LARGE SYNOPTIC SURVEY TELESCOPE

Large Synoptic Survey Telescope (LSST)

Concept of Operations for the LSST Data Facility Services

D. Petravick and M. Gelman

LDM-230

Latest Revision: 2017-07-03

revision: 2.1
status: draft

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
2.0	2016-05-8	Beginning to render working group schema as more complete view of operational need as a basis for planning.	D Petravick
2.1	2017-06-26	Import draft versions into single document, with updates based on evolved operational concepts.	M Gelman

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 ETL Service 10

2.4.1 Scope 11

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 Operations Data Service 15
2.6.1 Scope of Document 15
2.6.2 Overview 15
2.6.3 Operational Concepts 18
2.7 Observatory Operations QA and Base Computing Task Endpoint 18

3 Services for Offline Campaign Processing 18
3.1 Batch Production Services 19
3.1.1 Scope 19
3.1.2 Overview 19
3.1.3 Operational Concepts 20

4 Data Access Hosting Services for Authorized Users 27
4.1 User Data Access Services 27
4.2 Bulk Data Distribution Service 28
4.3 Hosting of Feeds to Brokers 28

5 Data, Compute, and IT Security Services 28
5.1 Data Backbone Services 29
5.1.1 Scope 29
5.1.2 Overview 29
5.1.3 Operational Concepts 30

5.2 Managed Database Services 31
5.2.1 Scope 31
5.2.2 Overview 32
5.2.3 Operational Concepts 33
5.3 Batch Computing and Data Staging Environment Services 33
5.3.1 Scope 33
5.3.2 Overview 34
5.3.3 Operational Concepts 35
5.4 Containerized Application Management Services 35
5.4.1 Scope 35
5.4.2 Overview 36
5.4.3 Operational Concepts 37
5.5 Network-based IT Security Services 37
5.5.1 Scope 37
5.5.2 Overview 37
5.5.3 Operational Concepts 40
5.6 Authentication and Authorizations Services 42

6 ITC Provisioning and Management 43

7 Service Management and Monitoring 44

7.1 Service Management Processes 44
7.1.1 Overview 44

7.2 Service Monitoring 45

7.2.1 Scope 45

7.2.2 Overview 45

7.2.3 Operational Concepts 47

8 acronyms **50**

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 Level 1 (L1) data products. These Level 1 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, Level 1 databases, interaction with the alert-to-broker distribution subsystem, and providing feedback to observers.
- A Prompt Processing Service for assessing the quality of nightly calibration exposures.
- A Prompt Processing Service for assessing exposures from the Collimated Beam Projector, used as part of telescope optical path calibration.
- An “Offline” L1 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 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.

- An Archiving Service for acquiring raw data from the spectrograph on the Auxiliary Telescope and ingesting it into the Data Backbone.
- 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 QA and Base Computing Task Endpoint that 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.
- 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.

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 QA of images, alert processing and delivery, returning image parameters to the Observatory, and populating the Level 1 Database.

2.1.2.3 Operational Context Prompt Processing is a service provided by the LSST Data Facility as part of the Level 1 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 FITS 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 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.

Another evisioned program is “deep drilling”, where many more exposures than the two exposure visit will be taken.

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.

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 no future knowlege about what exposures will be taken, until the “next visit” is determined.

During science data-taking the Prompt Processing Service computes and promptly returns QA parameters (referred to as “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 judgement.

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

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

The detected transients are passed off to another servince,which recored the data date in a catalog which can be queried offline and sends 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.

AP runs in the context of the wide-fast-deep survey, deep drilling program and TBD other programs. 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 Claibraitons include flats and biases. Dariks are nto anticipated. LSST has an additional calibration devvices in its baseline a collumanted Beam projector 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. As well, 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 QA parameters (referred to as "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 annual release processing.

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 about 10 broad-band flatfield exposures in each camera filter, about 10 bias frames acquired from rapid reads of an un-illuminated camera, and optional 10 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 as a single operational sequence that is 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 necessarily 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 1nm 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 QA codes to help assure that the data are appropriate for use. Note, the prompt processing system will not be used to construct these calibrations.

Consideration of the lengthy calibrations, and the complexity of scheduling them means that the system must be reasonably available when needed. An approach requires minimal coordination between the observing and the archive center, which as described below is responsible for maintenance of the system, is a default daily maintenance window, with deviations negotiated as needed.

