.1). Quality assurance here means

³WOM identifies this as the maintenance surge.

compliance with project guidelines for production, in out case of software production. A part of this is to have a verification/validation plan(s) which in and of itself is a major task (see Section 3.7).

3.6 Action item control

Actions in DM are tracked as Jira issues an periodically reviewed at DMLT meetings.

3.7 Verification and Validation

We intend to verify and validate as much of DM as we can before commissioning and operations. This will be achieved through testing and operations rehearsals/data challenges. The verification and validation approach is detailed in LDM-503 including a high level test schedule, the top level schedule is given in Figure 9.

4 Project Controls

DM follows the LSST project controls system, as described in LPM-98. Specific DM processes for project planning are elaborated in LDM-472.

The LSST Project Controller is responsible for the PMCS and, in particular, for ensuring that DM properly complies with our earned value management requirements. He is the first point of contact for all questions about the PMCS system.

4.1 Schedule

The entire LSST project schedule is held in Primavera. Tied to major project milestones we have a series of DM tests which need to be performed to show readiness for the different project phases. This is depicted in Figure 9.

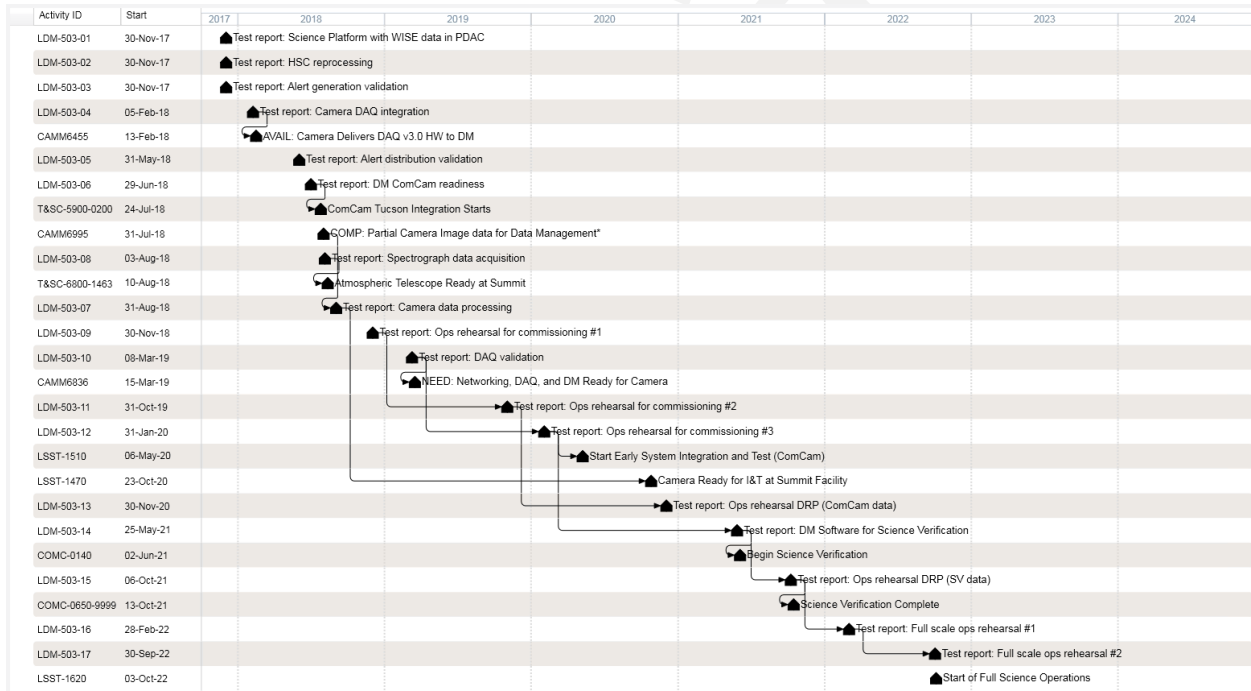


FIGURE 9: DM major milestones(LDM-503-x) in the LSST schedule.

TABLE 3: DM top level Work Breakdown Structure

WBS	Description	Lead Institution
1.02C.01	System Management	LSST Tucson
1.02C.02	Systems Engineering	LSST Tucson
1.02C.03	Alert Production	University of Washington
1.02C.04	Data Release Production	Princeton University
1.02C.05	Science User Interface	Caltech IPAC
1.02C.06	Science Data Archive	SLAC
1.02C.07	Processing Control & Site Infrastructure	NCSA
1.02C.08	International Communications. & Base Site	LSST Tucson
1.02C.09	Systems Integration & Test	NCSA & LSST Tucson
1.02C.10	Science Quality & Reliability Engineering	LSST Tucson

4.2 Work Breakdown Structure

The DM WBS is laid out in LPM-43 with definitions provided in LPM-44, the new WBS is currently described in Appendix B making LPM-43 out of date.

The WBS provides a hierarchical index of all hardware, software, services, and other deliverables which are required to complete the LSST Project. It consists of alphanumeric strings separated by periods. The first component is always "1", referring the LSST Construction Project. "02C" in the second component corresponds to Data Management Construction. Subdivisions thereof are indicated by further digits. These subdivisions correspond to teams within the DM project. The top level WBS elements are mapped to the lead institutes in Table 3, the lead institutions roles are outlined in Section 8. The various groups involved in the WBS are briefly described in Section 7.

5 Products

The products of DM are not the data products defined in LSE-163, rather they are the artifacts, systems and services, we need to produce those products. Section 2 outlines the highest level of this for dm while Appendix A defines the complete product tree for DM and it is pictorially represented at a trimmed level in Figure 10. LDM-148 provides a trace of products to requirements, while Appendix A provides a full list with technical manager WBS element and product owner for each. Our primary guiding requirements come from LSE-30 and Appendix D and Appendix E provide the traceability to and from the OSS requirements to the DM require-

ments LSE-61.

Figure 10 contains the WBS element associated with the component as well as an git repositories belonging to them. Since the figure stops at level 3 most git repos will only be found in full list in Appendix A.

Every git repo should appear in Appendix A and hence have a technical manager and product owner identified. The table is hierarchical hence if the manager/owner is not filled in (or gone) we may go to the parent element manager/owner.

Every JIRA component should map to one row in Appendix A thus providing a contact for that component.

6 Roles in Data Management

This section describes the responsibilities associated with the roles shown in Figure 5.

6.1 DM Project Manager (DMPM)

The DM Project Manager is responsible for the efficient coordination of all LSST activities and responsibilities assigned to the Data Management Subsystem. The DM Project Manager has the responsibility of establishing the organization, resources, and work assignments to provide DM solutions. The DM Project Manager serves as the DM representative in the LSST Project Office and in that role is responsible for presenting DM initiative status and submitting new DM initiatives to be considered for approval. Ultimately, the DM Project Manager, in conjunction with his / her peer Project Managers (Telescope, Camera), is responsible for delivering an integrated LSST system. The DM Project Manager reports to the LSST Project Manager. Specific responsibilities include:

- Manage the overall DM System
- Define scope and request funding for DM System
- Develop and implement the DM project management and control process, including earned value management

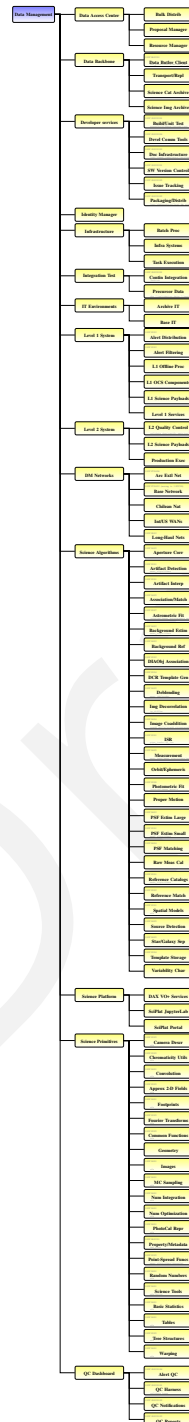


FIGURE 10: DM product tree. - there are over 200 products, this tree is to convey an idea of the products and is truncated to make it somewhat legible. The full list is given in Appendix A

- Approve the DM Work Breakdown Structure (WBS), budgets and resource estimates
- Approve or execute as appropriate all DM outsourcing contracts
- Convene and/or participate in all DM reviews
- Co-chair the DM Leadership Team (Section 7.3)

6.2 DM Deputy Project Manager (DDMPM)

The PM and deputy will work together on the general management of DM and any specific PM tasks may be delegated to the deputy as needed and agreed. In the absence of the PM the deputy carries full authority and decision making powers of the PM. The DM Project Manager will keep the Deputy Project Manager informed of all DM situations such that the deputy may effectively act in place of the Project Manager when absent.

6.3 DM Subsystem Scientist (DMSS)

The DM Subsystem Scientist (DMSS) has the ultimate responsibility for ensuring DM initiatives provide solutions that meet the overall LSST science goals. As such, they lead the definition and understanding of the science goals and deliverables of the LSST Data Management System, and are accountable for communicating these to the DM engineering team.

The DM Subsystem Scientist reports to the LSST Project Scientist. The DMSS are a member of the LSST Change Control Board and the Project Science Team. They chair and direct the work of the DM System Science Team (Section 7.1).

Specific responsibilities and authorities include (**cite the project-level R2A2 document, once issued**):

- Communicates with DM science stakeholders (LSST Project Scientist and Team, advisory bodies, the science community) to understand their needs and identifies aspects to be satisfied by the DM Subsystem.
- Develops, maintains, and articulates the vision of DM products and services responsive to stakeholder needs.

- Works with the LSST Project Scientist to communicate the DM System vision to DM stakeholders. Works with the DM Project Manager to communicate and articulate the DM System vision and requirements to the DM construction team.
- Regularly monitors DM construction team progress and provides feedback to the DM Project Manager to ensure the continual understanding of and adherence to the DM vision, requirements, and priorities.
- Develops and/or evaluates proposed changes to DM deliverables driven by schedule, budget, or other constraints.
- Provides advice to the DM Project Manager on science-driven prioritization of construction activities.
- Validates the science quality of DM deliverables and the capability of all elements of the DM System to achieve LSST science goals.
- Serves as Data Management Liaison as requested by LSST Science Collaborations
- Provides safe, effective, efficient operations in a respectful work environment.

Specific authorities include (**cite the project-level R2A2 document, once issued**):

- Defines the vision and high-level requirements of the DM products and services required to deliver on LSST science goals.
- Defines the science acceptance criteria for DM deliverables (both final and intermediate) and validates that they have been met (Science Validation).
- Hires or appoints DM System Science Team staff and other direct reports and defines their responsibilities.
- Advises and consents to the appointments of institutional DM Science Leads.
- Delegates authority and responsibility as appropriate to institutional Science Leads and other members of the DM System Science Team.
- Represents and speaks for the LSST Data Management.
- Convenes and/or participates in all DM reviews.
- Co-Chairs the DM Leadership Team

6.4 Project Controller/Scheduler

The DM Project Controller is responsible for integrating DM's agile planning process with the LSST Project Management and Control System (PMCS). Specific responsibilities include:

- Assist T/CAMs in developing the DM plan
- Synchronize the DM plan, managed as per LDM-472, with the LSST PMCS
- Ensure that the plan is kept up-to-date and milestones are properly tracked
- Create reports, Gantt charts and figures as requested by the DMPM

6.5 Product Owner

A product owner is responsible for the quality and acceptance of a particular product. The product owner should sign off on the requirements to be fulfilled in every delivery and therefore also on any descopes or enhancements. The product owner should define tests which can be run to prove a delivery meets the requirements due for that product.

6.6 Pipeline Scientist

Several DM products come together to form the LSST pipeline. The Pipeline Scientist is the product owner for the overall pipeline.

The Pipeline Scientist should:

- Provide guidance and test criteria for the full pipeline including how QA is done on the products
- Keep the big picture of where the codes are going in view. Predominantly the algorithms, but also the implementation and architecture (as part of the System Engineering Team Section 7.2).
- Advise on how we should attack algorithmic problems, providing continuing advice to subsystem product owners as we try new things.

- Advise on calibration issues, provide understanding of the detectors from a DM point of view
- Advise on the overall (scientific) performance of the system, and how we'll test it. Think about all the small things that we have to get right to make the overall system good.

6.7 Systems Engineer

With the System Engineering Team (Section 7.2) the Systems Engineer owns the DM entries in the risk register and is generally in charge of the *process* of building DM products.

As such, the Systems Engineer is responsible for managing requirements as they pertain to DM. This includes:

- Update and ensure traceability of the high level design & requirements documents: DMSR (LSE-61), OSS (LSE-30), and LSR (LSE-29)
- Oversee work on lower level requirements documents
- Ensure that the system is appropriately modeled in terms of e.g. drawings, design documentation, etc
- Ensure that solid verification plans and standards are established within DM

In addition, the Systems Engineer is responsible for the process to define & maintain DM interfaces (internal and external)

- Define and enforce standards for internal interfaces
- Direct the Interface Scientist's (Section 6.8) work on external ICDs

The Systems Engineer shall chair the DM Change Control Board (Section 7.4)

- Organize DMCCB processes so that the change control process runs smoothly

- Identify RFCs requiring DMCCB attention
- Shepherd RFCs through change control
- Call and chair DMCCB meetings, ensuring that decisions are made and recorded

Finally, the System Engineer represents DM on the LSST CCB.

