

ZTF Alert Packet Contents

Frank Masci

ZTF Public Alerts Readiness Review, May 2018



ZTF Science Data System (ZSDS)

background material

- Primary Data System document: design, pipelines, deliverables, product usage, and data access:
http://web.ipac.caltech.edu/staff/fmasci/ztf/ztf_pipelines_deliverables.pdf
- Presentation from March 2018 Science Team Meeting:
http://web.ipac.caltech.edu/staff/fmasci/home/miscscience/masci_pasadena_03.19.18.pdf
- Overview paper is currently in peer review (to be published this June):

THE ZWICKY TRANSIENT FACILITY: DATA PROCESSING, PRODUCTS, AND ARCHIVE

FRANK J. MASCI,¹ RUSS R. LAHER,¹ BEN RUSHOLME,¹ DAVID L. SHUPE,¹ STEVEN GROOM,¹ JASON SURACE,¹ EDWARD JACKSON,¹
SERGE MONKEWITZ,¹ RON BECK,¹ DAVID FLYNN,¹ SCOTT TEREK,¹ WALTER LANDRY,¹ EUGEN HACOPIANS,¹ VANDANA DESAI,¹
JUSTIN HOWELL,¹ TIM BROOKE,¹ DAVID IMEL,¹ QUAN-ZHI YE,^{1,2} HSING-WEN LIN,³ S. BRADLEY CENKO,⁴ GINNY CUNNINGHAM,⁴
UMAA REBBAPRAGADA,⁵ BRIAN BUE,⁵ ADAM A. MILLER,^{6,7} ASHISH MAHABAL,² ERIC C. BELLM,⁸ MARIA PATTERSON,⁸
RICHARD WALTERS,⁹ MATTHEW GRAHAM,² MANSI M. KASLIWAL,² RICHARD G. DEKANY,⁹ THOMAS KUPFER,² TOM BARLOW,⁹
ANGELA VAN SISTINE,¹⁰ REED RIDDLE,⁹ ROGER M. SMITH,⁹ GEORGE HELOU,¹ THOMAS A. PRINCE,² AND
SHRINIVAS R. KULKARNI²

What are alert packets?

- Self-contained files containing metadata for a single event extracted from a difference image
 - triggered from any flux transient, moving object, or artifact detected at some observed epoch
 - transmitted typically within 15 minutes of observation
- Metadata allows one to:
 - perform further vetting for reliability
 - filtering according to specific scientific interest
 - whether the event is worthy of follow-up
 - global or long-term statistical studies, clustering analyses to uncover exotic phenomena
- Packets are in the Apache Avro™ format – a serialized binary format for efficient distribution and parsing on multiple platforms, with a large software base for porting to other formats, databases...
- Alert packet contents have evolved considerably following feedback received from the partner science programs

Packet Structure

- Related metadata is grouped into fields, each defined by a JSON-based schema
- There are four overall fields:
 - *objectId* and *candid* – stores name and identifier for alert
 - *candidate* – stores metadata for actual event that triggered the alert according to some threshold
 - *prv_candidates* – stores metadata for historical positionally-associated events
 - *cutout[Science,Template,Difference]* – image cutouts from new, reference, and subtraction images

```
{
  "namespace": "ztf",
  "type": "record",
  "name": "alert",
  "doc": "avro alert schema",
  "version": "1.8",
  "fields":
  [
    1  {"name": "objectId", "type": "string", "doc": "add descriptions like this"},
      {"name": "candid", "type": "long"},
    2  {"name": "candidate", "type": "ztf.alert.candidate"},
    3  {"name": "prv_candidates", "type": "ztf.alert.prv_candidate"},
      {"name": "cutoutScience", "type": ["ztf.alert.cutout", "null"], "default": null},
    4  {"name": "cutoutTemplate", "type": ["ztf.alert.cutout", "null"], "default": null},
      {"name": "cutoutDifference", "type": ["ztf.alert.cutout", "null"], "default": null}
  ]
}
```

