The Large Synoptic Survey Telescope (LSST)

Big Data and Processing

Kian-Tat Lim
LSST Data Management System Architect
August 21, 2012

U.S. DEPARTMENT OF ENERGY
Office of Science

SLAC NATIONAL ACCELERATOR LABORATORY

Outline

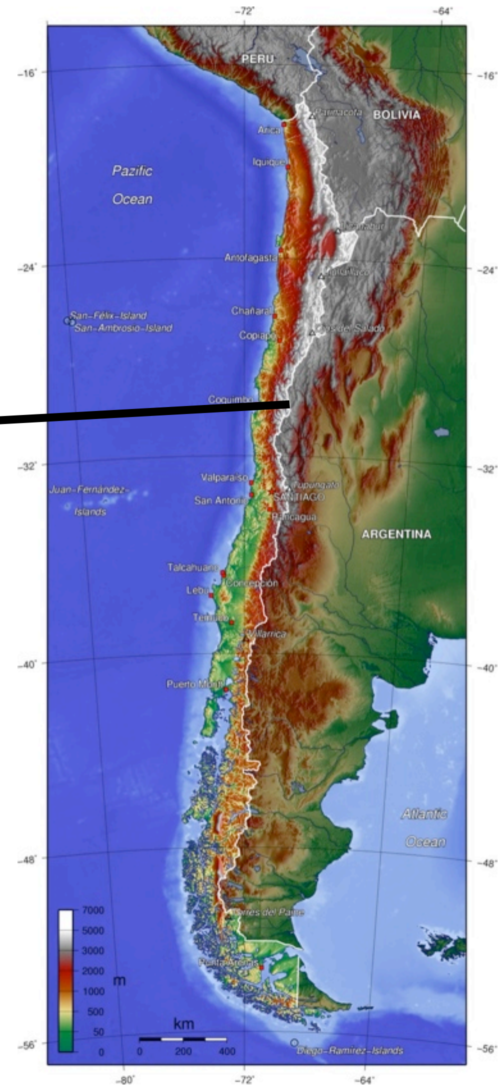
LSST: What

LSST: Why

LSST: Data

Data-Intensive Science

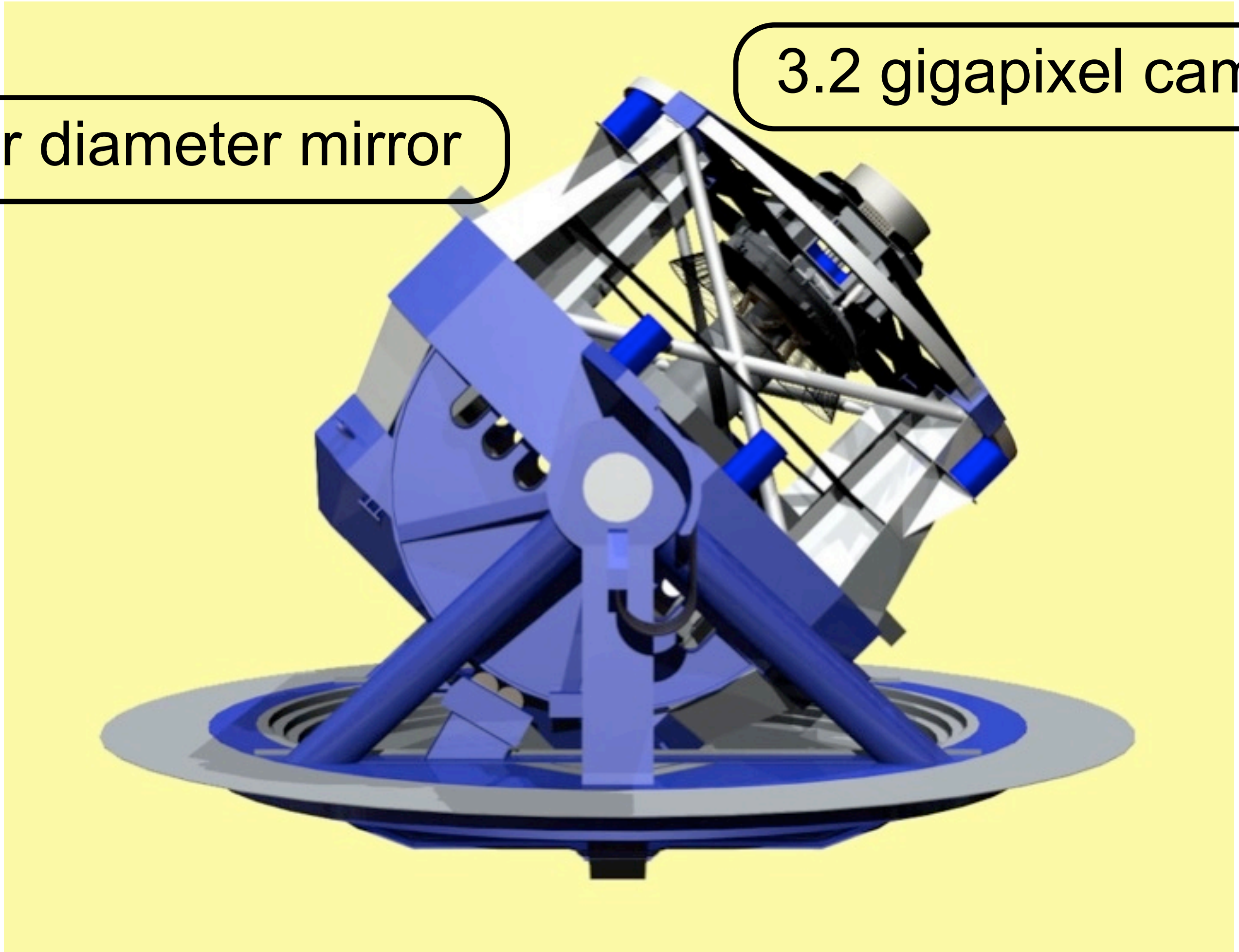
Telescope



Proposed telescope to be built in Chile

More than proposed: 26 tons of glass made into mirror, soon to be polished

Large



Synoptic Survey

Wide

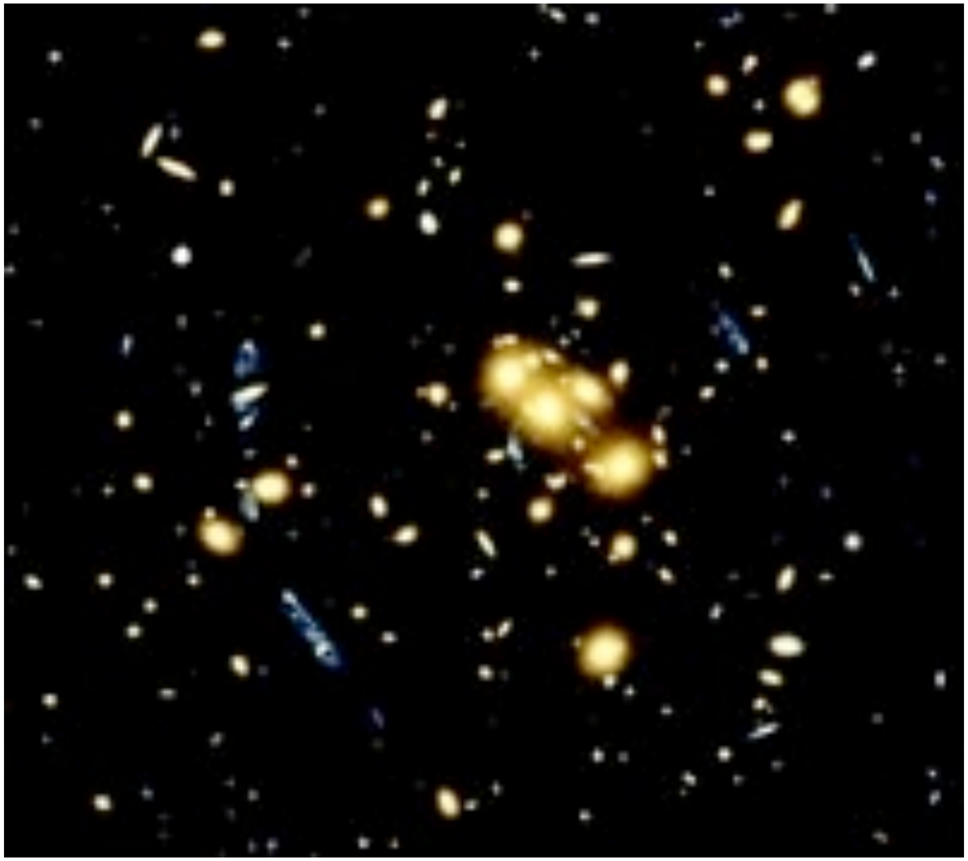
Deep

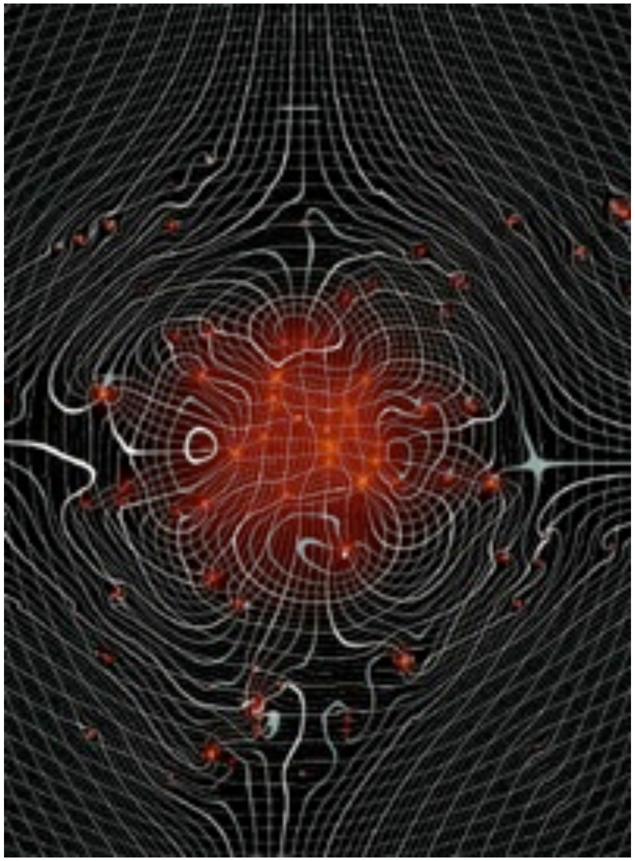
Fast

Wide: surveys entire sky, not just small area
Deep: gathers faint and distant objects
Fast: image every 15 seconds; entire sky twice a week
Brings in time domain, unlike almost all other surveys

