

The VLA Sky Survey and multimessenger astronomy



Mark Lacy & the
VLASS team

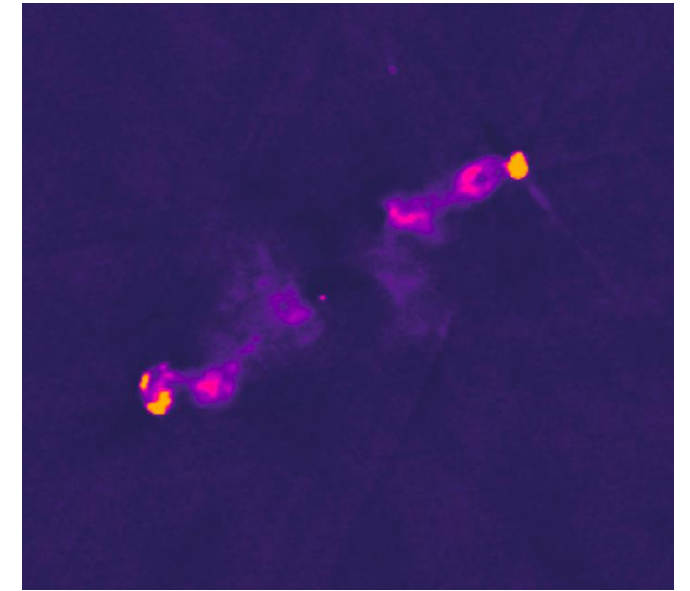


National Radio
Astronomy
Observatory

Overview

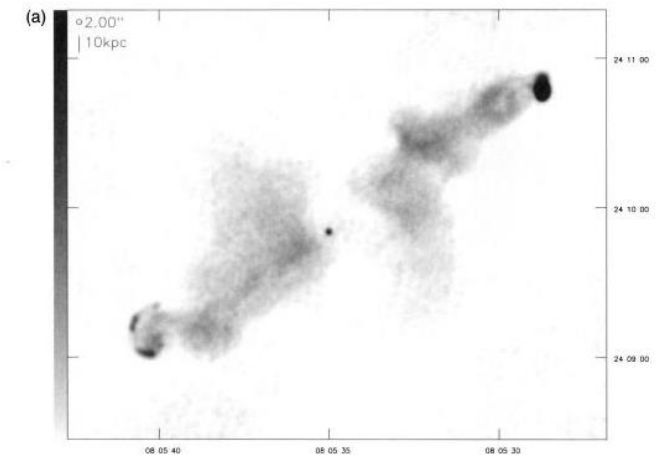
- Aims of VLASS:
 - The Very Large Array, built in the 1970s, got a major upgrade in ~2010, allowing much larger bandwidths and thus continuum sensitivity.
 - VLASS uses these enhanced capabilities to produce a survey of the whole sky visible to the telescope in the radio (2-4 GHz) that is close (2-3 arcsec) to the resolution of optical surveys.
 - Include multiple epochs, a wide bandwidth and polarization information.
 - Deliver images of the radio sky for use by non-expert radio astronomers.

3C192
VLASS
~5s
(Tasan
Smith-
Gandy)



