



## GPU Based Wideband Back-end for GMRT

S Harshavardhan Reddy

Sanjay Kudale

Nilesh Raskar

Ajith Kumar B

Irappa M. Halagali

Shelton Gnanaraj

Digital Back-end Group

Swinburne University, Australia

NVIDIA, India

**Objective :** To provide technical details, test results of GPU-based wideband back-end for Giant Metrewave Radio Telescope.

Revision	Date	Modification/Change
Ver.1	25 December 2013	Initial Version

**1. Introduction :** The GMRT consists of an array of 30 antennas, each of 45 m diameter, spread over a region of 25 km diameter, and operating at 5 different wave bands from 150 MHz to 1450 MHz. The maximum instantaneous operating bandwidth at any frequency band is 33 MHz. Each antenna provides signals in two orthogonal polarizations, which are processed through a heterodyne receiver chain and brought to the central receiver building, where they are converted to baseband signals and fed to the digital back-end consisting of correlator and pulsar receiver. The existing GMRT Software back-end (GSB) is built on software based approach designed from off-the-shelf components, PCI based ADC cards and a Linux cluster of 48 nodes with gigabit inter-node connectivity for real-time data transfer requirements.

GMRT is upgrading to uGMRT and the back-end systems are undergoing major changes to achieve the upgrade system specifications like increased bandwidth of 400MHz, direct processing of RF signals, increased dynamic range, improved channel resolution. As part of it, a 4-antenna dual polarization or 8-antenna single polarization correlator based on CPU-GPUs was built.

