VERA C. RUBIN OBSERVATORY

Distributing Alerts from the Rubin Observatory

Gregory Dubois-Felsmann for Eric Bellm

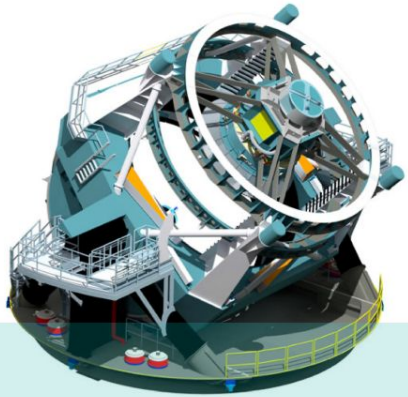
NSF AURA U.S. DEPARTMENT OF ENERGY Office of Science NOIR Lab SLAC CHARLES AND LISA SIMONYI FUND FOR ARTS AND SCIENCES LSST Discovery Alliance

Alert Production provides a world-public, near-real-time stream of LSST-identified transient, variable, and moving objects.

VERA C. RUBIN OBSERVATORY

Raw Data: 20TB/night

Sequential 30s images covering the entire visible sky every few days



LSST SCIENCE PIPELINES

Prompt Data Products

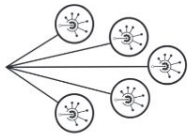
- Alerts incl. science, template and difference image cutouts — 60s — via Alert Streams
- Catalogs of detections incl. difference images, transient, variable & solar system sources — 24h — via Prompt Products
- Raw & processed visit images (PVIs), difference images — 80h — via Image Services

Data Release Data Products

Final 10yr Data Release:

- Images: 5.5 million x 3.2 Gpixels
- Catalog: 15PB, 37 billion objects

via Data Releases



Community Brokers

Rubin Data Access Centres (DACs)

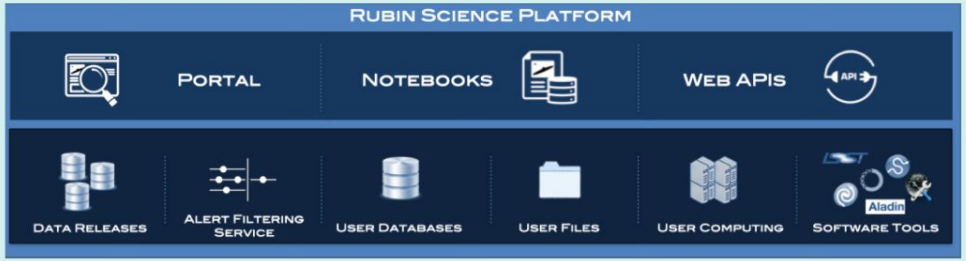
USA (USDF)
Chile (CLDF)
France (FRDF)
United Kingdom (UKDF)

Independent Data Access Centers (IDACs)

Access to proprietary data and the Science Platform require Rubin data rights

Rubin Science Platform

Provides access to LSST Data Products and services for all science users and project staff.



Credit: Leanne Guy

Rubin has adopted a "VO-first" approach to services.

We have committed to using IVOA standards wherever they reasonably fit our needs. E.g.,

- TAP/UWS and DataLink (service desc's) for catalog access; extensive UCD etc. annotations
- ObsTAP/DataLink/SODA for image access

We are contributing to standardization efforts to move things forward

- VOTable-in-Parquet, 'hipscat' for spatially-sharded Parquet catalogs
- P3T for a new generation of service standards
- Providing input on DataLink, DAP, etc.

The story has been a bit less straightforward for alerts.

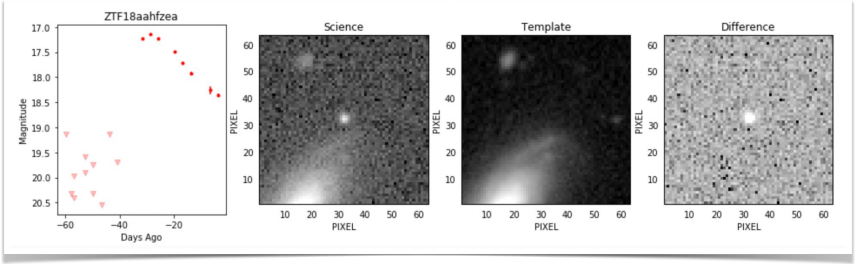
VOEvent 2.0 did not meet Rubin's needs for bulk alert transport to brokers.

