

## **Data Management Breakout: Technical**

#### William O'Mullane, AURA/LSST DM Project Manager

NSF/DOE Joint Status Review 28<sup>th</sup> Aug 2019





Joint Status Review • Tucson, AZ • August 27-30, 2019

100





- **DM Science**
- System Architecture
- **Science Pipelines**
- Science User Interface & Tools
- Science Data Archive & Application Services
- LSST Data Facility
- International Communications & Base Site
- Science Quality and Reliability Engineering
- Conclusion
- Reference material





- For each major team within DM (corresponding to third or fourth level of WBS, we present:
  - Highlights of significant achievements over the last year
  - Major upcoming activities







#### **DM Science**

- System Architecture
- **Science Pipelines**
- Science User Interface & Tools
- Science Data Archive & Application Services
- LSST Data Facility
- International Communications & Base Site
- Science Quality and Reliability Engineering
- Conclusion
- Reference material





The Data Management Science team includes the Subsystem Scientist, Deputy Subsystem Scientist, Pipelines Scientist, and Science Analysts. They are responsible for:

- Ultimate product ownership of the DM system
- Coordinating the activities of institutional science leads and the Subsystem Science Team
- Liaison with science collaborations and the external community
- Identifying and exploiting new scientific opportunities within DM
- Leading the DM Science Validation effort





- Workshops and Reviews
  - LSST Science Platform Final Design Review, Tucson, AZ 12-14 April 2019
  - Community Broker Workshop, Is.st/cbw, Seattle, WA 19-21 June 2019
- Science Studies
  - Data Management glossary
  - Alerts Key Numbers: ls.st/dmtn-102
  - Options for Alert Production in Year 1: Is.st/dmtn-107
  - Several updates and clarifications to the Date Products Definition Document
  - Next-to-data processing use cases: ls.st/dmtn-086
- Liaisons with Science Collaborations
  - Continual presence at stack-club to support LSST scientists and understand user needs





- SST liaisons to science collaborations attend meetings, workshops and present DM status
- DM-Camera-Commissioning bootcamp in December 2018
- PST science collaboration presentations on the LSST science platform and commissioning plans
- Verification and Validation
  - LSST Data Management Acceptance Test Specification Is.st/LDM-639
  - 1 new FTE working on DM verification since December 2018
  - Commissioning science verification workshop. joint with DM and Camera, June 2019





- Follow-up from the LSP Review, 28 recommendations in Jira
- Organize an LSST algorithms workshop, 2-4 December 2019, in the Princeton area
- Maintain a rotating fortnightly roster of SST members to attend and support Stack Club
- Study on Variability Characterization Parameters in the DIA/Object Catalog. A proposal will be prepared and feedback solicited from the science collaborations
- Advocate for increased resources and clearer requirements to improve the performance of the LSST Science Pipelines in crowded fields
- Study source detection efficiencies in difference images
- Investigate science use cases for the design of LSST data access services
- Increase efforts around science verification and validation in coordination with LSST commissioning
- Recruit a full time DM Science Validation scientist to be based in Tucson





#### **DM Science**

### System Architecture

- **Science Pipelines**
- Science User Interface & Tools
- Science Data Archive & Application Services
- LSST Data Facility
- International Communications & Base Site
- Science Quality and Reliability Engineering
- Conclusion
- Reference material





The DM System Architecture team includes the DM Software Architect, System Engineer, and Release Manager and two further engineers.

They are responsible for creating, maintaining, disseminating, and ensuring adherence to a common architecture for the Data Management System; for technological investigations, development, and prototyping; and for release management.





- Reviews
  - Participated in the LSST Science Platform external review
  - Took part in the PDR for the LSST Operations Visualization Environment
  - Chaired Identity Management internal review
- Requirements and Design
  - Substantially completed interface documents (LSE-68; LSE-72; LSE-131; LSE-140; LSE-400)
  - Updated DM System Design (LDM-148) including service levels and DM product tree (LDM-294)
  - Described data access and analysis for Observatory systems and Commissioning (DMTN-111)





- Release, Testing, Oversight, and Standards
  - Issued releases 16.0, 17.0.0, 17.0.1, 18.0.0
  - Defined release process (DMTN-106) and build/deploy environment handling (?)
  - Automated test documentation and verification control document generation
  - Defined deprecation procedure and tooling, switched to Python 3.7
- Investigation, Prototyping, and Implementation
  - Built LSP authentication and authorization (?)
  - Built WebDAV for LSP Web API access to the User File Workspace
  - Worked on Generation 3 Data Butler including in-memory, chained, object store datastores; metadata translation; instrument-specific packages including obs\_lsst; calibration datasets
  - Built catalog schema verification and documentation tool (https://felis.lsst.io) for the Science Data Model
  - Helped define and run Google Cloud and AWS proof-of-concept engagements





- Requirements and Design
  - Write guidelines for use of PyViz
  - Consult on Generation 3 Data Butler design
  - Work on new sizing/cost model
- Release, Testing, Oversight, and Standards
  - Provide detailed information on each DM product
  - Define and execute release processes
- Investigation, Prototyping, and Implementation
  - Assist with Amazon AWS proof of concept engagement
  - $\bullet$  Continue work on <code>obs\_lsst</code> and metadata translation
  - Continue work on Butler Datastores
  - Investigate and prototype changes to build and packaging tooling





#### **DM Science**

System Architecture

## **Science Pipelines**

- Science User Interface & Tools
- Science Data Archive & Application Services
- LSST Data Facility
- International Communications & Base Site
- Science Quality and Reliability Engineering
- Conclusion
- Reference material





"Science Pipelines" is an umbrella term covering the Alert Production (02C.03) and Data Release Production (02C.04) teams. They are responsible for designing and building the scientific algorithms which will be used to generate LSST's data products. The Alert Production team is responsible for:

- Detecting, formatting and issuing alerts of transient and variable sources
- Prompt image processing and catalog production
- Solar system object processing

The Data Release Production team is responsible for:

- Periodic ( $\sim$  annual) data release processing
- Calibration products production

The teams collaborate to construct common libraries, frameworks, and tools.





(These milestones were primarily completed by members of the Pipelines team.) **LDM-503-05: Alert Distribution** 

- Level 2 milestone demonstrating successful deployment and execution of the major components of the alert distribution system.
  - That is, transport of alert packets to science users via filters, at LSST-scale.
- Successful completion of the milestone documented in DMTR-91.

## LDM-503-07: Camera Data Processing

- Level 2 milestone demonstrating basic processing of data from *physical LSST* hardware through Science Pipelines.
  - As distinct from precursor or simulated data.
  - Includes availability of data at the Data Facility, access through the "Butler" I/O abstraction, scientific data processing, visualization of the results.
- Successful completion of this milestone documented in DMTR-112.





#### LDM-503-09a: Science Pipelines Fall 2018 Release

- Level 2 milestone capturing release of a pipeline system appropriate for carrying out the first Operations Rehearsal (LDM-503-09).
- Many new capabilities in this release:
  - Highlights include initial alert production capabilities, new middleware prototypes, improved reference catalogs, better instrument signature removal, etc.
  - Refer to https://pipelines.lsst.io/v/v17\_0/ for extensive notes.
- Successful completion of the milestone documented in DMTR-111.





- Alert Distribution
  - Proof-of-concept design and implementation for scalable Alert Filtering Service (formerly "mini-broker"); see DMTN-093.



Figure: Patterson.





