BIGCAT Overview (v0.5.1 12 March 2021)

Chris Phillips (Chris.Phillips@csiro.au)

BIGCAT (Broadband Integrated GPU Correlator for ATCA) is an upgrade of the CABB digital backend to hybrid FPGA/GPU based processing with twice the bandwidth of the current CABB system (increasing the instantaneous bandwidth from 4 to 8 GHz, dual polarization).

ATCA/CABB

The ATCA is a 6-element radio interferometer located near Narrabri, NSW, Australia. It is roughly 425 km North West of Sydney. The telescope operates between 1 and 105 GHz using 5 separate receivers, as shown in Table 1. In 2009 the original digital backend and receivers were updated to increase the instantaneously processed bandwidth as part of the "CABB" upgrade (CA BroadBand). Although the high frequency receivers have a very wide front-end frequency range, the existing down conversion system is limited to 8 GHz maximum instantaneous bandwidth. For the high frequency receivers, the RF signal is mixed with an LO signal to bring it to an IF range of 4-12 GHz, which matches the frequency range of the "4cm" receiver. The low frequency (1-3 GHz) "16cm" receiver is mixed with a LO to a *higher* IF frequency, to match the 4-12 GHz input range.

Receiver	Frequencies	SEFD
16 cm	1.1-3.1 GHz	45
4 cm	3.9-12.0 GHz	36
15 mm	16-25 GHz	60
7 mm	30-50 GHz	112
3 mm	83-105 GHz	724

Table 1 The ATCA Receiver fleet

For CABB, a second stage IF system splits the two IFs and each can be independently tuned within the 8 GHz range and filtered to a final 2 GHz band, which is then digitized and processed on the CABB correlator. The result is two dual polarization 2 GHz bands (ie a total of 4 GHz sky frequency coverage).

The CABB digital backend has a default spectral resolution of 1 MHz for standard continuum observations. For spectral line observations, 2 "zoom band" setups are available. These give 16 digitally tunable bands of 1 or 64 MHz across the 2 GHz bands, each with 2048 spectral points (giving spectral resolutions of 0.49 kHz and 31.25 kHz). The 16 bands can be concatenated in frequency to give continuous frequency/velocity coverage.

BIGCAT

The BIGCAT upgrade will replace all CABB digital equipment. This includes the digitizers and the correlators. Digitizers and coarse filterbank will be built around the Xilinx RFSoC, an FPGA with embedded high-speed samplers. The raw voltage data will then be streamed from the antenna to the screened room using commodity 100 Gbps Ethernet, utilizing existing optical fibres. In the screened room, the voltages from each telescope will be streamed to a small (16 node) GPU cluster for further processing (fine channelization, cross correlation and averaging). Computer server network speeds do not have the capacity to ingest the full 2 GHz bandwidth from 6 antennas, so the RFSoC board needs to form a coarse filterbank across the band to allow distributing the data in frequency across multiple GPUs. The flexibility of GPU programming means many new modes can in principle be easily added – e.g. “unlimited” spectral resolution, RFI blanking or active mitigation and multiple tied array beams.

To double the bandwidth, the 2x2 GHz tunable IFs from the CABB system will need to be entirely replaced. A new RF board will be developed which will connect to the 8 GHz IF from the receivers. 4x2 GHz bands will cover the 8 GHz range (dual polarization). Each of the bands will have limited tuning, to mitigate the problem of a specific spectral line (for example) sitting directly at the transition between IFs. The new scheme being developed will should hopefully allow tuning up to 12.3 GHz, giving access to the 12.2 GHz methanol transition for the first time[1]. The components of the BIGCAT system are shown in Figure 1.

