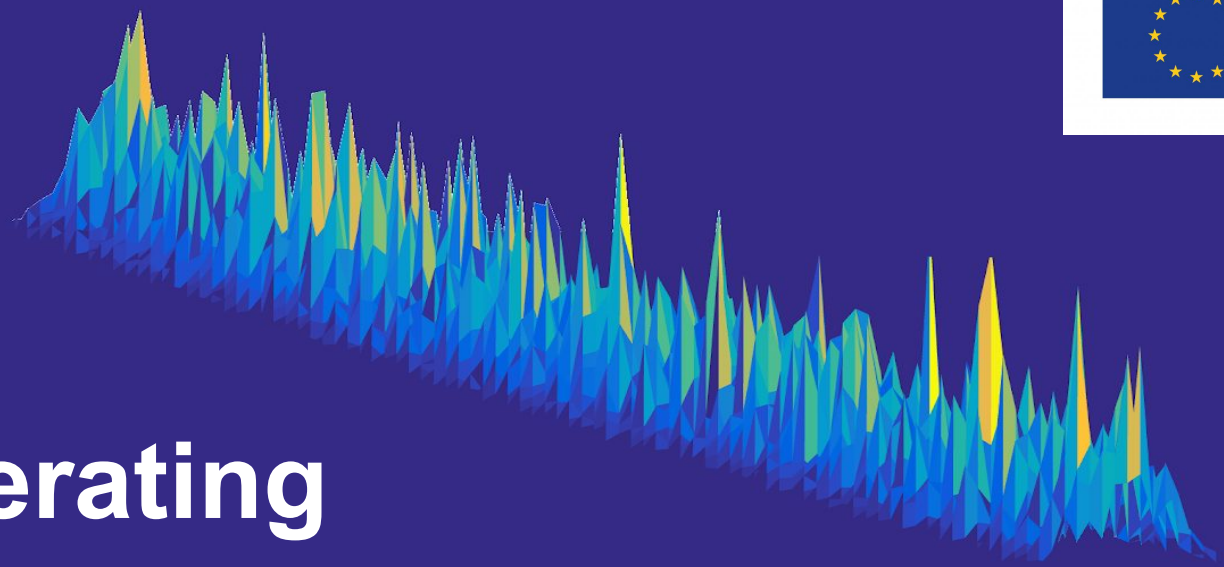
MANCHESTER 1824
The University of Manchester

MEERTRAP



European Research Council
Established by the European Commission

SKA AFRICA
SQUARE KILOMETRE ARRAY

Accelerating Radio Astronomy with HPC

From MeerTRAP perspective

Mateusz Malenta
on behalf of the
MeerTRAP project

The people

The University of Manchester:

- **Ben Stappers**
- Manisha Caleb
- Fabian Jankowski
- Mateusz Malenta
- Kaustubh Rajwade
- Sotiris Sanidas
- Laura Driessen (PG)
- Vincent Morello (PG)
- Tiaan Bezuidenhout (PG)

Oxford University and e-Research Centre:

- Karel Adamek
- Wes Armour
- Cees Carels
- Aris Karastergiou
- Jan Novotny
- Chris Williams

SARAO

- Anton Joubert, Lance Williams & CAM team
- Thomas Abbot
- Sarah Buchner
- Fernando Camilo
- Marisa Geyer
- Dave Horn
- Maciej Serylak
- **All the engineers, operators and on-site scientists - too many to name**

MPIfR:

- Ewan Bar
- Weiwei Chen
- Jason Wu

MeerKAT and MeerTRAP

MeerKAT telescope

- SKA precursor in SA
- 64, 13.96m **(not 13.5!)** offset Gregorian dishes

MeerTRAP - More TRAnsients and Pulsars with MeerKAT

- Fully commensal project - we piggyback on Large Survey Projects, such as MeerTIME
- 856MHz bandwidth, 1712MHz top frequency
- Up to 400 coherent beams, with total FOV ~0.1sq deg
- Incoherent beam ~1.27sq deg
- ~20,000 hours on the sky planned

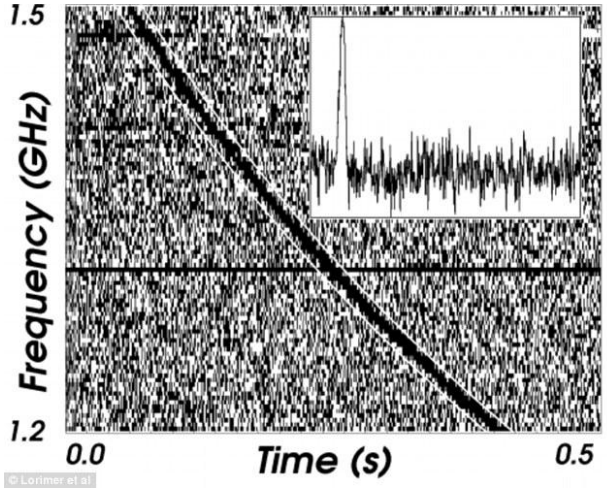
What are we looking for?