- Alert Generation Pipeline Integration & Test
  - The prototype Alert Generation Pipeline is part of the standard LSST Science Pipelines software distribution (since release 17.0 in March 2019).
  - Alert generation integration tests have been added to the DM continuous integration system; regularly tracking the computational and scientific performance of the system.
  - Introduced a system for collecting "metrics" generated by pipeline execution in a generic, reusable way.
- Solar System Object Processing
  - Refreshed conceptual design for LSST solar system object processing, including closer integration with the Minor Planet Center; see DMTN-087.
    - Includes a comprehensive re-evaluation of LSST data rates as regarding solar system objects (DMTN-109).
    - A formal change request describing this work will be forthcoming.
  - Incorporation of new, more-efficient linking algorithm "HelioLinC" (Holman et al., 2018).





- Source Association and DIAObject Generation
  - Integrated source association functionality with the Prompt Products Database.
  - Forced photometry and variability characterization metrics now produced during DIAObject generation.
- Differential Chromatic Refraction (DCR)
  - Tasks for creating image differencing templates corrected for DCR are now available as part of the Science Pipelines and extensively tested on existing data (DMTN-028).
  - Evaluation of DCR mitigation in the presence of a variable point spread function; results being used to inform future development and feed into survey strategy (DMTN-121).



The effect of DCR mitigation on subtraction residuals. Figure: Sullivan.





- Calibration and Jointcal
  - Deployed a new astrometric matcher which provides dramatically more reliable results in dense fields (DMTN-031).
  - Completed integration of the Jointcal tool with LSST pipeline processing.
    - Jointcal performs astrometric and photometric fitting incorporating catalogs derived from multiple visits.
    - It is now used by default in LSST processing of precursor data from Hyper Suprime-Cam.



Improvement in the AM1 relative astrometry metric achieved by using Jointcal in processing. Figure: AlSayyad & Parejko.





- Focus on at-scale testing of alert production pipeline.
  - Increase the range and size of test datasets.
  - Resolve remaining issues around image differencing artifacts and false positive rates.
- Pick up development of alert distribution system.
  - Building on existing designs and community involvement from June 2019 broker workshop.
  - Expect a new hire to drive this work forward.
- Finalize plans for DCR mitigation in the presence of a variable PSF.
- Finalize and submit to CCB plans for updated solar system objects processing.
- Convert AP pipelines to "Generation 3" middleware.





- Middleware development.
  - The DRP team has made major contributions to "Generation 3" middleware development, including both the "Butler" data access abstraction and the "PipelineTask" task definition system.
  - This work has been carried out in conjunction with the DAX and LDF teams, under the management of the DAX T/CAM (Mueller).
  - As of summer 2019, complete conversion to the new system is underway, with the hope that the old middleware can be retired late in 2019.
- Star/galaxy classification.
  - A new machine-learning based classification system has been completed.
  - This system can make use of both morphological and color information, and is based on an adaptive boosting ("AdaBoost") technique.





- Deblending
  - The DRP team undertook a very successful "sprint" to integrate the SCARLET deblender (Melchior et al., 2018) into the LSST Science Pipelines; it can now be run as part of regular LSST processing.
  - This makes it possible to evaluate SCARLET's performance, capabilities and failure modes in the context of LSST, and hence to evaluate its suitability for use during LSST data release processing.
  - In parallel, working is ongoing to experiment with multi-resolution deblending using SCARLET (i.e. combining ground-based and space-based data in a single deblending solution).





- Photometric Calibration
  - The DRP team have integrated a version of the Forward Global Calibration Model (FGCM; Burke et al., 2018) with the LSST Pipelines, and have extended it to provide absolute photometric calibration with respect to a reference catalog.
  - FGCM has been shown to provide exquisite photometric calibration on LSST precursor data.







- Science Data Model standardization
  - A generic system has been developed for transforming raw outputs from science algorithms into a form compliant with the Science Data Model being developed by the DAX and Architecture teams.
  - We expect to integrate this system with the database ingest procedure for both Data Release and Alert Production processing over the coming months.
- Atmospheric Characterization and the Auxiliary Telescope Data Processing
  - A first version of the one-dimensional spectral extraction pipeline designed to reduce data produced by the Auxiliary Telescope Spectrograph has been produced.
    - This is still under heavy development, and will be substantially refined before the system is fully operational.
  - Work has proceeded to establish techniques for incorporating data from the Auxiliary Telescope and other sources into an overall atmospheric model which will be used in LSST calibration.
    - A paper on incorporating data from the NASA MERRA-2 meteorological assimilation system into LSST's calibration system is currently under review by the LSST Publication Board.





- Detector Characterization & Instrument Signature Removal
  - The DRP team has been analysing lab data collected from Auxiliary Telescope and LSST sensors using the Science Pipelines codebase with the aim of characterizing and mitigating detector effects.
    - Delivered mitigation for the brighter-fatter effect (Antilogus et al., 2014).
    - Developed techniques for identifying defects (e.g. bright pixels) in Auxiliary Telescope data.
  - A large-scale refactoring of the code for applying these corrections to science data "instrument signature removal" was completed. This makes the ISR code more robust, more generic, and easier to follow, while also preparing it for conversion to the upcoming "Generation 3" middleware.





- Complete delivery of "Generation 3" middleware and conversion of existing code to make use of it.
  - Expect that the DRP team will continue to have a role in maintaining and developing the middleware in the long term, but that the bulk of work will be carried out elsewhere.
- Begin transition away from MultiFit approach to Object characterization.
  - Simultaneously fitting over multiple epochs is the baselined approach to detailed object characterization, but is computationally expensive and no longer reflects best practice in the community. The DRP team is developing alternative approaches which better match community expectations.
  - Research is ongoing; ultimately, this is expected to lead to both new development and a change in the baselined algorithm.
- At-scale deblender testing.
  - Having integrated SCARLET, we are now in a position to evaluate its performance, and that of alternative approaches, at scale.
- Single-frame processing restructuring.
  - Add full-visit processing; match the design in LDM-151.





#### **DM Science**

- System Architecture
- **Science Pipelines**
- Science User Interface & Tools
- Science Data Archive & Application Services
- LSST Data Facility
- International Communications & Base Site
- Science Quality and Reliability Engineering
- Conclusion
- Reference material





The SUIT team is responsible for development of the Portal Aspect of the LSST Science Platform, and for development of the Firefly framework which provides common UI infrastructure and functionality. In LCR-1578, the Project chose to exercise scope option DM-10 (LPM-72; DMTN-096). This deferred work on the Portal Aspect of the Science Platform to a later date. In response to this LCR, SUIT group staffing has been ramping down

throughout calendar 2019 and Portal development is in a "maintenance only" mode.





- Participated in the successful LSST Science Platform Final Design Review
- New Features in Portal for LSST Science Platform
  - VOTable support in Firefly, including support for the LINK element
  - Support any TAP searches compliant with VO TAP standard in Firefly
    - User can select/input a VO TAP service
    - Browse through the schema and tables
    - Search the table either by using the simple UI or typing in advanced ADQL
    - Display the catalog in tri-view of table, image, and scatter plot
    - Display images properly if the table contains basic ObsCore image meta data
  - Authentication and authorization login and token passing
  - Workspace access through portal is mostly done, needs more testing
- Visualization
  - Overlay MOC on HiPS images





- Dynamic switch between HiPS image and FITS image display when zooming in and out
- JupyterLab extension for FITS image display
- Build and deployment
  - Built and deployed the portal as integrated part of LSST Science Platform in Kubernetes
  - Redis was chosen, implemented, and deployed in lsp-int for alternate message passing to multi-cast
  - Updated 3rd party packages: React, Webpack ...
- Various improvements and bug fixes, 49 bugs fixed since October 2018, 11 known to remain



