Backend System Development for uGMRT

Harshavardhan Reddy, Sanjay Kudale, Nilesh Raskar, Shelton Gnanaraj, Navnath Shinde, Irappa Halagalli,
Sandeep Chaudhari, Mekhala Muley, Sweta Gupta, Dnyandeo Nanaware, Kaushal Buch, Ajith Kumar, Yashwant Gupta
Giant Metrewave Radio Telescope, NCRA-TIFR, Pune

Introduction

The Giant Metrewave Radio Telescope (GMRT) is a powerful Radio Telescope consisting of 30 parabolic dishes of 45 mtr diameter, spread over a region of 25 km diameter. The antennas are arranged in a Y shape with 12 antennas in a compact central array and 6 antennas in each arm of the "Y". The telescope currently operates at 5 different wave bands from 150 MHz to 1450 MHz with a maximum instantaneous bandwidth of 32 MHz. The signals from the antennas are down converted and processed using a Software based backend system (GSB) to generate interferometry and array outputs.

The GMRT is currently undergoing a major upgrade(uGMRT) to provide near seamless frequency coverage in the 50-1500MHz band and an instantaneous bandwidth of 400 MHz, with improved RFI rejection capabilities. The upgraded receiver will provide many features like high dynamic range, improved channel resolution, reduced system complexity, highly flexible software based backends etc. The Backend Receiver system is a major component being upgraded as part of uGMRT and consists of analog processing section GAB (GMRT Analog Backend) and a digital processing section GWB (GMRT Wideband Backend).

The Backend system combines the signals from all antennas and gives out visibilities for each baseline which are recorded for offline processing. It allows multi-subarray mode of observation with independent control on each subarray. For spectral observations narrow band mode is supported. In the beam mode, signals from selected antennas are added to generate PA and IA beams.

Analog Backend System

Frequency conversion of RF signals received at Central Electronics Building (CEB) to Baseband.

Switchable attenuator to equalise the power given to ADC (takes care of receiver gain variations).

Selectable filter bank to provide 100, 200, 400 MHz bandwidth signals to ADC.

LO frequency selection :
- option-1 common LO for all antennas
- option-2 individual for each antenna

Time & Frequency standard :
- option-1 GPS disciplined Rb oscillator
- option-2 Active Hydrogen Maser

Fig1. 30 Antenna GAB system

Specifications:

Number of antennas	:	30
Number of Polarisations	:	2
System control	:	online
GPS/Rb based standard		
Input Frequency range	:	50 - 1500 MHz
Input Power level	:	-24 dBm
Output Frequency range	:	400 MHz
Output Power level	:	-12 dBm
Output Bandwidth	:	100/200/400 MHz
Switchable attenuation	:	+/-16dB, 0.5dB step
LO common	:	10-1500MHz, 1Hz step
LO antenna specific	:	0.6 - 1.5GHz,0.5MHz step
Available Headroom	:	27 dB

Phase-2 Specifications:

LO selection	:	independent for each channel
RF Filter bank	:	7 band pass filter selection
Power Monitoring	:	input & output
Noise+CW cal signal at GAB input		
Noise cal signal at ADC input		
60-1 monitoring at GAB output		
System status monitoring		
Implementation of Maser standard		

