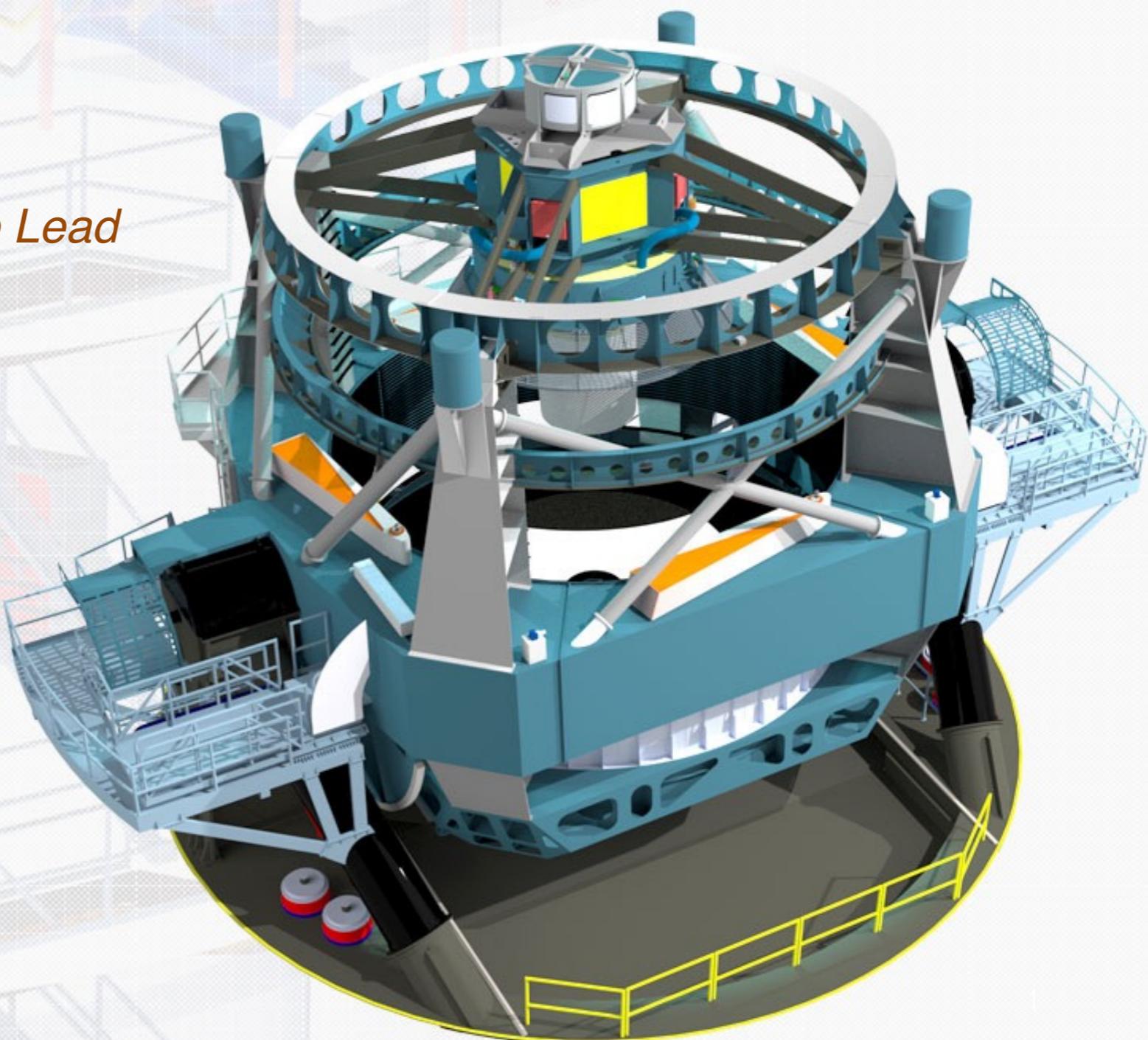


Alert Streams in the LSST Era: Challenges and Opportunities

Eric Bellm

University of Washington

LSST DM Alert Production Science Lead



The Large Synoptic Survey Telescope will produce an alert stream of greater scale and generality than any survey to date.



An automated 8.4 meter telescope that for 10 years will image half the sky every ~3 days, generate ~50 PB of (raw) imaging data, issue real-time alerts to any changes in the sky (~10 million/night), measure properties of ~40 billion objects in the sky (~1000 times each), and make the results available in a web-accessible database.

First Light: 2019
Operations: 2022



The summit, May 2018.



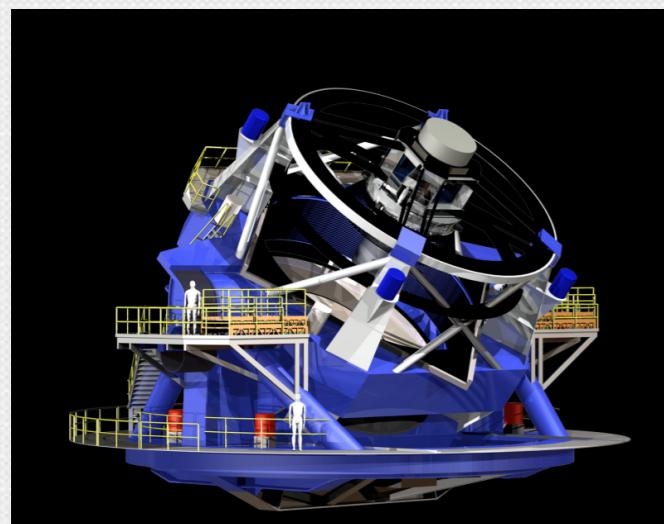
LSST is a database of the optical sky.



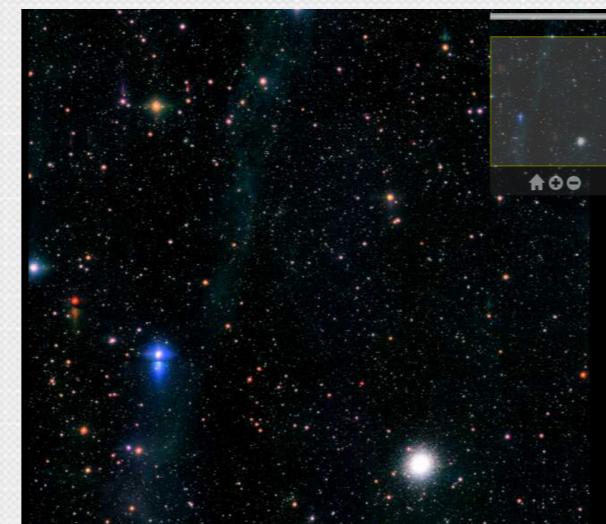
LSST data, including images and catalogs, will be available with no proprietary period to the astronomical community of the United States, Chile, and International Contributors. LSST's alerts are immediately world-public.

