CARMA

Combined Array for Research in Millimeter-Wave Astronomy

CARMA Memorandum Series #28

Revised CARMA Correlator Design Considerations

Kevin P. Rauch (UMD) and David W. Hawkins (Caltech)

April 25, 2006

ABSTRACT

This document discusses the high-level digital hardware design issues for the next-generation CARMA correlator that directly impact science return value. As prototyping of the revised hardware is in its initial phase, we consider only the most fundamental, design-neutral trade-offs. Our intent is to introduce and promote awareness of these design options, not (at this time) to provide a definitive list of actionable alternatives.

Change Record

Revision	Date	Author	Sections/Pages Affected	Remarks
1.0	2004-August-4	Kevin Rauch		Initial release.
1.1	2004-August-24	Kevin Rauch		Added sections 3.2 and 3.3 discussing options for digital downconversion and phase correction.
1.2	2006-April-25	Kevin Rauch		Updated resolution figures for COBRA-based bands and removed entries for the (unimplemented) 250 MHz and 125 MHz bandwidths.

1. The Revised CARMA Correlator

The 15-station CARMA first-light correlator (Beasley et al. 2003) will contain 3 bands (1.5 GHz total bandwidth in wideband mode) of COBRA hardware recycled from the original 8 band, 6-station OVRO system (including spares). This reconfiguration involved reprogramming the digital logic and developing specialized hardware (such as digitizer fanout cards) to enable calculation of the additional baselines. These efforts have been completed and final verification testing of the updated hardware is in progress.

An additional 5 correlator bands (2.5 GHz) are required to cover the full 4 GHz IF initially available with CARMA; this requires the purchase of new hardware. The baseline plan has been to replicate the existing COBRA design for the remaining 5 bands, and is provided for in version 6.2 of the CARMA budget. However, the A/D converter chips used in the COBRA digitizer modules are out of production and no longer available; therefore, a redesign of the digitizer boards prior to first-light is *mandatory* to achieve 4 GHz bandwidth coverage. Since a digitizer board contains a superset of the functionality of a correlator board, this task will result in a matching, updated correlator card design as well.

Preliminary design work on the revised CARMA correlator digital hardware is currently underway. Due to the continuation of Moore's Law, the new system can provide significantly improved performance at a total cost compatible with that originally budgeted. This memo examines the high-level design trade-offs which directly determine the system's scientific capabilities. As prototype hardware is not yet available, the focus here is on parameters that are largely design-independent. We stress that all quoted figures are initial estimates only, not guarantees, and that other considerations may ultimately limit final performance.

2. Design Performance Matrix

Given a fixed total IF bandwidth, the most important metrics determining correlator performance are operating efficiency and spectral resolution, both of which vary as a function of total observed bandwidth (i.e., the mode in which each correlator band is used). In digital hardware terms, the corresponding variables are the precision and throughput of the processing logic. Precision in this case refers mainly to the accuracy (sample bit-widths) of the digital filtering and correlation logic; the 2-bit deleted-inner-product correlation scheme used in COBRA, for example, provides a maximum observing efficiency of 87.2% (neglecting analog signal losses). Throughput is the total number of samples per second that can be processed, and depends critically on the maximum operating frequency of the digital logic and the total amount of logic available. Since precision and throughput are competing requirements, a suitable balance needs to be struck. The guide for achieving this is maximizing the instrument's science return value subject to cost constraints.

We expect the revised digitizer boards to support higher precision sampling (4-bit or more) than the 2-bit samples generated by COBRA hardware. At a minimum, this will allow digital filtering and decimation (used to implement the narrowband modes) to be performed with little or no loss in overall efficiency; for the COBRA-based bands, this loss will be a few percent. In principle, the correlation scheme itself can also be upgraded—efficiency would rise to 96.0% using 3-bit samples (a 10% improvement) or 98.8% using

4-bit samples (a 13% increase), compared to 87.2% for COBRA; the trade-off for implementing this option is a reduction in channel resolution, ranging from a factor of ≈ 1.5x for the narrowband modes (62 MHz and below) to a loss of 3x-4x for wideband (500 MHz). For continuum observations (now done in wideband mode), the loss in resolution is irrelevant and this high-efficiency observing option would be quite desirable. Unfortunately, supporting such a "quasi-continuum" mode is unlikely to be feasible with the upcoming revision, due to the (excessively) high rate of data transmission between digitizer and correlator boards it implies. For the spectral modes, this high-efficiency alternative is viable. Comparative resolution estimates will be given in the next section.

