



LARGE SYNOPTIC SURVEY TELESCOPE

# Large Synoptic Survey Telescope (LSST) Concept of Operations for the LSST Data Facility Services

**D. Petravick and M. Butler and M. Gelman**

**LDM-230**

**Latest Revision: 2018-07-17**

This LSST document has been approved as a Content-Controlled Document by the LSST DM Change Control Board. If this document is changed or superseded, the new document will retain the Handle designation shown above.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.



## Abstract

This document describes the operational concepts for the emerging LSST Data Facility, which will operate the system that will be delivered by the LSST construction project. The services will be incrementally deployed and operated by the construction project as part of verification and validation activities within the construction project.

## Change Record

Version	Date	Description	Owner name
1	2013-05-22	Initial release.	Kian-Tat Lim
1.1	2013-09-10	Updates resulting from Process Control and Data Products Reviews	Kian-Tat Lim
1.2	2013-10-10	TCT approved	R. Allsman
	2016-05-08	Beginning to render working group schema as more complete view of operational need as a basis for planning.	D. Petravick
	2017-06-26	Import draft versions into single document, with updates based on evolved operational concepts.	M. Gelman
	2017-12-15	Updates prior to submission to change control.	M. Gelman, D. Petravick
	2018-06-15	Updates based on comments in RFC-422	M. Butler
2.0	2018-07-17	Approved in RFC-422	W. O'Mullane

*Document source location:* <https://github.com/lstt/LDM-230>



# Contents

- 1 Scope of Document 1**
- 2 Services for Observatory Operations 1**
  - 2.1 LSSTCam Prompt Processing Services . . . . . 2
    - 2.1.1 Scope . . . . . 2
    - 2.1.2 Overview . . . . . 2
    - 2.1.3 Operational Concepts . . . . . 3
  - 2.2 LSSTCam Archiving Service . . . . . 7
    - 2.2.1 Scope . . . . . 7
    - 2.2.2 Overview . . . . . 7
    - 2.2.3 Operational Concepts . . . . . 8
  - 2.3 Spectrograph Archiving Service . . . . . 9
    - 2.3.1 Scope . . . . . 9
    - 2.3.2 Overview . . . . . 9
    - 2.3.3 Operational Concepts . . . . . 10
  - 2.4 EFD Transformation Service . . . . . 10
    - 2.4.1 Scope . . . . . 10
    - 2.4.2 Overview . . . . . 11
    - 2.4.3 Operational Concepts . . . . . 12
  - 2.5 OCS-Driven Batch Service . . . . . 14
    - 2.5.1 Scope . . . . . 14
    - 2.5.2 Overview . . . . . 14
    - 2.5.3 Operational Concepts . . . . . 15
  - 2.6 Observatory and Commissioning Operations Data Service . . . . . 15
    - 2.6.1 Scope of Document . . . . . 15
    - 2.6.2 Overview . . . . . 15
    - 2.6.3 Operational Concepts . . . . . 18
  - 2.7 Telemetry Gateway Service . . . . . 18
    - 2.7.1 Scope . . . . . 18

The contents of this document are subject to configuration control by the LSST DM Change Control Board.



2.7.2 Overview . . . . . 18

2.7.3 Operational Concepts . . . . . 19

**3 Services for Offline Campaign Processing 20**

3.1 Batch Production Services . . . . . 20

3.1.1 Scope . . . . . 20

3.1.2 Overview . . . . . 21

3.1.3 Operational Concepts . . . . . 21

**4 Data Access Services for Authorized Users 29**

4.1 User Data Access Services . . . . . 29

4.1.1 Scope . . . . . 29

4.1.2 Overview . . . . . 30

4.2 Bulk Data Export Service . . . . . 31

4.2.1 Scope . . . . . 31

4.2.2 Overview . . . . . 31

4.2.3 Operational Concepts . . . . . 32

4.3 Hosting of Feeds to Brokers . . . . . 32

4.3.1 Scope . . . . . 33

4.3.2 Overview . . . . . 33

4.3.3 Operational Concepts . . . . . 35

**5 Data, Compute, and IT Security Services 37**

5.1 Data Backbone Services . . . . . 37

5.1.1 Scope . . . . . 37

5.1.2 Overview . . . . . 38

5.2 Managed Database Services . . . . . 40

5.2.1 Scope . . . . . 40

5.2.2 Overview . . . . . 41

5.2.3 Operational Concepts . . . . . 42



- 5.3 Batch Computing and Data Staging Environment Services . . . . . 42
  - 5.3.1 Scope . . . . . 42
  - 5.3.2 Overview . . . . . 42
  - 5.3.3 Operational Concepts . . . . . 44
- 5.4 Containerized Application Management Services . . . . . 44
  - 5.4.1 Scope . . . . . 44
  - 5.4.2 Overview . . . . . 44
  - 5.4.3 Operational Concepts . . . . . 45
- 5.5 Network-based IT Security Services . . . . . 46
  - 5.5.1 Scope . . . . . 46
  - 5.5.2 Overview . . . . . 46
  - 5.5.3 Operational Concepts . . . . . 48
- 5.6 Authentication and Authorizations Services . . . . . 51
- 6 ITC Provisioning and Management . . . . . 51**
- 7 Service Management and Monitoring . . . . . 52**
  - 7.1 Service Management Processes . . . . . 52
    - 7.1.1 Scope . . . . . 52
    - 7.1.2 Overview . . . . . 52
    - 7.1.3 Operational Concepts . . . . . 53
  - 7.2 Service Monitoring . . . . . 53
    - 7.2.1 Scope . . . . . 54
    - 7.2.2 Overview . . . . . 54
    - 7.2.3 Operational Concepts . . . . . 56
- A References . . . . . 60**
- B Acronyms . . . . . 60**

# Concept of Operations for the LSST Data Facility Services

## 1 Scope of Document

This document describes the operational concepts for the emerging LSST Data Facility, which will operate the data management system as a set of services that will be delivered by the LSST construction project. These services will be incrementally stood up and operated by the construction project as part of validation and verification activities within the construction project.

## 2 Services for Observatory Operations

The LSST Data Facility provides a set of services that supports specific functions of Observatory Operations and generates prompt data products (L1)<sup>1</sup> These services include:

- A Prompt Processing Service for Alert Production for wide-field and targeted deep-drilling observing programs, including providing data support for difference image templates and calibrations, Prompt Products databases, interaction with the alert-to-broker distribution subsystem, and providing feedback to observing systems. Prompt Processing for science images and calibration images are separate services, but will be performed by the same system with different configurations.
- A Prompt Processing Service for assessing the quality of calibration exposures.
- An “Offline” Prompt Batch Processing Service, not commanded by OCS, to facilitate catch-up processing for use cases involving extensive networking or infrastructure outages, reprocessing of image parameters used by the Scheduler, pre-processing data ahead of release of a new version of Alert Production for broker training, and other emergent use cases as directed by project policy.
- An Archiving Service for acquiring raw image data from the LSST main camera and ingesting it into the Data Backbone (Section 5.1).

---

<sup>1</sup>Prompt products were previously referred to as Level 1 products and in this document Level 1 is often used in this context.

- An Archiving Service for acquiring raw data from the spectrograph on the Auxiliary Telescope and ingesting it into the Data Backbone (Section 5.1).
- An Extract, Transform, and Load (ETL) Service for data stored in the Engineering Facilities Database at the Observatory.
- An OCS-driven Batch Processing Service for Observatory Operations to submit batch jobs via the OCS environment either to NCSA or to the Commissioning Cluster at the Base Site.
- A Commissioning Cluster instance of the LSP allows fast and reliable access through the QA portal to recently acquired data from the Observatory instruments and other designated datasets, and ability to submit batch jobs, supporting operations at the Base Center. This endpoint is implemented as an instance of the LSST Science Platform. See LSE-319.
- An Observatory Operations Data Service that allows fast and reliable access to data recently acquired from LSST cameras and designated datasets held in the Data Backbone (Section 5.1).

The concept of operations for each of these services is described in the following sections.

## 2.1 LSSTCam Prompt Processing Services

### 2.1.1 Scope

This section describes the prompt processing of raw data acquired from the main LSST camera by the DM system.

### 2.1.2 Overview

**2.1.2.1 Description** During nightly operations, the DM system acquires images from the main LSST camera as they are taken, and promptly processes them with codes specific to an observing program.



**2.1.2.2 Objective** The LSSTCam Prompt Processing Services provide timely processing of newly acquired raw data, including characterization of image quality of images, alert processing and delivery, returning image parameters to the Observatory, and populating the Prompt Products Database.

**2.1.2.3 Operational Context** Prompt Processing is a service provided by the LSST Data Facility as part of the prompt processing system. It is presented to Observatory Operations as an OCS-commandable device. The Prompt Processing Service retrieves crosstalk-corrected pixel data from the main LSST camera at the Base Center, builds images, and sends them to NCSA for prompt processing.

## 2.1.3 Operational Concepts

### 2.1.3.1 Normal Operations

**2.1.3.1.1 Science Operations** Science data-taking occurs on nights when conditions are suitable. For LSST, this means all clear nights, even when the full moon brightens the night sky. Observing is directed by an automated scheduler. The scheduler considers observing conditions, for example, the seeing, the phase of the moon, the atmospheric transparency, and the part of the sky near the zenith. The scheduler is also capable of receiving external alerts, for example, announcements from LIGO of a gravitational wave event. The scheduler also considers required observing cadence and depth of coverage for the LSST observing programs.

About 90% of observing time is reserved for the LSST “wide-fast-deep” program. In this program, observations will be on the wide-field two-image-per-visit cadence, in which successive observations(only for the pair) will be in the same filter with no slew of the telescope. However, a new program, potentially with a new filter, a larger slew, a different observing cadence, or a different visit structure, can be scheduled at any moment for Special Programs. Alternative Science Visits are supposed to be per night.

Another envisioned program is “deep drilling”, where many more exposures than the two exposure visit will be taken. Deep Drilling is just one example of a Special Program.

In practice, science data-taking will proceed when conditions are suitable. Calibration data may be taken when conditions are not suitable for further observations, with science data-taking resuming when conditions again become suitable. Prompt Processing can be stopped when calibration data is being taken.

It follows that the desired behavior for science data-taking operations is to start the Prompt Processing system at the beginning of the night and to turn off the system after all processing for all science observations is finished.

The operational framework for observing discloses predictions for future exposures but they cannot be relied on until the “next visit”.

During science data-taking the Prompt Processing Service computes and promptly returns quality parameters (referred to as one of the forms of “telemetry”) to the observing system. The QA parameters are not specific to an observing program; examples are seeing and pointing corrections derived from the WCS. These parameters are not strictly necessary for observing to proceed – LSST can observe autonomously at the telescope site, if need be. Also note that the products are quality parameters, not the the “up-or-down” quality judgment.

The scheduler may be sensitive to a subset of these messages and may decide on actions, but a detailed description is TBD. The scheduler can make use of these parameters even if delivered quite late, since the scheduler uses historical data in addition to recent data.