LSST is a public facility: all science will be done by the community (not the Project!), using LSST's data products.

The ultimate deliverable of LSST is not the telescope, nor the instruments; it is the fully reduced data. LSST is a facility that delivers data products and data access and analysis services.



Telescope



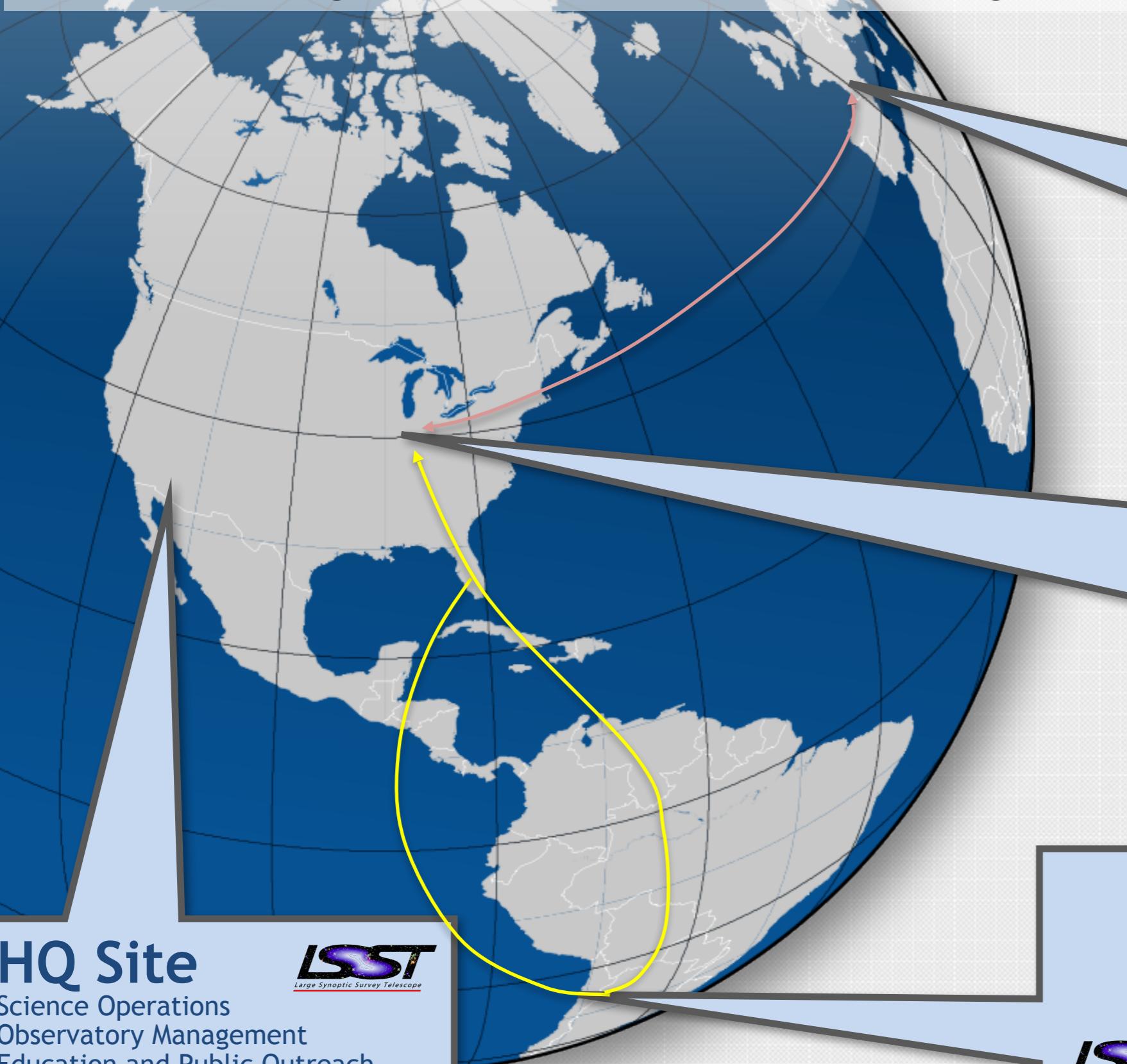
Images

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 <code>radec</code> .
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 <i>log</i> 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 <code>radecTai</code> , in each band.



Catalogs

We are building a multi-continent Data Management System.



Satellite Processing Center

(CC-IN2P3, Lyon, France)

Data Release Production (50%)



Archive Site



Archive Center

Alert Production

Data Release Production (50%)

EPO Infrastructure

Long-term Storage (copy 2)

Data Access Center

Data Access and User Services

Chilean Sites

Telescope and Camera

Data Acquisition

Crosstalk Correction

Long-term storage (copy 1)

Chilean DAC Entry-point

HQ Site

Science Operations
Observatory Management
Education and Public Outreach



LSST is planning a ten-year survey.

