Lasair: The transient alert broker for LSST:UK

K. W. Smith[1], R. D. Williams[2], D. R. Young[1], A. Ibsen[2], S. J. Smartt[1], A. Lawrence[2], D. Morris[2], S. Voutsinas[2], M. Nicholl[2]

[1]Astrophysics Research Centre, Queen's University Belfast, [2]Institute for Astronomy, University of Edinburgh, Royal Observatory

https://lasair.roe.ac.uk

When LSST begins surveying the sky, it is expected to produce a stream of up to 10^7 transients per night (LSST Science Collaboration et al. 2009). This time domain stream of data will need to be rapidly mined, with digestible alerts produced in a timely manner. Since early 2018, the Zwicky Transient Factory (ZTF) has been producing alerts at a rate of up to 10^6 per night (ZTF; Kulkarni 2018; Bellm et al. 2019). The alerts are publicly available to any brokers that are willing to consume and distribute them. We present Lasair[1] - the LSST:UK Transient Broker (Smith, K. W., Williams, R. D. et al 2019), a prototype interface to a relational database of explosive and eruptive transients based on the ZTF transient stream.

Introduction

ZTF emit a relatively clean source of transient data, with detections already assimilated into objects. Lasair consumes these via Kafka / Avro; a messaging protocol adopted by LSST[2] and developed by social media organisation LinkedIn (Kreps J et. al. 2011).

QUB has extensive experience in building large databases of difference detections from the Pan-STARRS and ATLAS surveys. In these surveys up to 50 million detections are emitted in any single night, and these need to be assimilated into objects and stored in a relational database. The vast majority of detections from these surveys is junk (e.g. cosmic rays, chip defects, crosstalk, etc), and implementing simple cuts (e.g. requesting that a minimum number of detections be spatially coincident) vastly reduces the number of subsequent post-ingest step that need to be performed. However, the large number of initial detections is still inserted into the database, and gives confidence that usage of relational databases even in the LSST era may still be a viable, scalable option.

Lasair has been constructed as an evolution of previous transient servers, but alternative "big data" technologies are also being actively explored.

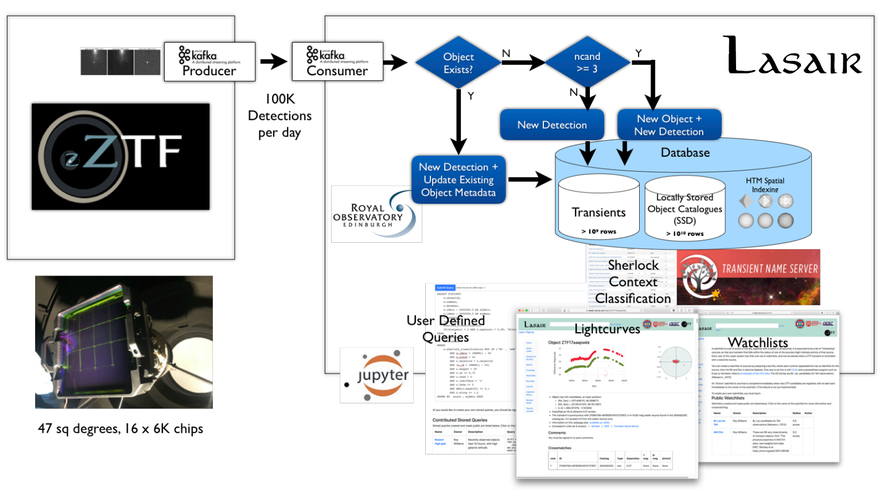

Figure 1. Lasair prototype with ZTF.

The Lasair-ZTF prototype

The Lasair prototype (Figure 1) has been running on relatively lightweight virtual machines at the Royal Observatory since the end of May 2018. On average, ZTF emits about 100,000 detections per night. ZTF alerts are transmitted typically within 13 minutes of the exposure and ingested into the Lasair database within 20 minutes.

Detections are consumed via Kafka and assimilated into objects (by name), allowing lightcurves to be easily rendered. Objects with three or more spatially coincident detections (Figure 2) are spatially crossmatched with multi-billion row locally curated catalogues. The catalogues (and the ingested ZTF detections) are spatially indexed using Hierarchical Triangular Mesh (Szalay A. S. et al. 2005).