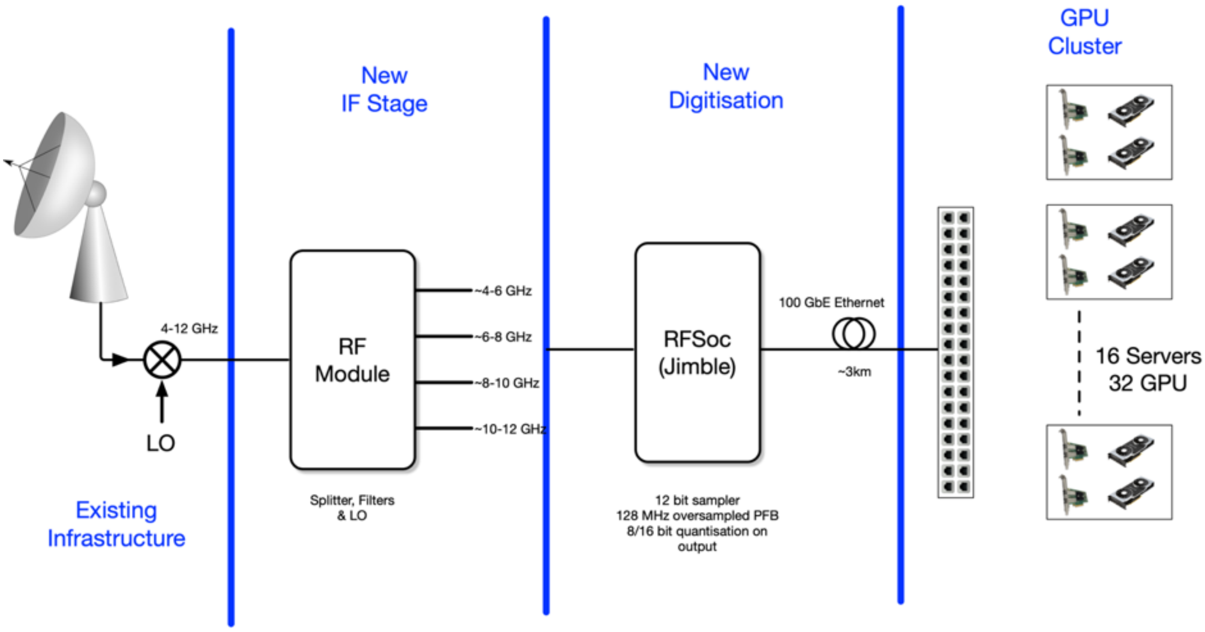

Figure 1: Overview of the BIGCAT signal chain

BIGCAT IF

The CABB IF system is built around an 8 GHz IF stage from 4-12 GHz. This corresponds to the RF output of the 4cm receiver. For the high frequency “mm” receivers the RF output is mixed with a reference LO and down converted to the 4-12 GHz range. This 8 GHz IF will be retained for BIGCAT, however all subsequent IF stages will be replaced. This will give BIGCAT an instantaneous bandwidth of (approximately) 8 GHz, double that of CABB. Because the 8 GHz IF stage is retained, it will not be possible to simultaneously observe frequencies with a separation > 8 GHz, even for receivers with a front-end bandwidth > 8 GHz. Note for the 1-3 GHz 16cm receiver, the RF will be upconverted to 9.64-11.64 GHz. Obviously, the 16cm receiver will be limited to a maximum bandwidth of 2 GHz.

The new RF stage will consist of 4 IF bands with an analogue bandwidth of 2048 MHz for each band. The bands will NOT be fully tuneable; however, each band will have a tuning range of 800 MHz in steps of 160 MHz. A schematic of this setup is show in Figure 2, see Table 2 for a list of frequency ranges for each of the 4 IFs. This arrangement will allow the astronomer to either choose to either align each IF to eliminate any gaps, or space the IFs as far as possible to maximum frequency separation.

[1] This may require replacement of the 4 cm receiver 4-12 GHz filter at a later date

Note while the analogue bandwidth of each IF is 2048 MHz with a sampling rate of 4096 MHz, filter roll off and properties of the digital polyphase filterbank means there is only 1920 MHz usable bandwidth (ie the bottom and top 64 MHz from each IF is unusable).

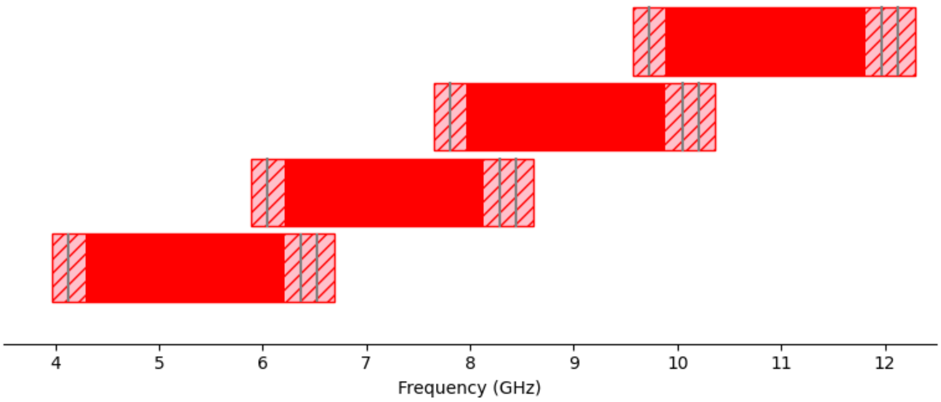

Figure 2: The proposed BIGCAT frequency bands. The solid red represents a “typical” observing setup giving continuous frequency coverage over 7.7 GHz. The “hashed” region indicated the possible tuning range of each IF (in steps of 160 MHz).