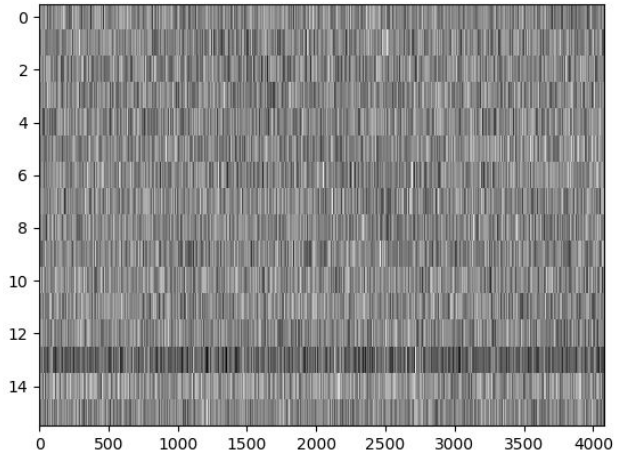
- Pulsars - periodic (1ms - 30s)
- RRATs - periodic, but not like pulsars
- FRBs - **mostly** not periodic (except 10 out of 97, so far)

What are we looking for?

- Pulsars - periodic (1ms - 30s)
- RRATs - periodic, but not like pulsars
- FRBs - **mostly** not periodic (except 10 out of 97, so far)

SINGLE PULSES!





‘Not great, not terrible’

A bit of science

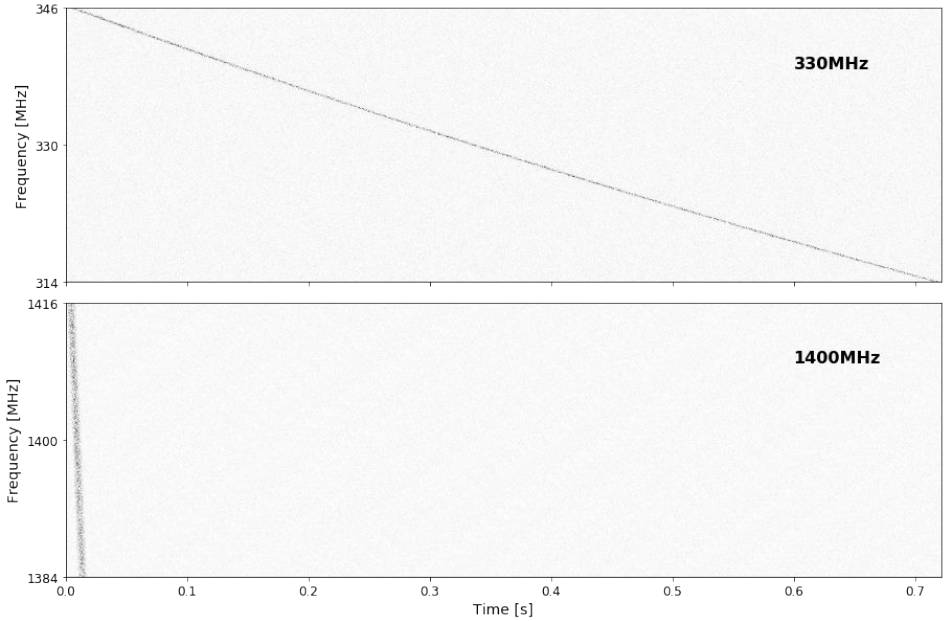
- ‘Light does not travel at the speed of light’
- Velocity of photons inversely proportional to the square of frequency - higher frequency radiation arrives first

$$\Delta t = 4.15 \times 10^6 \times \left(\frac{1}{f_b^2} - \frac{1}{f_t^2}\right)\left(\frac{\mathrm{DM}}{\mathrm{pc\,cm^{-3}}}\right)\ \mathrm{ms}$$