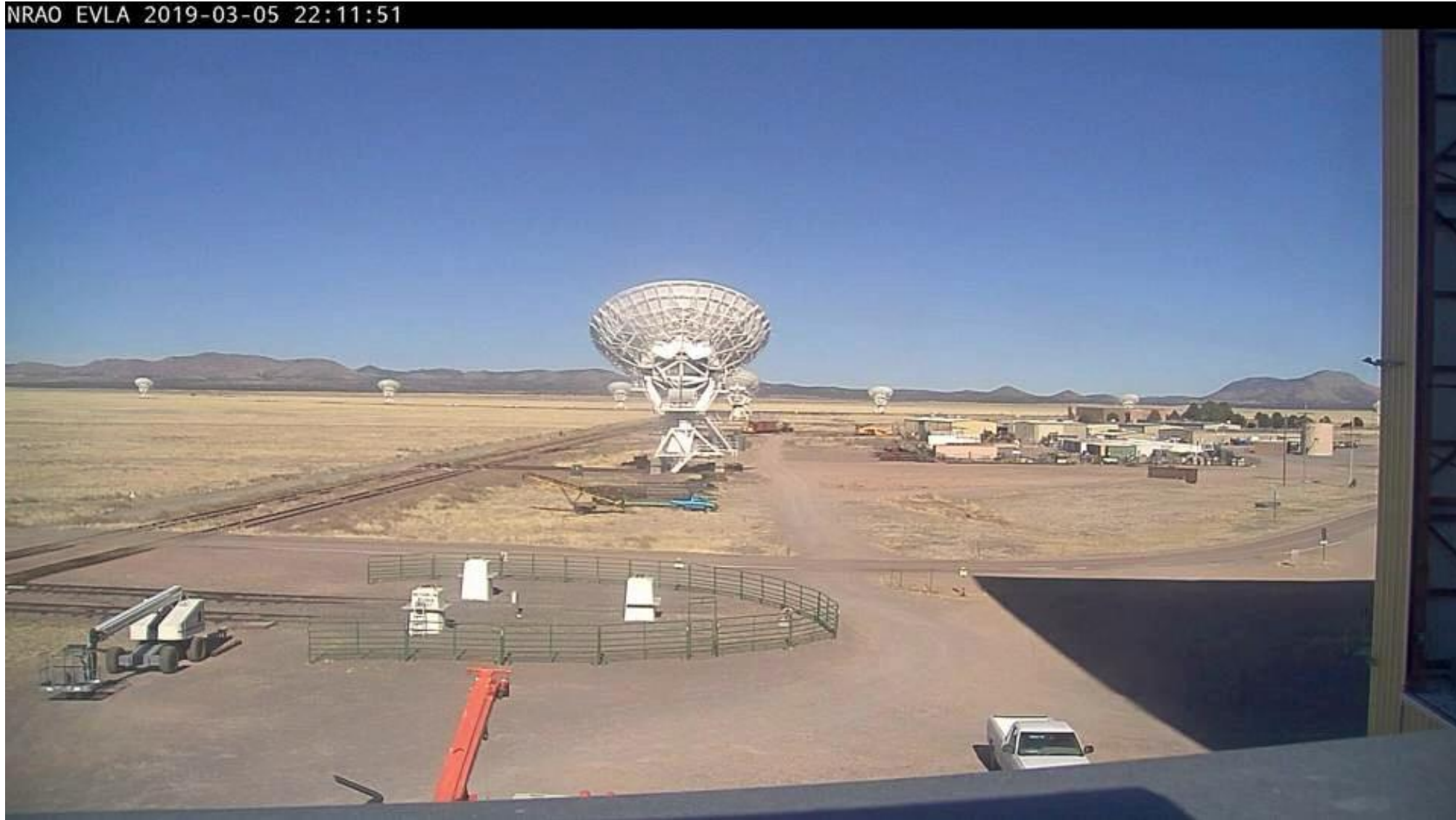
42 J. P. Leahy et al.

VLA X-band
~1995;
71min (B),
17min (C),
9min (D)