The Prompt Processing system also executes code that is specific to an observing program. For science exposures, the code is divided into a front-end – that is able to compute the parameters sent back to the observatory with single frame processing, and a back end , Alert Production (AP), is the specific science code that detects transient objects.

The detected transients are passed off to another service, which records the alert in a catalog which can be queried offline and sends events to an ensemble of transient brokers. Data are transmitted to end users either via feeds from an LSST-provisioned broker or via community-provided alert brokers.

Alert Production for deep drilling and other Special Programs is TBD. Other observing programs may also include AP as a science code, or may have codes of their own.

**2.1.3.1.2 Calibration Operations** In addition to collecting data for science programs, the telescope and camera are involved in many calibration activities.

The baseline calibrations include flats and biases. Darks are anticipated. LSST has a Collimated Beam Projector Calibrations Device as well.

Nominally, a three-hour period each afternoon is scheduled for Observatory Operations to take dark and flat calibration data. As noted above, calibration data may be taken during the night when conditions are not suitable for science observations. The LSST dome is specified as being light-tight, enabling certain calibration data to be collected whenever operationally feasible, regardless of the time of day.

Although there are standard cadences for calibration operations, the frequency of calibration data-taking is sensitive to the stability of the camera and telescope. Certain procedures, such as replacement of a filter, cleaning of a mirror, and warming of the camera, may subsequently require additional calibration operations. In general, calibration operations will be modified over the lifetime of the survey as understanding of the LSST telescope and camera improves.

The Prompt Processing Service computes and promptly returns quality parameters (referred to as one of the forms of “telemetry”) to the observing system. Note that the quality of calibrations needed for Prompt Processing science operations may be less stringent than calibrations needed for other processing, such as DRP.

An operations strawman, which illuminates the general need for prompt processing, is that there are two distinct, high-level types of calibrations.

- Nightly flats, biases and darks consist of a sequence of broad-band flatfield exposures in each camera filter, bias frames acquired from rapid reads of an un-illuminated camera, and optional dark images acquired from reads of an un-illuminated camera at the cadence of the expected exposure time. Observers will consider the collection of these nightly calibrations typically carried out prior to the start of nightly observing. The Prompt Processing system computes parameters for quality assessment of these calibration data, and returns the QA parameters to the observing system. Examples of defects that could be detected are the presence of light in a bias exposure and a deviation of a flat field from the norm, indicating a possible problem with the flat-field screen or its illumination. The sequence is considered complete when processing (which nec-

essarily lags acquisition of the pixel data) is finished or aborted.

- Narrow-band flats and calibrations involving the collimated beam projector help determine the response of the system, as a function of frequency, over the optical surfaces. The process of collecting these calibrations is lengthy; the bandpass over all LSST filters (760 nm) is large compared to the 1 nm illumination source, and operations using the CBP must be repeated many times as the device is moved to sample all the optical paths in the system. The length of time needed to collect these calibrations leads to the requirement that the Prompt Processing system be available during the day.

Time for an absolutely dark dome, which is important for these calibrations, is subject to an operational schedule. This schedule needs to provide for maintenance and improvement projects within the dome. These calibrations may be taken on cloudy nights or any other time. Because these operations are lengthy, and time to obtain the calibrations quite possibly precious, prompt processing is needed to run quality control codes to help assure that the data are appropriate for use. Note, the prompt processing system will not be used to construct these master calibrations.

Consideration of the lengthy calibrations, and the complexity of scheduling them means that the system must be reasonably available when needed.

### 2.1.3.2 Operational Scenarios

**2.1.3.2.1 Code Performance Problems** The Prompt Processing system provides a number of policies (which are TBD) to Observatory Operations that can be selected via the OCS interface. These policies are used to prioritize the need for prompt production of QA versus completeness of science processing, decide the conditions when science processing should abort due to timeout, and determine how any backlog of processing (including backlogs caused by problems with international networking) is managed. The policies may need to be sensitive to details of the observation sequence.

**2.1.3.2.2 Offline Backup** When needed, all of this processing, including both generating QA parameters and running science codes, can be executed under offline conditions at a

later time. The products of this processing may still be of operational or scientific value even if they are not produced in a timely manner. Considering Alert Production, for example, while alerts may not be transmitted in offline conditions, transients can still be incorporated into the portion of the Prompt Products Database that records transients. QA image parameters used to gauge whether an observation meets quality requirements can still be produced and ingested for use by the scheduler.

**2.1.3.2.3 Change Control** Upgrades to the LSSTCam Prompt Processing Services are produced in the LSST Data Facility. Change control of this function is coordinated with the Observatory, with the Observatory having an absolute say about insertion and evaluation of changes.

## 2.2 LSSTCam Archiving Service

### 2.2.1 Scope

This section describes the concept of operations for archiving designated raw data acquired from the main LSST camera to the permanent archive.

### 2.2.2 Overview

**2.2.2.1 Description** The LSSTCam Archiving Service acquires pixel and header data and arranges for the data to arrive in the Observatory Operations Data Server and in the Data Backbone.

**2.2.2.2 Objective** The objective of this system is to acquire designated raw data from the LSST main camera and header data from the OCS system, and to place appropriately formatted data files in the Data Backbone. The service needs to have the capability of archiving at the nominal frame rate for the main observing cadence, and to perform “catch up” archiving at twice that rate.

**2.2.2.3 Operational Context** LSSTCam Archiving is a service provided by LDF as part of the prompt processing system. It is presented to Observatory Operations as an OCS-commandable

device. The archiving system operates independently from related observatory services, such as data acquisition, as well as other prompt services, such as prompt processing. However, a normal operational mode is operation of the service such that data are ingested promptly into the permanent archive and into the Observatory Operations Data Server for fast-access by Observatory Operations staff.

### 2.2.3 Operational Concepts

**2.2.3.1 Normal Operations** The LSSTCam Archiving Service runs whenever it is needed. Operational goals are to provide prompt archiving of camera data and to provide expeditious catch-up archiving after service interruptions.

LSSTCam data is, by default, ingested into the permanent archive and into the Observatory Operations Data Server. However, while all science and calibration data from the main camera require ingest into the Observatory Operations Server, some data (e.g., one-off calibrations, engineering data, smoke test data, etc.) may not require archiving in the Data Backbone permanent store. Observatory Operations may designate data which will not be archived.

#### 2.2.3.2 Operational Scenarios

**2.2.3.2.1 Delayed Archiving** In delayed archiving, Observatory Operations may need to prioritize the ingestion of data into the archiving system based on operational dependencies with the Observatory Operations Data Service. The archiving service provides a number of policies (which are TBD) to Observatory Operations that can be selected via the OCS interface in order to prioritize data ingestion.

Other operational parameters of interest include rate-limiting when network bandwidth is a concern.

**2.2.3.2.2 Change Control** Upgrades to the LSSTCam Archiving Service are produced in the LSST Data Facility. Change control of this function is coordinated with the Observatory, with the Observatory having an absolute say about insertion and evaluation of changes.

## 2.3 Spectrograph Archiving Service

### 2.3.1 Scope

This section describes the concept of operations for archiving raw data acquired from instruments on the Auxiliary Telescope to the permanent archive.

### 2.3.2 Overview

**2.3.2.1 Description** The Auxiliary Telescope is a separate telescope at the Summit site, located on a ridge adjacent to that of the main telescope building. This telescope supports a spectrophotometer that measures the light from stars in very narrow bandwidths compared to the filter pass bands on the main LSST camera. The purpose of the spectrophotometer is to measure the absorption, which is how light from astronomical objects is attenuated as it passes through the atmosphere. By pointing this instrument at known “standard stars” that emit light with a known spectral distribution, it is possible to estimate the absorption. This information is used to derive photometric calibrations for the data taken by the main telescope.

The Auxiliary Telescope camera produces 2-dimensional CCD images, but the headers and associated metadata are different than the LSSTCam data because spectra, not images of the sky, are recorded.

From the point of view of LSST Data Facility Services for Observatory Operations, the spectrograph on the Auxiliary Telescope is an independent instrument that is controlled independently from the main LSSTCam. Thus, the operations of and changes to LSST Data Facility services for this instrument must be decoupled from all others.

The Auxiliary Telescope and its spectrograph are devices under the control of the observatory control system (OCS). The spectrograph contains a single LSST CCD. The Camera Data System (CDS) for the single CCD in the spectrograph uses a readout system based on the LSSTCam electronics and will present an interface for the Archiver to build files.

**2.3.2.2 Objective** The Spectrograph Archiving Service reads pixel data from the Spectrograph version of the CDS and metadata available in the overall Observatory Control System

and builds files. The service archives the data in a way that the data are promptly available to Observatory Operations via the Observatory Operations Data Service, and that the data appear in the Data Backbone.

### 2.3.3 Operational Concepts

Archiving is under control of OCS, with the same basic operational considerations as the CCD data from LSSTCam. Keeping in mind the differences between the two systems, the concept of operations for LSSTCam archiving apply (see section on LSSTCam Archiving Service). One differing aspect could be is that these data are best organized temporally, while some data from LSSTCam are organized spatially.

There is no prompt processing of Spectrograph data in a way that is analogous to the prompt processing of LSSTCam data.

**2.3.3.1 Normal Operations** Under normal operations the Spectrograph Archiving Service is under control of the Observatory Control System.

#### 2.3.3.2 Operational Scenarios

**2.3.3.2.1 Change Control** Upgrades to the Spectrograph Archiving Service are produced in the LSST Data Facility. Change control of this function is coordinated with the Observatory, with the Observatory having an absolute say about insertion and evaluation of changes.

## 2.4 EFD Transformation Service

### 2.4.1 Scope

The Engineering and Facility Database (EFD) See LTS-210 is a system used in the context of Observatory Operations. It contains data, apart from pixel data acquired by the prompt archiving systems, of interest to LSST originating from any instrument or any operation related to observing. The EFD is an essential record of the activities of Observatory Operations. It contains



data for which there is no substitute, as it records raw data from supporting instruments, instrumental conditions, and actions taking place within the observatory.

This section describes the concept of operations for ingesting the EFD data into the LSST Data Backbone and transforming this data into a format suitable for offline use.

## 2.4.2 Overview

**2.4.2.1 Description** The Original Format EFD, maintained by Observatory Operations, is conceived of as a collection of files and approximately 20 autonomous relational database instances, all using the same relational database technology. The relational tables in the Original Format EFD have a temporal organization. This organization supports the need within Observatory Operations to support high data ingest and access rates. The data in the Original Format EFD is immutable, and will not change once entered.

The EFD also includes a large file annex that holds flat files that are part of the operational records of the survey.

**2.4.2.2 Objective** The prime motivation behind the EFD Transformation Service is to be able to relate the time series data to raw images, and to hold these quantities in a manner that is accessible using the standard DM methods for file and relational data access.

The baseline design calls for all of the EFD relational and flat-file material to be ingested into what is called the Transformed EFD.

1. There is a need to access a substantial subset of the Original Format EFD data in the general context of data release production and in the Data Access Centers. This access is supported by a query-access optimized, separately implemented relational database, generally called the Transformed EFD. A prime consideration is to relate the time series data to raw images.
2. To be usable in offline context, files from the Original Format EFD need to be ingested into the LSST Data Backbone. This ingest operation requires provenance and metadata associated with these files.

