# Non-imaging Processing

Ben Stappers, for a cast of many contributors

# Outline

- What is non-imaging processing?
- Where are the problems?
- Connecting the science to the engineering
- What are the processing challenges?
- What is the present state-of-the art?
- What are the potential solutions?
- What tradeoffs need to still be considered?
- A tool for specifications
- Uniboard – an fpga based approach.

# What is non-imaging processing

- A working definition seems to be anything that doesn't use the correlator.
- Alternatively it is anything that processes data that comes out of the central beamformer.
- It is therefore predominantly associated with the processing of data in the time domain.
- The time resolution required typically far exceeds what the correlator can deliver.
- In general it relies on more than one beam being available at one time.

# Defining the problem.

- One of the key aims so far has been to better describe the requirements.
- The parameter space is so large that it is hard to choose an optimal set.
- So far this has predominantly been from the point of view of pulsar and fast transient processing.
- Strongly influenced also by design choices, in particular:
  - Types and numbers of receptors
  - Size of the core

# Science Goals: Gravitational Waves

- Required: additional MSPs widely distributed on the sky
- Implies: all-sky survey for MSPs
- Frequency: 400 MHz AAs and/or 800 MHz dishes, balance spectra and propagation effects
- Maximum DM trials: Depends on if in plane or not.
- Survey speed: but cannot trade (A/T) for FoV

- Required: Higher precision timing of known & future MSPs
- Implies: pointed observations toward "timers"
- Sensitivity: Maximum possible
- Field of View:  Larger + Beams gives greater efficiency.
- Frequency: 0.8-3 GHz

# Science Goals: Tests of GR

- Required: additional ultra-relativistic binaries
- Implies: low- and moderate Galactic latitude survey for binaries
- Survey speed: but cannot trade (A/T) for FoV
- Frequency: 1.4 GHz (and ~ 2 GHz?)
- maximum DM trials: Depends on BW/in Gal. plane.

- Required: higher precision timing of known and future ultra-relativistic binaries
- Implies: pointed observations toward "timers"
- Sensitivity: Maximum possible required
- Frequency range: 1.4 GHz and 2 GHz

# Begin with use cases

- Use cases provide one way to narrow down specifications and this is what we have done.
- By considering specific scenarios we have been able to describe an end to end description of the requirements.
- This has been done for both searching with the sparse AAs and pulsar timing.

# Processing Challenges

- Beamforming – sparse Array
- Dedispersion
  - Incoherent for pulsar surveys
  - Coherent for high precision timing
- Folding (only for known pulsars)

# Processing Challenges

- Beamforming – sparse Array
- Dedispersion
  - Incoherent for pulsar surveys
  - Coherent for high precision timing
- Searching:  i.e. Fourier Transforming
- Acceleration processing
- Candidate identification

# Pulsar Search Implications

- Beams - > few thousand
  - Depends on A/T, survey time, core size
- Dedispersion -> few thousand
  - Depends on frequency / Area of sky
- Searching:  i.e. Fourier Transforming
  - Requires accumulation step of data!
- Acceleration processing – few hundred
  - Frequency domain vs time domain
- Candidate identification
  - AI, Neural Nets or other pattern recognition.

# State of the Art Searching

- Parkes Multibeam
  - 13 beam analog until 2009
  - 13 beam digital since 2009
- PALFA – Arecibo
  - 7 beam digital (wide-band) since 2010(ish)
- WSRT interferometer
  - Up to a few thousand beams (formed offline)
- LOFAR
  - Few hundred beams simultaneous
- In all cases no real time data reduction (yet)

# Pulsar Timing Implications

- Simultaneously process 10-100 beams
- Directly related to number of pulsars found (20,000/2000), FoV and A/T
- Estimates of few days to get through all sources
- Need different approach for high precision and "normal".
- Coherent dedispersion needed for high precision
- BW ~ 1GHz, Tres ~ 0.2 – 100 us, Tint > 120s

# State of the Art Timing

All coherent dedispersion based

- CASPSR
  - ~ 1 GHz bandwidth/Direct sampled/FFT/GPU
- GUPPI
  - ~ 1 GHz bandwidth/PFF/GPU
- ROACH – Jodrell & Effelsberg
  - 512 MHz BW/PFF/CPUs -- x2 to get 1 GHz.

- All capable of real time but no multi-beaming.

# Potential solutions

- More of the same
  - Problems with the amount of data to store (linked to time for survey/beams etc..)
- New generation computers, e.g. BG
  - Problems with accumulating data
- Possibly/probably need to do something in real time so need a streaming solution.
- GPUs, FPGAs, ASICS, Hybrids

