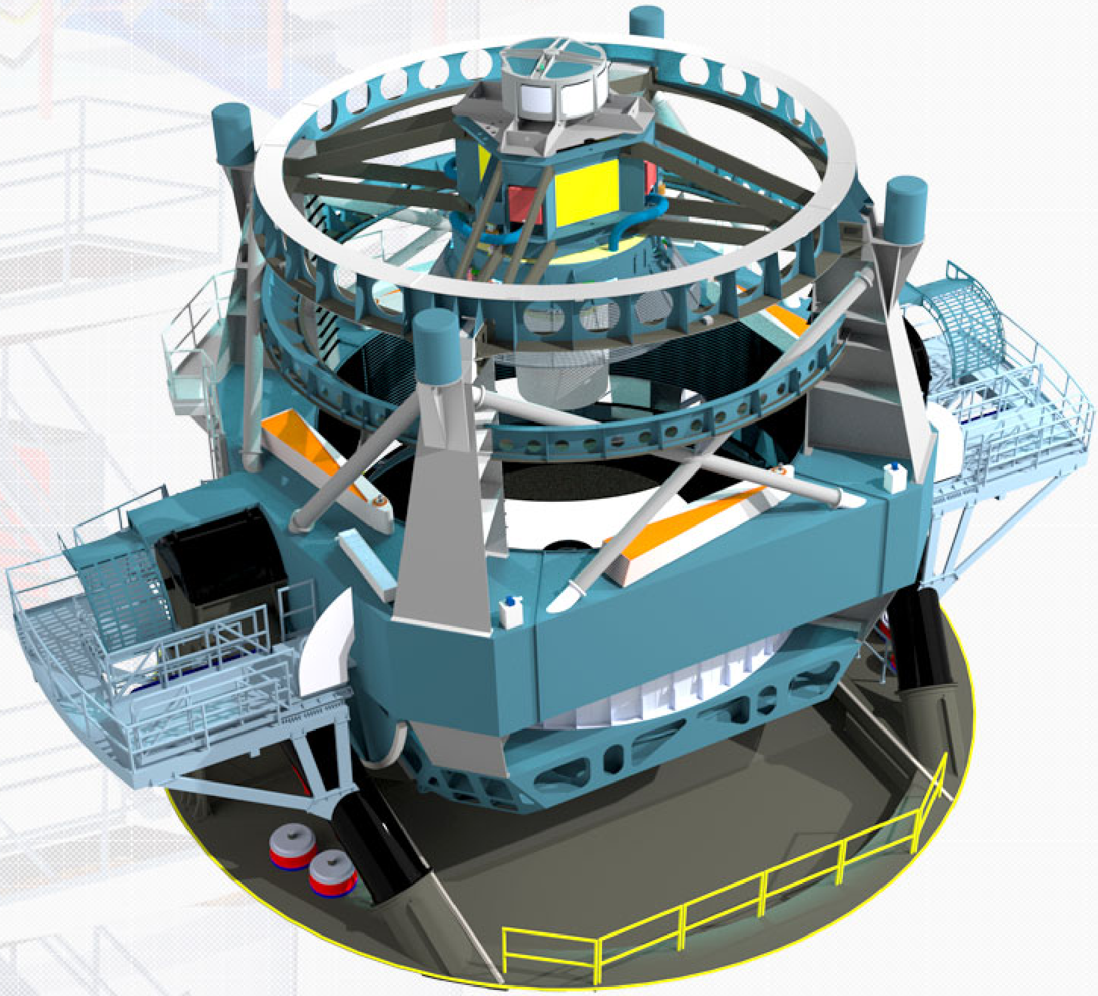
Computing for ngVLA: Lessons from LSST

Jeffrey Kantor,

LSST Data Management Project Manager

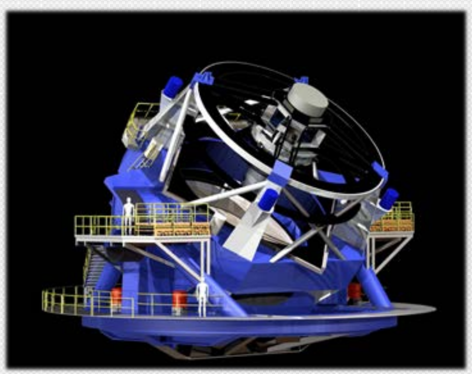


A Dedicated Survey Telescope

- **A wide (half the sky), deep (24.5/27.5 mag), fast (image the sky once every 3 days) survey telescope. Beginning in 2022, it will repeatedly image the sky for 10 years.**

- The LSST is an integrated survey system. The Observatory, Telescope, Camera and Data Management system are all built to support the LSST survey. There's no PI mode, proposals, or time.