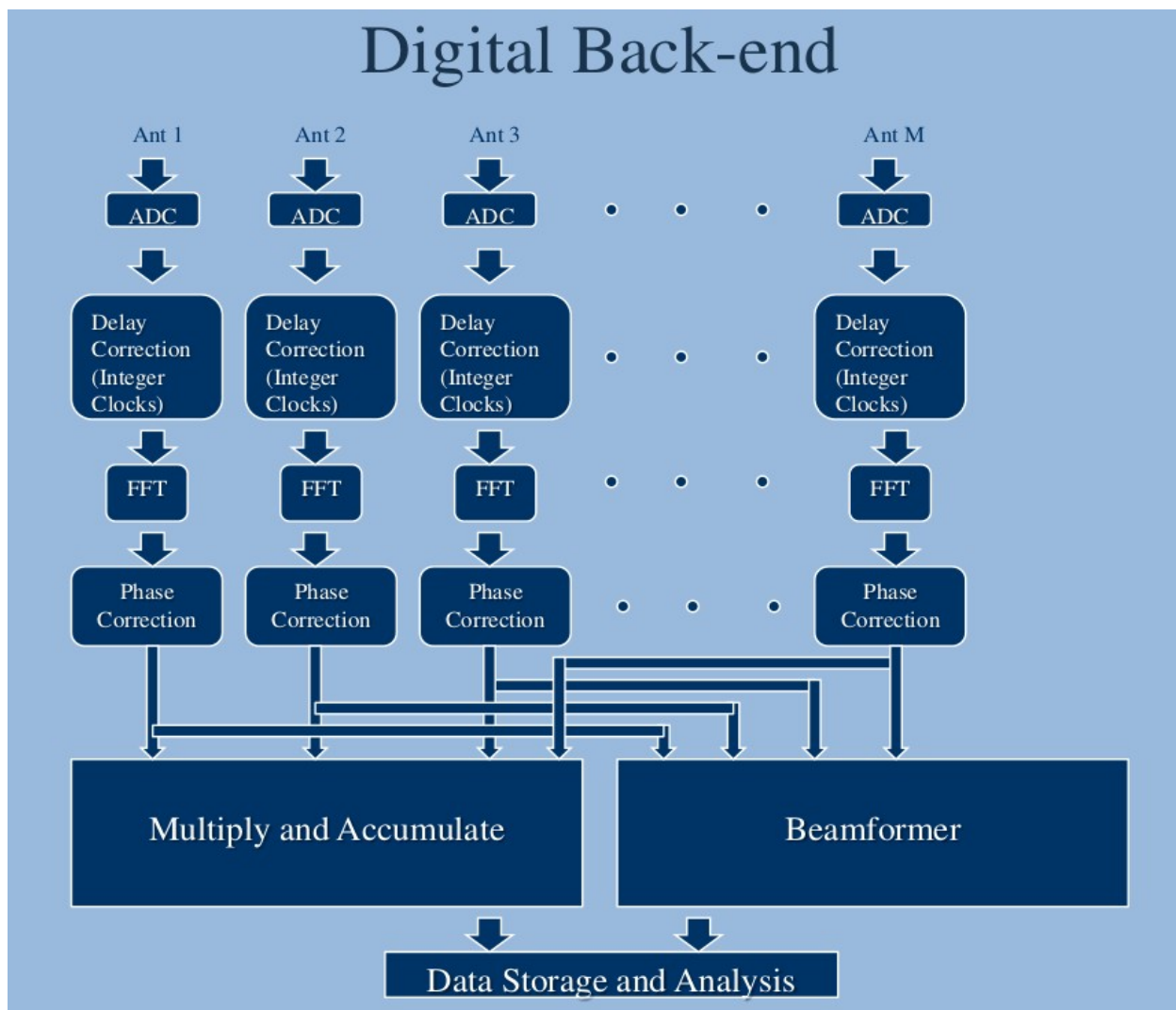


Fig 1 : Digital Back-end

## 2. Target specifications of uGMRT :

Number of Stations	:	32
Maximum instantaneous Bandwidth	:	400MHz
Number of spectral channels	:	2048 – 8192
Number of input polarizations	:	2
Full Stokes capability	:	Yes
Dump time	:	min. 128ms
Coarse and fine delay tracking	:	+/- 128 us
Fringe rotation	:	up to 5 Hz
Subarray support	:	Yes
Incoherent and Phased array beams(for pulsar work)		

## 3. Design description :

The design is a hybrid one using FPGAs and CPU-GPUs for various processes in the digital back-end chain. FPGAs connected with ADCs perform the digitisation and packetising the data while CPU-GPUs acquire the data, perform correlation and record the visibilities onto a disk for post-processing and analysis.

### Hardware :

1. ROACH boards:- ROACH stands for Reconfigurable Open Architecture Computing Hardware. The board is a standalone FPGA processing board built around Xilinx Virtex-5.
2. CPU-GPUs with Nvidia GPU (Tesla C2050 or above) with Myricom 10GbE NIC card and Infiniband NIC.
3. iADC boards :- 1x Atmel/e2V AT84AD001B 8-bit dual 1Gsps, with clock 10MHz – 1GHz 50Ω 0dBm.
4. Mellanox QDR 18-port Infiniband switch
5. Control PC for ROACH boards configuring and programming
6. PCs with Infiniband NIC as host, IA recording, PA recording nodes
7. Signal generator as clock source for the ADCs

### Software :

1. CASPER MSSGE Toolflow - Matlab-Simulink + System Generator +Xilinx 11.5 EDK
2. Centos Linux 5.6 (Kernel 2.6.18-194.32.1.el5 or higher on X86\_64)
3. Python 2.6 on Control PC for ROACH boards
4. Nvidia GPU device drivers + CUDA toolkit on CPU-GPUs
5. OpenMPI 1.4.5 or higher

**Implementation :** The 4-antenna dual polarization correlator is implemented using two ROACH boards each with two ADCs and four CPU-GPUs and an 8-port

infiniband switch(40 gbps).

Insert HERE a block diagram showing 4-ROACHes and CPU-GPUs with interconnectivity.

- 1. Digitisation and Packetising in FPGA :** The design is made of Matlab Simulink blocks, CASPER BEE-XPS block set. The ADC is an 8-bit sampler and is connected to the ROACH board through a 40-pin Z-DOK connector. Each ADC can sample two input signals. The sampled data from each channel is stored in the FIFOs. Each FIFO is 8192 bytes in size. The data from FIFOs is then packetised in the 10GbE block using control logic and is sent over 10GbE network. The packet size is 8242 bytes. The packet structure is 42 bytes of UDP packet header, 8 bytes of packet counter, 4096 bytes from channel 1 and 4096 bytes from channel 2. The sync signal to the ADC is a PPS/PPM signal to synchronize the data from many ROACH boards. For 200MHz BW, 8 bits per sample and two input channels, the data is 6.4 gbps.