Typically, the IF LOs will be set to allow continuous frequency coverage. This results in a total bandwidth of 7680 MHz. Standard configurations will “hide” this from the end user. In specific cases (e.g. experiments which want maximum continuum frequency range) the LOs can be configured to increase the total frequency coverage by 640 MHz (ie maximum span 8320 MHz). This scheme may also allow access to the 12.2 GHz methanol maser transition (if the front end CX 4-12 GHz filter allows). The specific frequency ranges for each band is shown in Table 3. It is intended that for each row in Table 3, the user can choose any one column as their frequency range, independent of their choice in the other rows (although this will be somewhat constrained as detailed later in this document).

	RF Low	RF High	BW (MHz)	LO (GHz)
Band 1	3968 MHz	6688 MHz	1920	8.00-8.80
Band 2	5888 MHz	8608 MHz	1920	9.92-10.72
Band 3	7648 MHz	10368 MHz	1920	11.84-12.64
Band 4	9568 MHz	12288 MHz	1920	13.76-14.56

Table 2: Proposed IF ranges for BIGCAT. Note the IFs are tuneable in steps of 160 MHz

	LO1	LO2	LO3	LO4	LO5	LO6
Band 1	3968-5888	4128-6048	4288-6208	4448-6368	4608-6528	4768-6688
Band 2	5888-7808	6048-7968	6208-8128	6368-8288	6528-8448	6688-8608
Band 3	7648-9568	7808-9728	7968-9888	8128-10048	8288-10208	8448-10368
Band 4	9568-11488	9728-11648	9888-11808	10048-11968	10208-12128	10368-12288

Table 3: Proposed discrete frequency bands for BIGCAT. For each band there are 6 possible LO settings, with 160 MHz step. Note frequencies above 12 GHz are not guaranteed because of the receiver front end filter.

DIGITIZERS

The BIGCAT digitizers are based around the Xilinx RFSoC FPGA. This device has 8 high-speed analogue-to-digital converters (ADC) and an FPGA fabric for control and initial digital signal processing. The RFSoC chip will be built into the “Jimble” digitizer board (shown in Figure 3). The Jimble is a custom designed board which will be used for the BIGCAT upgrade as well as CryoPAF system for Parkes and an update of the Parkes UWB receiving system. For BIGCAT, only 4 of the ADCs will be used due to resource limits on the FPGA fabric. This means two Jimbles will be needed per antenna.

The Jimble will each sample four 2 GHz IFs at 4096 MHz. A polyphase filterbank (PFB) will be used to coarsely channelize the data to channels with 128 MHz separation. To avoid digital filter roll off problems between the coarse channels, an oversampled PFB will be used with an oversampling ration of 32/27. This results in channels which are separated by 128 MHz, but with approximately an 18% overlap between channels. This overlap will be discarded after the fine channelization in the GPU correlator, resulting in continuous frequency coverage with no roll off between channels.