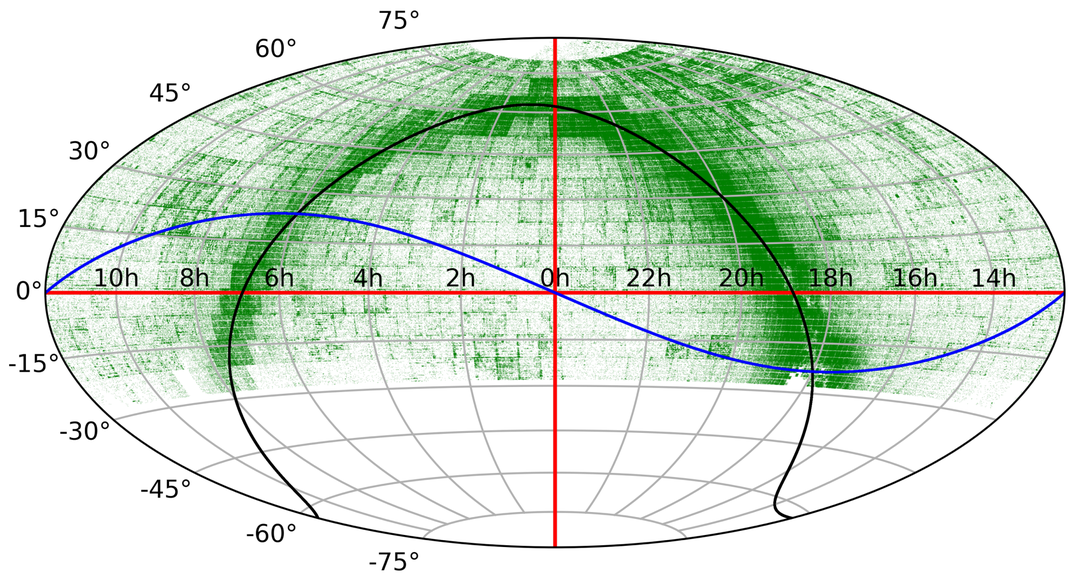

Figure 2. All (approximately 2 million) ZTF objects ingested into Lasair with 3 or more detections. The vast majority of detections are variable stars, mostly clustered around the galactic plane (black line).

Context Annotation: Sherlock

The context classifier designed to mine these catalogues, Sherlock, was developed for PESSTO, Pan-STARRS and ATLAS[3]. Although the code for Sherlock is written in python, the search algorithm is abstracted into a plain text, human readable YAML file. This facilitates exchangeable algorithms that can be adapted by anyone without having to rewrite any code. Sherlock uses star/galaxy separation, distances, and galaxy offsets in order to classify objects as likely supernovae, nuclear transients, or variable stars. It also matches against known AGN mostly with the Veron (Véron-Cetty & Véron 2010) and Milliquas (Flesch 2015) catalogues and with CV catalogues (Downes et al. 2001; Ritter & Kolb 2003). Any stationary, transient source which is not associated with a catalogued star or galaxy is classified as an "orphan." A ranked, user queryable table of nearby objects is stored for the transient (Figure 3) along with a summary statement (Figure 4(d)).

Web Interface

Lasair provides a user-friendly web interface to access public ZTF transient alerts. Each transient object has its own page, the details of which are shown in Figure 4.

Cone Searching and Watchlists

A search box enables simple cone searching (in decimal or sexagesimal or by object name) and registered users can also create watchlists of objects (e.g. a subset of galaxies) and be alerted if a transient is discovered near one of them.

Queries and Filters

A query builder is provided that allows a user to create custom queries on the database. Additionally, these queries can be turned into filters, which are run on the data as it is ingested (Figure 5). Example filtered streams of objects are provided by Lasair, and more complex queries can be performed via Jupyter notebooks to registered users.

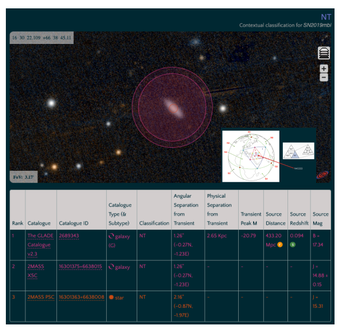

Figure 3. Sherlock context determination for the same object as shown in Figure 4. The large catalogues are indexed via the Hierarchical Triangular Mesh spatial indexing system (inset).

