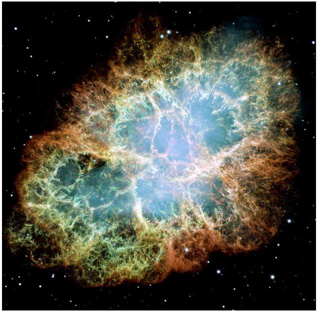
CAASTRO
ARC CENTRE OF EXCELLENCE
FOR ALL-SKY ASTROPHYSICS

Astroinformatics challenges for next generation radio transients surveys

Tara Murphy
University of Sydney

11th December 2013

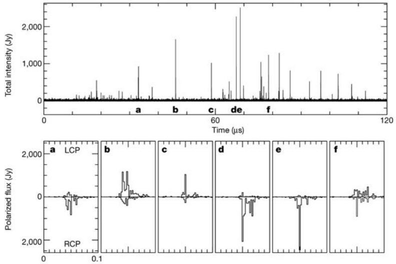
Transients capture our interest



http://www.hao.ucar.edu/education/archeoslides/slide_20.php

Transient behaviour is extreme

- An object can not change its brightness in an interval shorter than the time light takes to cross its diameter



- Giant pulses discovered in Crab pulsar
  Hankins et al. 2003, Nature, 422, 141
- ~ 2 ns structure in pulses $\implies d = c\Delta t = 1$ m!

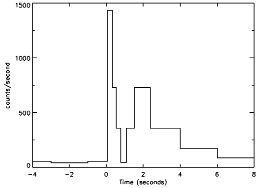
CAASTRO
ARC CENTRE OF EXCELLENCE FOR ALL-SKY ASTROPHYSICS

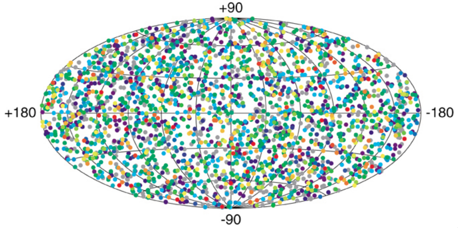
What causes transient behaviour?

1. **Explosions**
   - e.g. supernovae, Gamma-Ray bursts, orphan afterglows
2. **Propagation**
   - e.g. Extreme Scattering Events, intra-day variables
3. **Accretion**
   - e.g. neutron stars, black holes, quasars, X-ray binaries
4. **Magnetospheric**
   - e.g. magnetars, flare stars, planetary variability
5. **Unknown**
   - e.g. known unknowns, unknown unknowns. . .

Mysterious bursts from space

- Gamma-Ray Bursts were first detected by the US military Vela satellites
- Their goal was to monitor nuclear testing from space



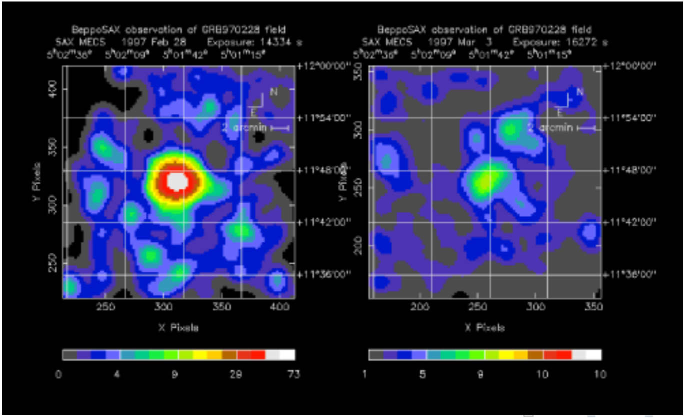


- BATSE detected 2704 GRBs in the 1990s
- Isotropic distribution implied extragalactic origin