3. Because the Original format EFD is the fundamental record related to instrumentation, the actions of observers, and related data, the data contained within it cannot be re-computed, and in general there is no substitute for this data. Best practice for disaster recovery is to not merely replicate the Original Format EFD live environment, but also to make periodic backups and ingests to a disaster recovery system.

**2.4.2.3 Operational Context** Ingest of data from the Original Format EFD into the Transformed EFD must be controlled by Observatory Operations, based on the principle that Observatory Operations controls access to the Original Format EFD resources. The prime framework for controlling operations is the OCS system. Operations in this context will be controlled from the OCS framework.

**2.4.2.4 Risks** The query load applied by EFD Transformation Service on the Original Format EFD at the Base Center may be disruptive to the primary purpose, serving Observatory Operations.

## 2.4.3 Operational Concepts

### 2.4.3.1 Normal Operations

**2.4.3.1.1 Original Format EFD Operations** Observatory Operations is responsible for Original Format EFD operations during LSST Operations. Observatory Operations will copy database state and related operational information into a disaster recovery store at a frequency consistent with a Disaster Recovery Plan approved by the LSST ISO. The LSST Data Facility will provide the disaster recovery storage resource. The DR design procedure should consider whether normal operations may begin prior to a complete restore of the Original Format EFD.

If future operations of the LSST telescope beyond the lifetime of the survey do not provide for operation and access to the Original Format EFD, the LSST Data Facility will assume custody of the Original Format EFD and arrange appropriate service for these data (and likely move the center of operations to NCSA) in the period of data production following the cessation of LSST operations at the Summit and Base Centers.

LSST staff working on behalf of any operations department will have access to the Original Format EFD at the Base Center for one-off studies, including studying the merits of data being loaded into the Transformed EFD.

**2.4.3.1.2 EFD Large File Annex Handling and Operations** Under control of the EFD Transformation Service, the LSST Data Facility will ingest the designated contents of the file annex of the Original Format EFD into the data backbone. The LSST Data Facility will arrange that these files participate in whatever is developed for disaster recovery for the files in the Data Backbone.

These files will also participate in the general file metadata and file-management service associated with the Data Backbone, and thus be available using I/O methods of the LSST stack.

#### **2.4.3.1.3 Transformed EFD Operations**

- The Transformed EFD is replicated to the US DAC and the Chilean DAC.
- LDF will extract, transform and load into the Transformed EFD pointers to files that have been transferred from the EFD large file annex into the Data Backbone.
- LDF will extract, transform and load all tabular data from the Original Format EFD into the Transformed EFDs residing in the Data Backbone at NCSA and the Base Center. Information in Transformed EFD that is private needs to be blocked on export.
- “Designated” data will include:
  - Any quantities used in a production process.
  - Any quantities designated by an authorized change control process.
- The information in the Transformed EFD is available to any authorized independent DAC which may choose to host a copy.

#### **2.4.3.2 Operational Scenarios**

**2.4.3.2.1 ETL Control** The Extract, Transform and Load operation is under the control of Observing Operations.

**2.4.3.2.2 Disaster Recovery and DR Testing for the Original Format EFD** Observing Operations will periodically test a restore in a disaster recovery scenario.

**2.4.3.2.3 Disaster Recovery and DR Testing for the Transformed EFD** Since the Transformed EFD relational database is reproducible from the Original Format EFD, disaster recovery is provided by a re-ingest from the original format. DR testing includes re-establishing operations of the Transformed EFD relational database and ETL capabilities from the Original Format EFD. Ingested files from the file annex can be recovered by the general disaster recovery capabilities of the Data Backbone.

**2.4.3.2.4 Change Control** Upgrades to the EFD ETL Service are produced by the LSST Data Facility. Change control of this function is coordinated with the Observatory, with the Observatory having an absolute say about insertion and evaluation of changes.

## 2.5 OCS-Driven Batch Service

### 2.5.1 Scope

The OCS-driven Batch Service provides an OCS-commandable device for Observatory Operations staff to submit batch jobs to the Commissioning Cluster, NCSA cluster, or Base cluster, and optionally return a small amount of data via the Telemetry Gateway.

### 2.5.2 Overview

The service provides a method for OCS script to invoke Batch production, allowing for the automation of batch processing into simple work flows involving LSST instrumentation.

**2.5.2.1 Description** The service provides OCS commands to initiate Batch Jobs and to optionally return with small amounts of data as telemetry.

**2.5.2.2 Objective** Primary usage is expected to be preparation of daily master calibration images. Full-frame wavefront sensing via pistoning of the Camera can also be processed in this mode. Other uses are expected during Commissioning. Examples of the type of workflows may be verification of flats or wave-front sensing application.

**2.5.2.3 Operational Context** The service is an OCS-commandable device which runs under the control of Observatory Operations.

### 2.5.3 Operational Concepts

**2.5.3.1 Normal Operations** During normal operations, the OCS scripts submit batch jobs. The subject batch systems attempt to dispatch jobs at high priority.

**2.5.3.2 Operational Scenarios** The promptness of the return of results varies according to the ability of the batch system to schedule the corresponding batch job promptly.

**2.5.3.2.1 Change Control** Upgrades to the OCS-driven Batch Service are produced by the LSST Data Facility. Change control of this function is coordinated with the Observatory, with the Observatory having an absolute say about insertion and evaluation of changes.

## 2.6 Observatory and Commissioning Operations Data Service

### 2.6.1 Scope of Document

This section describes the services provided to Observatory and Commissioning Operations to access data that satisfies the requirements that are unique to observing operations. These requirements include service levels appropriate for nightly operations.

### 2.6.2 Overview

**2.6.2.1 Description** The Observatory Operations Data Service provides fast-access to recently acquired data from Observatory instruments and designated datasets stored in the

LSST permanent archive.

**2.6.2.2 Objective** There is a need for regular and ad-hoc access to LSST datasets for staff and tools working in the context of Observatory Operations. The quality of service (QoS) needed for these data is distinct from the general availability of data via the Data Backbone. Access to data provided by the Observatory Operations Data Service is distinguished from normal access to the Data Backbone in the role the data play in providing straightforward feedback to immediate needs that support nightly and daytime operations of the Observatory. Newly acquired data is also a necessary input for some of these operations. The service must provide access methods that are compatible with the software access needs.

**2.6.2.3 Operational Context** The Observatory Operation Data Service is provided by the LSST Data Facility to Observatory Operations, and is used by observers and automated systems to access the data resident there. The service provides the availability and service levels needed to support Observatory Operations for a subset of the data that is critical for short-term operational needs.

The Observatory Operations Data Service supplements the more general Data Backbone by providing access to a subset of data at a QoS that is different (and higher) than the general Data Backbone. Less critical data is provided to Observatory Operations by the Data Backbone, which provides service levels provided generally to staff anywhere in the LSST project. For general access to the data for assessment and investigation at the Base Center, the service level is the same for any scientist working generally in the survey.

The Observatory Operations Data Service is instantiated at the Base Center. Therefore, the Observatory Operations Data Service does not directly support activities which must occur when communications between the Summit and Base are disrupted.

The service operates in-line with the Spectrograph and LSSTCam Archiving Services. Newly acquired raw data are first made available in the Observatory Operations server, and then are ingested into the Data Backbone permanent archive.

The intent is to provide access to:

- An updating window of recently acquired and produced data, and historical data identified by policy. An example policy is “last week’s raw data”.
- Other data as specifically identified by Observatory Operations. This may be file-based data or data resident in database engines within the Data Backbone.

A significant use case for the Observatory Operations Data Service is to provide near-realtime access to raw data on the Commissioning Cluster.

**2.6.2.3.1 Interfaces** File system export: The Observatory Operations Data Service provides access via a read-only file system interface to designated computers in the Observatory Operations-controlled enclaves.

Butler interfaces: Use of the LSST Stack is advocated for Observatory Operations, and so access to this data is possible via access methods supported by the LSST stack. The standard access method provided by the LSST stack is through a set of abstractions provided by a software package called the Butler. Whether the Observatory Operations Service provides a Butler interface is TBD.

Native interfaces: Not all needed application in the Observatory Operations context will use the LSST stack. These will not be able to avail themselves of Butler abstractions. The service accommodates this need by providing files placed predictably into a directory hierarchy.

Http(s) interface: The Observatory Operations Data Service also exposes its file system via http(s). Use of the Observatory Authentication and Authorization system is required for this access.

#### 2.6.2.4 Risks

- Concern: The need includes a continuously updated window of newly created data, in contrast to the other Butler use cases. How well the current set of abstractions work in a system that is ingesting new raw data is unknown to the author.
- Concern: Similarly, data normally resident in databases is part of the desiderata. Fulfilling these desiderata include solutions from an ETL into flat files, to establishing mirrored

databases. There are currently no actionable use cases for relational data. The technology to maintain subsets of relational data are distinct from the technologies to maintain subsets of files. It is likely that if relational data are needed, caches of relational data will need to be made by extract, transform and load into a file-format such as SQLite. Depending on Butler Registry implementation, this is TBD.

- Concern: This service needs to be available in Summit and Commissioning cluster enclaves. (and limited to those enclaves).

### 2.6.3 Operational Concepts

**2.6.3.1 Normal Operations** As indicated the OODS is a service used by the Spectrograph and LSSTCam archivers. The Availability targets for the Observatory operations server needs to accommodate the availability requirements of both systems.

**2.6.3.2 Operational Scenarios** Scheduling of changes to this system is controlled by the Observatory Change Control Authority.

## 2.7 Telemetry Gateway Service

### 2.7.1 Scope

The Telemetry Gateway is a service that receives information from trusted sources outside the Critical Operations Enclave and injects that information as telemetry into the OCS System.

### 2.7.2 Overview

**2.7.2.1 Description** The Telemetry Gateway is a service that receives information from trusted sources outside the Critical Operations Enclave and injects that information as telemetry into the OCS System.

**2.7.2.2 Objective** The prime use cases supported by the Telemetry Gateway are:



- The ingest of image parameters (WCS, PSF, or similar) from both prompt processing and offline processing use cases.
- Return of data from OCS-submitted batch jobs to the OCS System.

**2.7.2.3 Operational Context** The Telemetry Gateway Service is an OCS-commandable device which runs under the control of Observatory Operations.

The service transmits information originating from outside the Critical Operations Enclave to inside that enclave.

## 2.7.3 Operational Concepts

### 2.7.3.1 Normal Operations

**2.7.3.1.1 Operations During Prompt Processing** In normal operations, the service runs while prompt processing is running, or when OCS-driven batch jobs requesting return of parameters are running. The orchestration responsibility arranging for the telemetry gateway belongs to OCS.

**2.7.3.1.2 Operations Supporting Offline Processing** The service also runs at appropriate times to ingest delayed telemetry generated by offline alert and raw calibration validation processing. There is no need for the service to run concurrently with the offline processing, as offline processing is generated in a batch system, and the telemetry gateway runs according to the availability of the OCS system. This implies outputs from offline processing are buffered for ingest via the telemetry gateway service.

