SERENDIP

- "Classical" SETI thresholded spectrometer

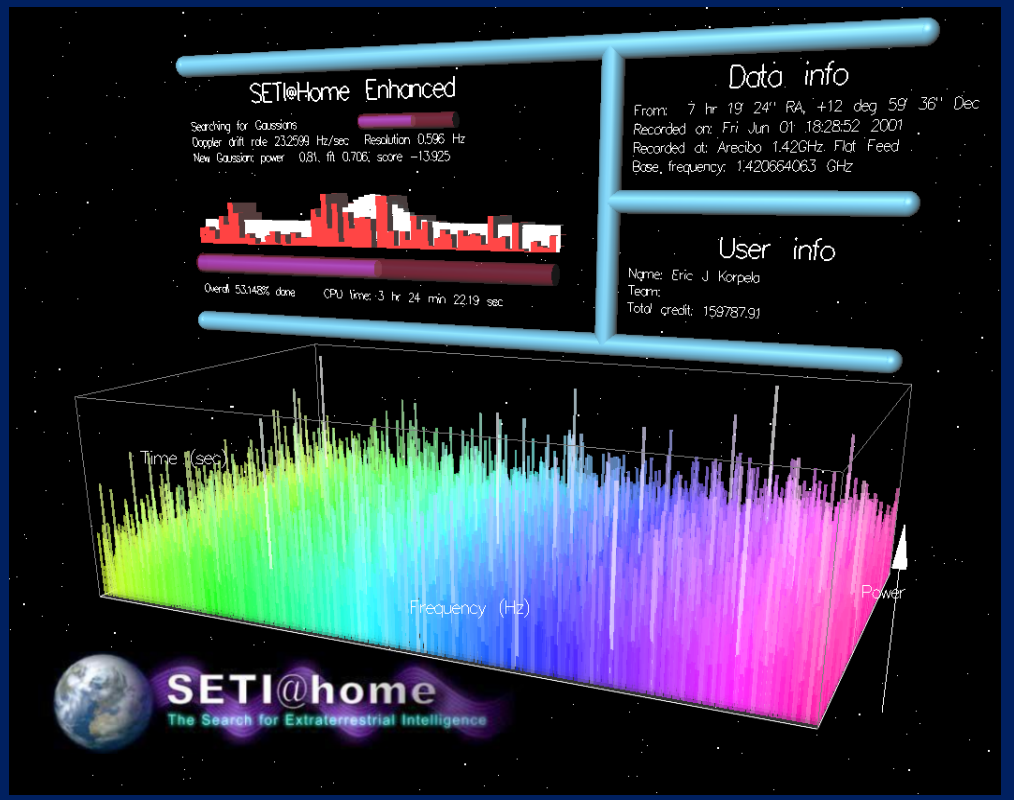
```
while (observing) {
    generate a power spectrum of the entire band
        (~1 sec integration, ~1 Hz resolution)
    find binds above threshold.
        Report time,  bin number, power
}
```

SERENDIP

- SERENDIP VI output
  - ETFITS format (1 file per observation or size limited)
    - FILE HEADER (site info)
    - BINTABLE header, 1 per integration
      - Time, Pointing, Receiver, IF, Spectrometer/ADC status, etc
    - BINTABLE of hits
      - Detected Power
      - Mean Power
      - Coarse Channel
      - Fine Channel

SETI@home

- Released May, 1999
- Second only to Napster in network usage at UCB at the time.
- Coherent Doppler drift correction
- Narrower Channel Width → Higher Sensitivity
- ~150,000 people,
  300,000 computers,
  ~1M CPUs
  >130,000 GPUs
  ~50,000 mobile phones
- 260 TFLOP (actual)
- Variable signal bandwidth/time resolution
- Search for multiple signal types
  - Traditional thresholded channels
  - Gaussian beam fitting
  - Search for repeating pulses
  - Auto-correlation to find repeated non-CW waveforms
- Sensitivity ~ 3×10^{-26} W/m^2