6.8 DM Interface Scientist (DMIS)

The DM Interface Scientist is responsible for all internal and external interfaces to the DM Subsystem. This includes ensuring that appropriate tests for those interfaces are defined. This is a responsibility delegated from the DM System engineer (Section 6.7).

6.9 Software Architect

The Software Architect is responsible for the overall design of the DM *software* system. Specific responsibilities include:

- Define the overall architecture of the system and ensuring that all products integrate to form a coherent whole
- Select and advocate appropriate software engineering techniques
- Choose the technologies which are used within the codebase
- Minimize the exposure of DM to volatile external dependencies

The Software Architect will work closely with the Systems Engineer (Section 6.7) to ensure that processes are in place for tracing requirements to the codebase and providing hooks to ensure that requirement verification is possible.

6.10 Operations Architect

Margaret or Don perhaps some text here ...

The DM Operations Architect is responsible for ensuring that all elements of the DM Sub-system, including operations teams, infrastructure, middleware, applications, and interfaces, come together to form an operable system.

Specific responsibilities include:

- Set up and coordinate operations rehearsals
- Ensure readiness of procedures and personnel for Operations
- Set standards for operations e.g. procedure handling and operator logging
- Participate in stakeholder and end user coordination and approval processes and reviews
- Serve as a member of the LSST System Engineering Team

6.11 Configuration Manager (CM)

The DM Configuration Manager (CM) is responsible for Configuration Management activities inside DM and NCSA(?). The following list is not exhaustive, but is intended as a guideline to the CM activities:

- Ensure that Configuration Management Plan (CMP) is correctly applied and provide appropriate reasons in case of non conformance's
- Define which Configuration Items are to be managed in the Configuration Item List
- Define the Product Baseline
- Support changes to Configuration Items within the DMCCB
- Manage the delivery of software products
- Maintain the Configuration Item List
- Manage the configuration control resources used by DM

- Maintain an awareness of the relationship between the elements of the Product Baseline (in order for instance to be able to answer the question: “What is the environment and which software is installed?”)
- Check that the Product Assurance and CMP procedures are correctly applied when Configuration Items are changed
- Participate in DMCCB activities

The Configuration Manager is the secretary of the CCB and works with the support of the Scientific and Technical Leaders and participates in the CCB monitoring the development and change control process.

6.11.1 Configuration Item List

The Configuration Item List (CIL) is the list of items that are maintained under configuration control. CUs and DPC need to report their configuration items in the CIL with an adequate level of details. CIL is part of the development plan but may be written in a separate document to which the development plan refers.

The Configuration Manager in charge has to identify the configuration items to include in the CIL, with the help of the technical leader and to maintain it when changes to the configuration items happen.

6.11.2 Release management

In DM usually each product will be released once per cycle. Additional releases may be done in case of bug fixing, urgent issues, or in case that the previous one is incomplete.

Each release needs to be identified with:

- Configuration Item
- Documentation:
 - User Manual: to be updated each major release

Requirements Specification: to be updated each major release

Test Specification: to be updated each major release

Release Note: new document each major release, updated for patch releases

Test Report: new document each major release, updated for patch releases

- Latest Release in the master branch in GitHub.

This information identifies a product baseline.

The release manager is in charge of preparing the release with the technical lead for the product. After CCB approval, the release will be delivered to NCSA.

6.11.3 Configuration Baseline

A Configuration Baseline (CB) represents the approved status of the project at key milestones like formal review or at the beginning of test activities.

Configuration Baselines are applicable to hardware and software, and will include the documents that describe the CIs and their status.

6.12 Lead Institution Senior Positions

Each Lead Institution (as defined in Section 8; see also Table 3) has a T/CAM and Scientific or Engineering Lead, who jointly have overall responsibility for a broad area of DM work, typically a Work Breakdown Structure (WBS) Level 2 element. They are supervisors of the team at their institution, with roles broadly analogous to those of the DM Project Manager and Project Scientist.

6.12.1 Technical/Control Account Manager (T/CAM)

Technical/Control Account Managers have managerial and financial responsibility for the engineering teams within DM. Each T/CAM is responsible for a specific set of WBS elements. Their detailed responsibilities include:

- Develop, resource load, and maintain the plan for executing the DM construction project within the scope of their WBS
- Synchronize the construction schedule with development in WBS elements managed by other T/CAMs
- Maintain the budget for their WBS and ensuring that all work undertaken is charged to the correct accounts
- Work with the relevant Science Leads and Product Owners (Section 6.5) to develop the detailed plan for each cycle and sprint as required
- Work with the DM Project Controller (Section 6.4) to ensure that all plans and milestones are captured in the LSST Project Controls system
- Perform day-to-day management of staff within their WBS
- Perform the role of “scrum-master” during agile development
- Report activities as required, including providing input for monthly status reports.

6.12.2 Institutional Science/Engineering Lead

The Institutional Science/Engineering Leads serve as product owners (Section 6.5) for the major components of the DM System (Alert Production, Data Release Production, Science User Interface etc).

In addition, they provide scientific and technical expertise to their local engineering teams.

They work with the T/CAM who has managerial responsibility for their product to define the overall construction plan and the detailed cycle plans for DM.

Institutional science leads are members of the DM System Science Team (Section 7.1) and, as such, report to the DM Subsystem Scientist (Section 6.3).

7 Data Management Groups/Bodies

Since the DM team is distributed in terms of geography and responsibility across the LSST partner and lead institutions, mechanisms are needed to ensure that the project remains on

track at all times. There are four primary coordinating bodies to ensure the management, technical, and quality integrity of the DM Subsystem.

7.1 System Science Team

Members of the DM System Science Team (SST) work together to define, maintain, and communicate to the DM System Engineering team a coherent vision of the LSST DM system responsive to the overall LSST Project goals, as well as scientifically validate the as-build system (LDM-503, Section 9.).

7.1.1 Organization and Goals

The System Science Team includes:

- DM Subsystem Scientist (chair)
- DM Science Validation Scientist
- DM Institutional Science Leads
- DM System Science Analysts
- DM Science Pipelines Scientist

The System Science Team has been chartered to:

- Support the DM Subsystem Scientist (as the overall DM Product Owner) in ensuring that Data Management Subsystem's initiatives provide solutions that meet the overall LSST science goals.
- Support the Institutional Science Leads in their roles as Product Owners for elements of the DM system their respective institutions have been tasked to deliver.
- Support the DM Science Validation Scientist, who organizes and coordinates the science validation efforts (LDM-503).
- Guide the work of System Science Analysts, who generally lead and/or execute studies needed to support SST work.

- Provide a venue for communication with the Science Pipelines Scientist, who broadly advises on topics related to the impact of science pipelines on delivered science and vice versa (Section 6.6).

The members of the System Science Team report to the DM Subsystem Scientist and share the following responsibilities:

- To communicate with the science community and internal stakeholders to understand their needs, identifying the aspects to be satisfied by the DM Subsystem.
- Liaison with the science collaborations to understand and coordinate any concurrent science investigations relevant to the DM Subsystem.
- Develop, maintain, and articulate the vision of DM-delivered LSST data products and services that is responsive to stakeholder needs, balanced across science areas, well motivated, and scientifically and technologically current.
- Work with the DM Project Manager and DM Technical Managers to communicate and articulate the DM System vision and requirements to the DM engineering team.
- Identify, develop, and champion new scientific opportunities for the LSST DM System, as well as identify risks where possible.
- Develop change proposals and/or evaluate the scientific impact of proposed changes to DM deliverables driven by schedule, budget, or other constraints.
- Lead the Science Verification of the deliverables of the DM subsystem.

7.1.2 Communications

DM System Science Team communication mechanisms are described on the SST Confluence page at <http://ls.st/sst>.

7.2 DM Systems Engineering Team

The Systems Engineering Team is led by the DMPM (§6.1) and looks after all aspects of systems engineering. It is comprised of not only the Systems Engineer (Section 6.7) but also the

Software Architect (Section 6.9), Operations Architect (Section 6.10), Pipeline Scientist (Section 6.6) and Interface Scientist (Section 6.8).

While the product owners (Section 6.5) help DM to create products, which are fit for purpose, the Systems Engineering Team must ensure we do it correctly. This group concerns its self with (sub)system wide decisions on architecture and software engineering.

The specific tasks of this group include:

- Formalize the product list for DM⁴
- Formalize the documentation tree for DM, defining which documents need to be produced for each product
- Agree the process for tracing the baseline requirements verification and validation status.
- Agree the formal versions of documents and software which form the technical baseline, individual items will go through the CCB for formal approval.
- ...

Some of these tasks are will be delegated to individual group members. These individuals also are the conduit to/from the rest of the DM team to raise ideas/issues with the engineering approach.

7.2.1 Communications

The System Engineering Team will only physically meet to discuss specific topics: there will not be a regular meeting of the group outside of the one to one meetings with the DM project manager for the individuals in the group. Discussions will be held via email until in person talks are required.

⁴In this sense, “products” are the software and systems which produce data products, rather than the data products themselves. See also §5.

7.3 DM Leadership Team

The purpose of the DM Leadership Team (DMLT) is to assist the DMPM establish the scope of work and resource allocation across DM and ensure overall project management integrity across DM. The following mandate established the DMLT:

- Charter/purpose
 - Maintain scope of work and keep within resource allocation across DM
 - Ensure overall project management integrity across DM
 - Ensure Earned Value management requirements are met
- Membership
 - Co-chaired by the DM Project Manager (Section 6.1) and DM Project Scientist (Section 6.3)
 - Lead Institution Technical/Control Account Managers (T/CAMs; Section 6.12.1)
 - Institutional Science or Engineering Leads (Section 6.12.2)
 - Members of the DM Systems Engineering Team (Section 7.2)
- Responsibilities
 - Prepares all budgets, schedules, plans
 - Meets every week to track progress, address issues/risks, adjust work assignments and schedules, and disseminate/discuss general PM communications

The DM Leadership Team and the DM System Engineering Team (Section 7.2) work in synchrony. The DMLT makes sure the requirements and architecture/design are estimated and scheduled in accordance with LSST Project required budgets and schedules.

7.3.1 Communications

A mailing list⁵ exists for DMLT related messages. On Mondays the DMLT hold a brief (30 to 45 minutes) telecon. This serves to:

⁵lsst-dmlt@listserv.lsstcorp.org

- Allow the Project manager and DM Scientist to pass on important project level information and general guidance.
- Raise any blocking or poorly issues across DM — this may result in calling a splinter meeting to further discuss with relevant parties.
- Inform all team members of any change requests (LCRs) in process at LSST level which may be of interest to or have an impact on DM
- Check on outstanding actions on DMLT members

Face to Face meetings of DM are held twice a year these are opportunities to:

- Discuss detailed planning for the next cycle
- Discuss technical topics in a face to face environment
- Work together on critical issues
- Help make DM function as a team

7.4 DM Change Control Board

The DMCCB has responsibility for issues similar to those of the LSST Change Control Board, but with its scope restricted to the DM Subsystem. The DMCCB reviews and approves changes to all baselines in the Subsystem, including proposed changes to the DM System Requirements' (DMSR), reference design, sizing model, i.e. any LDM-series document. The Technical Baseline, software/hardware and documentation is written by DM and controlled by the DM-CCB , DMCCB validates that the form and content of the Technical Baseline is consistent with LSST project standards such as the System Engineering Management Plan (SEMP) LSE-17.

- Charter/purpose
 - Ensure that the DM Technical Baseline (LDM-xxx) documents are baselined and subsequently changed only when necessary and according to LSST and DM configuration control processes

- Membership
 - Chaired by the System Engineer
 - Members include the DM Software Architect (§6.9), DM Operations Architect (§6.10), DM System Interfaces Scientist (§6.8), DM SQuaRE T/CAM (§6.12.1) and DM Project Manager (§6.1)
 - For on-line virtual meetings, if a consensus or quorum or is not reached within one week, the DM Project Manager will make a unilateral decision
- Responsibilities
 - Determines when specification and deliverables are of sufficient maturity and quality to be baselined (placed under configuration controlled status) or released.
 - Reviews and approves/rejects proposed changes to baselined items

8 Lead institutions in DM

8.1 LSST Tucson

The LSST Project Office in Tucson hosts the DM Project Manager (Section 6.1) and the Systems Engineer (Section 6.7). In addition, it is home to the Science Quality and Reliability Engineering(SQuaRE) group and LSST International Communications and Base Site (ICBS) groups, described below.

8.1.1 Science Quality and Reliability Engineering

The SQuaRE group is primarily charged with providing technical feedback to the DM Project Manager that demonstrates that DM is fulfilling its responsibilities with regard to quality — of both scientific data products and software — software performance, and reliability. As such, areas of activity include:

- Development of algorithms to detect and analyze quality issues with data⁶

⁶This may overlap with work carried out by the Science Pipelines groups (§§8.3.1 & 8.2.1). In some instances this will involve sharing code; in others, it may merit duplicating a metric to ensure that it is correct.

- Infrastructure development to support the generation, collection, and analysis of data quality and performance metrics
- DM developer support services to ensure DM is using appropriate tools to aid software quality
- DM documentation support, to include defining standards and providing tooling for documentation as well as some document writing
- Support of publicly released software products, including porting and distributing it according to the scientific community's needs

In the event that SQuaRE identifies issues with the performance or future maintainability of the DM codebase, it will bring them to the attention of the DM Software Architect. In the event that SQuaRE identifies issues with the quality of the data or algorithmic performance, it will bring them to the attention of the DM Project Scientist.