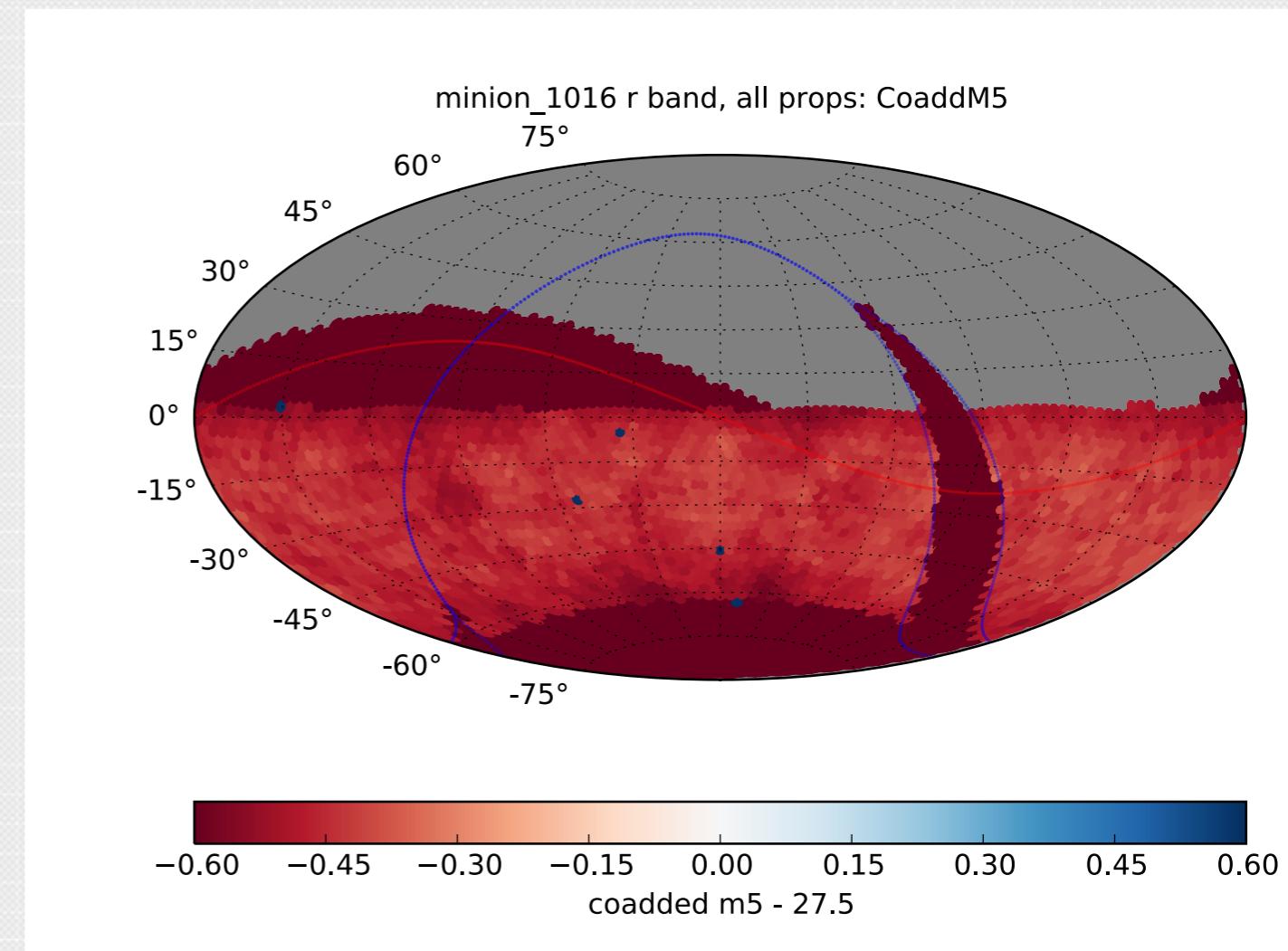


Survey in ugrizy bands, with
~825 visits per pointing

Wide-Fast-Deep:
2x/night every three nights
over 18,000 square degrees

Special programs:

- Deep Drilling
- Galactic Plane
- North Ecliptic Spur
- South Celestial Pole



Ongoing cadence development & evaluation:
[https://github.com/
LSSTScienceCollaborations/
ObservingStrategy](https://github.com/LSSTScienceCollaborations/ObservingStrategy)

LSST has three data processing modes.



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.

Prompt

For more details, see the “Data Products Definition Document”, <http://ls.st/dpdd>

LSST has three data processing modes.



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.

A catalog of ~37 billion objects (20B galaxies, 17B stars), ~7 trillion observations (“sources”), and ~30 trillion measurements (“forced sources”) accessible through online databases.

Reduced single-epoch, deep co-added images.

Prompt

Data Release

For more details, see the “Data Products Definition Document”, <http://ls.st/dpdd>

LSST has three data processing modes.



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.

A catalog of ~37 billion objects (20B galaxies, 17B stars), ~7 trillion observations (“sources”), and ~30 trillion measurements (“forced sources”) accessible through online databases.

Reduced single-epoch, deep co-added images.

Services and computing resources at the Data Access Centers enabling limited analysis, production, and federation of added value products.

Web APIs enabling the use of remote analysis tools.

Public LSST pipeline code for deeper insight into LSST data products.

For more details, see the “Data Products Definition Document”, <http://ls.st/dpdd>

Prompt: Time-Domain Alerts



We expect a high rate of alerts, **approaching 10 million per night**. We'll also provide an *alert filtering service*, to select subsets of alerts, as well as serve the full stream to external *event brokers*.

Each alert will include the following:

- Alert and database ID: IDs uniquely identifying this alert.
- The photometric, astrometric, and shape characterization of the detected source
- 30x30 pixel (on average) cut-out of the difference image (FITS)
- 30x30 pixel (on average) cut-out of the template image (FITS)
- The time series (up to a year) of all previous detections of this source
- Various summary statistics (“features”) computed of the time series

The goal is to quickly transmit nearly everything LSST knows about any given event, enabling downstream classification and decision making.

Prompt processing also includes nightly identification of Solar System Objects.

A series of software pipelines produces the LSST alert stream.