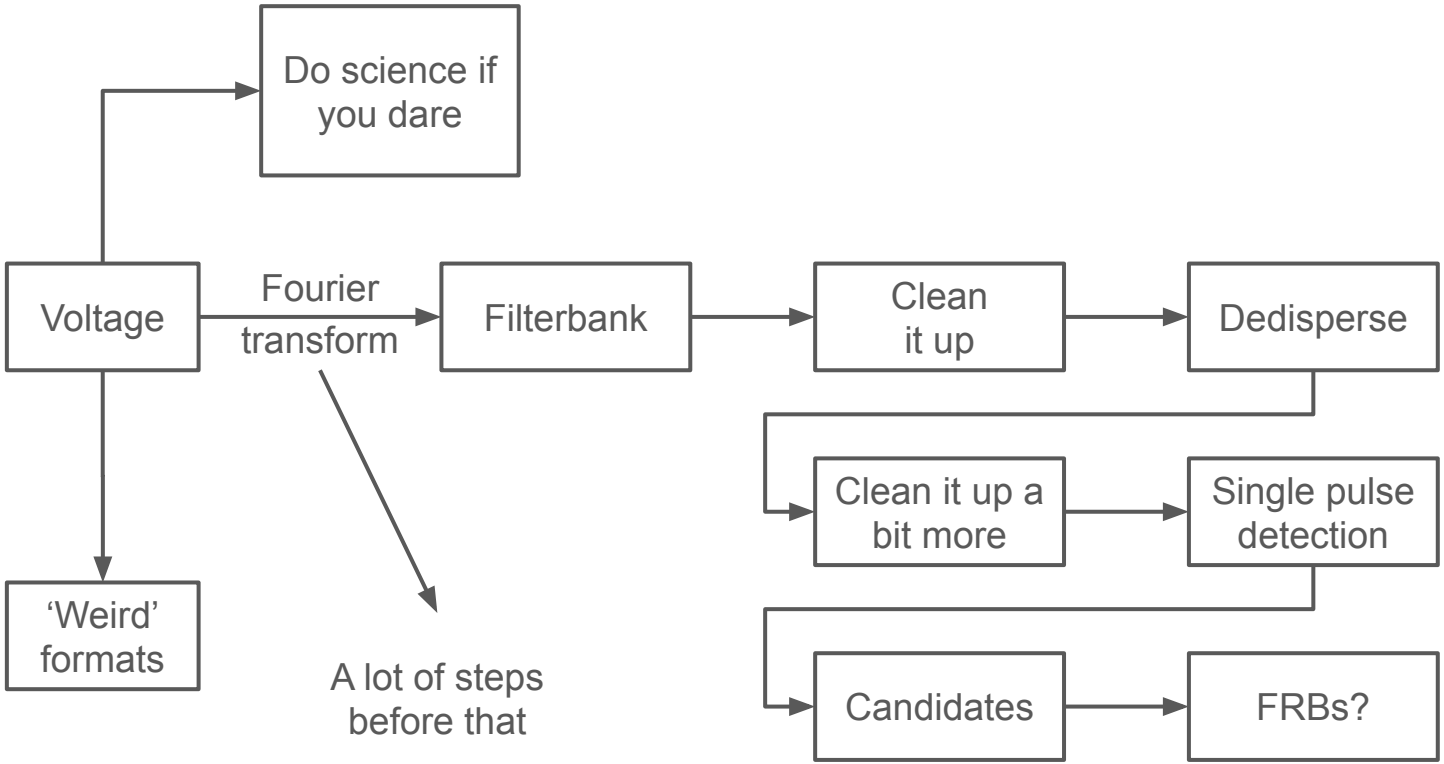
- Problematic when dealing with finite bandwidth, i.e. all the time in real life

$$\mathrm{DM} = \int_0^D n_e(l)\mathrm{d}l$$

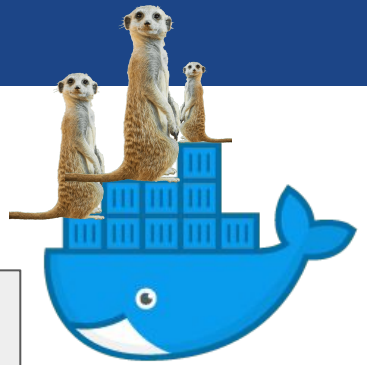
We can’t be worse than everyone else and not have our own distance proxy