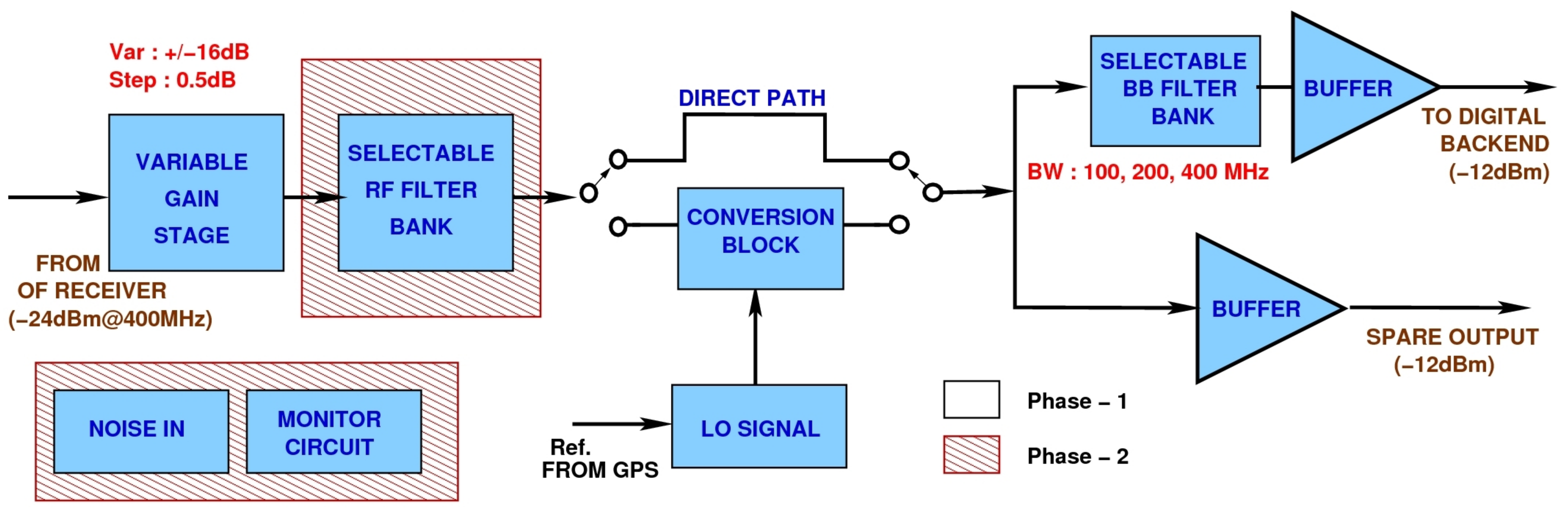


Fig2. Analog Backend block diagram

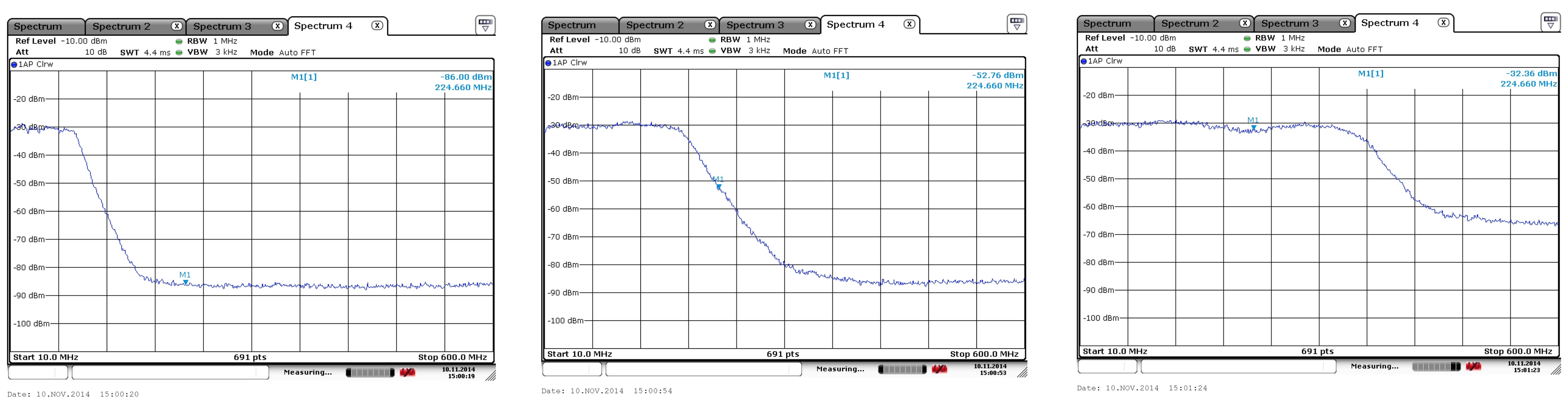


Fig3. GAB output bandwidth selection 100, 200, 400 MHz

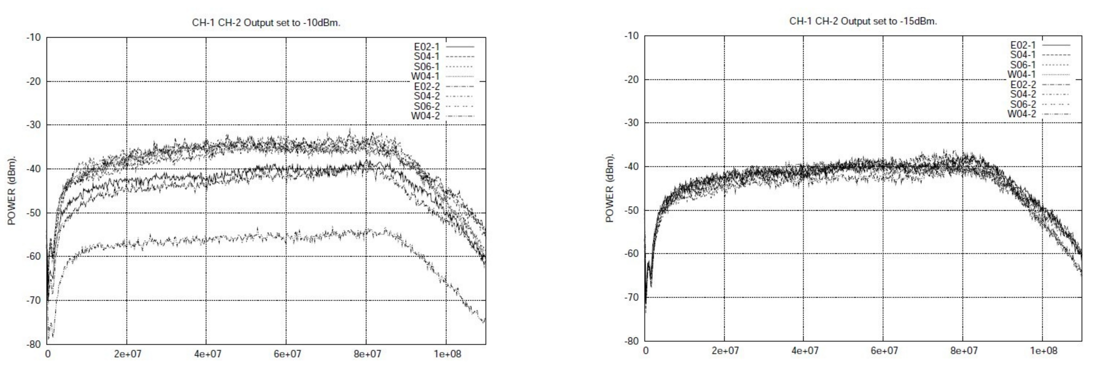


Fig4. GAB output before and after power equalisation

Digital Backend System

Specifications:

Number of stations	: 32
Number of input polarizations	: 2
Max instantaneous BW	: 400 MHz
Number of spectral channels	: 2048 - 32768
Full Stokes capability	: Yes
Dump time	: min 128 ms
Coarse and fine delay tracking	: (+/-128 uS)
Fringe rotation	: up to 5 Hz
Walsh switching	: Yes
RFI Filtering	: Yes
Subarray support	: Yes
Narrowband modes	: Yes

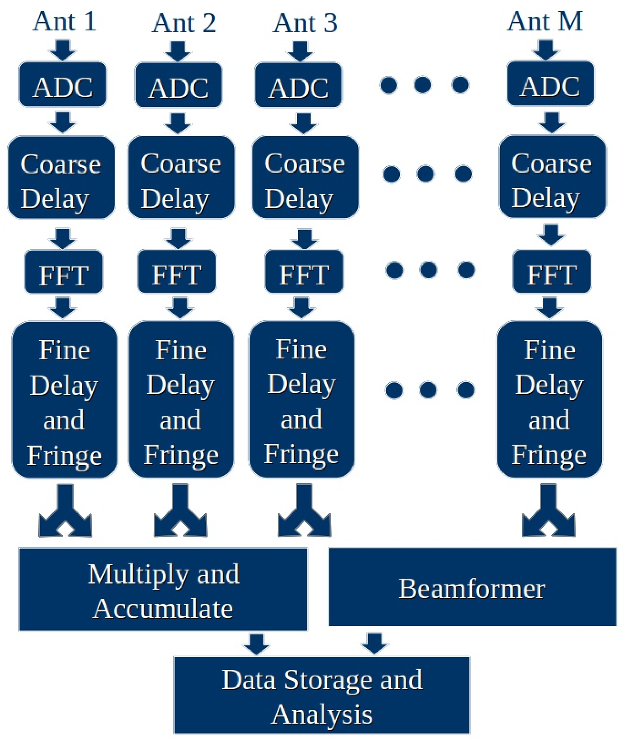


Fig5. General FX correlator

Design :

A hybrid design involving FPGAs and CPU-GPUs is used in GWB. FPGAs along with the ADCs perform the digitisation and packetising the data while CPU-GPUs perform correlation and record the visibilities onto a disk for post-processing and analysis. A scalable design is used in which FPGA and CPU-GPU boards can be added to take care of additional antennas when added to the system. The *figures 6* and *7* depict the basic design and data flow in a single FPGA-CPU-GPU pair respectively. The design takes advantage of the high data rate interconnect of Infiniband and faster processing power of GPUs.

