Data Management in Comissioning and Ops

William O'Mullane, AURA/LSST
DM Project Manager

DM All Hands
8th March 2018

LSST Operations

Commissioning

Reference material

LSST operations proposal

- Proposal was submitted in summer to NSF/DOE to fund LSST operations
- A joint agency review was completed with positive feedback Dec 7th 2017.
- LSST Operations key points:
  - Distributed over SLAC, Tucson, La Serena, NCSA Illinois
  - LSST Data Facility element
  - To fit in the new National Center for Optical and infrared Astronomy (NCOA) framework

What I will present are elements of the proposal - it is not accepted yet.

LSST Operations - Distributed

100 - 200 Gbps international links
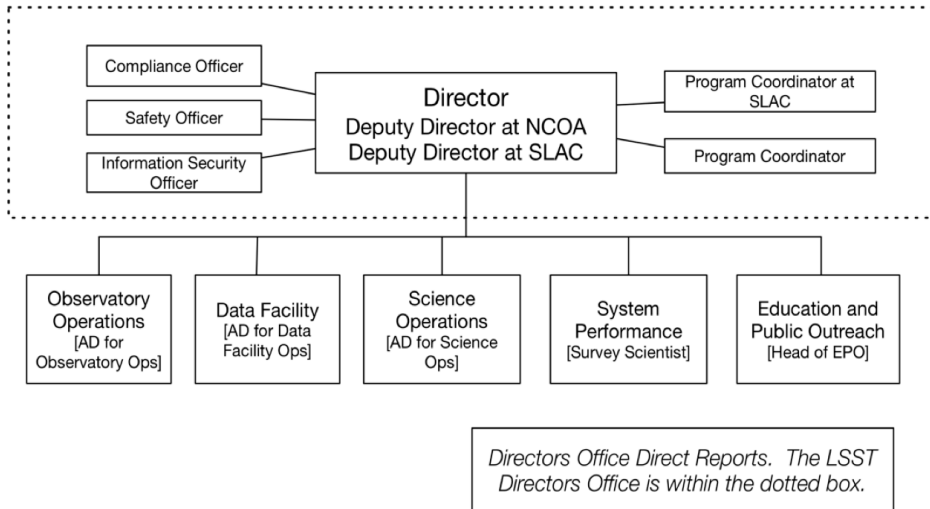
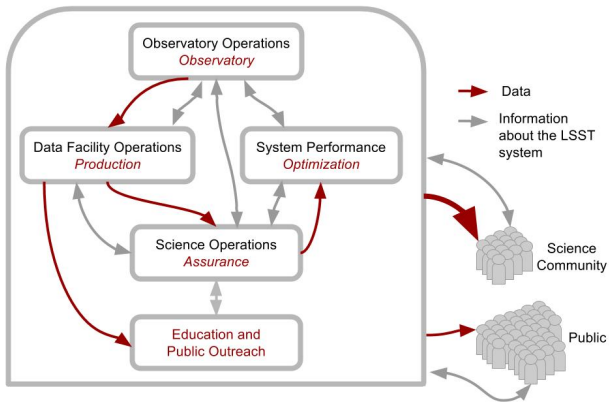40 - 200 Gbps summit base

See (LSE-78)

Jeff Kantor

LSST Operations: Sites & Data Flows

French Site
Satellite Processing Center
Data Release Production

HQ Site
Science Operations
Observatory Management
Education & Public Outreach

Archive Site
Archive Center
Alert Production
Data Release Production
Calibration Products Production
EPO Infrastructure
Long-term Storage (copy 2)
Data Access Center
Data Access and User Services

Base Site
Base Center
Long-term storage (copy 1)
Data Access Center
Data Access & User Services

Summit Site
Telescope & Camera
Data Acquisition
Crosstalk Correction

Emily Acosta

Director
Deputy Director at NCOA
Deputy Director at SLAC

Compliance Officer

Safety Officer

Information Security Officer

Program Coordinator at SLAC

Program Coordinator

Observatory Operations [AD for Observatory Ops]

Data Facility [AD for Data Facility Ops]

Science Operations [AD for Science Ops]

System Performance [Survey Scientist]

Education and Public Outreach [Head of EPO]

Directors Office Direct Reports. The LSST Directors Office is within the dotted box.

