New Zealand’s Contributions Toward the SKA Computer System

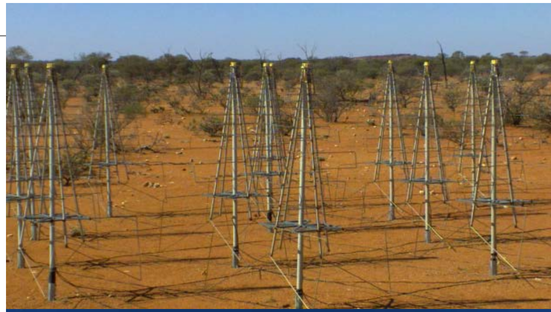
DR ANDREW ENSOR

DIRECTOR HPC RESEARCH LABORATORY/ DIRECTOR NZ SKA ALLIANCE

COMPUTING FOR SKA COLLOQUIUM 2016

11 FEBRUARY 2016

The Square Kilometre Array Project



Will be World's largest and most powerful radio telescope

Actually four types of telescope arrays:

- Low (sparse aperture array and later also dense aperture array)
- Mid (single pixel dishes)
- Survey (phased-array feed receivers)

Cost €650M for SKA phase 1, estimated €2- €6billion for SKA phase 2

Approximately 35-40% budget for computing hardware and software

Preconstruction design for SKA1 commenced Nov 2013

SKA1 design due to finish 2017

SKA1 construction 2018-2023

SKA Phase One

SKA design/build in phased approach

- SKA phase one represents 10% system, approx. 1% total compute

Dual site decision for SKA1

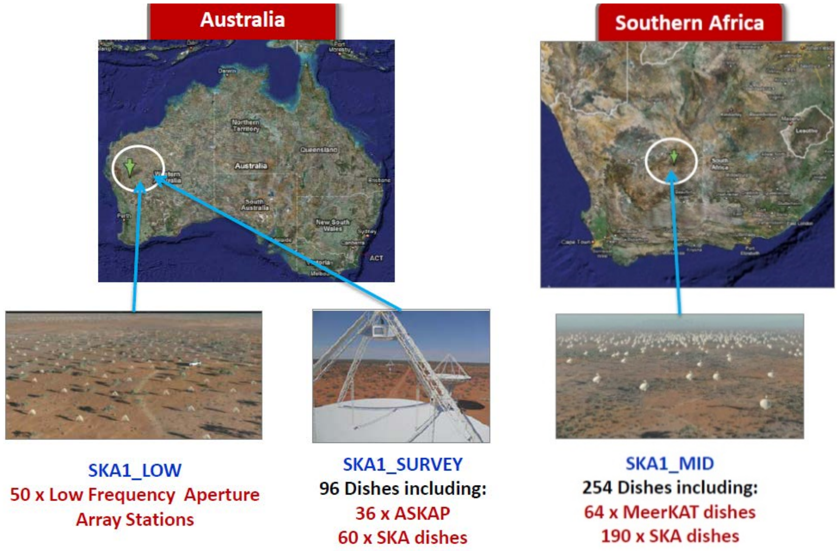
- Australia (Low) and South Africa (Mid)

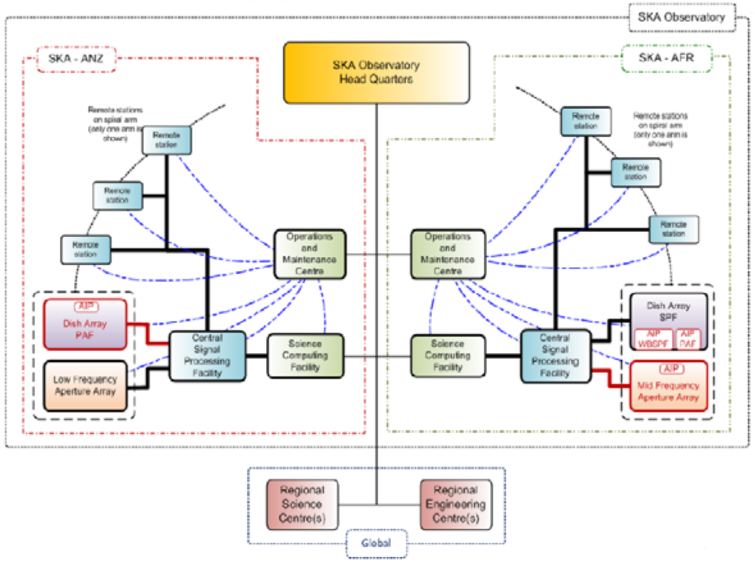
Two processing systems

International Work Package Consortia formed for RfP response for SKA1 design

Computing-Intensive work packages:

- Central Signal Processor (CSP) Consortium lead by NRC Canada and MDA Corporation, about 40 FTE
- Science Data Processor (SDP) Consortium lead by University of Cambridge, about 55 FTE





SKA Phase One Mid

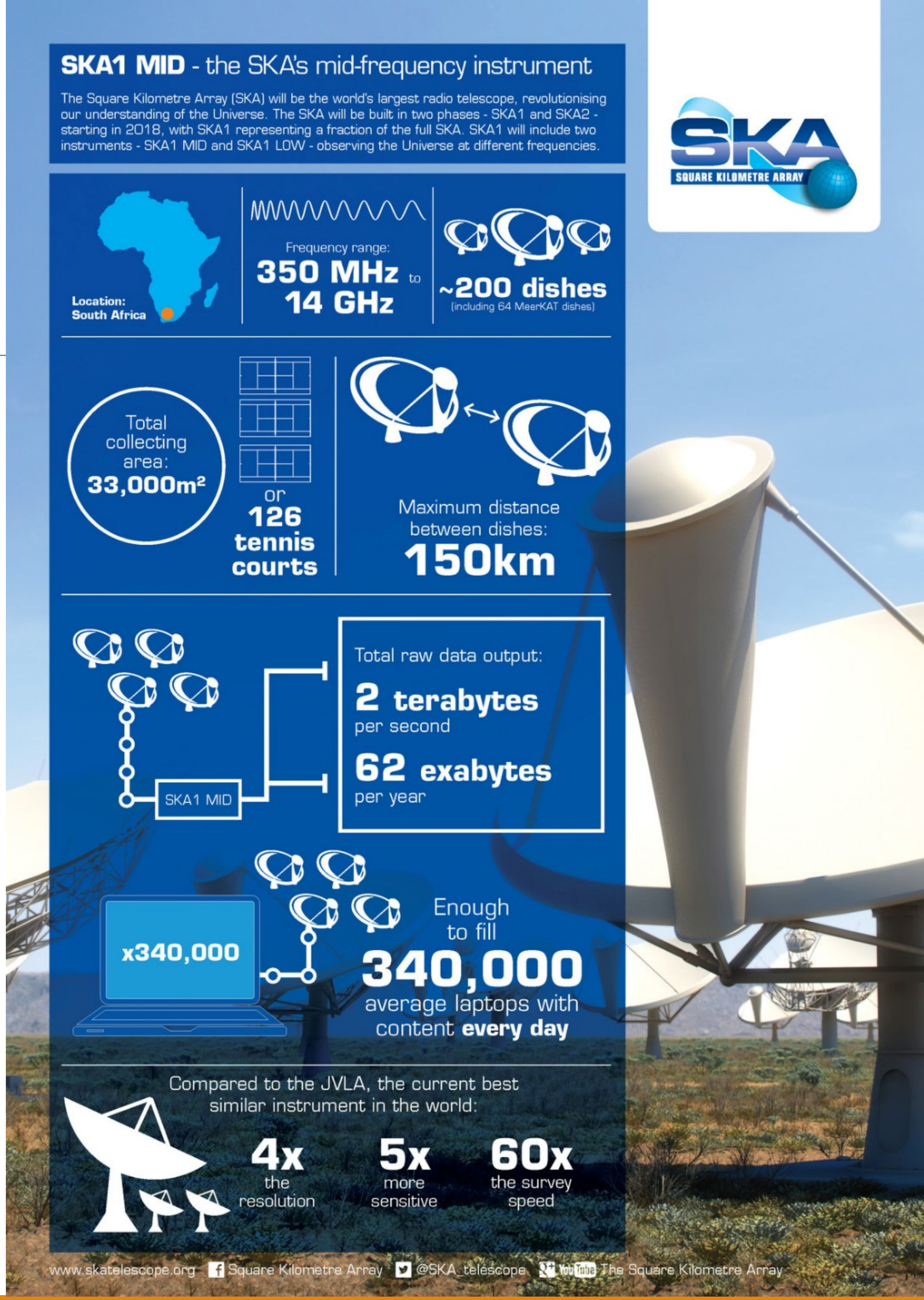
SKA1-Mid will be in South Africa

Particularly for pulsar observations, moderate red-shifted hydrogen, high sensitivity imaging