### 2.7.3.2 Operational Scenarios

**2.7.3.2.1 Change Control** Upgrades to the Telemetry Gateway Service are produced in the LSST Data Facility. Change control of this function is coordinated with the Observatory, with the Observatory having an absolute say about insertion and evaluation of changes.

## 3 Services for Offline Campaign Processing

The LSST Data Facility provides specific “offline” (i.e., not coupled to Observatory Operations) data production services to generate Data Release data products, as well as Prompt Data products-specific calibration data (e.g., templates for image differencing). Bulk batch production operations consists of executing large or small processing campaigns that use released software configured into pipelines to produce data products, such as calibrations and DRP products. Processing campaigns include

- **Data Release Processing:** Processing of payloads of tested work flows at NCSA and satellite sites through and including ingest of release products into file stores, relational databases, and the Data Backbone, including quality control and science validation.
- **Calibration Processing:** processing of payloads of tested work flows at NCSA through and including ingest of release products into file stores, relational databases, and the Data Backbone, including initial quality assurance. Calibration production occurs at various cadences from potentially daily to annual, depending on the calibration data product.
- **Special Programs and Miscellaneous Processing:** processing other than specifically enumerated.
- **Batch framework upgrade testing:** Test suites run after system upgrades and other changes to verify operations.
- **Payload Testing Verification and validation:** of work flows from the continuous build system on the production hardware for batch located at NCSA and satellite sites.

The concept of operations for batch production services serving Offline Campaign Processing is described in the following section.

### 3.1 Batch Production Services

#### 3.1.1 Scope

This section describes the operational concepts for batch production services, which are a set of services used to provide designated offline campaign processing.

## 3.1.2 Overview

**3.1.2.1 Description** Batch production service operations consists of executing large or small processing campaigns that use released software configured into pipelines to produce data products, such as calibrations and data release products.

**3.1.2.2 Objective** Batch production services execute designated processing campaigns to achieve LSST objectives. Example campaigns include calibration production, data release production, “after-burner” processing to modify or add to a data release, at-scale integration testing, producing datasets for data investigations, and other processing as needed. Campaign processing services provide automated quality control of data products.

- A campaign is a set of pipelines, a set of inputs to run the pipelines against, and a method of handling the outputs of the pipelines.
- A campaign satisfies a need for data products. Campaigns produce the designated batch data products specified in the DPDD [LSE-163], and other authorized data products.
- Campaigns can be large, such as an data release processing, or small, such as producing a few calibrations.

**3.1.2.3 Operational Context** Batch production services execute campaigns on computing resources to produce the desired LSST data products, which are measured against first-level quality criteria.

## 3.1.3 Operational Concepts

A pipeline is a body of code, typically maintained and originated within the Science Operations group. Each pipeline is an ordered sequence of individual steps. The output of one or more steps may be the input of a subsequent step downstream in the pipeline. Pipelines may produce final end data products in release processing, may produce calibrations or other data products used internally within LSST operations, may produce data products for investigations related to algorithm development, and may produce data products for testing purposes that cannot be satisfied using development infrastructure.

A campaign is the set of all pipeline executions needed to achieve a LSST objective.

- Each campaign has one or more pipelines.
- Each pipeline possesses one or more configurations.
- Each campaign has a coverage set, enumerating the distinct pipeline invocations. There is a way to identify the input data needed for each invocation.
- Each campaign has an ordering constraint that specifies any dependencies on the order of running pipelines in a campaign.
- Each campaign has an adjustable campaign priority reflecting LSST priority for that objective.
- Each pipeline invocation may require one or more input pipeline data sets.
- Each pipeline invocation produces one or more output pipeline data sets. Notice that, for LSST, a given file may be in multiple data sets.
- For each input pipeline data set there is a data handling scheme for handling that data set in a way that inputs are properly retrieved from the archive and made available for pipeline use.
- For each output pipeline data set there is a data handling scheme for handling that data set in a way that outputs are properly archived.

The key factor in the nature of the LSST computing problem is the inherent trivial parallelism due to the nature of the computations. This means that large campaigns can be divided into ensembles of smaller, independent jobs, even though some jobs may require a small number of nodes.

Batch Production Services are distinct from other services that may use batch infrastructure, such as Development Support Services. Also, there are other scenarios where pipelines need to be run outside the batch production service environment. For example, alternate environments include build-and-test, capable desk-side development infrastructure, and ad-hoc running on central development infrastructure.

From these considerations, the LSST Data Facility separates the concerns of a reliable production service from these other use cases, which do not share the concerns of production.

This also allows for supporting infrastructure to evolve independently. Example production service concerns include

- Supporting reliable operation of an ensemble of many campaigns, respecting priorities.
- Dealing with the problems associated with large-scale needs.
- Dealing with the campaign documentation, presentation, curation and similar aspects of formally produced data.

Computing resources are needed to carry out a production campaign. Batch processing occurs on LSST-dedicated computing platforms at NCSA, Base (Commissioning Batch and Chilean DAC) and CC-IN2P3, and potentially on other platforms. Resources other than for computation (i.e., CPU and local storage), such as custodial storage to hold final data products and network connectivity, are also needed to completely execute a pipeline and completely realize the data handling scheme for input and output data sets.

Computing resources are physical items which are not always fit for use. They have scheduled and unscheduled downtimes, and may have scheduled availability. The management of campaigns, provided by the Master Batch Job Scheduling Service requires:

1. the detection of unscheduled downtimes of resources
2. recovery of executing pipelines affected by unscheduled downtimes, and
3. best use of available resources.

One class of potential resources are opportunistic resources which may be very capacious but not guarantee that jobs run to completion. These resources may be needed in contingency circumstances. The Master Batch Job Scheduling Service is capable of differentiating kills from other failures, so as to enable use of these resources.

The types of computing platforms that may be available, with notes, are as follows.

Platform Type	Notes
---------------	-------

NCSA batch production computing system	Ethernet cluster with competent file systems.
NCSA Prompt computing for prompt processing	Shared nothing machines, available when not needed for observing operations.
NCSA US DAC computing	TBD
CC-IN2P3 bulk computing	Institutional experience is shared nothing machines + competent caches and large volume storage.
“Opportunistic” HPC	LSST type jobs running in allocated or in backfill context on HPC computers. [Backfill context implies jobs can be killed at unanticipated times].

An Orchestration system is a system that supports the execution of a pipeline instance. The basic functionality is as follows:

- Pre-job context:
  - Supports pre-handling of any input pipeline data sets when in-job context for input data is not required.
  - Pre-stages into a platform’s storage system, if available
  - Produces condensed versions of database tables into portable lightweight format (e.g., MySQL to SQLite, flat table, etc.)
  - Deals with TBD platform-specific edge services.
  - Identities and provides for local identity on the computing platforms.
  - Provides credentials and end-point information for any needed LSST services.
  
- In-job context:
  - Provides stage-in for any in-job pipeline input data sets
  - Provides any butler configurations necessarily provided from in-job context.
  - Invokes the pipeline and collects pipeline output status and other operational data
  - Provides any “pilot job” functionality.

- Provides stage-out for pipeline output data sets when stage-out requires job context.
- Post-job context:
  - Ingests any designated data into database tables.
  - Arranges for any post-job stage out from cluster file systems
  - Arranges for detailed ingest into custodial data systems
  - Transmits job status to workload management, defined below.

#### A Master Batch Job Scheduling Service:

- Considers the ensemble of available compute resources and the ensemble of campaigns.
- Dispatches pipeline invocations to an Orchestration System based on resource availability and considering priority of campaigns.
- Considers pipeline failures reported by the Orchestration System.
  - Identifies errors indicative of a problem with computing resources, and arranges for incident report.
  - Identifies some computational errors, and arranges for incident report.
  - Retries failed pipeline invocations, if appropriate.
- Exposes progress of the campaign to relevant entities.
- Provides appropriate logging and events (N.b. critical events can be programmed to initiate an incident).

#### Quality support:

Operations are supported by the following concepts, defined as follows for this document.

- Quality Assurance (QA) is what people do. This is identifying the issue and arranging for fixes. One source of input is quality controls, described below. Another source of input are the operational and scientific data products.

- A Quality Control (QC) is a software artifact that produces some sort of data that contains measure of quality. This data artifact may be
  - Simply produced, recorded and not used, because it seems useful for some future, likely retrospective purpose.
  - Displayed or presented for quality analysis.
  - Fed as input into active quality control which is software that automatically affects the execution of a campaign.
  - Fed into software that computes additional downstream quality control data.

**3.1.3.1 Normal Operations** During normal operations, Batch Production Services will conduct a number of concurrent campaigns that support LSST goals. These campaigns will be drawn from

- Runs to validate Data Release Processing,
- Data Release Processing itself.
- After-burner processing (to correct specific errors in not-yet-released data products).
- Calibration processing.
- Miscellaneous processing.

While Batch Production Services will use the majority of LSST batch capability, they may share the LSST batch infrastructure with certain Prompt services that require offline processing and with User Generated batch awardees. Resource conflicts are sorted out and expressed as priorities for each respective campaign.

The system is programmed to deal with anticipated errors. Human eye is applied during working hours, and can be summoned when events in the underlying systems generate incidents.

Each campaign is monitored for technical progress, both in in the sense of analyzing and responding to overtly flagged errors, and general monitoring and human assessment of the overall performance of the service.



First-order Quality Assurance is as follows:

1. Quality controls are considered by an LSST Data Facility Production Scientist and other staff. Data Facility staff apply any standard authorized mitigations, such as reprocessing, flagging anomalies, etc. The Production Scientist within the LSST Data Facility understands the full suite of quality controls, alerts Science Operations group to anomalies, and collaborates in diagnosis and mitigation of problems, as requested.
2. The service provided by the LSST Data Facility Production Scientist uses operational and scientific acumen to assess the data products at a first level, in addition to monitoring the extant quality controls. Particular attention is paid to
  - (a) operationally critical data (e.g., next night's flats needed for prompt processing)
  - (b) a processing campaign that is resource intensive, hence expensive to redo (or has expensive consequences)
  - (c) known problematic output data sets that are not adequately covered by existing quality controls.
  - (d) known problematic input data sets not adequately covered by existing quality controls.

Close collaboration is maintained between first order quality assurance and the broader scientific quality assurance in the project. Information obtained from first order quality assurance is continuously fed back to Science Operations.

Campaign closeout provides that all outputs are in final form, documentation and other artifacts have been produced, and all parties are actively notified about the status of a campaign.

**3.1.3.2 User Generated Products Batch Services** Authorized Science Users will also use batch services. This is a future development area.

**3.1.3.3 Development Support Batch Services** Small-scale ad-hoc batch services are also needed for use by developers and scientists on the operations team. Automated batch services used for regularly-scheduled test processing (i.e. continuous integration) is also needed. This is future development area.