2.1.3.2 Operational Scenarios

2.1.3.2.1 Code Performance Problems Should a code run longer than budgeted, and the pace of processing fail to keep up with the pace of images, operations input is needed because there are trade-offs, as this would affect the production of timely QA data. However, note that when this situation occurs no immediate human intervention is needed. 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 the Archive Center 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 L1 Database that records transients. QA image parameters used to gauge whether an observation meets quality requirements can still be produced and ingested to the OCS system.

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 coordiniated 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 Level 1 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 Level 1 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 Backbon 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 Archving Service are produced in the LSST Data Facility. Change control of this function is coordiniated 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 extinction. 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. The Auxiliary Telescope slews 1:1 with the main LSSTCam, which implies two exposures every 39 seconds.

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 FITS files. Telescope data

products are described **?** .

2.3.2.2 Objective The Spectrograph Archiving Service reads pixel data from the Spectrograph verison of the CDS and metadata available in the overall Observatory Control System and builds FITS 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 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 proceessing 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 coordiniated with the Observatory, with the Observatory having an absolute say about insertion and evaluation of changes.

2.4 EFD ETL Service

2.4.1 Scope

The Engineering and Facility Database (EFD) is a system used in the context Observatory Operations. It contains all data, apart from pixel data acquired by the Level 1 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 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 ETL Service is to be able to relate the time series data to raw images and computed entities produced by L1, L2, and L3 processing, 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 called for substantially all of the EFD relational and flat-file material to be ingested into what is callef the Reformatted EFD.

1. There is a need to access a substantial subset of the Original Format EFD data in the general context of Level 2 data production and in the Data Access Centers. This access

is supported by a query-access optimized, separately implemented relational database, generally called the Reformatted EFD. A prime consideration is to relate the time series data to raw images and computed entities produced by L1, L2, and L3 processing.

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 disater recovery system.

2.4.2.3 Operational Context Ingest of data from the Original Format EFD into the Reformatted 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 general staff 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 in the period where LSST Operations occurs. 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 Reformatted EFD. Data of ongoing interest will be loaded into the Reformatted EFD.

2.4.3.1.2 EFD Large File Annex Handling and Operations Under control of an OCS-commandable device, 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 Reformatted EFD Operations

- The Reformatted EFD is replicated to the US DAC and the Chilean DAC.
- LDF will extract, transform and load into the Reformatted EFD pointers to files that have been transferred from the EFD large file annex into the Data Backbone.
- LDF will extract, transform and load designated tabular data from the Original Format EFD into the Reformatted EFDs residing in the Data Backbone at NCSA and the Base Center.
- "Designated" data will include:
  - Any quantities used in a production process.
  - Any quantities designated by an authorized change control process.
- The information in the Reformatted 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 Reformatted EFD Should the Reformatted EFD relational database be 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 Reformatted 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 coordiniated 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, and optionally rendevous with a small amount of returned data via the Telemetry Gateway.

2.5.2 Overview

2.5.2.1 Description

2.5.2.2 Objective

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

2.5.3 Operational Concepts

2.5.3.1 Normal Operations

2.5.3.2 Operational Scenarios

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 coordiniated with the Observatory, with the Observatory having an absolute say about insertion and evaluation of changes.

2.6 Observatory Operations Data Service

2.6.1 Scope of Document

This section describes the services provided to Observatory 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. The Observatory Operations Data Service provides butler context, and updates that context continuously as new data (for example, new raw images) becomes available.

Native interfaces: Not all needed application in the Observatory Operations context will use the LSST stack and 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.

- Concern: This service needs to be available in TBD operational enclaves. (and limited to those enclaves).

2.6.3 Operational Concepts

2.6.3.1 Normal Operations

2.6.3.2 Operational Scenarios

2.7 Observatory Operations QA and Base Computing Task Endpoint

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 Level 2 data products, as well as Level 1-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

- Annual 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 system quality assurance.
- Calibration Processing: processing of payload 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 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 located of 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 first-order QA 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 DPPD (**?**), and other authorized data products.
- Campaigns can be large, such as an annual 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 campaign. Batch processing occurs on LSST-dedicated computing platforms at NCSA 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 cluster file system.
NCSA L1 computing for prompt processing	Shared nothing machines, available when not needed for observing operations.
NCSA L3 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 Level 1 services that require offline processing and with Level 3 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 L1 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 Operational Scenarios

3.1.3.2.1 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 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

3.1.3.2.2 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.

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

3.1.3.2.4 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).