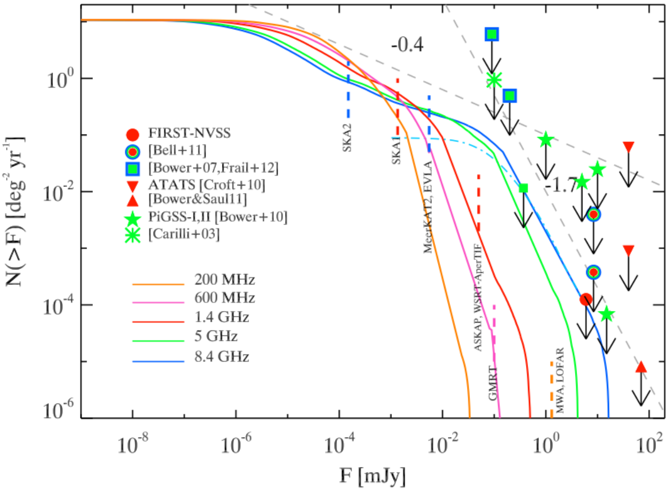
http://apod.nasa.gov/htmltest/jbonnell/www/grbhist.html

GRB afterglows — the missing link

- GRB970228 in February 1997 was detected by BeppoSax
- It was subsequently detected by Hubble



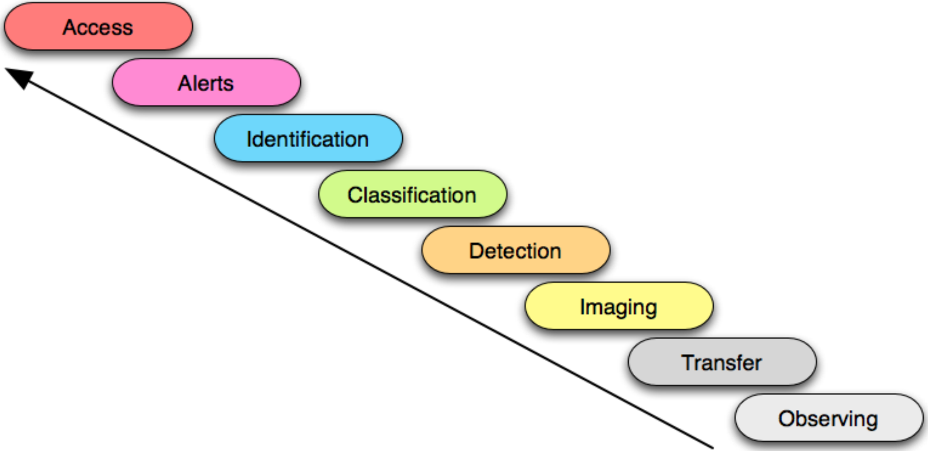
The search for orphan afterglows



Ghirlanda et al. 2014, submitted

The radio transients stack

The radio transients stack

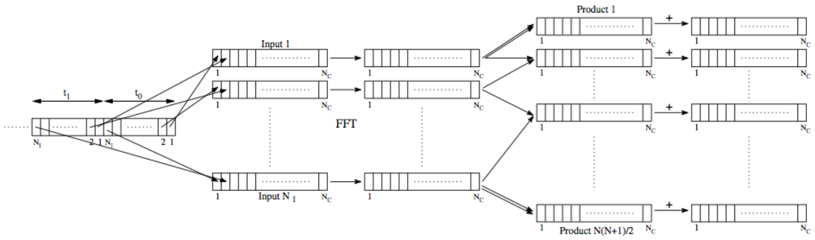


CAASTRO
ARC CENTRE OF EXCELLENCE FOR ALL-SKY ASTROPHYSICS

Step 1: Telescope

- **The challenge:** Schedule telescope including potential for automatic scheduling and triggers. Do rapid correlation and processing of raw data.
- **The technology:** Combination of hardware and software correlators. Custom scheduling software.
- **Where are we?** Innovation at the intersection of astronomy, computer science and engineering. FPGAs and GPUs new on the scene: limited but growing expertise in the astronomy community.
- **VO?** Probably not relevant for low level/early data products.