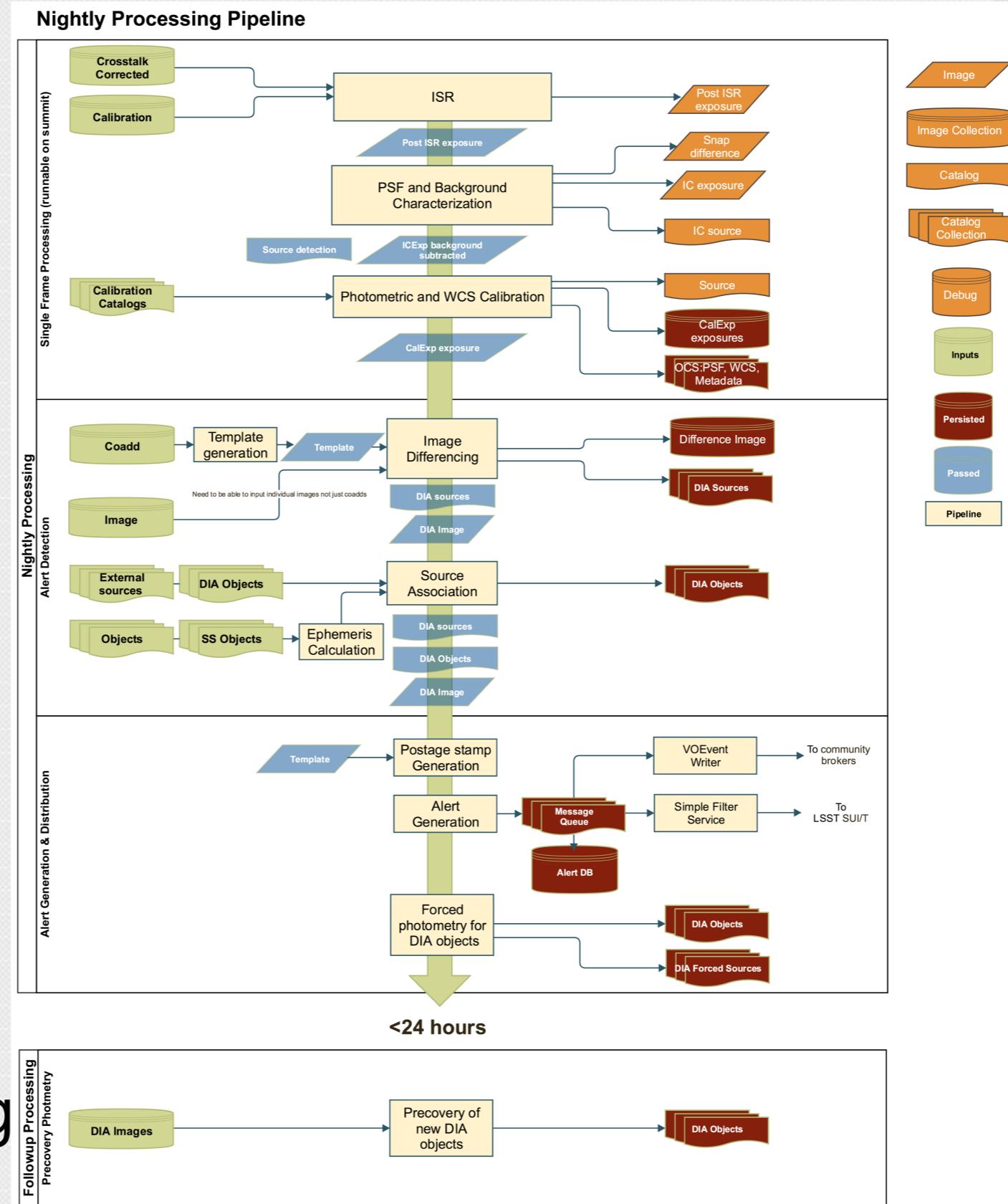


Single Frame Processing

Alert Generation

Alert Distribution

Forced Processing

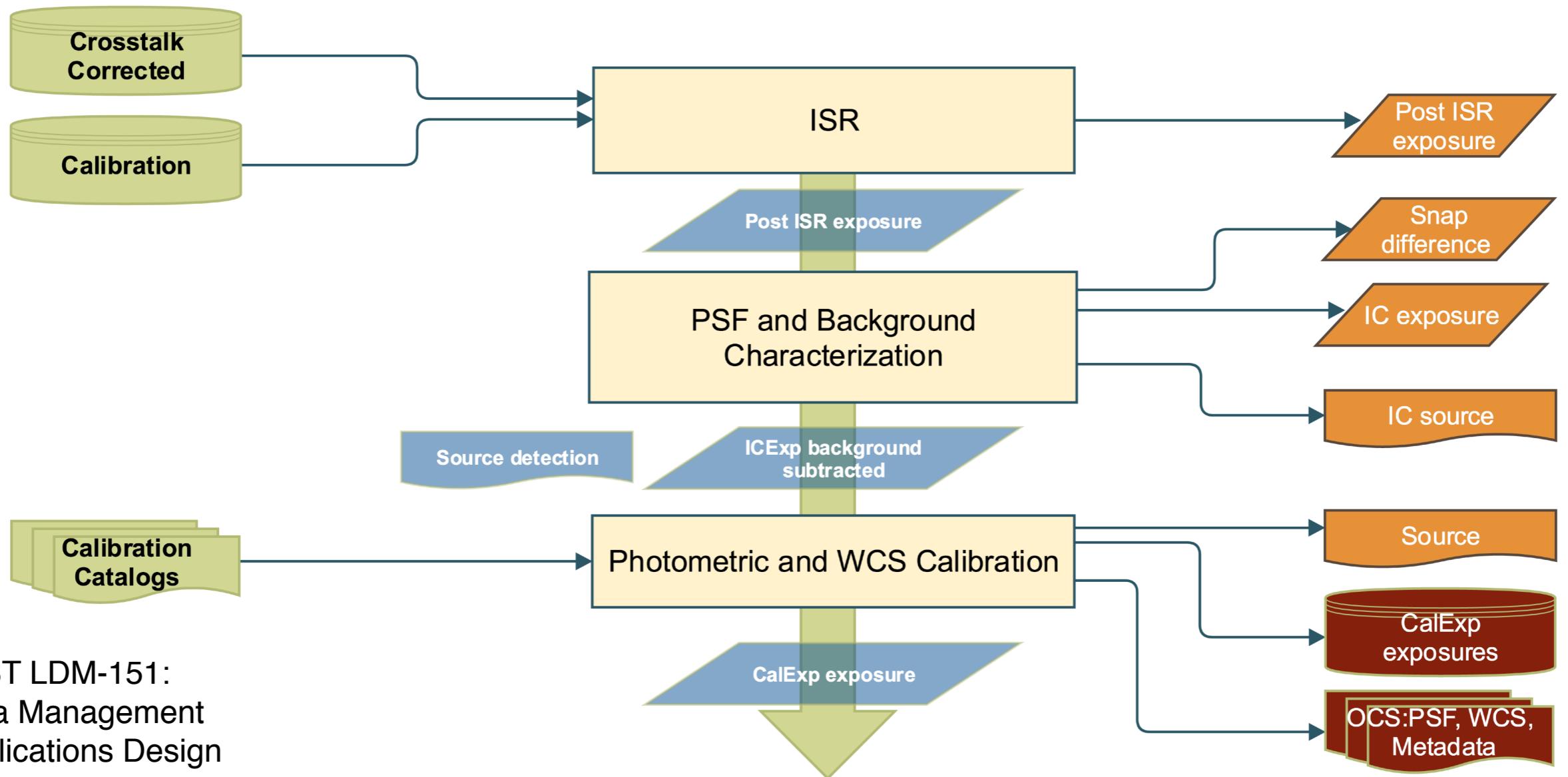


LSST LDM-151:
Data Management
Applications Design
ls.st/LDM-151

Single-Frame Processing provides calibrated exposures.



