The LSST Alert Stream

Eric Bellm, Alert Production Science Lead

LSST PST/SCC, April 24, 2018
Outline

Review alert stream components

Update on progress developing alert stream technologies

Highlight forthcoming “Broker Policy Document”

Discuss proposed plans for selecting community brokers

Discuss ongoing refinements of the LSST mini-broker concept
LSST’s Prompt Processing system will enable real-time science.

After an LSST image is taken, image subtraction will identify sources that move or vary within 60 seconds.

These Difference Image Analysis Sources (DIASources) will be packaged with contextual information into world-public alert packets for distribution.

The full stream of alerts will be sent to a small number of “community brokers”, who will add value to the stream and provide user access.

Additionally, LSST data rights holders will have access to a limited-capacity simple filtering service (the “mini-broker”).

There are many open questions about this ecosystem that we are now in the process of resolving; your feedback is welcome.
We have baselined a suite of open source technologies for distributing 10 million alerts/night.

Building a “bulk transport” system for distributing alerts to community brokers:

- Apache Avro (binary serialization format for alerts)
- Apache Kafka (realtime streaming platform; used for alert distribution)
- Apache Spark (cluster computing framework, may be used here for filtering)

LSST-scale experiments detailed in Data Management Technical Note #28 (dmtn-028.lsst.io)

acceptable performance at LSST scale

On track to meet NSF Level 2 milestone for Alert Distribution in July

A fork of this technology stack has recently been put into production by the Zwicky Transient Facility (ZTF) collaboration, with public distribution of alerts expected in June
We are developing a “Broker Policy Document” as a broad resource for understanding LSST’s alert plans.

Goal: In narrative language, outline LSST plans and policies for alert distribution for a broad audience.

Audience: LSST science users, community broker developers, funding agencies, LSST Project personnel


Timeline: complete draft to stakeholders by June 2018; CCB acceptance by AHM 2018
The Broker Policy document has four sections.

Components and Capabilities of the LSST Alert Distribution System

Data Rights to Alert Stream Components

Guidelines for Community Brokers

Guidelines for Users of the LSST Filtering Service and Science Platform
The Broker Policy document has four sections.

Components and Capabilities of the LSST Alert Distribution System

- Summarizes LSST data flow, data processing, and data products relevant to alert-driven science
- Provides links to governing design and requirements documents for interested readers
- **Answers questions like:**
  - What’s in an alert packet? (see ls.st/dpdd)
  - What happens in crowded fields? (Image subtraction proceeds as normal, but alerts may be throttled)
  - What are the bandwidth limits on alert distribution? (TBD)
  - Can DAC users query the Prompt and/or Data Release Production Difference Image Analysis (alert) databases? (yes)
The Broker Policy document has four sections.

Data Rights to Alert Stream Components

- Summarizes who can access which parts of the alert stream
- Broadly: alert packet contents are world-public; access to all other components requires data rights.
- **Answers questions like:**
  - *How does a scientist without data rights access the alert stream?* (through a community broker)
  - *Can community brokers access DAC resources that require data rights?* (Yes, but only if they can limit redistribution of LSST-proprietary data to users with data rights.)
The Broker Policy document has four sections.

Guidelines for Community Brokers

- Outlines a process by which community brokers will be chosen
- Suggests evaluation criteria and timeline—see subsequent slides.
- **Answers questions like:**
  - Are community brokers required to rebroadcast the alert stream to all? (not required, but proposals that do so will be stronger)
  - Is specific or minimum functionality (e.g., filtering, archiving) required to be a community broker? (Other than the technical capability to process the stream, no.)
  - Do community brokers have to receive the full alert stream? (Proposers can request pre-filtering using the technology employed by the LSST mini-broker if desired.)
We are suggesting an “open” call for community brokers.

Require demonstration of technical capability (see next slide) but no functional requirements, e.g.,

- No requirement to receive the full stream
- No requirement to redistribute the full stream
- No requirement to make products world public

But evaluation criteria (subsequent slides) are expected to favor proposals that offer these!