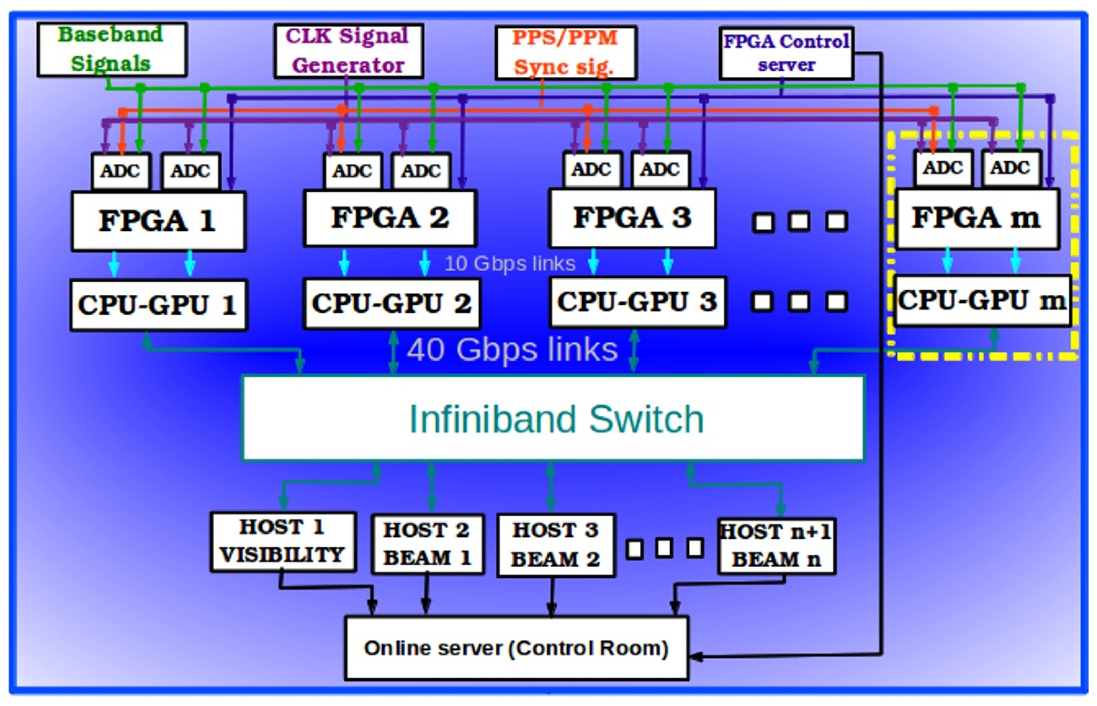


Fig6. GWB FPGA+GPU hybrid design for Digital Backend

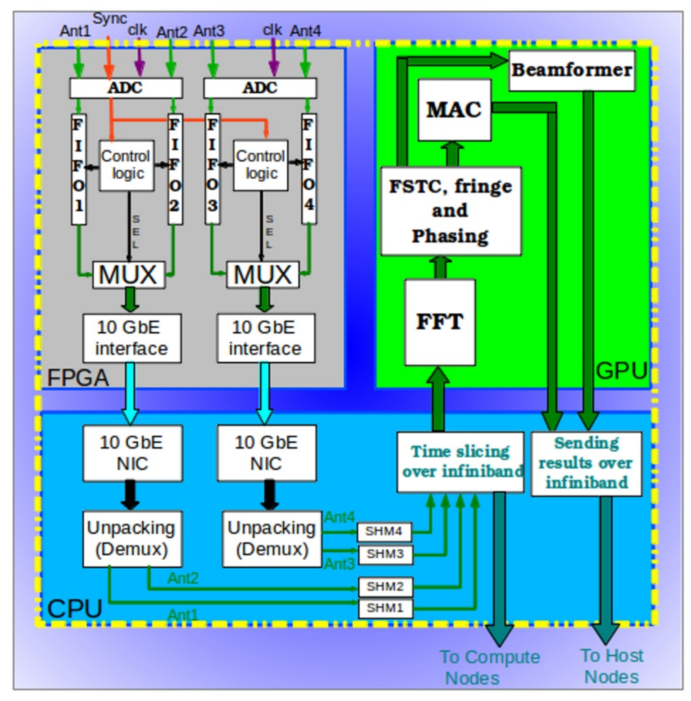


Fig7. Data flow in a single FPGA-CPU-GPU

Computation requirement at each stage :

No. of antennas, N_a = 30 No. of pols, N_p = 2 No. of spectral channels, N = 32768
Bandwidth, v = 400MHz No. of beams, N_b = 1 No. of stokes, N_s = 4
cops = complex operations per seconds

FFT computation load	=	$2vN_aN_p\log_2 N$ cops	=	3.1 Tflops
MAC computation load	=	$vN_sN_a(N_a+1)$ cops	=	6.6 Tflops
Fringe and FSTC correction	=	vN_aN_p cops	=	102.4 Gflops
Incoherent Beamformer	=	$vN_aN_p + vN_aN_p$ (multiply + add) cops	=	153 Gflops
Coherent Beamformer	=	$vN_pN_bN_s + v(N_a + N_s)N_pN_b$ (multiply + add) cops	=	166 Gflops
Total computation cost	=	**~ 10 Tflops**		

Data rate to be handled :

No. of CPU-GPUs, m = 16 No. of bits per sample, B = 4 Sampling frequency, F_s = 800MHz

Data rate between one FPGA board and a CPU-GPU, D_a	=	$N_aN_pF_sB$ bits/second	=	12.8 Gbps
Total data rate between FPGA boards and CPU-GPUs	=	D_am bits/second	=	204.8 Gbps
Bi-directional data rate between CPU-GPUs for time slicing and data sharing	=	$D_a(m-1)/m$ bits/second	=	12 Gbps

Why GPUs :

GPU-accelerated computing is the use of a graphics processing unit (GPU) together with a CPU to accelerate scientific, analytics, engineering, consumer, and enterprise applications. It offers unprecedented application performance by offloading compute-intensive portions of the application to the GPU, while the remainder of the code still runs on the CPU. A CPU consists of a few cores optimized for sequential serial processing while a GPU has a massively parallel architecture consisting of thousands of smaller, more efficient cores designed for handling multiple tasks simultaneously.