- Formal and informal channels in place
- Regular weekly daily meetings as well as Jira for tracking
- Slack, community.lsst.org, etc as well

Phil Marshall
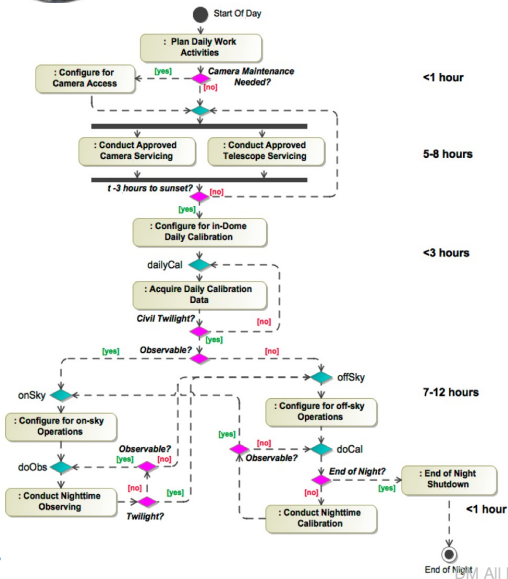
High level 24 hour activity (50FTE):

– Regular maintenance

– Evening calibrations

– Nightly observations

– Day crew, night crew shift 1, night crew shift 2
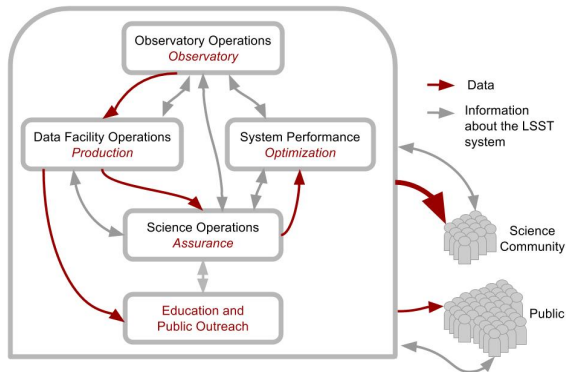
– Software

– ITC supports daily data transmission

Deliver the science products defined LSE-163 to the community.

That entails working with :

1. Chilean Operations to make sure the observations are scientifically good

2. the Data Facility to ensure production is running correctly

3. Survey performance to make sure over the longer period we will meet science goals.

4. EPO to communicate our successes **and possible more importantly our short comings**.

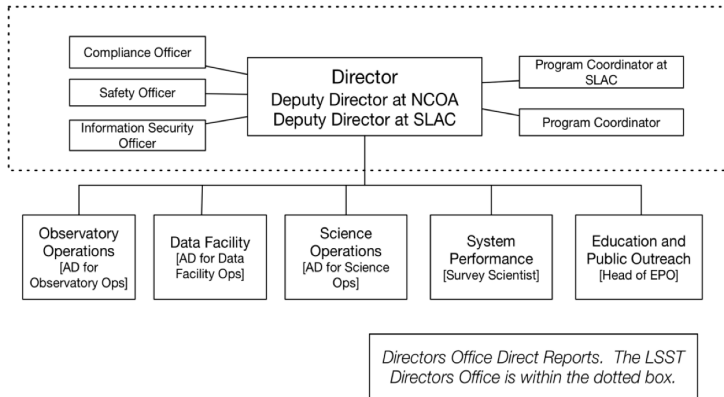**AND Maintain/improve software systems bought over from data management**

On a daily basis Science Operations Staff are looking at data quality from both instrument and software perspectives asking many questions (28FTE):

– Are the alerts as good as we can make them ?

– Are there any data products we could deliver/improve?

– Are the changes we should request on the telescope ?

– Was there some event (weather/hardware) affecting data we should be telling the community about ?

Longer term :

– Are there any disturbing trends in Key Performance Metrics?

– How is the Data Release Product quality ?

Compliance Officer

Safety Officer

Information Security Officer

Director
Deputy Director at NCOA
Deputy Director at SLAC

Program Coordinator at SLAC

Program Coordinator

Observatory Operations [AD for Observatory Ops]

Data Facility [AD for Data Facility Ops]

Science Operations [AD for Science Ops]

System Performance [Survey Scientist]

Education and Public Outreach [Head of EPO]

Directors Office Direct Reports. The LSST Directors Office is within the dotted box.