BERKELEY SETI RESEARCH CENTER

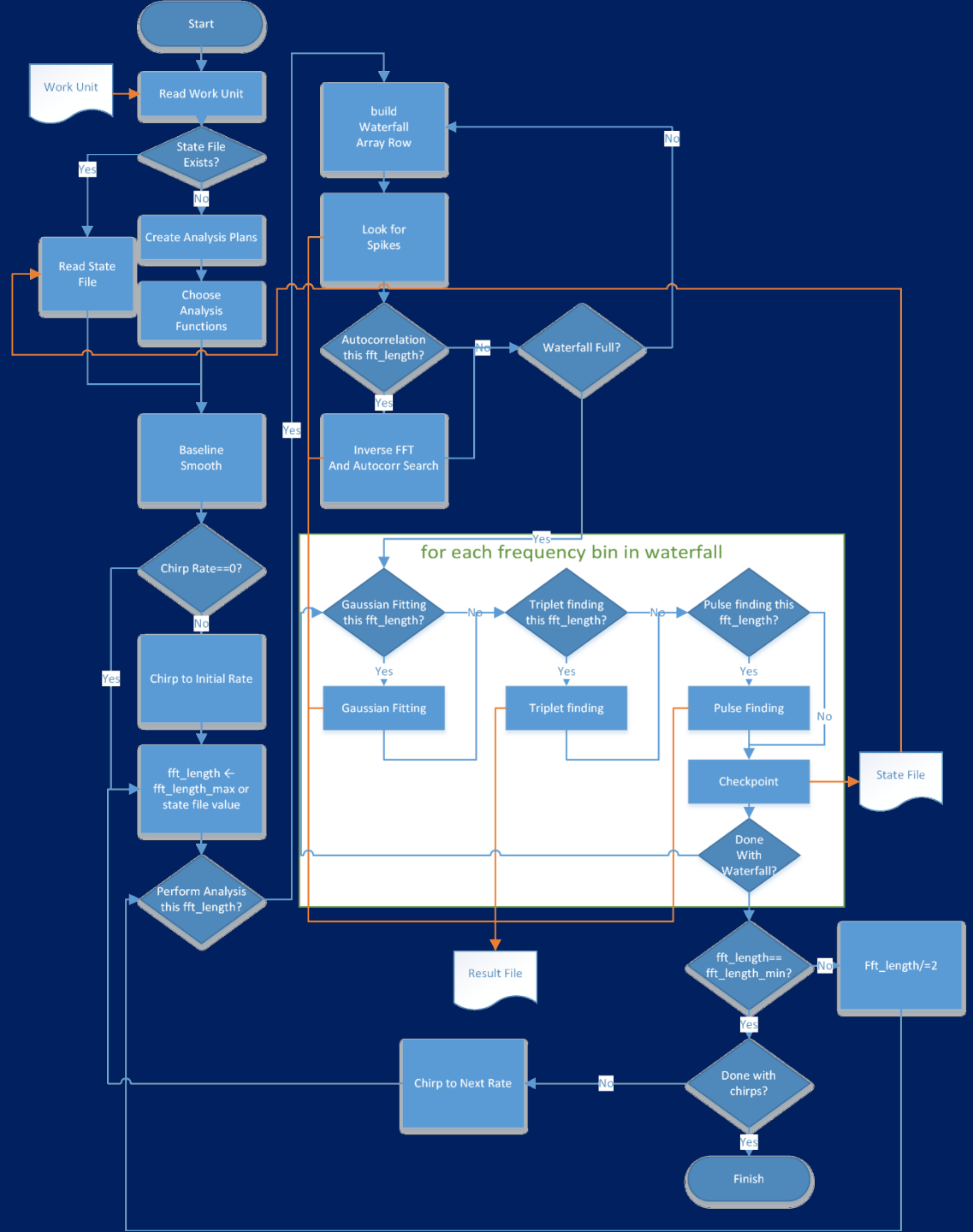
BOINC

SSL Berkeley · CASPER · BWRC · Berkeley Astronomy Department · Berkeley Lab · The University of California 1868

SETI@home Client Processing

for Doppler drift rates from -100 Hz/s to +100 Hz {

Search for signals using autocorrelation.
for bandwidths from 0.075 to 1220 Hz in 2X steps {
Generate time ordered power spectra.
Search for short duration signals above a constant threshold (spikes)
for each frequency in the power spectra {
Search for faint signals matching beam parameters (Gaussians)
Search for groups of three evenly spaced signals (triplets)
Search for faint repeating pulses (pulses)
}
}
}

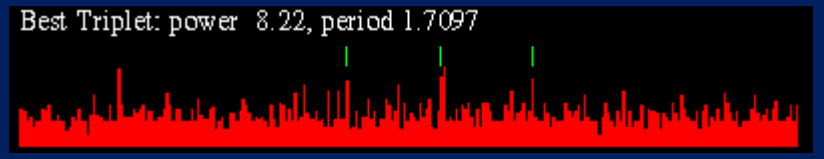


Signal Types

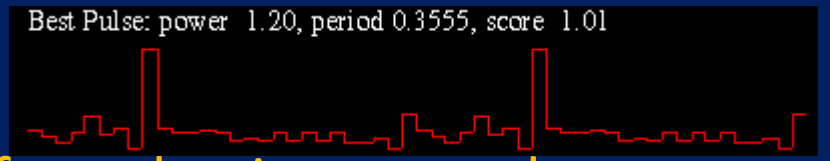
Spike – A single frequency bin at a specific time above a threshold

Gaussian – A power profile over time that matches the telescope beam width.

Triplet – Three spikes evenly spaced in time separated by ms to seconds.

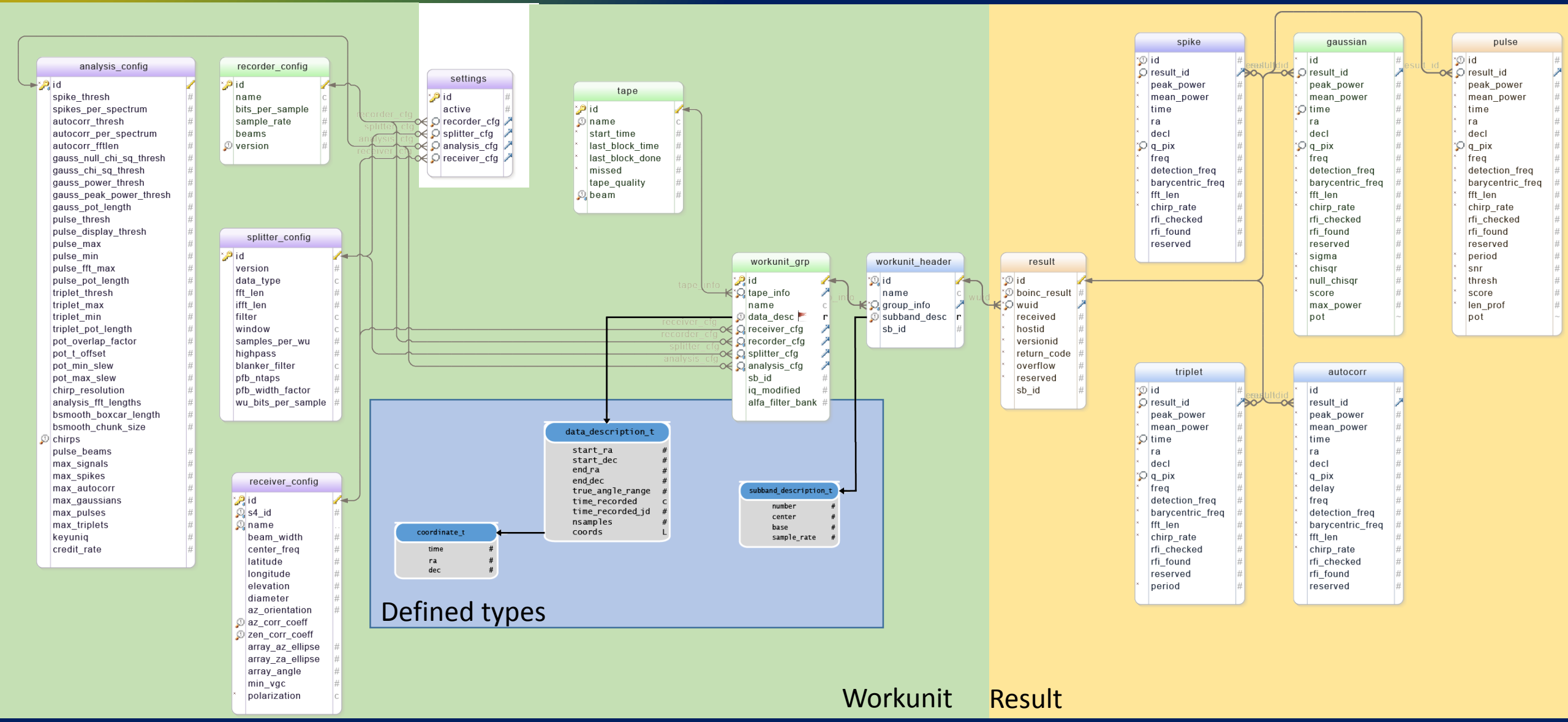


Pulse – Repeated pulsation on scales from ms to seconds.



Autocorrelation – Any waveform that is repeated one or more times with a characteristic delay (0.1 ms to 6.7s).