Dark Matter and Energy

SLAC

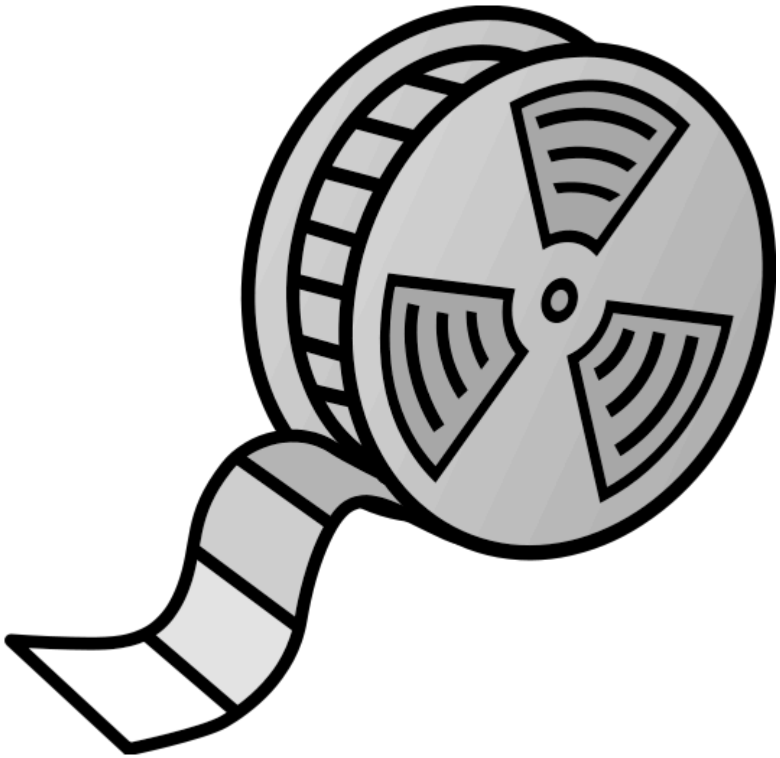




Credit: J. A. Tyson, W. Colley, E. L. Turner, and NASA

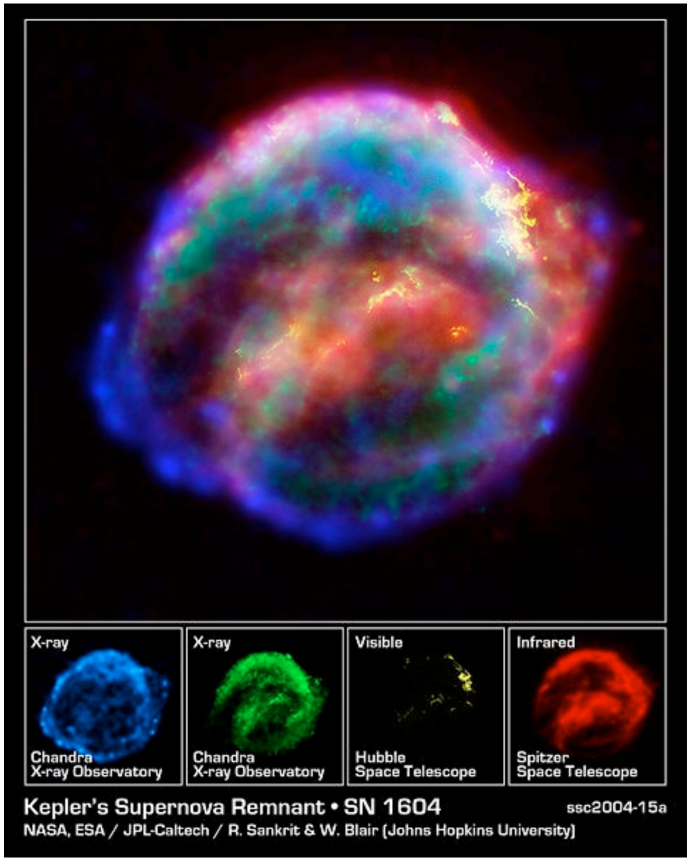
Looking at distortions of millions or billions of galaxies
Gives information to map dark matter and dark energy

Variable Objects



Get a thousand-frame movie of every visible object showing variations over time

Transient Objects



Supernovae appear unexpectedly
We will detect them as new objects that appear (or disappear)

Moving Objects

SLAC

Credit: D. Roddy, Lunar and Planetary Institute

Very interested in objects much closer — in the solar system — that move and might collide with us

Data Challenge

Images

Database

Two main types of data

Images - pixel data - raw and processed

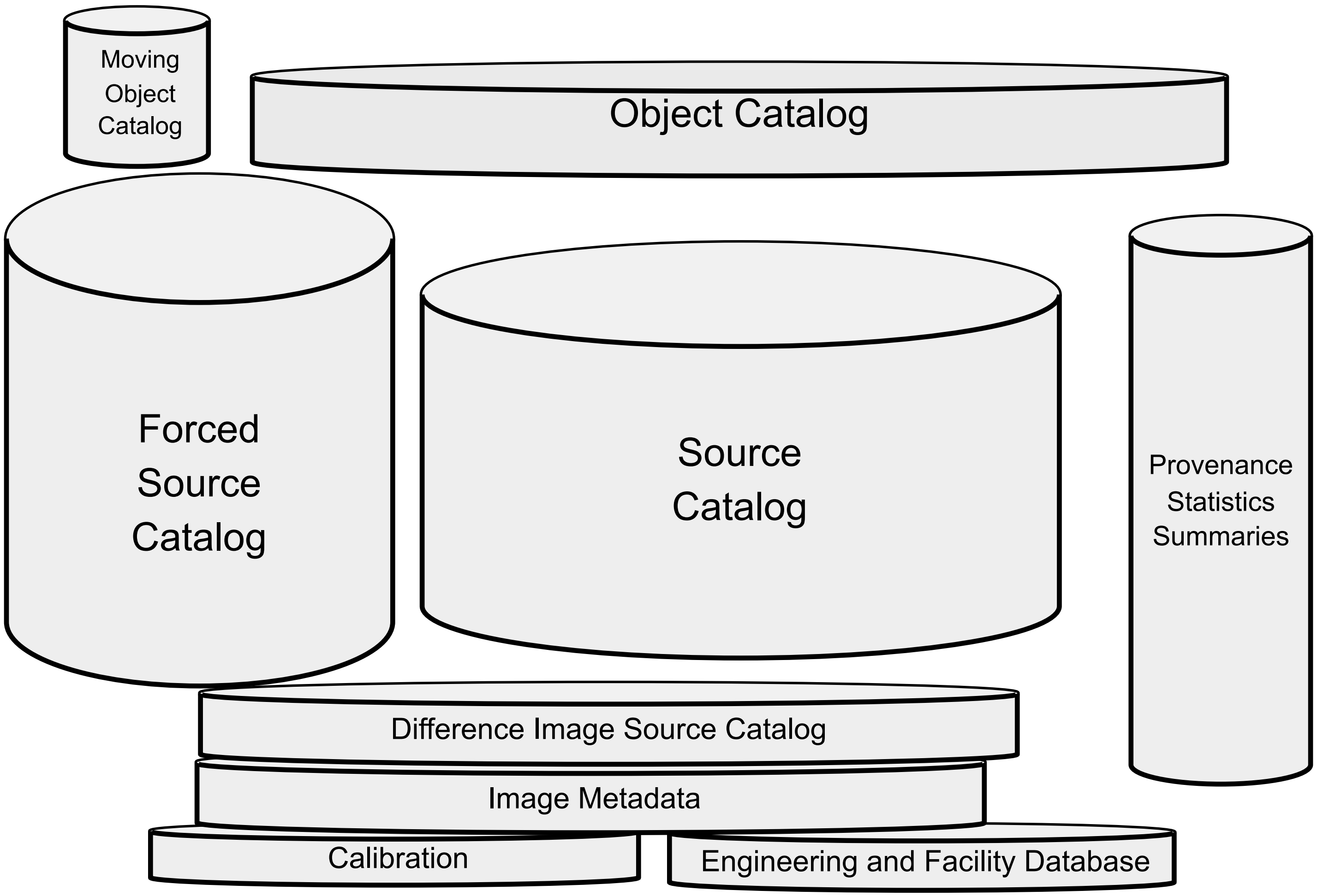
Database - everything else, including:

metadata and provenance

astronomical objects

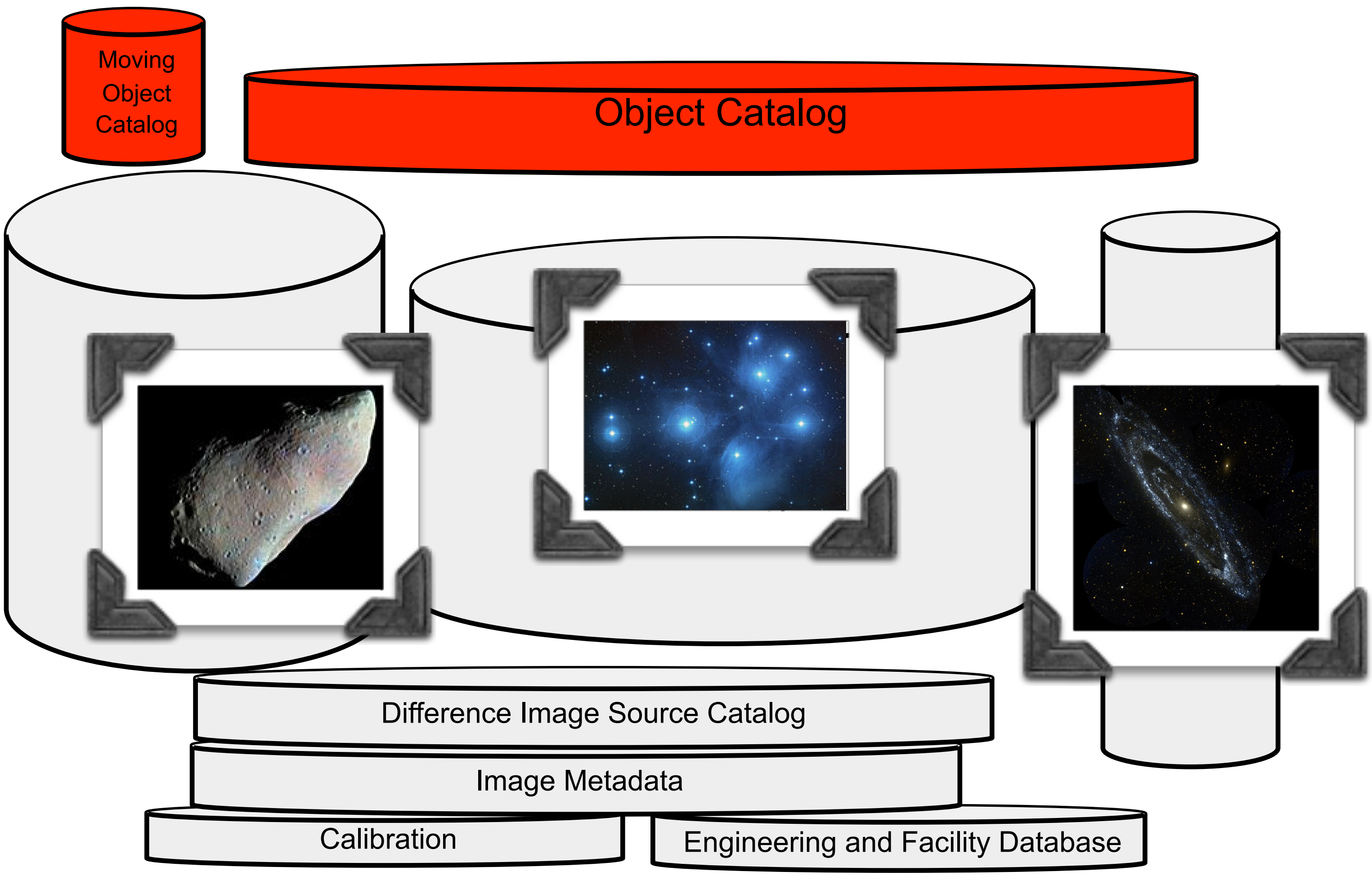
sources (or detections of objects)

Database: Components



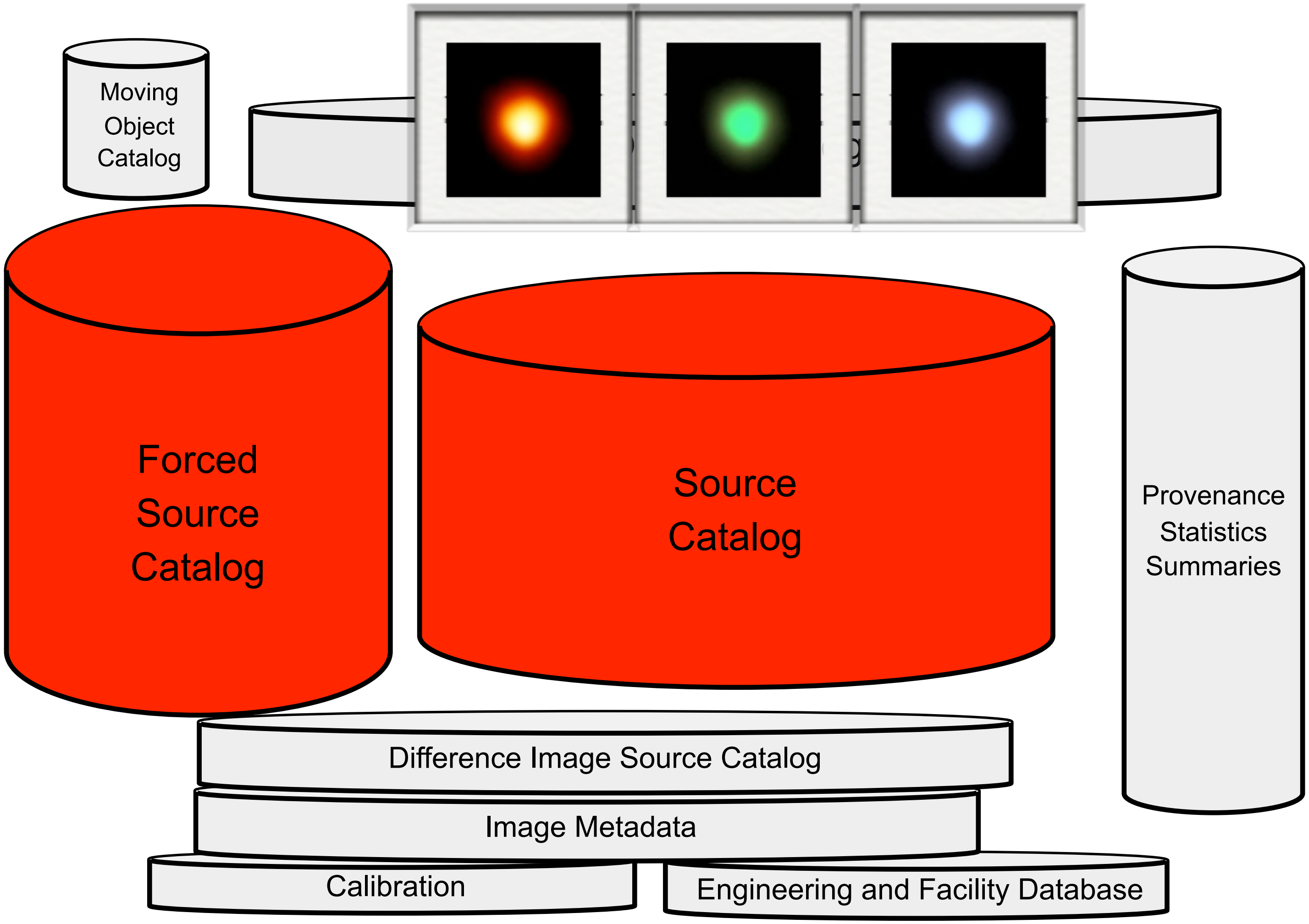
Database is composed of many catalogs, each of which has several tables