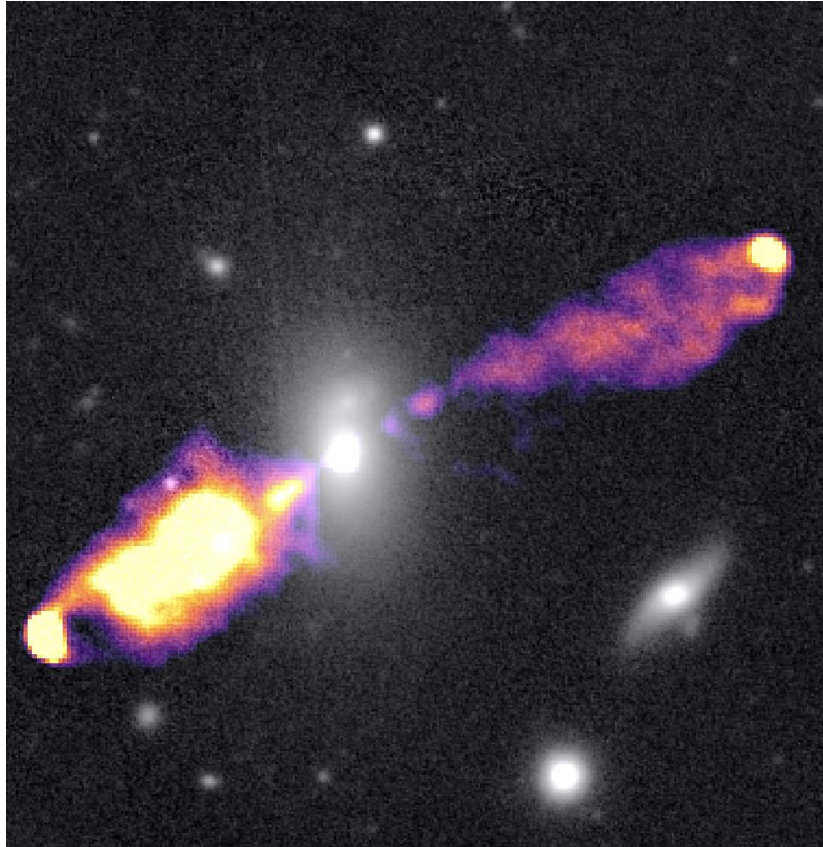


VLASS was taken in “on the fly” mode – the telescope moves in a raster pattern continuously during the observation. (movie courtesy G, Moellenbrock).

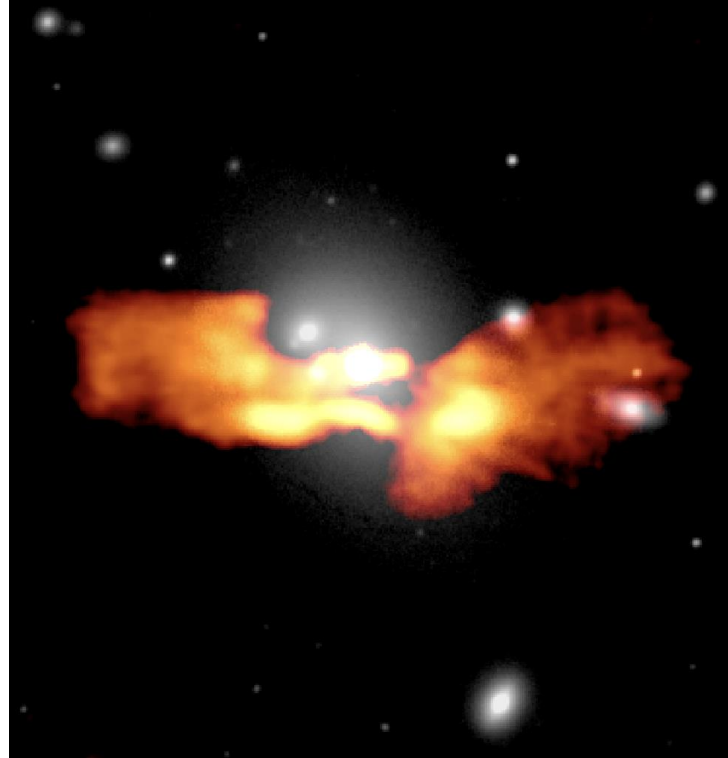
Enables rapid coverage of the sky.



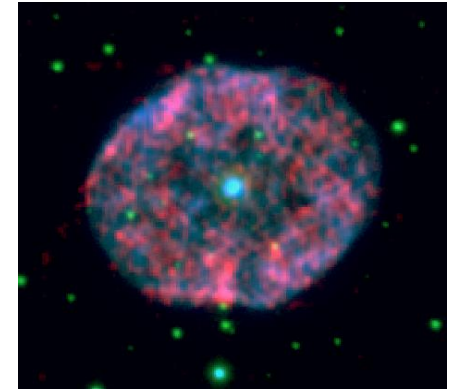
What do we see in VLASS



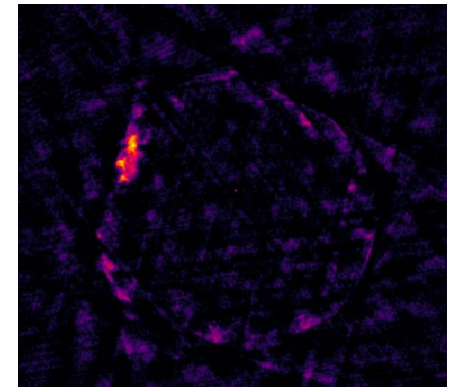
Radio galaxy J1155+5453



Radio galaxy NGC 6166



NGC 1555 (planetary nebula)