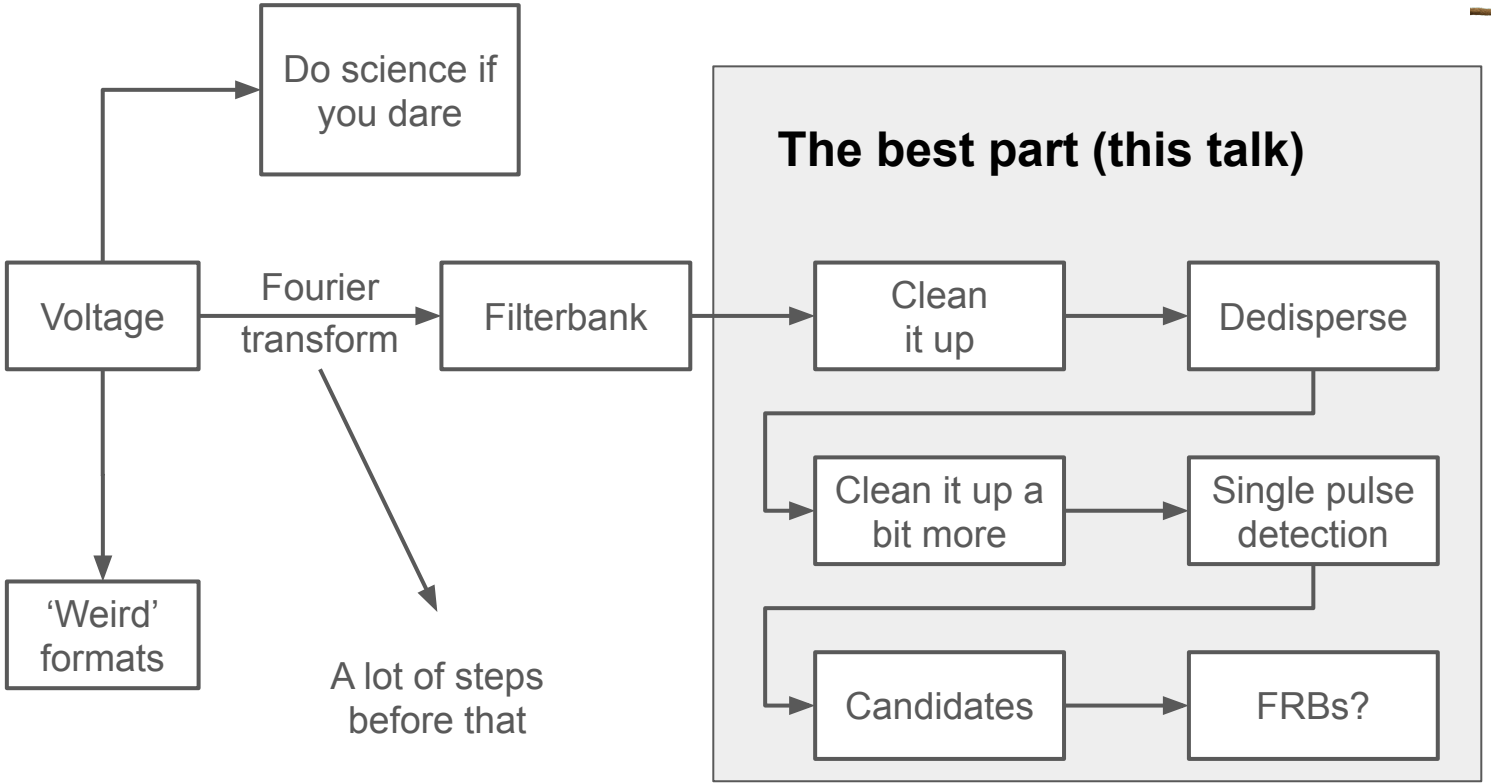


How are we searching?



How are we searching?



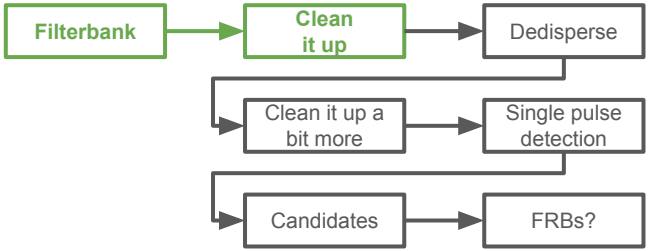


Hardware

- 66 compute nodes (+1 head)
- Intel Xeon E5-2620
- 2 x GTX 1080Ti
- 256GB RAM
- 500GB SSD storage for immediate results

- 500TB+ archive possible in the future

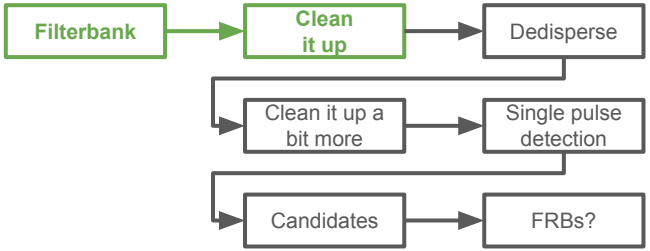
Pre-processing



- Multiple modes possible
- Currently used real-tile modes: 1024 channels @ 306us and 4096 channels at 306us
- ‘Save-to-disk’ mode available with sampling time as low as 76us
- **0.15 / 0.6 / 2.4Gbps between 6 beams**

- Need to receive the data, combine it and output in useful format, all in real time

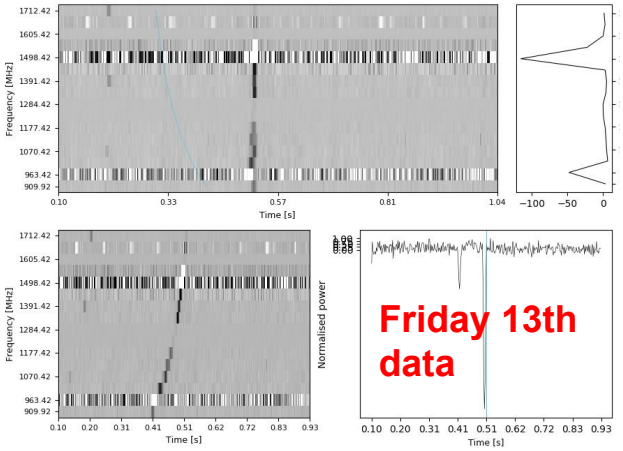
Pre-processing



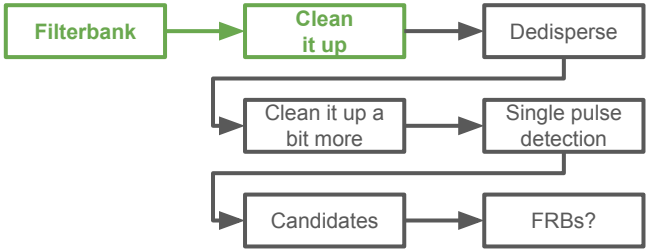
- Multiple modes possible
- Currently used real-tile modes: 1024 channels @ 306us and 4096 channels at 306us
- ‘Save-to-disk’ mode available with sampling time as low as 76us
- **0.15 / 0.6 / 2.4Gbps between 6 beams**

- Need to receive the data, combine it and output in useful format, all in real time

Not much cleaning up, right? It’s the middle of a desert!