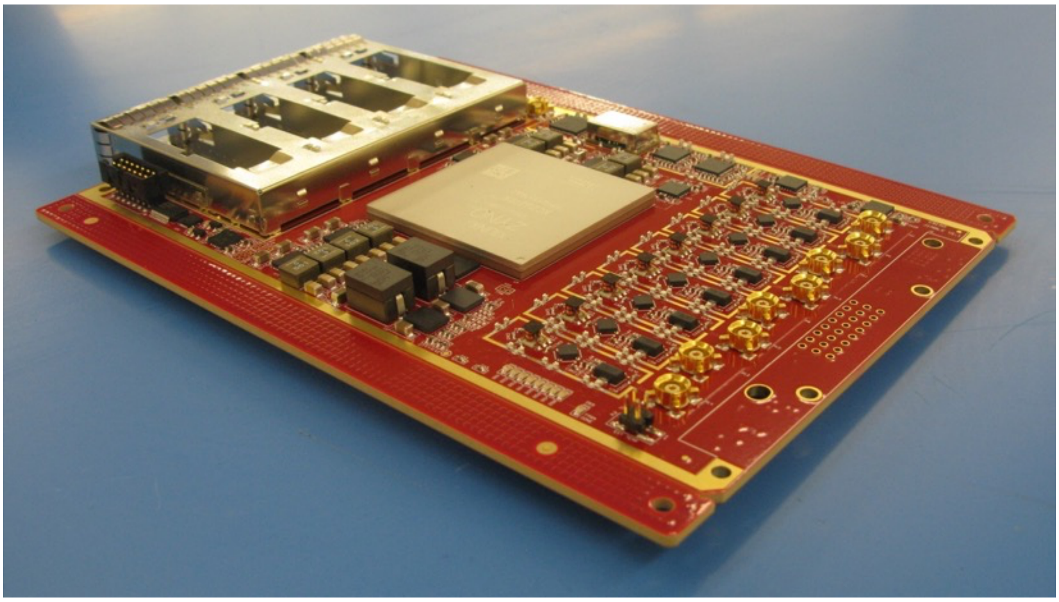

Figure 3: The CASS designed Jimble digitiser. The Jimble uses the Xilinx RFSoC chip for analogue to digital conversion and initial signal processing.

The voltage data for each coarse channel will be scaled then quantized to 8 bits (16 bits will be used for the 16cm system as it has less total bandwidth and requires higher dynamic range) and then sent to the screened room over commodity 100 Gbps Ethernet (GbE), using existing optical fibre. Each Jimble has two 100 GbE connections. At 8 bits/sample and 4 dual pol IFs, the total data rate is 73 Gbps per 100 GbE connection (291 Gbps/antenna). The total data rate going to the screened room from the antenna will be 1.75 Tbps.

8 bits per sample is required because there are only 4 optical fibres usable for the 100 GbE data connection. For the 16cm receiver, the total receiver bandwidth is only 2 GHz, so 16 bits per sample will be used.

GPU Correlator

Data received from the antennas will be connected to a 64 port 100 GbE ethernet switch. Connected to the switch will be a 16 server GPU cluster. Each server node will contain 2 GPUs (graphical processing unit) and two 100 GbE network cards. Each GPU/Network card pair will process independent data – for standard processing modes there will be no transfer of data between separate GPUs. For the following discussion a single GPU/network card is considered. The full system is 32x larger.

For processing, each 128 MHz coarse channel from each antenna must be sent to a single GPU. This gives a data rate of 29 Gbps. Each GPU will process data from 2 coarse channels (58 Gbps/GPU). Once the data is copied to the GPU, it needs a phase correction to allow for Earth rotation (fringe rotation), run through a fine channelization step and finally cross multiplication between antennas and polarizations and integration.

The current design plan will be to run the correlation at relatively high spectral resolution across the full 128 MHz band, for all coarse channels (approx. 9.3 kHz spectral resolution). This is approximately 4x higher spectral resolution than the CABB 64 MHz zoom mode. Continuum data will be formed by averaging many channels together, across the entire processed bandwidth. “Low Resolution” spectral line data will be generated by defining spectral “windows”, with a specified frequency range and resolution. Within reason, there should be no limit on the number of spectral windows defined – allowing for extremely flexible spectral setup. High resolution spectra line observations will require zoom modes. Zoom modes will be a limited bandwidth region within the 128 MHz bands, correlated with much higher spectral resolution (Velocity resolutions much smaller than 0.1 km/s should be possible). Multiple zoombands within a 128 MHz will be possible, with flexible bandwidth (ie velocity coverage and spectral resolution). It is unlikely that the entire 128 MHz band can be correlated with very high spectral resolution. The total number of spectral points, number of zoom bands will depend on available data rates and GPU compute available. Exact limits are not available yet.

The (relatively) simple programming nature of the GPU system will allow great flexibility. Specific modes (such as very high spectral resolution) can be added with minimum effort, and it is planned to add robust RFI mitigation techniques (including time and frequency blanking as well as opening the option for adaptive filtering using reference antennas).

Note the current plan is that the system will have a number of pre-configured modes from which the astronomer will choose. The astronomer will NOT be able to directly set the setup parameters. This includes the setting of the separation of the final LOs and spectral resolution. Changes to the setup from existing modes will require consultation with local staff to create and validate a new mode. However, the astronomer will be able to define an arbitrary number of spectral windows with custom width and frequency resolution to flexibly cover spectral lines they want to observe. There will be limits on peak and average data output rates for the visibilities (still to be determined).

