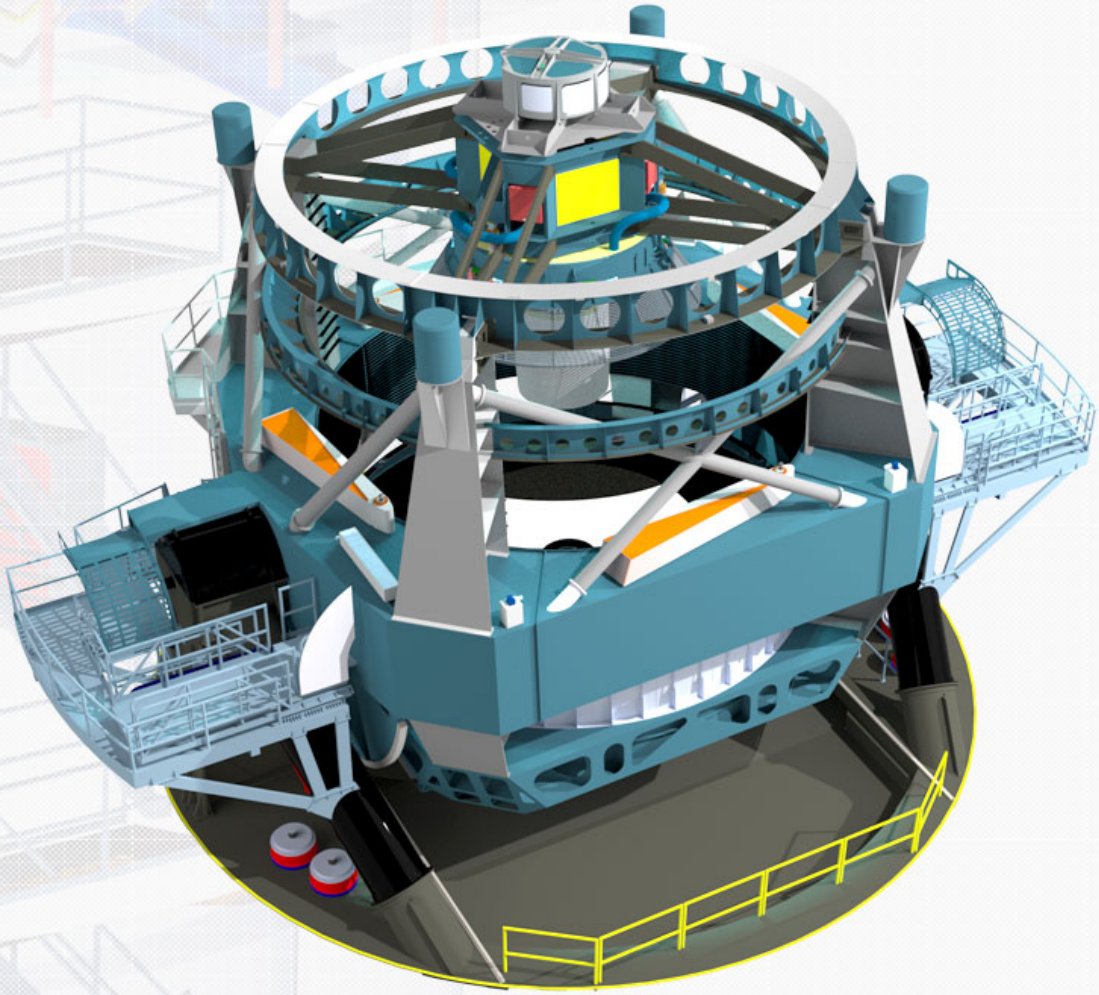
Discussion: Reporting Alerts



Reporting Time-domain Events

- For every source detected in a difference image, LSST will emit an "Event Alert" within 60 seconds of observation. **The primary use case is to enable real-time recognition and follow-up of transients of special interest.**

- Key elements of issued alerts:
  - 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 transmit nearly everything LSST knows about any given event, enabling downstream classification and decision making *without* the need to call back into LSST databases (thus introducing extra latency)**

LSST Alert Rates

- The alerts will be transmitted over the Internet in VOEvent format using community accepted protocols.

- Based on models of the rates of observed asteroids as well as stellar variability (dominant sources of event alerts), **we expect a high rate of alerts, approaching 10 million per night**.