- **The ultimate deliverable of LSST is not the telescope, nor the instruments; it is the <u>fully reduced data</u>.**
  - All science will be come from survey catalogs and images



Telescope ➔

Images ➔

Table 4: Level 2 Catalog Object Table

Name	Type	Unit	Description
psRadecTai	double	time	Point source model: Time at which the object was at position radec.
psPm	float[2]	mas/yr	Point source model: Proper motion vector.
psParallax	float	mas	Point source model: Parallax.
psFlux	float[ugrizy]	nmgy	Point source model fluxes[58].
psCov	float[66]	various	Point-source model covariance matrix[59].
psLnL	float		Natural *log* likelihood of the observed data given the point source model.
bdRadec	double[2]	degrees	B+D model[60]: (α, δ) position of the object at time radecTai, in each band.

Catalogs

Open Data, Open Source: A Community Resource

- LSST data, including images and catalogs, will be available with no proprietary period to the astronomical community of the United States, Chile, and International Partners

- Alerts to variable sources ("transient alerts") will be available world-wide within 60 seconds, using standard protocols

- LSST data processing stack will be free software (licensed under the GPL, v3-or-later)

- **All science will be done by the community (not the Project!), using LSST's data products.**

LSST From a Scientist's Perspective

- A stream of ~10 million time-domain events per night, detected and transmitted to event distribution networks within 60 seconds of observation.
- A catalog of orbits for ~6 million bodies in the Solar System.

Level 1

- A catalog of ~37 billion objects (20B galaxies, 17B stars), ~7 trillion observations ("sources"), and ~30 trillion measurements ("forced sources"), produced annually, accessible through online databases.
- Deep co-added images.

Level 2

- Services and computing resources at the Data Access Centers to enable user-specified custom processing and analysis.
- Software and APIs enabling development of analysis codes.

Level 3

Archive Site

NCSA

Archive Center

Alert Production
Data Release Production (50%)
Calibration Products Production
EPO Infrastructure
Long-term Storage (copy 2)

Data Access Center

Data Access and User Services

Satellite Processing Center
CC-IN2P3, Lyon, France

CCIN2P3 Data Release Production (50%)
French DAC

Dedicated Long Haul Networks

Two redundant 100 Gbit links from Santiago to Florida (existing fiber)
Additional 100 Gbit link (spectrum on new fiber) from Santiago – Florida
(Chile and US national links not shown)