Our 2018 Interop presentation summarized the major concerns:

- Rubin alerts are high volume & low-latency: bandwidth from the datacenter is the major constraint

  ⇒ An efficient binary serialization is necessary

  ⇒ VTP is too chatty for 10k alerts/39 seconds
- We need to embed cutout images in the alert packet



VOEvent solves a larger and more general problem and so sacrifices performance that Rubin requires for bulk alert transport.

Rubin has adopted Kafka + Avro for bulk alert transport to brokers.

Serialization: Apache Avro

Binary, strictly-schemaed serialization format

6x smaller than equivalent XML & 40x faster (de)serialization (Maeda 2012)

Transport: Apache Kafka

Distributed real-time message queue

Scales to trillions of daily messages

Stream consumers can replay topics as desired

These choices were made in 2018 and still seem defensible.

Kafka in particular continues to have broad utilization in industry.

ZTF has successfully used this stack since 2018.

Publications of the Astronomical Society of the Pacific, 131:018001 (7pp), 2019 January https://doi.org/10.1088/1538-3873/aae904

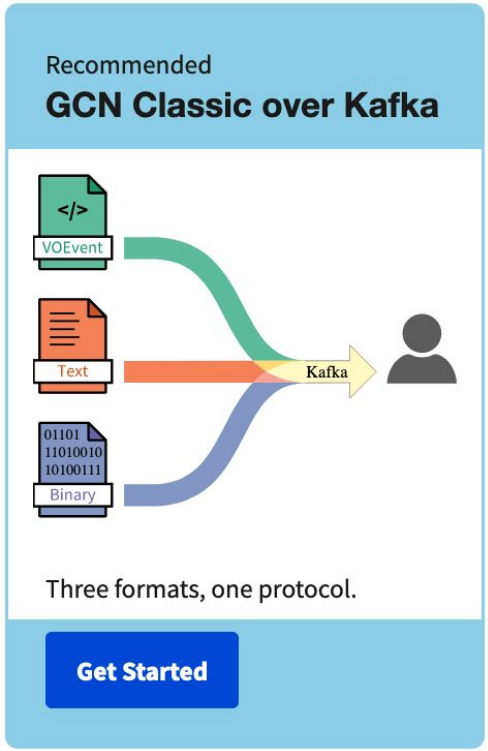
The Zwicky Transient Facility Alert Distribution System

Maria T. Patterson[1], Eric C. Bellm[1], Ben Rusholme[2], Frank J. Masci[2], Mario Juric[1], K. Simon Krughoff[3], V. Zach Golkhou[1,4,6], Matthew J. Graham[5], Shrinivas R. Kulkarni[5], and George Helou[2]
Zwicky Transient Facility Collaboration
[1] DIRAC Institute, Department of Astronomy, University of Washington, 3910 15th Avenue NE, Seattle, WA 98195, USA; mtpatter@uw.edu
[2] IPAC, California Institute of Technology, 1200 E. California Boulevard, Pasadena, CA 91125, USA
[3] LSST Project Office, 950 N. Cherry Avenue, Tucson, AZ 85719, USA
[4] The eScience Institute, University of Washington, Seattle, WA 98195, USA
[5] Division of Physics, Mathematics, and Astronomy, California Institute of Technology, Pasadena, CA 91125, USA
Received 2018 September 14; accepted 2018 October 10; published 2018 November 27