SETI@home Database Structure



SETI@home Database Structure

Common to all signals:

- result_id
- peak_power
- mean_power
- time
- ra
- decl
- q_pix
- freq
- detection_freq
- barycentric_freq
- fft_len
- chirp_rate
- score

Specific to gaussian:

- sigma
- chisqr
- null_chisqr
- max_power
- PoT[]

Specific to autocorr:

- delay

Specific to triplet:

- period

Specific to pulse:

- period
- snr
- thresh
- len_prof
- PoT[]

Current signal table counts

Spike:	7,285,376,286
Gaussian:	907,657,802
Pulse:	3,551,566,142
Triplet:	3,120,683,163
Autocorr:	729,089,931

How machine learning might help

Efficient RFI identification/removal

- Clusters in frequency and temporally but not spatially

Efficient Candidate identification

- Clusters in frequency and spatially
- (currently use fixed size frequency and spatial windows)

SETI@home Post Processing

Scoring

Based upon probability of arising due to random noise

$$S_{\text{spike}} = e^{-\frac{P}{<P>}} = Q(1, \frac{P}{< P >})$$

$$S_{\text{Gaussian}} = Q(\frac{63}{2}, \frac{\chi^2_{\text{null}}}{2})$$

$$S_{\text{Triplet}} = Q(3, \frac{P}{2 < P >})$$

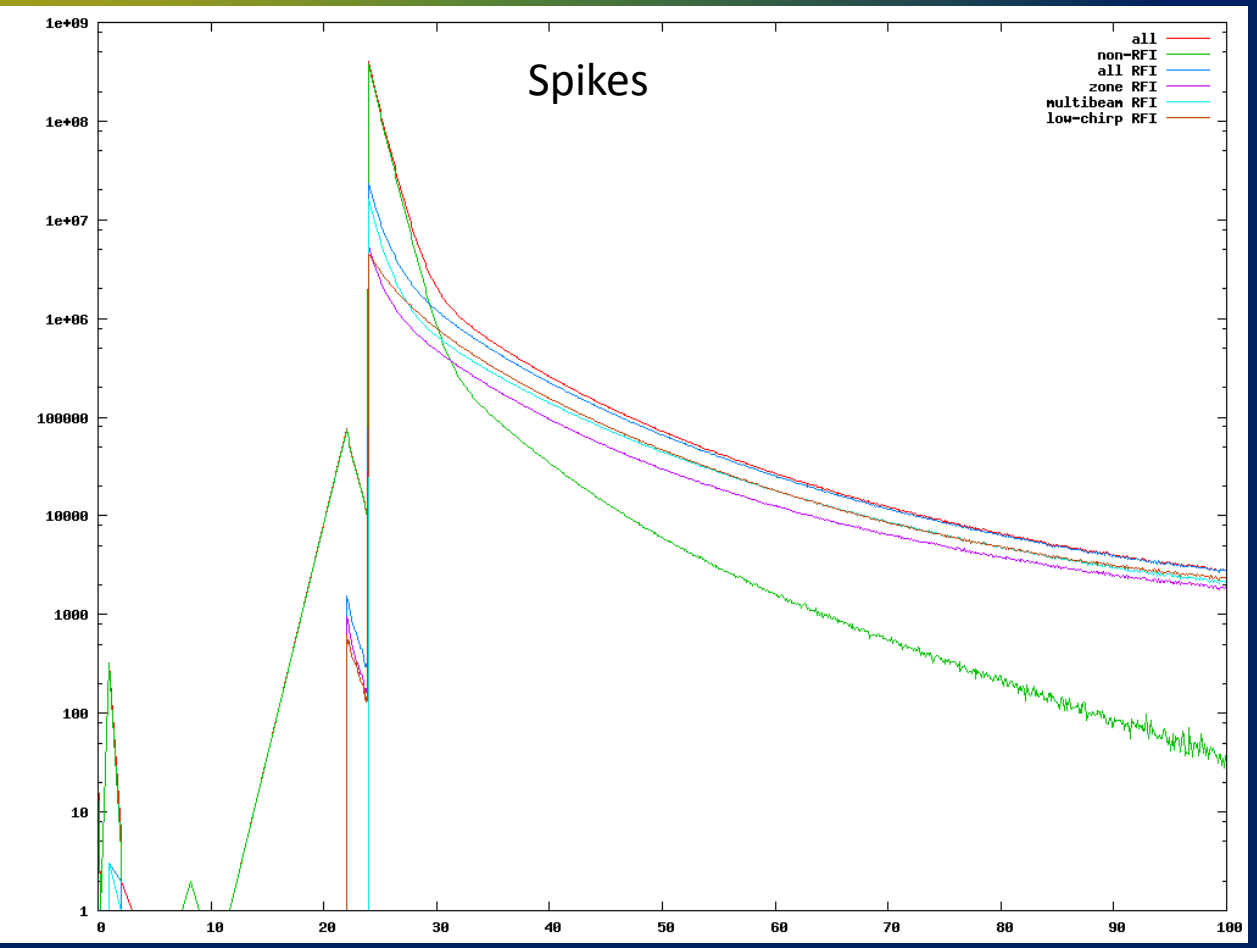
$$S_{\text{Pulse}} = mQ(n, n\frac{P}{< P >})$$

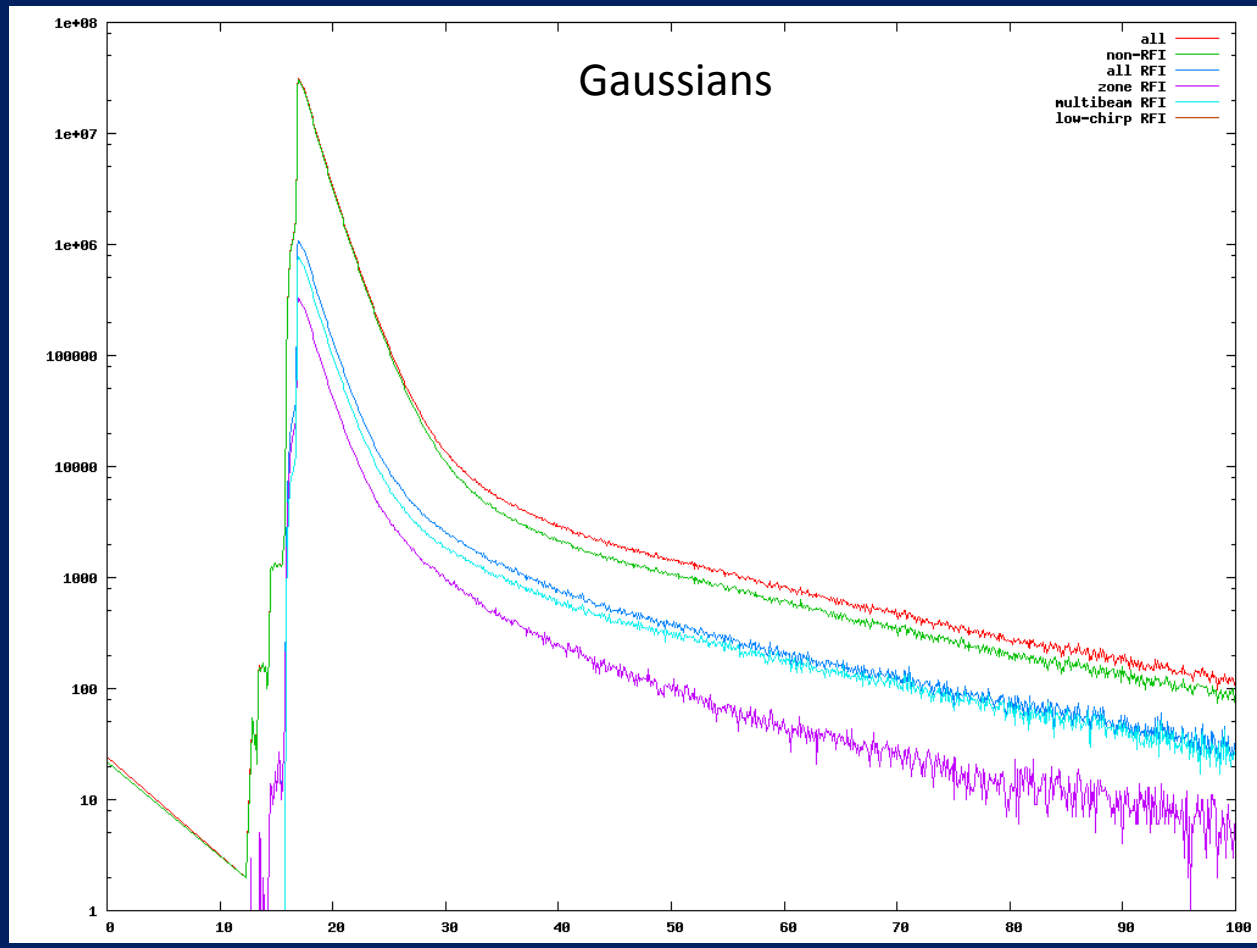
where m is the number of bins in the folded array, n is the number of times the array has been folded, and $\frac{P}{<P>}$ is the power of the pulse in the folded array.

