

Alert Production with the Vera C. Rubin Observatory

VERA C. RUBIN OBSERVATORY TEAM¹

¹*Vera C. Rubin Observatory Project Office, 950 N. Cherry Ave., Tucson, AZ 85719, USA*

(Dated: April 20, 2026)

ABSTRACT

Describes the LSST Alert Production pipelines, data products, execution environment, and early performance.

Todo list

8	Eric: more	2	42	overview of the algorithm, training, and performance of the ML spuriousness score; detailed discussion likely deferred to separate paper ..	4
9	Eric: add some examples: FBOTS, ISOs, asteroids interior to venus,	2	43	Eric: schema tables?	4
10	Eric: VOEvent, GCN, SciMMA/HopSkotch, TNS	2	44	Eric: sizing estimates here or in initial performance section? if not, collapse into next section	4
11	Eric: cites	2	45	Eric: cite patterson, GCN, Hopskotch, other Rubin usage... ..	4
12	Eric: confirm	2	46	Bri: details of the deployment	4
13	Eric: cites throughout	2	47	Bri: schema management	4
14	Eric: brief summary of observations: commissioning, first alerts, ongoing data taking. What we expect to be in the PPDB.	2	48	Eric: deduplication	4
15	Eric: cite	2	49	Ian: catchup, single frame detection, visit-level consolidation, metrics	4
16	Eric: image release?	2	50	Jake: completeness, astrometric accuracy	5
17	Eric: final date	2	51	Eric: provide citations	5
18	Eric: consider a figure or histogram showing the alert rate over time, perhaps a sky plot of alert density?? or colored by earliest alert date?	2	52	Melissa: ANTARES system, community alert filters	5
19	appropriate citation	2	53	Gregory: todo	5
20	Eric: scheduler discussion/citation? link to survey-strategy.lsst.io? mention status dashboards at prompt-products.lsst.io?	3	54	confirm IDAC section	5
21	Eric: Summary of relevant aspects of the DPDD (Jurić et al. 2023), including latency considerations and user access. Describe key requirements that drive technical decisions. . .	3	55	Melissa: prompt-products.lsst.io, pipelines.lsst.io, community.lsst.io	5
22	Eric: confirm contents	3	56	Eric: summarize	5
23	Eric: point to schema browser. forward reference the PPDB.	3	57	Eric: review and update citations to use current DOIs	5
24	consider a (simplified?) pipeline graph	3	58		
25	emphasize not a static/permanent release; evolving over time	3	59		
26	glint detection	4	60		
27	satellite filtering	4	61		
28			62		
29			63		
30			64		
31			65		
32			66		
33			67		
34			68		
35			69		
36			70		
37			71		
38			72		
39			73		
40			74		
41			75		
			76		
			77		

1. INTRODUCTION

Time-domain measurements are among of the foundational techniques of astronomical science: repeated observations of the night sky enable discovery of transient, variable, and moving objects. The cadence of the observations, the intrinsic luminosities and timescales of the phenomena, and the spatial volume probed by the observations set the rate of discovery (Bellm 2016). In recent years, scientists have conducted dedicated optical time-domain surveys with large CCD mosaic cameras (e.g., PanSTARRS, DECam, ZTF, HSC, Gaia, Kepler,

TESS) as well as distributed telescope networks (e.g., ASAS-SN, ATLAS, KMT-Net,

Eric: more

). These have yielded large new samples of supernovae, variable stars, active galactic nuclei, and solar system objects. Additionally, they have uncovered the first exemplars of rare new classes of objects

Eric: add some examples: FBOTS, ISOs, asteroids interior to venus, ...

