A GPU-based Real-time Software Correlation System for the Murchison Widefield Array Prototype

Randall Wayth — Harvard-Smithsonian Center for Astrophysics.*
Lincoln Greenhill — Harvard-Smithsonian Center for Astrophysics.
Frank Briggs — Research School for Astronomy & Astrophysics, Australian National University.
*Present address: Curtin Institute for Radio Astronomy, Curtin University. r.wayth@curtin.edu.au

http://www.mwatelescope.org

Abstract

At the heart of all interferometric radio telescopes is the correlator, which cross-correlates signals from all antennas. To support the testing and development of the MWA 32 antenna prototype system, a GPU-based correlator was developed and deployed on site while the full hardware correlator was still under development.

We evaluated several design options. Optimal performance for the correlation task is achieved by using the GPU shared memory to minimize redundant reads from GPU global memory[1].

Our GPU-based correlator outperforms a single-threaded CPU equivalent by a factor of 60, including all overhead, and runs on a single PC, which simplifies networking and synchronization issues compared to a distributed software correlation system.

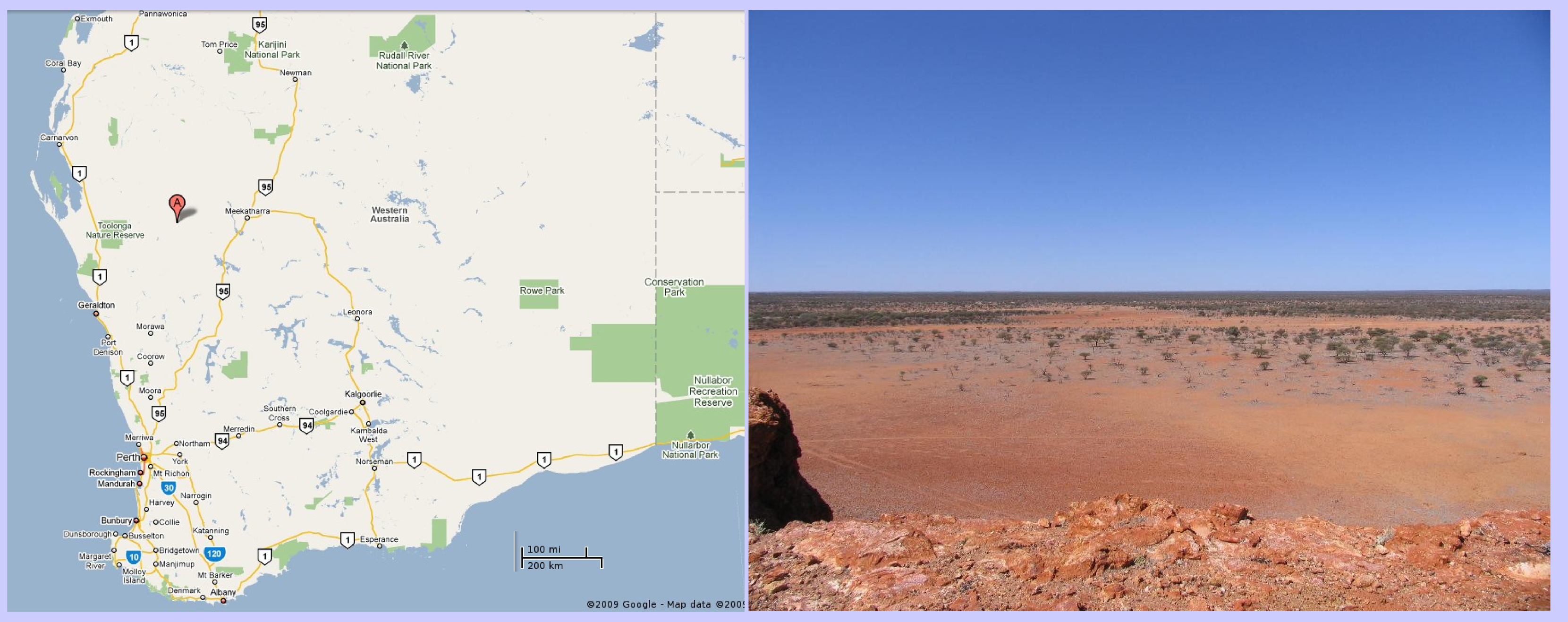


Fig 1: The site of the MWA, in the remote Murchison shire of Western Australia.

About the MWA

- It is a new, low frequency (80-300MHz) radio telescope under construction in the remote Murchison shire of Western Australia (see Fig 1).
- Presently the prototype array consists of just 32 antenna tiles (see Fig 2). The full array will consist of 512 tiles spread over an area of approximately 3 square kilometers.
- The MWA's key science goals include: detection of the power spectrum from the epoch of reionization (EoR), studying the Sun, heliosphere and ionosphere, and finding and studying variable and transient radio sources.
- The full MWA will generate data at a rate that is too large to be stored for offline processing. The data must be processed in real-time on a supercomputer. GPUs will be used for the real-time data processing system of the full array (see below, and talk by Richard Edgar).
- The frequency band of interest is the VHF band, which is heavily utilized by TV and radio stations, aircraft transponders and communications satellites. The remote site was chosen for its excellent radio-quietness, being very far from population centers.
- The MWA is a partnership between US, Australian and Indian institutions.