SETI@home Post Processing

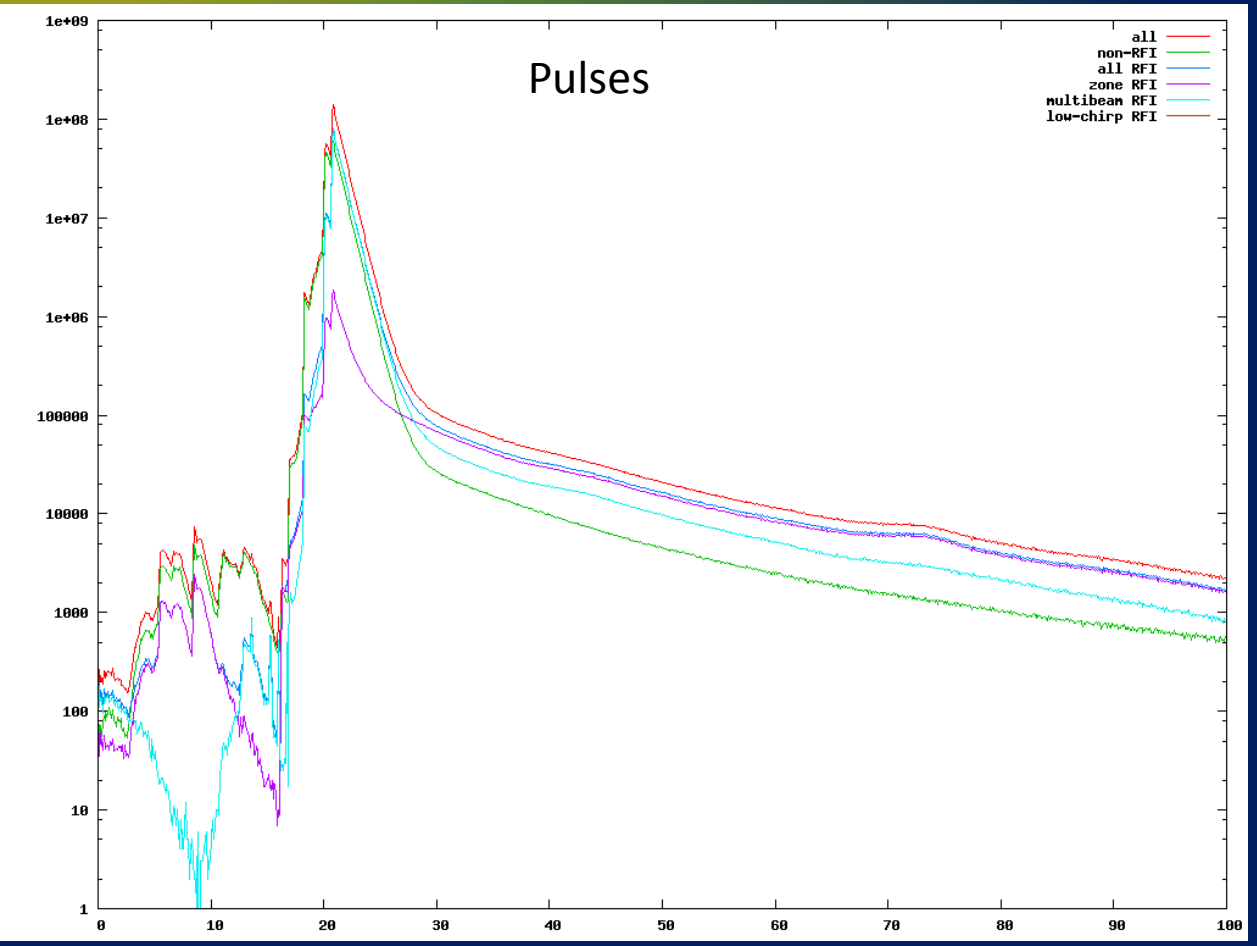
- Radar blanking (in splitter)
- Frequency, time, or period zones
- Cross beam rejection
- Cross polarization confirmation

