



# ZTF Alert Distribution System



Mario Juric, Zach Golkhou, Ben Rusholme, Eric Bellm, Maria Patterson, Frank Masci, et al.

University of Washington DIRAC Institute  
& the ZTF Partnership

# Overview: Distributing ZTF Alerts

- As ZTF observes, images are downloaded from the mountain, differenced, objects detected, measured, and associated.
- For each detection, an “alert packet” is produced, containing the current measurement, any measurements over the past 30 days, and image cutouts (FITS).
- This information is serialized in the AVRO format
- AVRO packets are posted to an Apache Kafka broker cluster at IPAC



kafka



<https://github.com/ZwickyTransientFacility/ztf-avro-alert>

<https://avro.apache.org/>

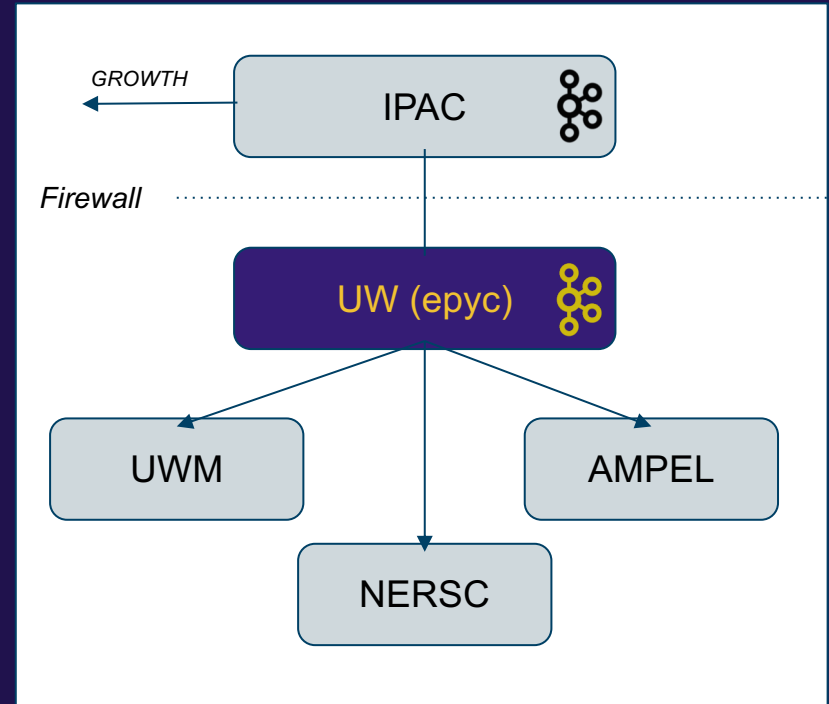
<https://kafka.apache.org/>

***The goal: reliably deliver notifications and information on newly detected events with ~20 minute latency.***

# ZTF Alert Distribution System: Partnership

For the past ~7 months, we've been operating an alert distribution service for the ZTF partnership.

The service runs at a machine at UW (epyc.astro.washington.edu), receives alerts from IPAC, and forwards them to partnership event consumers.