8.1.2 LSST International Communications and Base Site

The ICBS group spans both Tucson and La Serena, and is responsible for the design, procurement, installation, deployment, verification, and operating support during construction and commissioning of all data communications networks at the Summit and Base sites, as well as links between all the LSST Sites, with two exceptions: the Summit Network (WBS 1.04C.12.5) and the Archive External Network (1.02C.07.04.06). In the case of the exceptions, there are technical and managerial interfaces between the ICBS and the responsible parties, as well as overlaps of staff. The LSST Network Engineering Team (NET) spans all of these networking assignees and is chaired by the ICBS staff.

The ICBS group is also jointly responsible with the Data Facility Team at NCSA for procurement, installation, deployment, verification, and operating support during construction and commissioning of the computing and storage infrastructure at the Base Site.

Since a large majority of the ICBS work involves procurement and contracted services, the group works in close cooperation with AURA procurement and contracts, as well as with the following major sub awardees and their subcontractors:

REUNA - Chilean National Networks Florida International University/AmLight - International

Networks connecting Chile and the United States, and US National Networks.

8.2 Princeton University

Princeton University hosts the Pipelines Scientist (Section 6.6) and the Data Release Production group, described below.

8.2.1 Data Release Production

The Data Release Production (DRP) group has three major areas of activity within DM.

- Definition and implementation of the scientific algorithms and pipelines which will be used to generate LSST's annual data releases;
- Definition and implementation of the algorithms and pipelines which will be used to produce the "calibration products" (for example, flat fields, characterization of detector effects, etc) which will be used as inputs to the photometric calibration procedure in both nightly and annual data processing. This includes the development of the spectrophotometric data reduction pipeline for the Auxiliary Telescope;
- Development, in conjunction with the Alert Production team (AP; Section 8.3.1), of a library of re-usable software libraries and components which form the basis of both the AP and DRP pipelines and which are made available to science users within the LSST Science Platform.

Development of software in support of annual data releases and of reusable software components are carried out under the direction of the DRP Science Lead, who acts as product owner for this part of the system. The DRP Science Lead is ultimately responsible to both the Pipelines Scientist (Section 6.6) and DM Project Scientist (Section 6.3).

The product owner for calibration products is the LSST Calibration Scientist (who doubles as the Pipelines Scientist, Section 6.6). The Calibration Scientist liaises with other LSST subsystems and with the products owners of the annual and nightly data processing pipelines to ensure that appropriate calibration products are available to those pipelines to enable them to meet specifications.

Management of the group is the responsibility of the Science Pipelines T/CAM, reporting to the DM Project Manager (Section 6.1).

The DRP group is responsible for delivering software which adheres to the architectural and testing standard defined by the Software Architect (Section 6.9). In addition, the DRP group is responsible for testing each major product delivered to demonstrate its fitness for purpose, and working with the DM Project Scientist and DM System Science Team (Section 7.1) to define, run and analyze “data challenges” and other large scale tests to validate the performance of the data release production system.

8.3 The University of Washington

8.3.1 Alert Production

8.4 California Institute of Technology/IPAC

IPAC hosts the DM Interface Scientist (Section 6.8) and the Science User Interface and Tools (SUIT) group described below.

8.4.1 Science User Interface and Tools

The Science User Interface and Tools (SUIT) group has four major areas of activity within DM:

Design and develop the Firefly Web-based visualization and data exploration framework, based upon the the same software already in operations in other NASA archive services (i.e. IRSA's WISE Image Service) . The Firefly framework provides three basic components – image display and manipulation, tabular table display and manipulation, and 2D plotting – all of which work together to provide different views into the same data. Firefly also provides JavaScript and Python APIs to enable developers to easily use the components in their own Web pages or Jupyter notebooks.

Develop the interfaces needed to connect Firefly to the other LSST Science Platform components, e.g., connect to authentication and authorization, DAX services, user workspace, flexible compute system. Develop visualizations of the objects in the LSST Data Products data model, and support their metadata; e.g., Footprint, HeavyFootprint, WCS models. Provide

basic access to Firefly from the LSST stack via `afw.display`.

Design and implement the Portal Aspect of the LSST Science Platform for Data Access Center, based on Firefly, providing scientists an easy to use interface to search, visualize, and explore LSST data. The portal will enable users to do as much data discovery and exploration as possible through complex searches and facilitate data assessment through visualization and interaction. The Portal will assist users in understanding the semantic linkages between the various LSST data products. The Portal will guide users to documentation on the Science Platform itself, the LSST data products, and the processing that generated them. Support linkage between the Portal and Notebook aspects of the Science Platform, enabling users to switch between the aspects easily by providing tools to make data selected in the Portal readily available for further analysis in user notebooks.

Design and develop the LSST Alert Subscription web portal to enable scientists to access the alert system. The subscription service will enable users to register filters and destinations for alerts matching their interests. The Alert portal will also provide basic capabilities for searching alerts history and for exploring linkage between alerts and other data products.

8.5 SLAC

SLAC hosts the DM Software Architect (Section 6.9) and the Science Data Archive and Data Access Services group described below.

8.5.1 Science Data Archive and Data Access Services

The Science Data Archive and Data Access Services (DAX) group has the following major areas of activity within DM:

- Provides software to support ingestion, indexing, query, and administration of DM catalog and image data products, data provenance, and other associated metadata within the LSST Data Access Centers;
- Provides implementations of data access services (including IVOA services), as well as associated client libraries, to be hosted within the LSST Data Access Centers, which facilitate interaction between LSST data products and tools provided by both other parts

of the LSST project and by the astronomical research community at large;

- Provides a Python framework (the "Data Butler"), used by the LSST science pipelines, to facilitate abstract persistence/retrieval of in-memory Python objects to/from generic archives of those objects;
- Provides a Python framework ("SuperTask") which serves as an interface layer between pipeline orchestration and algorithmic code, and which allows pipelines to be constructed, configured, and run at the level of a single node or a group of tightly-synchronized nodes;
- Provides support for various middleware and infrastructure toolkits used by DM which would otherwise have no authoritative home institution within DM (e.g. logging support library, spherical geometry support library).

Management of the group is the responsibility of the DAX T/CAM, reporting to the DM Project Manager (Section 6.1).

The DAX group is responsible for delivering software which adheres to the architectural and testing standard defined by the Software Architect (Section 6.9). In addition, the DAX group is responsible for testing each major product delivered to demonstrate its fitness for purpose, and running and analyzing large scale tests to validate the performance of the science data archive and data access systems.

8.6 NCSA

NCSA hosts the LSST Project Office Information Security Officer and Computer Security group, as well as the DM group responsible for construction and integration of the LSST Data Facility (LDF), described below.

The LSST Data Facility group has the following major areas of activity within DM:

1. Construction of services, including software and operational methods, supporting observatory operations and nightly data production (Level 1 Services). Level 1 Services ingest raw data from all Observatory cameras and the Engineering and Facilities Database (EFD) into the central archive; provide a dedicated computing service controllable by the

- Observatory Control System (OCS) for prompt generation of nightly calibration assessments, science image parameters, and transient alerts; and provide computing services, data access, and a QA portal for Observatory staff.
2. Construction of services, including software and operational methods, for bulk batch data production. Batch Production Services execute processing campaigns, using resources at NCSA and satellite computing centers, to produce data release products, generate templates and calibrations, and perform scaled testing of science pipelines to assess production readiness.
 3. Construction of services, including software and operational methods, for hosting and operating data access services for community users. These services host the SUIT portal, manage the JupyterLab environment, provide computing and data storage for the Data Access Centers, enable bulk data export, and host the LSST limited alert-filtering service and feeds to community-provided brokers.
 4. Construction of services, including software and operational methods, for the Data Backbone. Data Backbone Services provide ingestion, management, distribution, access, integrity checking, and backup and disaster recovery for files and catalog data in the LSST central data archive.
 5. Construction and operation of services for LSST staff. Staff Services provide specific testing and integration platforms (e.g., a Prototype Data Access Center) and general computing and data services for LSST developers.
 6. Provisioning and management of hardware infrastructure at NCSA and the Chilean Base Center for all services described above, as well as infrastructure for project-wide network-based computer security services and authentication and authorization services.
 7. Construction and operation of a service management framework and methods to monitor operations of service elements in accordance with service level agreements, track issues, manage service availability, and support change management.
 8. Operation of services and IT systems during construction to support on-going development, integration, and commissioning activities.

The LDF group is responsible for delivering instantiated production services, which integrate software and hardware components developed across DM. The LDF group performs large-scale tests to integrate and verify production readiness of all components.

9 Development Process

DM is essentially a large software project, more we are developing scientific software with the in uncertainties that brings with it. An agile [23] is particularly suited to scientific software development. The development follows a six month cyclical approach and DM products are under continuous integration using the application software Jenkins. All code is developed in the GitHub open source repository under an open source license. Releases follow a six month cadence but the master is intended to be always working with a continuous integration system ensuring this.

How this fits with the Earned Value System is described in DMTN-020.

9.1 Communications

The main stories for the six month planning period are discussed at the DMLT F2F meeting near the beginning of the cycle (See Section 7.3).

The T/CAMs of each of the institutions meet via video on Tuesdays and Fridays for a short *standup* meeting to ensure that any cross-team issues are surfaced and resolved expeditiously. This meeting is chaired by the Deputy Project Manager. Each T/CAM notes any significant progress of interest to other teams and any problems or potential problems that may arise.

9.2 Conventions

Coding guidelines and conventions are documented online in <https://developer.lsst.io>

9.3 Reviews

The DM Project Manager and Subsystem Scientist will periodically convene internal reviews (following LSE-159) of major DM components as necessary to assess progress and maintain the integrity of the overall system. Planned DM reviews will be listed at the LSST Project Review Hub (<https://project.lsst.org/reviews/hub/>).

10 Data Management Problem/Conflict Resolution

The above organizational structure allocates significant responsibility to lead institutions. As such, when problems arise that cannot be solved with the responsibility and scope allocated to an institution, the path of escalation and resolution of such problems must be clear.

Any intra institutional problem should be brought as early as possible to the DM Project manager. The PM will attempt to mediate a resolution. The PM will consult with DMLT, DM System Science Team and DM System engineering if there are Scientific or technical impacts to be considered.

Should an issue need to be escalated the PM will bring it up in the weekly LSST project managers meeting. In that forum a way forward will be agreed with the LSST project manager and other sub system managers.

A DM Product List

WBS	Product	Description	Manager	Owner	Packages
.	Data Management	Data Management System			
.	Data Access Center	DAC Software			
	Bulk Distrib	Bulk Distribution System	Joel Plutchak		
	Proposal Manager	Proposal Manager	Joel Plutchak		
	Resource Manager	DAC Resource Manager	Joel Plutchak		
.	Data Backbone	Data Backbone System			
1.02C.06.02.01	Data Butler Client	Data Butler data access client library	Fritz Mueller		daf_persistence/ db/ daf_fmt_*
	Transport/Repl	File and database transport and replication with caching endpoints	Joel Plutchak		
.	Science Cat Archive	Science catalog archive			
1.02C.06.02.03	Qserv DBMS	Qserv distributed database system	Fritz Mueller		qserv/ partition/ scisql
.	Science Cat DBs	Science catalog databases			
1.02C.03.03	Alert DB	Alert database	Simon Krughoff	Eric Bellm	
1.02C.06.01.01	L1 Catalog DB	L1 catalog database	Fritz Mueller		cat
1.02C.06.01.01	L2 Catalog DB	L2 catalog database	Fritz Mueller		cat
.	Science Img Archive	Science image archive			
1.02C.06.02.05	Global Metadata	Global metadata service	Fritz Mueller		
1.02C.06.01.01	Provenance DB	Provenance database	Fritz Mueller		
.	Developer services	Developer services			
1.02C.10.02.03.01	Build/Unit Test	Build and unit test service	Frossie Economou		sconsUtils/ base/ lsstsw/ lsst_build
1.02C.10.02.03.04	Devel Comm Tools	Developer communication tools	Frossie Economou		
1.02C.10.02.03.03	Doc Infrastructure	Documentation infrastructure	Frossie Economou		lsst-texmf/ templates/ lsstDoxygen
1.02C.10.02.03.01	SW Version Control	Software version control system	Frossie Economou		
1.02C.10.02.03.05	Issue Tracking	Issue (ticket) tracking service	Frossie Economou		
1.02C.10.02.03.02	Packaging/Distrib	Packaging and distribution	Frossie Economou		lsst/ shebangtron/ lsst_dm_stack_demo
	Identity Manager	Identity (Authentication and Authorization) Manager	Joel Plutchak		
.	Infrastructure	Infrastructure Software Systems			
	Batch Proc	Batch Processing System	Joel Plutchak		
	Infra Systems	Filesystems/ provisioning/monitoring systems and system management	Joel Plutchak		
.	Task Execution	Task execution framework			