- Consequences:
  - A typical LSST user will not have sufficient bandwidth to receive this full stream.
  - LSST will not have enough outgoing bandwidth to transmit a full stream to everyone; data size and bandwidth in the current baseline allow for transmission of ~three copies of the full data stream.

LSST Alert Filtering Service

- Most end-users will **not** be interested in reception of the full stream, but **only a subset that matches their scientific interest** (e.g., potential SNe candidates, variable stars, or moving objects).

- To support selecting such subsets of alert candidates, **LSST will provide an alert filtering service**. This service will let astronomers create simple *filters* that limit which alerts are ultimately forwarded to them.

- These user defined filters will be possible to specify using an SQL-like declarative language, or short snippets of (likely Python) code.

- **This is the Level 1 analog of querying the database in Level 2**, a service we also provide. The (significantly smaller) returned subset will then be transmitted to the end-user for analysis.

Example of a User-Defined Filter

```
# Keep only never-before-seen events within two
# effective radii of a galaxy. This is for illustration
# only; the exact methods/members/APIs may change.

def filter(alert):
    if len(alert.sources) > 1:
        return False
    nn = alert.diaobject.nearest_neighbors[0]
    if not nn.flags.GALAXY:
        return False
    return nn.dist < 2. * nn.Re
```

The user will subscribe to the alert stream by specifying a filtering function such as the one shown above. Once specified, only the alerts for which the function returns True will be forwarded to the user's VOEvent client.

Limitations of the LSST Filtering Service

- The complexity and run time of user defined filters will be limited by available resources and may be throttled depending on load.

- The number of VOEvents transmitted to each user will be limited and dynamically throttled depending on load.
  - E.g., with a maximum of ~20 events per visit per user, we can serve about ~500 simultaneous users at any one time utilizing total bandwidth equivalent to one full stream.

- No information beyond what is contained in the VOEvent packet will be available to user-defined filters (eg., no cross-matches to other catalogs, or other alert streams).

- We will not provide any astrophysical classification (eg., “is the light curve consistent with an RR Lyra?”, or “a Type Ia SN?”).

- **We don’t expect these limitations will generally cause serious hardship. The numbers of alerts returned (>=20,000/night) should be sufficient for rough pre-selection, without rejecting many potentially interesting candidates.**

Public VOEvent Brokers and Networks

- **We also anticipate that advanced, public, filtering services – VOEvent brokers – will be established by the community.**