Astronomical Objects



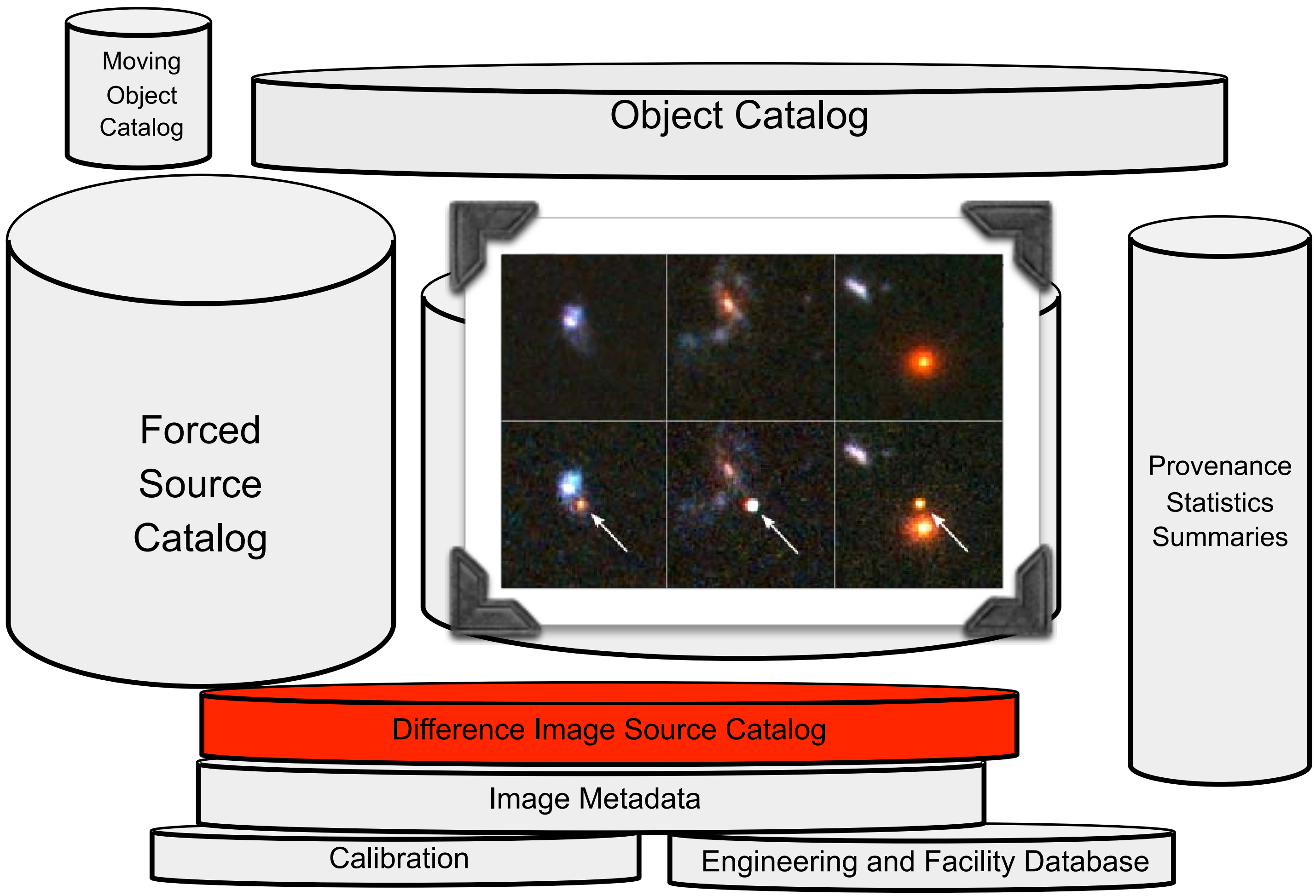
Asteroids, stars, and galaxies
Each has many attributes, some fixed, many averaged over time

Sources



Sources are observations or measurements of objects
At different times, through different filters
Forced Sources are low-signal-to-noise
Looks rather like commercial data warehouse