# SUIT Achievements III



	Paul Sain: 100,2000 Janger Baak TYAP Legasy POAC External Emages External Catalogs Add Chart Upload Background Monitor	-	LISST TAP	egacy PDAC External Images External	Catalogs	Add Chart Upload	Background Mor	*	
TAP Searches trivele ingets inger									
1. TAP Service (9)	laina chttas://ast-ise-stabis.nosa.ilinois.edu/asi/tas>. Banisos		Coverage			kst-kp-stable.ncsa.ilinois.edu/api × Isst-lsp-stable.ncsa.ilinois.e ×			
		4	Options: FITS HIPS Auto				1 - 30 of 30)	7 2 8 0 4 1	
2. Select Query Type ①	Single Table 📀 ADQL		2MASS K. s FOV: 6.4	×		ra decl	w1mpro w1sigmpro	w2mpro w2sigmpro v	
3. Advanced ADQL ⑦ AD	QL edits below will not be reflected in Single Table view			0	8	148.94009270 69.04463160 1	2.42900000 0.05300000 1	2.51200000 0.04400000 1	
Schema -> Table -> Column sdss_stripe82_01	ADQL Query: Clear					148.92930260 69.06290780 1 148.94464100 69.08071310 1	0.75500000 0.02300000 1 2.38600000 0.02800000 1	1.05000000 0.02100000 5 5 2.98500000 0.03800000 1 1.	
6wie,00 6tpschema 8aadd2 8575 8ws	SELECT diskutanhmina hisponnaðinna aðagrenu aðagrenu aðagrenu aðagrenu aðagrenu aðagrenu földt viskaða skeina a Contraktipkontrickin, n. Benjalekski citis, skaldoga, sandag, sandag, sandar skeina aðagrenna film skalda skein					148.3187396         69.07753410         1.2.3300000         0.230000           148.3195670         69.0755462         1.2.3100000         0.220000           148.31938770         69.07553420         12.31800000         0.220000           148.31938770         69.07553420         12.31800000         0.022000           148.31932830         69.07553420         12.31800000         0.022000           148.31973150         69.055050         1.000000         0.025000           148.31973150         69.05505470         12.37800000         0.0250000           148.3197430         69.05695470         12.4700000         0.0250000           148.3197430         69.0464220         12.45020000         0.0250000		11.5100000 0.02200000 10 11.9300000 0.02200000 9.4 13.0000000 0.02200000 11 13.31000000 0.02200000 10 13.31000000 0.02200000 10 13.02900000 0.02500000 10 12.74600000 0.02500000 11	
	Schema Browser Hints								
	Click on a Column node to insert the column's name at the Query input box's cursor.			×		×		×	
	People Functions	_			12	-	2.5 2.5 1.5 -0,5 0	·	
Search Cancel Row Limit:		Ð	wimpro (mag)	wimpro (m	ig)		w1mpro-w2mpr	0	
v1.8.1_Development Bult Do: 2019-69-15									





- Bug fixes and deployment support
- Documentation for portal build and deployment
- Firefly APIs in JavaScript and Python documentation update
- User manual and online help





#### **DM Science**

- System Architecture
- **Science Pipelines**
- Science User Interface & Tools
- Science Data Archive & Application Services
- LSST Data Facility
- International Communications & Base Site
- Science Quality and Reliability Engineering
- Conclusion
- Reference material





The DAX team is responsible for:

- Development of the Qserv database system which will be used to store and serve LSST Data Releases to the scientific community.
- Development of data access systems following community standards which form the "Web APIs" aspect of the Science Platform.
- Development of "middleware" abstractions which are used to interface the Science Pipelines codebase with the underlying Data Facility infrastructure.




- Catalog Database (Qserv) on toward 175 TB
  - Three 30-node clusters operating:
    - NCSA: science dataset (Stripe 82 + WISE; Gaia DR2 soon)
    - CC-IN2P3: synthetic datasets
  - Data replication / auto-recovery now online
  - Cloud Deployment Demonstration
    - Google; with Kubernetes
    - Within 10-15% perf. vs. dedicated hardware
  - Scale testing has outstripped tooling; in dev:
    - Parallel test-data synth
    - Parallel ingest
    - New load testing harness
  - 75% DR1 testing late 2019







- Web Services for Science Platform Standards Orientation
  - TAP/ADQL query service (backed by Qserv) now online
  - SODA image cutout service now online
  - AstroPy PyVO Python client library improvements
    - Key tech for accessing data APIs from LSP notebooks
    - Team member (Banek) now a package maintainer
  - Metadata services (ObsTAP, SIAv2) coming late 2019







- "Generation 3" Data Butler and PipelineTask: An object-oriented data access abstraction (Butler) and processing task framework (PipelineTask)
  - HSC RC2 data release now being periodically re-processed entirely with Gen3 middleware:
    - HSC processing tasks were ported to new PipelineTask framework
    - Tasks use Gen3 Data Butler, backed by Oracle db at NCSA
    - Reprocessing runs carried out at NCSA, w/ prototype work-flow system (BPS)
  - Pipelines groups are actively porting remaining legacy tasks to new framework through summer/fall 2019
  - Targeting retirement of legacy middleware at end of 2019





- **DM Science**
- System Architecture
- **Science Pipelines**
- Science User Interface & Tools
- Science Data Archive & Application Services
- LSST Data Facility
- International Communications & Base Site
- Science Quality and Reliability Engineering
- Conclusion
- Reference material





The LSST Data Facility is a distributed facility, centered on NCSA and calling on resources at the Base Site, CC-IN2P3 and commercial providers, which is preparing compute and storage facilities for LSST operations and providing a range of services to the LSST Construction Project, including:

- Acquisition of pixels from the LSST instruments
- Retention of LSST data
- Provision and hosting of developer support systems
- Execution of data processing campaigns





- Completed Level 2 milestones LDM-503-4, LDM-503-4b, LDM-503-8, LDM-503-8b
- Completed numerous Spectrograph end-to-end tests and integration tests in person and over video-calls. Successfully ran the prompt forwarder/archiver processes for extended periods on Spectrograph test stand in Tucson.
- Completed performance testing and tuning of the consolidated database system throughout the year as more schemas were created by new aspects using DB.
- Participated in numerous workshops, reviews and conferences throughout the year and supported >50 participants of the LSST Science Platform Jupyter Notebook at LDF. Supported the stack-club friendly users on the LSP.





- Integrated new services and enhanced existing services at the LDF over the year with the Kubernetes infrastructure, multiple Qserv head nodes, test HTCondor infrastructure, configured authorization and authentication support systems in Chile at the NOAO data center and at the summit, added additional development resources, implemented BRO network packet inspection environment in Chile, and completed the L1 test stand procurement and NCSA DAQ creating images for testing.
- Automatically transferred all data from the AuxTel test stand in Tucson to the LSST Data Facility (LDF), ingested it into a Gen2 butler which enables images to be viewed in the LSP. LDF is also transferring data from other test platforms at SLAC, ingested into Butler Gen2 enabling LSP to view those images or as a backup of the important data.





- Participated in "proof of concept" tests with Google and AWS through coordination of software runs or data movement examining how bursting to the cloud could be accomplished.
- Participated in a LSST network demonstration during the Supercomputing 2018 conference which transferred data from the Summit to LDF over "borrowed" networks in the US showing the large packets and speeds available.
- Completed the first version of the Observatory Operations Data System (OODS) software, which will provide low-latency access to images, other files, and metadata for use by Observatory systems and the Commissioning Cluster Science Platform instance.





