Large Synoptic Survey Telescope

Animated Video credit: SLAC National Accelerator Lab
Large Synoptic Survey Telescope
Annual Data Release Products
11 Data releases in 10 years.
Final catalog: 15PB
Final pixels: 500PB

Prompt Data Products via nightly alert streams
~10 million alerts per night
issued within 60 s of shutter close
LSST Data Management System
LSST Science Platform technology emerged
LSST Science Platform

Data access via IVOA-standard protocols

Same interfaces that support other aspects

Portal

Jupyter Notebooks

Web APIs

- Data access via IVOA-standard protocols
- Same interfaces that support other aspects
LSST DM software uses open source best practices

~500000 lines of Python & C++
Agile Principles
DevOps Engineering

Jenness, Economou et al. 2018
doi: 10.1117/12.2312157
LSST DM ~100 devs across 5 institutions

Princeton University
NCSA
University of Washington
SLAC
NOAO

Tucson - Project Office
PRINCETON UNIVERSITY
UNIVERSITY OF WASHINGTON

HOW2019 - MARCH 18 2019 - JEFFERSON LAB
Darling technologies come and go

Google Trends for svn vs. git
We get value from following open source trends

2011

2014

2016
LSST DM Developer Guide

This is an internal guide for LSST DM staff. It’s also openly available so that others can understand how we’re building the LSST’s data management subsystem.

This guide includes a mix of normative requirements and helpful, descriptive, pages. When it’s particularly important that you closely follow a standard, we include an annotation box at the top of the page.

Any member of DM can contribute to this guide. It’s published from the https://github.com/lsst-dm/dm_dev_guide GitHub repo. Check out the README to get started.
DM Development Workflow with Git, GitHub, JIRA and Jenkins

This page describes our procedures for collaborating on LSST DM software and documentation with Git, GitHub and JIRA:

1. Configuring Git for DM development.
2. Using JIRA for agile development.
3. DM GitHub organizations.
4. Policies for naming and using Git branches.
5. Preparing code for review.
6. Reviewing and merging code.

In appendices, we suggest some best practices for maximizing the usefulness of our Git development history:

- Commit organization best practices.
- Commit message best practices.
Our outstanding challenges

• **Workflow management**: taking shape on 3rd try based on Pegasus
• **Software**: Release management. Quick deployment during commissioning.
• **Data management**: How expectations around data access will change?
  • How to be flexible. Column stores?
• **Workforce** (e.g. careers of the postdocs writing the software)
If you see me later, ask:

- for a demo of the JupyterHub/Kubernetes-deployed Science Platform
- about adoption of Machine Learning in the astronomy community and plans in LSST
- why it's important to choose technologies supported by the private sector and open source communities

If you see Margaret Johnson (NCSA) around, ask:

- About data facility challenges, data storage and compute plans.
- See talk Wednesday 2-3:30 OSG parallel
In Summary

The HEP and Astro communities share challenges and can learn from each other.

Find us:
@lsst
github.com/lsst
www.lsst.io for docs
Data and compute sizes:
- Final volume of raw image data = 60 PB
- Final image collection (DR11) = 0.5 Exabytes
- Final catalog size (DR11) = 15 PB
- Final disk storage = 0.4 Exabytes
- Peak number of nodes = 1750 nodes
- Peak compute power in LSST data centers = about 2 PFLOPS
18 months until first light
Then we start 2 years of commissioning

So long M1M3! The @LSST 8.4-meter mirror left Tucson early this morning on a transport vehicle bound for Houston. With such a wide and fragile cargo load it will take about 10 days to make the trip! Read more at ow.ly/jp0V30o2SXm #NSFScience
LSST Operations: Sites & Data Flows

HQ Site
Tucson, AZ
- Science Operations
- Observatory Management
- Education & Public Outreach

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

French Site
CC-IN2P3, Lyon, France
- Satellite Processing Center
  - Data Release Production
  - Long-term Storage (copy 3)

LSST Data Facility
National Center for Supercomputing Applications (NCSA), Urbana-Champaign, IL
- Processing Center
  - Alert Production
  - Data Release Production
  - Calibration Products Production
  - EPO Infrastructure
  - Long-term Storage (copy 2)
- Data Access Center
  - Data Access and User Services

Summit Site
Cerro Pachón, Chile
- Telescope & Camera
  - Data Acquisition
  - CrossTalk Collection