HQ Site

LSST Large Synoptic Survey Telescope

Science Operations
Observatory Management
Education and Public Outreach

Summit and Base Sites

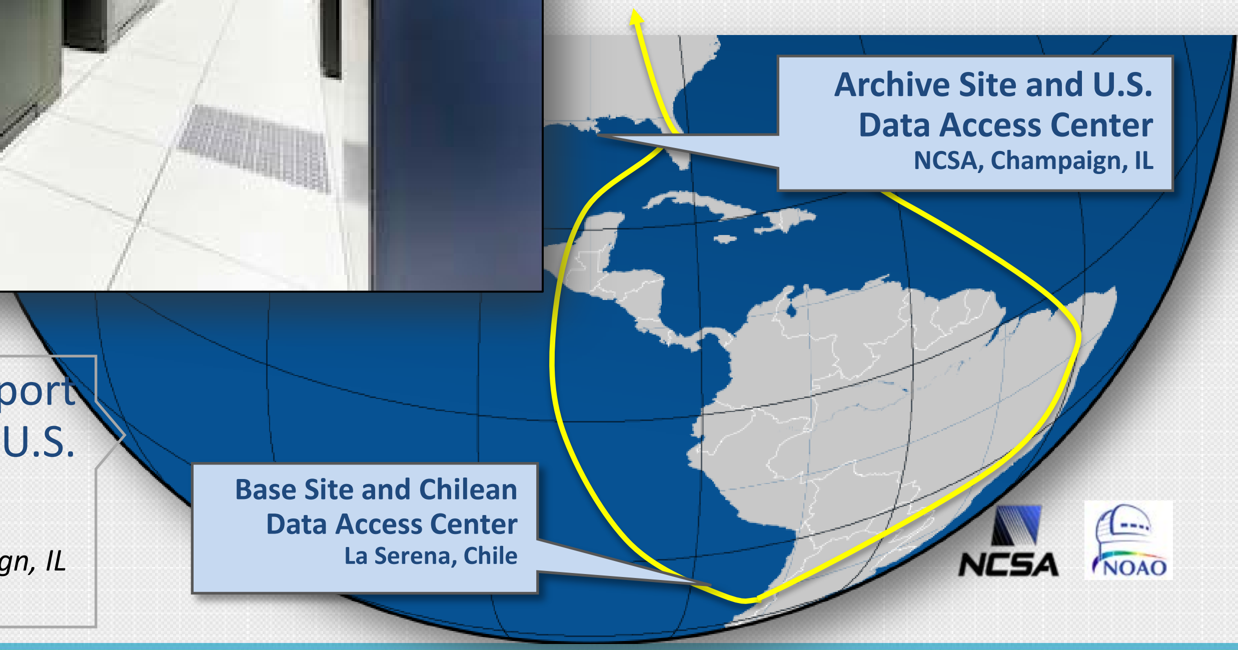
LSST Large Synoptic Survey Telescope

Telescope and Camera
Data Acquisition
Crosstalk Correction
Long-term storage (copy 1)
Chilean Data Access Center

Infrastructure: Petascale Computing, Gbit Networks

The computing cluster at the ***LSST Archive*** *(at NCSA) will run the processing pipelines.*

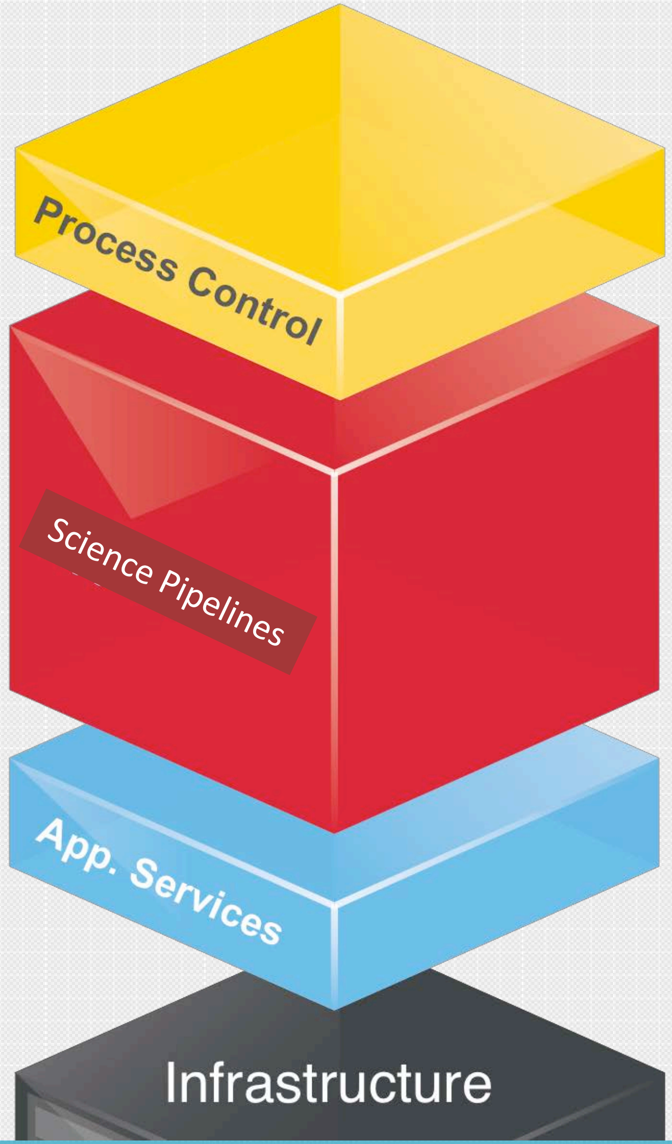
- *Single-user, single-application, dedicated data center*
- *Process images in real-time to detect changes in the sky*
- *Produce annual data releases*



Long Haul Networks to transport data from Chile to the U.S.