Single Frame Processing



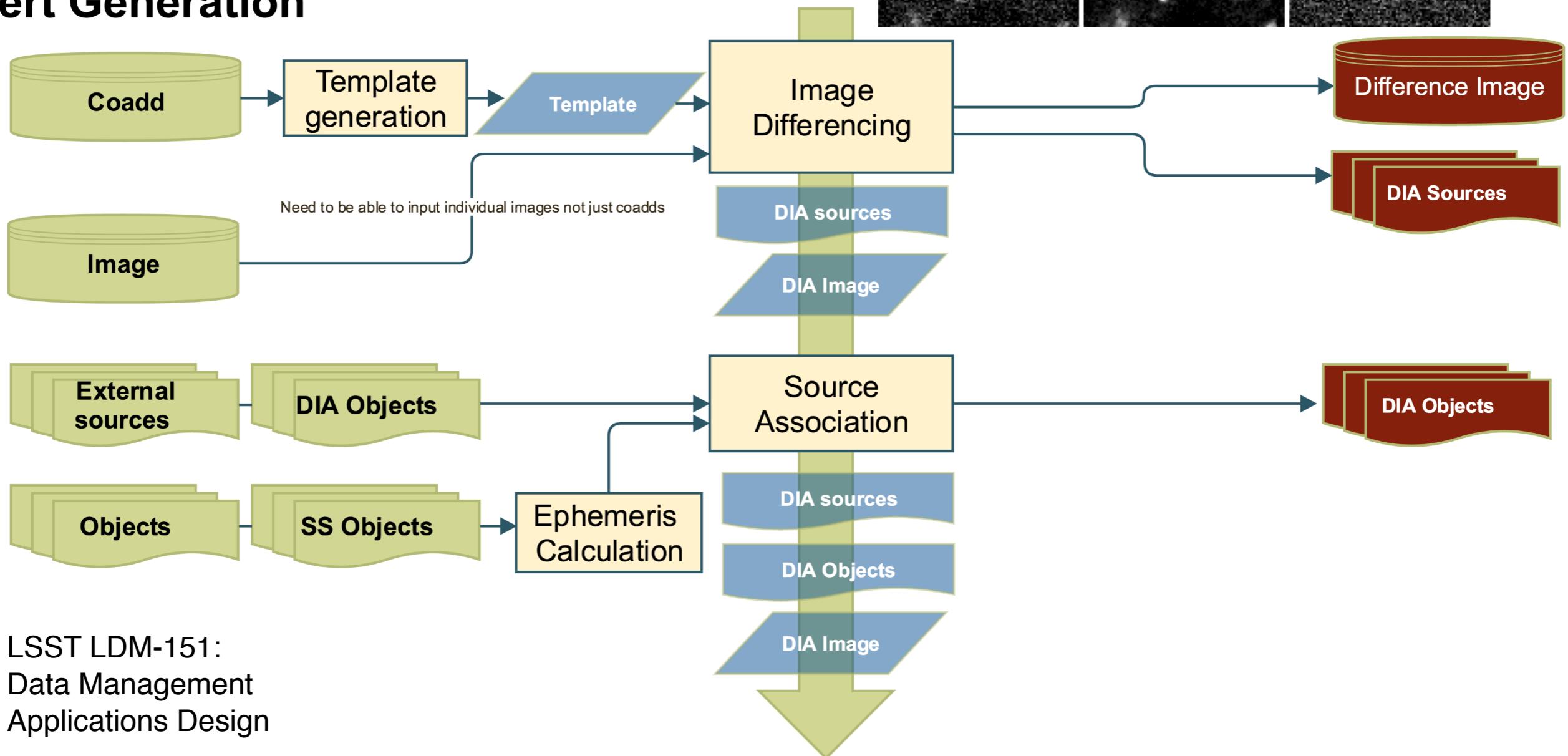
LSST LDM-151:
Data Management
Applications Design

ls.st/LDM-151

Alert Generation detects and associates transients.



Alert Generation



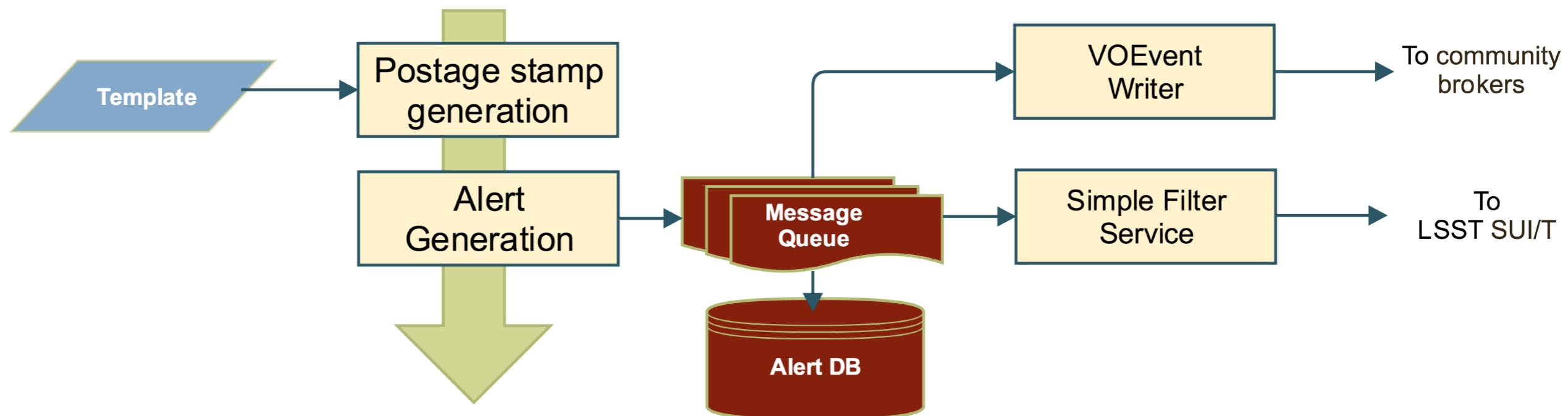
LSST LDM-151:
Data Management
Applications Design

ls.st/LDM-151

Alert Distribution packages and sends alerts.