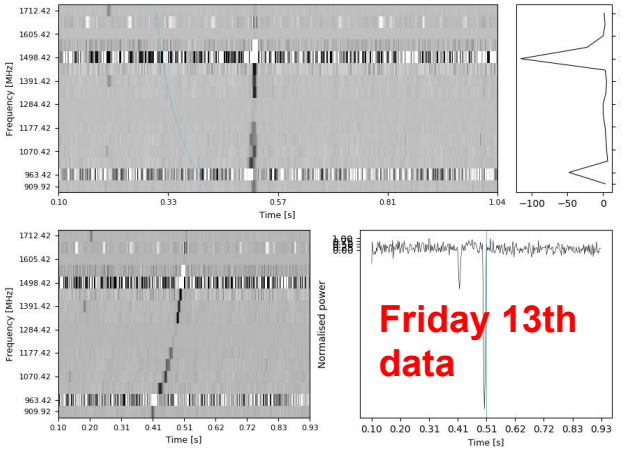
Pre-processing



- Multiple modes possible
- Currently used real-tile modes: 1024 channels @ 306us and 4096 channels at 306us
- ‘Save-to-disk’ mode available with sampling time as low as 76us
- **0.15 / 0.6 / 2.4Gbps between 6 beams**

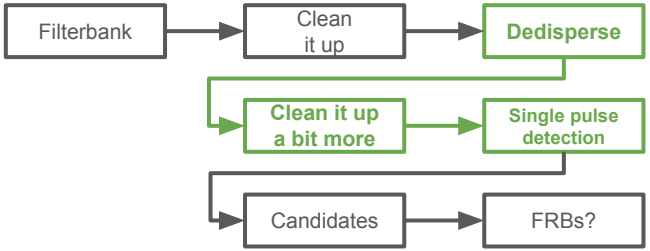
- Need to receive the data, combine it and output in useful format, all in real time

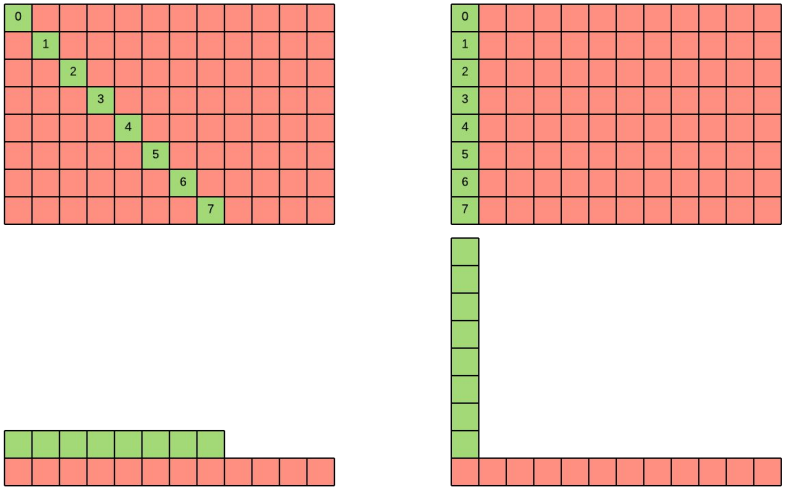
Not much cleaning up, right? It’s the middle of a desert!



- RFI environment is dynamic (planes, satellites, maintenance work)
- Currently static mask used - we are working on moving towards more robust solution

Processing - dedispersion



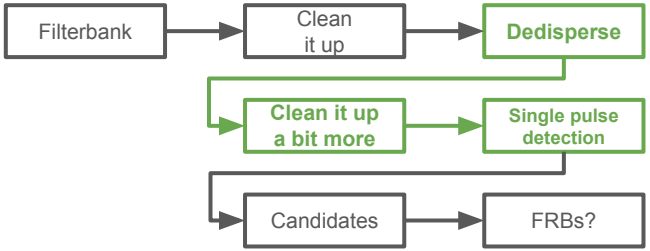


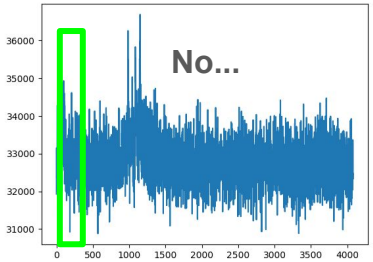
- Major problem - **we don't know the DM beforehand**
- DM step small enough that intra-channel smearing is not a major concern, but big enough so that we don't dedisperse unnecessary DMs
- End up with **(tens of) thousands** of DMs