Any institution worldwide can propose

Goal is to feed a rich broker ecosystem offering diverse capabilities
Brokers must demonstrate adequate resources to accomplish their proposed goals.

Large inbound and outbound network bandwidth (the full alert stream is a few TB/night)

Petabytes of disk capacity

Databases handling of billions of sources

Compute resources to handle sophisticated classification and filtering tasks in real time at scale

Appropriate personnel to develop and maintain the service

Institutional support to ensure the longevity and stability of the service.
We propose brokers be evaluated on their contribution to the scientific utilization of LSST.

Serve a large community
Enable high-profile science
Provide unique capabilities
Contribute to LSST’s four science pillars
Take advantage of the unique aspects of the LSST alert stream (real-time, world-public)
Suggested broker evaluation criteria have been drafted.

Scientific value --- see previous slides
Availability of data products and services to the world community
Scientific Validity --- demonstrated on precursor surveys or data
Integration with the Time-Domain Ecosystem --- e.g., followup telescopes, other services
Community adoption
Complementarity to other selected brokers
Applicable existing agreements, if any
The proposed timeline for selection is under discussion.

2018: “Plans and Policies for LSST Alert Distribution” issued
2019: Document bulk transport format and interface
       Set up canned alert stream with sample precursor data
       Issue call for proposals for community brokers
2020: Broker proposals due
       Produce sample commissioning alerts with ComCam
       with substantial latency
2021: Finalize number of full streams
       Selection of community brokers for early operations
       Produce sample commissioning alerts with LSSTCam
       with substantial latency
       Integrate selected brokers with the LSST Alert System
2022: Begin full LSST operations
The Broker Policy document has four sections.

Guidelines for Users of the LSST Filtering Service

- Describes capabilities and limitations of the LSST mini-broker
- Users may upload simple filters that operate on single LSST packets and receive the output, subject to capacity constraints
- Many details are being resolved—see next slides.
- Answers questions like:
  - How many alerts can a mini-broker filter pass? (Currently 20 per user per visit)
  - Can exceptions be made to this limit? (If the technical solution allows it, a proposal process could allow for additional resources to be allocated.)
  - Can users make the outputs of their filters accessible to others, potentially in a private group? (Hopefully—group permissions are an expected feature of the Science Platform)
The mini-broker is not fully defined by existing documents.

Formal requirements: (LSE-61):

- Accept user-defined filters that act on single alert packets
- Provide pre-defined filters
- Provide up to 20 full alerts per user per visit for up to 100 simultaneously-connected users
- No latency requirement
- Other limitations: no exclusive classification, no crossmatching, potential for a TAC-like process.

Not prepared to adjust these performance requirements today but would like to enable broader utilization if possible
We envision a mini-broker more tightly coupled to the Science Platform.

Want to share visualization tools, user groups and permissions, database storage, compute, etc. with other aspects of the Science Platform. Will provide a real-time stream interface to enable fast-response science

- User manages alert storage on their own computing
- Expect to provide VOEvent (whatever is current); might also offer LSST’s bulk transport protocol, if it differs

Access to alerts from within the Science Platform

- This may be degenerate with querying the Prompt Products DB
- Can we provide a way to run the same filters used by mini-broker streams?

How can we facilitate connections to other channels in the scientific ecosystem?

- email, GCN, ATEL, Transient Name Server, Twitter... ?
- not clear users can run automated/scheduled jobs on the Science Platform
We expect to make substantial progress this year.

May: alert standardization discussions at the IVOA Interop

June: Draft of Broker Policy Document to stakeholders for review

July: NSF milestone test for alert distribution performance

August: Public release of Broker Policy Document
Summary

Current efforts suggest our approach to bulk alert transport is adequate.

A forthcoming “Broker Policy Document” will describe how community brokers will be chosen.

Development of mini-broker use cases and architecture is ongoing.

Feedback on all aspects is welcome now and when the draft policy document is circulated this summer.