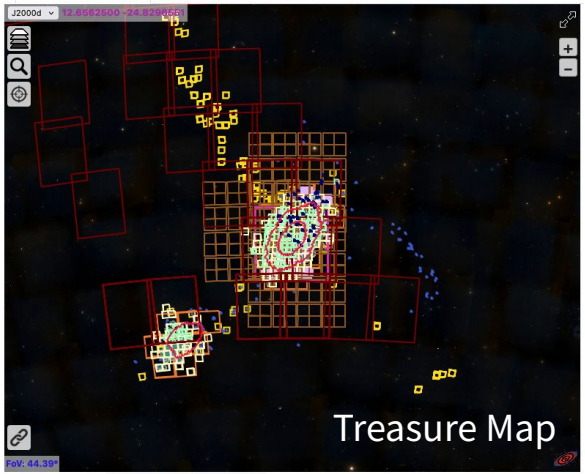
Patterson+19

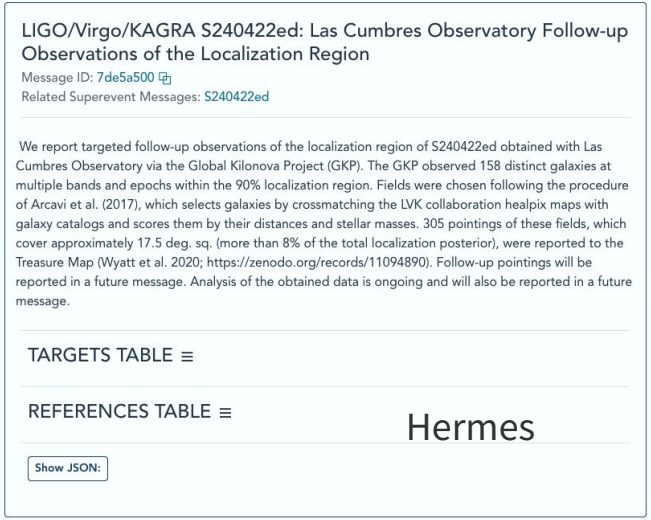
All seven Rubin-selected community alert brokers are currently receiving ZTF alerts via Kafka.

Many community time-domain services have been developed recently, largely outside of VOEvent.



HOPSKOTCH





These developments suggest some community preferences.

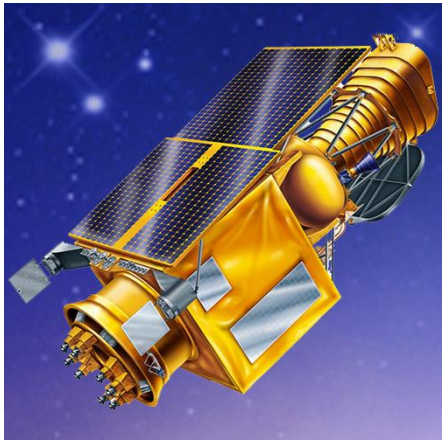
- Web-native APIs and services
- Modern development stacks and tooling
- Agility, rapid development, iteration, and experimentation

These concerns are largely disjoint from the technical challenges of Rubin bulk alert transport, and may suggest a community appetite for a broader reconsideration of the problem space.

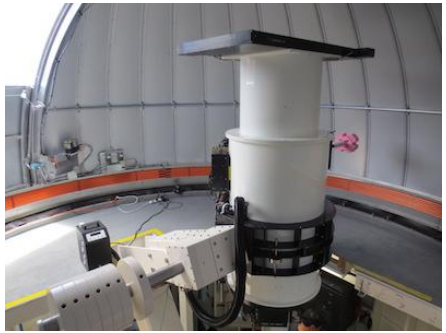
Most of these services operate in a low-event rate regime where VOEvent's affordances for authorship and citation are quite relevant.

The vision of interoperability that sparked VOEvent is more pressing than ever.

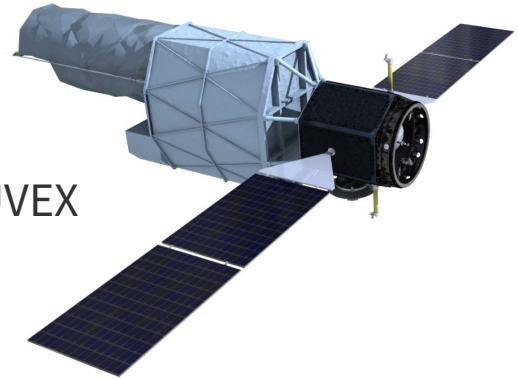
Alert-producing facilities are flourishing:



ULTRASAT



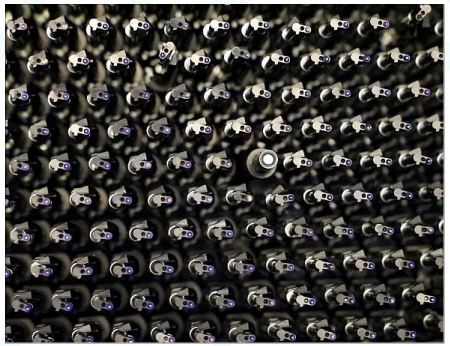
ATLAS



UVEX

Einstein Probe

DESI & other MOS surveys



What might VOEvent 3.0 look like?

An obvious and useful starting point would be to decouple the VOEvent data model from its serialization, and develop one or more modern serializations.

That could help solve Rubin's challenges with VOEvent 2.0, but that step alone seems unlikely to drive greater community adoption: Rubin-style alerts are only one (relatively simple!) corner of the problem space.

What are the carrots or "batteries included" that VOEvent can offer?

It seems valuable to build mindshare and consensus among a range of users and developers in the time domain ecosystem when envisioning the future evolution of VOEvent.