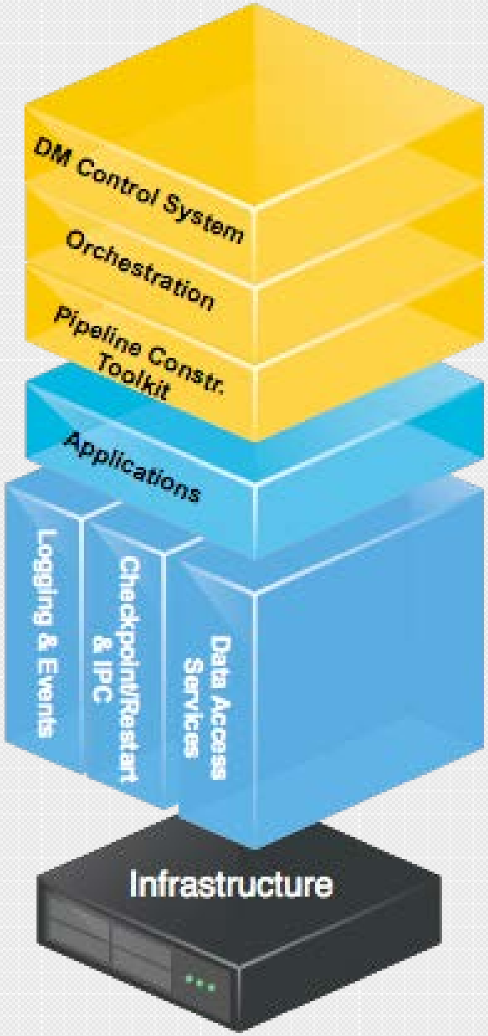
- *200 Gbps from Summit to La Serena (new fiber)*
- *2x100 Gbit (minimum) for La Serena to Champaign, IL (protected, existing fiber)*

Science Pipelines: Scientific Core of LSST DM



- *Science Pipelines* carry core scientific algorithms that process or analyze raw LSST data to generate output Data Products
- N.b. also referred to as "Applications"
- Variety of processing
  - Image processing
  - Measurement of source properties
  - Associating sources across space and time, e.g. for tracking solar system objects

Middleware Layer: Isolating Hardware, Orchestrating Software



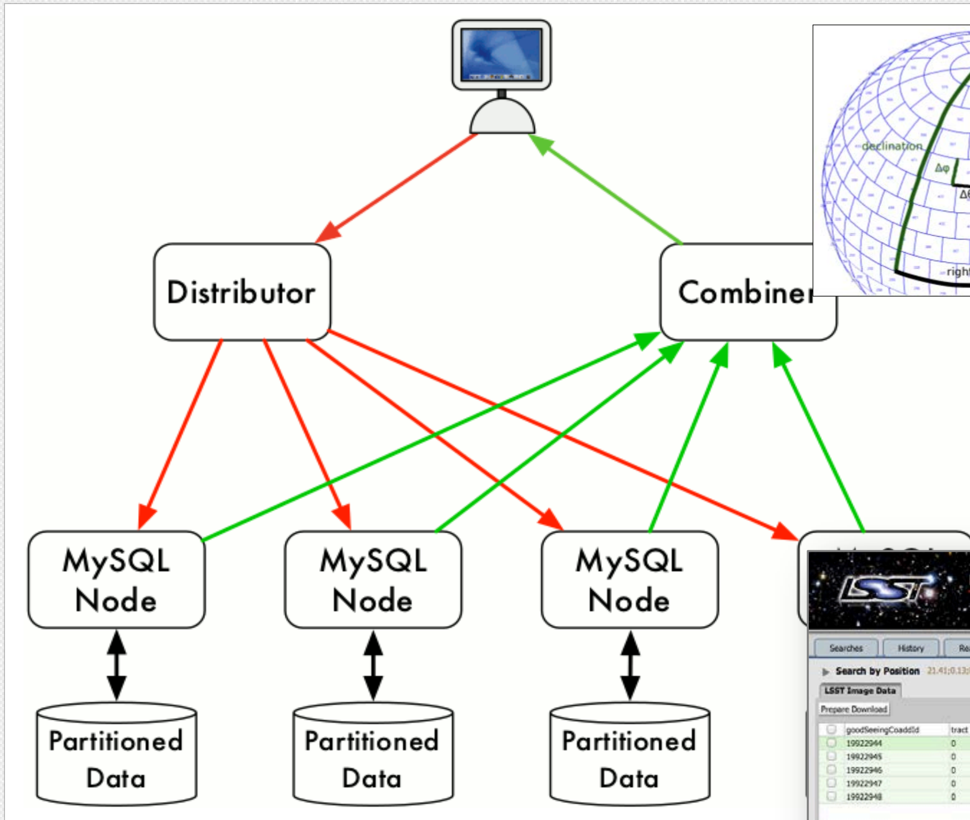
Enabling execution of science pipelines on hundreds of thousands of cores.