- Completed an in-depth demonstration of integration of the Pegasus Workflow Management System and Oracle database back-end with the current release of pipeline tasks and middleware. This demonstration included an in depth analysis of the Butler schema implemented in Oracle with numerous iterations of changes for performance and usability of the ButlerG3 middleware.
- Hosted and participated in a workshop that included installation of new Camera Data Acquisition (DAQ) hardware on the L1 Complete Test Stand at LDF as well as a tutorial for the new software interface.
- Base Data Center setup and configuration of test systems for databases, EFD+ETL systems, forwarder testing for images being sent to LDF, local file system needs, and data transfer nodes.





- Participated in Puppet workshops so that systems at the BDC, Tucson, and LDF have the same functionality, can be built and managed with the same tools ensuring uniformity across the platforms at all three locations.
- Led Operations Rehearsal #1 to prepare for commissioning by simulating nominal operations during a three-day observation period.
- Procured and configured test systems for ComCam and have sent them to Tucson for initial installation and testing.





- Observatory-oriented services, data management, and supporting infrastructure
  - Support multi-CCDs for ComCam in header, image build, data transfer
  - Archiving asynchronously from observing in "catch up" mode
  - Management/Monitoring of raw data archiving to LDF
  - Deploy spectrograph support environment at the summit for moving data to LDF
  - Test stand at LDF deployment for overall testing as systems are deployed on the summit
  - Build out of high availability enclave at the summit/BDC
  - Support for Spectrograph and ComCam on-site commissioning
  - Deploy EFD infrastructure for quick access and transfer to LDF
  - Base commissioning cluster and deployment infrastructure (xcat/Puppet/head nodes/Kubernetes) for on-site analysis
  - Security systems for authentication and authorization for all Chilean subsystems
  - Data Movement:





- ComCam, Spectrograph, other test stand data sent to LDF, ingested and tracked automatically
- Test long haul networks for throughput, stability and reliability
- Building on Rucio/Oracle for tracking multiple location files/datasets
- Build tools for data management/monitor workflow of raw images
- Production Data Processing
  - Integration of calibration and ISR pipelines for instrument commissioning
  - Testing development pipelines; merging new features, investigating and reporting errors, posting results, improving the workload management process
  - Integration of Butler G3 into production workload management system
  - Mock prompt processing workflow management including alert generation-to-distribution hand-off
  - Launching user applications from LSP through notebook to the compute clusters with accounting and allocation limitations





- Overall management and operations
  - Operations rehearsal #2 for commissioning
  - Ramp up staff for 7\*24 monitoring of the environment and building LDF knowledge base of problem resolutions
  - Manage overall data management workflows and services
  - LDF staff transition to commissioning and pre-operations





- **DM Science**
- System Architecture
- **Science Pipelines**
- Science User Interface & Tools
- Science Data Archive & Application Services
- LSST Data Facility
- International Communications & Base Site
- Science Quality and Reliability Engineering
- Conclusion
- Reference material





The International Communications and Base Site team ("Base & Networks") is responsible for the infrastructure for the Base Center and the national and international networks connecting the Mountain Summit, Base, Archive, and Headquarters sites.





- Summit LDF 44 Gbps First Optic Light Demonstration at Supercomputing 2017 (December 2017)
- Base LDF 100 Gbps Demonstration at Supercomputing 2018 (November 2018)
- Installed racks, fibers, PDUs, outlets in Base Data Center (enough for FY19)
- Installed network cabling in BDC offices and Base Office Building first floor
- Rerouted fibers on recinto to support move out of existing computer room to BDC
- Installed and tested LSST DWDM system from the Summit to Base
- Moved AURA and Reuna DWDM racks from the existing computer room to the BDC
- Installed switches for AuxTel, DIMMs, All Sky Camera
- Configured and tested Call Manager for new VOIP system





- Supported successful test controlling AuxTel Dome and Pointing
- Procured transceivers and modules to modify the LSST 10 $\times$ 10G ăto 100G link between La Serena and Santiago
- Installed and moved tenants onto new backup microwave system (Ubiquiti and Cambium)
- Completed repairs on Summit Base Fiber posts
- Finalized procurement for ESnet service, Atlanta Chicago
  - FY20 2×20 Gbps, FY21 2×100 Gbps



- 100 Gbps demonstration La Serena NCSA November 14 @Supercomputing 2018.
- 10 TB of DECam data via SCInet Miami Dallas Chicago.
- Jupyter notebook used to monitor from Data Transfer Node to disk@NCSA.







- Installation of summit networks (1st, 2nd, 3rd floors, casino, etc.)
- Support of T&S AIV, AuxTel team, DIMM testing
- Installation of networks in Base Office Building
- Installation of new 400 Mbps microwave backup system (Entel)
- Rack, PDU, network, server installations in Base Data Center
- Implementation of Phase 1 of Esnet service (2×20 Gbps) Atlanta Chicago
- Activation of Spectrum Link Santiago Boca Raton
- Activation of Boca Raton Miami diverse path
- Implementation of Virtual Network Operations Center (VNOC)





- **DM Science**
- System Architecture
- **Science Pipelines**
- Science User Interface & Tools
- Science Data Archive & Application Services
- LSST Data Facility
- International Communications & Base Site
- Science Quality and Reliability Engineering
- Conclusion
- Reference material





Science Quality and Reliability Engineering — "SQuaRE" – is responsible for:

- Automated software and science quality control (verification) services
- The Science Platform Notebook Environment for QA, commissioning & user science
- Developer infrastructure, software distribution, documentation and communications tooling.





- Current status and recent highlights:
  - The LSP Notebook Aspect is now a popular tool inside the project
    - DM and Commissioning teams make sustained internal use
    - Deployments on LDF Kubernetes Commons and Google Cloud Platform
    - Dask is a promising candidate for next-to-notebook processing
  - Continuous Integration and Developer Support
    - Continuous Integration agents now run under Kubernetes
    - Jenkins pipeline for Science Pipelines official releases completed
  - SQuaSH (metrics curation) now in InfluxDB
    - Metrics migrated to InfluxDB for off-the-shelf back ends
    - Django-based system being retired, notebooks replacing analytic views
    - Isst.verify being used by developers to ship metrics to the system
  - Documentation





- Task documentation framework for the Science Pipelines
- Technote, package, and more! creation via new slackbot ("sqrbot jr")
- Observatory Operations services (new!)
  - Kafka/InfluxDB based DM-EFD prototype promising, scales well even to full 50Hz
  - EFD deployments under kub3s in lab and k8s on Google Cloud Platform
  - Also new: automated report creation from notebook templates



## SQuaRE Achievements III









- Prioritize Notebook Aspect needs for Integration and Commissioning teams
- Work on Notebook Aspect stability in the LDF Kubernetes commons environment
- Single Sign On authentication support for Notebook Aspect
- Develop Kafka-based infrastructure to expose EFD telemetry to Science Platform
- Support recommendations from QA and Visualisation working groups
- Prototype infrastructure for notebook development support (CI, PRs, curation etc)
- Jenkins master port to Kubernetes
- Deployment automation and documentation for all services





- **DM Science**
- System Architecture
- **Science Pipelines**
- Science User Interface & Tools
- Science Data Archive & Application Services
- **LSST Data Facility**
- International Communications & Base Site
- Science Quality and Reliability Engineering
- Conclusion
- **Reference material**





- We have been tracking work more closely
- There are delays as in any project
- We are dealing with schedule slips
- We have been interacting in a more coordinated way with the community
- We hope the reviewers agree that DM technical progress is adequate.