Step 1: Telescope

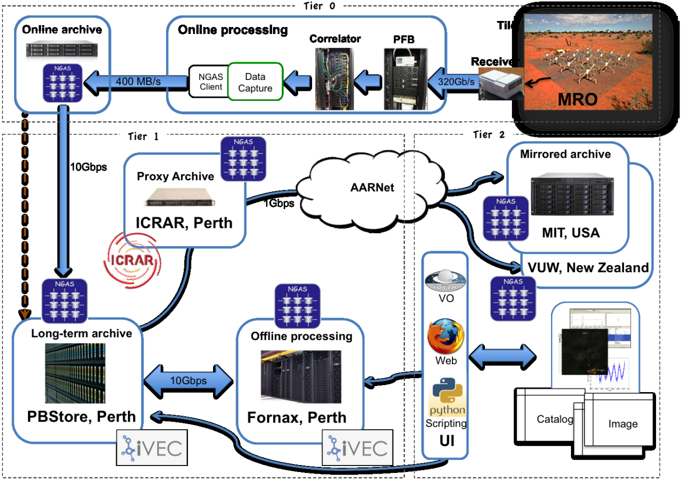


Wayth et al. 2009, PASP

Step 2: Data transfer and storage

- **The challenge:** Transfer large amounts (terabytes, petabytes) from the telescope to a data processing centre. Store this is a way that can be accessed by others.
- **The technology:** Large scale databases, high bandwidth data links, long term storage tapes.
- **Where are we?** Mostly in place and operational. Science teams transfer data to their own resources for further processing.
- **VO?** Plan for ASKAP Science Data Archive will include VO access using standard protocols.

Step 2: Data transfer and storage

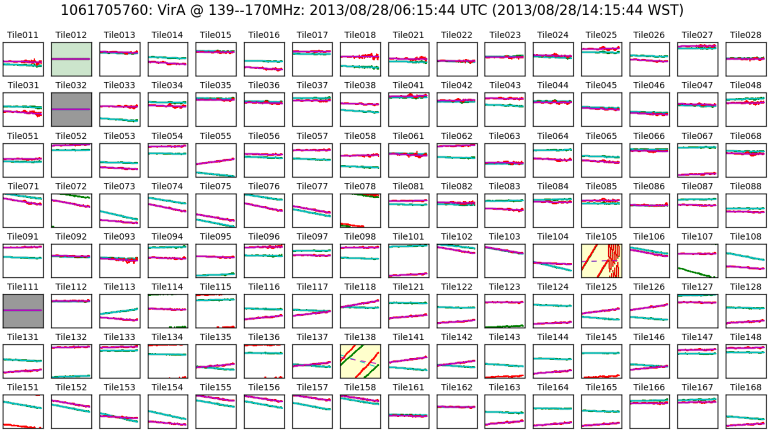


http://www.mwatelescope.org

Step 3: Calibration and imaging

- **The challenge:** Convert large amounts of raw data into science-ready data products such as images or spectral cubes.
- **The technology:** Community developed data reduction packages such as Miriad, AIPS, CASA. New custom packages such as ASKAPSoft, MWA-RTS.
- **Where are we?** Existing packages don't scale to supercomputing levels. Work on adapting these is ongoing.
- **VO?** Largely disconnected from web or other software packages.

Step 3: Calibration and imaging



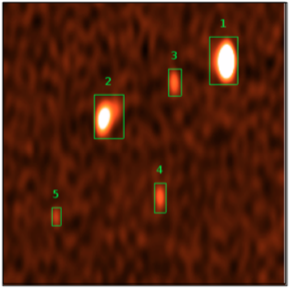
Credit: Natasha Hurley-Walker

Step 4: Source detection

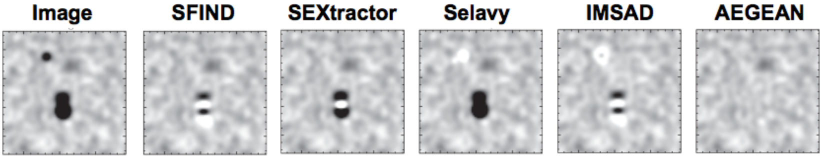
- **The challenge:** Find all astronomical objects in an image or data cube. Measure the properties of all objects.
- **The technology:** Generally a custom program written in Python or C. Store results in a database for rapid(ish) access.
- **Where are we?** Multiple packages exist. All solve different but related problems. Different levels of integration.
- **VO?** Not really. . .