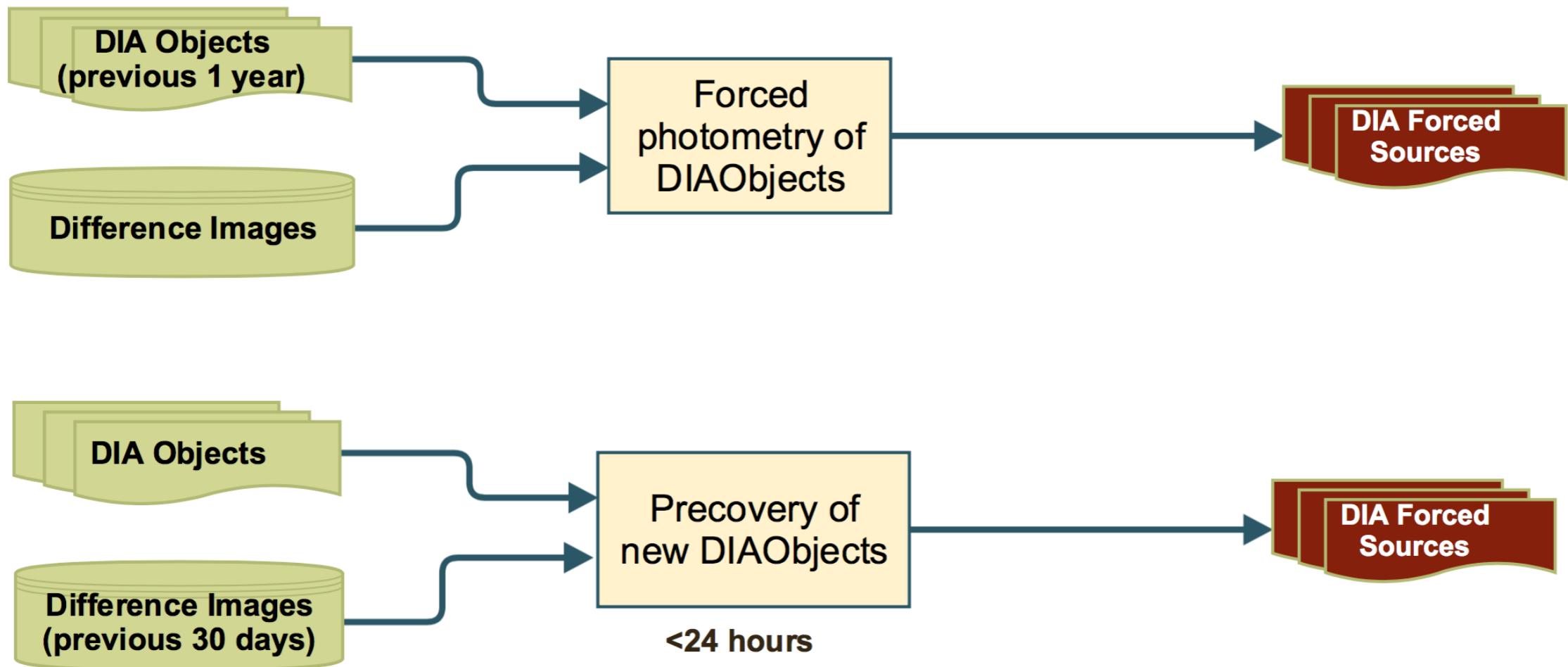
Alert Distribution



LSST LDM-151:
Data Management
Applications Design

ls.st/LDM-151

Recovery and Forced Photometry



LSST LDM-151:
Data Management
Applications Design

ls.st/LDM-151

**LSST's alert stream differs in scale and motivation
from current astronomical databases.**



Primary interface is an *alert stream*, not a *batch query*

Real-time, low-latency, naturally distributed & decentralized

All* subtraction candidates are streamed at low latency

“turn the database inside out”

(“alert” is somewhat of a misnomer...)

Events sent in (world-public!) rich alert packets

enable standalone classification

Users find events of interest through classification & filtering systems

full stream to community brokers: ANTARES, ALeRCE, LASAIR, etc.

simple LSST “mini-broker” filtering service

key decision: is this an object I want to follow up?

LSST uses rich alert packets to minimize followup queries.



Each alert

will at least include the following:

- *alertID*: An ID uniquely identifying this alert. It can also be used to execute a query against the Level 1 database as it existed when this alert was issued
- *Level 1 database ID*
- Science Data:
 - The DIASource record that triggered the alert
 - The entire DIAObject (or SSObject) record
 - All previous DIASource records -> last 12 months
 - A matching DIAObject from the latest Data Release, if it exists, and its DIASource records
- Cut-out of the difference image centered on the DIASource (10 bytes/pixel, FITS MEF)
- Cut-out of the template image centered on the DIASource (10 bytes/pixel, FITS MEF)

LSST LSE-163:
Data Products
Definition Document
ls.st/DPDD

DIA~~Source~~ and DIA~~Object~~ records contain a wide range of measurements.