#### DM and System Engineering teams at IPAC 2018

Joint Status Review • Tucson, AZ • August 27-30, 2019





- **DM Science**
- System Architecture
- **Science Pipelines**
- Science User Interface & Tools
- Science Data Archive & Application Services
- LSST Data Facility
- International Communications & Base Site
- Science Quality and Reliability Engineering
- Conclusion
- Reference material







Acronym	Description
ADQL	Astronomical Data Query Language
AIV	Assembly Integration and Verification
AP	Alerts Production
API	Application Programming Interface
AURA	Association of Universities for Research in Astronomy
AWS	Amazon Web Services
Alert	A packet of information for each source detected with signal-to-noise ratio > 5 in a difference image during Prompt Processing, containing measurement and characterization parameters based on the past 12 months of LSST observations plus small cutouts of the single-visit, template, and difference images, distributed via the internet
Alert Production	The principal component of Prompt Processing that processes and calibrates incoming images, performs Difference Image Analysis to identify DIASources and DIAObjects, packages and distributes the resulting Alerts, and runs the Moving Object Processing System
Archive	The repository for documents required by the NSF to be kept. These include documents related to design and development, con- struction, integration, test, and operations of the LSST observatory system. The archive is maintained using the enterprise content management system DocuShare, which is accessible through a link on the project website www.project.lsst.org
BAC	Budget At Completion
BDC	Base Data Center
BPS	Batch Production Service
Broker	Software which receives and redistributes Alerts, and may also perform processing such as filtering for certain characteristics, cross- matching with non-LSST catalogs, and/or light-curve classification, in order to identify and prioritize targets for follow-up and/or make scientific analyses.
Butler	A middleware component for persisting and retrieving image datasets (raw or processed), calibration reference data, and catalogs
С	Specific programming language (also called ANSI-C)







CAM	Control Account Manager
CCB	Change Control Board
CCOB	Camera Calibration Optical Bench
CI	Continuous Integration
CPI	Cost Performance Index
CPU	Central Processing Unit
Camera	The LSST subsystem responsible for the 3.2-gigapixel LSST camera, which will take more than 800 panoramic images of the sky every night. SLAC leads a consortium of Department of Energy laboratories to design and build the camera sensors, optics, electronics, cryostat, filters and filter exchange mechanism, and camera control system
Center	An entity managed by AURA that is responsible for execution of a federally funded project
Commissioning	A two-year phase at the end of the Construction project during which a technical team a) integrates the various technical components of the three subsystems; b) shows their compliance with ICDs and system-level requirements as detailed in the LSST Observatory System Specifications document (OSS, LSE-30); and c) performs science verification to show compliance with the survey perfor- mance specifications as detailed in the LSST Science Requirements Document (SRD, LPM-17)
Construction	The period during which LSST observatory facilities, components, hardware, and software are built, tested, integrated, and commis- sioned. Construction follows design and development and precedes operations. The LSST construction phase is funded through the NSF MREFC account
DAC	Data Access Center
DAQ	Data Acquisition System
DAX	Data Access Services
DB	DataBase
DCR	Document Change Request
DESC	Dark Energy Science Collaboration
DIA	Difference Image Analysis







DIAObject	A DIAObject is the association of DIASources, by coordinate, that have been detected with signal-to-noise ratio greater than 5 in at
	least one difference image. It is distinguished from a regular Object in that its brightness varies in time, and from a SSObject in that
	it is stationary (non-moving)
DIASource	A DIASource is a detection with signal-to-noise ratio greater than 5 in a difference image
DIMM	Differential Image Motion Monitor
DLP	DM Long Term Plan
DM	Data Management
DMCCB	DM Change Control Board
DMLT	DM Leadership Team
DMSR	DM System Requirements
DMTN	DM Technical Note
DMTR	Data Management Test Report
DOE	Department of Energy
DR	Data Release
DRP	Data Release Production
DWDM	Dense Wave Division Multiplex
Data Access Center	Part of the LSST Data Management System, the US and Chilean DACs will provide authorized access to the released LSST data
	products, software such as the Science Platform, and computational resources for data analysis. The US DAC also includes a service
	for distributing bulk data on daily and annual (Data Release) timescales to partner institutions, collaborations, and LSST Education
	and Public Outreach (EPO).
Data Backbone	The software that provides for data registration, retrieval, storage, transport, replication, and provenance capabilities that are com-
	patible with the Data Butler. It allows data products to move between Facilities, Enclaves, and DACs by managing caches of files at
	each endpoint, including persistence to long-term archival storage (e.g. tape)