– Science operations is one of the main pillars of LSST operations
– FY23 estimate is to have 28FTE in the department
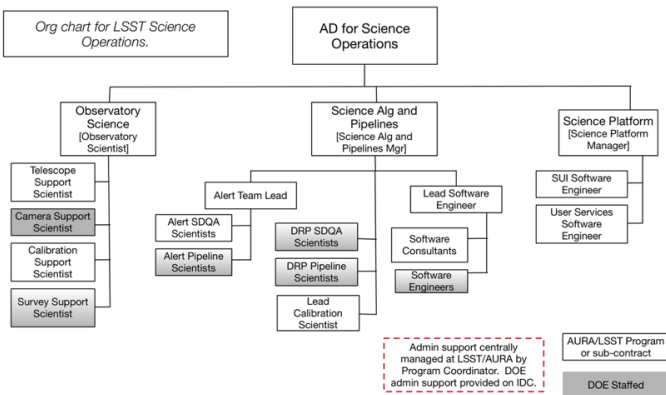– going down to 23FTE over 4 years

We are organized into three groups, which parallel the key components of science deliverables:

- Observatory Science
- Science Algorithms and Pipelines
- Science Platform.

Org chart for LSST Science Operations.

5.5 FTE starting in Chile to learn the instrument and them migrating to Tucson
This group should:

- Understand end-to-end impact of hardware and summit conditions on science
- Assure science images can deliver to LSSTs science requirements
- Track and document hardware issues
- Propose changes to the telescope, instrumentation or software for efficiency

Algorithms and Pipelines Group

8 FTE in various institutes eventually converging on Tucson.

- Tracking and documenting the data product quality for Alert Production using SDQA and other tools;

- Proposing when changes need to be made to all aspects of the Alert Production algorithms and pipelines (including distribution of alerts and orbits for solar system objects);

- Proposing when changes need to be made to all aspects of the annual data release algorithms (including calibration, Multifit, and deblender code); and

- Proposing to accept or reject software changes based on a scientific validation of new algorithms and an understanding of their impact on required computational resources.

4.5 FTE in Tucson

- – maintain and evolve the Science User Interface
- – maintain and evolve User Services

"Science Platform refers to the functions that will enable the community to conduct analyses and generate new data products, near the data, and will include Jupyter widgets to enable users to use Jupyter notebook with SUI visualization and a set of special visualization templates for interactive Quality Analysis (QA).

Science Operations Staffing

28 FTE starting operations going down to 23 FTE over 4 years.

Combination NCOA and SLAC positions.

Initially support also in institutes e.g. Princeton

Department	Position	NCOA	SLAC	Total
4 Science Operations				
4.1 AD of Science Operations	AD for Science Operations	1.00	-	1.00
4.2 Observatory Science	Calibration Support Scientist	1.00	-	1.00
	Camera Support Scientist	-	1.00	1.00
	Observatory Scientist	1.00	-	1.00
	Observatory Support Scientists	1.00	0.50	1.50
	Telescope Support Scientist	1.00	-	1.00
4.3 Science Platform	Science Platform Mgr	1.00	-	1.00
	Software Engineers	4.50	-	4.50
4.4 Science Algorithms & Pipelines	Science Algorithms and Pipelines Mgr	1.00	-	1.00
	Alert Production Team Leader	1.00	-	1.00
	Alert Production Pipeline Scientist	2.00	0.50	2.50
	SDQA Scientist for Alert Production	1.00	-	1.00
	Data Release Pipeline Scientists	2.00	1.50	3.50
	Data Release Systematics Scientist	-	1.00	1.00
	SDQA Scientist for Data Releases	1.00	-	1.00
	Lead Calibration Scientist	1.00	-	1.00
	Science Software Engineer Mgr	-	1.00	1.00
	Science Software Consultants	1.00	-	1.00
	Science Software Engineers	1.00	1.00	2.00
Total		**21.50**	**6.50**	**28.00**

LSST Data Facility Organisation

Organized to provide cohesive, reliable, and efficient production services

- Scientific Production Services and Data, Compute and IT Security Services provide the application-level elements of LDF services.

- Production Service Software maintains and evolves the LDF software.

- ITC and Facility Operations provide foundational hardware support, building on general facility staff.

Functional reporting structure for the Data Facility.

LDF responsibilities
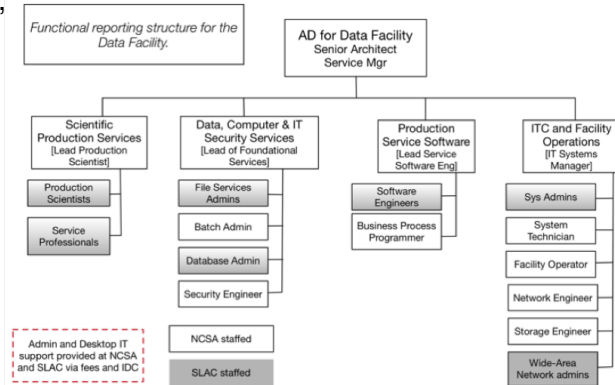
- Ingest data from all Observatory cameras and the EFD.