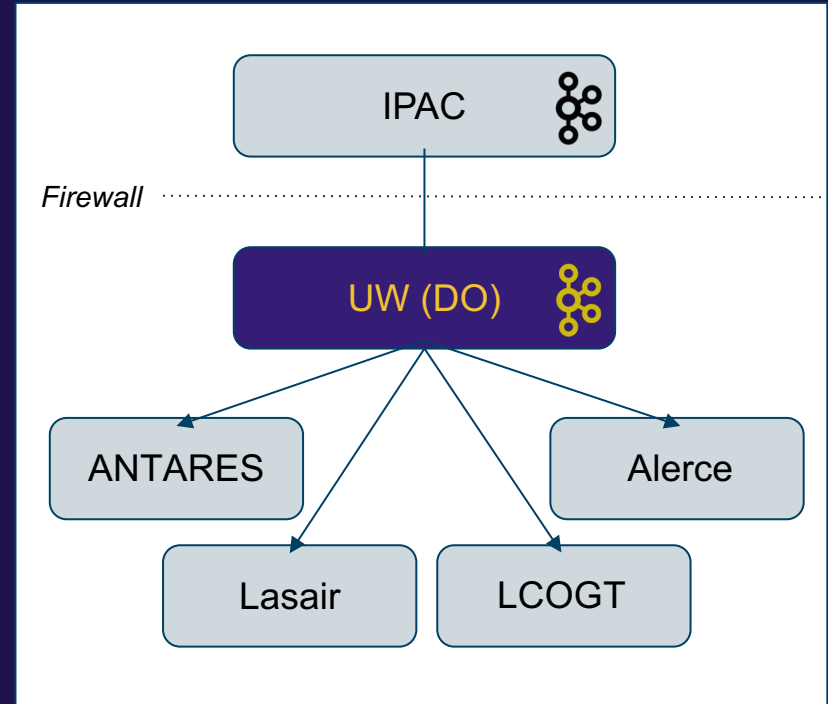


# ZTF Alert Distribution System: Public

Similarly, we run a public alert service, since June 4th.

The public service only redistributes the MSIP stream, and runs in the cloud. It's a "version 2.0" of the system, based on what we've learned by standing up the partnership service.

- Operational Goals:
  - Sufficient throughput for real-time operation: **1 MBytes/second/broker, sustained.**
  - 99% reliability: 1 full night/quarter downtime
    - Allowed for shorter night time downtimes, e.g. occasional 10-15 minute windows.



# Experience, ~7-months in

- Partnership endpoint
  - Generally stable operations (esp. the last 2-3 months)
  - The first version of kafka broker system; LSST-derived, brought over some unnecessary (for us) complexity
  - Configuration mismatches, hardware instability => duplicate messages
- Public endpoint
  - Very stable: no restart/issues since June 1<sup>st</sup> launch (\*)
  - Simpler setup, dedicated VM

*(\*) Restarted MirrorMaker once to clear up some cosmetic issues. No alerts were delayed because of it.*

# Monitoring ZADS Status (Public only)

- <http://monitor.alerts.ztf.uw.edu>
  - Password protected (see next slide)
- Development status:
  - Monitors upstream connection to IPAC
  - More metrics / dashboards to be added.

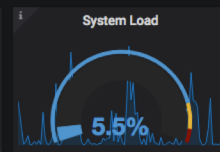
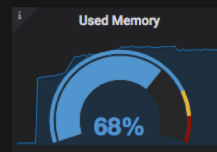
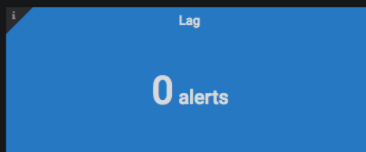
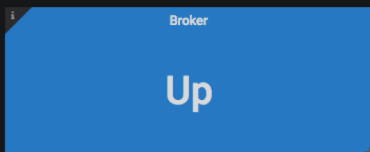
Night 20180726