### 3.1.3.4 Operational Scenarios

**3.1.3.4.1 Initiate campaign** Campaigns are initiated in response to an LSST objective, by specifying an initial set of pipelines, a coverage set, and an initial priority. The Batch Production Service staff is consulted with a reasonable lead time. Consistent with LSST processes, pipelines can be modified or added (for example, in the case of after-burners) during a campaign. These changes and additions are admitted when the criteria of change control processes are satisfied, including

- relevant build-and test criteria
- the impact of resource-intensive campaigns is approved and understood
- production-scale test campaigns

**3.1.3.4.2 Terminate failed campaign** Reasons for a campaign failure will be documented and submitted to Science Operations for review. Deletion of data products needs to be scheduled so that it occurs after the review is completed. This includes backing out files, materials from databases, and other production artifacts from the Data Backbone, and maintaining production records as these activities occur.

**3.1.3.4.3 Pause campaign** Stop a long running campaign from proceeding allowing for TBD interventions.

**3.1.3.4.4 Deal with problematic campaign** LSST is a large system. Pipelines will evolve and be maintained. There will be the campaigns, described in the operations documents. It is the nature of the system that as issues emerge extra resources will be needed to provide focused scrutiny on aspects of production for some pipeline. In many cases problems will be resolved by bug fixes, or addressed by quality controls and changes to processes. Any system needs to support mustering focused effort on quality analysis that is urgent, and lacks an adequate basis for robust quality controls. The LSST Data Facility Batch Production Services staff contribute effort to to solve these problems, in collaboration with Science Operations (or other parties responsible for codes).

**3.1.3.4.5 Deal with defective data** Production data may be deemed defective immediately as the associated pipelines terminate or after a period of time when inspection processes run. Such data need to be marked such that they will not be included in release data and will be set aside for further analysis.

**3.1.3.4.6 Deal with sudden lack (or surplus) in resources** As noted above, for large scale computing, the amount of resource available to support all campaigns will vary due to scheduled and unscheduled outages.

The technical system responds to an increase or decrease in resources by running more or few jobs, once the workload manager is aware of the new level of resources. The technical system responds to hardware failures on a running job in just like any other system – with the ultimate recovery being to delete an partial data and retry, while respecting the priorities of the respective campaigns.

## 4 Data Access Services for Authorized Users

The LSST Data Facility provides authorized users and sites access to data via a set of services that are integrated with the overall Authentication and Authorization (AA) System. These services are hosted by LSST Data Facility at the US and Chilean Data Access Centers and will include hosting elements of the LSST Science Platform. There is a MOU with the Chilean DAC that is not public and can not be cited here.

### 4.1 User Data Access Services

Service hosting elements of the LSST Science Platform.

#### 4.1.1 Scope

This section describes the concept of operations for hosting Data Access Center services in the US and Chile.

## 4.1.2 Overview

### 4.1.2.1 Description

Hosting of user data access services consists of:

- Providing Kubernetes for the LSST Science Platform, including the SUIT portal, Jupyter Notebooks, and DAX.
- Providing a Qserv instance for each data release.
- Loading a Qserv with the released data to the extent not already done during data release production.
- Providing support for users entering data into their MyDBs and entering sharded data into Qserv to the extent that is supported, as well as providing backup and disaster recovery for data entered into these databases.
- Providing user file space and backups and disaster recovery for the file space.
- Providing read-only access to data release products identified in the DPDD and other designated data resources, such as the Transformed EFD.
- Providing batch computing capabilities.
- Providing a second-tier help desk support to the primary help desk.
- Providing for special administration considerations of the Chilean DAC outlined in the Chilean MoU.
- Providing a security enclave appropriate for protecting LSST resources from actions of the user community.
- Integrating the LSST Authentication and Authorization system.
- Providing service-level monitoring to measure and assure a designated level of service.
- Providing users with a default level of access to computing resources.
- Providing users with a computing award with the designated amount of computing resources.

### 4.1.2.2 Objective

The objective is to host the data access services at the required availability for authorized users.

**4.1.2.3 Operational Context** The operational context is the LSST LDF infrastructure at the Base and at NCSA. An important element of the context is the elasticity of the supporting computing infrastructure. While, for example, the US Data Access Center consist of about 10% of the computing capacity at NCSA, it is desirable that the constructed system deliver to the end users 10% of the computing capacity averaged over a year, borrowing resources from the production system when demand on the Data Access Center is high, and returning resources to the production system during times when demand is low.

It is not required that the US and Chilean Data Access Centers be mirrors of each other, as this implies capacities may need to be nearly equal, which is not a requirement of the Chilean MoU. They are mirrors of each other for structure and configuration.

**4.1.2.4 Operational Concepts** TBD

## 4.2 Bulk Data Export Service

Service providing bulk data download to sites supporting groups of users.

### 4.2.1 Scope

This section describes the concept of operations for providing large-scale access to organizations.

### 4.2.2 Overview

**4.2.2.1 Description** The Bulk Export Service provides access to files containing LSST formal data products, as specified in the DPDD, for organized collaborations of authorized users and participating institutions supporting authorized users.

#### 4.2.2.2 Objective

1. Provide large-scale data access to organizations entrusted to hold large amounts of data before it is released to the worldwide astronomy community.

2. Provide large-scale data access to designated organizations for data after it is released to the worldwide astronomy community..

**4.2.2.3 Operational Context** The Bulk Export Service is available at NCSA only.

#### **4.2.2.4 Risks**

1. An organization may attempt to access data other than the data agreed to.
2. Information may be lost depending on off-site structure, compared to information available through the Data Access Centers.

### **4.2.3 Operational Concepts**

**4.2.3.1 Normal Operations** During normal operations, each site will have a MoU with LSST identifying the datasets and the due diligence required on the receiving site. The Authentication and Authorization system will be used to authenticate sites performing the data transfer.

**4.2.3.2 Operational Scenarios** Example use cases of the Bulk Export Service include:

1. Distribution of datasets of interest to the Dark Energy Science Collaboration.
2. Distribution of datasets of interest to the independent data access center at CC-IN2P3. However, forwarding of raw data and other data products to CC-IN2P3 is logically part of the data backbone functionality, and the return of data products computed at satellite computing facilities is the responsibility of the workload management system.

## **4.3 Hosting of Feeds to Brokers**

The LSST Data Facility hosts the alert distribution system and supports users of the LSST mini-broker, as well as providers of community brokers.

### 4.3.1 Scope

This section describes the concept of operations for hosting the system that distributes transient events to the LSST-provided and community-provided brokers..

### 4.3.2 Overview

#### 4.3.2.1 Description

Alert Distribution consists of:

- Receiving alerts from instances of the Alert Production Pipeline or backlogs of alerts from offline alert production.
- If needed, formatting received alerts into broker-specific syntax and forwarding the alerts to services which propagate alerts to the community. This includes community brokers selected by LSST and also the LSST alert “mini broker”, which is a limited functionality broker the LSST construction project is providing.
- Operating instances of the LSST alert “mini broker,” which includes accepting filters from individual authorized users.
- Providing status information about these activities to the community.
- Providing a clear interface including an audit trail for trouble-shooting, monitoring, statistics and other operational matters.
- Is capable of providing service to the prompt processing version and various batch configurations of Alert Processing.
- Alert distribution and alert generation have different operational needs and should be thought of 2 different environments.

**4.3.2.2 Objective** The objective is to provide a configurable layer that receives events from instances of the alert production pipeline and delivers alerts to event brokers ultimately resulting in end-user consumption while supporting the various operations scenarios enumerated below. This layer decouples event generators from the complexity of policy-defined event distribution.

**4.3.2.3 Operational Context** Alerts originate in the LSST Alert Production pipeline. The alert distribution system is configured into the overall pipeline process control system when running production modes at NCSA.

The context for this document is the distribution of the alerts between the following operational entities.

- A running alert pipeline which outputs alerts to an interface configured to pass alerts to the alert distribution service.
- The authoring interface for each supported alert broker instance. Alert Brokers transmit alerts to subscribed users, according to their own Service Level Agreements with their users. There are community-provided alert brokers and an LSST-provided “mini broker”.
- There may be feeds to multiple instantiations of a given broker. A use case that illustrates this need is to support testing upgrades of brokers, and the planned multiple instantiations of the LSST mini-broker. Note that community-provided alert brokers are not hosted on LDF computational hardware.
- As the design evolves, possibly serving as an intermediate buffer between the AP codes (which cannot block in a lengthy manner) and the database ingest for the Prompt database which stores records of the alert.
- A validation end point, which “looks like” a broker, but records the alerts sent to it, as a component for smoke testing, and other testing and operational needs of the alert distribution service itself.

**4.3.2.4 Risks** While the dominant method foreseen for alert distribution is the International Virtual Observatory Alliance (IVOA) VOEvent mechanism, practical brokers need to mature significantly to handle the LSST data rate. Moreover, it is likely that specialized brokers will serve specific astronomical interests as brokers can apply further science classification. Each event packet is large; not all information is of interest to every science topic. Providing for a way to reduce the packet size emitted at NCSA or allow brokers to filter packets before they are emitted at NCSA are risk mitigating features that need to be considered, and supported if consistent with budget. Some thought should be given in design for providing alerts to non-IVOA compliant entities. Attention should be paid to protocol and other issues related to scaling.



### 4.3.3 Operational Concepts

**4.3.3.1 Normal Operations** The normal operating scenario is Prompt Processing. In this scenario the alert distribution needs to be part of an overall system which normally presents an alert to a broker within 60 seconds of the data being read out.

The distribution system needs

- Only to present an alert to every broker instance, one and only once.
- For brokers that can ingest in a timely way, introduce no more than a 60-sec delay between production of the alert and presentation to the broker.
- Consider the effects of a given broker that is unable to ingest the stream of events subscribed to in a timely way.
- Need to protect the throughput of any one feed due to broker misbehavior from misbehavior of other brokers.
- Needs to accept alerts from the alert production pipelines.

**4.3.3.2 Operational Scenarios** Smoke testing: Smoke tests are end-to-end tests of the Prompt system. These tests are

- available to Observatory Operations to verify an Prompt service is functioning.
- used to validate changes to an Prompt service.

Testing of alert distribution is an element of smoke testing. The validation endpoint is used for this test. Testing of feeds to brokers is desired, but not required, as a valid system does not depend on the functioning of external components.

Offline processing: Offline alert distribution refers to distribution of alerts outside of the context prompt alert processing.

- When online processing fails, alert distribution may be configured into the system if offline processing occurs sufficiently recently after data is taken.
- Will likely occur when Alert Production algorithms change, due to the need to develop training sets for brokers with algorithms that need training, and where the software change in Alert Production may have affected that training. In this case alerts produced by the new software need to be conveyed appropriately to the brokers.
- Testing of upgrades of the alert distribution service itself with downstream brokers.

Availability: The availability requirements for the Alert Production system are quite high. The availability of Alert Distribution is a component of that availability.

- Alert distribution needs to be tested as a separate component from Alert Production AP software.
- Needs to be instantiated as part of the L1 Complete Test Stand.