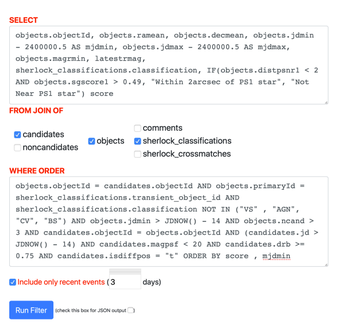

Figure 5. The Lasair query builder page. Users can enter a query to be run against a subset of database tables. The output can be delivered in JSON if necessary and the query itself can be stored and run as a user defined transient filter as data is ingested.

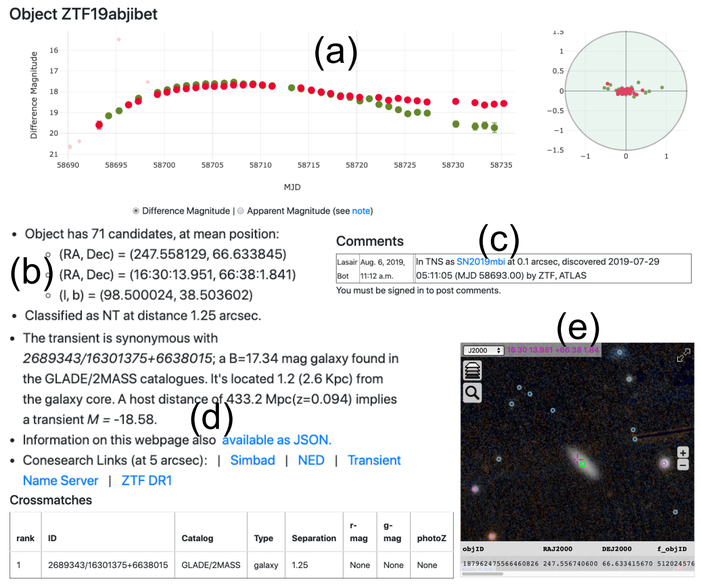

Figure 4. Anatomy of the Lasair page for the object ZTF19abjibet. (a) Light curve and scatter plot (units of arcseconds). (b) Coordinate information. (c) Crossmatch with the Transient Name Server[4] and user added comments. (d) Contextual information from Sherlock describing which object the transient is likely to be associated with, and what kind of transient (in this case NT = nuclear transient). (e) AladinLite (Bonnarel et al. 2000; Boch & Fernique 2014) view showing Pan-STARRS imagery and overlaid PS1 and Gaia catalogues (blue and pink respectively). The green box is the highlighted Pan-STARRS object whose table entry is shown at the foot of the image. (f) ZTF photometric information and image cutouts for each candidate comprising this object.

GW Events

As previously done with Pan-STARRS and ATLAS transients, Lasair will be able to trigger alerts for ZTF transients within the high probability localisation regions of GW event skymaps (Figure 6). It is also possible to superimpose ZTF events that are temporally and spatially coincident.

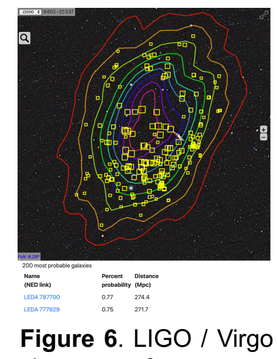

Figure 6. LIGO / Virgo skymap for event S190814bv. The yellow boxes represent possible galaxies that may be candidate hosts for a possible EM counterpart.

Lasair Prototype Status and Future

Lasair is being actively used to discover interesting transient events (Figure 7). It will will continue to evolve and serve as a broker for ZTF data while the LSST broker is developed in parallel.

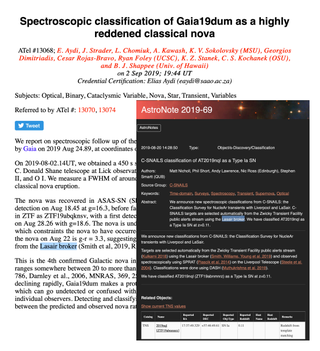

Figure 7. A recent ATel and AstroNote referencing Lasair as the discovery engine.

Scaling up Lasair to LSST

LSST will emit up to 10 million detections per night; 100 times more data, on average, than is generated by ZTF. The LSST UK Data Access Centre has access to the IRIS peta scale infrastructure[5] being developed by STFC upon which the current version of Lasair-ZTF will be deployed. The current relational database technology may also present issues of scalability and latency when operating at the scale of LSST. Although these may be tractable, other technologies are being explored for both crossmatching and data curation (e.g. Astronomy Extensions for Spark[6], Cassandra[7]). A schematic representing proposed LSST functionality is shown in Figure 8.