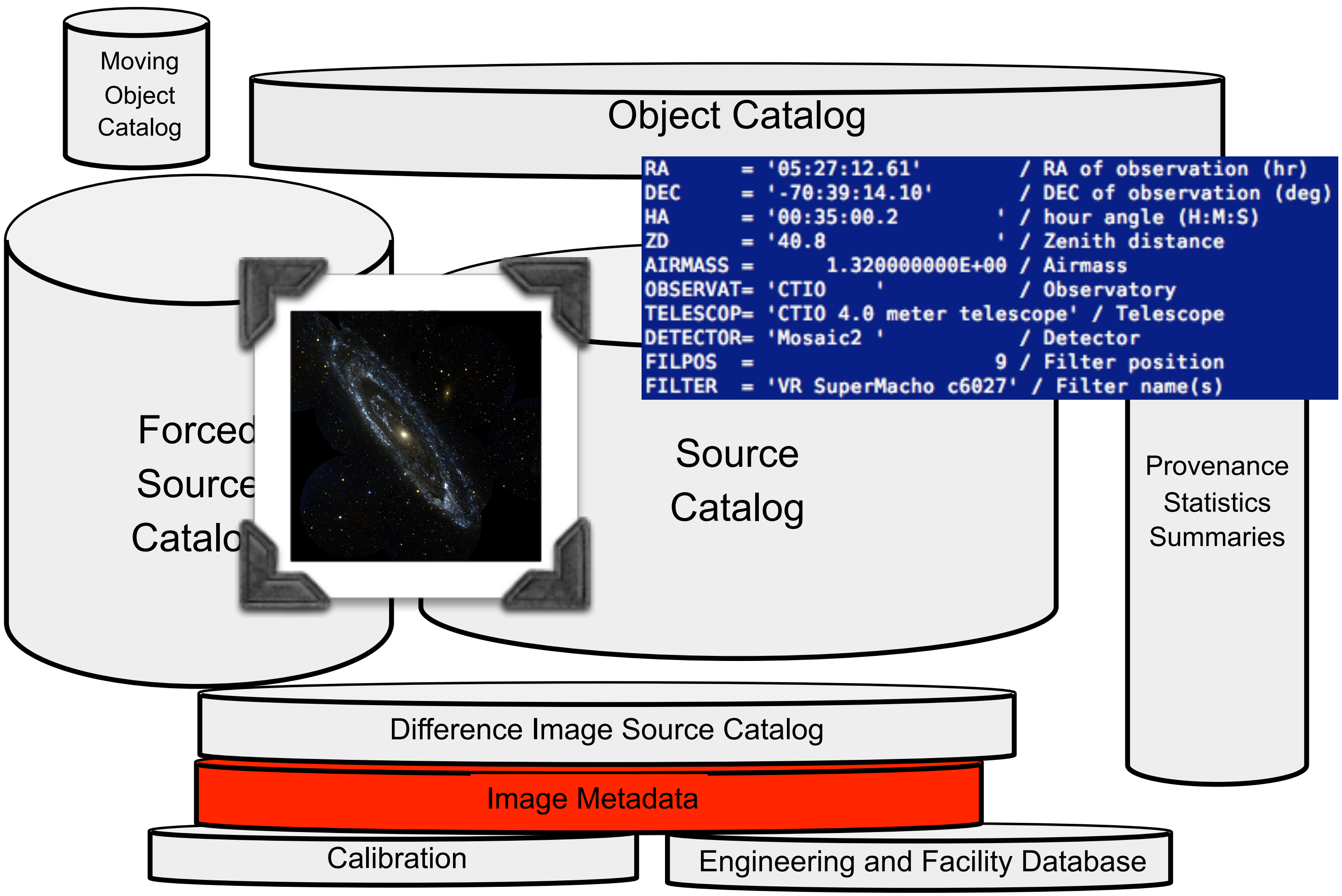
Changes



Difference images are how we detect changes
Subtract an average known template from the current image
Anything that has changed stands out
Sources from these difference images are kept separately