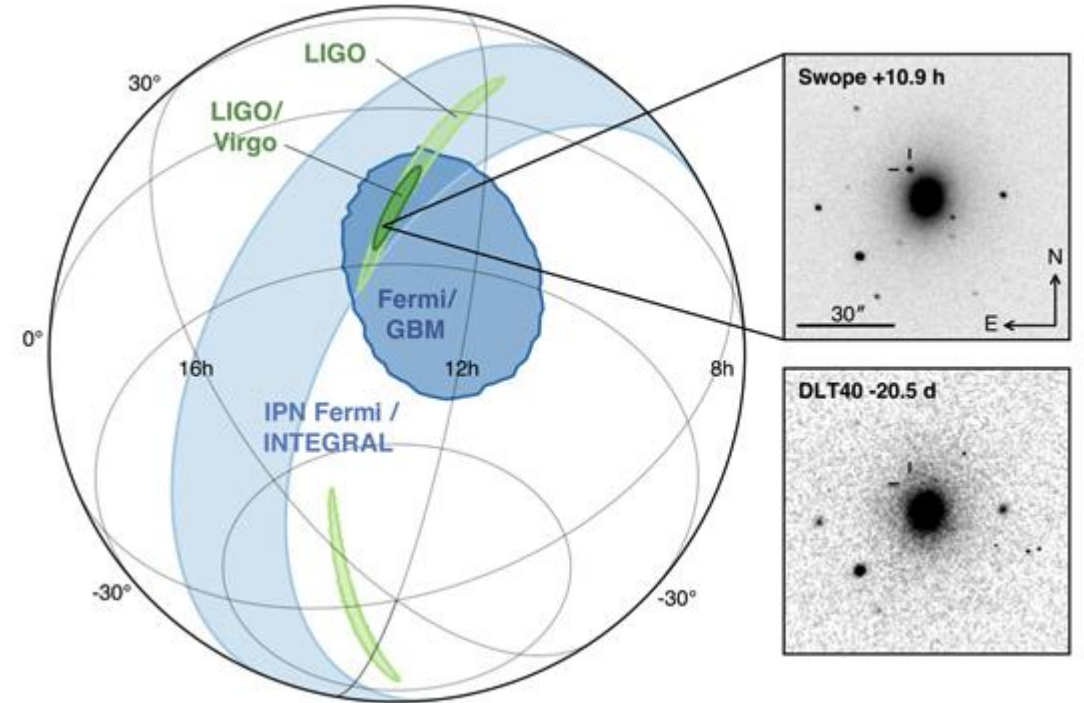
Tycho's supernova remnant

Multi-messenger astronomy with VLASS

- VLASS is most useful as a reference survey – a snapshot of the sky that can be used to compare with new data taken after an event such as a neutrino, gravitational wave or cosmic ray burst.
 - New, or highly-varying objects that might be associated with such a signal can be identified.
 - Higher frequency than other all-sky radio surveys.
- Has also been used for statistical associations of MMA events.

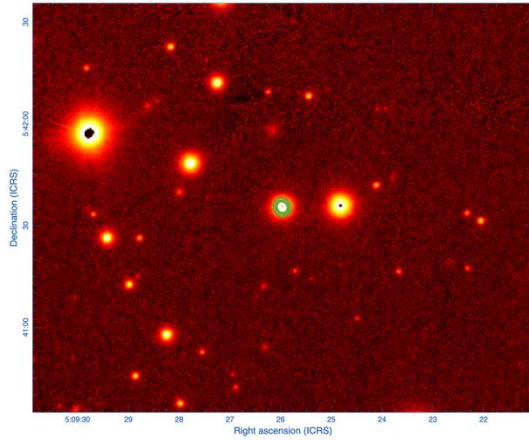
Gravitational waves

- Reference survey to compare to e.g. ns-ns GW event follow-up searches.
 - Unfortunately, GRB 170817A (neutron star – neutron star merger) was only 30 microJy 40Mpc away (compared to VLASS 5-sigma ~ 400 microJy) so any of these would have to be within ~ 10 Mpc for VLASS to be useful as a reference. Also, some events may be radio-brighter.
- May be able to find pulsars buried in the galactic plane for pulsar timing work.

