

The CARMA Software System

Stephen L. Scott¹, N. S. Amarnath⁵, Andrew D. Beard¹, Paul Daniel¹,
Chul Gwon⁵, Rick Hobbs¹, J. Colby Kraybill³, Erik Leitch⁴, David M.
Mehringer², Raymond Plante², Marc W. Pound⁵, Kevin P. Rauch⁵,
Peter J. Teuben⁵

¹ *California Institute of Technology/Owens Valley Radio Observatory*

² *NCSA/University of Illinois*

³ *University of California, Berkeley*

⁴ *University of Chicago*

⁵ *University of Maryland*

Abstract. CARMA combines the existing OVRO and BIMA millimeter-wave arrays with the new SZA array at a high altitude site. The array will have a total of 23 antennas of three different sizes, providing heterogeneous imaging capabilities at millimeter and centimeter wavelengths, along with a maximum bandwidth of 8 GHz and resolution of 0.1 arcseconds. The software system encompasses a monitor and control system, an archive, and an imaging pipeline. A well defined software process is used for a distributed software team that is spread across five sites. The university based nature of CARMA will provide hands on training for young astronomers and serve as a testbed for technical innovation.

1. Introduction

The Combined Array for Research in Millimeter-Wave Astronomy (CARMA) combines two existing millimeter-wave arrays, Caltech's Owens Valley Radio Observatory (OVRO) array and the Berkeley-Illinois-Maryland Association (BIMA) array at Hat Creek, adds the new Sunyaev-Zeldovich Array (SZA), and introduces significant new hardware. The hardware modifications to the existing antennas are extensive, including replacement of the local oscillator and IF systems to use common ones for interferometry. The new antenna hardware brings with it the next generation of technology, with embedded microprocessors interconnected by CANbus controlling individual hardware modules. All of the interferometric aspects of the array are also new, including the correlators. Although many algorithms and some code can be reused, much of the CARMA software system is new, driven by the hardware changes. The largest software component is a new monitor and control system, while the other major components are the archive and the imaging pipeline.

2. CARMA

CARMA¹ is a collaboration of six universities, composed of the five universities represented by the authors of this paper, with Columbia University as the sixth. The site for the array is Cedar Flat, located in the Inyo National Forest, about 16 miles on paved road from OVRO. At an altitude of 7200 feet (2200 meters), the site will allow routine operation in the 230 GHz band. There will be 55 stations with a maximum baseline of 2 km, giving a resolution of 0.1 arcseconds. CARMA will operate in three frequency bands, 27-36 GHz, 70-116 GHz, and 210-270 GHz. At first light there will be two correlators: an 8 station/8 GHz wide coarse resolution unit for continuum, and a 15 station/4 GHz bandwidth unit for spectral work. These two correlators define the basic science sub-arrays, but there are three other sub-arrays available for engineering use. The CARMA correlators are based on the hardware and software of the COBRA correlator (Scott et al. 2003), with expansion of the number of stations, bandwidth and resolution. Hardware and software for CARMA has been under development for over a year. The site permit has been granted and civil construction will begin early in 2004, with initial operations in 2005, and completion in 2006. The tight schedule to first light leaves little room for distractions.

Table 1. CARMA Antennas (from Scott et al. 2003)

Number of antennas	Diameter(m)	Status	Organization
6	10.4	Exists at OVRO	Caltech
9	6.1	Exists at Hat Creek	BIMA
8	3.5	New	U. Chicago

The different sizes of the 23 antennas make CARMA a heterogeneous array, and while this poses some technical challenges, it has advantages in image reconstruction as shown by Wright (1999) and Mundy & Scott (2000). Heterogeneous imaging is just one of several unique aspects of CARMA. University culture makes it ideal for training young astronomers, and the readily accessible site and relatively small number of antennas promotes instrumentation development. The northern hemisphere coverage of CARMA will complement ALMA.

3. Monitor and Control System

The monitor and control system is implemented as a distributed computing system, with the Array Control Computer (ACC) playing a central role. The next layer in the computing hierarchy is composed of Intel/Linux nodes distributed with the hardware, such as in an antenna or in a crate of correlator hardware. These nodes in turn control the hardware via embedded micros over CANbus

¹<http://www.mmarray.org/>

(as in the antennas) or over the PCI bus (for correlators). Communication between the nodes is done with CORBA. Coding is done in C++ to allow a single language to be used from the the device driver on up through the system, and sufficient compute power and queueing obviate the need for an RTOS. The system is designed to tolerate failures in the hardware (and even software!) and to fail gracefully. Ideally, only the data from failed components are affected and the data collection continues from the rest of the array while the observer and technical staff are notified of a fault so that repair can be initiated. The monitor and control parts of the system have fundamentally different data flow requirements which their implementation reflects. More on the control and monitor systems are found in Gwon et al. (2004) and Amarnath et al. (2004).