1.02C.06.03	Activator Bases	Activator base and Command Line Activator	Fritz Mueller		
1.02C.06.03	Pipeline Config	Pipeline configuration	Fritz Mueller		pex_config
1.02C.06.04.01	Logging	Logging	Fritz Mueller		log
1.02C.06.03	Multi-Core Task	Multi-core Task API	Fritz Mueller		
1.02C.06.03	Multi-Node Task	Multi-node Task API	Fritz Mueller		pipe_base/ ctrl_pool
1.02C.06.03	SuperTask	SuperTask	Fritz Mueller		pipe_supertask/ pipe_base/ pex_exceptions
	Integration Test	Integration and test			
1.02C.10.02.03.01	Contin Integration	Automated integration and test services	Frossie Economou		Jenkins
	Precursor Data	Precursor data for development and testing			obs_*/ validation_data_*/ testdata_*/ afwdata
.	IT Environments	Computing and Storage Infrastructure including provisioning			
.	Archive IT	Archive IT Environments			
	Archive Center Env	Archive Production Center environment	Joel Plutchak		
	Archive DAC Env	Archive DAC environment	Joel Plutchak		
	DAC Integ Env	DAC Integration environment (PDAC)	Joel Plutchak		
	Archive DBB Env	Archive Data Backbone endpoints and storage	Joel Plutchak		
	DBB Integ Env	Data Backbone Integration environment	Joel Plutchak		
	Dev Env	Developer environment	Joel Plutchak		
	L1 Integration Env	Level 1 Integration environment	Joel Plutchak		
	L2 Integration Env	Level 2 Integration environment	Joel Plutchak		
	Satellite Env	Satellite compute environment	Joel Plutchak		
	Science Valid Env	Archive science validation environment	Joel Plutchak		
.	Base IT	Base IT Environments			
	Base Center Env	Base Production Center environment	Joel Plutchak		
	Base DAC Env	Base DAC environment	Joel Plutchak		
	Base DBB Env	Base Data Backbone endpoints and storage	Joel Plutchak		
	Comm Cluster Env	Commissioning Cluster environment	Joel Plutchak		
.	Level 1 System	Level 1 System			
1.02C.03.03	Alert Distribution	Alert distribution service	Simon Krughoff	Eric Bellm	
1.02C.03.03	Alert Filtering	Alert filtering service	Simon Krughoff	Eric Bellm	
	L1 Offline Proc	L1 Offline Processing System	Joel Plutchak		
.	L1 OCS Components	Level 1 Online (OCS-connected) components			
	Archiver	Archiving Commandable SAL Component	Joel Plutchak	Felipe Menanteau	ctrl_iip
	Catchup Archiver	Catch-up Archiving Commandable SAL Component	Joel Plutchak	Felipe Menanteau	ctrl_iip
	EFD Transform	EFD Transformation Commandable SAL Component	Joel Plutchak	Felipe Menanteau	
	Header Generator	Header Generator Commandable SAL Component	Joel Plutchak	Felipe Menanteau	
	OCS Batch Proc	OCS-Driven Batch Processing Commandable SAL Component	Joel Plutchak	Felipe Menanteau	ctrl_iip
	Pointing Publisher	Pointing Prediction Publishing Commandable SAL Component	Joel Plutchak	Felipe Menanteau	
	Prompt Proc	Prompt Processing Commandable SAL Component	Joel Plutchak	Felipe Menanteau	ctrl_iip
	Telem Gateway	Telemetry Gateway Commandable SAL Component	Joel Plutchak	Felipe Menanteau	ctrl_iip
.	L1 Science Payloads	L1 science payloads			
.	Offline Alert Prod	Offline Alert Production payload			
1.02C.03.03	Offline Alert Gen	Offline alert generation pipeline	Simon Krughoff	Eric Bellm	
1.02C.03.06	Moving Object	Moving object pipeline	Simon Krughoff	Eric Bellm	mops_daymops
1.02C.03.04	Precovery	Precovery and forced photometry pipeline	Simon Krughoff	Eric Bellm	
1.02C.03.01	Offline SFP	Offline single frame processing pipeline	Simon Krughoff	Eric Bellm	
.	Prompt Alert Prod	Prompt Processing Alert Production payload			
1.02C.03.03	Alert Gen Pipe	Alert generation pipeline	Simon Krughoff	Eric Bellm	
1.02C.03.01	Single Frame Pipe	Single frame processing pipeline	Simon Krughoff	Eric Bellm	pipe_drivers
1.02C.04.02	Aux Tel Spec Pipe	Offline Auxiliary Telescope spectrograph pipeline	John Swinbank	Jim Bosch	
1.02C.04.02	Daily Calibration	OCS-Controlled batch daily calibration update payload	John Swinbank	Jim Bosch	
1.02C.04.02	Offline Calibration	Offline calibration single frame processing pipeline	John Swinbank	Jim Bosch	pipe_drivers



1.02C.04.02	Prompt Calibration	Prompt Processing raw calibration validation payload	John Swinbank	Jim Bosch	pipe_drivers
1.02C.04.02	CBP Control	OCS control scripts for collimated beam projector control	John Swinbank	Jim Bosch	
	L1 Quality Control	L1 QC measurement generators	Simon Krughoff	Eric Bellm	
.	Level 1 Services	Level 1 Services			
.	Archiving Services	Image and EFD Archiving Services			
	Aux Tel Archiver	Auxiliary Telescope Archiving Service			
	ComCam Archiver	ComCam Archiving Service			
	LSSTCam Archiver	LSSTCam Archiving Service			
	ComCam Catchup	ComCam Catchup Archiving Service			
	LSSTCam Catchup	LSSTCam Catchup Archiving Service			
	EFD Transf Service	EFD Transformation Service			
	Pointing Service	Pointing Prediction Publishing Service			
.	Prompt Services	Prompt Processing Services			
	ComCam Prompt	ComCam Prompt Processing Service			
	LSSTCam Prompt	LSSTCam Prompt Processing Service			
.	Level 2 System	Level 2 System			
	L2 Quality Control	L2 QC measurement generators	John Swinbank	Jim Bosch	validate_drp/ verify_metrics/ ci_hsc
.	L2 Science Payloads	L2 science payloads			
1.02C.04.02	CPP Quality Control	CPP QC measurement generators	John Swinbank	Jim Bosch	
1.02C.04.02	Periodic Cal Prod	Periodic CPP payload	John Swinbank	Jim Bosch	
1.02C.04.02	Annual Cal Prod	Annual CPP payload	John Swinbank	Jim Bosch	
.	Data Release Prod	Annual mini-DRP and DRP payload			
1.02C.04.04	Coadd and Diff	Image coaddition and differencing	John Swinbank	Jim Bosch	pipe_drivers
1.02C.04.05	Coadd Processing	Coadd processing	John Swinbank	Jim Bosch	pipe_drivers
1.02C.04.06	DRP Postprocessing	DRP Postprocessing	John Swinbank	Jim Bosch	
1.02C.04.03	Image Char and Cal	Image characterization and calibration	John Swinbank	Jim Bosch	pipe_drivers
1.02C.04.06	Object Char	Multi-epoch object characterization	John Swinbank	Jim Bosch	
1.02C.04.05	Overlap Resolution	Overlap resolution	John Swinbank	Jim Bosch	
1.02C.06.01.01	DRP-Internal DB	DRP-internal database	Fritz Mueller		daf_ingest
1.02C.03.04	Template Gen Prod	Template generation payload	Simon Krughoff	Eric Bellm	
.	Production Exec	Production Execution System			
	Campaign Manager	Campaign Manager	Joel Plutchak		
	Job Activator	Job Activator	Joel Plutchak		
	Pre-Flight Activator	Pre-flight Activator	Joel Plutchak		
	Workflow Manager	Workflow Manager/Orchestrator	Joel Plutchak		ctrl_orca/ ctrl_execute/ ctrl_provenance
	Workload Manager	Workload Manager	Joel Plutchak		ctrl_platform_*/ ctrl_stats/
.	DM Networks	Data Management Provided Networks			
1.02C.07.04.06	Arc Extl Net	Archive External Network	Don Petravick	D Wheeler	
1.02C.07.04.03 (moving to 1.02C.08)	Base Network	Base Local Area Network	Don Petravick (moving to Jeff Kantor)	Jeff Kantor/Don Petravick	
.	Chilean Nat	Summit - Gatehouse La Serena - Gatehouse/ La Serena - Santiago Networks DWDM Equipment			
1.02C.08.03.01.03	Summit - AURA	Summit - AURA Gatehouse Network	Jeff Kantor	Jeff Kantor	
1.02C.08.03.01.04	DWDM Equipment	DWDM Equipment	Jeff Kantor	Jeff Kantor	
1.02C.08.03.01.01	La Serena - AURA	La Serena - AURA Gatehouse Network	Jeff Kantor	Jeff Kantor	
1.02C.08.03.01.01	La Ser - Santi	La Serena - Santiago Network	Jeff Kantor	Jeff Kantor	
.	Int/US WANS	International WAN/US WAN			
1.02C.08.03.02.01	SCL - MIA 100 Gbps	Santiago - Miami 100 Gbps Ring	Jeff Kantor	Jeff Kantor	
1.02C.08.03.02.02	Network Mgmt	Network Management	Jeff Kantor	Jeff Kantor	
1.02C.08.03.02.03	SCL - BR Spectrum	Santiago - Boca Raton Spectrum	Jeff Kantor	Jeff Kantor	
1.02C.08.03.02.01	US National WAN	US National WAN	Jeff Kantor	Jeff Kantor	
1.02C.08.03	Long-Haul Nets	Summit - Base/ Base - Archive/ US Networks	Jeff Kantor	Jeff Kantor	
.	Science Algorithms	Common science algorithmic components			
1.02C.04.05	Aperture Corr	Aperture correction	John Swinbank	Jim Bosch	
1.02C.03.01	Artifact Detection	Artifact detection	Simon Krughoff	Eric Bellm	meas_algorithms
1.02C.03.01	Artifact Interp	Artifact interpolation	Simon Krughoff	Eric Bellm	
1.02C.04.05	Association/Match	Association and matching	John Swinbank	Jim Bosch	
1.02C.03.07	Astrometric Fit	Astrometric fitting	Simon Krughoff	Eric Bellm	jointcal/ meas_astrom/ meas_mosaic
1.02C.04.03	Background Estim	Background estimation	John Swinbank	Jim Bosch	meas_algorithms
1.02C.04.03	Background Ref	Background reference	John Swinbank	Jim Bosch	



1.02C.03.02	DIAObj Association	DIAObject association	Simon Krughoff	Eric Bellm	
1.02C.03.04	DCR Template Gen	DCR-corrected template generation	Simon Krughoff	Eric Bellm	
1.02C.04.05	Deblending	Deblending	John Swinbank	Jim Bosch	meas_deblender
	Img Decorrelation	Image decorrelation	Simon Krughoff	Eric Bellm	ip_diffim
1.02C.04.04	Image Coaddition	Image coaddition	John Swinbank	Jim Bosch	coadd_utils/ coadd_chisquared
1.02C.03.01	ISR	ISR	Simon Krughoff	Eric Bellm	pipe_tasks/ ip_isr
1.02C.04.05	Measurement	Measurement	John Swinbank	Jim Bosch	meas_base/ meas_algorithms/ meas_extensions_*/ meas_modelfit
.	Orbit/Ephemeris	Orbit tools			
1.02C.03.06	Attribution/Precov	Attribution and precovary	Simon Krughoff	Eric Bellm	mops_daymops
1.02C.03.06	Ephemeris Calc	Ephemeris calculation	Simon Krughoff	Eric Bellm	mops_night
1.02C.03.06	Orbit Fitting	Orbit fitting	Simon Krughoff	Eric Bellm	
1.02C.03.06	Orbit Merging	Orbit merging	Simon Krughoff	Eric Bellm	
1.02C.03.06	Tracklet Gen	Tracklet generation	Simon Krughoff	Eric Bellm	mops_daymops
1.02C.03.08	Photometric Fit	Photometric fitting	Simon Krughoff	Eric Bellm	jointcal/ meas_mosaic
	Proper Motion	Proper motion and parallax	Simon Krughoff	Eric Bellm	
1.02C.04.03	PSF Estim Large	PSF estimation (visit)	John Swinbank	Jim Bosch	
1.02C.03.01	PSF Estim Small	PSF estimation (1 CCD)	Simon Krughoff	Eric Bellm	meas_algorithms
1.02C.04.04	PSF Matching	PSF matching	John Swinbank	Jim Bosch	
	Raw Meas Cal	Raw measurement calibration	John Swinbank	Jim Bosch	
1.02C.03.01	Reference Catalogs	Reference catalogs	Simon Krughoff	Eric Bellm	meas_algorithms
1.02C.03.02	Reference Match	Matching to reference catalogs	Simon Krughoff	Eric Bellm	
1.02C.03.01	Spatial Models	Spatial models	Simon Krughoff	Eric Bellm	afw
1.02C.04.05	Source Detection	Source detection	John Swinbank	Jim Bosch	
1.02C.04.05	Star/Galaxy Sep	Star/galaxy classification	John Swinbank	Jim Bosch	
1.02C.03.04	Template Storage	Difference template storage/retrieval	Simon Krughoff	Eric Bellm	
	Variability Char	Variability characterization	Simon Krughoff	Eric Bellm	
.	Science Platform	Science Platform			
.	DAX VO+ Services	DAX VO+ services			
1.02C.06.02.05	Catalog Access	Catalog access	Fritz Mueller		dax_dbserve
1.02C.06.02.05	Cat Metadata Acc	Catalog metadata access	Fritz Mueller		dax_metaserv
1.02C.06.02.05	Img Metadata Acc	Image metadata access	Fritz Mueller		dax_metaserv
1.02C.06.02.04	Image Access	Image access	Fritz Mueller		dax_imgserv
1.02C.06.02.02	Web Framework	Web services framework	Fritz Mueller		dax_webserv/ dax_webservcommon
.	SciPlat JupyterLab	Science Platform JupyterLab component			
1.02C.10.02.02.05	JupyterLab Activator	JupyterLab Activators	Frossie Economou		
1.02C.10.02.02.06	JupyterLab Deploy	JupyterHub deployment	Frossie Economou		
1.02C.10.02.02.01	JupyterLab Env	Basic JupyterLab environment	Frossie Economou		
1.02C.05.07.04	JupyterLab SUIT Intf	JupyterLab visualization widgets and other JupyterHub/Portal bridges	Xiuqin Wu		
1.02C.10.02.02.04	JupyterLab SW Env	JupyterLab software environments	Frossie Economou		
.	SciPlat Portal	Science Platform portal component			
1.02C.05.07.03	Firefly Python APIs	Low-level Python API to Firefly	Xiuqin Wu		firefly_client
1.02C.05.06	Firefly	LSST-independent Firefly framework and visualization capabilities	Xiuqin Wu		firefly
1.02C.05.09	SUI Alert Interfaces	Portal alert interfaces to configure alert subscriptions	Xiuqin Wu		
1.02C.05.08	Portal Applications	Web application(s) implementing the Portal	Xiuqin Wu		
user workspace	Portal Interfaces	Interfaces to DAX	identity management	1.02C.05.07	Xiuqin Wu
1.02C.05.07.03	Visualizers	Firefly components to visualize LSST Science Pipelines data objects	Xiuqin Wu		
.	Science Primitives	Science software primitives			
1.02C.03.05	Camera Descr	Camera descriptions	Simon Krughoff	Eric Bellm	afw
1.02C.03.05	Chromaticity Utils	Chromaticity utilities	Simon Krughoff	Eric Bellm	afw
1.02C.04.01	Convolution	Convolution kernels	John Swinbank	Jim Bosch	afw
1.02C.03.05	Approx 2-D Fields	Interpolation and approximation of 2-D fields	Simon Krughoff	Eric Bellm	afw
1.02C.04.01	Footprints	Footprints	John Swinbank	Jim Bosch	afw
1.02C.03.05	Fourier Transforms	Fourier transforms	Simon Krughoff	Eric Bellm	afw
1.02C.03.05	Common Functions	Common functions and source profiles	Simon Krughoff	Eric Bellm	afw
.	Geometry	Geometry primitives			
1.02C.03.05	Cartesian Geom	Cartesian geometry	Simon Krughoff	Eric Bellm	
1.02C.03.05	Coord Transforms	Coordinate transformations	Simon Krughoff	Eric Bellm	afw/ astshim
1.02C.06.04.03	Spherical Geom	Spherical geometry	Fritz Mueller		sphgeom/ skypix/ skymap/ geom/ afw