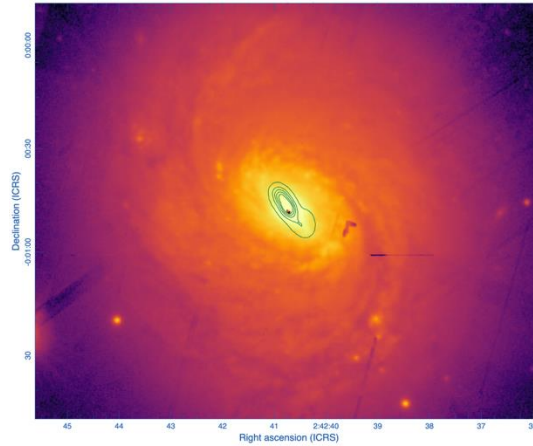


Abbott et al. 2017

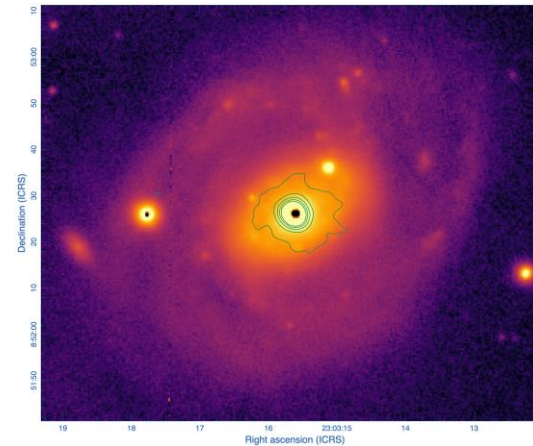
Some VLASS candidate neutrino emitters.



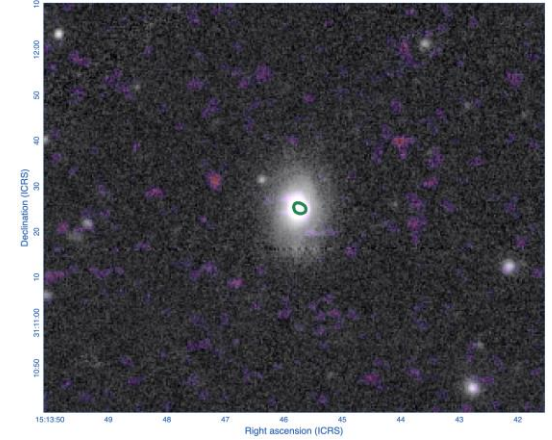
TXS 0506+056
Bright (2Jy) , emission
beamed towards us.



NGC 1068, bright nearby
AGN, extended, not
beamed



NGC 7469(?), nearby
AGN (2 photons)

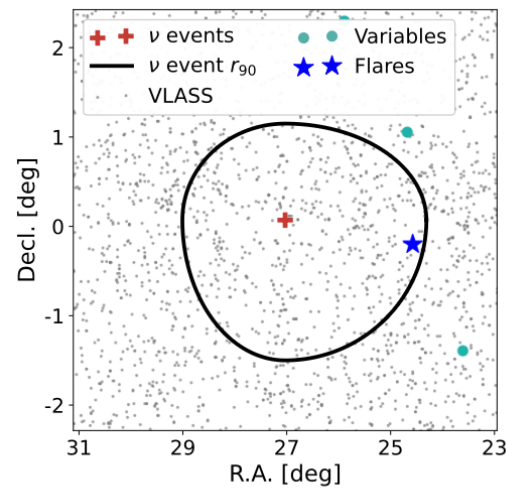


Tidal disruption event?
(1 photon; Zhou et al. 2026)

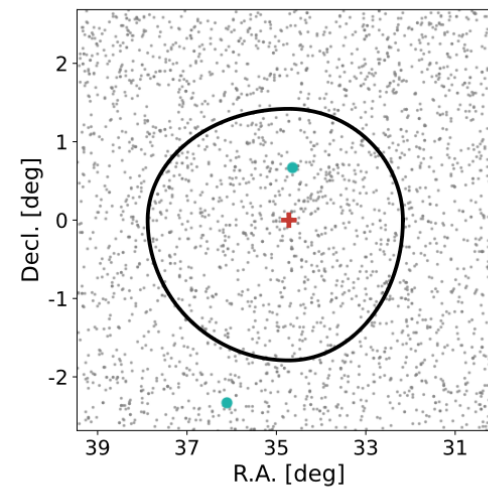
- Typical IceCube high energy neutrino localization ~ 1 deg.
- A handful of sources have been (more-or-less) identified as candidate neutrino emitters and have a very high VLASS detection rate (though possibly biased..)

Statistical studies

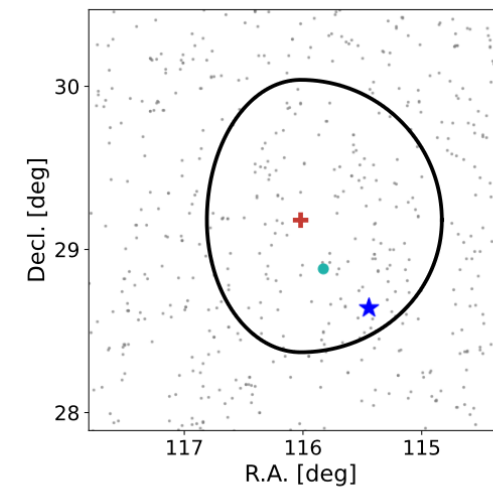
- Gordon et al. 2025. ~ 2 -sigma association of IceCube neutrino sources with flaring VLASS radio sources.