Step 4: Source detection

- Identification of interesting events will need to be catalogue-based, not image-based
  - missed/blended sources will trigger huge numbers of false alarms
  - 99% accuracy is not good enough!
- BLOBCAT (Hales, Gaensler et al. 2012)
  - flood-fill: superior to gaussian fitting
- AEGEAN (Hancock, Gaensler et al. 2012)
  - Laplacian: robust component separation



Hales et al. (2012)



Hancock et al. (2012)

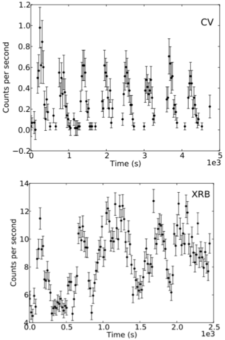
CAASTRO
ARC CENTRE OF EXCELLENCE FOR ALL-SKY ASTROPHYSICS

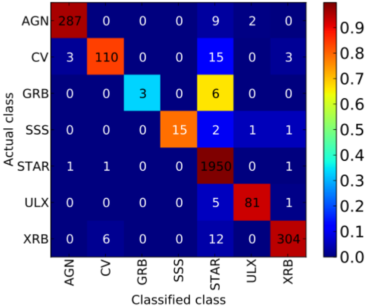
Step 5: Light curve classification

- **The challenge:** Classify light curves (flux vs. time) to determine what kind of transient or variable behaviour we are seeing.
- **The technology:** Wide range of off-the-shelf machine learning packages in a variety of languages.
- **Where are we?** Off-the-shelf solutions are OK. Not well integrated with data analysis pipelines. Limits to scalability.
- **VO?** Starting. . . (DAME?)

Step 5: Light curve classification

- Supervised learning methods (e.g. random forest)
- Require automatic calculation/extraction of features
- Successful use in optical, X-ray



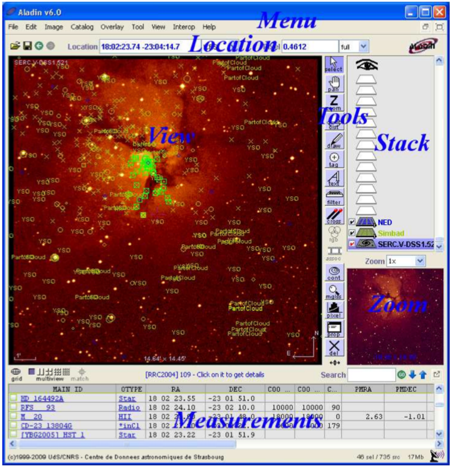


Lo et al. 2013, ApJ, submitted

Step 6: Transient identification

- **The challenge:** Identify object by cross-matching with existing surveys and archival images, spectra and other data, or searching literature.
- **The technology:** Databases that store this information. VO protocols and other methods for querying it automatically. Online archives (NED, SIMBAD, CDS, Vizier).
- **Where are we?** The databases are well established and many have the capacity to automatically query. Access reliability is an issue — 'best' solution is still to locally download.
- **VO?** Partial integration. Storing data locally is still easier.

Step 6: Transient identification



http://aladin.u-strasbg.fr

Step 7: Publication and alerts

- **The challenge:** Alert the wider transients community as soon as possible after the event is detected.
- **The technology:** Ranges from 'old school' Astronomers Telegrams to protocols such as VO Event to social media such as Twitter.
- **Where are we?** A range of mechanisms exist. Most VOEvents generated by a few surveys. Will become more interesting as more produce data.
- **VO?** Well established connectivity with VOEvent XML as the underlying standard.