# Acronyms IV



dataset to the scientific community and public. The DM team is responsible for the DMS architecture, applications, middleware, infrastructure, algorithms, and Observatory Network Design. DM is a distributed team working at LSST and partner institutions, with the DM Subsystem Manager located at LSST headquarters in Tucson         Data Management System       The computing infrastructure, middleware, and applications that process, store, and enable information extraction from the LSST dataset; the DMS will process peta-scale data volume, convert raw images into a faithful representation of the universe, and archive the results in a useful form. The infrastructure layer consists of the computing, storage, networking hardware, and system software. The middleware layer handles distributed processing, data access, user interface, and system operations services. The applications layer includes the data pipelines and the science data archives' products and services         Data Release       The approximately annual reprocessing of all LSST data, and the installation of the resulting data products in the LSST Data Access Centers, which marks the start of the two-year proprietary period         Data Release Production       An episode of (re)processing all of the accumulated LSST images, during which all output DR data products are generated. These episodes are planned to occur annually during the LSST survey, and the processing will be executed at the Archive Center. This includes Difference Imaging Analysis, generating deep Coadd Images, Source detection and association, creating Object and Solar System Object cataloos and relator metadata
infrastructure, algorithms, and Observatory Network Design. DM is a distributed team working at LSST and partner institutions, with the DM Subsystem Manager located at LSST headquarters in Tucson           Data Management Sys- tem         The computing infrastructure, middleware, and applications that process, store, and enable information extraction from the LSST dataset; the DMS will process peta-scale data volume, convert raw images into a faithful representation of the universe, and archive the results in a useful form. The infrastructure layer consists of the computing, storage, networking hardware, and system software. The middleware layer handles distributed processing, data access, user interface, and system operations services. The applications layer includes the data pipelines and the science data archives' products and services           Data Release         The approximately annual reprocessing of all LSST data, and the installation of the resulting data products in the LSST Data Access Centers, which marks the start of the two-year proprietary period           Data Release Produc- tion         An episode of (reprocessing all of the accumulated LSST images, during which all output DR data products are generated. These episodes are planned to occur annually during the LSST survey, and the processing will be executed at the Archive Center. This includes Difference Imaging Analysis, generating deep Coadd Images, Source detection and association, creating Object and Solar System Object cataloos, and related metadata
the DM Subsystem Manager located at LSST headquarters in Tucson           Data Management System         The computing infrastructure, middleware, and applications that process, store, and enable information extraction from the LSST dataset; the DMS will process peta-scale data volume, convert raw images into a faithful representation of the universe, and archive dataset; the DMS will process peta-scale data volume, convert raw images into a faithful representation of the universe, and archive the results in a useful form. The infrastructure layer consists of the computing, storage, networking hardware, and system software. The middleware layer handles distributed processing, data access, user interface, and system operations services. The applications layer includes the data pipelines and the science data archives' products and services           Data Release         The approximately annual reprocessing of all LSST data, and the installation of the resulting data products in the LSST Data Access Centers, which marks the start of the two-year proprietary period           Data Release Production         An episode of (re)processing all of the accumulated LSST images, during which all output DR data products are generated. These episodes are planned to occur annually during the LSST survey, and the processing will be executed at the Archive Center. This includes Difference Imaging Analysis, generating deep Coadd Images, Source detection and association, creating Object and Solar System Object cataloors and related metadata
Data Management System       The computing infrastructure, middleware, and applications that process, store, and enable information extraction from the LSST dataset; the DMS will process peta-scale data volume, convert raw images into a faithful representation of the universe, and archive the results in a useful form. The infrastructure layer consists of the computing, storage, networking hardware, and system software. The middleware layer handles distributed processing, data access, user interface, and system operations services. The applications layer includes the data pipelines and the science data archives' products and services         Data Release       The approximately annual reprocessing of all LSST data, and the installation of the resulting data products in the LSST Data Access Centers, which marks the start of the two-year proprietary period         Data Release Production       An episode of (re)processing all of the accumulated LSST images, during which all output DR data products are generated. These episodes are planned to occur annually during the LSST survey, and the processing will be executed at the Archive Center. This includes Difference Imaging Analysis, generating deep Coadd Images, Source detection and association, creating Object and Solar System Object cataloos, and related metadata
tem       dataset; the DMS will process peta-scale data volume, convert raw images into a faithful representation of the universe, and archive the results in a useful form. The infrastructure layer consists of the computing, storage, networking hardware, and system software. The middleware layer handles distributed processing, data access, user interface, and system operations services. The applications layer includes the data pipelines and the science data archives' products and services         Data Release       The approximately annual reprocessing of all LSST data, and the installation of the resulting data products in the LSST Data Access Centers, which marks the start of the two-year proprietary period         Data Release Production       An episode of (re)processing all of the accumulated LSST images, during which all output DR data products are generated. These episodes are planned to occur annually during the LSST survey, and the processing will be executed at the Archive Center. This includes Difference Imaging Analysis, generating deep Coadd Images, Source detection and association, creating Object and Solar System Object cataloors and related metadata
the results in a useful form. The infrastructure layer consists of the computing, storage, networking hardware, and system software. The middleware layer handles distributed processing, data access, user interface, and system operations services. The applications layer includes the data pipelines and the science data archives' products and services Data Release The approximately annual reprocessing of all LSST data, and the installation of the resulting data products in the LSST Data Access Centers, which marks the start of the two-year proprietary period Data Release Produc- tion An episode of (re)processing all of the accumulated LSST images, during which all output DR data products are generated. These episodes are planned to occur annually during the LSST survey, and the processing will be executed at the Archive Center. This includes Difference Imaging Analysis, generating deep Coadd Images, Source detection and association, creating Object and Solar System Object cataloos, and related metadata
The middleware layer handles distributed processing, data access, user interface, and system operations services. The applications layer includes the data pipelines and the science data archives' products and services           Data Release         The approximately annual reprocessing of all LSST data, and the installation of the resulting data products in the LSST Data Access Centers, which marks the start of the two-year proprietary period           Data Release         An episode of (re)processing all of the accumulated LSST images, during which all output DR data products are generated. These episodes are planned to occur annually during the LSST survey, and the processing will be executed at the Archive Center. This includes Difference Imaging Analysis, generating deep Coadd Images, Source detection and association, creating Object and Solar System Object cataloos. and related metadata
layer includes the data pipelines and the science data archives' products and services           Data Release         The approximately annual reprocessing of all LSST data, and the installation of the resulting data products in the LSST Data Access Centers, which marks the start of the two-year proprietary period           Data Release Produc- tion         An episode of (re)processing all of the accumulated LSST images, during which all output DR data products are generated. These episodes are planned to occur annually during the LSST survey, and the processing will be executed at the Archive Center. This includes Difference Imaging Analysis, generating deep Coadd Images, Source detection and association, creating Object and Solar System Object cataloos. and related metadata
Data Release         The approximately annual reprocessing of all LSST data, and the installation of the resulting data products in the LSST Data Access Centers, which marks the start of the two-year proprietary period           Data Release Produc- tion         An episode of (reprocessing all of the accumulated LSST images, during which all output DR data products are generated. These episodes are planned to occur annually during the LSST survey, and the processing will be executed at the Archive Center. This includes Difference Imaging Analysis, generating deep Coadd Images, Source detection and association, creating Object and Solar System Object cataloos, and related metadata
Centers, which marks the start of the two-year proprietary period Data Release Produc- tion An episode of (re)processing all of the accumulated LSST images, during which all output DR data products are generated. These episodes are planned to occur annually during the LSST survey, and the processing will be executed at the Archive Center. This includes Difference Imaging Analysis, generating deep Coadd Images, Source detection and association, creating Object and Solar System Object cataloos, and related metadata
Data Release Produc- tion An episode of (re)processing all of the accumulated LSST images, during which all output DR data products are generated. These episodes are planned to occur annually during the LSST survey, and the processing will be executed at the Archive Center. This includes Difference Imaging Analysis, generating deep Coadd Images, Source detection and association, creating Object and Solar System Object catalons, and related metadata
tion episodes are planned to occur annually during the LSST survey, and the processing will be executed at the Archive Center. This includes Difference Imaging Analysis, generating deep Coadd Images, Source detection and association, creating Object and Solar System Object catalogs and related metadata
includes Difference Imaging Analysis, generating deep Coadd Images, Source detection and association, creating Object and Solar System Object catalogs and related metadata
System Object catalogs, and related metadata
Cystem Cejset datalogo, and related metadata
Differential Chromatic The refraction of incident light by Earth's atmosphere causes the apparent position of objects to be shifted, and the size of this shift
Refraction depends on both the wavelength of the source and its airmass at the time of observation. DCR corrections are done as a part of DIA
DocuShare The trade name for the enterprise management software used by LSST to archive and manage documents
Document Any object (in any application supported by DocuShare or design archives such as PDMWorks or GIT) that supports project man-
agement or records milestones and deliverables of the LSST Project
EFD Engineering Facilities Database
EIA Early Integration Activity
ETC Estimate To Complete







EVMS	Earned Value Management System
Earned Value	A measurement of how much work has been completed compared to how much was expected to have been completed at a given
	point in the project
FGCM	Forward Global Calibration Model
FITS	Flexible Image Transport System
FPA	Focal Plane Array
FTE	Full Time Equivalent
Firefly	A framework of software components written by IPAC for building web-based user interfaces to astronomical archives, through which
	data may be searched and retrieved, and viewed as FITS images, catalogs, and/or plots. Firefly tools will be integrated into the
	Science Platform
GUI	Graphical User Interface
HEP	High Energy Physics
HSC	Hyper Suprime-Cam
ID	Identifier (Identification)
IPAC	Infrared Processing and Analysis Center
ISR	Instrument Signal Removal
IT	Integration Test
IVOA	International Virtual-Observatory Alliance
Instrument Signature	Instrument Signature Removal is a pipeline that applies calibration reference data in the course of raw data processing, to remove
Removal	artifacts of the instrument or detector electronics, such as removal of overscan pixels, bias correction, and the application of a flat-field
	to correct for pixel-to-pixel variations in sensitivity
L1	Level 1 (ambiguous could mean milestone or processing)
L2	Level 2 (ambiguous could mean milestone or processing)
LATISS	LSST Atmospheric Transmission Imager and Slitless Spectrograph







LCR	LSST Change Request
LDF	LSST Data Facility
LDM	LSST Data Management (handle for controlled documents)
LIT	LSST Issue Tracker
LOE	Level Of Effort
LPM	LSST Project Management (Document Handle)
LSE	LSST Systems Engineering (Document Handle)
LSP	LSST Science Platform
LSR	LSST System Requirements; LSE-29
LSST	Large Synoptic Survey Telescope
MERRA	Modern-Era Retrospective analysis for Research and Applications
MOC	Multi Ordered Catalogue
MOPS	Moving Object Pipeline System
MOU	Memo Of Understanding
MYDB	My Database, the notion of having a local storage beside the queriable database to store either temporary tables or uploaded
	catalogs
NASA	National Aeronautics and Space Administration
NCOA	National Center for Optical-Infrared Astronomy
NCSA	National Center for Supercomputing Applications
NET	NETworking
NOAO	National Optical Astronomy Observatories (USA)
NSF	National Science Foundation
OCS	Observatory Control System
OODS	Observatory Operations Data Service







