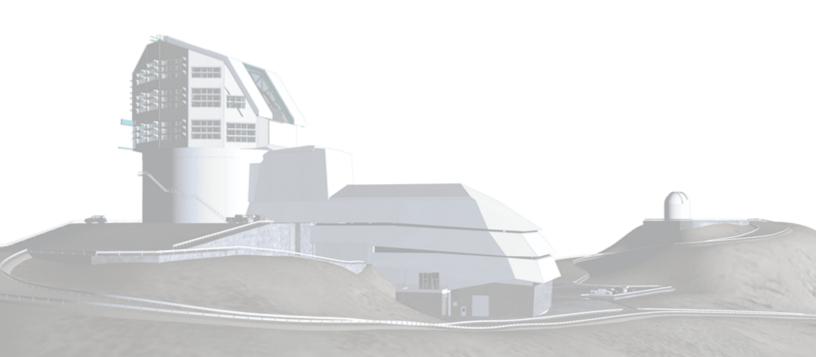**Vera C. Rubin Observatory
Data Management**

Implementing the Rubin/LSST Alert Filtering System with ANTARES

Eric C. Bellm, Leanne Guy

DMTN-226

Latest Revision: 2024-05-06

Abstract

This document presents a strategy for using the ANTARES alert broker as the Rubin Observatory Alert Filtering System.

Change Record

Version	Date	Description	Owner name
1	2022-08-02	Initial draft.	Eric Bellm
2	2023-06-14	Update implementation section.	Eric Bellm

Document source location: `https://github.com/lsst-dm/dmtn-226`

Contents

Implementing the Rubin/LSST Alert Filtering System with ANTARES

1 Background

During steady-state operations, Rubin Observatory's LSST will produce about ten million world-public alerts of transients, variable, and moving objects. Science users will work with community alert brokers to crossmatch, filter, and classify these alerts in order to identify the subset that require real-time followup observations. At the time the vision for this ambitious and comprehensive alert system was developed, no community alert brokers were functional. Accordingly, the Rubin project provided a fallback: a simple "alert filtering system" (AFS; §2) that the project would provide. The AFS would allow Rubin data-rights holders to apply simple user-defined or pre-defined filters to the Rubin alert stream. This minimized the risk that community alert brokers would not be available, or not provide the functionality required to fulfill users' needs.

Today, the situation is quite different. Seven community alert brokers[1] have been approved for direct access to the full Rubin alert stream, with two more planning to operate downstream of a full-stream broker. The brokers have mature systems and active user bases and are operating on live Zwicky Transient Facility alerts. Several brokers have functionality which mirrors that planned for the Rubin AFS. Accordingly, the rationale for a project-provided AFS is less clear.

However, it is reasonable for the Rubin project to ensure that Rubin data-rights holders have access to AFS-like capabilities throughout the survey, even if that service is provided by another entity. In this technote we propose a partnership with the ANTARES alert broker (Matheson et al., 2021), one of the seven approved full-stream brokers. Like the Rubin Operations project, ANTARES operates as part of NOIRLab.

In §2 we review the formal requirements for the Rubin AFS. §3 describes the ANTARES alert broker. In §4 we discuss steps necessary to allow ANTARES to supply AFS functionality to the Rubin data rights community.

---

[1] https://www.lsst.org/scientists/alert-brokers

2   Rubin Requirements for Alert Filtering

The top-level Science Requirements Document [LPM-17] describes the AFS in narrative language:

> The users will have an option of a query-like pre-filtering of this data stream in order to select likely candidates for specific transient type... Several pre-defined filters optimized for traditionally popular transients, such as supernovae and microlensed sources, will also be available, as well as the ability to add new pre-defined filters as the survey continues.

This is followed down into formal requirements as follows.

LSR, [LSE-29]:

LSR-REQ-0025 is supposed to be related but seems to be broken in current LSR

> Pre-defined filters optimized for traditionally popular transients shall be made available. It shall be possible for the project to add new pre-defined filters as the survey progresses.