candidates schema in alert packet

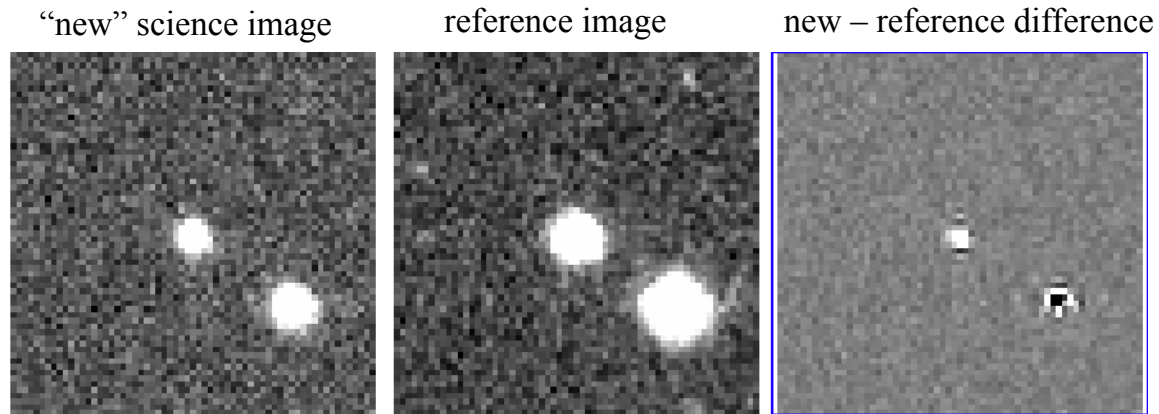
- An objectId (or name, e.g., ZTF18abcdefg) is assigned to an alert if none was previously assigned; name is reused if a new alert is positionally coincident with a previously named alert
- 82 source-extraction and image-based metrics: source shape; photometry; timing and position information; science, reference, and difference image quality metrics; input span used to generate reference image; identifiers to enable retrieval of ancillary products from archive
- *RealBogus* reliability score preassigned by machine-linear classifier; optimized for point-source events; this classifier was pre-trained on a subset of the above features
- Number of previous events detected on/near alert position going back to beginning of survey
- Number of historical images covering the alert position
- The nearest known Solar System object within a specific radius; if found, metadata is reported.
- The nearest reference-image detected source within a specific radius, with metadata.
- The closest, second closest, and third closest Pan STARRS1 source falling within a specific radius. If found, report distances, magnitudes, PS1 catalog IDs, and star-galaxy classification scores.

previous (historical) candidates schema

- All previous extracted events falling on/near the event that triggered the alert going back 30 days – constrained by database/compute resources, and overall runtime of realtime pipeline
- Contains 51 source-based and image-based metrics – a subset of the primary *candidate* metrics
- Includes detections from any filter (g, R, i), not only filter of observation that triggered event
- Includes detections from both *positive* (new – reference) and *negative* (reference – new) difference
- If a historical image covered the alert position (regardless of filter) but no association with a difference-image detection was found; a flux upper limit is reported, with nulls for other metrics
- Nearest Solar System object with metadata
- Nearest reference-image source with metadata