Image Metadata

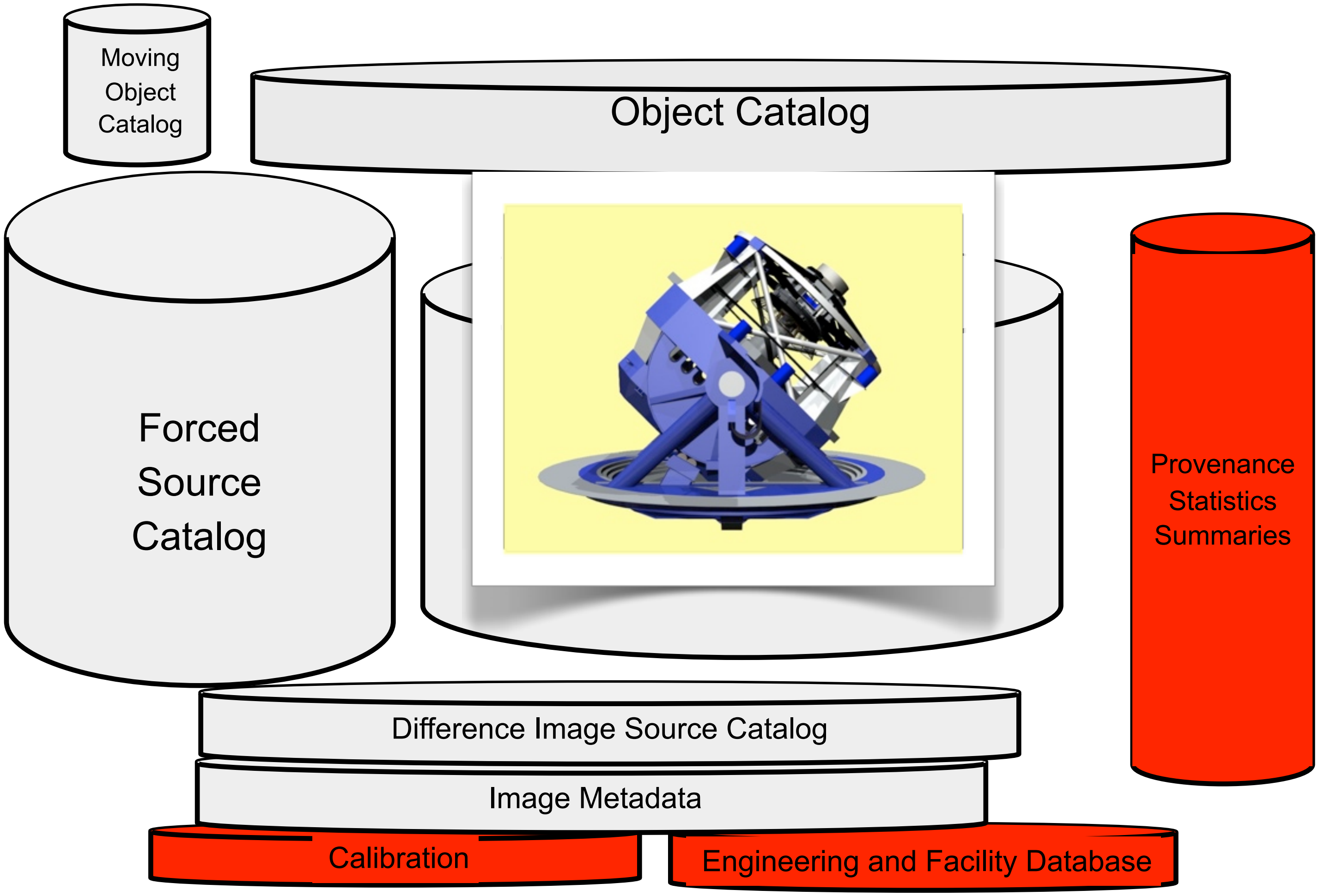
SLAC



Each image has lots of metadata

- Observation conditions: pointing, time, filter, atmospheric conditions
- Image analysis for artifacts: cosmic rays, satellites and airplanes
- Quality control flags
- Location on disk or tape

Calibration and Facility



Calibration information and information about the telescope itself must be retained

Also need provenance of all elements in database, various aggregate statistics

Very Big Numbers

38 *billion* objects

5 *trillion* sources

32 *trillion* forced sources

8 *quadrillion* pixels

May have even more sources if signal-to-noise threshold is reduced
May have even more objects and sources if we point closer to the galactic plane

Data Locations

SLAC

Mountain

Base Site

Archive Site

Data Access Center

EPO Center

Where is the data?
Five locations
Mountain is where images are generated
Base Site is where images are sent; real-time processing for alert generation may occur here
Archive Site is permanent repository, science processing
Data Access Center is where end users (US/Chile/Intls) query the DB and access images
Education and Public Outreach is specialized for the general public (Google Sky/WWT)

Data Size

SLAC

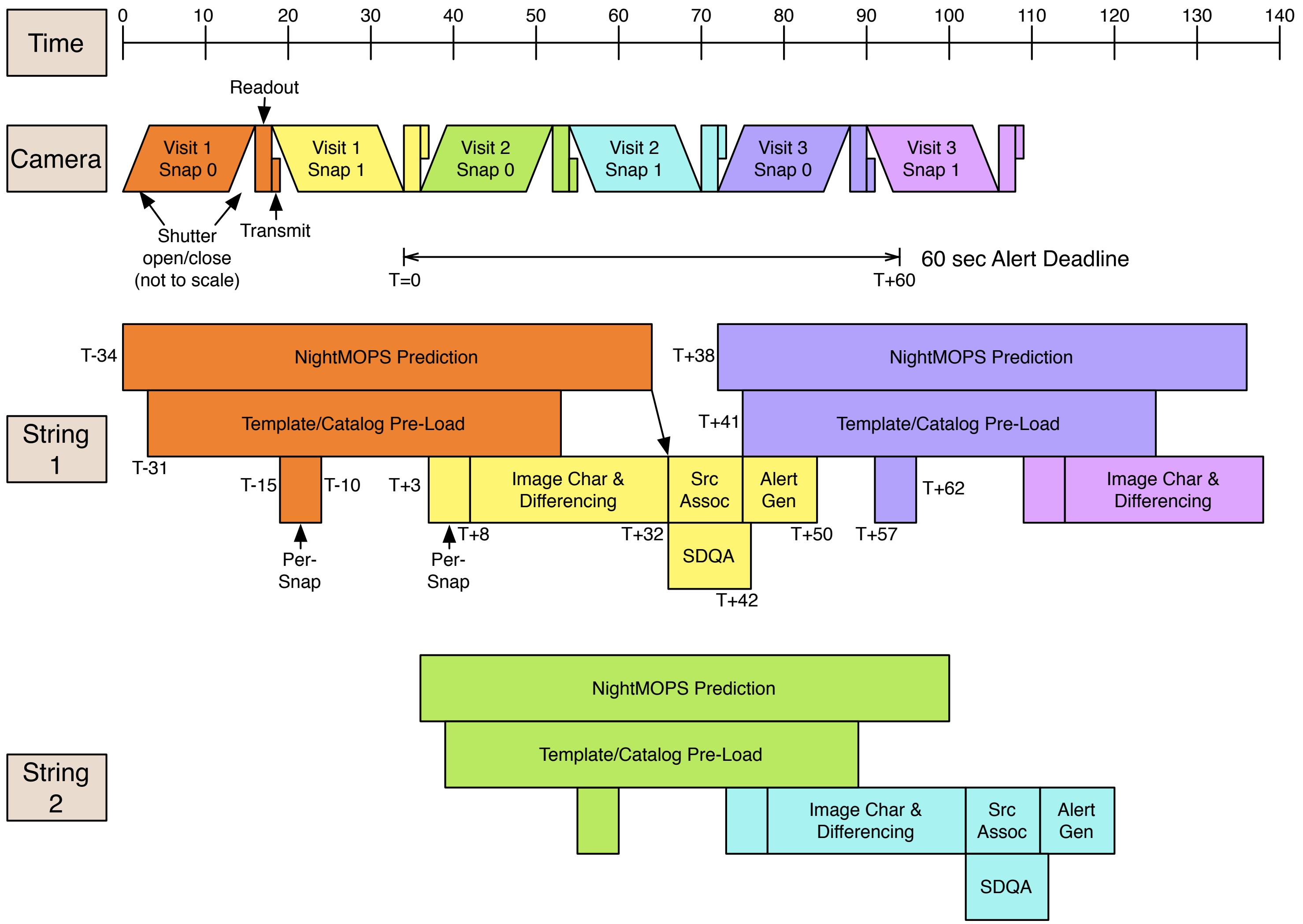
		Archive Site	Base Site
Compute	Teraflops (sustained)	175 → 1000 (required) 270 → 1600 (on-the-floor)	30 → 45 (required) 50 → 80 (on-the-floor)