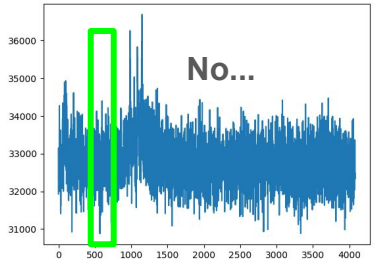
This is the most compute-intensive part of the pipeline

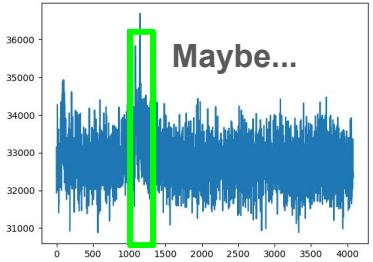
Has to be run on GPUs, using highly-optimised AstroAccelerate algorithms - up to **20x faster than real-time on a single GPU**.

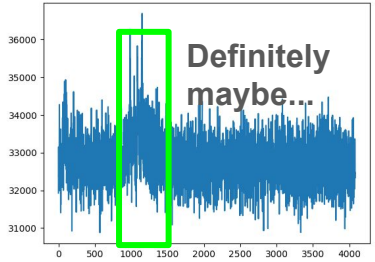
Processing - search for single pulses







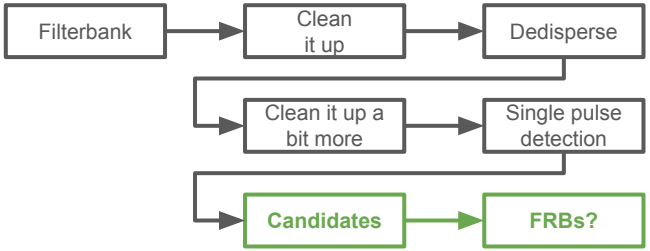




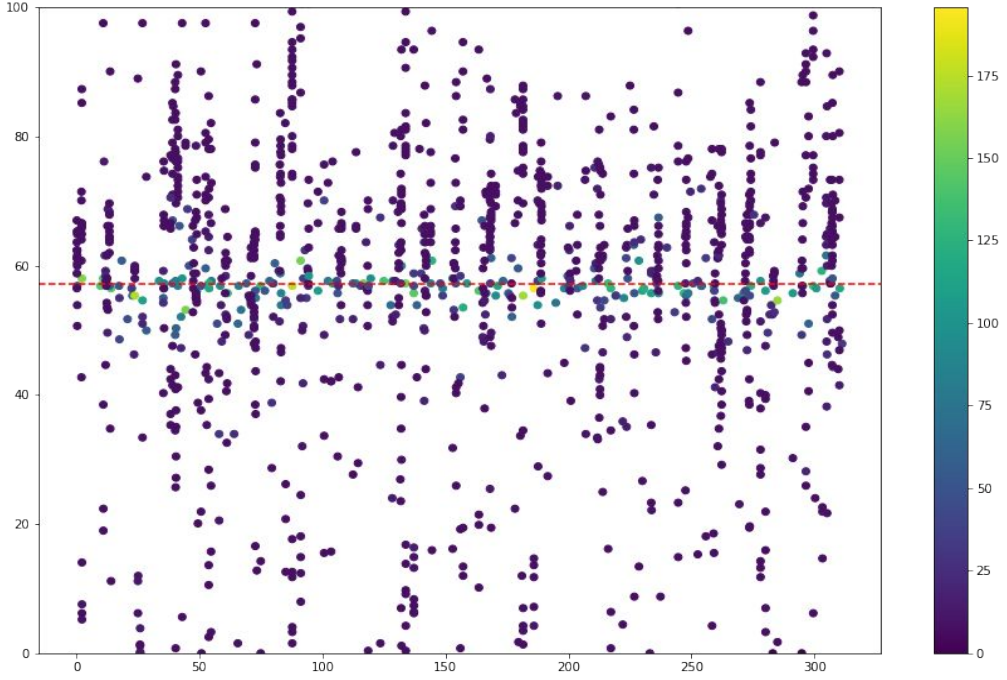
- Boxcar filter - convolve the signal with a simple rectangular function
- **1000s of DMs x tens / hundreds of widths**
- Boxcar filtering is fast, (relatively) easy to implement, has a lot of room for optimisations and gives good enough results
- Use a GPU again