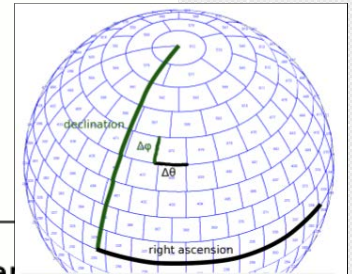
- Frameworks to construct pipelines out of basic algorithmic components
- Orchestration of execution on thousands of cores
- Control and monitoring of the whole DM System

Isolating the science pipelines from details of underlying hardware

- Services used by applications to access/produce data and communicate
- "Common denominator" interfaces handle changing underlying technologies

Database and Science UI: Delivering to Users





*Massively parallel, distributed, fault-tolerant **relational database**.*

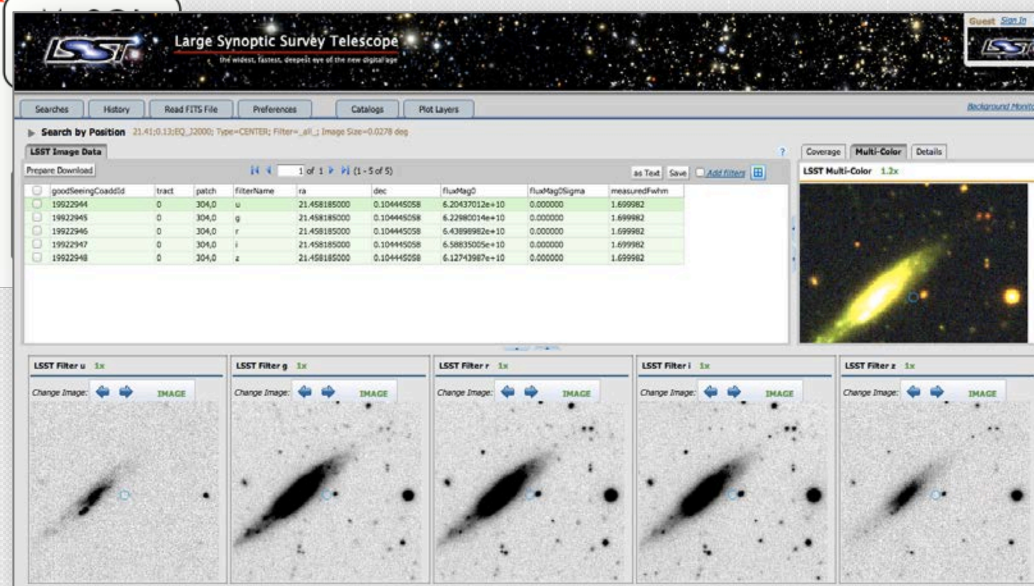
- To be built on existing, robust, well-understood, technologies (MySQL and xrootd)
- Commodity hardware, open source
- Advanced prototype in existence (qserv)

SLAC

***Science User Interface** to enable the access to and analysis of LSST data*

- Web and machine interfaces to LSST databases
- Visualization and analysis capabilities

ipac



Distributed Team, SQuaREd

LSST

ipac — UI

SLAC — Database

UNIVERSITY of WASHINGTON — Core Algorithms (“Apps”)