The maximum operating frequency of FPGA (programmable logic) devices, which provide the logic resources for implementing the correlator design, is limited by semiconductor process technology. Improving net performance therefore requires increasing the total (effective) logic available per antenna baseline, which in turn is set by the total number of observing bands and the number of FPGAs devoted to each baseline within a band, as well as by the effectiveness of the design (the amount of logic needed to accomplish a specific task). In our case the latter parameter, design efficiency, is dominated by the sampling rate (i.e., total bandwidth) of the band in question. If the sample frequency (1 GHz in wideband mode) exceeds the logic operating frequency (125 MHz for COBRA hardware), parallel logic structures operating on time-demultiplexed samples must be used to cope with the high input data rate. This seriously erodes the spectral resolution of wide bandwidth modes (greater than 62.5 MHz). In terms of science, it is therefore of prime importance to have a clear notion of the useful channel resolution *as a function of total bandwidth* for the anticipated key projects.

3. Practical Design Options

3.1. Programmable Logic Capacity

The FPGAs under consideration for the revised correlator (the Altera Stratix II family) are offered in a range of logic densities, from 3 to 30 times as dense as the FLEX 10K100E devices used by COBRA. The latter sell for $\approx$ \$240 each and account for approximately 35% of the total cost of each COBRA-based band, analog hardware excluded. Based on Stratix II pricing at the time of writing, the 3x density devices are somewhat cheaper (\$175 ?) and the 6x variant a little more expensive (\$300 ?) than a FLEX 10K100E FPGA. Hence it is likely that one of these will be used to produce a cost-effective correlator design. An interesting possible design (and science) trade-off therefore presents itself. Namely, since FPGAs represent a substantial fraction of the total hardware budget, and since the cost of an FPGA is roughly proportional to the amount of logic it contains, it would be possible to use the cheaper, less dense FPGAs (or fewer total FPGAs) in exchange for purchasing one or more additional bands, at comparable total cost. This would reduce the resolution proportionately in all modes; however, the increased observing flexibility provided by the additional bands could outweigh this loss, *if* few projects benefit from high resolution over a wide bandwidth. Which to purchase depends on which case is most typical. Depending on available funds and the primary scientific goals, therefore, the high-level design options are as follows:

1. *Lowest cost:* Purchase only 5 additional bands (the minimum required to reach 4 GHz total coverage), using the cheapest, 3x density FPGAs. Would provide a 3x increase in channels relative to the COBRA-based bands and savings of up to $100,000 compared to the original budget.

2. *Neutral cost I:* Purchase 5 additional bands, using the 6x density chips. Offers double the channel resolution of option (1) at a cost similar to that already budgeted.

3. *Neutral cost II:* Purchase 6 additional bands, using the cheapest FPGAs. Offers additional observing flexibility compared to option (1) at a cost similar to that already budgeted.

Note that the ability to purchase additional bands (given additional funds), either up front or at a later time, exists with all three alternatives. Also keep in mind that 3 COBRA-derived bands will be available whichever alternative is chosen. Comparing options (2) and (3), the former would be preferred by projects requiring high resolution over a large bandwidth, such as spectral band scans within our galaxy. Option (3) (or adding more bands in general) offers more flexibility; for example, 8 bands could be used to cover the 4 GHz IF and maximize continuum sensitivity, with the remaining, added bands available to target individual spectral lines. Recall that the frequency and bandwidth of each CARMA band will be individually tunable. To put these choices into perspective, Tables 1-3 list current resolution estimates for the COBRA-based bands, and for option (2) optimized for either resolution or efficiency, respectively. Note that the figures are single-window values; all bands contain two sidebands. The uncertainty in the COBRA figures is less than 10%; for the revised hardware, roughly 50%. Estimated channels for option (1) or (3) are simply half those of (2).