# Zwicky Transient Facility Alert Distribution System Monitor (monitor.alerts.ztf.uw.edu)

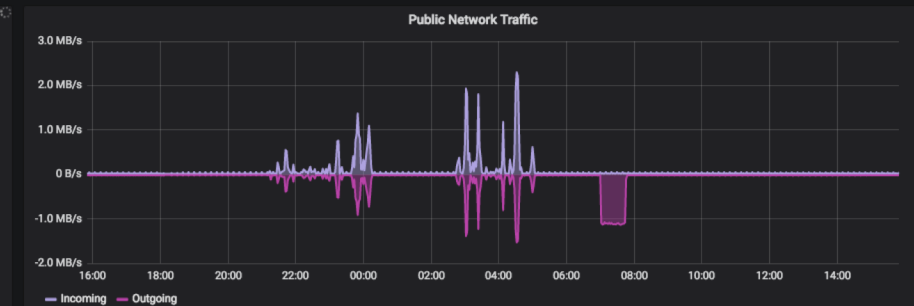
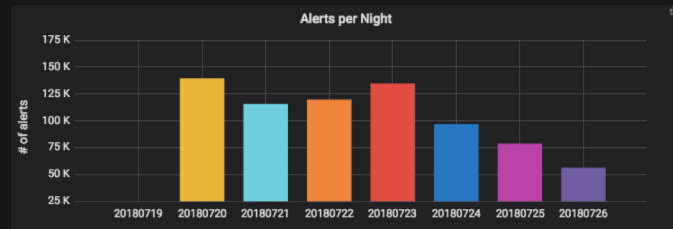
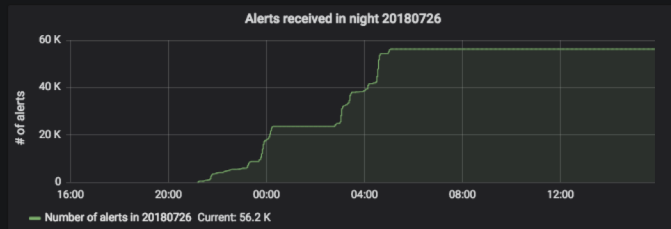
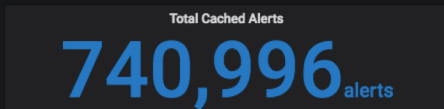
<https://monitor.alerts.ztf.uw.edu>

u/p: **ztf** / **fullofstars**

## System Health



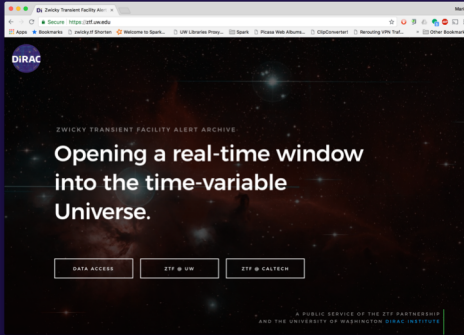
## Science Report



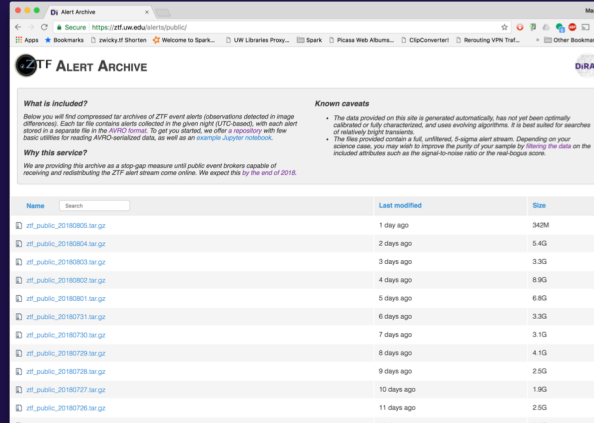


# ZTF Alerts Archive

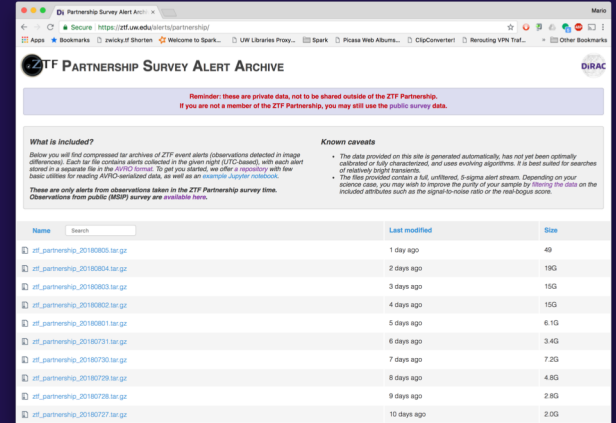
*If you don't care about real-time response, this is what you should use (or the database: see next session).*



<https://ztf.uw.edu>



<https://ztf.uw.edu/alerts/public/>



<https://ztf.uw.edu/alerts/partnership/>

u/p: [ztf](#) / [16chipsOnPalomar](#)





# LCO Interface to Public Alerts

<https://mars.lco.global>

MARS - Las Cumbres Observa x Mario

Secure <https://mars.lco.global>

Apps ★ Bookmarks 📄 zwicky.tf Shorten 🌟 Welcome to Spark... 📄 UW Libraries Proxy... 📄 Spark 📄 Picasa Web Albums... 📄 ClipConverter! 📄 Rerouting VPN Traf... 📄 Other Bookmarks