LSR-REQ-0026

DMSR, [LSE-61]:

> A basic, limited capacity, alert filtering service shall be provided that can be given user defined filters to reduce the alert stream to manageable levels.

DMS-REQ-0342

> Users of the LSST Alert Filtering Service shall be able to use a predefined set of simple filters.

DMS-REQ-0348

> The LSST alert filtering service shall support `numBrokerUsers` (100) simultaneous users with each user allocated a bandwidth capable of receiving the equivalent of `numBrokerAlerts` (20) alerts per visit.

DMS-REQ-0343

Finally, the DPDD [LSE-163] provides a narrative description of the expected functionality from an end-user perspective.

> To directly serve the end-users, LSST will provide a basic, limited capacity, alert filtering service. This service will run at the LSST U.S. Archive Center (at NCSA). It will let astronomers create simple filters that limit what 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. For example, here's what a filter may look like: *(Code snippet removed)*
>
> We emphasize that this LSST-provided capability will be limited, and is not intended to satisfy the wide variety of use cases that a full-fledged public Event Broker could. For example, we do not plan to provide any exclusive classification to a unique category of object. Following the SRD specification, however, we will provide a limited number of pre-defined filters for a small number of object types of common interest. These will answer non-exclusive questions such as "is the light curve consistent with an RR Lyra?", and will have potentially highly overlapping selections, designed to provide good completeness but perhaps only very modest purity. No information beyond what is contained in the VOEvent packet will be available to the pre-defined or user-defined filters (e.g., no cross-matches to other catalogs). The complexity and run time of user defined filters will be limited by available resources. Execution latency will not be guaranteed. The number of VOEvents transmitted to each user per visit will be limited as well (e.g., the equivalent of 20 full-size alert packets per visit per user, dynamically throttled depending on load). Finally, the total number of simultaneous subscribers is likely to be limited—in case of overwhelming interest, a TAC-like proposal process may be instituted.

Implicit in the discussion above is that the AFS should be broadly useful the widest potential range of alert-driven science. Additionally, consistent with the Rubin Data Policy [RDO-013], the AFS should be accessible to any Rubin Data Rights holder (subject to the capacity limitations described above).

While the alert data itself is world-public, access to Rubin-hosted services like the AFS is restricted to Data Rights holders. However, there is no restriction on third parties offering similar filtering services.

LDM-554 provides several requirements on the Portal aspect of the Rubin Science Platform to provide appropriate user interfaces to the AFS.

The Portal aspect shall provide an interface to the alert subscription service that allows authenticated users with LSST data rights to subscribe to a stream of alert events.

<div align="right">DMS-PRTL-REQ-0119</div>

The Portal aspect shall provide an interface to permit alert subscriptions to be configured with Project-provided alert filters.

<div align="right">DMS-PRTL-REQ-0120</div>

The Portal aspect shall provide an interface to permit alert subscriptions to be configured with user-defined alert filters.

<div align="right">DMS-PRTL-REQ-0121</div>

The Portal aspect shall report feedback about the status and performance of a user's filters in the alert subscription service.

<div align="right">DMS-PRTL-REQ-0127</div>

During Rubin Project development of the AFS, additional unofficial "desirements" were collected[2].

3   The ANTARES Alert Broker

ANTARES is a community alert broker that has been operating on alerts from the Zwicky Transient Facility (ZTF) since 2018 and is approved for full stream access to LSST alerts. ZTF's public alert stream anticipates Rubin's both scientifically and technically and so provides a useful precursor for validating alert stream use cases prior to the start of the LSST. ANTARES offers a wide range of services to registered users from around the world, including alert visualization; rich positional and property searches; output Kafka streams; user watchlists; and rapid alerting via Slack. ANTARES also serves as the "upstream" broker for the SNAPS solar system broker.

ANTARES uses a "Locus" abstraction for timeseries data. New alerts are positionally associated with previous alerts at that location with a 1'' tolerance. If no previous detections are present,

---

[2] https://confluence.lsstcorp.org/display/~ebellm/Mini-Broker+Use+Cases+and+Draft+Requirements

a new Locus object is formed. Using a Locus allows users to make use of all of the historical lightcurve for an object rather than the limited window provided in each new alert. It also makes it convenient to attach multiwavelength crossmatch information.