Frequently, time-critical followup observations are necessary to classify and characterize objects discovered in time-domain surveys. Historically, human-composed circulars or telegrams were used to disseminate discoveries to the wider community. The increasing rate of transient candidates as well as the desire for rapid automated followup motivated machine-readable alternatives

Eric: VOEvent, GCN, SciMMA/HopSkotch, TNS

. These trends culminated in the public “alert stream” paradigm employed by ZTF (Patterson et al. 2019), in which hundreds of thousands of unfiltered difference image sources are shipped along with historical lightcurves and image cutouts to third-party alert brokers for classification, filtering, and followup. This approach has enabled fully automated identification, reporting, and followup of supernova candidates

Eric: cites

The Legacy Survey of Space and Time (LSST) to be conducted by the Vera C. Rubin Observatory promises an order of magnitude increase in transient discovery. Rubin’s large collecting area, wide field of view, and fast readout and slew will deliver nearly a thousand

Eric: confirm

exposures across a wide swath of the Southern Hemisphere sky to unprecedented depths. This capability motivated the development of a rapid data processing pipeline to identify and publicize time-variable phenomena in LSST images: the Rubin Alert Production System (AP). Along with the annual Data Release Processing (DRP), these productions make use of the Rubin Science Pipelines software as well as the larger systems and infrastructure of Rubin Data Management (DM).

Eric: cites throughout

In this paper, we describe the design, implementation, and initial performance of the Rubin Alert Production system. In §2 we detail the alert-producing observations taken to date. §3 presents the science data products re-

sulting from alert production and §4 describes the data processing pipelines which produce them². We summarize the data processing environment and infrastructure in §5. We characterize the early scientific and technical performance of alert production in §6. In §7 and §8 we describe how data are served to scientists and the support channels available to users. We conclude in §9.

2. OBSERVATIONS

Eric: brief summary of observations: commissioning, first alerts, ongoing data taking. What we expect to be in the PPDB.

Unlike Rubin’s Data Releases (e.g., DP1, DP2

Eric: cite

), which contain a fixed dataset acquired before the processing begins, Alert Production provides a continuous, ongoing release of data obtained by Rubin that is released within seconds to days of observation. During commissioning, alerts were transmitted to community alert brokers for technical integration tests beginning on 2025-09-05. Since data from this period are available in alert history and the Prompt Products Database (PPDB; §7.3.4)

Eric: image release?

, we include them in this paper, although they were not available publicly at the time of processing. Real-time public alert release began on 2026-02-24. In this paper we include data obtained until

Eric: final date

. During the commissioning and early operations phases, scientific data taking was intermixed with ongoing technical commissioning activities that did not produce data suitable for alert production. Calibration, delivered image quality, and pipeline processing were under substantial development. Additionally, template images were only available initially for a small subset of the sky area. Accordingly, the quantity and quality of early alerts produced are not representative of steady-state performance, and not all data products are available. We specify the range of data included in our performance analysis (§6).

² Discovery of new solar system objects from daytime processing of the prompt data products is not described here but is presented in

appropriate citation

Eric: consider a figure or histogram showing the alert rate over time, perhaps a sky plot of alert density?? or colored by earliest alert date?

Eric: scheduler discussion/citation? link to survey-strategy.lsst.io? mention status dashboards at prompt-products.lsst.io?

3. PROMPT DATA PRODUCTS

Eric: Summary of relevant aspects of the DPDD (Jurić et al. 2023), including latency considerations and user access. Describe key requirements that drive technical decisions.

3.1. Alerts

Alert packets are the primary data product for time-critical LSST science. Each alert packet is triggered by a single positive or negative detection on a difference image. This measurement is packaged with a variety of contextual information in order to enable downstream users to quickly identify events of interest without requiring followup queries to Rubin databases.

Each alert packet contains

- The triggering DIASource record
- The corresponding DIAObject or SSOBJECT record
- Any DIASource and DIAForcedSource records that exist, and difference image noise estimates where they do not, taken from the previous 12 months.
- Cutout images of the science, template, and difference images.

Eric: confirm contents

The cutout images are packaged in the FITS format using the `astropy CCDData` class and include mask and variance planes, an estimate of the PSF, a World Coordinate System, and other metadata. Most cutouts are 30×30 pixels but are expanded to a larger size for extended sources. The maximum cutout size of 102×102 pixels allows cutouts to contain trailed Near Earth Asteroids traveling slower than ten degrees per day while avoiding the performance impacts that would result from larger cutouts.