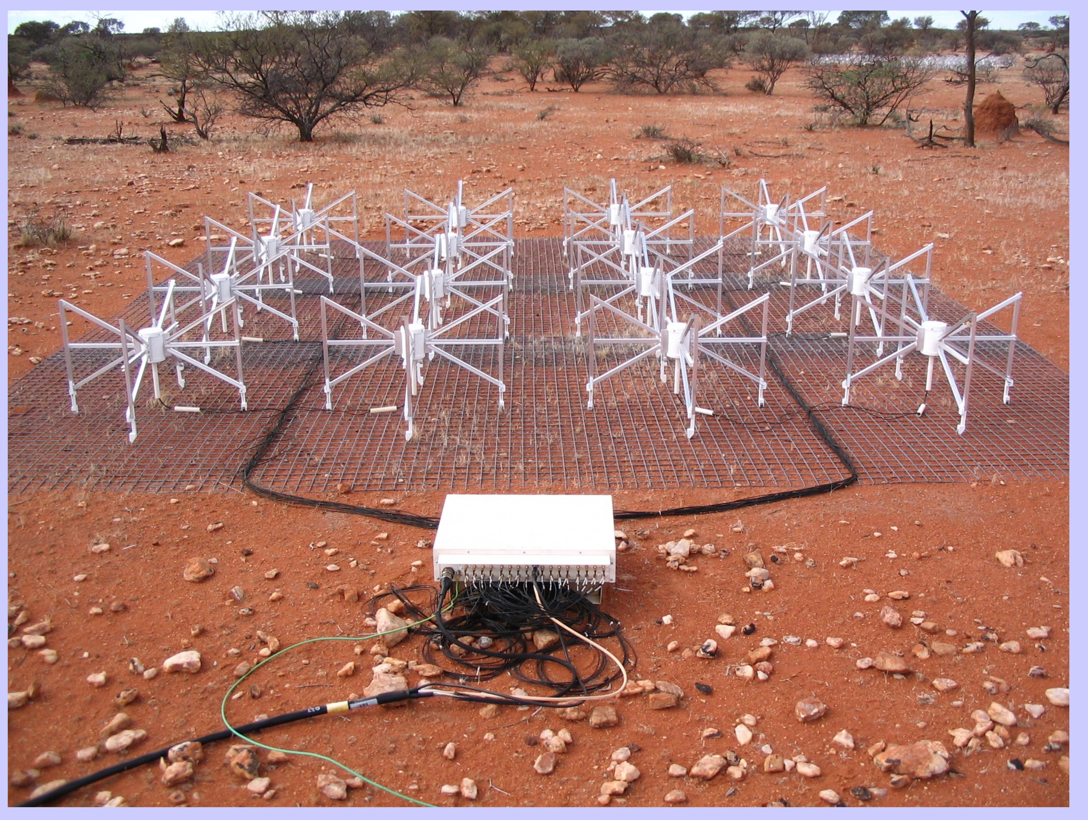


Fig 2: An MWA antenna "tile", consisting of 16 dual linear polarization dipoles and an analog beamformer

Correlation:

Signal correlation for an `FX' style correlator in radio astronomy consists of three basic steps, as shown in Fig 3. The data streams arrive ordered by antenna, then time and are usually represented as integers with only a few bits.

1. Unpack the data into floating-point format and reorder it so that samples for a single antenna are grouped together for channelization.
2. Channelize the data, which is usually performed via FFT, but more sophisticated schemes using polyphase filter banks are possible.
3. Form the correlation products by multiplying each antenna with all others (parallel by frequency) and accumulating the results.

Details of design options for correlators can be found in [1] and [2].

The number of samples per second into the correlator scales linearly with the signal bandwidth and number of antennas, N_{ANT}. The number of complex multiply and accumulate (CMAC) operations *per sample* scales as log2 (N_{CHAN}) for the FFT operation, and as approximately $N_{ANT}/2$ for the cross-correlation. Hence, cross-correlation dominates the arithmetic operation count for large (>10) N_{ANT}.