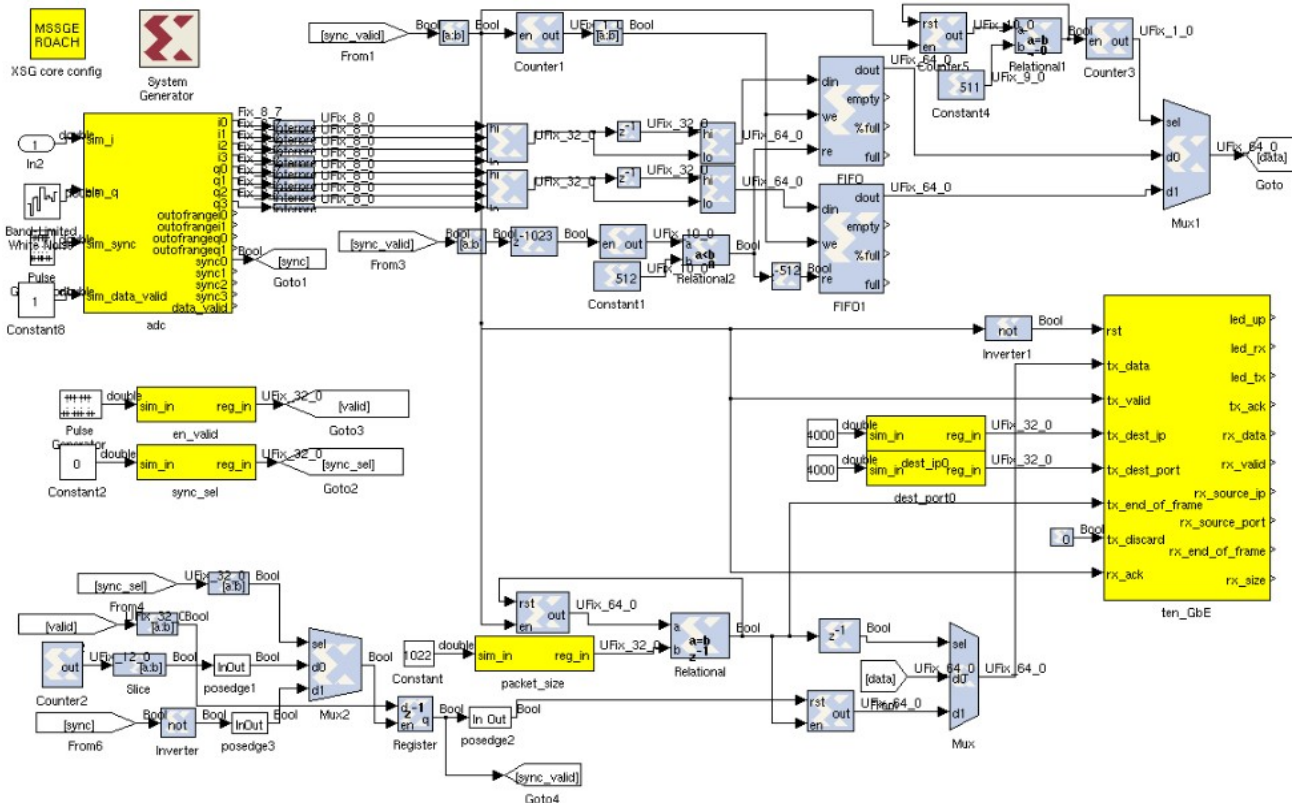


Fig 2 : FPGA design

- 2. Acquisition and Shared Memory buffering :** The packets are received on the CPU side through the Myricom 10GbE NIC. The data thus received is written into shared memories. Two shared memories are created on each node for each input channel. Each shared memory is of four buffers, each buffer size is 256 MB. The data before writing into the shared memory is corrected for coarse delay. Also the time when the first packet is received in the first acquisition node is noted and shared with other nodes for time

synchronization. The nodes are NTP synchronized with the GPS time server. The coarse delay is corrected for every buffer of 256 MB.