Alerts are transmitted and stored in Apache Avro³ format, a strongly-schemaed compact binary serialization. However, the Avro standard includes its schemas as uncompressed ASCII. To conserve bandwidth, we therefore send alerts without including the schemas required

to deserialize them. Instead, the schemas are stored in a separate schema registry and are referenced by a unique identifier in the alert packet. (Patterson et al. 2024) describes the alert packaging and schema versioning used in the alerts. Avro alert schemas are programatically generated from the database schema definitions stored in the `lsst.sdm.schemas` package and can be found in the `lsst.alert.packet` repository⁴.

3.2. Images

3.3. Prompt Catalogs

Eric: point to schema browser. forward reference the PPDB.

3.4. Data Rights

Blum & the Rubin Operations Team (2020) describes the Rubin Data Rights policies. In brief, alert packets are world-public and can be freely shared with anyone. Similarly, the contents of the PPDB are freely sharable, although direct access to the PPDB is restricted to Data Rights holders. All other data products are proprietary.

4. PIPELINE

The Alert Production Pipelines are a specialized configuration of the LSST Science Pipelines. ? provides a general overview of their design and capabilities. In this section we describe the detailed implementation used for Alert Production.

consider a (simplified?) pipeline graph

emphasize not a static/permanent release; evolving over time

4.1. Template Generation

4.2. Preload

4.3. Single Frame Processing

ISR, PSF and Background fitting, photometric and astrometric calibration. Likely major commonalities with DRP; perhaps point to DP2 paper and describe any differences

4.4. Image Differencing

including background subtraction

4.5. Source Detection and Measurement

Including discussion of point source, dipole, and trail detection on difference images.

³ <https://avro.apache.org/>

⁴ https://github.com/lsst/alert_packet/

4.6. Initial Filtering

glint detection

satellite filtering

`FilterDiaSourceCatalogTask` in `lsst.ap.association`

4.7. Reliability scoring

Image-differencing searches for transients typically contend with high rates of false positives. Among the raw detections, artifacts may dominate real astrophysical sources by an order of magnitude. These may be due to imperfectly-corrected instrument signatures, astrometric mismatches between the template and the science image, cosmic rays, or algorithmic failures in the differencing. Modern surveys employ machine-learned classifiers to winnow these candidates (e.g., Bailey et al. 2007; Bloom et al. 2012; Brink et al. 2013; Wright et al. 2015; Goldstein et al. 2015; Duev et al. 2019). These have steadily improved in performance and sophistication as the classifiers transitioned from Random Forest models trained on manually-constructed feature sets to deep neural networks trained directly on image pixels. Some approaches have combined both detection and scoring without using the difference image (Sedaghat & Mahabal 2018; Acero-Cuellar et al. 2023).

overview of the algorithm, training, and performance of the ML spuriousness score; detailed discussion likely deferred to separate paper

(Acero-Cuellar 2026)

4.8. Catalog Transformation

`TransformDiaSourceCatalogTask` in `lsst.ap.association`

Eric: schema tables?

4.9. Source Association

Due to AP’s real-time nature, it is not possible to perform *post-facto* source association as in the annual data releases. Instead, spatial association is performed on-the-fly by the `DiaPipelineTask` within the `lsst.ap.association` package. A dedicated Alert Production Database (APDB; §5.2) holds the current state of the system.

During the preload step (§4.2), `DIAObjects` and `SSObjects` overlapping the expected field of view of the image are stored in the local prompt processing worker butler repository. The new `DIASources` are first spatially associated with the existing `DIAObjects` by finding the closest match within a maximum distance of one arcsecond. Pairs with the closest spatial separations are joined first to minimize misassociations. When a match is found, the `DIASource` is added to the `DIAObject`’s list of measurements. Unassociated `DIASources`

are then spatially associated with the predicted positions of the `SSObjects` at the time of the exposure. For `DIASources` with no matching `DIAObject` or `SSObject`, a new `DIAObject` is created and the `DIASource` is added to it.

4.10. Alert Generation

Each `DIASource` not filtered for quality purposes (§4.6) produces an alert (§3.1).

Eric: sizing estimates here or in initial performance section? if not, collapse into next section

4.11. Alert Distribution

We use the Kafka distributed streaming platform⁵ to transmit alerts to brokers. Kafka provides a fault-tolerant, low-overhead mechanism to transmit large volumes of data to multiple consumers and is widely used both in industry and in astronomy.