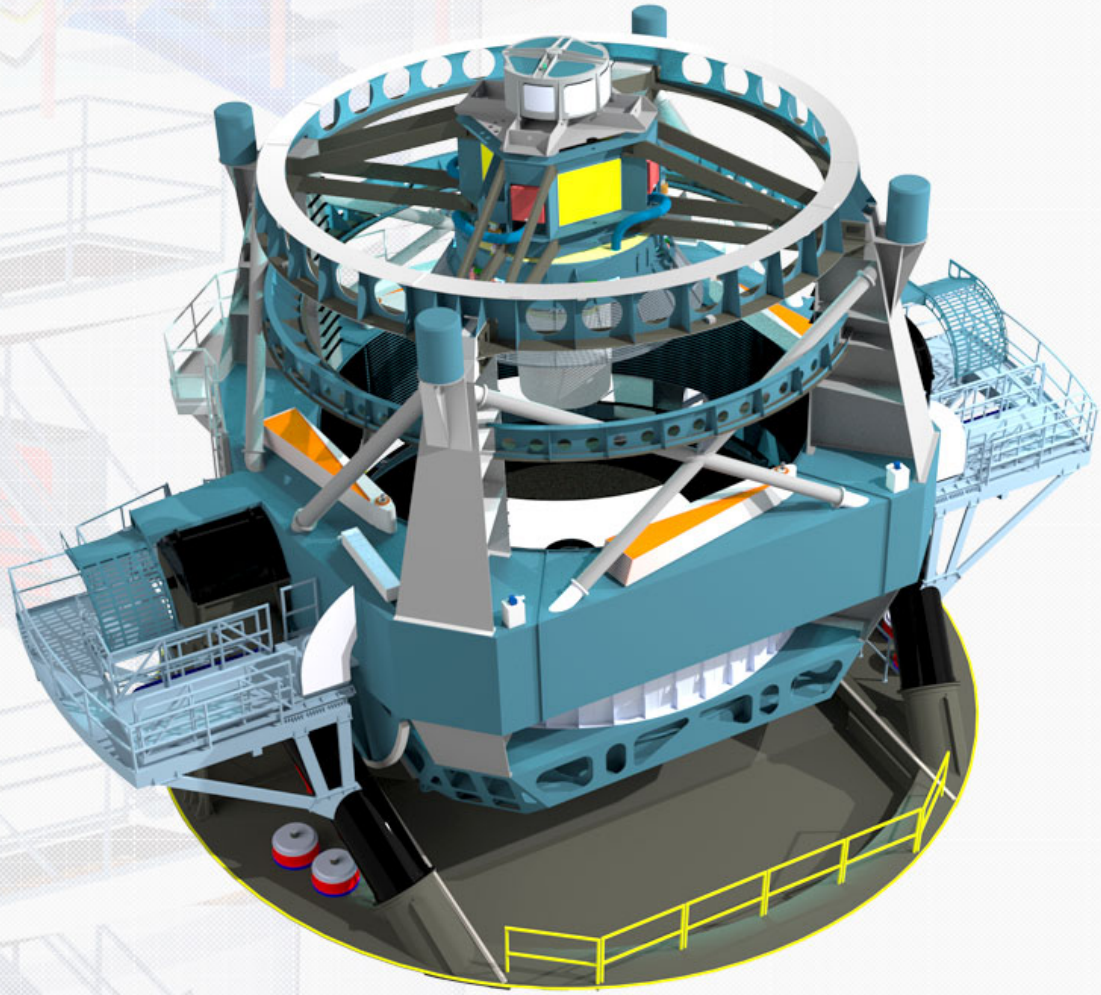
- Compared to the LSST filtering service, these may provide advanced functionality such cross-correlation of LSST alerts with external catalogs and other alert streams, classification engines, more extensive annotation of alerts, coordination of follow-up groups, and (more generally) incorporation of other contextual information needed to decide on whether a transient is worth following up.

- **Given the current baseline, we have reserved sufficient bandwidth to serve two to three public VOEvent brokers** (exact number is implementation dependent).

- If established, these will further simplify for the end-user the task of winnowing the alert stream to only contain the objects of interest for their science case. Because of their advanced functionality, we expect they will be preferred by the end-users, compared to the more limited LSST filtering service.

- **However, if such public brokering facilities fail to materialize, we expect the filtering service provided by LSST will be sufficient to enable initial Level 1 science.**

Discussion: The availability of older Data Releases in Operations



Current baseline

- **Catalogs:**
  - The catalogs for the latest data release ("DR N"), and the previous data release ("DR N-1"), are kept loaded in the database.
  - The catalogs from previous releases are stored on tape. They will be made available as bulk downloads.

- **Images:**
  - All raw images are available at all times (from tape)
  - All-sky co-adds are kept on disk only for DR N and N-1; others are on tape.
  - Reduced images are not stored at all, and are regenerated on the fly.
    - Regeneration with the software used for DR N and N-1 will be possible.

Keeping all DRs: Hardware Costs

- We have performed the preliminary analysis using the current sizing model
- The results are at https://jira.lsstcorp.org/browse/DM-3624

	v142	DR on Disk
Grand Total for Both Construction and Operations	$110,824,289	$112,390,714
Grand Total for Construction	$52,372,759	$52,419,559
DM Infrastructure Total	***$52,372,759***	***$52,419,559***
Annual Cost for Operations	***$5,845,153***	***$5,997,116***

- Roughly $150k/yr for ops, and $50k for construction
  - The reason for the small increment is our need to overprovision the storage to meet the I/O requirements

- **This is a very approximate analysis; more work is needed.**

Keeping all DRs: Development Costs

- Keeping more than just two (approximately equal in size) data releases on disk will change I/O access patterns. It also makes data replication and management more complicated.

- We currently estimate at least ~1 FTE-year of time in construction to analyze and implement changes necessary to support more than two DRs on spinning disk
  - Roughly ~$250,000 (a fully loaded sr. software developer)

- It's likely that there would also be some marginal increase in FTEs required for maintenance in operations.

- **<u>This is a preliminary estimate of development costs</u>**