We stress that the numbers presented here are initial estimates only, and that other constraints (such as total power dissipation and data processing implications) must also be respected during the design process, and may limit the final performance figures. As a practical matter, for first-light the decision to optimize for resolution or efficiency is also mutually exclusive. Although no hardware changes are implied either way, the additional effort to program, test, and integrate system support for *both* possibilities would divert attention from critical first-light tasks. One optimization should be chosen for initial implementation; the other can be investigated at a suitable time after first-light. Maximum resolution (Table 2) represents the least effort as it duplicates the existing COBRA scheme. In any event, a representative list from the observer community of target resolution as a function of bandwidth would ensure that design choices are tuned to science objectives to the fullest extent.

3.2. Digital Downconversion

The COBRA hardware requires analog spectral line downconverters to implement spectral line modes for CARMA, due to the limited logic capacity of these boards. These downconverters contain four individual analog filters (500, 250, 125, and 62 MHz wide) costing around $200 each, making them more expensive than wideband downconverters, which contain only a 500 MHz filter. Digital filtering of the 62 MHz band by the digitizers is used to create all narrower observing bands (Rauch 2003).

The increased logic density in the revised digitizers implies that wider bandwidth modes can also be digitally manufactured, obviating the need for some or all of the sub-500 MHz analog filters. The increase required to accomplish this is easy to estimate. The FIR filters in the COBRA boards can operate at a clock frequency of 62 MHz, allowing bandwidths of 31 MHz and less to be created digitally (only one such FIR filter will fit in a COBRA digitizer FPGA). Sample place-and-routes indicate that filters in the new digitizer FPGAs can operate at a 125 MHz clock rate, hence the 62 MHz band can be created with no additional logic, assuming the same input sample bit-width and number of coefficients (filter taps) is used. Digital generation of the 125 and 250 MHz observing bands would require 2 or 4 FIR filters operating in parallel, respectively. The implied logic increase is well within the density range offered by the new FPGAs, making this a viable option. Note that in reality this factor is a lower limit: since both the digitizer sample bit-width and channel resolution are expected to increase, it will be desirable to increase the length (number of taps) and accuracy of the FIR filter to keep pace. A density increase of 8 or more is therefore advisable. The Stratix II development boards contain 12x density FPGAs, allowing full-scale testing to be done.

Although it will therefore be *possible* to replace spectral line downconverters with wideband downconverters using the revised hardware, this is not the optimal solution in terms of signal fidelity. In particular, the 500 MHz analog filter has been optimized for wideband efficiency, at some cost in band flatness (2 dB of response variation). All digitally created bands will inherit these response ripples, though the narrowest bands will largely resolve them out. Ideally, one additional analog filter should be used, a 250+ MHz filter optimized purely for response flatness; since steep bandpass edges are **not** required of it (the edges will be sculpted digitally), it should be easy to design. In this approach, spectral line downconverter modules would still be employed, but they would be loaded with only two analog filters instead of four. In budgetary terms, the added cost of using two filters instead of one is approximately $3,400 per band ($200 per filter by 17 downconverter modules); the gain would be a nearly 2 dB decrease in bandpass ripple in the 250 MHz and narrower observing bands. Note that a savings of about $6,800 per band is realized even in this case relative to fully loaded downconverter modules (with 4 analog filters), as are required for the COBRA-based bands. This savings would substantially offset the extra cost of the denser FPGAs required to implement this advanced processing option (see below).

3.3. Phase Flattening

COBRA correlator and digitizer cards contain a general purpose DSP chip that is responsible for such tasks as retrieving the correlation lags from the FPGAs, computing and normalizing their spectra via an FFT, and accumulating and later transferring the spectra to the Linux crate computer. DSP workload is dominated by the FFT of the lags, which must be performed after every 6.25ms integration to apply phase corrections before the spectra are accumulated. The corrections remove both a phase slope (due to sub-ns antenna delays) and offset (due to lobe rotation differences) from the spectra, which typically vary from one integration to the next.