# Things still to consider: Tradeoffs

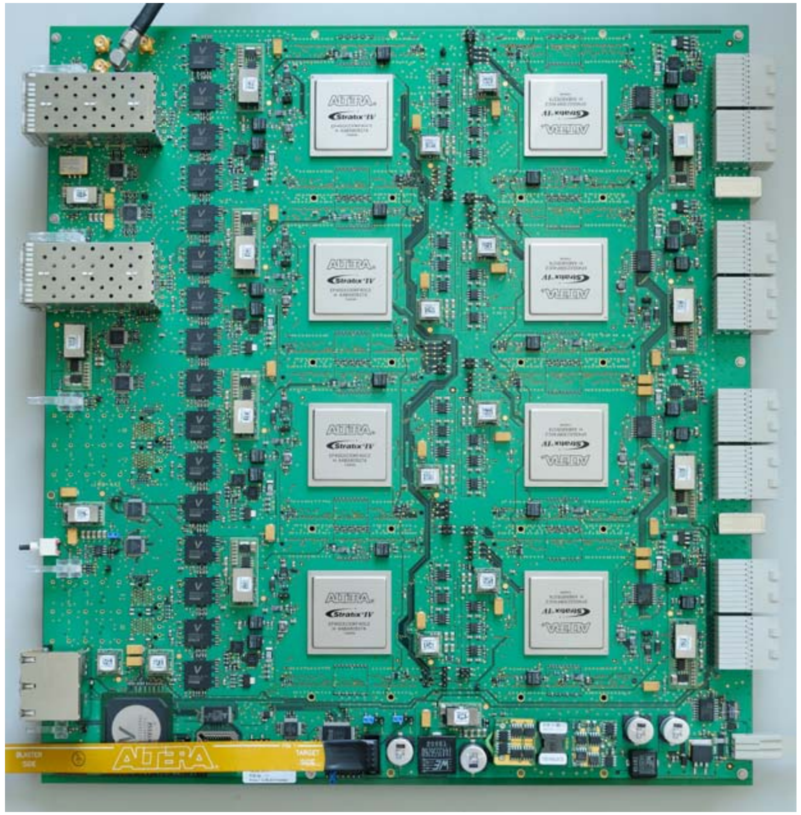
- Complex as many considerations
  - beams vs bandwidth vs integration time vs acceleration vs gain
  - number of bits, many places in processing chain.
  - channel widths vs time resolution
  - timing accuracy
  - polarisation purity

# Developing a tool

- Complex systems which are highly interdependent.
- Tool already includes:
  - Dish/AA size, number of elements
  - Size of core, size of stations
  - Observing frequency / BW
  - Dispersion measures
  - Data rates / processing
- Still needs linkage between all, acceleration, processing architectures.

# Uniboard for Pulsar Backend

**A RadioNet FP7 Joint Research Activity,**
9 partners




Multi-purpose, scalable, high performance, generic interfaces (10GE, DDR3)

Per board: 8 Altera Stratix IV FPGAs (40 nm), 2 × 16 × 10 Gbps, each front node to all back node mesh

Many applications, you'll here more about the board and the applications in the correlator section.

Consider here its application for high precision pulsar timing/coherent dedispersion and pulsar searching/incoherent dediseprsion.

# Uniboard: Pulsar Timing

- Take up to 1 GHz of BW and coherently dedisperse and fold it in real time
  - most demanding aspect is the requirement to perform large FFTs
  - Can use PPF to divide up the band BUT have limit of desired time resoln. and throughput
  - We have implemented a 2D-FFT method

***BW x log2(T x BW / N) <= Ffpga x P x B / 4***

For example if *T* = 1 second, *Ws* = 8 bits and *B* = 2 then:

| ***N*** | 1 | 2 | 4 | 8 | 16 | 32 | 64 | 128 | 256 | 512 | 1024 | 2048 | 4096 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ***BW (MHz)*** | 230 | 238 | 247 | 256 | 266 | 277 | 289 | 302 | 316 | 331 | 348 | 366 | 387 |

Since there are 4 rows of FPGAs on a Uniboard then the above bandwidths can be multiplied by 4 giving the following bandwidths for 1 Uniboard:

| ***N*** | 1 | 2 | 4 | 8 | 16 | 32 | 64 | 128 | 256 | 512 | 1024 | 2048 | 4096 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ***BW (MHz)*** | 921 | 954 | 989 | 1026 | 1067 | 1110 | 1157 | 1209 | 1265 | 1326 | 1393 | 1467 | 1548 |

# Uniboard Pulsar Searching

- Two main problems:
  - – dedispersion causes data size to expand
  - – Need to accumulate data if doing an FFT

$T_{obs}$ = 600s
Ws - word size
N - total number of channels
B - number of bands

Memory is the key
Processing cycles left over possibly use for acceleration?

If Ws = 4, *Nt* = 8192, *M* = 9.15E6, and *K* = 64 then *P* = 128:

| **Number of UNIBOARDs** | 1 | 2 | 4 |
|---|---|---|---|
| **B** | 8 | 16 | 32 |
| **N** | 1024 | 512 | 256 |
| **External memory size (GB)** | 13.1 | 6.6 | 3.3 |
| **De-dispersion Computation time (sec)** | 562.5 | 281.3 | 140.7 |
| **FFT Computation time (sec)** | 16.9 | 8.5 | 4.25 |
| **Viability** | No | No, not enough memory | **Yes** |

# Status of Project

- Algorithm in place for coherent dedispersion, currently working on how best to incorporate it within the data flow model.
  - Test set up using similar Altera fpga on a board to go in a PC almost complete
  - Boards @ ASTRON (see AS talk)
  - Implementation complete by mid-Autumn
- Methodology and sizing just completed for the pulsar searching

# The way forward

- See summary talk later this session.