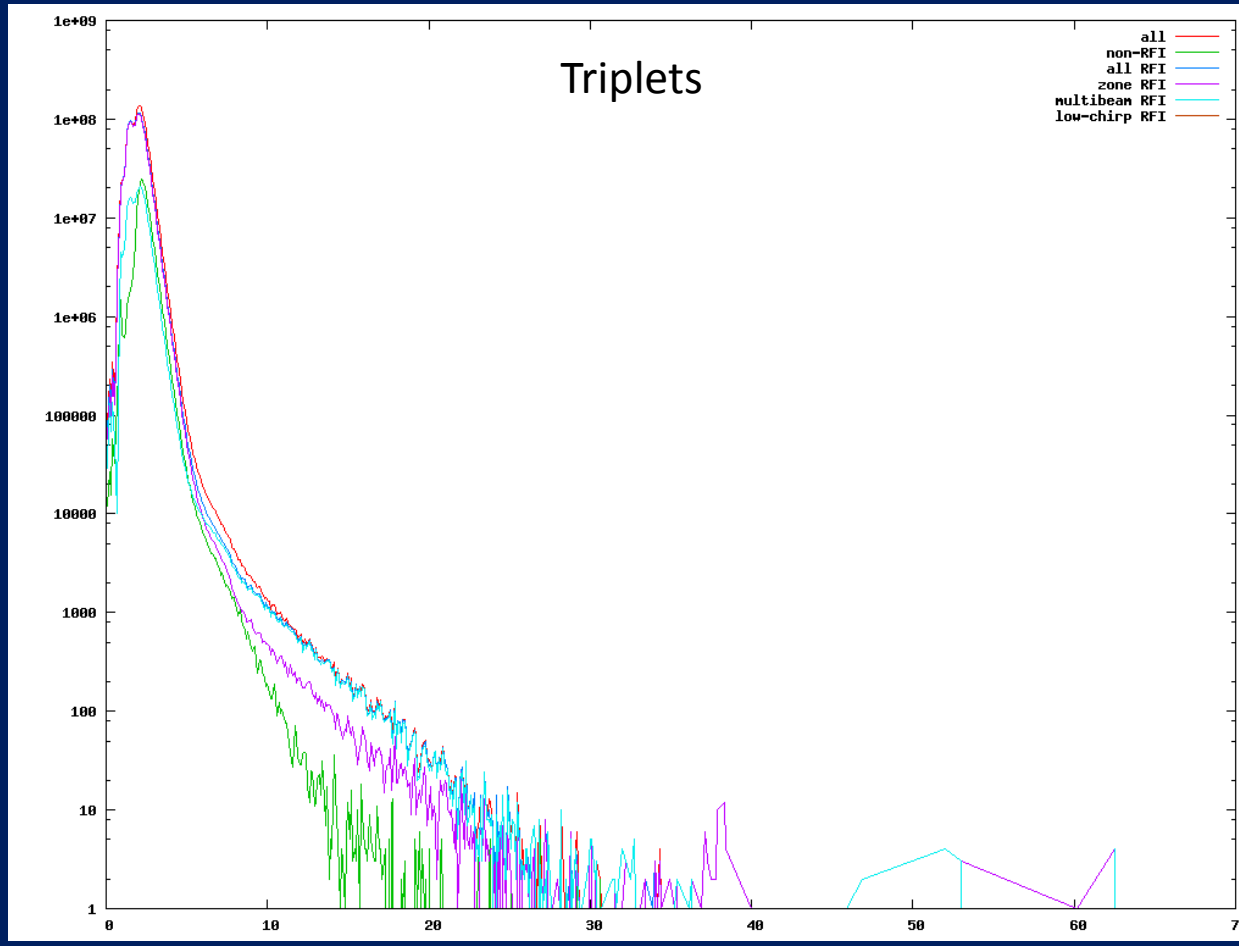
SETI@home Post Processing



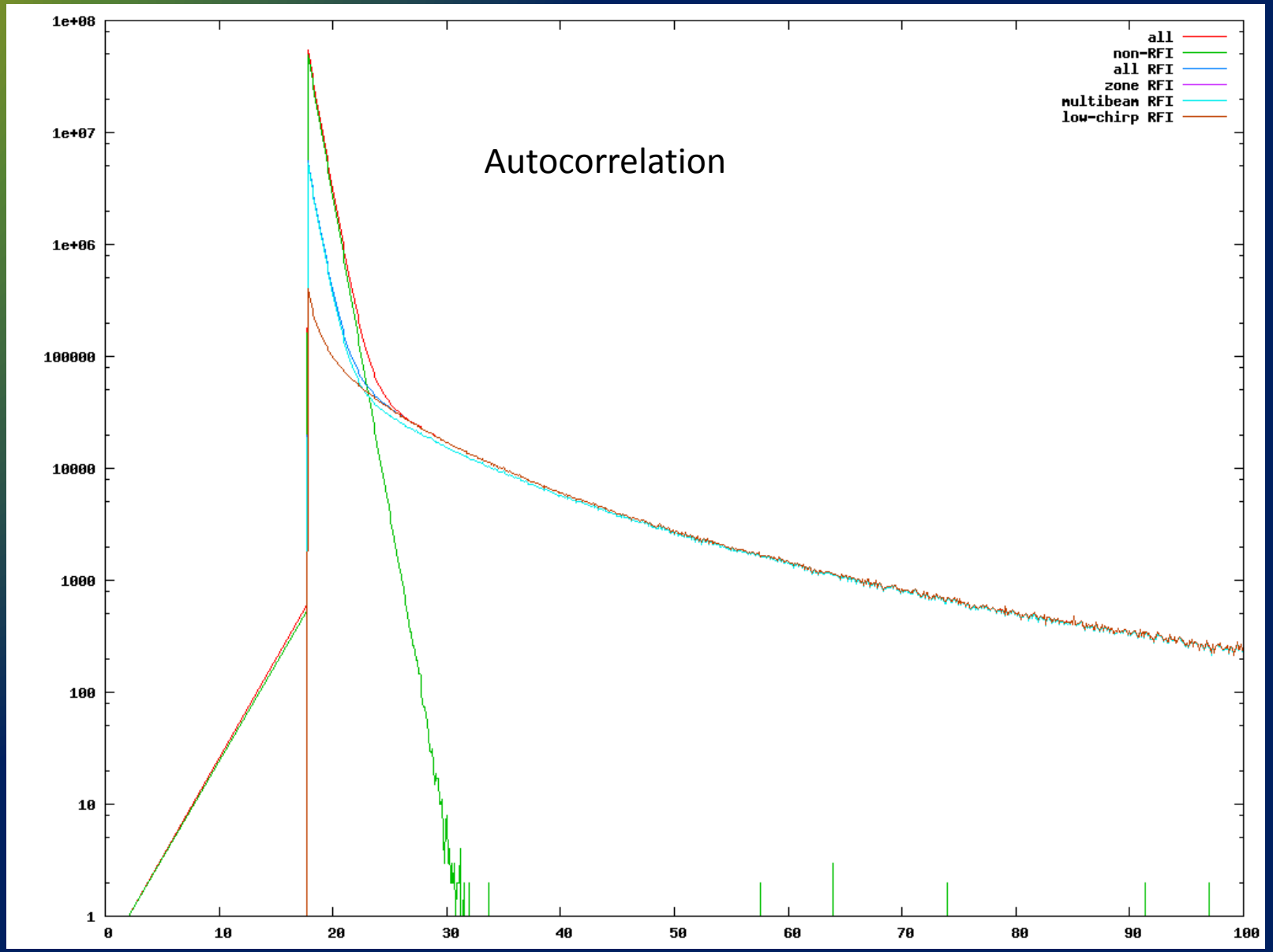


SETI@home Post Processing





SETI@home Post Processing



Candidate Identification

- Look for clustering (multiplets) in position and frequency
  - But must cluster in a different way than observing time does
- Current method:
  - Fixed frequency and position windows to detect multiplets
  - Multiplet score based on probability of random coincidence
  - Score each pixel on the sky according to its signal and multiplet content