- Process data promptly for alert generation and in batch for daily to yearly calibrations and data releases.

- Provide alerts for transient sources to event brokers for distribution.

- Oversee and manage the hardware and data archive, including data integrity and disaster recovery, at both NCSA and the Chilean DAC.

- Monitor and coordinate the wide-area network infrastructure.

- Provide infrastructure support for data access services to the user community.

- Provide central Authentication and Authorization (AA) security-related services, including access control for science data rights holders and network security controls for the operational facilities.

Scientific Production Services Group (6.75 FTE)

Ensures the reliable and timely generation of prompt, nightly, and annual data products.

- Provides first-order tactical scientific data quality assurance (SDQA), focused on an initial diagnosis of issues during processing.
- Communicates operational feedback to Science Operations on status of data products production.
- Receives pipeline codes from Science Operations and verifies that processing is operationally feasible at production scale.

Staffed by Lead Production Scientist (1.0 FTE), Production Scientists (3.0 FTE), and Computational Facility Scientists (2.75 FTE).

Data, Compute, and IT Security Services Group (8.5.2)

Provides operational support for foundational services.

- – File management services on the LSST data backbone including ingestion, distribution, migration, replication, deletion, backup, disaster recovery.
- – Managed database administration.
- – Batch computing and containerized application management.
- – Authentication and Authorization (AA) and network security monitoring across LSST.

Staffed by Foundational Services Manager (0.75 FTE), File Services Administrators (1.75 FTE), Database Administrators (3.75 FTE), Workflow Services Administrators (0.75 FTE), Operational Security Engineers (1.5 FTE).

Production Service Software Group (4.3 FTE) LSST

Maintains, enhances, and evolves supporting software infrastructure required for production, including

- Qserv, the large-scale custom scientific database.
- Image acquisition and archiving.
- Near real-time execution orchestration, batch processing, data staging.
- LSST data backbone management.
- Data access, bulk export, hosting the LSST Science Platform.
- General LDF operational and service management software.

Staffed by Lead Software Engineer (1.0 FTE), Service Software Engineers (3.25 FTE), and Business Process Programmer (0.05 FTE).

ITC and Facility Operations Group (12.25 FTE)

Provides system administration and physical system operations of the LSST Data Facility computing, storage, and communications infrastructure.

- Operation of systems at NCSA and Base Site, including 24x7 support.
- Deployment and decommissioning of ITC infrastructure, physical space planning, coordination with Base and satellite facilities.
- Operation, management and support of Chilean and U.S. national and international networks for LSST usage.

Staffed by IT Center Systems Manager (1.0 FTE), IT System Administrators (3.25 FTE), IT System Technician (1.0 FTE), Computing Facility Operator (0.5 FTE), Facility Network Engineers (1.75 FTE), Storage Engineers (4 FTE), WAN Technical Manager (0.5 FTE), and WAN Architect (0.25 FTE).
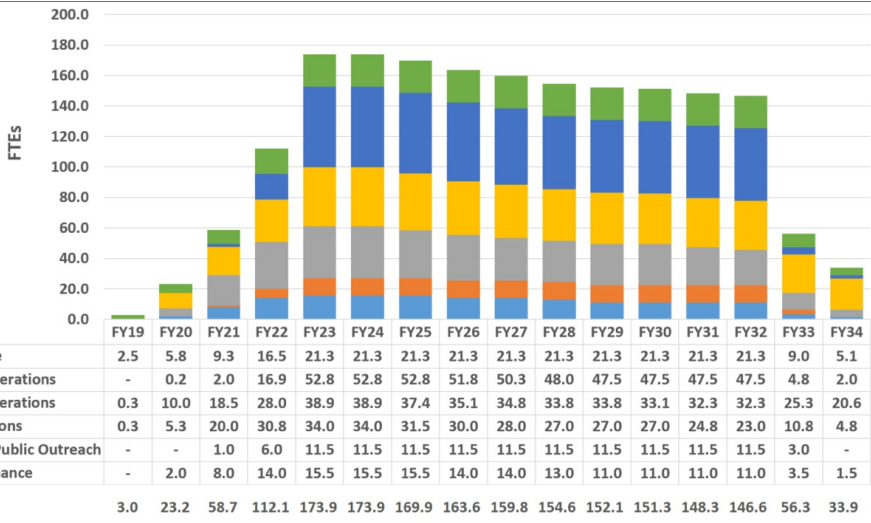
NCSA scientific and technical staff drawn from existing core areas of Center expertise.