OSS	Operations Support System
Object	In LSST nomenclature this refers to an astronomical object, such as a star, galaxy, or other physical entity. E.g., comets, asteroids
	are also Objects but typically called a Moving Object or a Solar System Object (SSObject). One of the DRP data products is a table
	of Objects detected by LSST which can be static, or change brightness or position with time
Offer	A response to a solicitation that, if accepted, would bind the offeror to perform the work described in resultant contract. Responses
	to sealed bidding are offers that are often referred to as 'bids' or 'sealed bids;' responses to a request for proposals (RFP, negotiated-
	type procurements) are offers often referred to as 'proposals' responses to a request for quotations (RFQ) are not offers and are
	generally called 'quotes'
Operations	The 10-year period following construction and commissioning during which the LSST Observatory conducts its survey
Operations Rehearsal	A data management system prototype project employing the same methods, tools, personnel, and technologies as the real system
	in order to introduce and validate new algorithms, functionality, and infrastructure. Previously referred to as a data challenge
Opportunity Manage-	The proactive art and science of planning, assessing, and handling future events to seek favorable impacts on project, cost, schedule,
ment	or performance to the extent possible. Opportunity management is a structured, formal, and disciplined activity focused on the
	necessary steps and planning actions to determine and exploit opportunities to the extent possible
PB	PetaByte
PDR	Preliminary Design Review
PDU	Power Distribution Unit
PSF	Point Spread Function
PST	Project Science Team
Parsl	Parallel Scripting Library http://parsl-project.org/
PipelineTask	A special kind of Task that can read its inputs and write its outputs using a Butler, in addition to being able to have them passed
	in and out directly as Python objects. PipelineTasks may be connected together dynamically and executed by a generic workflow
	system. PipelineTasks typically (but not always) delegate most of their work to nested regular Tasks






Project Manager	The person responsible for exercising leadership and oversight over the entire LSST project; he or she controls schedule, budget,
	and all contingency funds
Project Science Team	an operational unit within LSST that carries out specific scientific performance investigations as prioritized by the Director, the Project
	Manager, and the Project Scientist. Its membership includes key scientists on the Project who provide specific necessary expertise.
	The Project Science Team provides required scientific input on critical technical decisions as the project construction proceeds
Prompt Processing	The processing that occurs at the Archive Center on the nightly stream of raw images coming from the telescope, including Difference
	Imaging Analysis, Alert Production, and the Moving Object Processing System. This processing generates Prompt Data Products
QA	Quality Assurance
Qserv	Query Service, Proprietary LSST Database system
RM	Release Manager
RMS	Root-Mean-Square
Release	With regard to data pipelines or data products, a version that is cleared for distribution (i.e., has met QA specifications), is assigned
	a version identifier (e.g., 2.1), and does not evolve in the future to enable provenance
Review	Programmatic and/or technical audits of a given component of the project, where a preferably independent committee advises further
	project decisions, based on the current status and their evaluation of it. The reviews assess technical performance and maturity, as
	well as the compliance of the design and end product with the stated requirements and interfaces
Risk	The degree of exposure to an event that might happen to the detriment of a program, project, or other activity. Alt is described by
	a combination of the probability that the risk event will occur and the consequence of the extent of loss from the occurrence, or
	impact.ăRisk is an inherent part of all activities, whether the activity is simple and small, or large and complex
Risk Management	The art and science of planning, assessing, and handling future events to avoid unfavorable impacts on project cost, schedule, or
	performance to the extent possible. A Risk management is a structured, formal, and disciplined activity focused on the necessary
	steps and planning actions to determine and control risks to an acceptable level. Risk Management is an event-based management
	approach to managing uncertainty
SAC	Science Advisory Committee







SC	Science Collaboration
SDSS	Sloan Digital Sky Survey
SLAC	No longer an acronym; formerly Stanford Linear Accelerator Center
SODA	Server-side Operations for Data Access
SPI	Schedule Performance Index
SPIE	the international society for optics and photonics
SQuaRE	Science Quality and Reliability Engineering
SQuaSH	Science Quality Analysis Harness
SST	System Science Team
SUIT	Science User Interface and Tools
Science Collaboration	An autonomous body of scientists interested in a particular area of science enabled by the LSST dataset, which through precursor
	studies, simulations, and algorithm development lays the groundwork for the large-scale science projects the LSST will enable. In
	addition to preparing their members to take full advantage of LSST early in its operations phase, the science collaborations have
	helped to define the system's science requirements, refine and promote the science case, and quality check design and development
	work
Science Pipelines	The library of software components and the algorithms and processing pipelines assembled from them that are being developed by
	DM to generate science-ready data products from LSST images. The Pipelines may be executed at scale as part of LSST Prompt
	or Data Release processing, or pieces of them may be used in a standalone mode or executed through the LSST Science Platform.
	The Science Pipelines are one component of the LSST Software Stack
Science Platform	A set of integrated web applications and services deployed at the LSST Data Access Centers (DACs) through which the scientific
	community will access, visualize, and perform next-to-the-data analysis of the LSST data products
Scope	The work needed to be accomplished in order to deliver the product, service, or result with the specified features and functions
Solar System Object	A solar system object is an astrophysical object that is identified as part of the Solar System: planets and their satellites, asteroids,
	comets, etc. This class of object had historically been referred to within the LSST Project as Moving Objects







Source	A single detection of an astrophysical object in an image, the characteristics for which are stored in the Source Catalog of the DRP
	database. The association of Sources that are non-moving lead to Objects; the association of moving Sources leads to Solar System
	Objects. (Note that in non-LSST usage "source" is often used for what LSST calls an Object.)
Source Association	The process of associating source detections on multiple images taken at different epochs, or in multiple passbands, with a single astronomical Object
Specification	One or more performance parameter(s) being established by a requirement that the delivered system or subsystem must meet
Stripe 82	A 2.5ř wide equatorial band of sky covering roughly 300 square degrees that was observed repeatedly in 5 passbands during the course of the SDSS, In part for calibration purposes
Subsystem	Aăset of elements comprising a system within the larger LSST system that is responsible for a key technical deliverable of the project
Subsystem Scientist	The principal science advisor to a Subsystem Manager; he or she ensures that the subsystem specifications are appropriated for
	achieving the project's goals
Summit	The site on the Cerro Pachón, Chile mountaintop where the LSST observatory, support facilities, and infrastructure will be built
T&S	Telescope and Site
T/CAM	Technical/Control (or Cost) Account Manager
ТАР	Table Access Protocol
ТВ	TeraByte
Task	Tasks are the basic unit of code re-use in the LSST Stack. They perform a well defined, logically contained piece of functionality.
	Tasks come standard with configuration, logging, processing metadata, and debugging features. For further details, see How to
	Write a Task in the source code documentation. Tasks can be nested, providing a natural way to structure - and configure - high level
	algorithms that delegate work to lower-level algorithms
UI	User Interface
US	United States
VCD	Verification Control Document
VNOC	Virtual Network Operations Center