Step 7: Publication and alerts

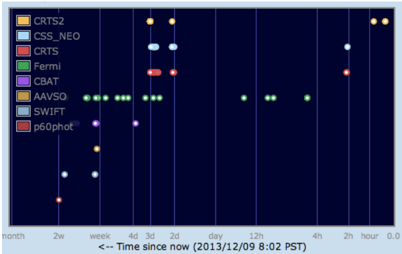
```
TITLE:   GCN CIRCULAR
NUMBER:  13180
SUBJECT: GRB120327A - ATCA radio afterglow detection at 34GHz
DATE:    12/03/31 21:01:15 GMT
FROM:    Paul Hancock at U of Sydney  <hancock@physics.usyd.edu.au>

P. Hancock, T. Murphy, B. Gaensler, M. Bell, D. Burlon (University of
Sydney/CAASTRO), A. de Ugarte Postigo (Dark Cosmology / IAA)

We observed GRB120327A (GCN13123) with the Australia Telescope Compact
Array for 40 mins centered on 17:55UT on March 31 2012 (T0+4.625days).

We find an unresolved radio source at ra=16:27:27.4, dec=-29:24:54.0
with a flux of 0.72+/-0.03 mJy. This position is consistent with the
optical position of GCN13138.

Further observations are planned.
```



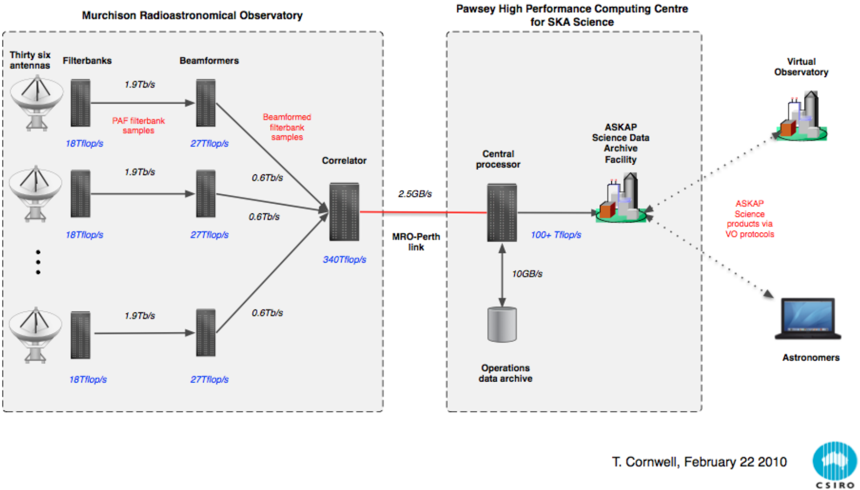
http://gcn.gsfc.nasa.gov http://skyalert.org

CAASTRO ARC CENTRE OF EXCELLENCE FOR ALL-SKY ASTROPHYSICS

Step 8: Data and archival access

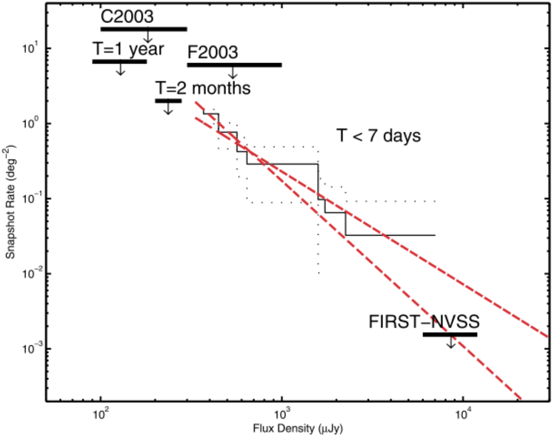
- **The challenge:** Make large and information rich scientific data products available to the collaboration and wider community.
- **The technology:** Large scale databases, VO protocols, web 3.0 interfaces.
- **Where are we?** MWA data archive is currently 'do it yourself'. ASKAP data archive is in the planning stages.
- **VO?** Should be possible to utilise existing VO protocols.