Expands on the MeerKat 64 dish array with another 133 offset Gregorian 15m dishes up to 75 km radius

Each dish produces and sends around 10 GB/s to Central Signal Processor

Will be expanded by order of magnitude in SKA phase two



SKA Phase One Low

SKA1-Low will be in Western Australia

Particularly suited for highly red-shifted hydrogen

Primarily for Epoch of Reionization investigation, looking far back in time to the early universe

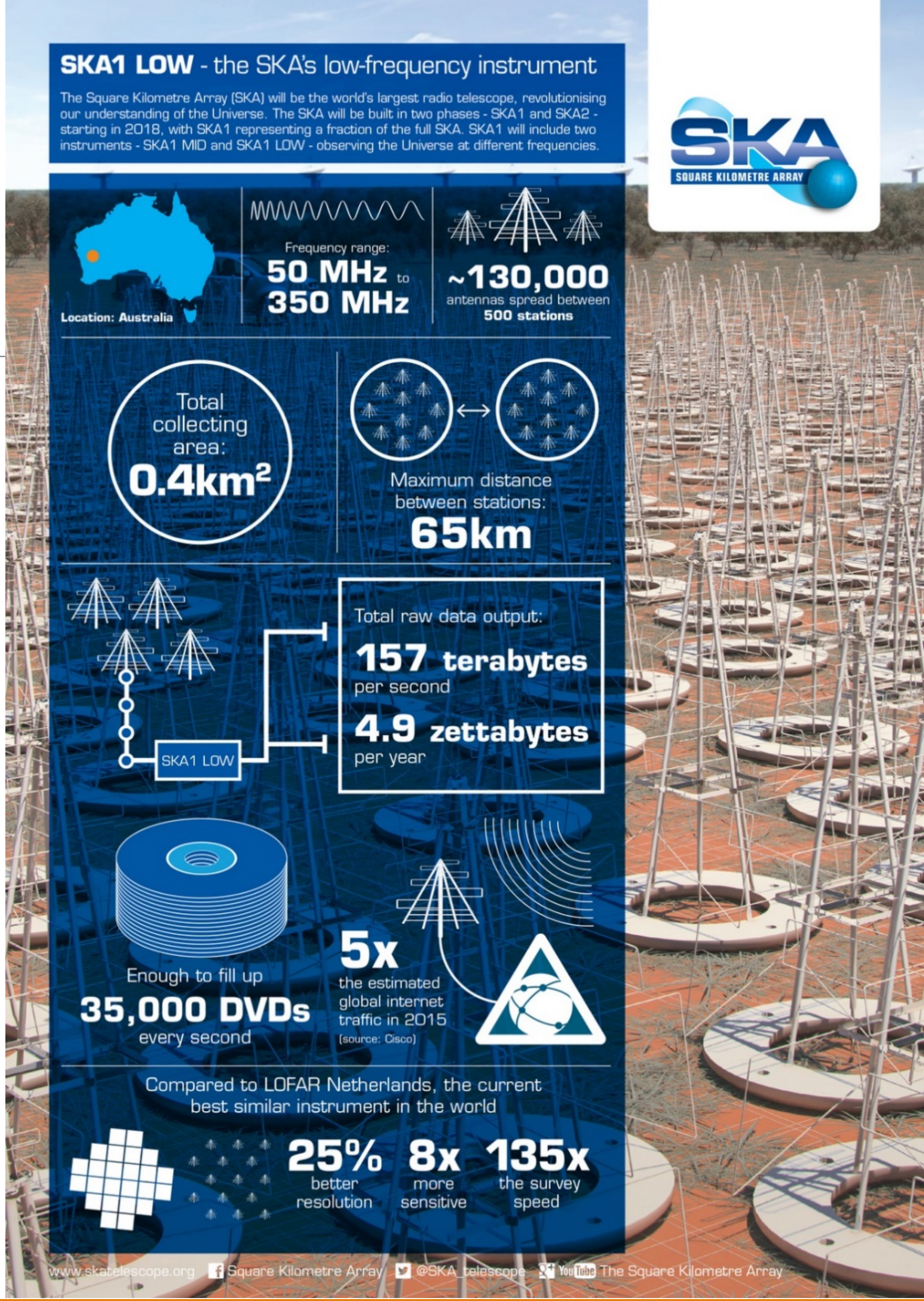
Test station recently set up in Murchison, Western Australia