FNAL due to proximity and existing collaborative relationships (e.g., DES).

SLAC due to intimate knowledge of software created during the LSST construction project. Currently includes QSERV

Department	Position	NCOA	NCSA	SLAC	Total
3 Data Facility Operations					
3.1 AD of Data Facility Operations	AD of Data Facility Operations	-	1.00	-	1.00
	Senior Data Facility Architect	-	1.00	-	1.00
	Data Facility Service Manager	-	1.00	-	1.00
3.2 Science Products Generation	Computation Facility Scientists	-	1.75	1.00	2.75
	Production Scientists	-	3.00	1.00	4.00
3.3 Data, Compute and IT Security Servic	Lead of Data, Compute, and IT Services	-	0.75	-	0.75
	Database Administrators	-	1.75	2.00	3.75
	File Service Administrators	-	1.00	0.75	1.75
	Operational Security Engineer	-	1.50	-	1.50
	Workflow Services Administrator	-	0.75	-	0.75
3.4 Production Service Software	Lead Service Software Engineer	-	1.00	-	1.00
	Service Software Engineers	-	1.25	2.00	3.25
	IT Business Process Programmer	-	0.05	-	0.05
3.5 ITC and Facilities	IT Center Systems Manager	-	1.00	-	1.00
	Computing Facility Operator	-	0.50	-	0.50
	Facility Network Engineers	-	1.75	-	1.75
	IT Systems Administrators	-	1.25	2.00	3.25
	IT Systems Technician	-	1.00	-	1.00
	Storage Engineers	-	4.00	-	4.00
	Wide Area Network Architect	-	-	0.25	0.25
	Wide Area Network Techical Manager (DOE ESN	-	-	0.50	0.50
Total		-	25.30	9.50	34.80

LSST Operations - Proposed Staffing profile

	FY19	FY20	FY21	FY22	FY23	FY24	FY25	FY26	FY27	FY28	FY29	FY30	FY31	FY32	FY33	FY34
1 Director's Office	2.5	5.8	9.3	16.5	21.3	21.3	21.3	21.3	21.3	21.3	21.3	21.3	21.3	21.3	9.0	5.1
2 Observatory Operations	-	0.2	2.0	16.9	52.8	52.8	52.8	51.8	50.3	48.0	47.5	47.5	47.5	47.5	4.8	2.0
3 Data Facility Operations	0.3	10.0	18.5	28.0	38.9	38.9	37.4	35.1	34.8	33.8	33.8	33.1	32.3	32.3	25.3	20.6
4 Science Operations	0.3	5.3	20.0	30.8	34.0	34.0	31.5	30.0	28.0	27.0	27.0	27.0	24.8	23.0	10.8	4.8
5 Education and Public Outreach	-	-	1.0	6.0	11.5	11.5	11.5	11.5	11.5	11.5	11.5	11.5	11.5	11.5	3.0	-
6 System Performance	-	2.0	8.0	14.0	15.5	15.5	15.5	14.0	14.0	13.0	11.0	11.0	11.0	11.0	3.5	1.5
Total	3.0	23.2	58.7	112.1	173.9	173.9	169.9	163.6	159.8	154.6	152.1	151.3	148.3	146.6	56.3	33.9