Eric: cite patterson, GCN, Hopskotch, other Rubin usage...

The Alert Distribution Kafka system is deployed at the USDF using Kubernetes.

Bri: details of the deployment

Bri: schema management

Because of the large size of the alert packets and the short latency requirement, bandwidth out of the US Data Facility (USDF) at SLAC limits the number of direct recipients of the full alert stream. Seven community alert brokers receive the full alert stream directly from the USDF (§7.1).

4.12. Forced Photometry

4.13. Source Injection

4.14. Metrics

5. DATA PROCESSING

5.1. Prompt Processing Framework

(Findeisen et al. 2026)

5.2. Alert Production Database

5.3. Metrics and Monitoring

5.4. Morning Batch Processing

Eric: deduplication

Ian: catchup, single frame detection, visit-level consolidation, metrics

⁵ <https://kafka.apache.org/>

6. INITIAL PERFORMANCE

6.1. *Template Properties*

6.2. *Alert Rate*

6.3. *Alert Purity*

Raw and after ML scoring

6.4. *Alert Completeness*

6.5. *Alert Latency*

6.6. *Photometric Precision*

6.7. *Astrometric Precision*

6.8. *Association Accuracy*

6.9. *Association of Known Solar System Objects*

Jake: completeness, astrometric accuracy

6.10. *Direct Source Catalogs*

6.11. *Crowded Field Performance*

7. DATA ACCESS SERVICES

7.1. *Community Alert Brokers*

Near real-time access to alert packets is available through services provided by community alert brokers. Due to limited bandwidth from the USDF, seven brokers were selected through a proposal process (Bellm et al. 2020) to receive the alert stream directly from the alert distribution system at the USDF (§4.11):

- ALerCE: Automatic Learning for the Rapid Classification of Events
- AMPEL: Alert Management, Photometry, and Evaluation of Light Curves
- ANTARES: Arizona-NOIRLab Temporal Analysis and Response to Events System
- Babamul
- Fink
- Lasair
- Pitt-Google

As alert packets are world-public and freely sharable (§3.4), additional brokers and other services may operate “downstream” by connecting to full or filtered streams provided by the upstream brokers. As of this writing, downstream brokers include

- Point of Interest Broker
- SNAPS: Solar System Notification Alert Processing System

Eric: provide citations

7.2. *Alert Filtering Service*

Melissa: ANTARES system, community alert filters

7.3. *Rubin Science Platform*

Gregory: todo

7.3.1. *Alert Archive*

7.3.2. *Image Services*

7.3.3. *Butler Data Products*

7.3.4. *Prompt Products Database*

7.4. *IDACs*

confirm IDAC section

8. SUPPORT FOR COMMUNITY SCIENCE

Melissa: prompt-products.lsst.io, pipelines.lsst.io, community.lsst.io

9. CONCLUSION

Eric: summarize

ACKNOWLEDGMENTS

This material is based upon work supported in part by the National Science Foundation through Cooperative Agreements AST-1258333 and AST-2241526 and Cooperative Support Agreements AST-1202910 and AST-2211468 managed by the Association of Universities for Research in Astronomy (AURA), and the Department of Energy under Contract No. DE-AC02-76SF00515 with the SLAC National Accelerator Laboratory managed by Stanford University. Additional Rubin Observatory funding comes from private donations, grants to universities, and in-kind support from LSST-DA Institutional Members.

Facility: Rubin:Simonyi (LSSTCam), Rubin:USDAC

Eric: review and update citations to use current DOIs

Software: Rubin Data Butler (Jenness et al. 2022), LSST Science Pipelines (?), LSST Feature Based Scheduler v3.0 (Yoachim et al. 2024; ?) Astropy (Astropy Collaboration et al. 2013, 2018, 2022)