1.02C.04.01	Images	Images	John Swinbank	Jim Bosch	afw
1.02C.04.01	MC Sampling	Monte Carlo sampling	John Swinbank	Jim Bosch	afw
1.02C.04.01	Num Integration	Numerical integration	John Swinbank	Jim Bosch	afw
1.02C.04.01	Num Optimization	Numerical optimization	John Swinbank	Jim Bosch	afw
1.02C.04.01	PhotoCal Repr	Photometric calibration representation	John Swinbank	Jim Bosch	afw
1.02C.06.02.01	Property/Metadata	Multi-type associative containers	Fritz Mueller		daf_base
1.02C.03.05	Point-Spread Funcs	Point-spread functions	Simon Krughoff	Eric Bellm	meas_algorithms/ shapelet
1.02C.04.01	Random Numbers	Random number generation	John Swinbank	Jim Bosch	afw
1.02C.04.01	Science Tools	Science tools	John Swinbank	Jim Bosch	afw/ utils
1.02C.04.01	Basic Statistics	Basic statistics	John Swinbank	Jim Bosch	afw
1.02C.04.01	Tables	Tables	John Swinbank	Jim Bosch	afw
1.02C.03.05	Tree Structures	Tree structures (for searching)	Simon Krughoff	Eric Bellm	afw
1.02C.04.01	Warping	Warping	John Swinbank	Jim Bosch	afw
.	QC Dashboard	QC measurement collection/storage/dashboard service			
1.02C.10.02.01.04	Alert QC	Alert stream QC harness	Frossie Economou		
1.02C.10.02.01.01	QC Harness	QC harness	Frossie Economou		validate_base
1.02C.10.02.01.02	QC Notifications	QC threshold notification framework	Frossie Economou		
1.02C.10.02.01.03	QC Reports	QC verification reporting	Frossie Economou		

B Proposed FY2018 WBS: Data Management Construction

This WBS element provides the complete LSST Data Management System (DMS). The DMS has these main responsibilities in the LSST system:

- Process the incoming stream of images generated by the Camera Subsystem during observing to generate and archive the LSST nightly data products.
- Provide real-time information on data quality to the Observatory Control System (OCS) during observing.
- Process the entire survey data each year to produce deep catalogs of objects and precise measurements of those objects.
- Capture and process calibration images from the Camera Subsystem. Incorporate pipeline improvements and correct errors.
- Provide a VO-compliant interface that makes publicly available all generated data products.

The following products (per Section 5) are defined at this level of WBS:

- Data Management System
- Level 1 System
- Level 1 Online (OCS-connected) components
- Archiving Commandable SAL Component
- Catch-up Archiving Commandable SAL Component
- EFD Transformation Commandable SAL Component
- Prompt Processing Commandable SAL Component
- OCS-Driven Batch Processing Commandable SAL Component
- Telemetry Gateway Commandable SAL Component

- Header Generator Commandable SAL Component
- Pointing Prediction Publishing Commandable SAL Component
- Level 1 Services
- Image and EFD Archiving Services
- Auxiliary Telescope Archiving Service
- ComCam Archiving Service
- LSSTCam Archiving Service
- ComCam Catchup Archiving Service
- LSSTCam Catchup Archiving Service
- EFD Transformation Service
- Pointing Prediction Publishing Service
- Prompt Processing Services
- ComCam Prompt Processing Service
- LSSTCam Prompt Processing Service
- L1 Offline Processing System
- L1 science payloads
- Prompt Processing Alert Production payload
- Offline Alert Production payload
- L1 QC measurement generators
- Identity (Authentication and Authorization) Manager
- DAC Software
- Proposal Manager
- DAC Resource Manager
- Bulk Distribution System

- Infrastructure Software Systems
- Batch Processing System
- Data Backbone System
- File and database transport and replication with caching endpoints
- Science image archive
- Science catalog archive
- Science catalog databases
- Filesystems/ provisioning/monitoring systems and system management
- Computing and Storage Infrastructure including provisioning
- Archive IT Environments
- Developer environment
- Level 1 Integration environment
- Level 2 Integration environment
- DAC Integration environment (PDAC)
- Data Backbone Integration environment
- Archive Production Center environment
- Archive science validation environment
- Archive Data Backbone endpoints and storage
- Archive DAC environment
- Satellite compute environment
- Base IT Environments
- Base Production Center environment
- Base Data Backbone endpoints and storage
- Base DAC environment

- Commissioning Cluster environment
- Data Management Provided Networks
- Summit - Base/ Base - Archive/ US Networks
- Base Local Area Network
- Archive External Network
- Summit - Gatehouse La Serena - Gatehouse/ La Serena - Santiago Networks DWDM Equipment
- La Serena - Santiago Network
- La Serena - AURA Gatehouse Network
- Summit - AURA Gatehouse Network
- DWDM Equipment
- International WAN/US WAN
- Santiago - Miami 100 Gbps Ring
- Network Management
- Santiago - Boca Raton Spectrum
- US National WAN
- Level 2 System
- Production Execution System
- Campaign Manager
- Workload Manager
- Workflow Manager/Orchestrator
- Pre-flight Activator
- Job Activator
- L2 science payloads

- Annual mini-DRP and DRP payload
- L2 QC measurement generators
- Common science algorithmic components
- Image decorrelation
- Variability characterization
- Proper motion and parallax
- Raw measurement calibration
- Orbit tools
- Science software primitives
- Geometry primitives
- Science Platform
- DAX VO+ services
- Task execution framework
- QC measurement collection/storage/dashboard service
- QC harness
- QC threshold notification framework
- QC verification reporting
- Alert stream QC harness
- Science Platform JupyterLab component
- Basic JupyterLab environment
- JupyterLab software environments
- JupyterLab Activators
- JupyterHub deployment
- Developer services

- Software version control system
- Build and unit test service
- Packaging and distribution
- Documentation infrastructure
- Developer communication tools
- Issue (ticket) tracking service
- Integration and test
- Automated integration and test services
- Precursor data for development and testing
- Science Platform portal component

B.1 System Management

This WBS element provides all activities related to the management and administration of the Data Management WBS elements. This includes all activities and support to maintain a core team responsible for the execution of all Data Management Tasks. Effort in this task insures compliance with Project level controls, documentation, and reporting. This also includes overseeing the management of DMS physical assets (hardware, software, facilities, documents) and support for travel and communications during the Construction Phase, as well as preparing for the same activities during the Operations Phase. This includes:

- Project management of resources, schedules, tasks, and deliverables.
- Quality assurance of the DM project, including Science Data Quality Assessment.
- Configuration control of all DMS elements and asset management for all DMS physical assets.

No products are defined at this level of the WBS.

B.1.1 Project Management

This WBS element includes Project Management staff and supporting tools for the project. The primary deliverables are the Data Management sections of overall LSST PMCS-based project plans, schedules, budgets, and reports. Communications and collaboration tools for team management and coordination are also within this WBS element.

No products are defined at this level of the WBS.

B.2 Data Management Science

This WBS element provides for the scientific leadership of the Data Management Subsystem. Specifically, it provides:

- Ultimate ownership of all DM products (per §5);
- Coordination of the Institutional Science Leads;
- Communication with the external scientific community and internal stakeholders to understand their needs, and, where applicable, ensure they are satisfied with by the DM Subsystem;
- Liaison with science collaborations;
- Resources to identify, develop and champion new scientific opportunities for the LSST DM System, as well as identify risks where possible;
- Leadership of the DM Science Validation effort.

No products are defined at this level of the WBS.

B.2.1 1.02C.02.02: DM System Architecture

The Architecture Team is a subset of the DM Systems Engineering Team that implements decisions of that team by creating, maintaining, disseminating, and ensuring adherence to a common, consistent system architecture for the Data Management System.

The Architecture Team monitors construction activities to ensure consistency with the defined architecture and performs investigations needed to support its core responsibilities.

No products are defined at this level of the WBS.

B.2.1.1 1.02C.02.02.01: System Architecture Definition This WBS element includes all activities related to documenting the high-level architecture of the LSST Data Management System. This includes writing and maintaining documents that define and describe the DMS's high-level components and their interfaces, both internal and external, as well as how they work together and are operated to meet the DM System Requirements. This work is performed in conjunction with the technical leadership of the DM teams as well as that of other LSST subsystems.

No products are defined at this level of the WBS.

B.2.1.2 1.02C.02.02.02: System Architecture Oversight This WBS element includes all activities related to ensuring that the constructed LSST Data Management System, including the computing and storage systems, the processing systems, and the science pipelines, adheres to its architectural principles and standards and that the Data Management development processes are followed. It involves tracking software development; leading, advising, and educating during design, code, sprint, and other reviews; contributing to the completeness of verification testing; maintaining the DM Risk Register; and communicating the DMS architecture internally and externally. This WBS element also involves making decisions on design and process changes to ensure emergent properties of the system such as usability, reliability, understandability, and maintainability. The Architecture Team provides input to decision-making personnel and bodies but does not supervise, directly control, or exercise a veto over development work except where explicitly delegated that role. One such delegation is the Release Manager role which oversees and coordinates the preparation for each software release. Architecture Team input about low-level code is conveyed to individual developers during reviews. Input about refinement of designs is conveyed to technical leads and the NCSA Steering Committee. Input about revisions to designs or plans is conveyed to technical managers and the NCSA Steering Committee for incorporation into prioritization. Interactions with LSST System Engineering, Operations Planning, Risk Management, and Change Control are contained within this WBS, as is architectural representation in the DM

Systems Engineering Team and Change Control Board.

No products are defined at this level of the WBS.

B.2.1.3 1.02C.02.02.03: Architecture Investigation This WBS element includes all activities related to obtaining the data necessary to make architectural decisions, including literature research, prototyping, and model-building.

No products are defined at this level of the WBS.

B.3 1.02C.03: Alert Production

No products are defined at this level of the WBS.

B.3.1 1.02C.03.00: Management & Leadership

This WBS element covers project management and scientific leadership of the Alert Production group. It includes effort to develop the overall plan for the work of the group, to schedule that work, to perform day-to-day technical and control account management of the team, to coordinate development with other parts of the LSST WBS, and to contribute to the operation of the DM Subsystem Science Team.

No products are defined at this level of the WBS.

B.3.2 1.02C.03.01: Single Frame Processing

This WBS element covers the construction of the software, together with its associated configuration files, test suites and documentation, used to produce the single frame processing pipeline. This pipeline produces calibrated images from raw images. The focus is on a pipeline that produces science frames ready for image differencing, but is expected to result in many pieces that can be re-used in the data release production system.

The following products (per Section 5) are defined at this level of WBS:

- Single frame processing pipeline
- Offline single frame processing pipeline
- Reference catalogs
- ISR
- Artifact detection
- Artifact interpolation
- Spatial models
- PSF estimation (1 CCD)

B.3.3 1.02C.03.02: Catalog Association for Alert Production

This WBS element covers the construction of the software, together with its associated configuration files, test suites and documentation, which is used to associate DIASources with other entities required for alert packet construction. Specifically, this will include association with solar system objects and previously constructed DIAObjects.