Image cutouts in packets

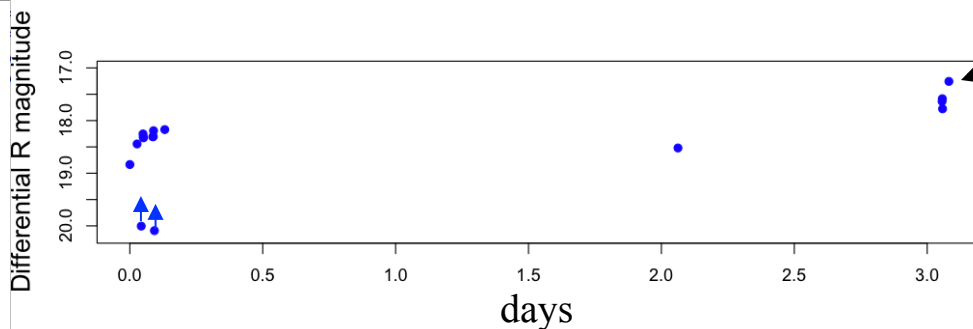
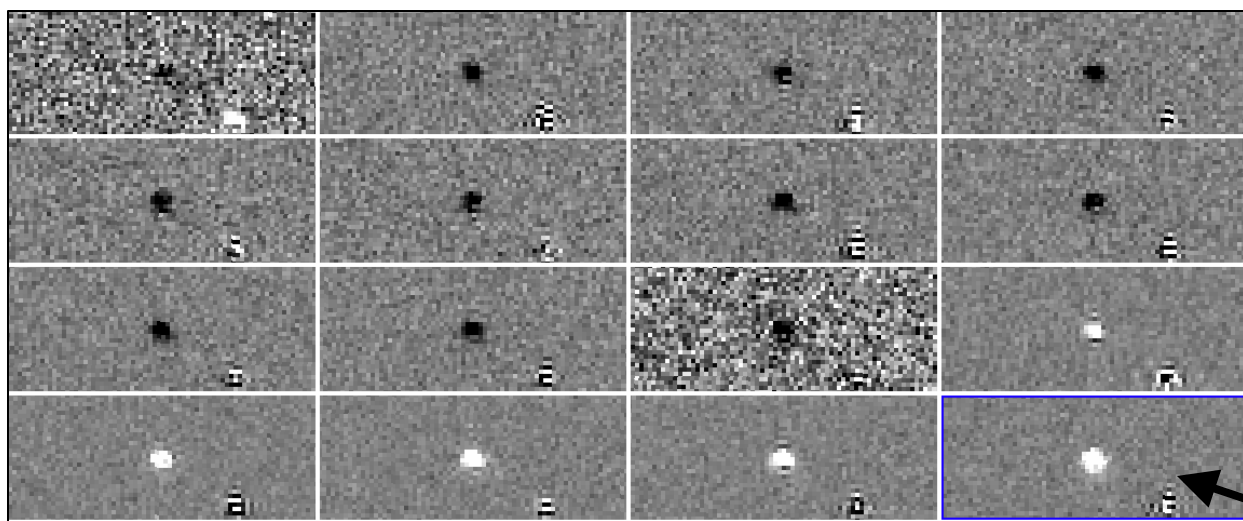
- Three image-cutouts centered on alert position in the new, reference, and difference images
- 32-bit floating point images, retaining full precision from input images in *.fits.gz* format
- Cutouts are 63 x 63 pixels square where 1 pixel \sim 1 arcsec.



- Archive product ID information is provided in packet to enable retrieval of full images and/or associated source catalogs, or perform custom cutouts on any of the historical overlapping images

Example Alert packet: variable star trigger

- Example with event triggered on a variable star (bottom right-most cutout)
- Associated history includes metadata on detections from both *positive* (science – reference) & *negative* (reference – science) difference images
- Historical image cutouts here were retrieved from ZTF archive
- Each historical event may have triggered its own separate alert packet