DIASources:

- Position
- aperture/PSF/dipole/trailed fluxes
- moments
- likelihoods, extendedness, spuriousness

DIAObjects:

- linkages to DIASources [→ light curve], Data Release Objects
- time series statistics

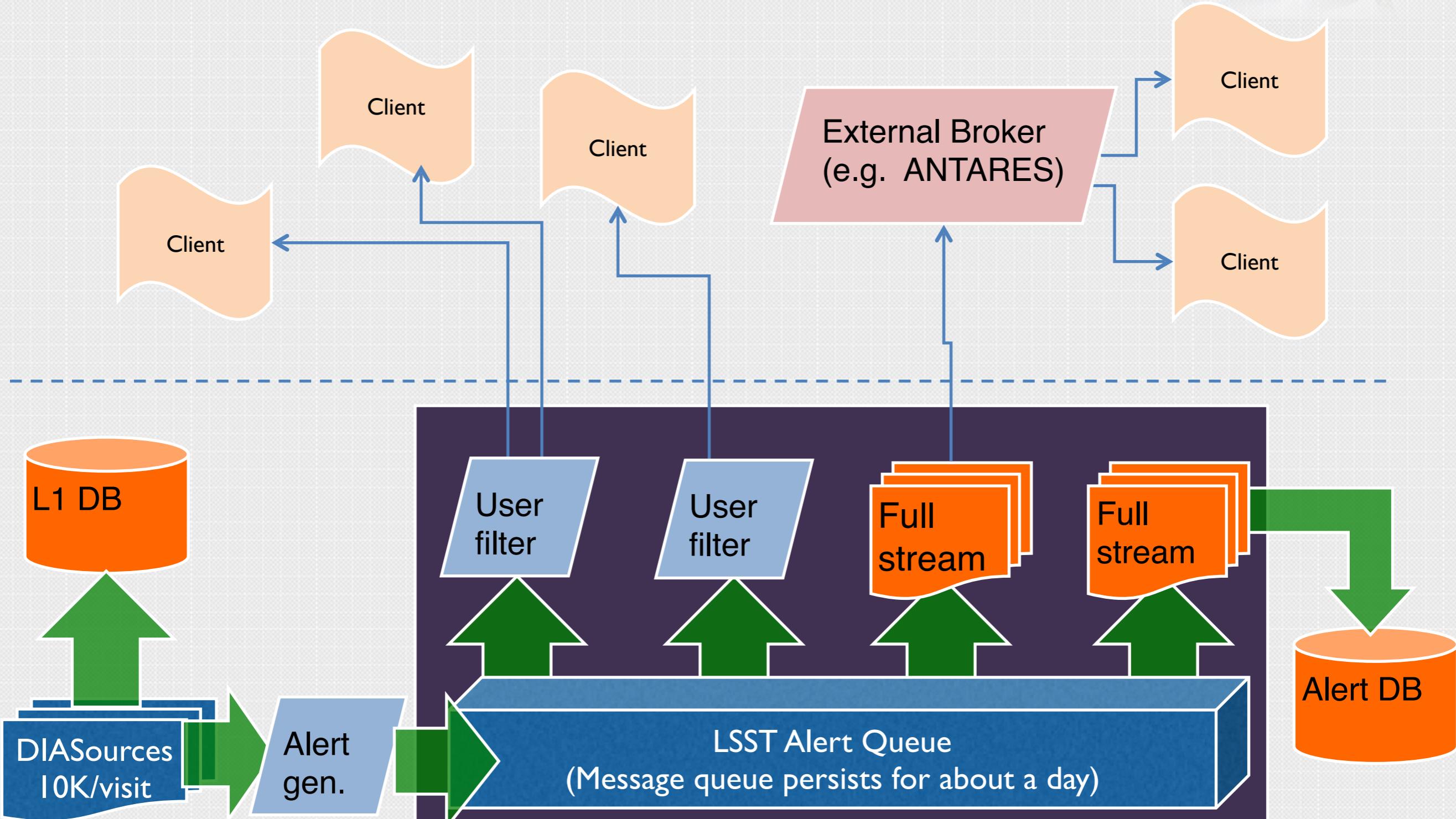
SSObjects:

- linkages to DIASources
- variety of solar system parameters

LSST LSE-163:
Data Products
Definition Document

ls.st/dpdd

LSST alert distribution requires a new community ecosystem.



At ~20 full sized events per visit per user (or summarizing the lightcurve for all events in ~40 numbers) we can serve ~500 simultaneous users for the cost of a single full data stream

Community brokers will enhance the LSST alert stream.



- **cross-match with other catalogs and alert streams**
- **classify events** (the LSST Project can only characterize)
- redistribute alert packets
- filter alerts
- provide user interfaces
- enable community coordination
- trigger followup resources and manage that data
- provide storage and archiving
- provide annotation & citation
- manage “discovery”
- ...probably more?

A finite number of brokers will be selected by a proposal process to receive the full stream.