ANTARES offers a capability for user filtering of the alert stream. Users may write filters in Python using the ANTARES development kit[3]. Filters may access all of the Locus information, including full timeseries history, catalog crossmatches, tags, and LIGO/Virgo detection information. Filters may provide outputs to the ANTARES database, Slack, and Kafka streams. These features provide valuable additional functionality beyond the minimal requirements of the Rubin AFS, which operate only on single alert packets with no crossmatching. ANTARES also provides access to predefined alert filters.

Filters are executed serially by "Pipeline Worker" jobs run inside Kubernetes, each of which attaches to a Kafka partition. Due to this design, user filters are not sandboxed from each other. Filters executed downstream can see tags applied by upstream filters and may be slowed by upstream filter execution. The ANTARES team mitigates these issues by reviewing filters before they are put into production and monitoring the operation of the filtering system. This increases the staff effort required to operate the ANTARES filtering system in steady state, but avoids the challenging problem of constructing a self-serve, fully sandboxed filtering system. Filter installation and updates thus cannot be conducted without staff intervention (i.e. within the night), although the ANTARES toolkit allows offline testing and tuning on sample data.

While having filters operate on Locus objects enables more scientific utility for stationary transients and variables, it limits the applicability of the ANTARES filtering system for Solar System Objects. Packet-based filtering of attributed Solar System Objects would be sufficient to satisfy the requirements of the Rubin AFS. This is not a current capability of ANTARES, although ANTARES does provide a stream of alerts from known Solar System Objects to the SNAPS solar system broker.

Currently, ANTARES executes 21 public and private filters on typically half a million alerts nightly. Filters in use today pass an average of 11% of alerts. (Since the filter outputs are implemented as tags, however, there are not strong technical pressures to limit the number of alerts which pass a filter, except in cases where the alerts are sent downstream through secondary Kafka streams.) These results suggest that ANTARES filtering can support the Rubin-required `numBrokerUsers` and `numBrokerAlerts`, although the ZTF-scale testing does not conclu-

---

[3]https://noao.gitlab.io/antares/filter-documentation/devkit/index.html

sively demonstrate performance at the scale of 10 million nightly Rubin alerts. However, the ANTARES team has performed scale testing that suggests acceptable performance at Rubin scale with planned hardware upgrades.

We thus conclude that it is highly likely that the ANTARES filtering service can meet the technical requirements of the Rubin AFS. The use of the Locus aggregation as an input to the filters provides scientifically valuable capabilities beyond the minimum requirements, but makes handling Solar System alerts more difficult.

4 Implementing the Rubin AFS with ANTARES

In January 2020, Rubin staff met with ANTARES personnel in Tucson to discuss whether the ANTARES filtering system could serve as the Rubin AFS, either as a software deliverable from the ANTARES team to be operated by the Rubin Project or as a service run by ANTARES. A follow-on video conference was held in April 2022. Participants in the latter teleconference agreed that the simplest and thus most cost-effective course of action would be to explore how the ANTARES-hosted filtering service could provide the functionality of the Rubin AFS.

Subsequent dialogue achieved convergence on the following topics.

4.1 Alert Throughput

ANTARES report cloud scaling tests run in GCP show that the ANTARES system can handle the average LSST alert rate (10,000 alerts every 39 seconds) with 100–150 nodes. Performance optimization is ongoing.

4.2 Access to Rubin Data Rights Holders

Currently ANTARES services are open to all users. The ANTARES team expect to easily support the number of Rubin Data Rights holders (100) specified in DMS-REQ-0343 [LSE-61] . In case of significant demand in excess of ANTARES capacity, particularly from non-Data Rights holders, a resource allocation committee might be required to determine the best way to ensure the required level of service.