512 stations, each with 256 fixed dipole antennas

Each station produces 256 GB/s raw data

Each station sends around 1 GB/s to Central Signal Processor

Most stations within 2 km radius, some out to 35 km along spiral arms



Why is SKA of interest in Computing?

Besides being the biggest mega-science project in the world, it’s the biggest BIG Data problem in history

SKA Phase 1 must grapple with BIG Data on a scale that ICT Industry will be facing in 5-10 years time

SKA Phase 2 will be an order of magnitude BIGGER yet

Mega-science projects attract clever people to innovate, give many spin-offs, help steer technology evolution

High potential for producing technology disruption, new IP

Large international collaboration, multinational companies keen to be involved

NZ SKA Alliance Organisations

AUT University

Catalyst

Compucon NZ

Massey University

Nyriad

Open Parallel

University of Auckland

Over thirty academics and industry partners in NZ working on SKA design for Central Signal Processor and Science Data Processor

Focused on computer system design for SKA

One of largest SKA computing groups worldwide

Very strong international collaborations

Key Parts to SKA1

Telescope receivers

Receiver beamformers (some NZ involvement)

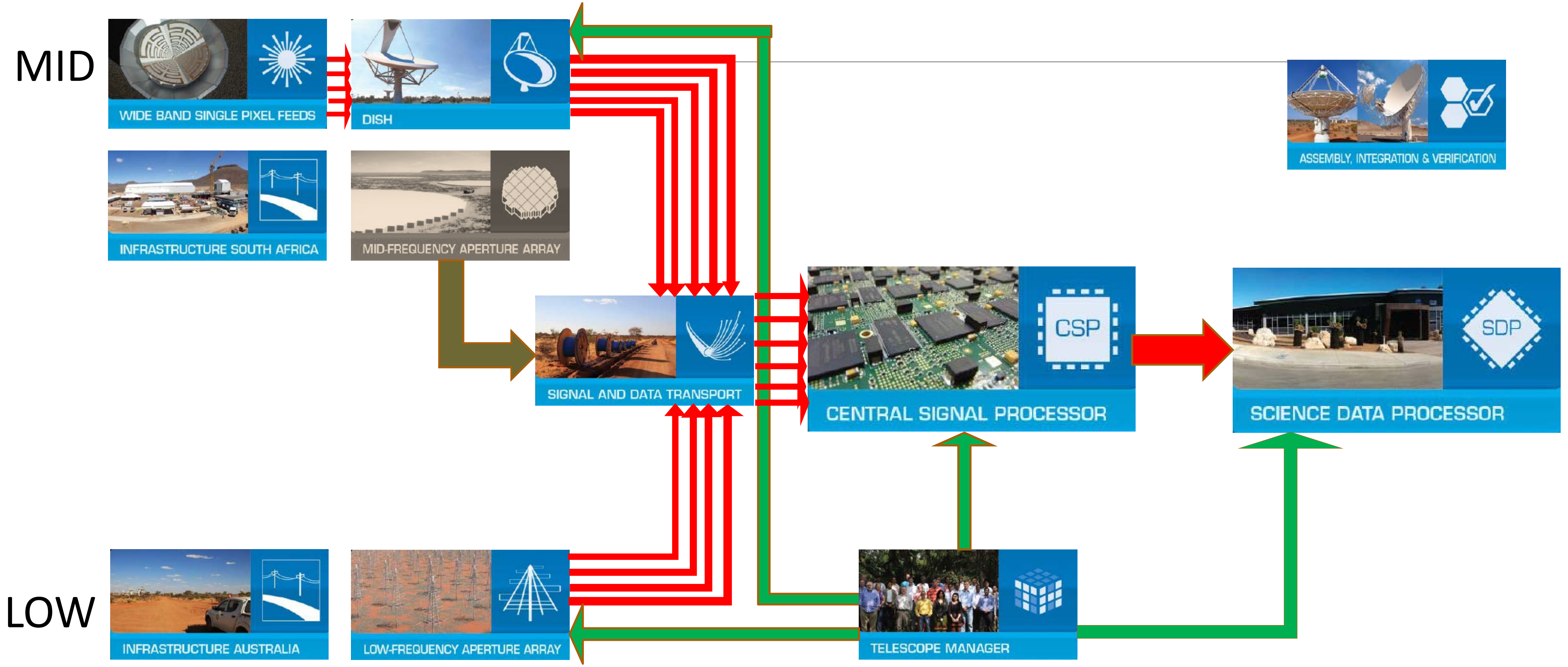
Signal transport

Correlators (large NZ involvement)

Pulsar Search (NZ involvement)

Imaging Pipeline (large NZ involvement)

SKA Phase One Data Flow and Consortia Teams



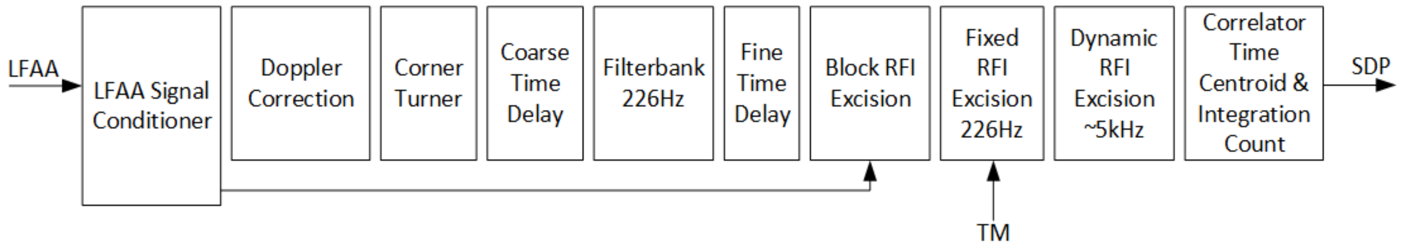
Correlator Overview

Correlator responsible for combining together signals from each receiver in real-time

Four main roles:

- Channelize wideband data from MID dish or LOW station
- Correlate and accumulate each pair of channelized data (very heavy compute)
- Beamform channelized data for Pulsar Searching/Timing
- RFI analysis/flagging

Pipeline has subtleties and enormous data rates



Correlator Challenges

Exascale Big Data project:

- Large Hadron Collider expected to produce about 15 PetaByte data per year
- Correlators will produce about 40 PetaByte per hour

Unprecedented I/O challenge:

- Each telescope array requires up to 8 TeraByte/s channelization and correlation and outputs up to 4 TeraByte/s (Cray Titan I/O about 1.4 TeraByte/s total)
- Data throughput a major concern

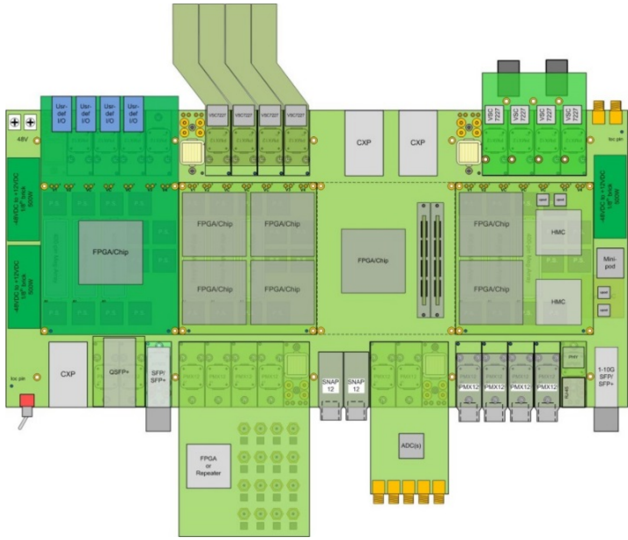
Green computing requirements:

- Low array will be deep in Australian outback, no power grid so energy consumption an overriding concern

Correlator Designs

MID Correlator developing new high data throughput board design called PowerMX

- Open specification at www.powermx.org
- Motherboard currently being prototyped, features four mezzanine sites
- Design underway for Arria 10/Stratix 10 and MPPA mezzanines
- Joint NRC Canada and NZA design



New processor interconnect comm standard under development:

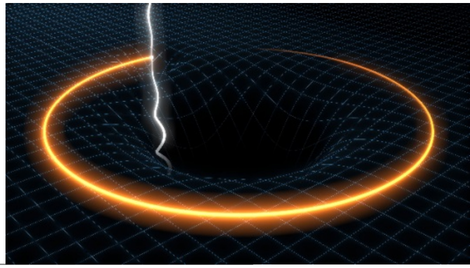
- AXI over Ethernet
- NZA design

LOW Correlator

- Single-site boards with Xilinx Virtix Ultrascale+ 16nm and HMC memory
- Full optical interconnect
- Joint CSIRO, ASTRON, NZA design

ASIC correlator design work at JPL and at Massey University

Pulsar Search Overview



Pulsars will be used for many key science experiments in SKA1

- Need to discover many new pulsars

Task to find weak signals from unknown pulsars deeply buried in noise

- Ultimate needle-in-the-haystack search

Simultaneously search 2000 beams

Each beam has 4000-15000 channels, sampled every 50μs (700Gsample/s total) for 7 min

Pulsar Search Challenges

Unknown pulsar(s) so don’t know pulse period

Signal suffers frequency-dependent dispersion as travels through interstellar matter

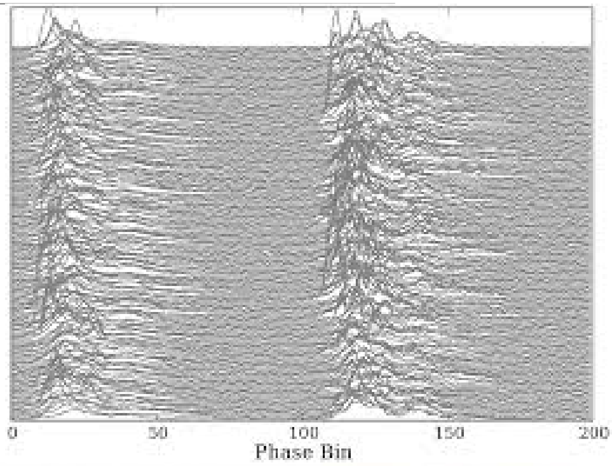
- Time delay proportional to $1/f^2$ but unknown scale due to matter density

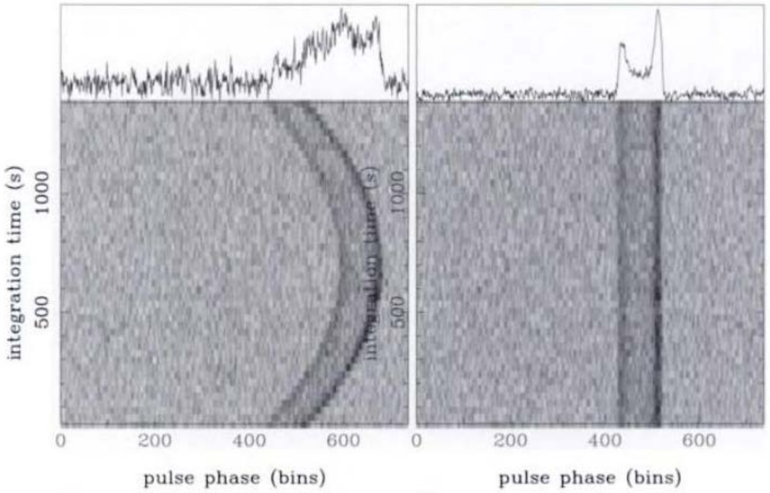
Really want pulsars orbiting eg black holes so pulse has Doppler shift