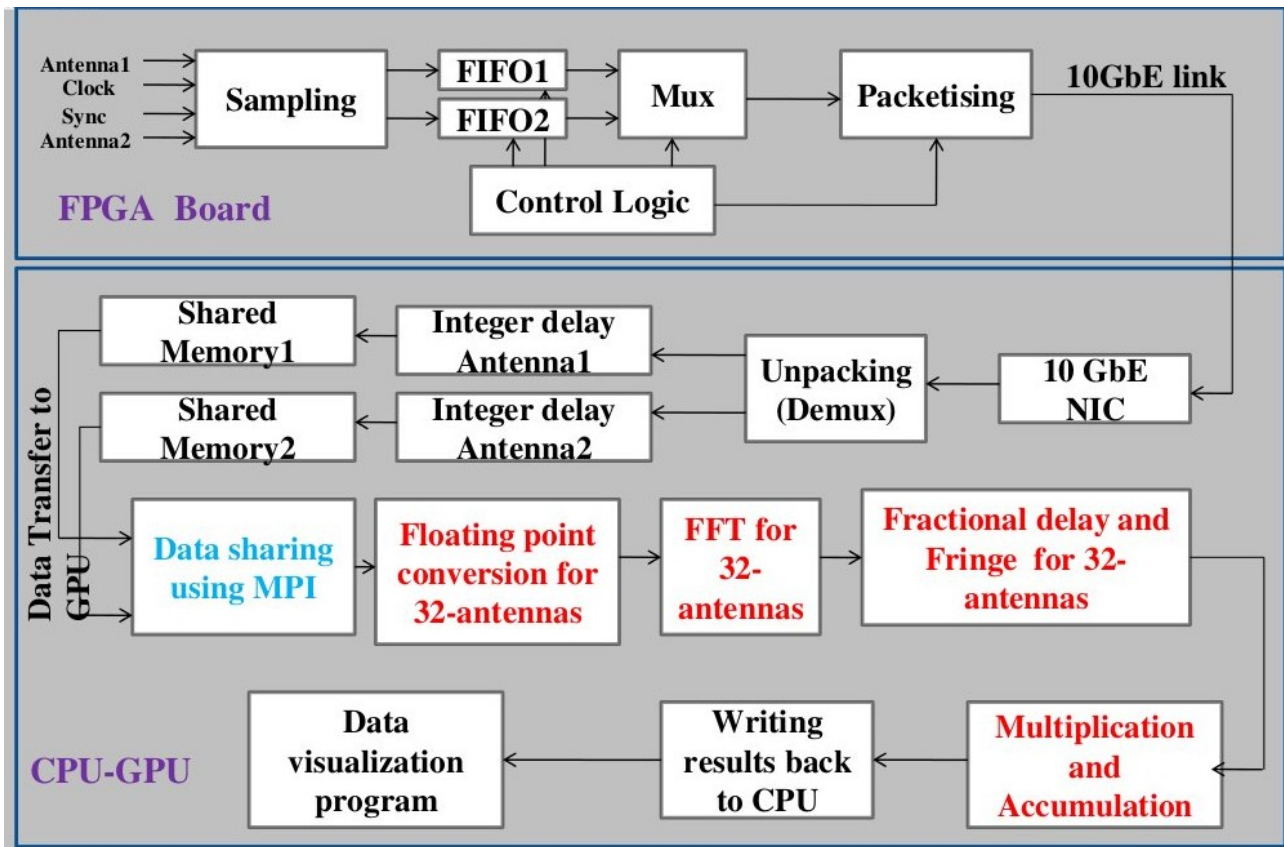


Fig 3 : Data flow on single ROACH and CPU-GPU

- 3. Time slicing and data sharing using Infiniband :** The data from the shared memories is read and is shared over infiniband network with the other nodes. The shared memory buffer is read and sliced into number of nodes times slices. Node 0 keeps the first slice and sends the second slice to node1, third slice to node2 and so on. At the same time, node 0 receives slice 0 from node1, node 2 and so on. Similarly, node 1 keeps slice 1, sends the remaining slices to corresponding nodes and receives slice 1 from all the other nodes. This is happened over all the nodes at the same time. To keep the nodes busy all the time, non-blocking MPI calls are used for this communication.
 

Insert model code of the algorithm HERE.