NCSA, SLAC — Middleware

NCSA, NOAO, CCIN2P3 — Infrastructure

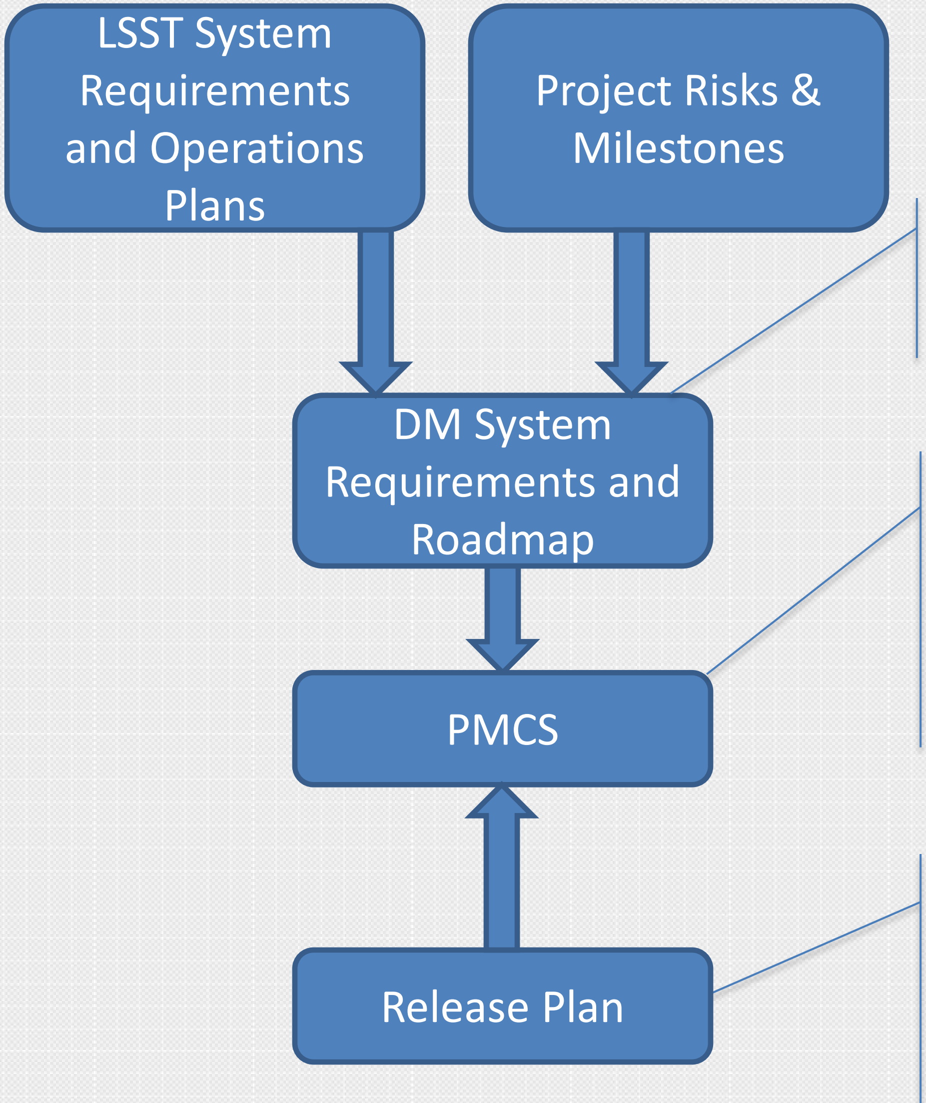
LSST Large Synoptic Survey Telescope

Science QA & Reliability Engineering (SQuaRE)

Tools and Processes

- Mailing lists, wikis, Confluence, telecons, videocons, …
- Github
- Jenkins
- JIRA planning and Earned Value
- Licensing and contributions
- Investigating Slack and Discourse

Data Management Planning Process/Levels



- DM Roadmap (LDM-240)
- High-level picture
- Granularity: 12 months

- Institutional-level resource assignments
- Budgeting
- Earned-value management
- Basis of monthly reports
- Granularity: 3 - 6 months
- 2 x 6-month development/release cycle per year
- 10-20 activities for each cycle

- Detailed 6-month plan
- Developed from Agile Process (Epics, Stories)
- Individual resource assignments
- ~100 activities per cycle
- Granularity of activities: 1 – 3 months
- Granularity of steps within activities: 2 – 20 days

LSST Software Development Roadmap (LDM-240)

LSST Data Management Software Development Roadmap

LDM-240 v20

Release	Institution	R4.0	R5.0/5.1	R6.0/6.1	R7.0/7.1	R8.0/8.1 - #1 Release for ComCam	R9.0/9.1 – #2 Release for ComCam	R10.0/10.1 – LSST Camera
Finish Date		8/31/14	8/31/15	8/31/16	8/31/17	8/31/18	8/31/19	8/31/20
Year during which work is done		FY14	FY15	FY16	FY17	FY18	FY19	FY20
02C.03.01 Single Frame Processing Pipelines	University of Washington		flat, bias) Source detection and measurement of PSF flux, aperture flux, adaptive moments WCS determination module capable of fitting WCS to multiple CCDs at a time (in a single visit). Snap processing, CR detection and removal Improved sky background estimation Preliminary deblender	WCS terms needed to describe distortions at edges of chips and near the bloom stop Astrometric registration of a stack of images. Pixel-level intensity-dependent PSF correction Multi-CCD sky background detemination Cross-talk removal Fringe removal	Recognition and retention of trails (moving objects) in visit creation ISR functional with as-built CCD and raft characteristics (based on characterization of sensor/raft prototypes) Measureme functional with as CD and raft ch s.			
02C.03.02 Association Pipeline	University of Washington	Source association and matchi refacto				Association of DIASources to DIAObjects Association of DIAObjects with Objects	Scalability and speed satisfy transient alert latency requirement	
02C.03.03 Alert Generation Pipeline	University of Washington						Transient Alert Packet creation and transmission capability Initial implementation of simple Event Broker	Scalability and speed satisfy transient alert latency requirement Full implementation of simple Event Broker, ability to receive and broker events under expected event and user load.
				Source detection on difference images using	artifacts (e.g., bright stars or ghosts)	Spuriousness metric	Selection criteria and strategy for building templates Forced photometry on	