The burden on the DSP can be greatly reduced by applying the phase corrections to each antenna prior to correlation, which allows the lags for each baseline to be accumulated directly. The FFT then needs to be done only once every 100ms or 500ms, the time interval on which lag spectra are transferred to the host. This can be accomplished through additional signal processing in the digitizer FPGAs. The sub-ns delay component can be removed using fractional sample delay (FD) filters, for which numerous design techniques exist (see Laakso, Valimaki, and Karjalainen 1996). The lobe rotation offset can be removed using a numerically controlled oscillator (NCO) with adjustable phase, for which the FPGA vendor provides a black-box design component.

Shifting phase corrections to the digitizer FPGAs is highly desirable for the revised hardware. Since the DSP load increases super-linearly with the number of lags (FFT length), an expensive, powerful processor would be required to apply baseline phase corrections on the revised boards, due to their much improved spectral resolution. We have performed simulations to determine the amount of logic required to implement this option in a Stratix II FPGA. To remove phase offsets, an NCO is used to modulate the input stream, which is then filtered/decimated in the same manner by which the COBRA-based bands create their narrow-band modes. Altera provides an optimized NCO component whose logic usage is readily determined. An instantiation suitable for CARMA (-40 dB intermodulation noise, 1 kHz frequency resolution) consumes 200 logic cells, 1 M4K memory block, and 1/2 of a DSP block. Eight NCOs operating in parallel are needed to support wideband mode (demux-by-8 @ 125 MHz), altogether adding 3% to total logic and memory usage in a 12x density Stratix II device (EP2S60). Additional DSP blocks can be used to perform the actual modulation (multiplications), allowing maximum logic to be devoted to the filtering and decimation; a 12x density FPGA is sufficient to support wideband mode, the most challenging to implement due to the high degree of parallelism required.

The resources needed to implement phase slope correction was estimated by designing an FD filter meeting the CARMA phase error requirement (0.5 degrees per baseline, or 0.25 degrees per antenna), which was then synthesized in a Stratix II device. The performance of the chosen filter set is presented in Figure 1. The figure displays frequency response in the top panel, and phase error (relative to the desired delay) in the lower window. Horizontal lines demarcate the CARMA accuracy goals; the vertical line marked "Analog BW" indicates the -15 dB rejection bandwidth of the 500 MHz analog filter: response and delay variations beyond this line are of little interest. A different set of filter coefficients is required for each possible fractional delay; the solid lines display the performance at one particular delay, while the dotted envelope shows the maximum error over all possible delays. Within the usable analog bandwidth, the FD filter response varies by only a few hundredths of a dB, and delay error remains below the CARMA limit. This filter requires 400 logic cells and either 16 M512 or 8 M4K memory blocks per bit of input to implement. Supporting demux-by-8 with 4-bit samples will therefore consume 20% of the logic and 60% of the M512/M4K memory in a 12x density device (EP2S60); with 6-bit samples—sufficient to support 4-bit cross-correlations—the figures rise to 30% and 90%, respectively. One M-RAM block is also needed to store the filter coefficient set (an EP2S60 contains two).

We conclude that two 12x density FPGAs per antenna—one devoted to NCO modulation and filtering/decimation, the other to sub-ns delay correction, auto-correlation calculation, etc.—can support digital downconversion and phase correction in all observing modes. This is the same number of FPGAs devoted to each antenna