VO	Virtual Observatory
VOIP	Voice Over Internet Protocol
Validation	A process of confirming that the delivered system will provide its desired functionality; overall, a validation process includes the
	evaluation, integration, and test activities carried out at the system level to ensure that the final developed system satisfies the intent
	and performance of that system in operations
Verification	The process of evaluating the design, including hardware and software - to ensure the requirements have been met; verification (of
	requirements) is performed by test, analysis, inspection, and/or demonstration
WBS	Work Breakdown Structure
WCS	World Coordinate System
WISE	Wide-field Survey Explorer
algorithm	A computational implementation of a calculation or some method of processing
arcmin	arcminute minute of arc (unit of angle)
astrometry	In astronomy, the sub-discipline of astrometry concerns precision measurement of positions (at a reference epoch), and real and
	apparent motions of astrophysical objects. Real motion means 3-D motions of the object with respect to an inertial reference frame;
	apparent motions are an artifact of the motion of the Earth. Astrometry per se is sometimes confused with the act of determining a
	World Coordinate System (WCS), which is a functional characterization of the mapping from pixels in an image or spectrum to world
	coordinate such as (RA, Dec) or wavelength
background	In an image, the background consists of contributions from the sky (e.g., clouds or scattered moonlight), and from the telescope
	and camera optics, which must be distinguished from the astrophysical background. The sky and instrumental backgrounds are
	characterized and removed by the LSST processing software using a low-order spatial function whose coefficients are recorded in
	the image metadata
brighter-fatter effect	The common term used to refer to one of the photometric qualities of the LSST camera: sources with a higher flux have a broader
	PSF. This is accounted for during calibration







calibration	The process of translating signals produced by a measuring instrument such as a telescope and camera into physical units such
	as flux, which are used for scientific analysis. Calibration removes most of the contributions to the signal from environmental and
	instrumental factors, such that only the astronomical component remains
configuration	A task-specific set of configuration parameters, also called a 'config'. The config is read-only; once a task is constructed, the same
	configuration will be used to process all data. This makes the data processing more predictable: it does not depend on the order in
	which items of data are processed. This is distinct from arguments or options, which are allowed to vary from one task invocation to
	the next
jointcal	The jointcal package optimizes the astrometric and photometric calibrations of a set of astronomical images that cover a sky tract
	and were obtained as a series of visits, which may be spread out in time. The jointcal algorithms incorporates object matching both
	between visits and to reference star catalogs, and produces more accurate distortion and throughput models than if the astrometry
	and photometry were fit independently. Jointcal is a part of the Science Pipelines
metadata	General term for data about data, e.g., attributes of astronomical objects (e.g. images, sources, astroObjects, etc.) that are charac-
	teristics of the objects themselves, and facilitate the organization, preservation, and query of data sets. (E.g., a FITS header contains
	metadata)
metric	A measurable quantity which may be tracked. A metric has a name, description, unit, references, and tags (which are used for
	grouping). A metric is a scalar by definition. See also: aggregate metric, model metric, point metric
monitoring	In DM QA, this refers to the process of collecting, storing, aggregating and visualizing metrics
pipeline	A configured sequence of software tasks (Stages) to process data and generate data products. Example: Association Pipeline
point spread function	The point-spread function (PSF) is the distribution of intensity on a sensor (or image) originating from an unresolved point-source
	(i.e., a star). Often the PSF is not the same Airy shape as would be expected from a finite-aperture optical system, owing primarily
	to atmospheric effects and imperfections in the optical system and the detector
stack	a grouping, usually in layers (hence stack), of software packages and services to achieve a common goal. Often providing a higher
	level set of end user oriented services and tools
transient	A transient source is one that has been detected on a difference image, but has not been associated with either an astronomical
	object or a solar system body







- Antilogus, P., Astier, P., Doherty, P., Guyonnet, A., Regnault, N., 2014, Journal of Instrumentation, 9, C03048 (arXiv:1402.0725), doi:10.1088/1748-0221/9/03/C03048, ADS Link
- [DMTR-91], Bellm, E., 2018, LDM-503-05 (Alert Distribution Validation) Test Report, DMTR-91, URL https://ls.st/DMTR-91
- Burke, D.L., Rykoff, E.S., Allam, S., et al., 2018, AJ, 155, 41 (arXiv: 1706.01542), doi:10.3847/1538-3881/aa9f22, ADS Link
- [DMTN-106], Comoretto, G., 2019, DM Release Process, DMTN-106, URL http://DMTN-106.lsst.io
- [LSE-68], Dubois-Felsmann, G., 2015, Camera Data Acquisition Interface, LSE-68, URL https://ls.st/LSE-68
- [LSE-140], Dubois-Felsmann, G., 2016, Auxiliary Instrumentation Interface between Data Management and Telescope, LSE-140, URL https://ls.st/lse-140
- [LSE-72], Dubois-Felsmann, G., Schumacher, G., Selvy, B., 2014, OCS Command Dictionary for Data Management, LSE-72, URL https://ls.st/LSE-72
- [DMTN-109], Eggl, S., 2019, LSST discovery estimates for Solar System Objects, DMTN-109, URL http://dmtn-109.lsst.io
- Holman, M.J., Payne, M.J., Blankley, P., Janssen, R., Kuindersma, S., 2018, AJ, 156, 135, doi:10.3847/1538-3881/aad69a, ADS Link
- [LSE-131], Jacoby, S., Emmons, B., Selvy, B., 2017, Interface between Data Management and Education and Public Outreach, LSE-131, URL https://ls.st/LSE-131
- [DMTN-087], Juric, M., Jones, L., 2018, Proposed Modifications to Solar System Processing and Data Products, DMTN-087, URL https://dmtn-087.lsst.io, LSST Data Management Technical Note
- [LPM-72], Krabbendam, V., 2015, Scope Options, LPM-72, URL https://ls.st/LPM-72
- [DMTN-111], Lim, K.T., 2019, DM Usage in Observatory Operations, DMTN-111, URL http://dmtn-111.lsst.io
- [LSE-400], Lim, K.T., 2019, Header Service Interface, LSE-400, URL https://lse-400.lsst.io
- Melchior, P., Moolekamp, F., Jerdee, M., et al., 2018, Astronomy and Computing, 24, 129 (arXiv:1802.10157), doi:10.1016/j.ascom.2018.07.001, ADS Link

[DMTN-031], Morrison, C.B., 2018, Pessimistic Pattern Matching for LSST, DMTN-031, URL https://dmtn-031.lsst.io, LSST Data Management Technical Note



## References II



[DMTN-096], O'Mullane, W., Swinbank, J., Guy, L., 2018, Implementation and impacts of DM scope options, DMTN-096, URL https://dmtn-096.lsst.io, LSST Data Management Technical Note

[LDM-294], O'Mullane, W., Swinbank, J., Jurić, M., DMLT, 2018, Data Management Organization and Management, LDM-294, URL https://ls.st/LDM-294

- [DMTN-093], Patterson, M., Bellm, E., Swinbank, J., 2018, Design of the LSST Alert Distribution System, DMTN-093, URL https://dmtn-093.lsst.io, LSST Data Management Technical Note
- [DMTN-028], Patterson, M.T., 2018, Benchmarking a distribution system for LSST alerts, DMTN-028, URL https://dmtn-028.lsst.io, LSST Data Management Technical Note

[DMTN-121], Sullivan, I., 2019, Impact of variable seeing on DCR coadd generation, DMTN-121, URL http://dmtn-121.lsst.io

[DMTR-111], Swinbank, J., 2018, LDM-503-09a (Science pipelines fall 2018 release) Test Plan and Report, DMTR-111, URL https://ls.st/DMTR-111

[DMTR-112], Swinbank, J., 2018, LDM-503-07 (Camera data processing) Test Plan and Report, DMTR-112, URL https://ls.st/DMTR-112

[LDM-151], Swinbank, J.D., et al., 2017, Data Management Science Pipelines Design, LDM-151, URL https://ls.st/LDM-151