	Nodes	1700 → 1400 (on-the-floor)	300 → 60 (on-the-floor)
	Cores	70K → 280K (on-the-floor)	12K → 13K (on-the-floor)
	Memory Bandwidth	25 → 130 TB/s (required)	4 → 6 TB/s (required)
	File I/O Bandwidth	120 → 425 GB/s (required)	25 → 30 GB/s (required)
Storage	Capacity	13 → 25 PB (required) 19 → 100 PB (on-the-floor)	8 → 8 PB (required) 12 → 23 PB (on-the-floor)
	Drives	1500 → 1100 (on-the-floor)	950 → 275 (on-the-floor)
	Disk Bandwidth	120 → 425 GB/s (required)	27 → 31 GB/s (required)
Mass Storage	Capacity	7 → 75 PB (required) 8 → 83 PB (on-the-floor)	7 → 75 PB (required) 8 → 83 PB (on-the-floor)
	Tapes	1000 → 3800 (near line) 1000 → 3800 (offsite)	1000 → 3800 (near line)
	Tape Bandwidth	6 → 24 GB/s (required)	6 → 24 GB/s (required)
Database	Teraflops (sustained)	16 → 193 PB (required) 16 → 199 PB (on-the-floor)	12 → 126 PB (required) 12 → 133 PB (on-the-floor)
	Storage Capacity	4 → 50 PB (required) 10 → 128 PB (on-the-floor)	3 → 40 PB (required) 7 → 95 PB (on-the-floor)
	Disk Drives	1400 → 2600 (on-the-floor)	1000 → 2000 (on-the-floor)
	Disk Bandwidth (sequential)	125 → 625 GB/s (required)	95 → 425 GB/s (required)
	Database Nodes	100 → 190 (on-the-floor)	80 → 130 (on-the-floor)
Facilities	Floorspace	950 → 850 sq ft (1100 high water mark)	460 → 350 sq ft (470 high water mark)
	Power	720 → 640 kW (880 high water mark)	180 → 110 kW (210 high water mark)
	Cooling	2.4 → 2.1 mmbtu (3.0 high water mark)	0.6 → 0.4 mmbtu (0.7 high water mark)