Fig8. CPU vs GPU

Implementation :

A 16 ant, 2 pol, 200MHz bandwidth digital backend system has been implemented on Nvidia Tesla K20c GPUs hosted by eight Dell T620 servers with CASPER based Xilinx Virtex5 FPGA and ADC boards.

*Fig9.*16 antenna backend

A 32 ant, 2 pol, 400MHz bandwidth digital backend system is being developed with Walsh demodulation, RFI Filtering, Narrow band and multi-subarray modes.

Results :

The backend system is used to process the GMRT antenna signals and the results are found to be as expected.

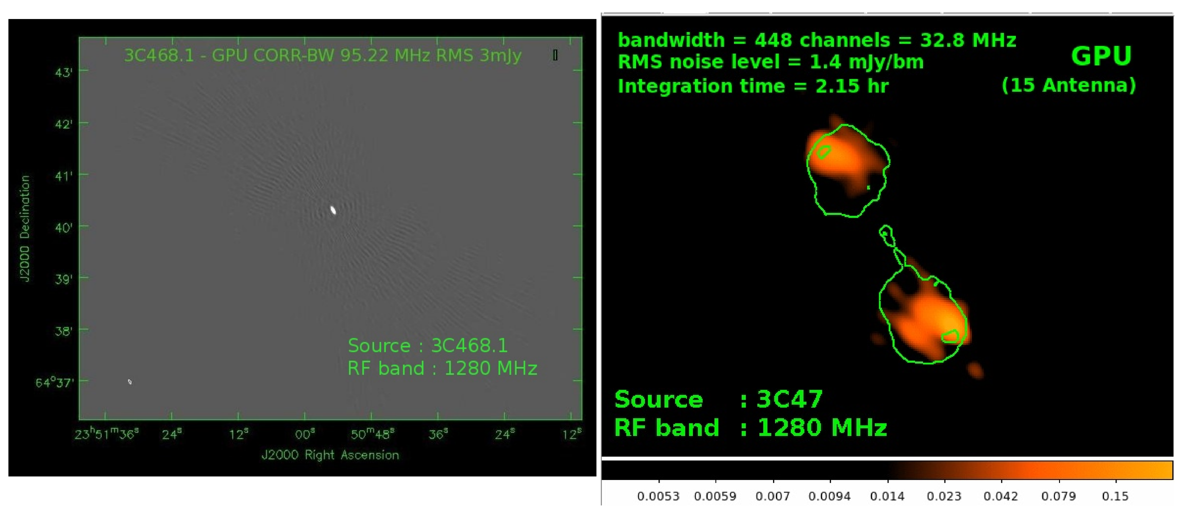
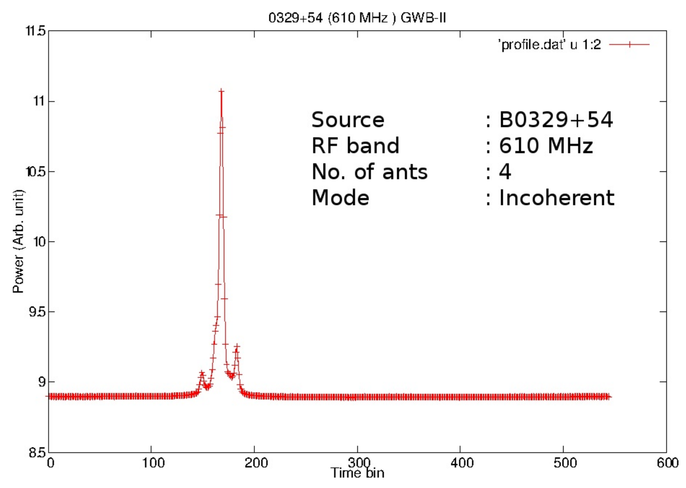


Fig10. Some early results (Courtesy : Dharam Vir Lal)



Acknowledgements:

1. Swinburne University, Australia
2. Nvidia, India
3. CASPER Group
4. Backend Group, GMRT
5. Mangesh Umbarje, GMRT, NCRA-TIFR
6. Sumit Mirajkar, GMRT, NCRA-TIFR
7. Jayanta Roy, NCRA-TIFR
8. Dharam Vir Lal, NCRA-TIFR

References :

1. A real-time software backend for the GMRT, Jayanta Roy et.al 2010
2. http://www.nvidia.com/object/what-is-gpu-computing.html
3. https://casper.berkeley.edu/