Broker test/broker support: LSST has a notion of a limited number of supported brokers. In this model, the set of authorized brokers will change over time. Each broker will have a service level (or similar) agreement with the LSST project that provides information about the needed level of interface. The alert distribution system needs to provide a vocabulary of support actions, in addition to providing the real time stream of alerts. This support is envisioned as:

1. Alert replay, including full-rate replay, to support resolution of end-to-end problems involving rate.
2. Concurrent operations of two feeds to support major upgrades of a broker's infrastructure.
3. Pushing training sets to community-provided alert brokers through out-of-band means, for example in the case that our data model changes, and the classification algorithms in the target broker need to see training set data, processed by the new LSST algorithms.

## 5 Data, Compute, and IT Security Services

The LSST Data Facility provides a set of general IT services which support the LSST use-case-specific services mentioned in previous sections. These “undifferentiated heavy lifting” services include

- Data Backbone Services providing file ingestion, management, movement between storage tiers, and distribution to sites.
- Managed Database Services providing database administration for all database technologies and schema managed for the project.
- Batch Computing and Data Staging Environment Services providing batch capabilities on each LSST-provided platform at NCSA and the Base Center.
- Containerized Application Management Services providing elastic capabilities for deploying containerized applications at NCSA and the Base Center.
- Network-based IT Security Services providing project-wide intrusion detection, vulnerability scanning, log collection and analysis, and incident and event detection, and verification of controls.
- Authentication and Authorization (AA) Services providing central management of identities, supporting workflows and various authentication mechanisms, and operating AA endpoints at the Summit, Base, and Archive Sites.

The concept of operations for each of these services is described in the following sections.

### 5.1 Data Backbone Services

#### 5.1.1 Scope

The Data Backbone is a set of data storage, management and movement resources distributed between the Base Center and NCSA. The scope of the Data Backbone includes both files and data maintained by relational and other database engines holding the record of the survey and used by Prompt Products, Data Release Products, and Data Access Center services. There is not a Data Backbone endpoint at CCIN2P3. Out-of-band data transfer will

occur depending on what data is needed for the processing taking place and what is already at CCIN2P3.

The Data Backbone provides read-only data service to the US and Chilean DACs, but does not host data stores where DAC users create state. This is done to create a hard and easily enforceable separation of technologies, where no flaw in a DAC can corrupt the data produced by Prompt and Data Release production systems. For example, DAC resources such as Qserv and user databases, colloquially known as MyDBs, are provisioned in the context of a data access center, not the Data Backbone.

The Data Backbone ensures that designated data sets are replicated at both sites. The data backbone provides an enclave environment that is oriented toward protecting data by management, operational, and technical controls, including processes such as maintaining disaster recovery copies.

## 5.1.2 Overview

**5.1.2.1 Description** Files in the data backbone are presented as file system mounts and data access services. Database-resident data are presented as managed database services.

### 5.1.2.2 Objective

- Replication of designated file data within LSST Data Facility sites at NCSA and the Base Center.
- Replication of designated relational tables and data maintained in other database engines at NCSA and the Base Center.
- Implementation of policy-based flows to the disaster recovery stores. At the time of this writing, disaster recovery stores include the NCSA tape archive, CC-IN2P3, and commercial providers.
- Ingest of images produced by the Spectrograph, ComCam, and LSSTCam instruments.
- Ingest of the Engineering and Facility Database and associated Large File Annex.
- Ingest of data products from Prompt and Data Release production processing.

- Ingest of data from TBD other sources, approved by a change control process.
- Serving data to Prompt, Data Release, and other approved production processes.
- Serving data to the US and Chilean Data Access Centers.
- Integrity checking and other curation activities related to data integrity and continuity of operations.

### 5.1.2.2.1 Operational Concepts

**5.1.2.2.2 Files** Files in the Data Backbone possess path names which are subject to change through the lifetime of the LSST project, which at the time of this writing is seen as serving the last data release through 2034.

Robust identification of a file involves

- Obtaining a logical file name through querying metadata and provenance.
- Possibly migrating a file from a medium where the file is not directly accessible, such as tape, to a medium where the file is accessible.
- Selecting a distinct instance of the file from possibly many replicas.
- Accessing the file through an access method such as a file system mount or URL(s).

The project has identified several caches of data that are used in production circumstances. The distinguishing circumstances for these caches involve quality of service requirements for performance and availability. Absent sophisticated QoS in file systems, performance requirements are met by controlling access to the underlying storage via caching. Availability is assured by decoupling the cache from the database providing metadata, provenance, and location information. Application-level cache management provides path names within the cache to the application.

Casual use of data for short periods may rely on knowledge such as file paths, but is subject to disruption when paths are re-arranged, or should the underlying storage technology change, such as introduction of object stores.

**5.1.2.2.3 Data Managed by Databases** Replication of database information is specific to the database technology involved. Databases identified as holding permanent records of the survey are in the Data Backbone in the sense that they are instantiated in the context of a security enclave with management, operational, and technical controls needed to assure preservation of this data, and that the principal concern of enclave management is that data reside at the Base and at NCSA driven by business need.

**5.1.2.2.4 Managed Lifetime of Data** Data in the Data Backbone are preserved and replicated under processes guided by an overall Data Management Policy. Elements of the policy include specifications for replication supporting disaster recovery, and the permissibility of deleting file based data, and purging database records. Examples of deletion of data likely in such a policy are: Removal of production data from faulty production campaigns, removal of data produced in the production system that are not part of the retained record of the LSST survey, and removal of data in shakedown and debug processing.

### 5.1.2.3 Operational Scenarios

**5.1.2.3.1 Availability and Change Management** Catalog-based access systems such as indicated for the Data Backbone are limited by database availability as well as the availability of the file store and its access methods.

Time-critical applications involving the Observatory Operations Data Service and access to alert templates for prompt processing protect themselves by having caches as described above.

## 5.2 Managed Database Services

### 5.2.1 Scope

Managed Database Services provide database access to data that reside in relational or non-relational databases and generally meet at least one of the following criteria:

- Data originate outside of the LSST Data Facility, but are (or are potentially) used in

Prompt and Data Release processing, especially in the sense of data inputs needed to reproduce or refine a Prompt or Data Release computation.

- Data originate as data produced within the Prompt or Data Release production processes, and are meant to be retained for some period of time.
- Data are production-related metadata.
- Data are used as data coupling for processes involved in maintaining Prompt and Data Release products or other aspects of the LSST Data Facility.

## 5.2.2 Overview

In LSST, a distinction is made between patterns of storage of data in a database engine (schema, for purposes here) and an implementation of the schema in a database engine which stores the data. In LSST, common schema are used and shared in many scenarios in distinct but schema-compatible databases.

As an example, a common relational database schema can be used in development, unit test, integration, and production, but realized in different relational database software, e.g., SQLite3 in development and a heavy commercial database in production.

### 5.2.2.1 Description

**5.2.2.2 Objective** The primary focus of Managed Database Services is, as outlined, not the support of developers, but the support of production and data needing custody or curation. While some database schema design is performed in managed database services, by no means is all schema designed by Managed Database Services. Managed Database Services does have a role in determining the fitness for use of any schema present in databases it operates.

**5.2.2.3 Operational Context** The operational context for Managed Database Services is the context of the LSST Data Backbone within the LSST Data facility.

Part of the context is to consolidate database technologies where appropriate.

### 5.2.3 Operational Concepts

- Select technology appropriate for managed database instances
- Present a managed database service hosting the required schema
- Support the evolution of schema in a managed database service
- Provide the level of service needed for each managed database instance
- Provide capacity planning
- Provide installation
- Provide configuration
- Provide data migration
- Provide performance monitoring
- Provide security
- Provide troubleshooting
- Provide backup and data recovery
- Provide data replication where needed

## 5.3 Batch Computing and Data Staging Environment Services

### 5.3.1 Scope

Batch Computing and Data Staging Services provide primitives used by the Master Batch Job Scheduling Service. Batch Computing and Data Staging Services are provided in a distinct implementation for that is tailored for each batch system deployment.

### 5.3.2 Overview

Batch Computing and Data Staging Services are provided at NCSA and the Base Center.

Analogous (but not identical) services are provided by MoU to the LSST Data Facility by CC-IN2P3, as well as by any commercial batch provisioning and agency resources, such as XSEDE.



**5.3.2.1 Description** Both NCSA and the Base Center will have a core batch infrastructure that uses batch system logic to partition a pool of batch resources to various enclaves at the respective sites, with policies that govern priorities and file systems exposed for batch nodes running in the context of each enclave.

At the Base Center, Batch Services are supplied to the Commissioning Cluster and the Chilean Data Access Center from this pool.

At NCSA, Batch Services are supplied to the Development, Integration, General Production, Prompt and US Data Access Centers enclaves from this pool.

An additional pool of batch resources at each site is drawn from idle nodes in the core Kubernetes provisioning. An enhanced goal is the unification of resource management of the Kubernetes nodes and the batch pool.

Data staging refers to mechanisms needed to move data, primarily files, between the Data Backbone and the storage used by batch programs. This may be as simple as a copy operation between mounted file systems, or as complex as a staging via http or FTP.

**5.3.2.2 Objective** Batch Computing and Data Staging Services support LSST batch operations by providing a batch system supported by data movement primitives.

- Provide a batch scheduler.
- Provide any enclave-specific resources. An example is distinct head nodes for different enclaves.
- Provide enclave-specific configurations, including configurations needed for information security and work processes.
- Integrate ITC into the batch system.

**5.3.2.3 Operational Context** Batch Computing and Data Staging Services use resources in the master provisioning enclave and expose them to a given enclave, implementing policies appropriate to that enclave.

### 5.3.3 Operational Concepts

**5.3.3.1 Operational Scenarios** An important consideration is that these resources do not have a constant level of use within each enclave, and that over time the hardware resources needed for batch operations in an enclave will change.

Operating conditions may change as well. For example even with container abstractions, it may be necessary to partition the batch resources to support two versions of an operating system.

Somewhat analogously, the batch system may opportunistically use idle nodes provisioned for elastic Kubernetes computing.

Lastly, NCSA has substantial resources for prompt processing, such as alert production. Scheduling jitter and performance may preclude using a single batch scheduler for general offline production and prompt processing. Batch Computing and Data Staging Services covers having multiple scheduler instances.

## 5.4 Containerized Application Management Services

### 5.4.1 Scope

Containerized Application Management Services provide an elastic capability to deploy containerized applications. These services are provided with distinct configurations tailored for each enclave, but are provisioned on a common pool of ITC resources residing in the Master Provisioning Enclaves at each site.

### 5.4.2 Overview

There are two instances of Containerized Application Management Services - one at the Base Center and one at NCSA. These instances are the basis for servicing elastic computing needs at each site and a portable abstraction for symmetric deployment on commercial provisioning.

**5.4.2.1 Description** Both NCSA and the Base Center will have a containerized infrastructure that logically partitions a pool of Kubernetes resources to various enclaves at the respective sites, with policies that govern priorities and file systems exposed to applications running in the context of each enclave.

At the Base Center, Containerized Application Management Services are supplied to the Commissioning Cluster and the Chilean Data Access Center from this pool.