(a)



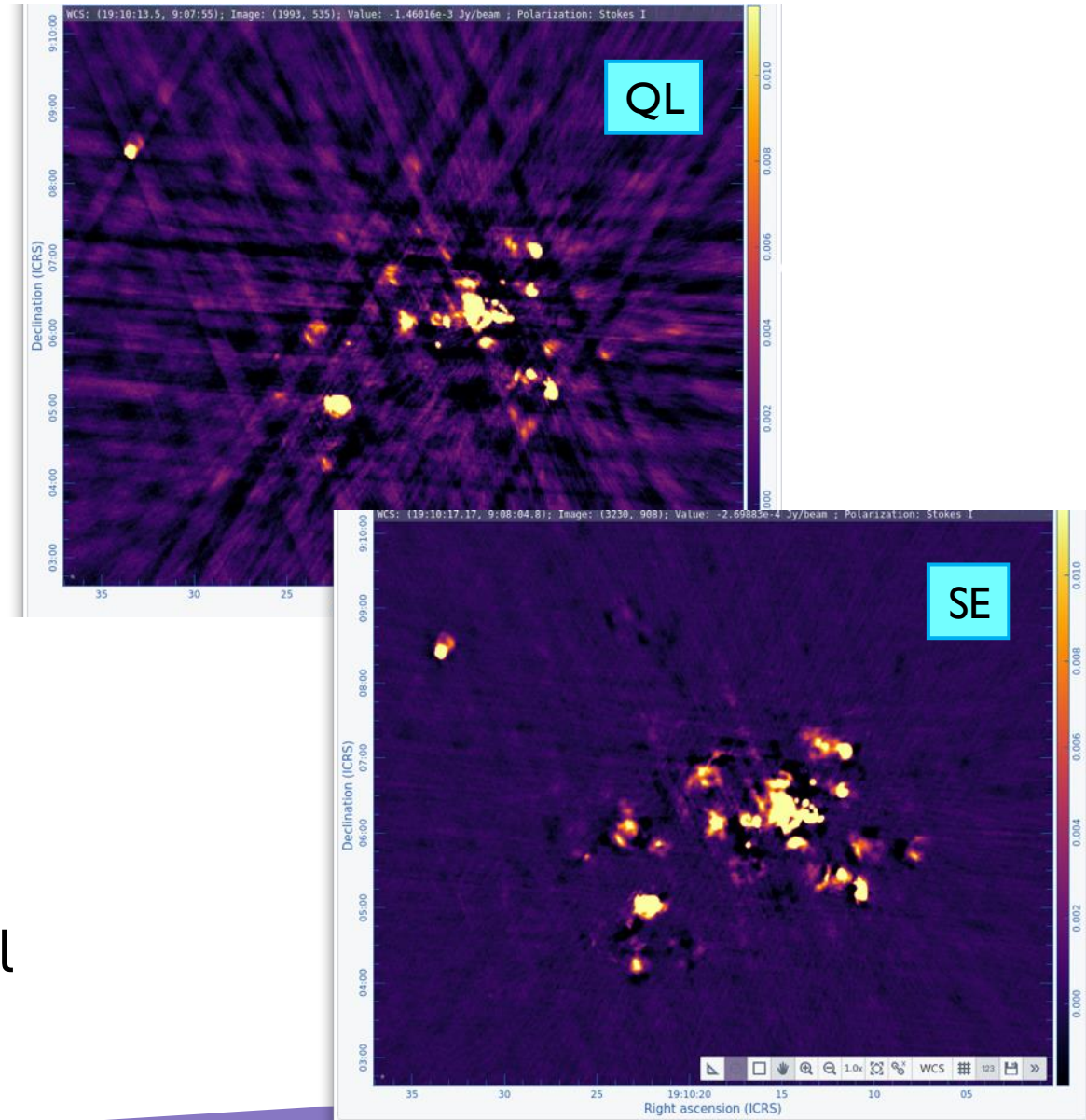
(b)



(c)

VLASS data processing

- “Quick Look” images were produced quite rapidly after observation (average < 2 weeks).
- Higher quality images (“Single Epoch”) require much more processing.
- The algorithmic and computing needs of these products have led to several year delays in making them.
- Algorithmic issues now (mostly) fixed, but computing resource needs are still daunting.