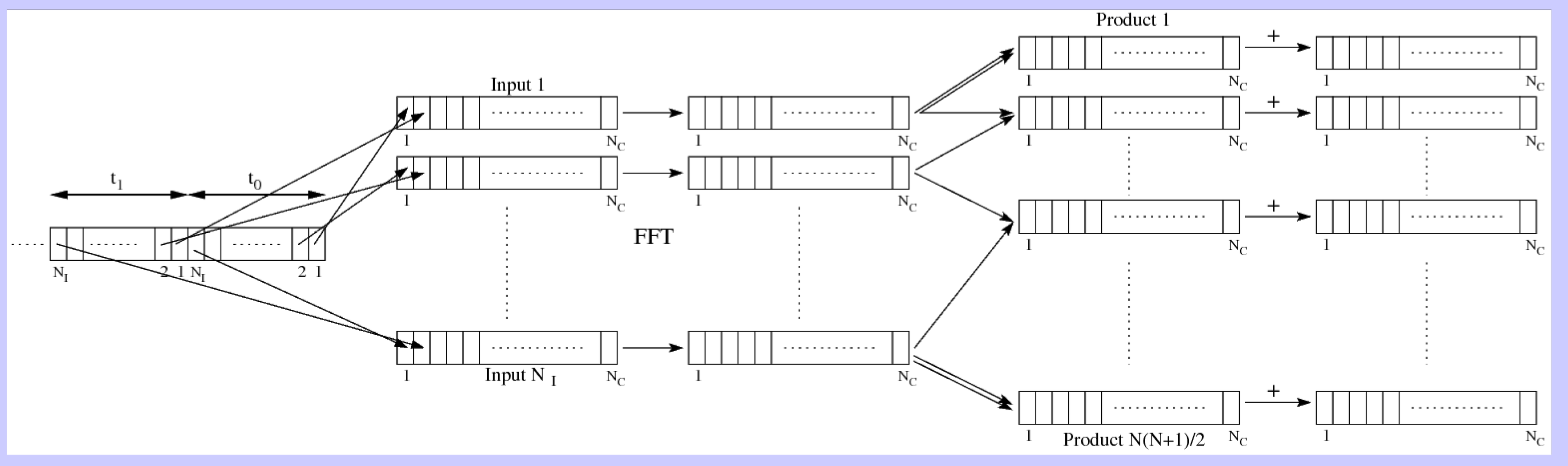


Fig 3: Data flow for an FX correlation system.

Performance

We compare execution times for a single threaded CPU version of a correlator, running on a workstation with a dual-core AMD Opteron 1214 to an NVIDIA C1060 GPU, using our best performing design. The comparison was for 1.28MHz bandwidth data and 64 "antennas" (32 tiles X 2 polarizations). Times given are to process 1 second's worth of data.

Task	CPU(ms)	GPU(ms)
Stage data	117	117
Transfer to GPU		57
Unpack data	1276	53
FFT	1840	53
Cross-correlate	20950	185

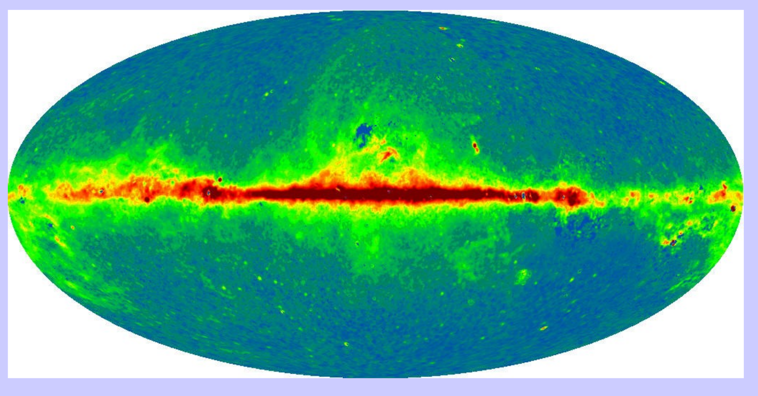
The GPU correlator has other advantages besides raw speed compared to the CPU version. Data can be staged and transferred to the GPU in parallel to the cross-multiply task. In addition, since the bulk of the processing work is done on the GPU, the host PC is lightly loaded, which allows it to perform networking and staging tasks optimally.

GPUs for the MWA real-time data processing system

The full 512 tile MWA system will have a purpose-built hardware correlator, generating correlation products at a rate of several GB/s. This data rate is too large to store, hence the MWA will be one of the first "next generation" radio telescopes to have real-time calibration and imaging system.

The basic steps of the MWA's real-time system (RTS) are quite well suited to GPUs, as outlined below.

- Calibration: scalar multiplication, inner products and discrete Fourier transforms on large vectors of complex numbers.
- Gridding: accumulating weighted convolution kernels in a grid. Analogous to "splatting".
- Imaging: 2D FFTs, scalar multiplication of images
- Reprojecting: interpolation of an image with distorted coordinate frame onto a regular grid, conserving flux. This operation might simply require bilinear interpolation, which is implemented very quickly in GPU hardware.



References:

[1] Wayth, Greenhill & Briggs (2009). PASP 121, 857. Preprint: http://adsabs.harvard.edu/abs/2009PASP..121..857W
[2] Harris, Haines & Staveley-Smith (2008). ExA 22, 129. Preprint: http://adsabs.harvard.edu/abs/2008ExA....22..129H
[3] Thompson, Moran & Swenson. *Interferometry and Synthesis in Radio Astronomy, 2nd edition.* 2001, Wiley. ISBN : 0471254924