Controls are initiated by observer commands, either interactively or through scripts. These commands generally need to be distributed to different parts of the array, such as the antennas, but a few will initiate procedures associated with data collection that may last for minutes. There are also a few pieces of state information that must be periodically recomputed and sent out to the hardware, such as source positions and frequencies, but in general the rate of command flow is quite low. The antenna API presents a uniform interface to the control system for all of the antennas, thus simplifying its task.

The monitor system works in the opposite direction from the control system, with data regularly flowing from the distributed components back to the ACC. The basic design rationale is to monitor everything possible. Both the astronomical visibility data and the monitor data are collected on synchronized half second frames, although monitor sample rates of up to 100 Hz translate to multiple samples in the frame. This allows precise collation of both streams, enabling debugging of instrumental problems that could otherwise prove difficult. The monitor data is available in the ACC as a source for operator and engineering displays and as input into the fault diagnostic system.

A tight coupling of the visibility data and the monitor data is an integral part of the system design. The continuum visibility data and all monitor data points are written to database storage on every half second frame. This fast sampled data store is not persistent but is recycled on the timescale of about a month, allowing problems requiring high time resolution to be addressed. The permanent archive has the average, minimum and maximum values for all monitor points on both a one minute timescale and for each requested astronomical integration. The one minute data guarantees monitor data even when the instrument is not taking visibility data (slews, bad weather).

4. Archive and Imaging Pipeline

The average data rate for CARMA is about 14 GB/day with a peak that is ten times the average. The format of the archive data has headers and monitor data in an RDBMS and the visibility data in flat files. The use of an RDBMS allows searching and other functions to use the inherent query facilities, while storing the visibility data in flat files conserves space for bulky spectra that are only accessed in conjunction with headers. There will be a temporary archive at the CARMA high site to store data until it is moved over the Internet to the University of Illinois/NCSA permanent archive. Users will obtain their data

from the permanent archive in their choice of supported export format: miriad, FITS, or mir. The archive and data transport mechanisms are based on the current BIMA archive system, with plans for substantial reuse.

The imaging pipeline will also be implemented at NCSA. When a PI specifies an experiment, data processing information will be included which is then passed on to the pipeline. The pipeline will be implemented using the miriad data processing package. Images will be available along with the raw u,v data, allowing the PI to determine if further image processing is necessary. After a proprietary period, the images will be available through an image archive.

5. Software Development

Because of the distributed nature of the software team, the development process has been clearly defined. All work passes through three design/review milestones: conceptual, preliminary, and critical. There is a design emphasis on interfaces. Code is reviewed for adherence to the CARMA coding standard, and unit test code is required with a goal of 75% coverage. The concurrent versioning system (CVS) is used for revision control and code distribution while doxygen is used for embedded documentation extraction. Tinderbox is used as a continuous build and test system and Bugzilla for a defect tracking system. Communication is critical to a distributed team, and we use weekly telecons, email exploders, and face to face meetings. The general philosophy is to treat software engineering very much like hardware engineering.

References

- Amarnath, N. S., Scott, S. L., Kraybill, J. C., Beard, A. D., Daniel, P., Gwon, C., Hobbs, R., Leitch, E., Mehringer, D., Plante, R., Pound, M. W., Rauch, K. P., Teuben, P. J. 2004, "The CARMA Monitor System", this volume, 720
- Gwon, C., Beard, A. D., Daniel, P., Hobbs, R., Scott, S. L., Kraybill, J. C., Leitch, E., Mehringer, D., Plante, R., Amarnath, N. S., Pound, M. W., Rauch, K. P., Teuben, P. J. 2004, "The CARMA Control System", this volume, 708
- Mundy, L. G., Scott, S. L. 2000, "CARMA: Combined Array for Millimeter-wave Astronomy", *Imaging at Radio through Submillimeter Wavelengths* ed. J. Mangum & S.J.E. Radford, ASP Conf. Series, 217, 306
- Scott, S. L., Hobbs, R., Beard, A. D., Daniel, P., Mehringer, D. M., Plante, R., Kraybill, J. C., Wright, M., Leitch, E., Amarnath, N. S., Pound, M. W., Rauch, K. P., Teuben, P. J. 2003, "The COBRA/CARMA Correlator Data Processing System", in ASP Conf. Ser., Vol. 295, ADASS XII, ed. H. E. Payne, R. I. Jedrzejewski, & R. N. Hook (San Francisco: ASP), 265
- Wright, M. C. H. 1999, BIMA Memo Series², Memo No. 73

²<http://bima.astro.umd.edu/bima/memo/>