- More complex shapes? More complex algorithms

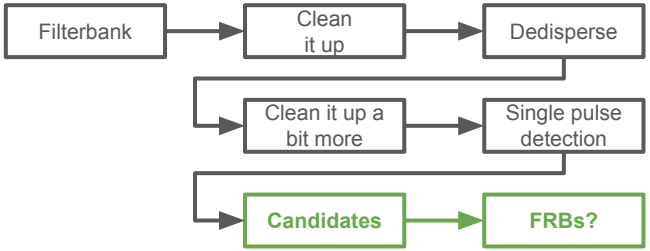
Post-processing - close to ideal



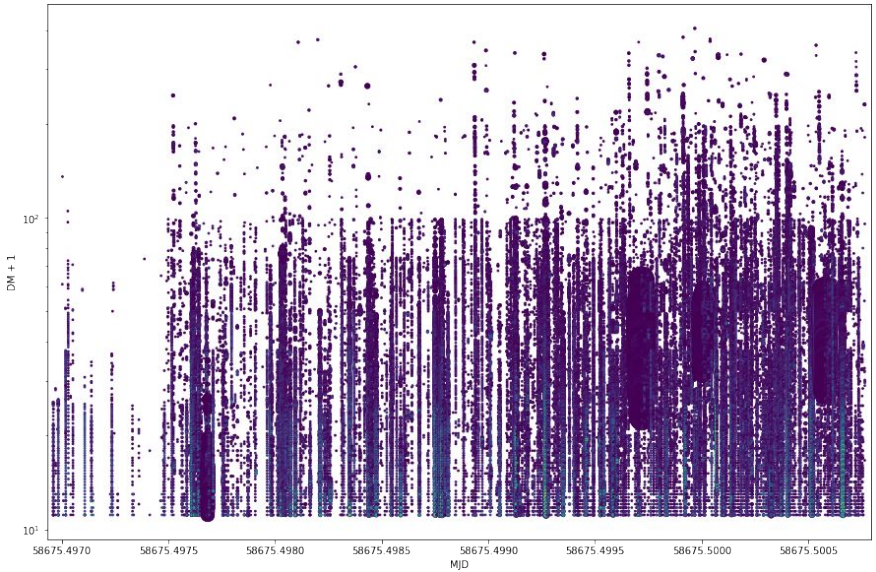
Nice and easy **when you know what you are looking for**



Post-processing - reality

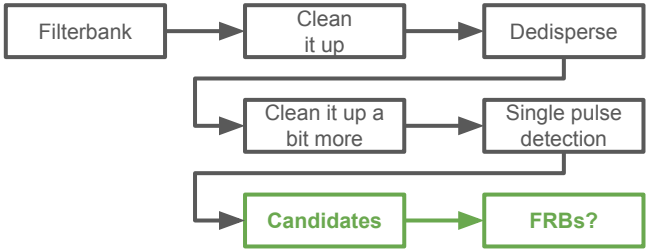


- A lot of students to go through the data or a single one for 4 years
- Use a GPU, or a very fast CPU
- Simple thresholding, clustering or machine learning
- Need to get better at rejecting false-positives, while keeping **false-negatives close to 0** - we don't want to remove real events that never repeat

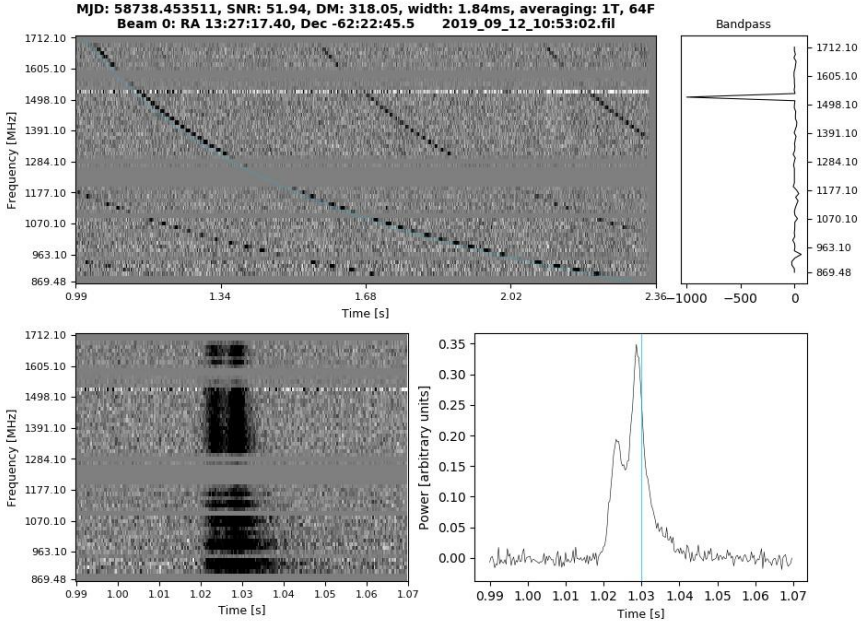


None of these are 'real'