This will also cover the work necessary to carry out the updating of DIAObjects with the addition of another DIASource.

The following products (per Section 5) are defined at this level of WBS:

- Matching to reference catalogs
- DIAObject association

B.3.4 1.02C.03.03: Alert Distribution System

This WBS element covers the construction of the software, together with its associated configuration files, test suites and documentation, which will make up the alert distribution system. DIAObjects and ancillary data necessary for alert packet assembly will be delivered to this system. There are three parts to the alert distribution system:

- Robust, redundant message queue – DIAObjects and ancillary data will be delivered to the message queue by the alert generation pipeline.
- Flexible stream filtering system – Will operate on the packaged alert stream. This will provide the interface to both community broker and to the minimal LSST provided filtering system.
- Alert database – All alerts will be dumped (possibly verbatim) to a database that can be replayed later from any point in the stream.

The following products (per Section 5) are defined at this level of WBS:

- Alert generation pipeline
- Offline alert generation pipeline
- Alert distribution service
- Alert filtering service
- Alert database

B.3.5 1.02C.03.04: Alert Generation Pipeline

This WBS element covers the construction of the software, together with its associated configuration files, test suites and documentation, which is used to produce all information necessary to build the LSST alert stream. This will include work in the areas of:

- Template optimization – An example area of optimization is that of reducing false positives due to the presence of differential chromatic refraction in the data used to build the templates.
- Image differencing – Production of algorithms needed to produce optimal image differences in all contexts LSST expects to operate in (i.e. both low and high galactic latitude).
- Difference image measurement – Dipole, point source (positive and negative), and trailed source measurement are all needed.

The following products (per Section 5) are defined at this level of WBS:

- Precorrection and forced photometry pipeline
- Template generation payload
- Difference template storage/retrieval
- DCR-corrected template generation

B.3.6 1.02C.03.05: Tools for Science Pipelines

This WBS element covers the construction of the software, together with its associated configuration files, test suites and documentation, which are used to provide tooling, software primitives, and software upkeep necessary to execute the science pipelines successfully.

The following products (per Section 5) are defined at this level of WBS:

- Cartesian geometry
- Coordinate transformations
- Chromaticity utilities
- Interpolation and approximation of 2-D fields
- Common functions and source profiles
- Camera descriptions
- Point-spread functions
- Fourier transforms
- Tree structures (for searching)

B.3.7 1.02C.03.06: Moving Object Processing System (MOPS)

This WBS element covers the construction of the software, together with its associated configuration files, test suites and documentation, which are used to produce the moving object processing system. This system is responsible for producing high quality orbital properties for solar system objects observed by LSST.

The following products (per Section 5) are defined at this level of WBS:

- Moving object pipeline
- Ephemeris calculation
- Tracklet generation
- Attribution and precovery
- Orbit fitting
- Orbit merging

B.3.8 1.02C.03.07: Transform fitting on stacks of images

This WBS element covers the construction of the software, together with its associated configuration files, test suites and documentation, which will allow for high precision photometric and astrometric calibration of objects using multiple apperitions of each object.

Major features of this work will be:

- Photometric zeropoint fitting on scales larger than a chip.
- Support photometric zeropoints which vary on scales smaller than a chip.
- Support fitting composable astrometric models that can represent distortions on a variety of scales from multiple different sources.
- The above will allow the distortions from the optical system to be separated from the distortions imprinted by the atmosphere.

The following products (per Section 5) are defined at this level of WBS:

- Astrometric fitting

B.3.9 1.02C.03.08: Integration

This WBS element covers the construction of the software, together with its associated configuration files, test suites and documentation, necessary for the integration of the alert production system and MOPS. This includes pipeline assembly, integration with the workflow system, definition and implementation of interfaces with other subsystems, and documentation of the system to a level where it can be operated by non-Ap team members.

The following products (per Section 5) are defined at this level of WBS:

- Photometric fitting

B.4 1.02C.04: Data Release Production

This WBS element covers three broad areas of work:

- The development of scientific algorithms and pipelines which will be used to generate LSST's annual data releases;
- The development of algorithms and pipelines which will be used to generate the calibration products required in both nightly and annual data processing;
- The development of reusable algorithmic and software primitives which will be used in the construction of both nightly and annual data processing pipelines.

No products are defined at this level of the WBS.

B.4.1 1.02C.04.00: Management & Leadership

This WBS element covers project management and scientific leadership of the Data Release Production group. It includes effort to develop the overall plan for the work of the group, to schedule that work, to perform day-to-day technical and control account management of the team, to coordinate development with other parts of the LSST WBS, and to contribute to the operation of the DM Subsystem Science Team.

No products are defined at this level of the WBS.

B.4.2 1.02C.04.01: Software Primitives

This WBS element covers the construction of low-level, re-usable software primitives which form the core libraries underlying the LSST Science Pipelines. It includes the production of test suites demonstrating the correct operation of these primitives and technical, developer-focused documentation describing their use.

The following products (per Section 5) are defined at this level of WBS:

- Images
- Tables
- Footprints
- Basic statistics
- Photometric calibration representation
- Convolution kernels
- Numerical integration
- Random number generation
- Numerical optimization
- Monte Carlo sampling
- Warping
- Science tools

B.4.3 1.02C.04.02: Calibration Products

This WBS element covers the construction of the software, together with its associated configuration files, test suites and documentation, which is used to produce the LSST calibration products. Calibration products are used in the LSST Science Pipelines (both Alert Production and Data Release Production) to:

- Characterize detector anomalies;
- Correct for sensor cross-talk;
- Perform photometric calibration through understanding the throughput of the LSST system and the transmissivity of the atmosphere.

The following products (per Section 5) are defined at this level of WBS:

- Prompt Processing raw calibration validation payload
- OCS control scripts for collimated beam projector control
- Offline Auxiliary Telescope spectrograph pipeline
- Offline calibration single frame processing pipeline
- OCS-Controlled batch daily calibration update payload
- Periodic CPP payload
- Annual CPP payload
- CPP QC measurement generators

B.4.4 1.02C.04.03: Image Characterization

This WBS element covers the construction of the software, together with its associated configuration files, test suites and documentation, which is used to characterize and calibrate each exposure as part of the Data Release Production processing. This will include:

- Estimation of the point spread function;
- Modelling the background;
- Developing astrometric and photometric calibration solutions.

The following products (per Section 5) are defined at this level of WBS:

- Image characterization and calibration
- Background estimation
- Background reference
- PSF estimation (visit)

B.4.5 1.02C.04.04: Coaddition

This WBS element covers the construction of the software, together with its associated configuration files, test suites and documentation, which are used to generate co-added and differenced images as part of the Data Release Production processing.

The following products (per Section 5) are defined at this level of WBS:

- Image coaddition and differencing
- PSF matching
- Image coaddition

B.4.6 1.02C.04.05: Detection & Deblending

This WBS element covers the construction of the software, together with its associated configuration files, test suites and documentation, which are used to detect sources on astronomical images, and to decompose detections which consist of multiple overlapping astronomical

objects into their constituent parts (“deblending”). It also includes functionality to merge redundant processing carried out in the overlapping regions of the LSST sky tessellation.

The following products (per Section 5) are defined at this level of WBS:

- Coadd processing
- Overlap resolution
- Source detection
- Deblending
- Measurement
- Aperture correction
- Star/galaxy classification
- Association and matching

B.4.7 1.02C.04.06: Characterization & Measurement

This WBS element covers the construction of the software, together with its associated configuration files, test suites and documentation, which are used to characterize objects detected in LSST images. This includes:

- Application of fundamental measurement algorithms;
- Fitting of source models to data;
- Multi-epoch measurement;
- Forced measurement;
- Star/galaxy classification.

The following products (per Section 5) are defined at this level of WBS:

- Multi-epoch object characterization
- DRP Postprocessing

B.4.8 1.02C.04.07: Maintenance & Documentation

This WBS element covers holistic documentation, verification and maintenance tasks that pertain to the pipelines and algorithms developed in the other 02C.04 WBS elements and elsewhere in the project. It includes work to construct and maintain an end-to-end Data Release Production test system and ongoing maintenance to adapt the pipelines to changes elsewhere in the system. It also includes providing high-level, scientist and operator facing documentation describing the the system as delivered.

No products are defined at this level of the WBS.

B.5 1.02C.05: Science User Interface and Tools

No products are defined at this level of the WBS.

B.5.1 1.02C.05.00: Management, Leadership, & Other LOE

This WBS element covers Project Management, Control Account Management, and Scientific Leadership of the Science User Interface and Tools group. It includes effort to develop the overall schedule and cost for the work, perform day-to-day technical management of the team, coordinate with other parts of the WBS, provide support for invoices, contribute to the operation of the DM Subsystem Science Team, and LOE (for meetings ...) for team members.

No products are defined at this level of the WBS.

B.5.2 1.02C.05.06: Client-server Query & Visualization Framework

This WBS element covers the construction of low-level, re-usable software components which form the core libraries (Firefly) underlying both the client and server sides of the LSST Science User Interface Portal. It includes software packaging, releases, and deployment, the

production of test suites demonstrating the correct operation of the components, and the development of technical, developer-focused documentation describing their use.

Firefly comprises a variety of capabilities, including the following:

- Libraries for data display and visualization for tabular data and astronomical images, including various 2D charts, and a shared data model supporting data overlays on images, and brushing and linking among related displays;
- Abstract search processor interface;
- Libraries for data query, retrieval, and export, including an abstract search processor interface as well as implementations providing common astronomical archive search forms and interfaces to standard (including VO) data query APIs;
- Support for various common formats for astronomical tabular and image data;
- JavaScript and low-level Python APIs, used internally as well as providing for user control and customization;
- Identity and preferences management;
- Load balancing.

The following products (per Section 5) are defined at this level of WBS:

- LSST-independent Firefly framework and visualization capabilities

B.5.3 1.02C.05.07: LSST Science Platform Interfaces

This WBS element covers the construction of software components that provide the connection between the core data search/visualization components and the rest of the "LSST Science Platform": supporting the LSST data model, data-access services, and compute and storage resource access. It includes the development of associated documentation and test suites.

It incorporates the following:

- Interface to LSST-specific user identity and management services;
- Access to LSST user workspace;
- Search processors matching the DAX interfaces;
- Search processors for the Engineering and Facilities Database;
- Interface for invoking LSST stack Python code to perform services needed by the Portal;
- Support for reading and displaying LSST-specific data formats (such as afw.table-format files) and data objects (e.g., masks, Footprints, PSF models);
- Support for the afw.display interface to Firefly.

The following products (per Section 5) are defined at this level of WBS:

- JupyterLab visualization widgets and other JupyterHub/Portal bridges
- Low-level Python API to Firefly
- Interfaces to DAX, user workspace, SuperTask, identity management
- Firefly components to visualize LSST Science Pipelines data objects

B.5.4 1.02C.05.08: Applications

This WBS element covers the construction of the Portal Aspect of the LSST Science Platform software, including online help, deployment instructions, and other documents. The Portal uses the core Firefly components and the LSST-specific software interfaces in 02C.05.07, relying on infrastructure and data access services provided under other WBS elements, to deliver a portal for the science community to access, discover, explore, analyze, and download the LSST data. It covers the following:

- The overall user interface layout and structure;
- Basic access to all LSST catalog and image data (Level 1, Level 2, user/Level 3, calibration, and Engineering and Facilities Database);

- Scientifically motivated workflows guiding users to the available data and illuminating connections among tables and between tables and images;
- All-sky displays allowing exploration of the image data and of maps of the properties of the LSST survey across the sky;
- Access to the user workspace, supporting data sharing and and collaboration;
- Deployment packaging and configuration management.

The following products (per Section 5) are defined at this level of WBS:

- Web application(s) implementing the Portal

B.6 1.02C.06: Science Data Archive and Application Services

The Science Data Archive and Data Access Services provides the ability to ingest, index, federate, query, and administer DM data products on distributed, heterogeneous storage systems and data servers. All services will be implemented to provide reasonable fault-tolerance and autonomous recovery in the event of software and hardware failures.

No products are defined at this level of the WBS.

B.6.1 1.02C.05.09: Alert Interfaces

This WBS element covers the construction of a Web-based user interface for alert subscription and searches. It is based on the core Firefly package and uses interfaces provided by SUIT, DAX, and AP. It covers:

- Alert subscription, setting filters and alert stream destination;
- Access to user management system;
- Alert searches and filtering.

The following products (per Section 5) are defined at this level of WBS:

- Portal alert interfaces to configure alert subscriptions

B.6.2 1.02C.05.10: Integration & Test

This WBS element covers the integration and higher-level testing of all the deliverables from the 02C.05 sub-WBSs, including manual UI tests, tracking and reporting bugs in the software from other teams, possible automated UI test setup and instructions, and deployment of the software at the LSST Data Access Centers in collaboration with NCSA.

No products are defined at this level of the WBS.

B.6.3 1.02C.06.00: Management, Leadership, and Other LOE

This WBS element covers Project Management, Control Account Management, and technical leadership of the Science Data Archive and Data Access Services group. It includes effort to develop the overall schedule and cost for the work, perform day-to-day technical management of the team, coordinate with other parts of the WBS, provide support for invoices, and LOE (meetings, etc.) for team members.

No products are defined at this level of the WBS.

B.6.4 1.02C.06.01: Science Data Archive

This WBS element is a summary element that includes software programs, database tables, configuration files, unit tests, component integration tests, and documentation that implement structures and tools to manage the LSST Data Products in databases and files, including defining schemas and ingesting tables and files and their metadata and provenance into the archive.