Astropulse

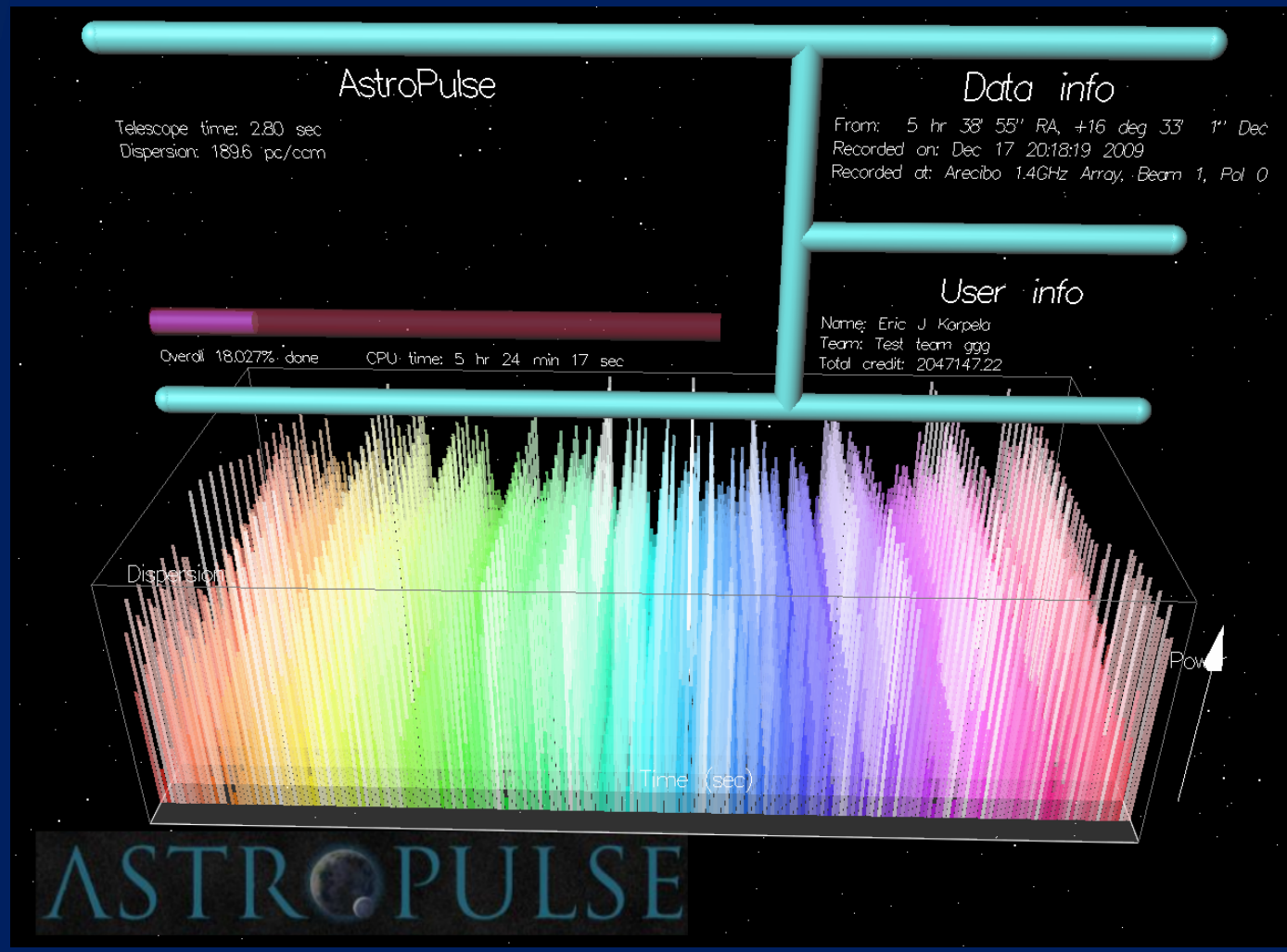
- “SETI@home for pulses”
- Coherent Dedispersion
- 2.5 MHz Bandwidth

(0.4 μs samples)

- 28416 DMs

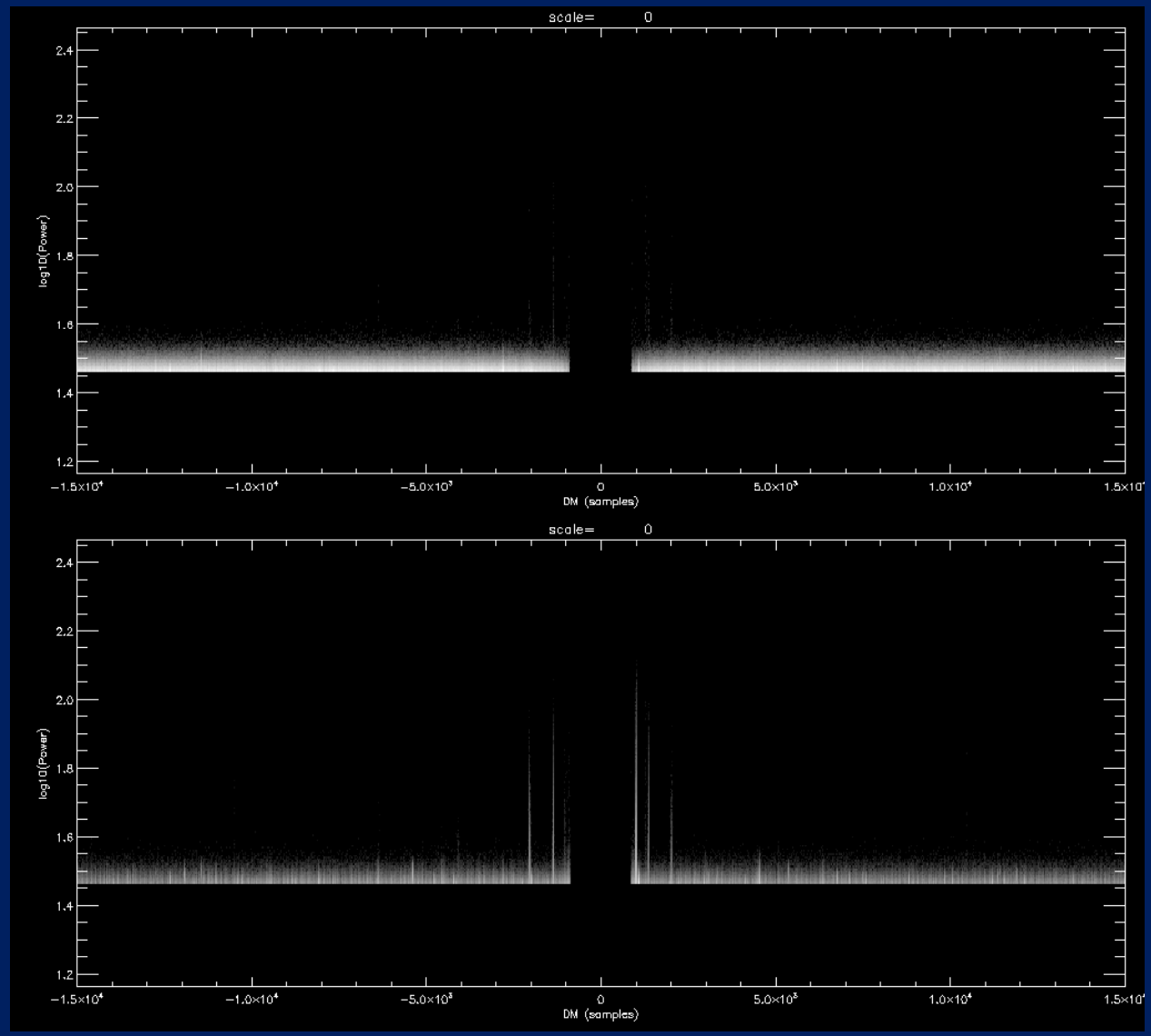
  54 cm^{-3} pc < | DM | < 1200 cm^{-3} pc
- Sensitivity ~ 54 Jy μs
- Repeating and single pulse searches
  - 10 octaves pulse duration