DMS-REQ-0343

4.3 Filtering of Solar System Objects

The ANTARES architecture is built around non-moving objects but can easily extract light curves of known objects from our its locus database containing alerts. Solar System filtering to capture all known objects from ZTF is running nightly already with no further ANTARES development required. The SNAPS downstream broker currently is receiving this stream of Solar System Objects, and that is where filtering for activity or behavioral change would best be run.

4.4 Predefined Filters

The ANTARES system already contains public, predefined filters which would enable formal verification of the relevant requirement (LSR-REQ-0026; LSE-29). Additional collaboration between the Rubin Project, the ANTARES team, and interested science users would be valuable to broaden the range and quality of the predefined filters.

LSR-REQ-0026

4.5 Linkages between ANTARES and the Rubin Science Platform

The LDM-554 Portal requirements listed above should be reviewed. Some could be descoped in favor of existing ANTARES functionality.

4.6 Plan for Commissioning the ANTARES-AFS

The ANTARES-provided filtering system must be formally commissioned. We propose to use the same activities planned for integrating community brokers more generally; see RTN-011.

4.6.1 Verification

The major verification activity will be to develop and load 20 ANTARES filters configured to run on Rubin alerts; to send a simulated LSST alert stream based on DP0.2 processed images from the USDF; and to ensure that the throughput requirements are met.

4.6.2 Validation

Existing handling of the ZTF alerts by the ANTARES system already provides a great deal of real-world scientific validation. We will have limited ability to further validate the ANTARES alert filtering system on Rubin data in advance of public LSST alerts. RTN-061 describes the process for beginning public Alert Production; we expect alerts to be released to all community brokers publically as soon as they can be generated reliably. Sample alerts from Rubin-processed precursor surveys can serve as test datasets for development of predefined filters.

4.6.3 Timeline

The relevant DM milestone DM-AP-14[4] is currently due in January 2024.

4.7 Agreements

We expect the above understanding to be codified in a memo between the Rubin and CSDC Directors.

A References

[RTN-061], Bellm, E., 2023, Planning for the First Public Release of LSSTCam Alerts, URL `https://rtn-061.lsst.io/`,
Vera C. Rubin Observatory Technical Note RTN-061

[RDO-013], Blum, R., the Rubin Operations Team, 2020, Vera C. Rubin Observatory Data Policy,
URL `https://ls.st/RDO-013`,
Vera C. Rubin Observatory RDO-013

[LSE-29], Claver, C.F., The LSST Systems Engineering Integrated Project Team, 2017, LSST System Requirements (LSR), URL `https://ls.st/LSE-29`,
Vera C. Rubin Observatory LSE-29

---

[4]`https://dmtn-158.lsst.io/#dm-ap-14`

[LSE-61], Dubois-Felsmann, G., Jenness, T., 2019, Data Management System (DMS) Requirements, URL `https://lse-61.lsst.io/`,
Vera C. Rubin Observatory LSE-61

[LDM-554], Dubois-Felsmann, G., Ciardi, D., Mueller, F., Economou, F., 2019, Data Management LSST Science Platform Requirements, URL `https://ldm-554.lsst.io/`,
Vera C. Rubin Observatory Data Management Controlled Document LDM-554

[RTN-011], Guy, L.P., Bechtol, K., Bellm, E., et al., 2024, Rubin Observatory Plans for an Early Science Program, URL `https://rtn-011.lsst.io/`,
Vera C. Rubin Observatory Technical Note RTN-011

[LPM-17], Ivezić, Ž., The LSST Science Collaboration, 2018, LSST Science Requirements Document, URL `https://ls.st/LPM-17`,
Vera C. Rubin Observatory LPM-17

[LSE-163], Jurić, M., Axelrod, T., Becker, A., et al., 2023, Data Products Definition Document, URL `https://lse-163.lsst.io/`,
Vera C. Rubin Observatory LSE-163

Matheson, T., Stubens, C., Wolf, N., et al., 2021, AJ, 161, 107 (`arXiv:2011.12385`), doi:10.3847/1538-3881/abd703, ADS Link

B   Acronyms