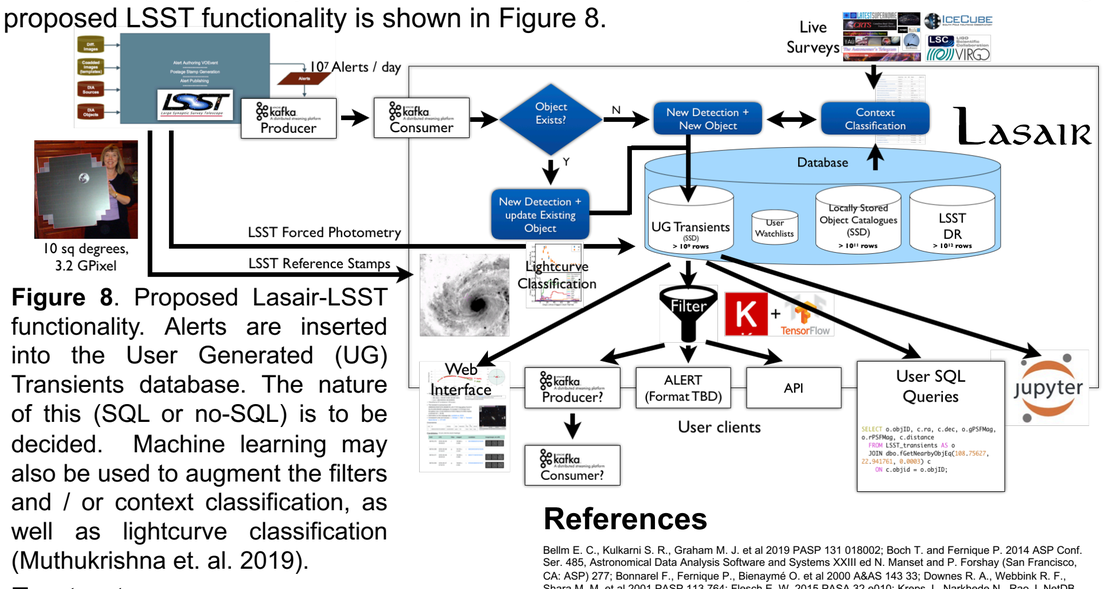

Figure 8. Proposed Lasair-LSST functionality. Alerts are inserted into the User Generated (UG) Transients database. The nature of this (SQL or no-SQL) is to be decided. Machine learning may also be used to augment the filters and / or context classification, as well as lightcurve classification (Muthukrishna et. al. 2019).

Footnotes

[1]The word "lasair" means flame or flash in Scots and Irish gaelic. [2]https://ldm-612.lsst.io/LDM-612.pdf [DRAFT]. [3]https://github.com/thespacedoctor/sherlock [4]https://wis-tns.weizmann.ac.il/ [5]https://www.iris.ac.uk/ [6]https://github.com/dirac-institute/AXS. [7]http://cassandra.apache.org

References

Bellm E. C., Kulkarni S. R., Graham M. J. et al 2019 PASP 131 018002; Boch T. and Fernique P. 2014 ASP Conf. Ser. 485, Astronomical Data Analysis Software and Systems XXIII ed N. Manset and P. Forshay (San Francisco, CA: ASP) 277; Bonnarel F., Fernique P., Bienaymé O. et al 2000 A&AS 143 33; Downes R. A., Webbink R. F., Shara M. M. et al 2001 PASP 113 764; Flesch E. W. 2015 PASA 32 e010; Kreps J., Narkhede N., Rao J. NetDB workshop 2011 https://www.microsoft.com/en-us/research/wp-content/uploads/2017/09/Kafka.pdf; Kulkarni S. 2018 ATel 11266; LSST Science Collaboration et al. 2009 arXiv:0912.0201; Muthukrishna, D., Narayan, G., Mandel, K. S.,Biswas, R., & Hlo˙zek, R. 2019, arXiv e-prints,arXiv:1904.00014; Patterson M. T., Bellm E. C., Rusholme B. et al 2019 PASP 131 018001; Ritter H. and Kolb U. 2003 A&A 404 301; Smith K. W, Williams R. D. et. al. 2019 RNAAS 3 26; Szalay A. S., Gray J., Fekete G. et al. 2005, Tech. Rep. MSR-TR-2005-123; Véron-Cetty M.-P. and Véron P. 2010 A&A 518 A10