3.1.3.2.5 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.

3.1.3.2.6 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 Hosting 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.

4.1 User Data Access Services

Service hosting elements of the LSST Science Platform.

4.2 Bulk Data Distribution Service

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

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.

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 heaving 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 L1, L2, and Data Access Center services.

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 L1 and L2 production systems. For example, DAC resources such as Qserv and user databases, colloquially know 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 L1 and L2 production processing.
- Ingest of data from TBD other sources, approved by a change control process.
- Serving data to L1, L2, and other approved production processes.
- Serving data to the US and Chilean Data Access Centers.
- Integrity checking and other curation activates related to data integrity and continuity of operations.

5.1.2.3 Operational Context

5.1.3 Operational Concepts

5.1.3.0.1 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 though an access method such as a file system mount or Http(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.

5.1.3.0.2 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.

5.1.3.1 Operational Scenarios

5.1.3.1.1 Availablity 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 L1 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 L1 and L2 processing, especially in the sense of data inputs needed to reproduce or refine a L1 or L2 computation.
- Data originate as data produced within the L1 or L2 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 L1 and L2 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 provides 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, L1 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, L1 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 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 Level 1 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 (citedsLPM-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.

Production Service	**Security Enclave**
Level One 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 elments 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 unpatched 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 **?** .

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 "pair-of-hands" support staff in Chile.

The operational enclaves are as follows:

- Master Provisioning Enclaves, each at NCSA and the Base Facility. The Master Provisioning Enclaves provide administrative, security, and core computing infrastructure that can be provisioned for many enclaves.
- Level 1 Enclave, which spans Chile and NCSA to support prompt processing and archiving services.
- General Production Enclave, which hosts production infrastructure for offline processing, hosting VOEvent distribution, and Bulk Export Service presentation nodes.
- General Base Enclave, which provides general computing and data access for investigations by Observatory Operations.
- US Data Access Center Enclave, which presents for each authorized users the ability to query the Qserv database, make custom state in MyDB databases and user areas on file systems, access a custom JupyterHub, access files via shell, and submit batch jobs.
- Chilean Data Access Center Enclave, which presents for each authorized users the ability to query the Qserv database, make custom state in MyDB databases and user areas on file systems, access a custom JupyterHub, access files via shell, and submit batch jobs.
- Data Backbone Enclave, spanning Chile to NCSA, which hosts the primary record of the survey and synchronizes across sites. Managed data includes raw data acquired by Level 1 services, data produced by project processes, files from the large file annex of the EFD, and relational databases that are not part of the DACs.
- Wide Area Network, which provides connectivitiy between border routers of La Serena, NCSA, CC-IN2P3 and other designated sites.

LDM-129, the DM Infrastructure Design Document, is now considered obsolete.

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 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.
    - *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 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.

    - *Request Response*
    - *Incident Response*
    - *Problem Management*

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 bespoken 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 (pex.log), event package (ctrl_events), L1 logging, L1 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 Level 1 Services Level 1 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 4: 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 L1 infrastructure subsystem to Chilean border router	Observatory Operations
Network transport on UIUC campus to L1 installation area in NPCF	U of I networking, NCSA networking
OCS interfaces (bridge, telemetry injection)	Observatory Operations
CDS interface	Observatory Operations

Base Center computing room resources	Observatory Operations, Computing Facility manager
"Pair of Hands" assistance in Chile	TBD (if any)
ITC for L1 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 L1 internals	Interfaces provided by LDF software group

Table 5: 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 having reliances on NSCA need 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 Batch Production Services

7.2.3.1.3 Data Backbone Services

7.2.3.1.4 Data Access Hosting Services

7.2.3.1.5 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. LDF service monitoring taps into this information base and collects and forwards it to the central console for health and status monitoring.

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

- Generic User
- System Admininistrator
- Science User
- LSST DM Admininistrator
- 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.

References

[1] **[LDM-294]** 20, *placeholder for DM management plan*, LDM-294, URL `https://ls.st/LDM-294`

8 acronyms

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

Acronym	Description
CAM	CAMera
DAC	Data Analysis Consortium (obsolete; superseded by DPAC)
DM	Data Management
DMLT	DM Leadership Team
DPAC	Data Processing and Analysis Consortium
ESA	European Space Agency
LSST	Large-aperture Synoptic Survey Telescope
NCSA	National Center for Supercomputing Applications
TBD	To Be Defined (Determined)
US	United States