Step 8: Data and archival access



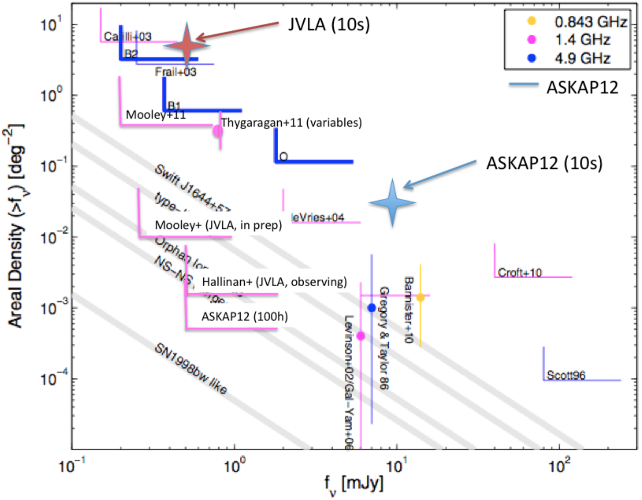
ASKAP Science Processing Document

Transient snapshot rates (c. 2007)



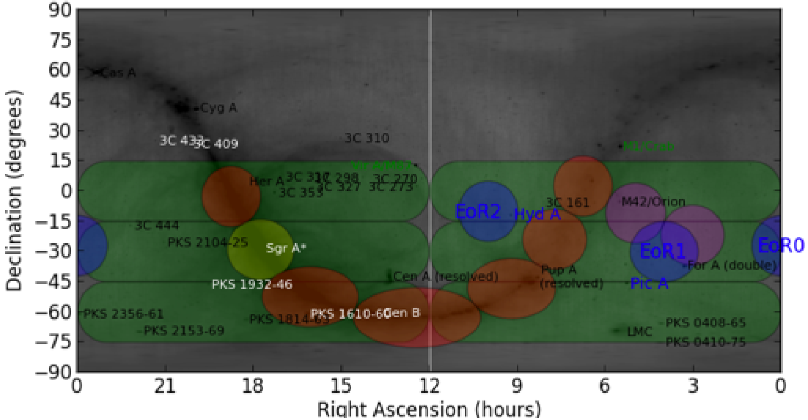
Bower et al. 2007, ApJ, 666, 346

Transient snapshot rates (c. 2013)



Keith Bannister, adapted from Frail et al. 2012 , ApJ, 747, 70

The 'next generation' is here



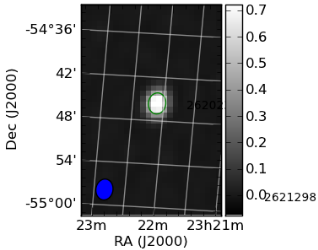
Credit: David Kaplan
(See movie from Martin Bell)
(See Galactic Plane image from David Kaplan)

Transient detection pipeline

Source 200201 MWA_232206-544535

RA 23:22:06.69 Dec -54:45:37.07 search SIMBAD NED
Cross-match this source with the imported survey catalogues. View position plot.

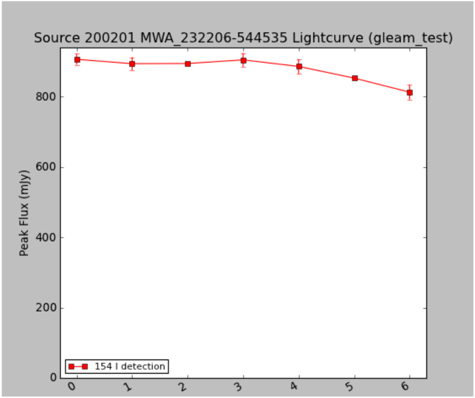
Quality source: None [set to True | False | Remove]



First Detection

Bands: 154

Polarizations: I



Challenges for transient detection

1. Don't want to average your data
2. Want to analyse in (close-to) real time
3. RFI is similar to what you're looking for
4. Confirmation is harder if object is transient
5. Need to mobilise resources quickly

Challenges for transient detection

1. Don't want to average your data
2. Want to analyse in (close-to) real time
3. RFI is similar to what you're looking for
4. Confirmation is harder if object is transient
5. Need to mobilise resources quickly

And an observation...
After 32 years, FITS is still the glue!