    (0.4 – 409.6 μs)
- 389,892,349 pulses in database
- 50% single, 50% repeating

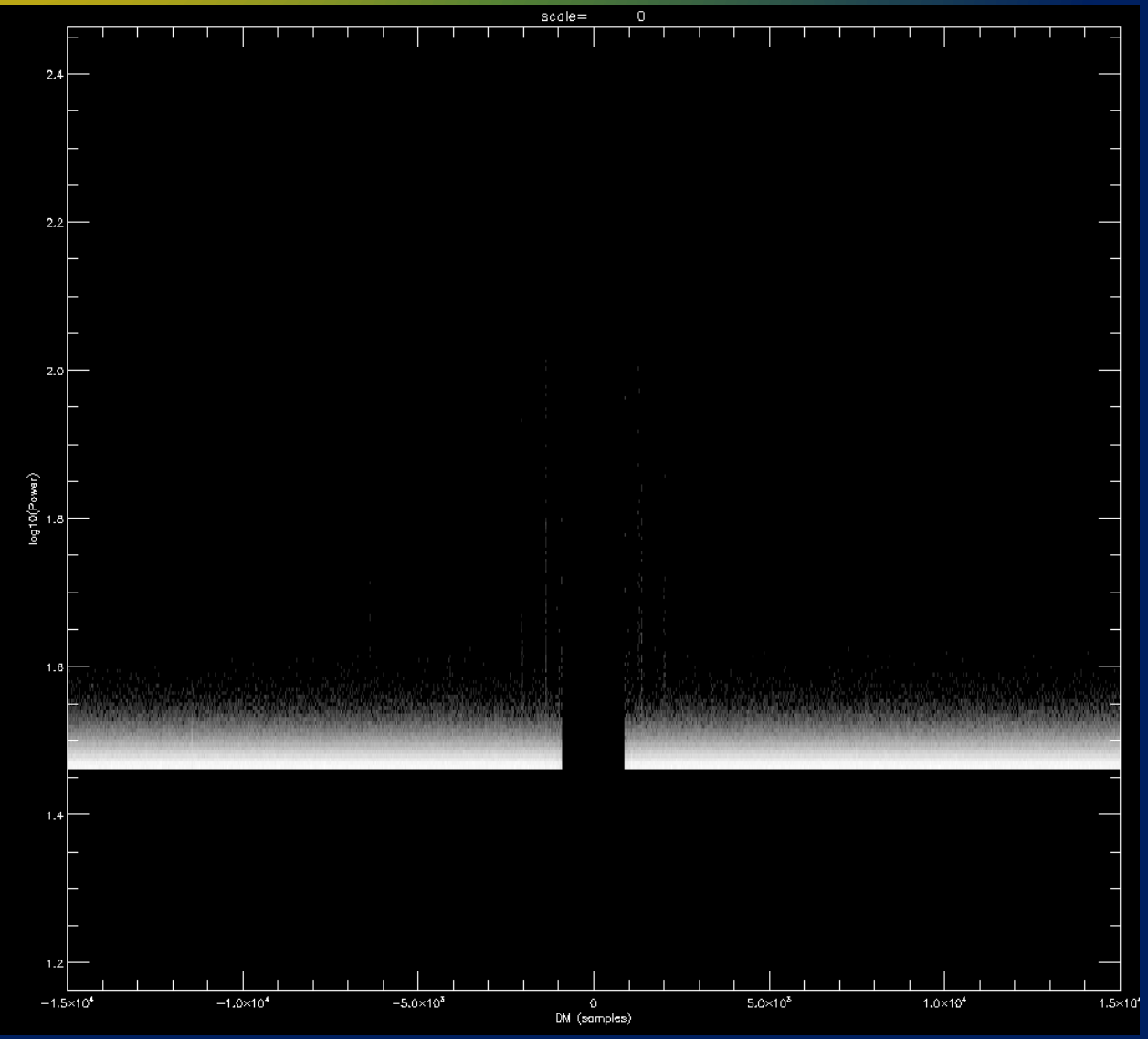


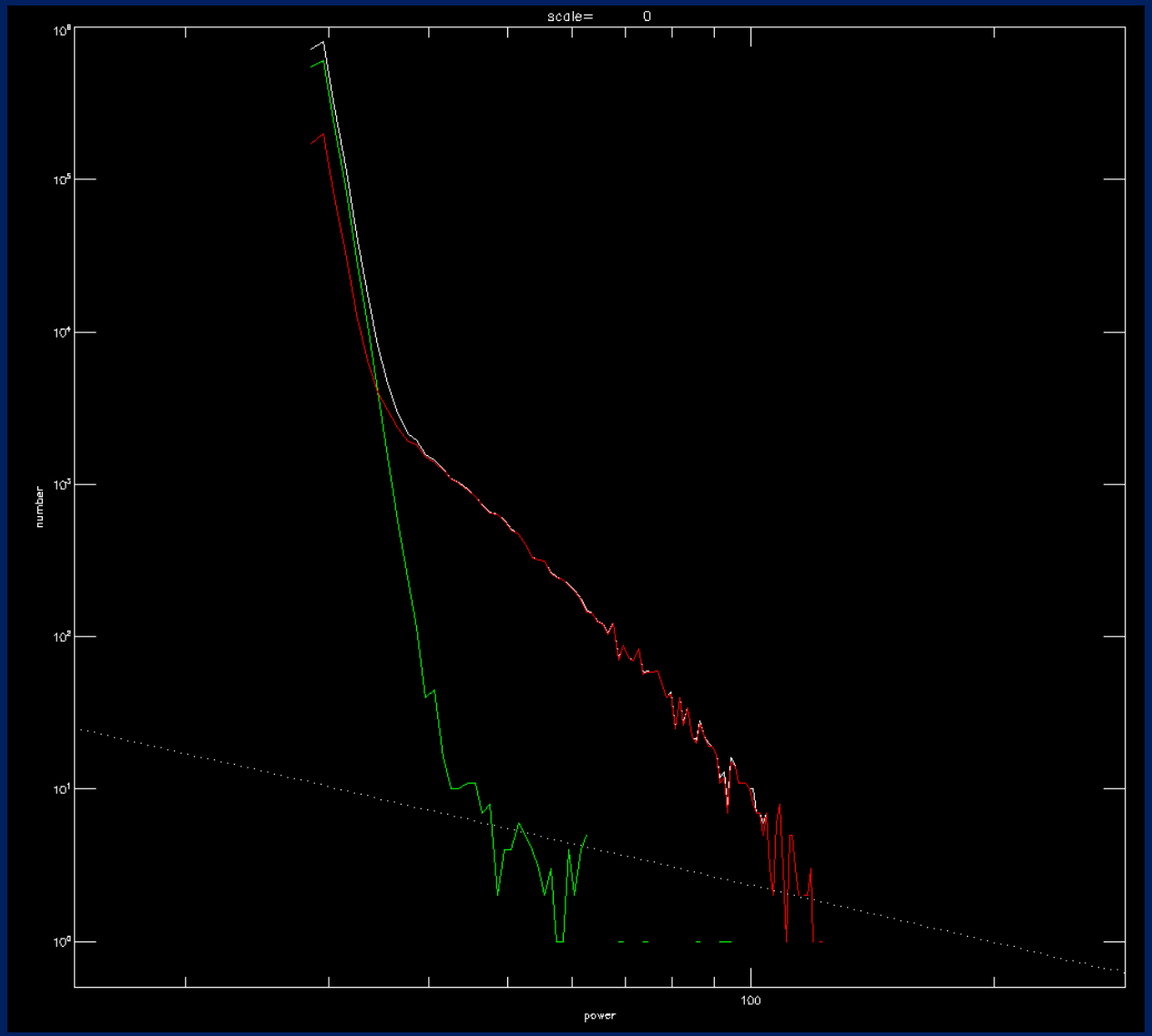
Astropulse Post Processing

- Radar blanking (in splitter and in client application)
- Frequency, time, or period zones
- Cross beam rejection