LSST will provide a “mini-broker” service



User-defined filters that act *only* on alert packet contents

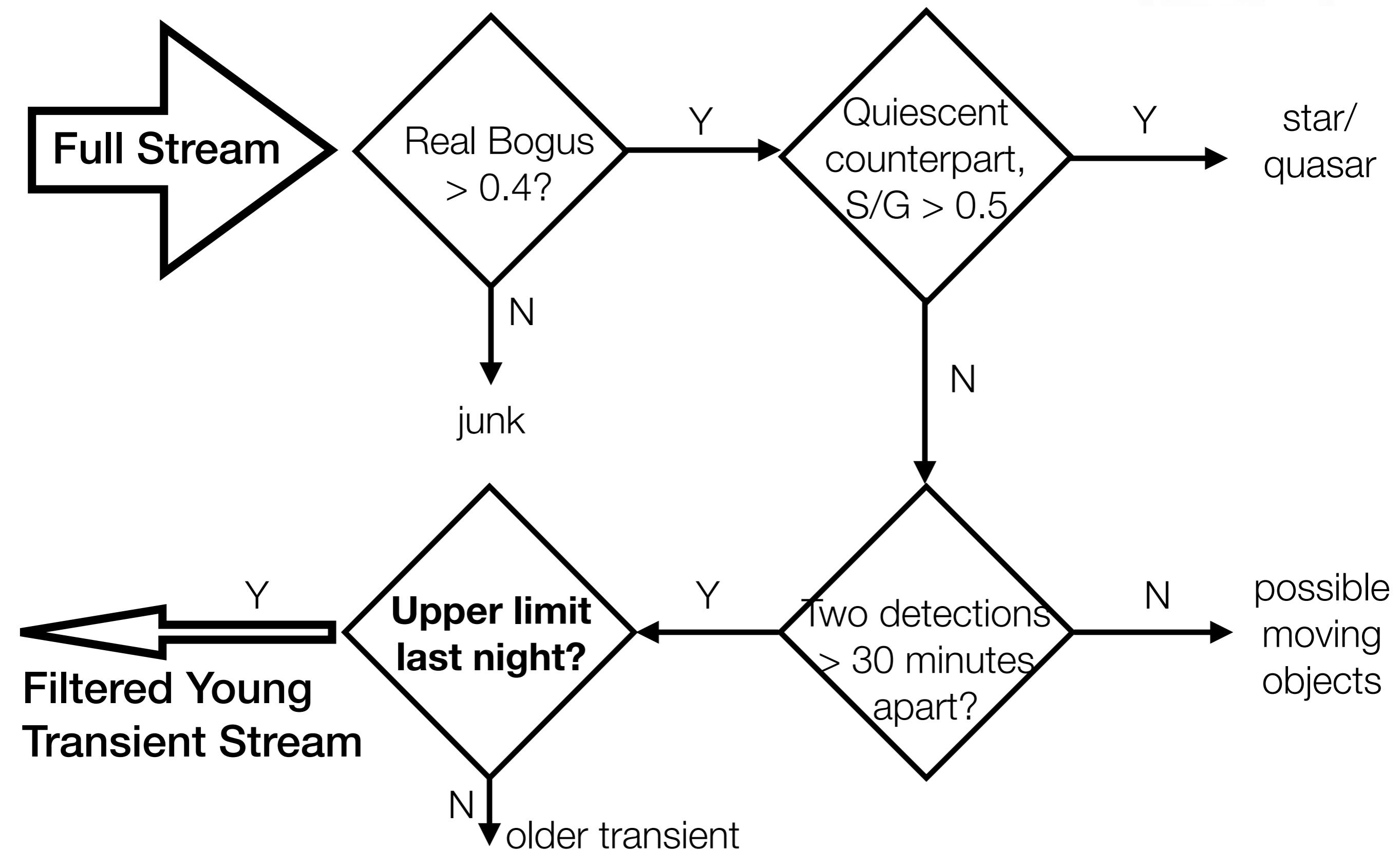
Access to the filtered stream through LSST’s Science Platform

Cap of ~20 alerts per user per visit; some limits on computing capacity

LSST LSE-163:
Data Products
Definition Document

ls.st/DPDD

Simple single-alert filters can enable a lot of science.



The LSST alert stream presents both opportunities and challenges.



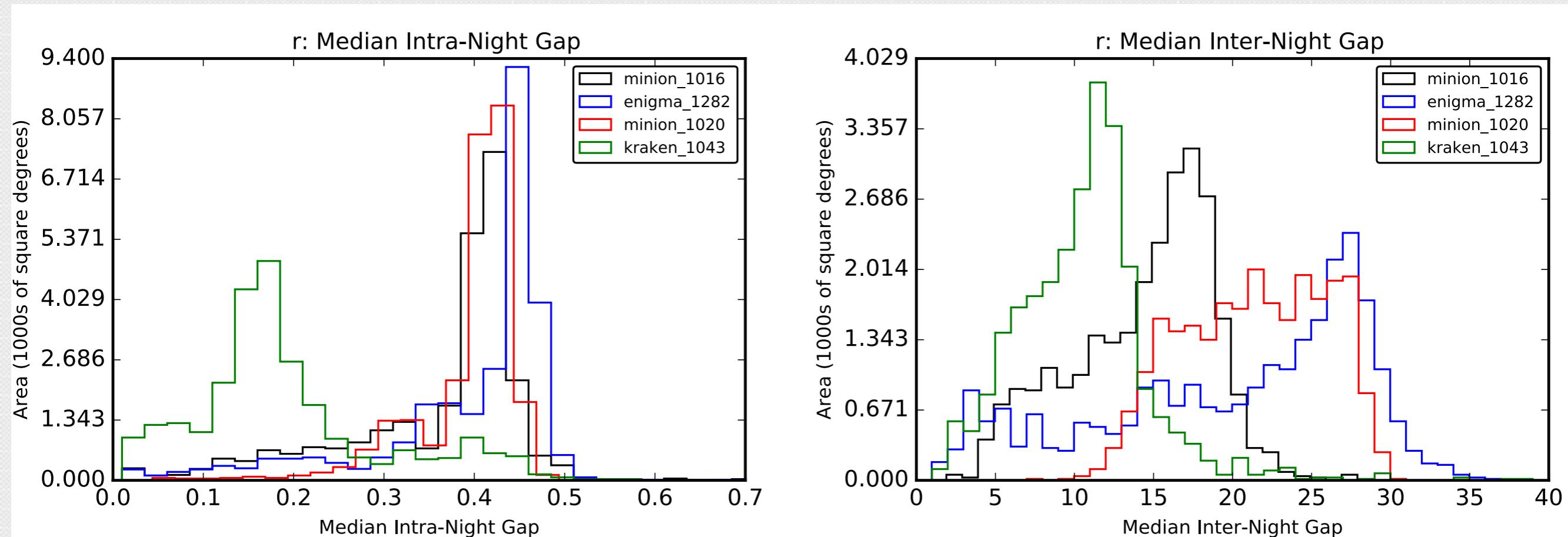
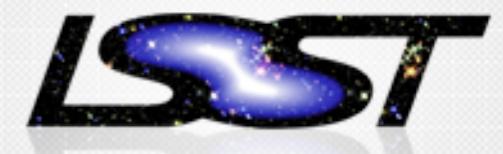
Opportunities

- a powerful new facility; huge discovery space
- rich data products to enable general-purpose inference: “batteries included”
- naturally distributed, BYOC

Challenges

- large data volumes and event rates
- **sparse & irregular sampling due to LSST cadence**
- faint targets; limited followup resources
- need to join with heterogenous data sets, other alert streams
- LSST survey and tools must serve many science goals
- key scientific capabilities delegated to community brokers not directed by the LSST Project
- how is information shared in a distributed ecosystem?

How can changes to the cadence improve LSST's discovery space for time-domain science?



Little model-independent information between 30 minutes and 2 weeks.
Hard to filter!
Different rolling cadence behaviors by band?

Conclusions



LSST will deliver an alert stream of unprecedented scale and great scientific potential.

Cadence decisions will have a major effect on your ability to select objects for rapid followup.

TVS science is uniquely sensitive to cadence choices and will need to advocate coherently for its goals.