- Pulse period changing

Three-dimensional search space amongst noise:

- Unknown pulse period
- Unknown dispersion
- Unknown source acceleration





Pulsar Search Designs

Apply many dispersion measures and add channels together for each to increasesignal:noise

- Repeat search for each potential dispersion measure
- Most faint pulsars will still be buried in noise even after dedispersion

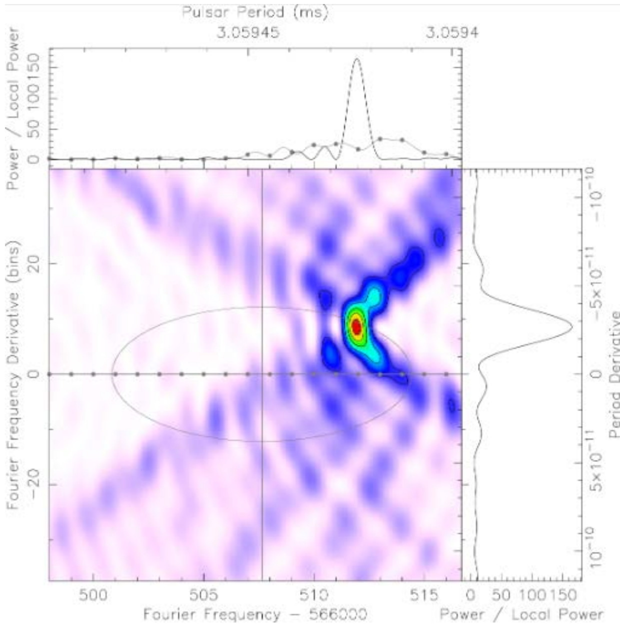
Apply FFT and perform harmonic summing

- Fainter pulsars hopefully rise above noise

Two alternative approaches for finding accelerating pulsars:

- Time domain acceleration search: guess time domain resampling and repeatedly search
- Frequency domain acceleration search: matched filtering by correlating templates

Prototyping on GPU and/or FPGA (OpenCL)

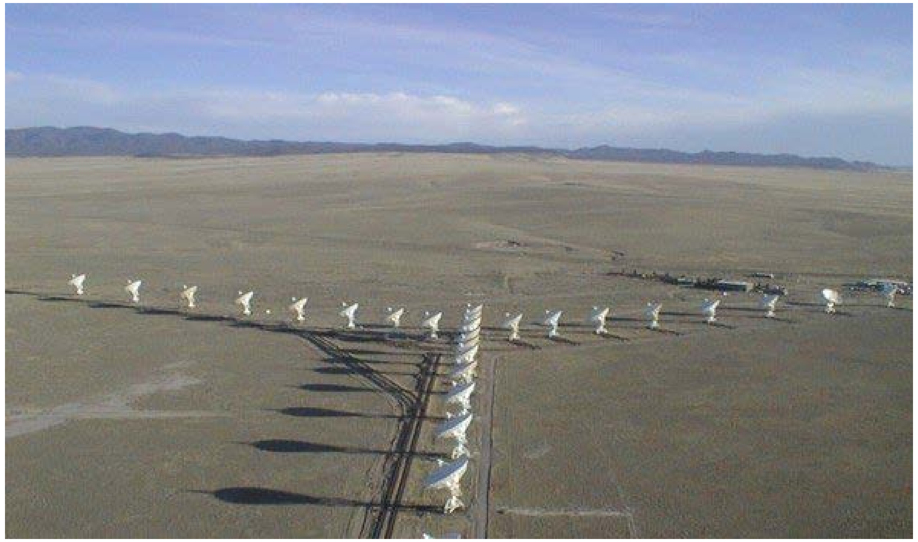


Imaging Overview

Correlator output for each pair of receivers called a visibility, is sample of sky image in Fourier plane at some (u,v) given by baseline

Want good coverage of Fourier plane:

- Collect many visibilities
- Grid visibility data
- Inverse Fourier transform
- Deconvolve effect of point spread function



Imaging Challenges

Gridding and (inverse) FFT are enormous compute challenges

Using up to 150kx150k grid and including each of up to 2^{16} channels in the imaging pipeline requires supercomputing performance:

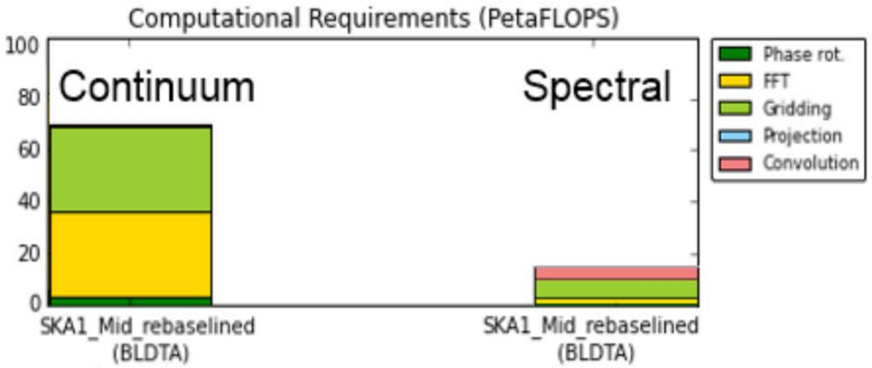
- SDP estimated to require 80 PetaFLOP for Mid in SKA1 (Cray Titan is 17.59 PetaFLOP), within reach by 2018 but everything must be done in real time 24-7
- Power requirements determine ultimate costs, SDP computing need to fit within 5-10 MW (NZ entire installed capacity about 10GW, more renewable sources than Australia/South Africa)

SKA Phase Two will increase computing demands by approximately 100 fold

Imaging Design

UV coverage not uniform, is Fourier transform of sky image multiplied by impulse-response of array

- Grid visibilities (NZA heavily involved)
- Apply inverse 2D FFT (NZA heavily involved)
- Apply deconvolution
- Repeat via Major Cycles



SDP supercomputer primarily needed for Imaging Pipeline

- COTS-based but potentially custom Compute Node design
- Highly parallel computations
- Independent Compute Islands
- Each Compute Island has Compute Nodes with hardware accelerators (CPU, GPU, FPGA)
- Runs 24/7 so power is a major driver

Previous Contributions of Note

Bit precision savings in correlators to reduce visibility data volumes

Data compression savings for DISH/LFAA->CSP and CSP->SDP

Successful CSP multicore single stage channeliser prototype

ASIC correlator and beamformer development

Low power SDP hardware alternatives (ARM, multicore, FPGA)

Gridding investigation

Software Development Environment for CSP

Software Developments and Considerations for SDP

Survey correlator and CSP element costings for BDR

Survey correlator PIP

Current Contributions

More FFT prototyping to reduce power for CSP and SDP

Large effort on Mid/Low correlator designs and prototyping

- Mid correlator X-step FPGA and multicore implementation
- Low correlator contributions

Responsible for CSP end-to-end signal processing modelling and thermal modelling

Project management responsibility for CSP Low

Imaging Pipeline role

- Imaging Pipeline end-to-end model
- Precision analysis (eg single versus double precision)
- Comparison of Imaging techniques and parameters

Continue SDP Compute Platform design, prototyping and specification

- Particularly role of hardware accelerators and HPC software
- Middleware, OpenStack, Software Development Environment
- Continue working closely with industry

Increase SDP prototyping to help consortium needs

- Become an Open Architecture Lab?

Working toward an NZ Big Data Centre of Technology Excellence

NZ Benefits of Involvement in SKA

SKA represents some firsts for NZ

- Biggest BIG data project in the world
- Largest Science project NZ ever substantially involved in
- Possibly one of largest NZ involvements in international ICT collaboration
- Long-term multiway academic-industry collaboration

Early exposure to new ICT technologies, builds capability important for SKA1 tender

SKA very high profile

- Good for attracting talented people as capable people drawn to interesting work
- Encourages science/technology/engineering/maths students

NZ on world stage

- Fosters international relationships
- Foot in door for future big projects

Building a system that will answer fundamental questions about Universe over next 50 years

Questions?

Thank You

Andrew Ensor

Director HPC Research Laboratory/ Director NZ SKA Alliance

andrew.ensor@aut.ac.nz +64-9-9219999 ext8485