

Basic Operations of the **C**ommensal **O**pen **S**ource **M**ultimode **I**nterferometer **C**luster (COSMIC)

Jack Hickish

June 30, 2022

On behalf of
The COSMIC Collaboration

For more information, contact

Dr Andrew Siemion

`asiemion@seti.org`

Dr Jack Hickish

`jack@realtimeradio.co.uk`

Abstract

The **C**ommensal **O**pen **S**ource **M**ultimode **I**nterferometer **C**luster (COSMIC) system is a newly developed digital signal processing instrument deployed at the Jansky Very Large Array (VLA), which is designed to enable Search for Extra-Terrestrial Intelligence (SETI) observations to be conducted at the VLA in parallel with “Primary User” PI-led Science observations. The COSMIC system capabilities are still under active development. However, in this document we attempt to summarize the planned operating modes of the COSMIC system in order to highlight opportunities where PIs might be able to leverage the power of the COSMIC system for their own observations, and also to provide insight as to where conflicts may exist between PI-led science and COSMIC.

1 What Is COSMIC?

The **C**ommensal **O**pen **S**ource **M**ultimode **I**nterferometer **C**luster (COSMIC) system is a cluster of FPGA, CPU, and GPU hardware, which at all times receives a copy of digitized data streams from each of the VLA’s antennas. The data received by COSMIC necessarily reflects antenna-pointing, LO-tuning and digitization settings commanded by VLA observers, but otherwise operates completely independently. In particular, COSMIC is unaffected by any of the interferometric phasing, channelization, or correlation processes of the VLA’s facility instrument: WIDAR.

COSMIC comprises three main components. These are:

1. A **fiber optic splitter**, which duplicates digitized data streams from the VLA antennas and delivers these to COSMIC digital instrumentation.
2. A small **FPGA Cluster**, which receives VLA antenna digitized data streams, channelizes these to 1 MHz frequency resolution, and transmits these data streams over multiple industry-standard 100 Gb/s Ethernet links.
3. A **CPU/GPU Cluster** which receives 1 MHz resolution data over Ethernet and performs a variety of digital signal processing operations, ultimately outputting SETI signals of interest (often referred to as “hits”) into a database for further analysis.

As deployed, the COSMIC fiber and FPGA subsystems are capable of handling the full observing bandwidth of the VLA. However, for the duration of 2023, we expect the COSMIC GPU cluster to be able to process no more than 2 GHz of bandwidth from each of the VLA antennas. Over time, the cluster will be built out so as to be able to handle the array’s full bandwidth.

2 COSMIC Digital Signal Processing

The COSMIC cluster is designed to exclusively utilize off-the-shelf hardware, and can perform a variety of software defined signal processing tasks. These tasks are undertaken on both the FPGA and CPU/CPU elements of the cluster.

2.1 FPGA Processing

The COSMIC FPGA subsystem receives custom-format digitized data streams from the VLA antennas. As defined by the telescope’s primary observer, these data streams comprise multiple IF tunings, each with either 1 GHz, or 2 GHz of bandwidth. The first stage of FPGA processing filters the incoming data streams into multiple 1 MHz channels, and applies delay and phase tracking to these channelized data streams to track a phase center (currently assumed to invariably be the pointing-direction of the array’s antennas). Channelized, “fringe-stopped” data streams are then transmitted from the FPGAs over a

collection of 100 Gb/s Ethernet ports as a series of UDP/IP packets. COSMIC may be dynamically configured to transmit all, or a subset, of 1 MHz channels. Data may also be dynamically steered to different CPU/GPU servers for processing, using standard Ethernet packet switching, or IP routing.

2.2 CPU/GPU Processing

The COSMIC CPU/GPU pipeline implements a beamformer, high-resolution channelizer, doppler-drift search, and calibration correlator. The configuration parameters of these algorithms may be varied in order to best leverage the array’s behaviour as dictated by the Primary Observer.

The COSMIC CPU/GPU processing pipeline is still under development, but will have the following capabilities:

1. Buffering of ~ 5 minutes voltage data for all antennas in the VLA
2. Beamforming of up to 32 coherent tied-array beams.
3. Incoherent beamforming.
4. Configurable frequency resolution between 1 MHz and 1 Hz.
5. Inter-antenna correlation (used for gain calibration observations) with millisecond time resolution.
6. Doppler drift searching for narrow band signals (in antenna or beam data) with linear drift rates between -50 Hz/s and $+50$ Hz/s

In normal operation, the COSMIC system will search for candidate signals with non-zero linear drift rates, and will permanently archive small data snippets (< 100 kHz bandwidth and < 5 minutes duration) around these events.

3 Observing Strategy

The COSMIC observing strategy needs to accommodate the behaviour of the array, which varies depending on the PI science being undertaken. However, the basic COSMIC behaviour will generally be:

1. Obtain the current pointing of the VLA from the facility meta-data broadcast system.
2. Configure FPGAs to transmit desired frequency channels, as dictated by the Primary User’s IF tuning settings.
3. If the telescope is pointed at a calibrator, configure the CPU/GPU system in “correlator” mode, generating full visibility matrices for all VLA antennas and polarizations, to be used to generate per-antenna gain solutions.

4. If the telescope is not pointed at a calibrator, look up SETI candidate stars in the Gaia¹ archive which fall in the VLA’s field of view.
5. Record antenna data to disk for the length of the VLA pointing, or a maximum of ~ 5 minutes.
6. Form synthesized beams on SETI target stars using antenna voltage buffer as input.
7. Generate ~ 1 Hz-resolution power spectra for each beam.
8. Use a Doppler-drift search algorithm to identify any narrow-band signals present in data which drift in frequency in a linear manner over time.
9. Where drifting signals are found, archive the corresponding antenna voltages around the immediate time and frequency of the drifting signal.

The COSMIC system will archive ~ 1 PB/year of candidate events on site. Post-processing of candidate data will include the use of voltage domain (for example direction finding, modulation classification, or other) algorithms on antenna-level archived data products.

4 Future Observing Opportunities

The COSMIC system is deliberately flexible, and its computing and storage resources may be repurposed through the development of new processing software. For example, future observing functionality which could be leveraged by non-COSMIC scientists might include:

1. Deep (many minutes, or hours) time buffers for single antenna or single beam voltage data streams, over restricted bandwidth.
2. Continuous, high time-resolution visibility matrix recording.
3. Real-time RFI studies.
4. Searches for non-linear frequency drift signatures.

The COSMIC team is keen to work with scientists in the community to develop auxiliary COSMIC observing modes which may augment the existing capabilities of the VLA’s existing facility and guest instruments.

¹<https://gea.esac.esa.int/archive/>