No products are defined at this level of the WBS.

B.6.4.1 1.02C.06.01.01: Database Catalogs, Alerts, and Metadata This WBS element includes software programs, database tables, configuration files, unit tests, component integration tests, and documentation that implement Database Catalogs, Alerts, and Metadata

capabilities.

It implements all database catalogs: L1 Alert Production and User Database, L2 Internal DRP, L2 Data Release Catalogs, Level 3 Catalogs, Calibration Database, Restructured Engineering Facilities Database, and Deep Drilling Database. It includes schemas and structures (partitioning, replication, distribution models, L1 production/user, L2 swap/release), and tools for manipulating the catalogs, such as managing ingest, replication, hot swap, recovery and import/export.

It implements data-product-specific metadata and provenance for all LSST data product catalogs and images. It includes schemas and structures (partitioning, replication, distribution models), and tools for manipulating the metadata, such as managing ingest, replication, hot swap, recovery and import/export. Global metadata that spans multiple data products is handled through 02C.06.02.05.

It implements structure for alerts, as well as tools for persisting and efficiently retrieving them.

The following products (per Section 5) are defined at this level of WBS:

- L1 catalog database
- L2 catalog database
- Provenance database
- DRP-internal database

B.6.4.2 1.02C.06.01.02: Image and File Archive This WBS element includes software programs, database tables, configuration files, unit tests, component integration tests, and documentation that implement the Image and File Archive, including tools for managing image and files (ingestion, import/export).

No products are defined at this level of the WBS.

B.6.5 1.02C.06.02: Data Access Services

This WBS element is a summary element that includes software programs, database tables, configuration files, unit tests, component integration tests, and documentation that implement Data Access Service capabilities.

No products are defined at this level of the WBS.

B.6.5.1 1.02C.06.02.01: Data Access Client Framework This WBS element includes software programs, database tables, configuration files, unit tests, component integration tests, and documentation that implement overall Client Framework for Data Access Services.

It implements capability to store and retrieve LSST Data Products in terms of their application level "astronomical" semantics, mapping those semantics to physical, persistent versions of those data products in databases and files.

It provide capabilities to run user data analysis close to the data.

The following products (per Section 5) are defined at this level of WBS:

- Data Butler data access client library
- Multi-type associative containers

B.6.5.2 1.02C.06.02.02: Web Services Framework This WBS element includes software programs, database tables, configuration files, unit tests, component integration tests, and documentation that implement an overall framework for running database, metadata and image cutout services.

The work involves providing IVOA standard service interfaces where applicable.

The following products (per Section 5) are defined at this level of WBS:

- Web services framework

B.6.5.3 1.02C.06.02.03: Query Services This WBS element includes work needed to come up with a DBMS that meets LSST user query analysis needs. Such DBMS should include standard off-the-shelf DBMS capabilities including advanced features such as scalability to petabytes, incremental scaling, parallel queries, shared scans, fault tolerance, resource management, as well as LSST-specific features such as efficient support for spatial and temporal data at scale.

The work involves customizing, optimizing, improving and gluing together relevant components, building missing features, configuration files, unit tests, integration tests, and documentation. It also includes building tools for maintaining, configuring, and administering such system.

The following products (per Section 5) are defined at this level of WBS:

- Qserv distributed database system

B.6.5.4 1.02C.06.02.04: Image and File Services This WBS element includes software programs, database tables, configuration files, unit tests, component integration tests, and documentation that implement image and file services. Image and File Services provide the capability to access and manipulate image and file-based data, manage file caches, and recreate images on demand.

The following products (per Section 5) are defined at this level of WBS:

- Image access

B.6.5.5 1.02C.06.02.05: Catalog Services This WBS element includes software programs, database tables, configuration files, unit tests, component integration tests, and documentation needed to build web services on top of all LSST database products (all levels, all metadata).

It includes work on global metadata structures for all LSST data products, including all data releases, L3 user data and all images. Data-product-specific metadata is handled through 02C.06.01.01.

The following products (per Section 5) are defined at this level of WBS:

- Global metadata service
- Catalog access
- Image metadata access
- Catalog metadata access

B.6.6 1.02C.06.03: Task Framework

This WBS element includes software programs, database tables, configuration files, unit tests, component integration tests, and documentation needed to build the Task Framework.

The Task Framework is a Python class library that provides a structure (standardized class entry points and conventions) to organize low-level algorithms into potentially-reusable algorithmic components (Tasks; e.g. dark frame subtraction, object detection, object measurement), and to organize tasks into basic pipelines (SuperTasks; e.g., process a single visit, build a coadd, difference a visit). The Task Framework allows the pipelines to be constructed, configured, and run at the level of a single node or a group of tightly-synchronized nodes. In addition to multi-node Tasks, it also allows for sub-node parallelization across multiple cores.

Pipeline configuration includes configuring parameters for scientific algorithms, allowing overrides of defaults based on camera/survey, computing environment, or user choice. It also includes configuration of debugging capabilities used during pipeline development.

The Task Framework serves as an interface layer between orchestration and the algorithmic code. It exposes a standard interface to "activators" (command-line runners as well as the orchestration layer and QA systems), which use it to execute the code wrapped in tasks. The Task Framework exposes to the orchestration system needs and capabilities of the underlying algorithmic code (e.g., the number of cores needed, expected memory-per-core, expected need for data). It may also receive from the orchestration layer the information on how to optimally run the particular task (i.e., which level of intra-node parallelization is desired).

This WBS includes construction of basic implementations for these components. More complex (or custom) implementations and alternative backends for the APIs and components

above (e.g., a special backend to retrieve a configuration from a central database, or a backend to send logs to a database instead of files, or a MultiCore API backend that's better aware of local machine architecture) are out of scope.

The following products (per Section 5) are defined at this level of WBS:

- SuperTask
- Activator base and Command Line Activator
- Pipeline configuration
- Multi-node Task API
- Multi-core Task API

B.6.7 1.02C.06.04: Middleware, Infrastructure, and Toolkits

This WBS element is a summary element that includes software programs, database tables, configuration files, unit tests, component integration tests, and documentation that implement basic middleware infrastructure primitives.

No products are defined at this level of the WBS.

B.6.7.1 1.02C.06.04.01: Logging This WBS element includes software programs, database tables, configuration files, unit tests, component integration tests, and documentation that implement a set of classes/functions enabling tasks to log diagnostic messages about their execution.

The following products (per Section 5) are defined at this level of WBS:

- Logging

B.6.7.2 1.02C.06.04.02: Daf_base and Utilities This WBS element is a summary element that includes software programs, database tables, configuration files, unit tests, component

integration tests, and documentation that implement small common utility classes for memory management, key/value storage, dates and times, RA/declination formatting, etc.

No products are defined at this level of the WBS.

B.6.7.3 1.02C.06.04.03: Sphgeom This WBS element is a summary element that includes software programs, database tables, configuration files, unit tests, component integration tests, and documentation pertaining to the sphgeom spherical geometry library.

The following products (per Section 5) are defined at this level of WBS:

- Spherical geometry

C DM Discussion and Decision Making Process

The Escalation process only occurs when the issue cannot be resolved within the DM, i.e. when the following internal discussion and decision making process has failed to yield a decision.

C.1 Empowerment

All DM team members are empowered by the DM Project Manager (PM) and DM Subsystem Scientist (SS) to make decisions on any DM-internal matter, including technical/algorithm issues, process improvements, tool choices, etc., when:

- A) they are willing and able to do the work to implement the decision or with people who agree with the team member,
- B) they (collectively) are willing and able to fix any problems if it goes wrong, and
- C) they believe that all affected parties (including your immediate manager) would not seriously object to your decision and implementation.

C.2 RFC Process

If the above three criteria are not met, perhaps because the team member doesn't know all the affected parties or because they don't know their positions, the team member should publish the proposed decision and implementation as a JIRA issue in the Request For Comments (RFC) project with a component of "DM".

It is usually difficult to determine all the affected parties for published package interfaces. Changes to interfaces should thus typically go through this process.

It's a good idea to contact any known affected parties before starting this process to check that the resolution is sensible. The institutional technical manager is always affected, as she or he is responsible for tracking the work schedule. If work for others is being proposed, they are obviously affected. The institutional scientist, the DM Software Architect (SA), the DM Interface Scientist (IS), and the DM Subsystem Scientist (SS) are also valuable resources for determining affected parties.

The purpose of an RFC is to inform others about the existence and content of the proposed decision and implementation in order to allow them to evaluate its impact, comment on it, refine it if necessary, and agree (implicitly or explicitly) or object (explicitly) to its execution.

The discussion of the RFC takes place in the medium of the requestor's choosing (e.g., a specific mailing list, the RFC JIRA issue itself, a Slac Channel, a convened videocon, some combination of those, etc.), but the requestor should be open to private communications as well.

In the RFC process, the opinions of those who will be doing the work (and fixing any problems if something goes wrong) are given more weight. In some cases, this may mean that the RFC issue's Assignee passes to someone else. The opinions of more senior people or people more experienced in the area should also be given more weight and may also result in the Assignee changing.

The Assignee is responsible for determining when no serious objections remain. In particular, there is no need to call for a formal vote on the (refined) resolution. If no explicit objections have been raised within, typically, 72 hours for "ordinary" issues and 1 week for "major" issues, the Assignee should assume that there are none. This is known as "lazy consensus". When this state has been reached, the Assignee is responsible for ensuring that the final

consensus has been recorded in the RFC issue before closing it and proceeding with implementation of the decision.

The requestor must be especially careful about not making irreversible changes in the "lazy consensus" time period unless they are absolutely certain there's a general agreement on the stated course of action. If something is broken, the requestor must be ready to fix it. It is critical to apply sound reasoning and good judgment about what may be acceptable and what might be not. Mistakes will happen; accept that occasionally there will be a requirement to revert an action for which it was thought agreement existed.

C.3 Exceptions and Appeals

Some proposed resolutions may require changes to one or more of the baselined, change-controlled documents describing the Data Management system (those in DocuShare with an LDM- handle or marked as change-controlled in Confluence). Note that major changes to budget or scope will almost certainly affect one or more LDM- documents. In this case only, the DM Configuration Control Board (DMCCB) (Section 7.4) may empanel an ad hoc committee including the lead author of the document and other relevant experts. This committee or the TCT itself must **explicitly** approve the change.

Change-controlled documents with other handles, such as LSE- or LPM-, including inter-subsystem interfaces, have project-wide change control processes. Please consult the DM PM, SA, or IS for more information. At least one member of the DM TCT will read each RFC to determine if it might affect a change-controlled document.

If the DM team can't converge on a resolution to an RFC that has no serious objections but the requestor still feel that something must be done, the request will be escalated. In most non-trivial cases, they will, with the advice of the SA, empanel a group of experts to which they will delegate the right to make the decision, by voting if need be.

C.4 Formalities

For project management purposes, RFCs are formally proposals made to the DM PM and PS who by default are responsible for everything in DM (they "own" all problems). As owners, they have the final word in accepting or rejecting all proposals. Functionally, they delegate

that ownership, the right and responsibility to make decisions – to others within the team (e.g. the SA, IS, group leads, etc.) who are expected to delegate it even further. Notifying the institutional technical manager about an RFC serves to inform the DM PM.

Draft

D Traceability matrix of DMSR requirements to OSS Requirements

DMS	OSS
DMS-REQ-0002	OSS-REQ-0127, OSS-REQ-0184
DMS-REQ-0004	OSS-REQ-0127
DMS-REQ-0006	
DMS-REQ-0008	
DMS-REQ-0009	
DMS-REQ-0010	OSS-REQ-0129
DMS-REQ-0018	OSS-REQ-0114
DMS-REQ-0020	OSS-REQ-0316
DMS-REQ-0022	OSS-REQ-0114, OSS-REQ-0127
DMS-REQ-0024	OSS-REQ-0114
DMS-REQ-0029	
DMS-REQ-0030	
DMS-REQ-0032	
DMS-REQ-0033	
DMS-REQ-0034	OSS-REQ-0339
DMS-REQ-0042	
DMS-REQ-0043	
DMS-REQ-0046	
DMS-REQ-0047	
DMS-REQ-0052	
DMS-REQ-0059	
DMS-REQ-0060	
DMS-REQ-0061	
DMS-REQ-0062	
DMS-REQ-0063	
DMS-REQ-0065	
DMS-REQ-0068	OSS-REQ-0122
DMS-REQ-0069	OSS-REQ-0129
DMS-REQ-0070	
DMS-REQ-0072	
DMS-REQ-0074	
DMS-REQ-0075	
DMS-REQ-0077	
DMS-REQ-0078	
DMS-REQ-0089	

Continued on next page

DMS	OSS
DMS-REQ-0094	
DMS-REQ-0096	
DMS-REQ-0097	OSS-REQ-0131
DMS-REQ-0098	
DMS-REQ-0099	OSS-REQ-0131
DMS-REQ-0100	
DMS-REQ-0101	OSS-REQ-0131
DMS-REQ-0102	
DMS-REQ-0103	OSS-REQ-0136
DMS-REQ-0106	OSS-REQ-0122
DMS-REQ-0119	
DMS-REQ-0120	
DMS-REQ-0121	
DMS-REQ-0122	
DMS-REQ-0123	
DMS-REQ-0124	
DMS-REQ-0125	
DMS-REQ-0126	
DMS-REQ-0127	
DMS-REQ-0128	
DMS-REQ-0130	
DMS-REQ-0131	
DMS-REQ-0132	
DMS-REQ-0155	
DMS-REQ-0156	
DMS-REQ-0158	
DMS-REQ-0160	
DMS-REQ-0161	
DMS-REQ-0162	
DMS-REQ-0163	
DMS-REQ-0164	
DMS-REQ-0165	
DMS-REQ-0166	
DMS-REQ-0167	
DMS-REQ-0168	
DMS-REQ-0170	
DMS-REQ-0171	
DMS-REQ-0172	
DMS-REQ-0173	