← Feature Implementation →

← Performance Improvement →

- Major axes are 12-month cycles x WBS
- Cells represent milestones/major results
- Cells to left are predecessors to cells to the right

http://ls.st/ldm-240

LDM-240 DM Key Performance Metrics

LSST Data Management Key Performance Metrics

Release	Unit	Contributing Elements	R4.0	R5.0/5.1	R6.0/6.1	R7.0/7.1	R8.0/8.1 - #1 Release for ComCam	R9.0/9.1 - #2 Release for ComCam	R10.0/10.1 - LSST Camera	Operational Readiness
Elapsed time (years)			1	2	3	4	5	6	7	
Finish Date			8/31/2014	8/31/2015	8/31/2016	8/31/2017	8/31/2018	8/31/2019	8/31/2020	
Year during which work is done			FY14	FY15	FY16	FY17	FY18	FY19	FY20	
Photometric repeatability										
procCalRep	mmag	SFM	15	13	10	5	5	4	3.5	3
PA1gri	mmag	SFM, CPP, PhotoCal	15	13	12	8	8	6	5.5	5
PA1uzy	mmag	SFM, CPP, PhotoCal	15	13	14	12	12	8	8	7.5
Photometric Spatial Uniformity										
PA3u	mmag	PhotoCal					40	30	20	20
PA3 (g)	mmag	PhotoCal					30	15	10	10
PA3 (y)	mmag	PhotoCal					30	15	10	10
Color Zero-point Accuracy										
PA5	mmag	PhotoCal					30	20	10	5
PA5u	mmag	PhotoCal					50	30	15	10
Absolute Photometry Accuracy										
PA6	mas	PhotoCal						30	20	10
Relative Astrometry										
AM1	mas	SFM, ObjChar	80	60			30	20	15	10
AM2	mas	SFM, ObjChar	80	60			30	20	15	10
AM3	mas	SFM, ObjChar, AstroCal	100							
AB1	mas	SFM, ObjChar, AstroCal								
Absolute Astrometry										
AA1	mas	AstroCal								
Residual PSF Ellipticity Correlations										
TE1	-	PsfEst, ObjChar, Coadd								
TE2	-	PsfEst, ObjChar, Coadd								
Moving Object Linkage Efficiency										
orbitCompleteness	%	MOPS								
Spuriousness Metric Efficiency										
transCompletenessMin	%	Diffim								
transPurityMin	%	Diffim								
mopsCompletenessMin	%	Diffim								
mopsPurityMin	%	Diffim								
Computational Performance Metrics										
OTT1	seconds	All AP Pipelines + MW + Infra		240	240	240	180	120	90	60
AP computational budget	TFLOPS	All AP Pipelines + MW		231	193	154	77	58	58	39
DRP computational budget (DR1)	TFLOPS	All DRP+AP Pipelines		645	215	161	151	129	108	108
LV query rate	simult. queries	qserv		50	50	50	70	80	100	100
LV query response time	seconds	qserv		18	15	12	12	12	10	10
HV shared scan duration	hours	qserv		24	24	20	18	14	12	12
HV simultaneous queries	queries	qserv		6	6	6	8	10	12	12
System Reliability Metrics										
sciVisitAlertDelay	%	All AP Pipelines						10	5	1
sciVisitAlertFailure	%	All AP Pipelines						2	0.5	0.1
sciImageLoss	%	Image and File Archive						2	1	1

- Major axes are 12-month cycles x metric
- Cells represent planned achievement of metric
- Progressively increasing performance from to left to right
- SRD minimum spec is met first, then design spec, then (if applicable) stretch