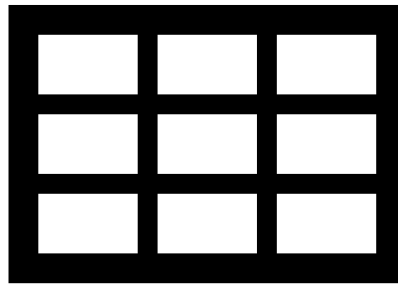


Radio interferometric imaging in one slide

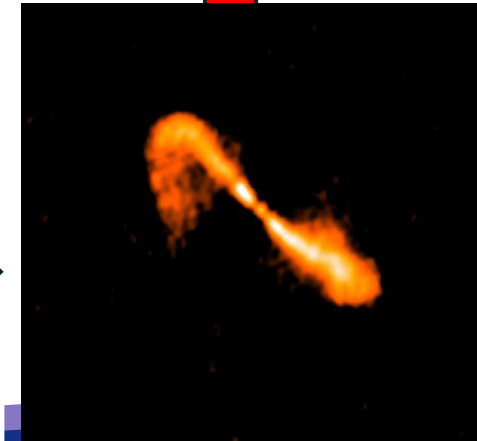


Signals from the VLA's 27 antennas are cross-correlated to produce complex "visibilities" of interferometric fringes.

Visibility values (phase and amplitude) are then convolved with the antenna response and interpolated onto a grid



Fourier Transform



Iteratively deconvolved in a "minor cycle"

Model from deconvolution is Fourier transformed back into visibilities and subtracted from data. Repeat these "major cycles" until convergence.

Major Cycle



Gridding

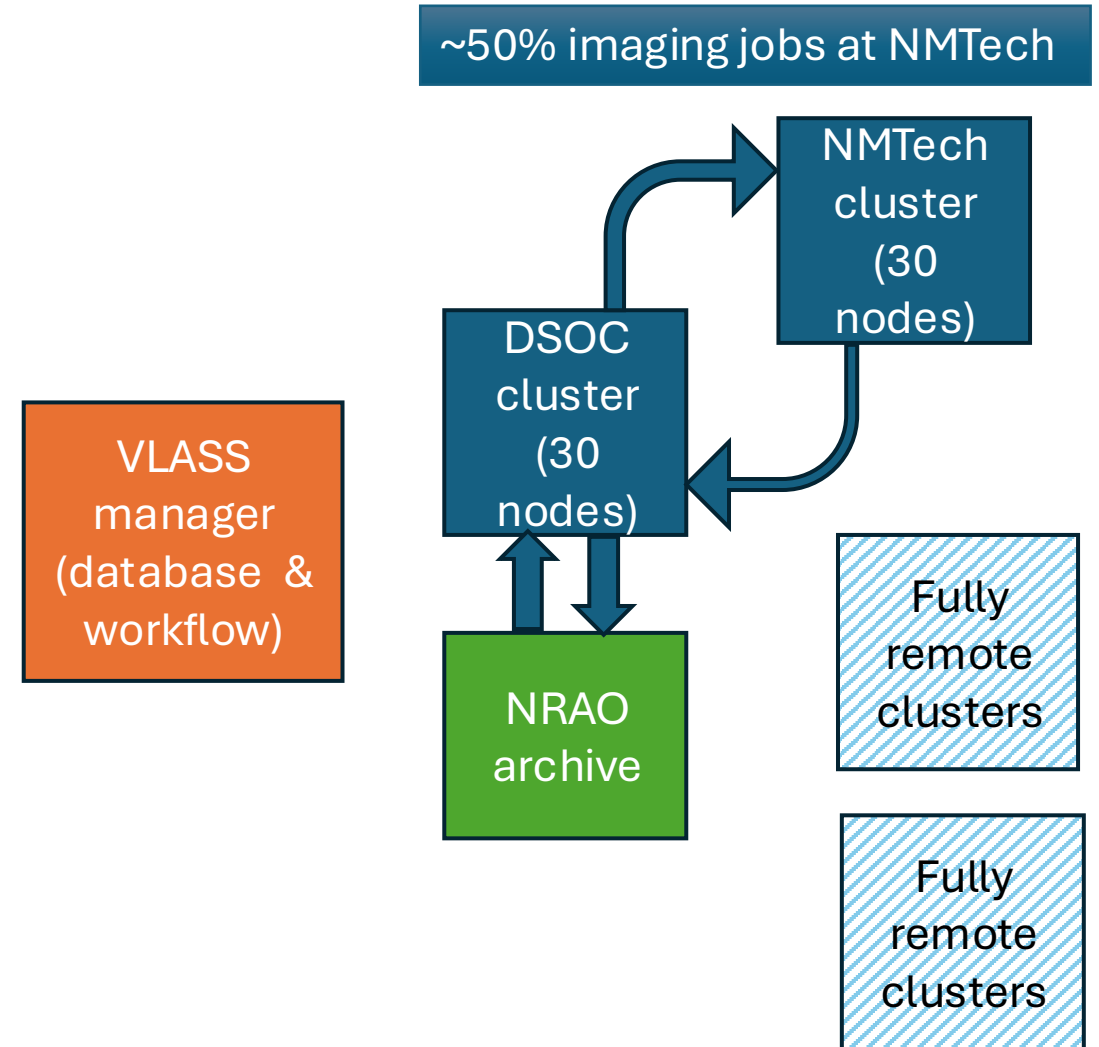
- Gridding is conceptually simple, but computationally expensive.
- For wide-band, wide field images such as VLASS, the curvature of the sky needs to be accounted for at low elevations, needing w -terms in a 3rd dimension (in adding to u, v in 2D), multiplies compute needs by $\sim x100$.
- GPU and parallelized CPU-based gridders take $\sim 20-60$ minutes per major cycle (with a typical VLASS image needing 50-100 major cycles (0.6-4 days) x multiple images per run).
- Compared to this, the minor cycles are short (few minutes to converge per major cycle).