## LCO MARS

Make Alerts Really Simple Help Contact

MARS provides access to all public alerts issued by ZTF since the start of the public alert stream on June 1, 2018. Subsets of the alerts, filtered by selectable constraints, may be identified and downloaded, either through this webpage or using the underlying API. Alerts are ingested as they are generated by the ZTF survey and are made available immediately, which is reflected by the "Latest Alert" value below. Users are advised to limit their request frequency to a reasonable time period, preferably allowing at least 5 minutes between requests. In addition to our own [help](#) page, users should refer to the [ZTF website](#) and the [ZTF Alert Archive](#) for documentation on ZTF and the generation of alerts.

The following table lists ZTF alerts in descending order by JD. Use the filters on the right to narrow down the results to interesting candidates. When the results look good, add `?format=json` to the url. You can now access this url to retrieve the full data and use it in your scripts. You can access an alert's previous alerts by visiting `/<id>/` where id is the value of the lco\_id key in the json view or by clicking the id link in the table.

See the [help](#) page for descriptions of the table values and available filters.

Select Reset

Prev Next Results: 5841854 Pages: 58419

Latest Alert: 2018-08-04 07:20:49 UTC

Sort By	id	objectId	time	filter	ra	dec	magpsf	magap	distnr	Δmaglatest	Δmagref	rb
time	<a href="#">6121378</a>	ZTF18ablonfj	2018-08-04 07:20:49	g	269.91461	66.96017	18.26	18.76	6.367			0.290
Sort Order	<a href="#">6121764</a>	ZTF18abloniq	2018-08-04 07:20:49	g	265.89783	66.70604	18.78	18.98	5.880			0.290
objectId	<a href="#">6121943</a>	ZTF18ablonlr	2018-08-04 07:20:49	g	257.63087	68.18566	18.12	18.56	11.711			0.270
candid	<a href="#">6121964</a>	ZTF18aakvzfc	2018-08-04 07:20:49	g	256.65589	67.64165	18.00	18.26	0.588	0.12	-2.66	0.497
time (lower)	<a href="#">6121989</a>	ZTF18aaoechd	2018-08-04 07:20:49	g	273.53155	70.73565	16.75	16.84	0.452	-0.06	-1.31	0.560
time (upper)	<a href="#">6121998</a>	ZTF18aamyhuq	2018-08-04 07:20:49	g	257.13025	68.88617	16.77	16.88	0.596	0.18	-2.96	0.330
jd	<a href="#">6122003</a>	ZTF18aapsnka	2018-08-04 07:20:49	g	257.34833	67.93375	19.63	19.76	0.535	0.20	-2.47	0.690
	<a href="#">6122022</a>	ZTF18ablonlq	2018-08-04 07:20:49	g	271.74009	69.55966	18.34	18.76	4.095			0.353



# Connecting and Consuming ZTF Alerts (0/3)

- Who can connect?
- Anyone with a clear need who talks to us and promises not to blow up the system (no permissions implemented yet!)
- Marshal builders, broker builders, etc..

# Connecting and Consuming ZTF Alerts (1/3)

- Broker mirrors and keeps the last 7 days\* of public alerts. Downstream users should make sure to download the data before it expires.
- Each night is a separate Kafka topic: `ztf_20180725_programid1`
- Bootstrap server: `epyc.astro.washington.edu:9092`  
require Kafka clients v1.0+
- Authentication: IP Based (note: pwd auth TBD)
- Consumer group ID: `<PROJECT> [ -<foo> ]`  
example: `LCOGT, LCOGT-test`

# Connecting and Consuming ZTF Alerts (2/3)

- A graduated checklist is at <https://zwicky.tf/g2o>
  - This links to <https://github.com/dirac-institute/zads-terraform/blob/master/documentation/connection-checklist.md>
- Parsing the alerts:
  - Each alert is transmitted in *AVRO Object Container File* (<https://zwicky.tf/2aa>) format. Any AVRO library should be able to deserialize it. The ZTF schema is described at <https://zwicky.tf/4t5>.
  - We provide an example Python reader at [https://github.com/ZwickyTransientFacility/alert\\_stream](https://github.com/ZwickyTransientFacility/alert_stream). You can use this as a basis to develop your clients.