EV Data Flow

Budget BCWS

1) Create Epics In JIRA

Epic #	Description	Freemon M LST NCSA	Pietrowicz S LSEN NCSA	Petravick D LMA NCSA	Wefel P LTE NCSA	resource loading [L-month]	sprint 3	sprint 4	sprint 5	sprint 6	sprint 7	sprint 8	WBS
DM 2206	Deploy and test network emulation for nightly p	160	160	0	0	320	0	0	160	160	0	0	02C.07.01.01
DM-2207	Alert Production Simulator	0	384	0	0	384	96	96	96	96	0	0	02C.07.01.01
DM-2208	Complex Event Processing	0	124	0	0	124	0	62	62	0	0	0	02C.07.02.01
DM-2209	OCS Software	0	40	0	0	40	0	0	40	0	0	0	02C.07.01.01
DM-1120	State Diagram (DM-1120)	0	24	0	0	24	24	0	0	0	0	0	02C.07.01.01
DM-1121	Refactor Event Services (DM-1121)	0	8	0	0	8	8	0	0	0	0	0	02C.07.02.01
DM 1122	Write log4cxx handler that emits events for log	0	8	0	0	8	8	0	0	0	0	0	02C.07.01.03.02
DM-2210	Configuration Management (Puppet)	0	0	0	0	14.4	7.2	7.2	0	0	0	0	02C.07.02.04
DM-1270	VMware Infrastructure (DM-1270)	0	0	0	0	15.6	0	0	16	0	0	0	02C.07.02.04
DM-1274	Server Room Refresh (DM-1274)	0	0	0	0	26	0	0	0	13	13	0	02C.07.02.04
DM-1273	OpenStack Deployment Plan (DM-1273)	0	0	0	32	115.2	0	0	29	29	29	29	02C.07.02.04

2) Complete Resource Loading Workbook

3) Epics and resources are loaded into Primavera

4) Resource time-phasing is Baselined in Cobra as the Budget (BCWS)

Earned BCWP

1) Status Stories in JIRA Completed

2) Stories are loaded in Primavera as activity steps

3) Status is Integrated into Cobra at the Work Package level and the basis for the Earned (BCWP) Calculation

Actuals ACWP

1) Accounting records all actual costs

Institution Invoice and WBS/Account Breakout

2) Monthly export of Actuals is loaded into Cobra and is the basis for Actuals (ACWP)

What systems supply these numbers???

	Cumulative to Date				
WBS / WP	Budget BCWS	Earned BCWP	Actuals ACWP	SV	CV
1.02C	**$5,460,515**	**$4,625,024**	**$4,842,382**	**($835,491)**	**($217,358)**
1.02C.01	**$858,053**	**$844,155**	**$975,825**	**($13,898)**	**($131,670)**
1.02C.01.01	$516,460	$745,968	$741,571	$229,508	$4,397
1.02C.01.02	$341,593	$98,188	$234,255	($243,406)	($136,067)
1.02C.02	**$468,091**	**$381,645**	**$136,697**	**($86,446)**	**$244,949**
1.02C.02.01	$102,727	$102,727	$5,340	$0	$97,387
1.02C.02.02	$365,365	$278,919	$131,357	($86,446)	$147,562
1.02C.03	**$339,445**	**$151,181**	**$262,466**	**($188,264)**	**($111,285)**
1.02C.03.00	$64,192	$60,027	$131,347	($4,165)	($71,320)
1.02C.03.01	$31,882	$31,882	$10,548	$0	$21,334

Roadmap

- 2016: Finish end-to-end, start refining components, developer multipliers, run easily on precursor data
- 2017: Reprocess DES, continuous simulated Alert Production
- 2018: Ready for ComCam
- 2019: Ready for full Camera
- 2020: Ready for Science Verification

Team Culture (emphasis added, truncated)

The ***LSST DM community includes LSST-paid staff at multiple institutions and external contributors*** from around the world. We have built a strong team that collaborates well. As we hire new people, we want to make sure we are maintaining a healthy, supportive, productive culture. While ***culture is best transmitted by daily example***, having some formal standards for conduct can aid newcomers and reinforce good patterns.

Since LSST is an AURA center, the ***AURA standards of workplace conduct*** provide a starting point. As stated there, we in DM dedicate ourselves to fostering a ***civil and inclusive community characterized by mutual respect for the contributions of all individuals***. As a community, we embrace the values in that document, in addition to any ***local institutional standards***.

Lessons Learned

- AURA/LSST is in many ways a completely new entity, with inherent start-up activities
- Hiring has been slow, particularly for lead roles
- Team is not yet developing at maximum productivity
- It is hard to mesh scientists, managers, and engineers on one team
- Algorithmic progress is hard to measure and even harder to predict
- Architectures are evolving rapidly: e.g. Cloud, AstroPy, Hadoop ecosystem
- It is hard to balance technical priorities, trading off between performance/scale versus adaptability and usability

- Reorganize work with delayed staffing, relieve current leads of less-critical tasks
- Focus on developer multipliers
- Establish team culture and norms
- Focus on "end-to-end" system
- Document design path and replan
- Track new architectures, maintain portability
- Make sure we don’t accumulate too much “technical debt”

18