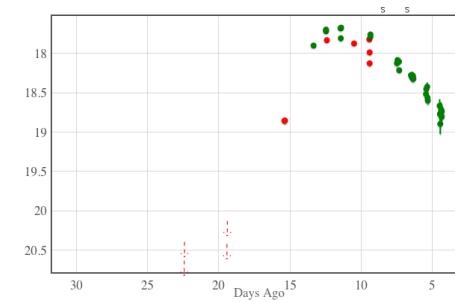
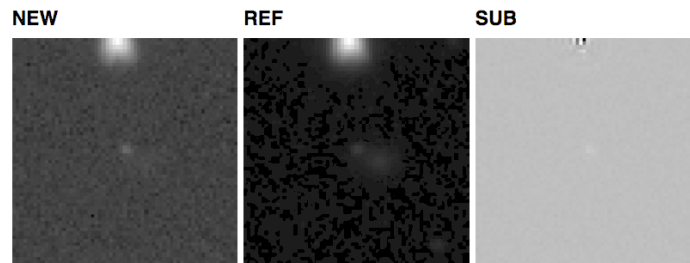
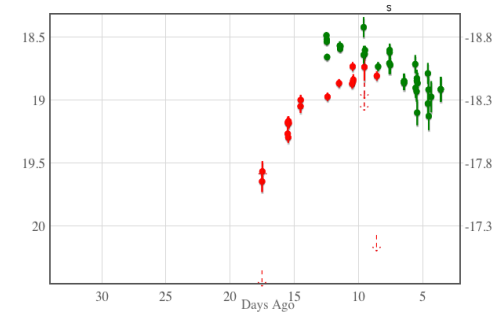
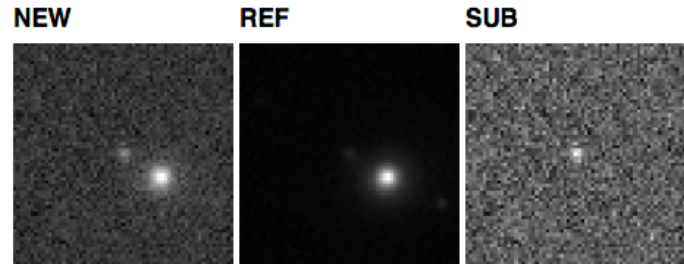


Alert was triggered on this

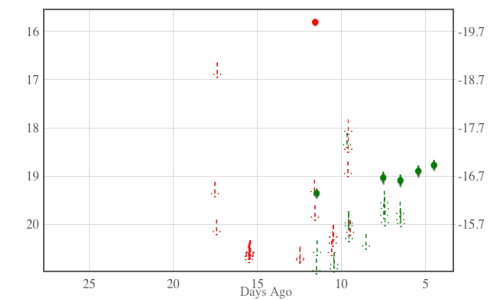
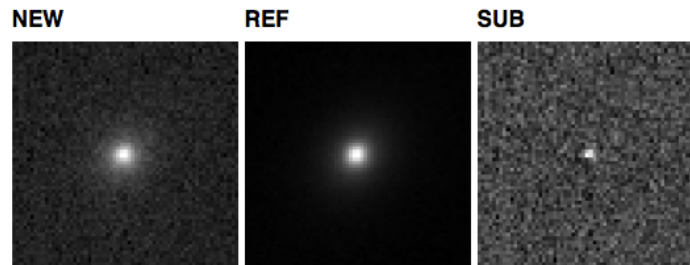
Examples from ZTF science programs

From the *Global Relay of Observatories Watching Transients Happen* (GROWTH) scanning page:

Type Ia SNe

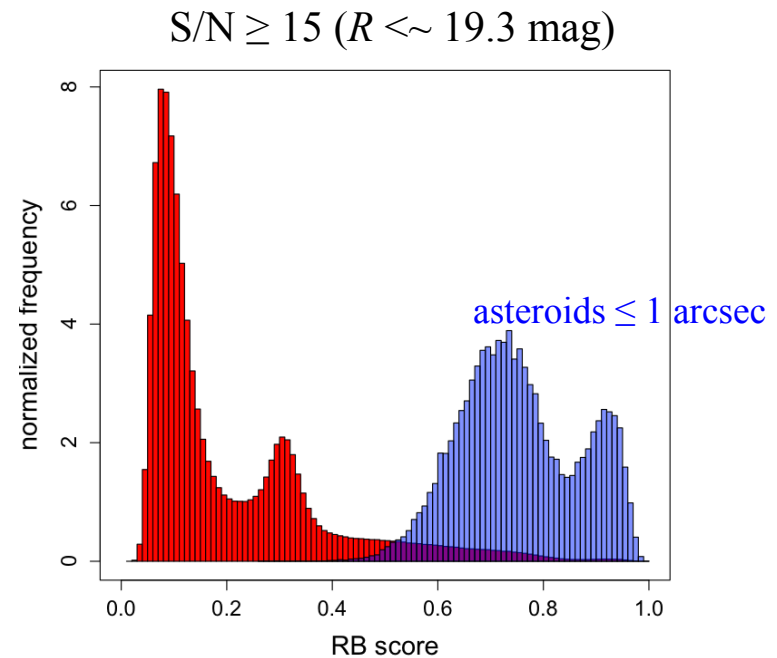
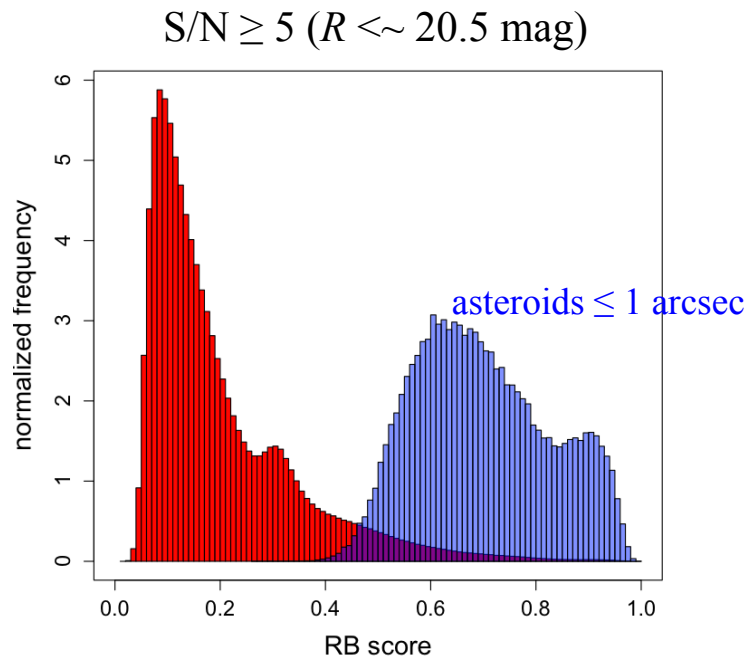


Nuclear transients
AGN



Analysis of alert stream

- Used associations with known asteroids to explore cuts on reliability score from machine-learning
- Asteroids provide an ideal truth set
- Use *RealBogus* (*RB*) scores computed using latest ML model: deployed on April 5, 2018



Recommendation for *initial* public alerts

- From visual examination of alerts and numbers, picked a working cutoff of $RB = 0.85$ (tail in red histograms)
- Number of alerts with RB score > 0.85 and $S/N \geq 5$ from April 5 – 30: **15,967**
- Number of alerts with RB score > 0.85 and $S/N \geq 15$ from April 5 – 30: **12,643**
~ 80% of these are associated with known asteroids
- **Recommendation: $RB > 0.85$ and $S/N \geq 15$**
~ 100 – 500 alerts per night on average for **public** stream

Improvements and progress

- Difference-image photometry sometimes shows large intra-night scatter; analysis in progress.
- An appreciable number of Pan STARRS1 associations are missing star/galaxy classification scores
 - an issue with construction of the original (raw) PS1 catalogs when merging individual epochs
- Additional desired metadata:
 - propagate uncertainty information from absolute photometric calibration
 - report location or presence of missed bright (saturated) stars within some distance of an alert

Near term:

- Refine thresholds that define the public alert stream as ML classifiers improve
 - primary focus for now is reliability
 - assess scientific usefulness of resulting stream