# Software

- Alerts are stored in AVRO format – how do you read them?
- Connecting to the real-time stream:
  - [https://github.com/ZwickyTransientFacility/alert\\_stream/](https://github.com/ZwickyTransientFacility/alert_stream/)
- Reading from files:
  - [https://github.com/zgolkhou/ztf\\_public\\_alerts](https://github.com/zgolkhou/ztf_public_alerts)
- Upcoming:
  - A new, conda-installable, client to make all this much easier (“st”)

```

10. root@ksql:~/streamtool (ssh)
[root@ksql streamtool]# ./st --help
usage: st [-h] [--from {json,avro}] [--to {json,avro}] [--schema SCHEMA]
         [--save-schema SAVE_SCHEMA] [--count COUNT] [--progress]
         sources [sources ...] dest

Consume AVRO or JSON-formatted messages from a source, and produce them to a
destination.

positional arguments:
  sources              Data source.
  dest                Data destination

optional arguments:
  -h, --help          show this help message and exit
  --from {json,avro}  Input format
  --to {json,avro}    Input format
  --schema SCHEMA     Output schema
  --save-schema SAVE_SCHEMA
                     Save input schema into a file
  --count COUNT       Number of messages to consume
  --progress          Show progress indicator
[root@ksql streamtool]# ./st --progress "kafka://epyc.astro.washington.edu/^ztf_20180806_programid.$" mydata.json --to json --count 160
Generated fake groupid root-DMWY969SH9W3ERD2U548 and commit=False
0000000000000000000000000000000000000000000000000000000000000000 [      80 @ Aug 06 2018 07:30:41 ]
0000000000000000000000000000000000000000000000000000000000000000 [     160 @ Aug 06 2018 07:30:41 ]

Sources read:
  kafka://epyc.astro.washington.edu/ztf_20180806_programid1
  kafka://epyc.astro.washington.edu/ztf_20180806_programid2
[root@ksql streamtool]# █

```

10. root@ksql:~/streamtool (ssh)

```
[root@ksql streamtool]# head mydata.json | jq -r ' ' | head -n 28
{
  "schemavsn": "3.0",
  "publisher": "ZTF (www.ztf.caltech.edu)",
  "objectId": "ZTF18ablqdor",
  "candid": 582172701815015000,
  "candidate": {
    "jd": 2458336.6727083,
    "fid": 1,
    "pid": 582172701815,
    "diffmaglim": 20.38648796081543,
    "pdiffimfilename": "ztf_20180806172708_000477_zg_c05_o_q3_scimrefdiffimg.fits",
    "programi": "Kulkarni",
    "programid": 1,
    "candid": 582172701815015000,
    "isdiffpos": "t",
    "tblid": 3,
    "nid": 582,
    "rcid": 18,
    "field": 477,
    "xpos": 1634.05615234375,
    "ypos": 1735.678955078125,
    "ra": 215.0934073,
    "dec": 3.1143465,
    "magpsf": 17.11831283569336,
    "sigmapsf": 0.07662621885538101,
    "chipsf": 40.73303985595703,
    "magap": 17.052600860595703,
    "sigmagap": 0.027899999171495438,
  }
}
[root@ksql streamtool]# █
```

# Communications & Team

- Mailing list:

- [ztf-broker-ops@uw.edu](mailto:ztf-broker-ops@uw.edu)

- Slack:

- <https://ztf-broker-ops.slack.com/>

- Alert Archive

- <https://ztf.uw.edu>

***I'd like the partnership consumers to join these mailing lists, slack, etc! Info on how will be added to the wiki, but for now please talk to me to add you!***

- Team

- Mario Juric, Zach Golkhou, Eric Bellm, Maria Patterson (UW)
- Ben Rusholme, Frank Masci (IPAC)