Continued on next page

DMS	OSS
DMS-REQ-0174	
DMS-REQ-0175	
DMS-REQ-0176	
DMS-REQ-0177	
DMS-REQ-0178	
DMS-REQ-0180	
DMS-REQ-0181	
DMS-REQ-0182	
DMS-REQ-0183	
DMS-REQ-0185	
DMS-REQ-0186	
DMS-REQ-0187	
DMS-REQ-0188	
DMS-REQ-0189	
DMS-REQ-0190	
DMS-REQ-0191	
DMS-REQ-0193	
DMS-REQ-0194	
DMS-REQ-0196	
DMS-REQ-0197	
DMS-REQ-0265	
DMS-REQ-0266	OSS-REQ-0130
DMS-REQ-0267	OSS-REQ-0137
DMS-REQ-0268	OSS-REQ-0137
DMS-REQ-0269	OSS-REQ-0130
DMS-REQ-0270	OSS-REQ-0166
DMS-REQ-0271	OSS-REQ-0130
DMS-REQ-0272	
DMS-REQ-0273	OSS-REQ-0130
DMS-REQ-0274	OSS-REQ-0128
DMS-REQ-0275	OSS-REQ-0137
DMS-REQ-0276	
DMS-REQ-0277	
DMS-REQ-0278	OSS-REQ-0136
DMS-REQ-0279	OSS-REQ-0136
DMS-REQ-0280	OSS-REQ-0136
DMS-REQ-0281	OSS-REQ-0136
DMS-REQ-0282	
DMS-REQ-0283	

Continued on next page

DMS	OSS
DMS-REQ-0284	
DMS-REQ-0285	
DMS-REQ-0286	
DMS-REQ-0287	
DMS-REQ-0288	
DMS-REQ-0289	
DMS-REQ-0290	
DMS-REQ-0291	
DMS-REQ-0292	
DMS-REQ-0293	
DMS-REQ-0294	
DMS-REQ-0295	
DMS-REQ-0296	
DMS-REQ-0297	
DMS-REQ-0298	
DMS-REQ-0299	
DMS-REQ-0300	
DMS-REQ-0301	
DMS-REQ-0302	
DMS-REQ-0303	
DMS-REQ-0304	
DMS-REQ-0305	
DMS-REQ-0306	
DMS-REQ-0307	
DMS-REQ-0308	
DMS-REQ-0309	
DMS-REQ-0310	
DMS-REQ-0311	
DMS-REQ-0312	
DMS-REQ-0313	
DMS-REQ-0314	
DMS-REQ-0315	
DMS-REQ-0316	
DMS-REQ-0317	OSS-REQ-0130
DMS-REQ-0318	OSS-REQ-0373
DMS-REQ-0319	
DMS-REQ-0320	
DMS-REQ-0321	
DMS-REQ-0322	

Continued on next page

DMS	OSS
DMS-REQ-0323	
DMS-REQ-0324	
DMS-REQ-0325	
DMS-REQ-0326	
DMS-REQ-0327	
DMS-REQ-0328	
DMS-REQ-0329	
DMS-REQ-0330	
DMS-REQ-0331	
DMS-REQ-0332	
DMS-REQ-0333	
DMS-REQ-0334	
DMS-REQ-0335	
DMS-REQ-0336	
DMS-REQ-0337	
DMS-REQ-0338	
DMS-REQ-0339	
DMS-REQ-0340	
DMS-REQ-0341	
DMS-REQ-0342	
DMS-REQ-0343	
DMS-REQ-0344	
DMS-REQ-0345	
DMS-REQ-0346	
DMS-REQ-0347	
DMS-REQ-0348	
DMS-REQ-0349	
DMS-REQ-0350	
DMS-REQ-0351	

E Traceability matrix of OSS requirements to DMSR requirements

OSS	DMS
OSS-REQ-0114	DMS-REQ-0024, DMS-REQ-0018, DMS-REQ-0022
OSS-REQ-0122	DMS-REQ-0068, DMS-REQ-0106
OSS-REQ-0127	DMS-REQ-0022, DMS-REQ-0004, DMS-REQ-0002
OSS-REQ-0128	DMS-REQ-0274
OSS-REQ-0129	DMS-REQ-0069, DMS-REQ-0010
OSS-REQ-0130	DMS-REQ-0266, DMS-REQ-0269, DMS-REQ-0271, DMS-REQ-0273, DMS-REQ-0317
OSS-REQ-0131	DMS-REQ-0097, DMS-REQ-0099, DMS-REQ-0101
OSS-REQ-0136	DMS-REQ-0279, DMS-REQ-0280, DMS-REQ-0281, DMS-REQ-0278, DMS-REQ-0103
OSS-REQ-0137	DMS-REQ-0267, DMS-REQ-0275, DMS-REQ-0268
OSS-REQ-0166	DMS-REQ-0270
OSS-REQ-0184	DMS-REQ-0002
OSS-REQ-0316	DMS-REQ-0020
OSS-REQ-0339	DMS-REQ-0034
OSS-REQ-0373	DMS-REQ-0318

F References

- [1] **[LSE-130]** *List of Data Items to be Exchanged Between the Camera and Data Management*, LSE-130, URL <https://ls.st/LSE-130>
- [2] **[LSE-131]** *Data Management Interface Requirements to Support Education and Public Outreach*, LSE-131, URL <https://ls.st/LSE-131>
- [3] **[LSE-140]** *Auxiliary Instrumentation Interface between Data Management and Telescope*, LSE-140, URL <https://ls.st/LSE-140>
- [4] **[LSE-72]** *OCS Command Dictionary for Data Management*, LSE-72, URL <https://ls.st/LSE-72>

- [5] **[LSE-75]** *Control System Interfaces between the Telescope and Data Management*, LSE-75, URL <https://ls.st/LSE-75>
- [6] **[LSE-76]** *Infrastructure Interfaces between Summit Facility and Data Management*, LSE-76, URL <https://ls.st/LSE-76>
- [7] **[LSE-77]** *Infrastructure Interfaces between Base Facility and Data Management*, LSE-77, URL <https://ls.st/LSE-77>
- [8] **[LSE-159]**, Angeli, G., 2013, *Reviews Definitions, Guidelines, and Procedures*, LSE-159, URL <https://ls.st/LSE-159>
- [9] **[Document-11920]**, Angeli, G., McKercher, R., 2013, *Document Cover Page and Style Guide*, Document-11920, URL <https://ls.st/Document-11920>
- [10] **[Document-9224]**, Angeli, G., McKercher, R., 2013, *Change Controlled Document Cover Page and Style Guide*, Document-9224, URL <https://ls.st/Document-9224>
- [11] **[DMTN-020]**, Becla, J., Economou, F., Gelman, M., et al., 2016, *Project Management Guide*, DMTN-020, URL <https://dmtn-020.lsst.io/>
- [12] **[LDM-472]**, Becla, J., Economou, F., Mueller, F., et al., 2017, *LSST DM Project Management and Tools*, LDM-472, URL <https://ls.st/LDM-472>
- [13] **[LSE-17]**, Claver, C., Angeli, G., Selvy, B., 2016, *Systems Engineering Management Plan*, LSE-17, URL <https://ls.st/LSE-17>
- [14] **[LSE-29]**, Claver, C.F., The LSST Systems Engineering Integrated Project Team, 2016, *LSST System Requirements*, LSE-29, URL <https://ls.st/LSE-29>
- [15] **[LSE-30]**, Claver, C.F., The LSST Systems Engineering Integrated Project Team, 2016, *LSST System Requirements*, LSE-30, URL <https://ls.st/LSE-30>
- [16] **[LSE-69]**, Dubois-Felsmann, G., 2014, *Interface between the Camera and Data Management*, LSE-69, URL <https://ls.st/LSE-69>
- [17] **[LSE-68]**, Dubois-Felsmann, G., 2015, *Camera Data Acquisition Interface*, LSE-68, URL <https://ls.st/LSE-68>
- [18] **[LSE-61]**, Dubois-Felsmann, G., 2016, *LSST Data Management Subsystem Requirements*, LSE-61, URL <https://ls.st/LSE-61>

- [19] **[LSE-163]**, Juric, M., et al., 2017, *LSST Data Products Definition Document*, LSE-163, URL <https://ls.st/LSE-163>
- [20] **[LDM-148]**, Kantor, J., Axelrod, T., 2013, *Data Management System Design*, LDM-148, URL <https://ls.st/LDM-148>
- [21] **[LDM-138]**, Kantor, J., Axelrod, T., Lim, K.T., 2013, *Data Management Compute Sizing Model*, LDM-138, URL <https://ls.st/LDM-138>
- [22] **[LPM-20]**, Krabbendam, V., Selvy, B., 2015, *Risk & Opportunity Management Plan*, LPM-20, URL <https://ls.st/LPM-20>
- [23] Larman, C., Basili, V.R., 2003, *Computer*, 36, 47, doi:10.1109/MC.2003.1204375
- [24] **[LPM-98]**, Long, K.E., 2016, *LSST Project Controls System Description*, LPM-98, URL <https://ls.st/LPM-98>
- [25] **[Document-13760]**, McKercher, R., 2012, *Travel Request Instructions*, Document-13760, URL <https://ls.st/Document-13760>
- [26] **[LPM-51]**, McKercher, R., 2013, *Document Management Plan*, LPM-51, URL <https://ls.st/LPM-51>
- [27] **[LPM-43]**, McKercher, R., 2016, *WBS Structure*, LPM-43, URL <https://ls.st/LPM-43>
- [28] **[LPM-44]**, McKercher, R., 2016, *WBS Dictionary*, LPM-44, URL <https://ls.st/LPM-44>
- [29] **[LDM-503]**, O'Mullane, W., Jurić, M., Economou, F., 2017, *Data Management Test, Verification and Validation Plan*, LDM-503, URL <https://ls.st/LDM-503>
- [30] **[LDM-493]**, Sick, J., 2016, *Data Management Documentation Architecture*, LDM-493, URL <https://ls.st/LDM-493>
- [31] **[LPM-55]**, Sweeney, D., McKercher, R., 2013, *Project Quality Assurance Plan*, LPM-55, URL <https://ls.st/LPM-55>
- [32] **[LDM-141]**, Unknown, 2013, *Data Management Storage Sizing and I/O Model*, LDM-141, URL <https://ls.st/LDM-141>
- [33] **[LDM-144]**, Unknown, 2016, *Site Specific Infrastructure Estimation Model*, LDM-144, URL <https://ls.st/LDM-144>

G Acronyms

The following table has been generated from the on-line Gaia acronym list:

Acronym	Description
AP	Alerts Production
API	Application Programming Interface
AURA	Association of Universities for Research in Astronomy
CB	Configuration Baseline
CCB	Change Control Board
CI	Configuration Item
CIL	Configuration Item List
CM	Configuration Management
CMDB	Configuration Management DataBase
CMP	Configuration Management Plan
CU	Coordination Unit (in DPAC)
DAC	Data Access Center
DAX	Data access services
DDMPPM	Data Management Deputy Project Manager
DM	Data Management
DMCCB	DM Change Control Board
DMIS	DM Interface Scientist
DMLT	DM Leadership Team
DMPM	Data Management Project Manager
DMSR	DM System Requirements
DMSS	DM Subsystem Scientist
DMTN	DM Technical Note
DPC	Data Processing Centre
DRP	Data Release Production
EFD	Engineering Facilities Database
ICBS	International Communications and Base Site
ICD	Interface Control Document
IPAC	No longer an acronym
IS	Interface Scientist

IT	Integration Test
ITC	Information Technology Center
IVOA	International Virtual-Observatory Alliance
JIRA	issue tracking product (not an acronym, but a truncation of Gojira, the Japanese name for Godzilla)
LCR	LSST Change Request
LDF	LSST Data Facility
LDM	Light Data Management
LPM	LSST Project Management (Document Handle)
LSE	LSST System Engineering (Document Handle)
LSST	Large-aperture Synoptic Survey Telescope
LaTeX	(Leslie) Lamport TeX (document markup language and document preparation system)
NASA	National Aeronautics and Space Administration (USA)
NCSA	National Center for Supercomputing Applications
NET	Not Earlier Than
NSF	National Science Foundation
OCS	Observatory Control System
OSS	Operations Support System
PDF	Portable Document Format
PM	Project Manager
PMCS	Project Management Control System
PS	Project Scientist
PST	Project Science Team
QA	Quality Assurance
RFC	Request for Comments
SA	Science Alert(s)
SAT	Science Archives Team (at ESAC)
SEMP	System Engineering Management Plan
SLA	Service Level Agreement
SLAC	Stanford Linear Accelerator Center
SS	Subsystem Scientist
SST	Space Surveillance Telescope
SUI	Science User Interface

SUIT	Science User Interface Team
TCT	Technical Control Team (Obsolete - now DMCCB)
US	United States
WBS	Work Breakdown Structure
WCS	World Coordinate System
WISE	Wide-field Survey Explorer

Draft