REFERENCES

- 398 Acero-Cuellar, T. 2026, Performance of Machine-Learned
399 Reliability Scoring for Image Differencing, Data
400 Management Technical Note DMTN-337, NSF-DOE Vera
401 C. Rubin Observatory. <https://dmtn-337.lsst.io/>
- 402 Acero-Cuellar, T., Bianco, F., Dobler, G., et al. 2023, AJ,
403 166, 115
- 404 Astropy Collaboration, Robitaille, T. P., Tollerud, E. J.,
405 et al. 2013, A&A, 558, A33,
406 doi: [10.1051/0004-6361/201322068](https://doi.org/10.1051/0004-6361/201322068)
- 407 Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M.,
408 et al. 2018, AJ, 156, 123, doi: [10.3847/1538-3881/aabc4f](https://doi.org/10.3847/1538-3881/aabc4f)
- 409 Astropy Collaboration, Price-Whelan, A. M., Lim, P. L.,
410 et al. 2022, ApJ, 935, 167, doi: [10.3847/1538-4357/ac7c74](https://doi.org/10.3847/1538-4357/ac7c74)
- 411 Bailey, S., Aragon, C., Romano, R., et al. 2007, ApJ, 665,
412 1246
- 413 Bellm, E., Blum, R., Graham, M., et al. 2020, Plans and
414 Policies for LSST Alert Distribution, Data Management
415 Controlled Document LDM-612, NSF-DOE Vera C.
416 Rubin Observatory, doi: [10.71929/rubin/2998139](https://doi.org/10.71929/rubin/2998139)
- 417 Bellm, E. C. 2016, PASP, 128, 084501
- 418 Bloom, J. S., Richards, J. W., Nugent, P. E., et al. 2012,
419 PASP, 124, 1175
- 420 Blum, R., & the Rubin Operations Team. 2020, Vera C.
421 Rubin Observatory Data Policy, Data Management
422 Operations Controlled Document RDO-013, NSF-DOE
423 Vera C. Rubin Observatory. <https://ls.st/RDO-013>
- 424 Brink, H., Richards, J. W., Poznanski, D., et al. 2013,
425 MNRAS, 435, 1047
- 426 Duev, D. A., Mahabal, A., Masci, F. J., et al. 2019,
427 MNRAS, 489, 3582
- 428 Findeisen, K., Lim, K.-T., Speck, D., et al. 2026, arXiv
429 e-prints, arXiv:2603.19541,
430 doi: [10.48550/arXiv.2603.19541](https://doi.org/10.48550/arXiv.2603.19541)
- 431 Goldstein, D. A., D’Andrea, C. B., Fischer, J. A., et al.
432 2015, AJ, 150, 82
- 433 Jenness, T., Bosch, J. F., Salnikov, A., et al. 2022, in
434 Society of Photo-Optical Instrumentation Engineers
435 (SPIE) Conference Series, Vol. 12189, Software and
436 Cyberinfrastructure for Astronomy VII, 1218911,
437 doi: [10.1117/12.2629569](https://doi.org/10.1117/12.2629569)
- 438 Jurić, M., Axelrod, T. S., Becker, A. C., et al. 2023, Data
439 Products Definition Document, Systems Engineering
440 Controlled Document LSE-163, NSF-DOE Vera C. Rubin
441 Observatory, doi: [10.71929/rubin/2587118](https://doi.org/10.71929/rubin/2587118)
- 442 Patterson, M. T., Bellm, E. C., Swinbank, J. D., Nelson, S.,
443 & Smart, B. M. 2024, Design of the LSST Alert
444 Distribution System, Data Management Technical Note
445 DMTN-093, NSF-DOE Vera C. Rubin Observatory,
446 doi: [10.71929/rubin/2586493](https://doi.org/10.71929/rubin/2586493)
- 447 Patterson, M. T., Bellm, E. C., Rusholme, B., et al. 2019,
448 PASP, 131, 018001
- 449 Sedaghat, N., & Mahabal, A. 2018, MNRAS, 476, 5365
- 450 Wright, D. E., Smartt, S. J., Smith, K. W., et al. 2015,
451 MNRAS, 449, 451
- 452 Yoachim, P., Jones, L., Eric H. Neilsen, J., & Becker, M. R.
453 2024, `lsst/rubin_scheduler`: v3.0.0, v3.0.0, Zenodo,
454 doi: [10.5281/zenodo.13985198](https://doi.org/10.5281/zenodo.13985198)