- Cross polarization confirmation

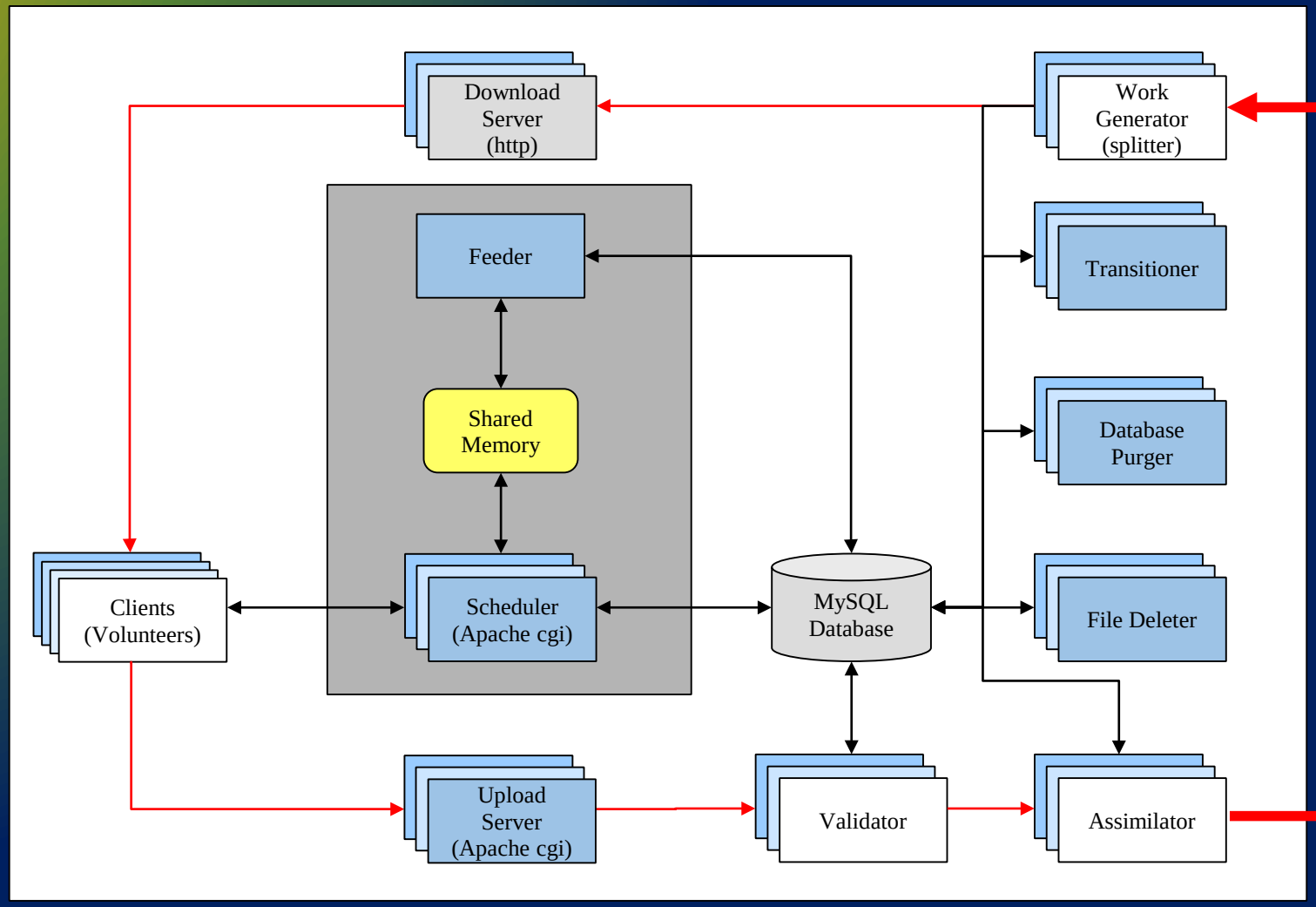


Astropulse Post Processing





BOINC - Berkeley Open Infrastructure for Network Computing



From Arecibo

Standard Open Source Software

Standard BOINC Components

Customized BOINC Components

To Nobel Prize Committee

BERKELEY SETI RESEARCH CENTER