Why are VLASS (and radio interferometric datasets in general) such a bad fit to HPC?

- Legacy code (CASA) was written when only small bandwidths (~100MHz) could be correlated. Did not scale well to the ~GHz bandwidths of modern facilities. More modern code is 5-10 years from production and VLA data not prioritized. Not written for parallel processing – speed up not linear with #cores. More modern code is 5-10 years from production.
- A fundamental issue, difficult to overcome in any code, is the need to hold data in memory while the Fourier Transform is performed – can't easily split the data into smaller pieces like e.g. optical detectors.
- The deconvolution process is iterative, making run times unpredictable.
- **Need for large (~16-64GB) memory/core and long and variable run times (3-10d) make these data poor fits to traditional HPC.**

Logistical challenges

- Besides the amount of computation, VLASS has significant logistical challenges.
- 3 epochs x 35,500 images/epoch x 2 products = 210,000 imaging jobs (plus ~20% reruns).
 - Data (~20GB/image, expanding to ~150GB in the pipeline) needs to be sent to remote clusters and products retrieved automatically using a workflow manager.
 - Have done this with clusters we manage (DSOC & NMTech) by using HTCondor, but need to be able to use non-NRAO facilities too.



Remaining processing needs

Parameter	SE-continuum* (no w-projection)	SE-continuum (GPU) (w-projection)	SE-Continuum (CPU) (w-projection)	SE-Cube	Total
Number of jobs	28,000	30,000	30,000	120,000	220,000
Total CPU hrs	2.5 million	2.5 million	28 million	36 million	79 million
Total GPU hrs	-	5 million	-	-	5 million
Total data in	0.28 PB	0.25 PB	0.25 PB	1PB	2 PB
Total products	28 TB	25 TB	25TB	2 PB	2.1 PB
RAM/CPU core	64 GB	128 GB	16 GB	30 GB	-
CPU Cores/job [†]	1	1	11	4	-
GPUs /job (L4)	-	1	-	-	-
Job duration [‡]	6 days	3.5 days	3.5 days	3 days	-
Storage/job	150 GB	150 GB	150GB	170 GB	-

Table 2. Remaining processing needs for VLASS, including a 13% contingency for reruns and assuming 50% of the w-projection imaging is via CPUs (awp2) and 50% via GPUs (awphpg) using 32 w-planes for awphpg and 8 w-planes for awp2.

* Excludes the $\approx 32,000$ images already made.

[†] Core numbers can be varied for CPU-based jobs

[‡] Job durations can run 2-3 times longer on complicated fields with bright emission.

VLASS and High Throughput Computing

- HTC is a much better fit for VLASS than HPC.
- Considerable challenges remain
 - High memory/core.
 - Long and unpredictable runtimes.
 - Challenging logistics.
- We are also very constrained with respect to available engineering effort at NRAO to help with data management and logistics.
- Nevertheless, we are very keen to work with the CHTC!