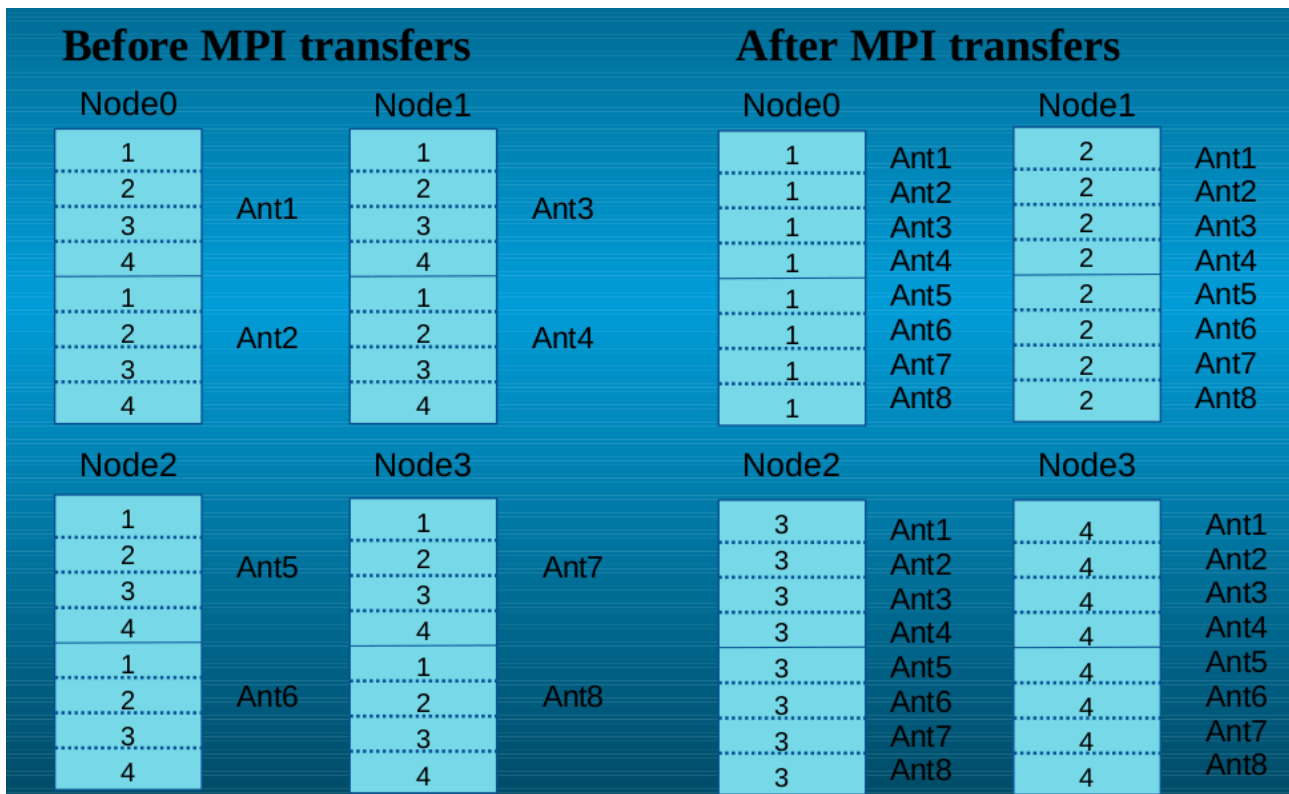


Fig 4 : Data sharing

4. **Correlation in GPU :** A slice of data of all antennas in each node is copied to GPU for correlation process. FFT is performed using CUFFT library. The fourier transformed data is fine delay corrected and compensated for fringe rotation. The delay cal function calculated the fstc and fringe values from the starting time which was noted when the first packet was received.

Insert HERE detailed text about FFT, phase shifting and MAC. Insert HERE MAC table for blocks and threads.

5. **Software Model :** The various processes mentioned above in the CPU-GPU are performed parallelly using nested openMP sections. The main thread creates two threads. One for data acquisition and shared memory buffering. The other thread creates four more threads for (1)data reading from shared memories, (2)MPI sharing of data, (3)correlation in GPU and (4)writing the visibilities to shared memory to be written by a host node. Double buffering ping-pong scheme has been implemented to enable the parallel processing.
6. **Online control and visibility recording :** From Sanjay and Nilesh
7. **Computation and IO requirements :** For 400 MHz bandwidth and 4 antennas, the total computation requirement for the correlation process is nearly 10 Teraflops. Of these 10 teraflops, the maximum chunk is taken by MAC (6.6 Gflops) , followed by FFT (2.9 Teraflops) and phase shifting ( 0.1 Teraflops). The IO to be processed is nearly 25 GB/s.

**4. Status and Results :** Include first results and some images. Get a summary from DVL

## **Appendix – I**

SOP for GWB phase 1 to be added HERE