At NCSA, Containerized Application Management Services are supplied to the Development, Integration, General Production, Prompt and US Data Access Center enclaves from this pool.

**5.4.2.2 Objective** The objective of these services is to support LSST elastic operations by providing a containerized application management system compatible with LSST requirements.

- Provide containerized application management to each enclave, respecting enclave-specific controls including information security and work rules.
- Provide storage for containers and container management.
- Provide adequate capacity for each site.

**5.4.2.3 Operational Context** These services use resources in the master provisioning enclave at each site and expose them to a given enclave, implementing policies appropriate to that enclave.

### 5.4.3 Operational Concepts

**5.4.3.1 Operational Scenarios** An important consideration is that these resources do not have a constant level of use within each enclave, and that over time the hardware resources needed for elastic services in an enclave will change.

Operating conditions may change as well. For example, even with container abstractions, it may be necessary to partition the hardware resources to support two versions of an operating system.

## 5.5 Network-based IT Security Services

### 5.5.1 Scope

This section describes network-based operational information security services supporting the Observatory Operations and the LSST Data Facility.

### 5.5.2 Overview

**5.5.2.1 Description** The LSST Network-based IT Security Service provides technical controls for operational security assurance. These controls provide data that support the LSST Master Information Security Plan and IT security processes such as incident detection and resolution.

**5.5.2.2 Objective** The objective of the Network-based IT Security Service is to provide:

- Network Security Monitoring, including monitoring of high-rate data connections for data transfer across the LDF system boundaries (but excluding certain high rate transfers, such as the Prompt service access to the Camera Data System), including deployment of technologies for Active Response and Blocking of Attacks.
- Vulnerability Management for computers and application software in the enclaves.
- The technical framework to facilitate efficient Incident Detection and Response, including central log collection/event correlation for security purposes.
- Management of certain access controls, such as firewalls and bastion hosts, used for administrative access.
- Host-based intrusion detection clients deployed on end systems as appropriate.
- Security configuration management and auditing to baseline standards.

**5.5.2.3 Operational Context** The general approach to operational information security is that there is a LSST Information Security Officer (ISO) who reports to the Head of the LSST construction project, and will transition to report to the Head of the Operations project.

The ISO drafts a Master Security Program [LPM-121] plan, which the Head approves of as appropriately mitigating the Information Security risk. The Head then assumes responsibility for the residual risk of the plan. This is the security risk that remains, given faithful execution of the plan. The ISO oversees implementation and evolution of the plan, seeing that it is faithfully implemented and noting when mitigation and changes are needed. The ISO does any required staff work for the Head; for example, running staff training. The ISO is informed of and keeps records on security incidents, and is responsible for evolution of the security plan and evolution of security threats.

The ISO is responsible for a Information Security Response Team, which deals with actual or latent potential breaches in information security. The Incident Response Team is made of a set of draftees from the various operations departments, with the draft weighted towards departments with expertise and responsibility for critical operations and critical information security needs.

The ISO runs the annual security plan assessment. The management of each construction subsystem and operations department is responsible for annual revision of a Departmental Security Plan that complies with the Master Plan. These departmental plans include

- A comprehensive list of IT assets, applications and services.
- A list of security controls the department applies to each asset (technical and operational).
- A list of controls supplied by others that are relied on.

These controls apply to all offered services and all supported ITC. Reporting is easiest if the systems offered are under good configuration control. Under a good system, the security plans are living and updated by an effective change control process.

Verification: The ISO oversees a group that provides network-based security services described in the Objective part of this concept of operations.

A general approach to LSST-specific networking is the use of software-defined networking. This provides for isolation of networking supporting security enclaves. In particular, this allows for the separation of critical infrastructure for Observatory Operations and the LSST Data

Facility from office or other routine networking.

The context for these security enclaves cover the following production services in the LSST project, though other enclaves may join if feasible and desired by the relevant operational partners. These networks may participate in this infrastructure, but are currently seen as the responsibilities of AURA and NCSA.

<b>Production Service</b>	<b>Security Enclave</b>
Prompt Services	Split Between NCSA and Chile
Batch Production services	NCSA portion, excluding satellite center
US Data Access Center Services	NCSA
Critical Observing Enclave Services	Summit and Base Center
Chilean Data Access Center Services	Base Site
Data Backbone Services	Base Center and NCSA

**5.5.2.4 Risks** These are the standard elements of an information security infrastructure which are needed for a credible IT security project. Certain elements of the system are near the state of the art due to the data rates involved. Lack of credible infrastructure in this area will be seen as a flaw in the overall construction plan, preparing the LSST MREFC for operations.

### 5.5.3 Operational Concepts

**5.5.3.1 Normal Operations** The following elements provide the functionality needed to implement the network-based security elements of the LSST Master Information Security Plan:

- Intrusion Detection Systems (IDS) detect patterns of network activity that indicate attacks on systems, compromise of systems, violations of Acceptable Use of systems, abuse of systems, and other security-related matters.
- Vulnerability Scanning detects software services with vulnerable configurations or un-

patched versions of software via network fingerprinting. The system scans designated systems subject to a black-list. In addition to scanning for vulnerabilities, port scanning for firewall audits and ARP scanning for network asset management can also be conducted.

- Central Log Collection and Event Generation collects syslogs and other designated logs for storage (making logs invulnerable for an attacker to modify) and processes the logs to detect signatures indicating a compromise or poor security practices.
- Firewalls and bastion hosts provide a layer of active security. A typical use of a bastion host is to provide a layer of security between networks used to administer computers and more general networks.
- Host-based Intrusion Detection complements network monitoring by detecting actions within a system not visible from the network with tools such as auditd and OSSEC. This component also monitors the filesystems and checks for filesystem integrity.
- Active Response blocks communication with entities outside the observation site networks. This component is typically used to block “bad actor” entities outside the observation site network.
- Central Configuration Management enforces a baseline security and configuration on all systems.

New systems being deployed must be “hardened” to a security baseline and vetted by security professionals before moving into operations or after major configuration changes.

**5.5.3.2 Operational Scenarios** Vulnerability scanning periodically assesses designated ports on designated computers sensing vulnerabilities. An example of when this service is applied in a crucial case is assessing the effectiveness of the program of work patching a critical vulnerability.

Intrusion detection can detect, for example, an attempt to compromise a system. The detection system interacts with the active response system to cut off the attacker’s access to the computer. The intrusion detection system can also be used to aid in the investigation of an attack during the incident response and handling of a security incident.

Host-based Intrusion Detection checks for attacks against a host from the perspective of the host. Examples include multiple failed remote logins as reported by the host, or reports of file system changes that do not accompany an approved request for change or do not fall within a maintenance window.

The networks at the Observatory site must be monitored by intrusion detection systems. Acceptable IDS solutions include Bro and Snort. These systems must be able to handle the traffic load from various network segments at 10GB to 100GB speeds. The IDS systems must be placed at strategic locations and account for any expansion or changes in the network without the need to completely retool the IDS systems.

The information produced by the system is accessible by LSST staff involved in LSST information security, “landlords” hosting systems, and other parties with a valid interest in the data, to the extent required by the site’s specific security plans.

Active Response is typically referred to as a Black Hole Router (BHR) since it peers with the border routers on a network and offers the shortest router to destinations being blocked. Quagga and ExaBGP are two examples of BHR software.

The central Configuration Management System will enforce a security baseline and configuration on all systems. Examples of this technology are Puppet and Chef. In the event that Windows systems will be deployed onsite, a system enforcing Group Policy Objects and WSUS for patching will have to be available. It is required that this system would also be the Domain Controller using Active Directory and federate with LSST’s Unified Identity Management system.

The central log collectors are responsible for collecting and archiving all logs collected as described in the previous section. The collectors must be able to store at minimum six months of logs with a rotating windows deleting the oldest logs to maintain disk space. In addition to the log collectors, there is a SEIM/analysis system. This system is used for real-time log alerts, searches, and visualization. ElasticSearch, Kibana, and OSSEC are three examples of such software. This server spools a copy of the logs from the central collectors but may not be able to keep the full time window due to overhead or log metadata storage.

All systems, both workstations and servers, are required to send system logs to a central collector. For Linux systems, syslog must be configured to send a copy of all logs in realtime



to the central collector. For Windows systems, software such as Snare will be installed to send Windows Event logs to the central collector using the syslog protocol. Other alternatives exist for log collection such as logstash an open source log collection tool that can be used to collect logs from a wide variety of platforms.

Network devices are also required to send system logs to the central collector. Note that this is different than any network logs a switch or router would send. The system logs refer to events such as device logins, configuration changes, and other system specific events. Note that this is a requirement only if the device has this feature available.

Network devices such as routers or firewalls that are placed on the ingress/egress points of a vlan or the network must send firewall or router ACL logs to the central collector.

It is best practice for network devices to also send netflow to a central collector. However, if netflow is collected and forwarded to the central collector it must not be at the expense of network or device performance.

Any other devices not classified as a server, workstation, or networking device must be configured to send logs to the central collector if this feature is available. An example of a device that falls into this category is a VOIP appliance or a VPN appliance.

## 5.6 Authentication and Authorizations Services

See LSE-279.

## 6 ITC Provisioning and Management

ITC is managed in distinct enclaves. Enclaves are defined based on administrative and security controls, and operational availability requirements. Enclaves may span geographic sites, with elements in both the Base Facility in La Serena and at NCSA. Enclaves may share computing and other resources. Central administration is operated by NCSA staff, including remote administration of the Base Facility, with local support staff in Chile.

See LDM-129, the LSST Data Facility Logical Infrastructure Technology and Communications (ITC) Design Document, for more information.

## 7 Service Management and Monitoring

The LSST Data Facility provides a set of services supporting overall management of services, as well as monitoring infrastructure which collects information about running services for service delivery, incident response, planning for future upgrades, and supporting change control. Service management processes are drawn from the ITIL IT Service Management vocabulary.

### 7.1 Service Management Processes

#### 7.1.1 Scope

This section describes operational concepts of systems-level and service-level management for services operated by the LSST Data Facility.

#### 7.1.2 Overview

This section briefly describes functions and processes of service management that are implemented across all service and ITC layers of the LSST Data Facility. These elements were drawn from the Information Technology Infrastructure Library (ITIL) which is an industry-standard vocabulary for IT service management.

IT Service Management processes include

1. Service Design: Building a service catalog and arranging for changes to the service offering, including internal supporting services.
2. Service Transition: Specifying needed changes, assessing the quality of proposed changes, and controlling the order and timing of inserting changes into the system. This can also include changes that "undo" other previous changes.
  - *Change Management* provides authorization for streams of changes to be requested, for the insertion of changes into the reliable production system, and for the assessment of the success of these changes.
  - *Release Management* interacts with a project producing a specific change to ensure that a complete change is presented to change management for approval into the

live system. Examples areas that are typically a concern are accompanying documentation and security aspects.

- *Configuration Management* provides an accurate model of the components in the live system sufficient to understand changes, and support operations.

3. Service Delivery: operating the current set and configuration of production services. Service delivery processes must satisfy the detailed service delivery concepts presented elsewhere in this document.