First number is survey year 1; second number is survey year 10

Requirements go up because data size goes up and processing rate goes up

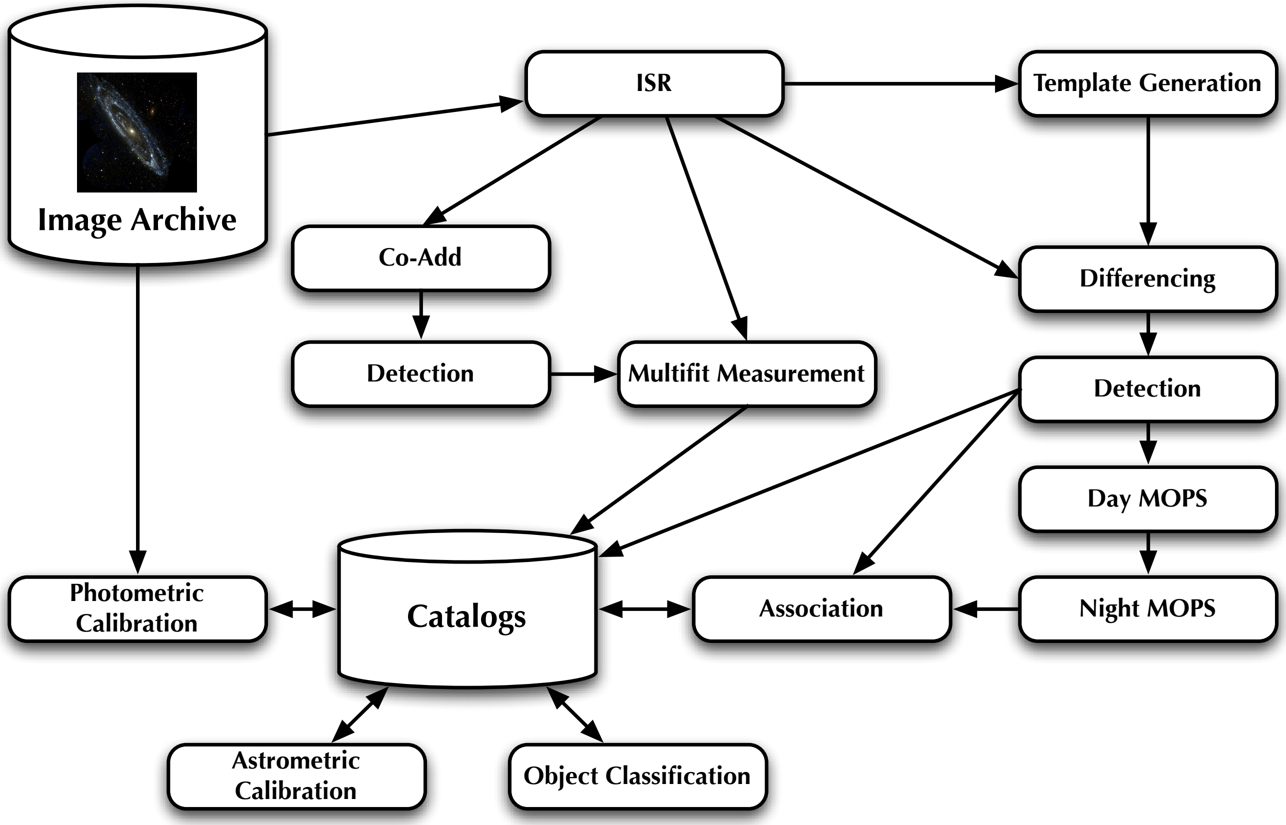
In last year, need to process 10 years of data in just one year

Alert Production



Alerts issued within 60 seconds of taking picture
Near-realtime processing involving complex image processing

Data Release Production



Reprocess entire data set every year to create consistent, best-available catalogs

Usual Suspects

Scalable

Fast

Fault-tolerant

Cost-effective

Open Source

Reuse existing components where possible:

xrootd for distributed processing and communication

Fermi/GLAST workflow software contending for workflow management

Data-Intensive Science

Computing, but not super-computing

Credit: NCSA/University of Illinois

Mining large databases

Credit: National Institute for Occupational Safety and Health

Types of Queries

- All about an object
- All objects meeting criteria
- All objects near objects meeting criteria
- All objects with interesting time series
- All pairs of objects with similar time series

DISC

Bandwidth

Bandwidth

Bandwidth

For input
For (re-)processing
For queries

Summary

LSST is an example of a Big Data experiment requiring data-intensive scientific computing