Expected Modes

Standard Interferometry

- Track and correlate “sidereal” objects
- Integration times of a few seconds to 10sec
- 8 GHz bandwidth
- Spectral resolution of up to 10kHz over entire band
- Spectral resolution < 0.5 kHz over portion of band
- Full Stokes products (parallel and cross polarisations)
- Autocorrelations
- Continuous Tsys measurements
- Wideband online calibration
- Mosaicing

Specialized Requirements

- Fast dump visibilities?
- Ultra-high spectral resolution
- Programmatic observation control (control of caobs from user python script rather than schedule)?

Voltage Recording (VLBI)

- Recording of up to 8 GHz bandwidth (Limited by available recording devices)
- 2, 4, 8 and 16 bit quantization
- Output channel bandwidths of a few to 128 MHz
- Phased array of 1 to all antennas
    - Record subsets of antennas as separate phased array sub arrays (same pointing centre). Subsets with separate fringe rotation reference. Potentially record 6 antenna independently, bandwidth limited by disk recording limits.
- Multiple tied array beams, with different pointing centres

Bi-static Radar/SSA

- Specialized modes, or rely on voltage streaming?

Non-sidereal observations (asteroids, satellites, planets, moon)

- Able to track and correlate non-sidereal objects
- Alternative geometric model input than "CALC"
- Handling near-field, rapidly moving objects

Pulsars

- Accumulate visibilities into N evenly spaced bins across the pulse period
    - N to be determined (>=64)
    - Bin length >= 1 usec
    - Tempo2 predictors
    - No de-dispersion?

Rapid Response

- "Hands off" correlator reconfiguration
- High reconfiguration reliability

Glossary of terms

- BIGCAT: Broadband Integrated-GPU Correlator for the ATCA; it is referred to as BIGCAT, not the BIGCAT correlator.
- CCC: Correlator Control Computer; is the computer which manages the configuration and operation of the individual correlator nodes, and receives the data from each node, and handles formatting the data into the output correlated data files.
- Coarse channels: the complex voltage output stream from the digitizer; these are oversampled chunks of bandwidth, and each digitizer produces many coarse channels.
- Configuration state: all the settings that need to be saved/recalled for scans at different frequencies and using different modes; things like attenuation settings, delay and phase offsets, amplitude scaling coefficients.
- Correlator: the entire cluster of GPU compute nodes.

- Correlator Mode: a fixed configuration of the correlator, which defines the channelisation, the steps involved in the digital signal processing, and what can be output by the correlator. Each mode is defined by a correlator profile.
- Correlator Profile: a pre-made file which specifies all the settings defining a correlator mode. These files are not user-configurable but are available to the users through the scheduling tool.
- Digitizer: the combination of analogue-to-digital converter and coarse-channeliser, and ethernet formatter; it is not to be called a sampler.
- DSP module: a logical part of the Digital Signal Processing chain, which may contain one or more kernels. Examples include the fringe-rotation module, the cross-correlator, RFI mitigation, tied array summer.
- GPU: Graphics Processing Unit, which will be used to do the massively parallel mathematics in the correlator.
- IF: Intermediate Frequency; this is the analogue signal after mixing and before the digitizer. There can be several different IF stages.
- Ingest: the process of moving the data from the digitizer into the CPU/GPU memory on a node.
- Integration: the amount of time contained in each visibility written out to disk by the CCC.
- Kernel: a block of code that is run in parallel on the GPU that does some processing.
- Mosaic: a way to observe more than one primary-beam of sky with low overheads
- Node: a single computer chassis containing a CPU(s) and multiple GPUs.
- Operator API: a service that a user can connect to if they would like to programmatically control the telescope instead of using a fixed schedule.
- O/S PFB: Over-Sampled Polyphase FilterBank, the device that does the coarse channelisation within the digitizer.
- Polarisation: one of the two receptors in each receiver and are labelled as X or Y (not A or B).
- Receiver: one of 16cm, 4cm, 15mm, 7mm or 3mm systems on one antenna.
- Receiver package: one of L/S, C/X or mm; the package includes one or more receivers and electronics in a single dewar.
- RF: Radio Frequency; this is the analogue signal straight from a receiver before any mixing.
- Scan: a single observation with a defined source (or mosaic), duration and configuration.
- Scheduling tool: the software that the end user will use to create fixed telescope schedules.
- Sky frequency: the user-specified frequency of interest, likely at the centre of the bandwidth made available in a correlator mode.
- Sub-integration: the amount of time processed in the GPU at once; the smallest amount of time that can be flagged without affecting other data.
- TASR: Tied-Array Stream Recorder; the machine(s) that receive the network traffic containing the tied array voltages and put them on disk.
- Tied array: coherently summed antenna voltages; the correlator may produce multiple tied arrays simultaneously.