- *Incident Response* is invoked when a service does not perform as specified according to its version. The goal of incident response is to allow work to continue, either by restoring the service to its original working form, or by providing a work-around. Each LSST service has a level of incident response support, which depends on the operational criticality of the service. For example, support of observatory services is likely more critical that support of a developer test stand.
- *Request Response* Request response are for normal operation activities and covers both structured requests and unstructured requests. An example of structured requests is to on-board or off-board a user. Examples of unstructured requests are to answer a question about the implementation of a service, or to produce an informative report.
- *Problem Management* Not all issues with a service can be addressed at a given time. These issues are aggregated into a problem backlog. The problem back log contains unaddressed, valid issues know to service users, and problems generated by internal staff. Problem management is the process where items are selected from the backlog to be addressed, root cause determined, followed by a change request to implement a fix.

### 7.1.3 Operational Concepts

TBD

## 7.2 Service Monitoring

## 7.2.1 Scope

This section describes operational concepts of systems-level and service-level monitoring for services operated by the LSST Data Facility.

## 7.2.2 Overview

**7.2.2.1 Description** The service monitoring system is the source of truth for the health and status of all operational services within its scope. The monitoring system deals with quality controls related to service delivery. These data have both retrospective and real-time uses:

- Acquires data from subordinate monitoring systems within components that are not bespoke LSST software. These monitoring systems may have an API, log files, SNMP, and TBD other interfaces.
- Acquires data from native LSST interfaces, including interfacing to the logging package (lsst.log), Prompt logging, Prompt events, scoreboards (Redis), TBD Qserv, and data from other independent packages.
- Probes services from monitoring agents and ingests quality control parameters.
- Synthesizes new quality control data from existing quality control data (for example, correlating a series of events before generating an event that will issue a page).
- Can generate events based on performance or malfunction which can trigger incident response for services and ITC, including to a non-NCSA incident response software.
- Can generate reports used for problem management, availability management, capacity management, vendor management and similar processes.
- Provides dashboard (or comfort) displays satisfying the use cases defined below.
- Provides for instantiation of displays anywhere within the LSST operational environment (concerns porting vs. remote display, paint display with high latency).
- Provides for publicly visible displays and displays visible only to those authorized by the LSST Authentication and Authorization system.
- Is sensitive to dynamic deployment of services to ITC resources.

- Is sensitive to modes of deployment and test, integration, development when generating alerts, painting displays, and recording data for retrospective use (concerns segregation and separation).
- Is itself highly reliable and available.
- Provides for disconnected operations between geographic sites (Summit, Base and NCSA) and enclaves (e.g., Observing Critical and non-Observing Critical).

**7.2.2.2 Objective** The set of services and infrastructure relied on by LSST Data Facility operations is inherently distributed due to the distributed deployment of the LDF services. Reliable operations of LDF services involves components instantiated (at least) in Chile, at NCSA, and at CC-IN2P3, as well as the networks between these sites.

A dataset based on the operational characteristics of the facilities, hardware, software and other elements of service infrastructure is needed to support service management, service delivery, service transition, and ITC-level activities, as well as to provide health and status information to the users of the systems. This dataset must be substantially unified, so that all activities are supported by a single source of truth. From a unified dataset, for example, staff concerned with availability management of a service can obtain records that consistently reflect availability information generated by incident response activities, while staff concerned with capacity management can obtain information on how capacity is provided by ITC activities.

In general, service management needs both a subset of the data that is needed for ITC management and data which may not be supplied by traditional ITC monitoring. Examples of data not supplied by ITC monitoring include the end-to-end availability of a service that tolerates hardware faults, user-facing comfort displays which address specific areas of interest, and controls that monitor data flow into disaster recovery stores for consistency with creation of data.

**7.2.2.3 Operational Context** LDF services rely on ITC hosted at NCSA, the Chilean Base Center, satellite computing centers, test stands at LSST Headquarters, wide area networks, and possibly other sources of infrastructure. Each of these sources possesses organization-specific (non-uniform) ITC monitoring and service management information on which LDF

services rely. In all cases the LSST Data Facility needs to centrally acquire sufficient data to provide for management of LDF services, while minimizing coupling to the ITC or service provisioning from these sites. The coupling should be defined in an internal Service Level Agreement (SLA) or similar written instrument.

## 7.2.3 Operational Concepts

### 7.2.3.1 Normal Operations

**7.2.3.1.1 Example: Prompt Products Services** Prompt Products services are instantiated at NCSA and the Chilean Base Center. The services the LDF relies on that may provide monitoring information are described in the table below. The monitoring system may acquire additional essential data by agents, consistent with SLAs and systems engineering best practice.

Table 3: Sources of Monitoring Data

Reliance	Subordinate monitoring interfaces provided
WAN from Chilean border router to UIUC	Wide area network activity area of LDF
Network transit from Chilean prompt processing infrastructure subsystem to Chilean border router	Observatory Operations
Network transport on UIUC campus to prompt processing installation area in NPCF	University of Illinois networking, NCSA networking
OCS interfaces (bridge, telemetry injection)	Observatory Operations
CDS interface	Observatory Operations
Base Center computing room resources	Observatory Operations, Computing Facility manager
Local assistance in Chile	TBD (if any)
ITC for Prompt system, exclusive of reliances listed otherwise	LDF ITC group or relied upon NCSA groups
NCSA/NPCF facility resource management	NCSA/NPCF facility management

Service-specific code and service-level performance as a part of the overall system, component-level aspect of prompt processing internals	Interfaces provided by LDF software group
--	---

Table 4: Uses of Monitoring Service Data Products

Entity	Need	Notes
Incident response	Events indicating service faults.	TBD these directly generate notifications (page), (and have the right filtering semantics)
Problem management	Incident information and information about marginal or near-miss events detected.	
Observing Operations, DPP Staff, and LSST HQ	Comfort displays indicating real-time status of services used.	NCSA staff should be able to see the same information as Observatory Operations staff to prevent confusion in incident response. It is important to note that monitoring relied on by Observatory Operations in Chile needs to be independent of NCSA. Monitoring at each site needs to operate and provide appropriate subsets of information to each site, should the connectivity between sites be disrupted.
Alert users	Information about when alerts are being exported and flows to various broker-like entities.	
Availability management	Queries, reports and displays focused on historical contributions to failures by reliance.	

Capacity management	Queries, reports and displays focused on historical usage of resources.	
Contract and SLA management	Queries, reports, and displays of quantities related to performance, e.g., response times, quality of materials or services.	
ITC staff	Supplemental information to ITC monitoring.	

**7.2.3.1.2 Example: Wide Area Networks** Many of the hardware components which make up the WAN will be managed by different entities (ISPs) based on who owns the particular section of the network. Typically all ISPs run their own SNMP network to monitor the health of the devices.

**7.2.3.2 Operational Scenarios** A predefined hierarchy of roles which will include different levels of users as listed below:

- Generic User
- System Administrator
- Science User
- LSST DM Administrator
- Hardware Operator (ISPs)
- Camera Control System Administrator
- Observatory Control System Administrator
- Super User



The level of access and response capabilities will be as defined in the user-profile. In the case of a “Generic User,” it may be necessary only to show if the LSST system is up and running. There can be a graphical representation of the status of the systems and subsystems.

For a “Super User” who will have access to detailed status information on the systems and subsystems, will be able to see in-depth event history and status reports (through log-scraping and fault databases). The Super User will also be able to access the logs database through the same portal.

In-between levels of access will be defined as per the definition of the roles and responsibilities of the user.

## A References

- [1] **[LSE-319]**, Jurić, M., Ciardi, D., Dubois-Felsmann, G., 2017, *LSST Science Platform Vision Document*, LSE-319, URL <https://ls.st/LSE-319>
- [2] **[LSE-163]**, Jurić, M., et al., 2017, *LSST Data Products Definition Document*, LSE-163, URL <https://ls.st/LSE-163>
- [3] **[LTS-210]**, Mills, D., 2015, *Engineering and Facility Database Design Document*, LTS-210, URL <https://ls.st/LTS-210>
- [4] **[LDM-129]**, Petravick, D., Johnson, M., Butler, M., 2018, *LSST Data Facility Logical Information Technology and Communications Design*, LDM-129, URL <https://ls.st/LDM-129>
- [5] **[LPM-121]**, Petravick, D.L., Withers, A., 2016, *LSST Master Information Security Policy*, LPM-121, URL <https://ls.st/LPM-121>
- [6] **[LSE-279]**, Withers, A., 2017, *Concept of Operations for Unified LSST Authentication and Authorization Services*, LSE-279, URL <https://ls.st/LSE-279>

## B Acronyms

Acronym	Description
ACL	Access Control List
AP	Alerts Production
API	Application Programming Interface
ARP	Antenna Reference Point
AURA	Association of Universities for Research in Astronomy
BHR	Black Hole Router
CBP	Collimated Beam Projector
CCD	Charge-Coupled Device
CDS	Camera Data System
CPU	Central Processing Unit
DAC	Data Access Center

DAX	Data Access Services
DM	Data Management
DPDD	Data Product Definition Document
DPP	Data Processing Production
DR	Disaster Recovery
DRP	Data Release Production
EFD	Engineering Facilities Database
ETL	Extract Translate and Load
FTP	File Transfer Protocol
HPC	High-Performance Computing
HQ	Head Quarters
IDS	Intrusion Detection Systems
ISO	International Organisation for Standardisation (Geneva, Switzerland)
IT	Integration Test
ITC	Information Technology and Communications
ITIL	Information Technology Infrastructure Library
IVOA	International Virtual-Observatory Alliance
L1	Level 1 (ambiguous could mean milestone or processing)
LDF	LSST Data Facility
LDM	LSST Data Management (handle for controlled documents)
LIGO	Laser Interferometer Gravitational-Wave Observatory
LPM	LSST Project Management (Document Handle)
LSE	LSST Systems Engineering (Document Handle)
LSP	LSST Science Platform
LSST	Large Synoptic Survey Telescope
LTS	LSST Telescope & Site (Document handle)
MOU	Memo Of Understanding
MREFC	NSF's Major Research Equipment and Facilities
NCSA	National Center for Supercomputing Applications
NPCF	National Petascale Computing Facility
OCS	Observatory Control System
OODS	Observatory Operations Data Service
OSSEC	Operating Systems SECurity
PSF	Point Spread Function

QA	Quality Assurance
QC	Quality Control
Qserv	Query Service, Proprietary LSST Database system
RFC	Request for Comments
SEIM	Security Event and Incident Management
SLA	Service Level Agreement
SNMP	Simple Network Management Protocol
SUIT	Science User Interface and Tools
TBD	To Be Defined (Determined)
TCT	Technical Control Team (Obsolete - now DMCCB)
UIUC	University of Illinois—Urbana-Champaign
URL	Uniform Resource Locator
US	United States
VOIP	Voice Over IP
VPN	Virtual Private Network
WAN	Wide Area Network
WCS	World Coordinate System
WSUS	Windows Server Update Services
XSEDE	Extreme Science and Engineering Discovery Environment