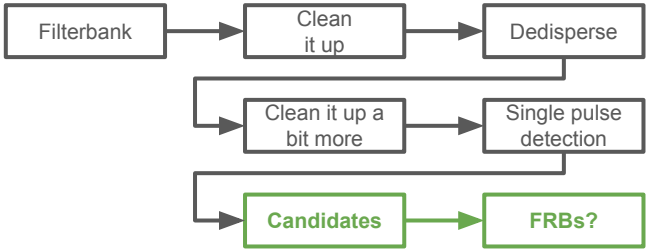
Post-processing - visual inspection



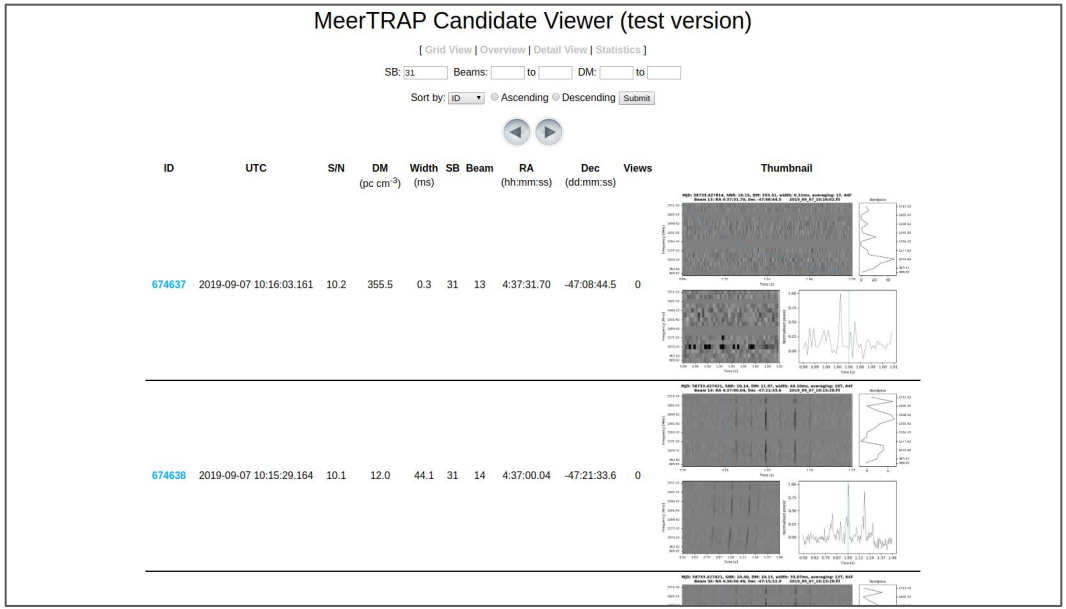
- Should be a manageable size at this stage, but sometimes isn't **(220k+ plots per day)**
- Relatively simple internal website used to view candidates
- Code for marking and **filtering out multibeam RFI** and known sources coming soon
- Can be used as a **training set** in the future - we need to do the dirty work first



Post-processing - visual inspection



- Should be a manageable size at this stage, but sometimes isn't **(220k+ plots per day)**
- Relatively simple internal website used to view candidates
- Code for marking and **filtering out multibeam RFI** and known sources coming soon
- Can be used as a **training set** in the future - we need to do the dirty work first



MeerTRAP future

We are working on many improvements:

- Better time resolution - trying to squeeze as much as possible from algorithms to go down to 152us or even the ‘native’ 76us
- Better candidate filtering - investigating ML with raw candidates or final data products
- Dynamic RFI mitigation
- ‘Raw’ voltage saves

- Pipeline automation / automated monitoring
- Separating monolithic containers and moving towards orchestration

MeerTRAP future

We are working on many improvements:

- Better time resolution - trying to squeeze as much as possible from algorithms to go down to 152us or even the ‘native’ 76us
- Better candidate filtering - investigating ML with raw candidates or final data products
- Dynamic RFI mitigation
- ‘Raw’ voltage saves

- Pipeline automation / automated monitoring
- Separating monolithic containers and moving towards orchestration

MeerKAT detections of FRB 121102 at L-band

ATel #13098; *Manisha Caleb (University of Manchester), Benjamin Stappers (University of Manchester), Ewan Barr (MPIfR), Mechiel Christiaan Bezuidenhout (University of Manchester), Laura Driessen (University of Manchester), Fabian Jankowski (University of Manchester), Michael Kramer (MPIfR), Mateusz Malenta (University of Manchester), Vincent Morello (University of Manchester), Kaustubh Rajwade (University of Manchester), Sotiris Sanidas (University of Manchester), Weiwei Chen (MPIfR), Jason Wu (MPIfR), Sarah Buchner (SARAO), Rob Fender (University of Oxford), Lauren Rhodes (University of Oxford), Maciej Serylak (SARAO), Lee Townsend (UCT), Patrick Woudt (UCT), Julio Andrianjafy (University of Mauritius/DARA), Nalini Heeralall-Issur (University of Mauritius), Divya Hurwanth (University of Mauritius/DARA)*
*on **11 Sep 2019; 13:24 UT***
Credential Certification: Manisha Caleb (manishacaleb@gmail.com)

Subjects: Radio, Transient, Fast Radio Burst

Hopefully many more of these soon!

Thank you!