LSST Operations

Commissioning

Reference material

Operations Rehearsals

Date/Freq	Location	Title, Description
Oct 2018	NCSA	**Operations rehearsal for commissioning** With TBD weeks commissioning (lets say a week) – pick which parts of plan we could rehearse. Chuck suggests Instrument Signal Removal should be the focus of this (or the next rehearsal).
Oct 2019	NCSA	**Operations rehearsal #2 for commissioning** More complete rehearsal – where do the scientist look at quality data? How do they feed it back to the Telescope ? How do we create/update calibrations ? Exercises some of the control loops.
Jan 2020	Base	**Operations rehearsal #3 for commissioning** Dress rehearsal – Just like it will be April for the actual commissioning.
Dec 2020	NCSA	**Operations rehearsal data release processing (commissioning data)** Dress rehearsal – Just like it will be April for the actual commissioning.
2021	NCSA	**Operations rehearsal for data release processing (regular data).**
Feb 2022	NCSA/Base	**Operations rehearsal** Rehearsals for real operations which start Oct 2022
Sept 2022	NCSA/Base	**Operations rehearsal** Full Dress rehearsal for real operations which start Oct 2022
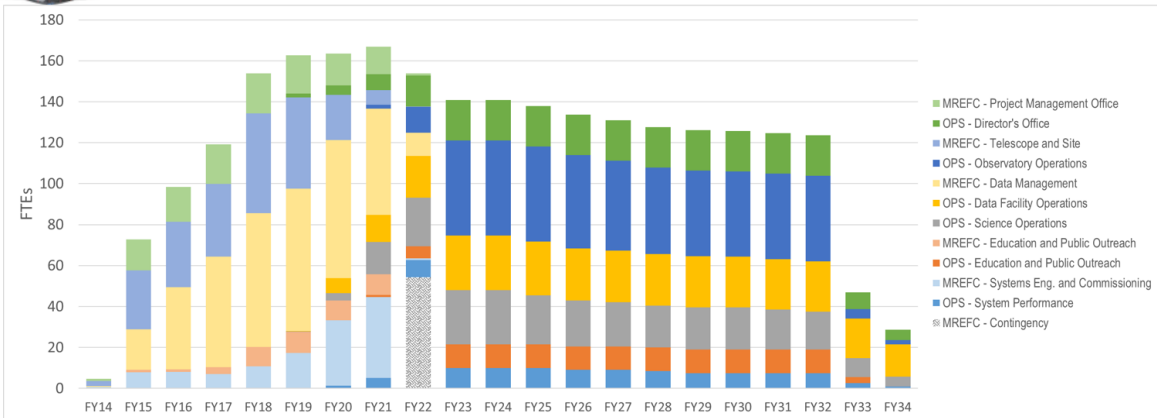
DM Transition

- DM –> Science Operations and LSST Data Facility
- Starting to look at specific roles and the transition of DM staff into them
- As you see in the previous chart there will be a problem in commissioning to find enough FTEs
- We may have to extend DM constructing further
- A few DM people will be in in Chile for some period of commissioning (Jim, Merlin, Robert . . .)
  - based on need to see the hardware
  - We are exploring this in the commissioning rehearsals
  - Most of us can work from our home institutes or perhaps congregate in Illinois or Tucson for specific events.

LSST Operations

Commissioning

Reference material

Acronym	Description
AP	Alerts Production
API	Application Programming Interface
CCB	Change Control Board
CI	Configuration Item
DAC	Data Access Center
DAQ	Data AcQuisition (system)
DAX	Data Access Services
DM	Data Management
DRP	Data Release Production
DWDM	Dense Wave Division Multiplex
EFD	Engineering Facilities Database
FITS	Flexible Image Transport System
GPFS	General Parallel File System
HSC	Hyper Suprime-Cam
ICD	Interface Control Document
IPAC	Infrared Processing and Analysis Center
ISR	Instrument Signal Removal
ITC	Information Technology Center
IVOA	International Virtual-Observatory Alliance
KPM	Key Performance Metric
LAN	Local Area Network

LCR	LSST Change Request
LDF	LSST Data Facility
LSST	Large Synoptic Survey Telescope
LaTex	(Leslie) Lamport TeX (document markup language and document preparation system)
NCSA	National Center for Supercomputing Applications
NOAO	National Optical Astronomy Observatories (USA)
OCS	Observatory Control System
PDAC	Prototype Data Access Center
PDR	Preliminary Design Review
PM	Project Manager
PSF	Point Spread Function
QA	Quality Assurance
Qserv	Query Service, Proprietary LSST Database system
SDSS	Sloan Digital Sky Survey
SLAC	Stanford Linear Accelerator Center
SODA	SCOS ORATOS Distributed Access
SQuaRE	Science Quality and Reliability Engineering
SST	Space Surveillance Telescope
SUIT	Science User Interface Team
TB	TeraByte
TBD	To Be Defined (Determined)
TCS	Telescope Control System
TCT	Technical Control Team (Obsolete – now DMCCB)
UCD	Unified Content Descriptor

US	United States
UW	University of Washington
WBS	Work Breakdown Structure
WCS	World Coordinate System
WG	Working Group
WISE	Wide-field Survey Explorer
WP	Work Package

[LSE-163], Jurić, M., et al., 2017, *LSST Data Products Definition Document*, LSE-163, URL https://ls.st/LSE-163

[LSE-78], Lambert, R., Kantor, J., Huffer, M., et al., 2017, *LSST Observatory Network Design*, LSE-78, URL https://ls.st/LSE-78