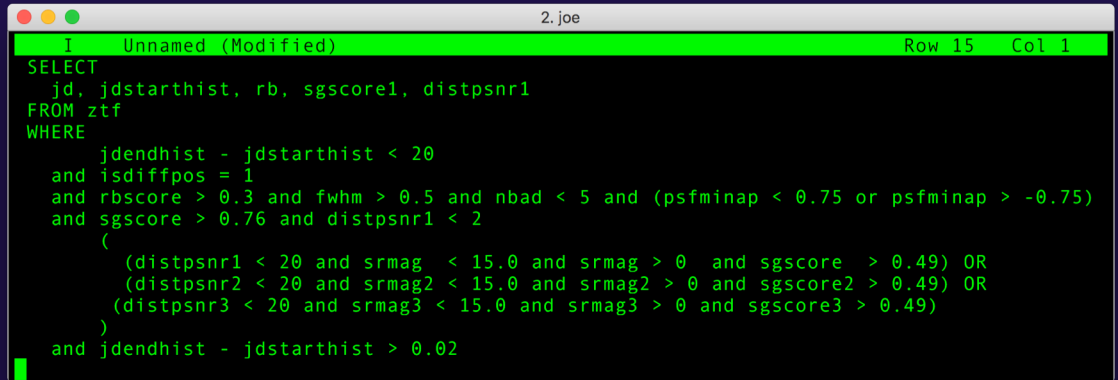
- Currently connected public brokers

- ANTARES (NOAO)
- ALERCE (CMM-Chile)
- LASAIR (Edinburgh/Belfast)
- MARS (LCO)



# Experiments: Kafka SQL Server

- The ability to write alert filters in SQL – select only the alerts of interest to your science case.
- Consumes the alerts from one stream, filters it (potentially joins with a table), produces to another stream
- Scalable, robust
- Of interest to nex-gen marshals
- Of interest for algorithm development



```
2. joe
I Unnamed (Modified) Row 15 Col 1
SELECT
  jd, jdstarthist, rb, sgscore1, distpsnr1
FROM ztf
WHERE
  jdendhist - jdstarthist < 20
  and isdiffpos = 1
  and rbscore > 0.3 and fwhm > 0.5 and nbad < 5 and (psfminap < 0.75 or psfminap > -0.75)
  and sgscore > 0.76 and distpsnr1 < 2
  (
    (distpsnr1 < 20 and srmag < 15.0 and srmag > 0 and sgscore > 0.49) OR
    (distpsnr2 < 20 and srmag2 < 15.0 and srmag2 > 0 and sgscore2 > 0.49) OR
    (distpsnr3 < 20 and srmag3 < 15.0 and srmag3 > 0 and sgscore3 > 0.49)
  )
  and jdendhist - jdstarthist > 0.02
```

*Work by Marwan Shehri (summer project @ UW, UMd student)*

# (Technical) Work ahead:

- Move the partnership endpoint to "ZADS 2.0"
- Proper authentication and authorization (required to broadly open it up)
- Monitoring for the partnership endpoint
  
- Establish clear procedures, responsibilities, and lines of communication (e.g., schema changes, exceptional situations, data transfer issues, cybersecurity issues, etc.)
  
- Consider on-the-wire format change to reduce packet sizes
- Documentation, documentation, documentation!

Keeping a list of prioritized issues in:  
<https://github.com/dirac-institute/zads-terraform>

# (Science) Discussions Ahead

- Overall:
  - What is the feedback on ZADS? We'd like to hear more from our downstream partnership users.
- Filtering:
  - There's an ongoing debate of what gets alerted on – only "high quality" alerts, or all the alerts.
  - Current cuts make some science difficult to do:
    - E.g., comets are lost
    - No streaked objects in the alert stream
  - Proposal: let's try to alert on *everything*, and we'll filter it at the broker into multiple streams of different quality.
- Topic names/usage and late processing
  - Some alerts are missed they're processed late, due to how we've configured topic names & mirroring policies.



# Discussion

---

Contact: [mjuric@astro.washington.edu](mailto:mjuric@astro.washington.edu)  
*@mjuric on ZTF Slack*