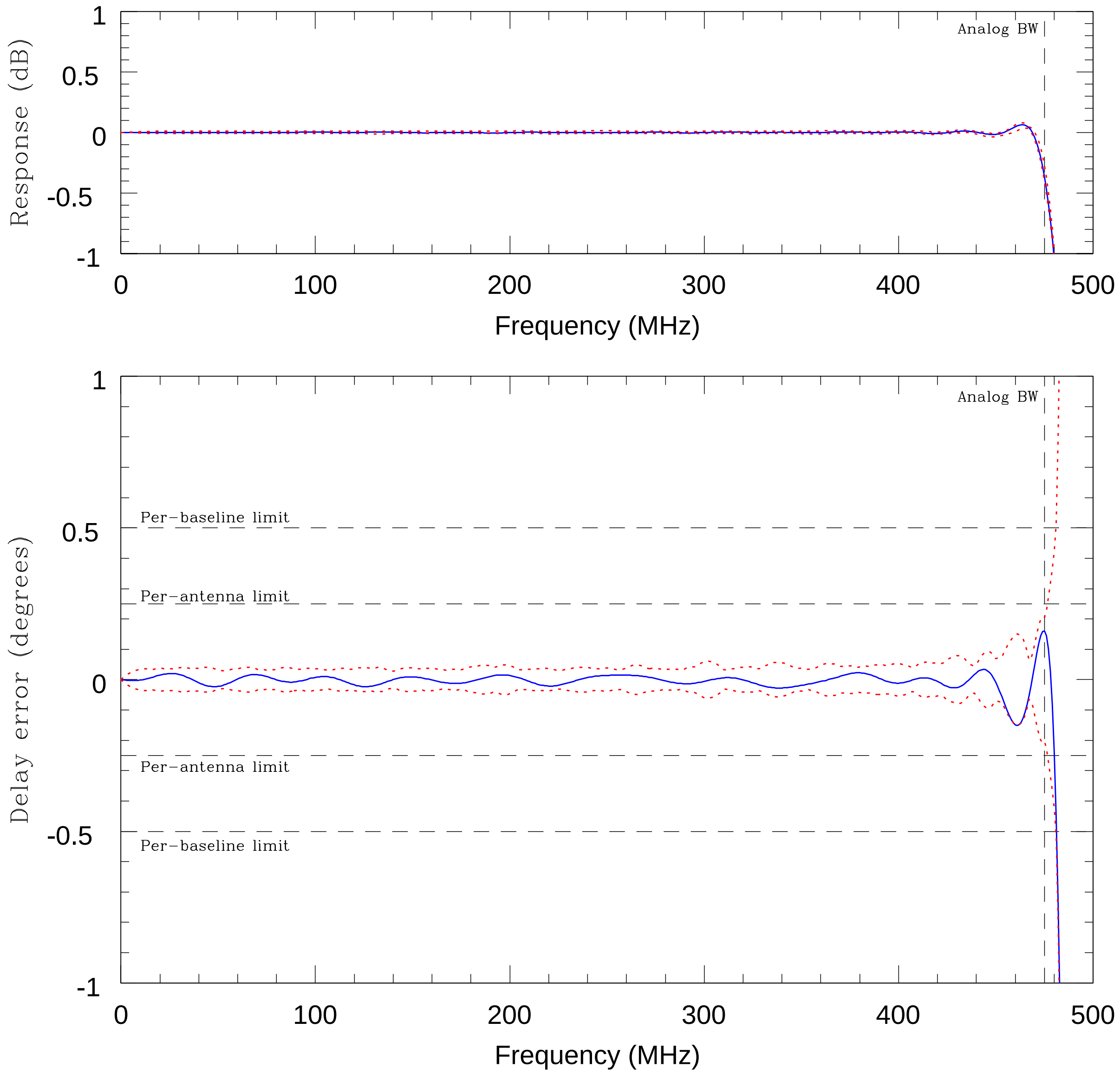

Fig. 1.— Fractional sample delay filter performance for a 64-tap filter set suitable for use with CARMA. The top and bottom windows display frequency and delay responses, respectively. See § 3.3 for details.

in the COBRA digitizers. Assuming the cost per logic cell is constant across devices, and that 6x density devices are used in the matching correlator boards, the cost of this support is equivalent to four 6x FPGAs, or about $1200 per digitizer card. After accounting for the savings it enables (fewer analog filters, a low-cost CPU or DSP, and simplified board design), we estimate little to no net cost increase to implement this enhanced functionality. More precise figures are unavailable until a prototype board design is completed. It may also be possible to apply the sub-ns delays (only) directly in the 1 GHz clock generation hardware; this possibility will be investigated during hardware development testing.

REFERENCES

Beasley, A.J., Hawkins, D.W., Rauch, K.P., and Woody, D.P. 2003, CARMA Memo 11.

Laakso, T.I., Valimaki, V., and Karjalainen, M. 1996, IEEE Signal Processing Magazine, vol. 13, no. 1, pp. 30-60.

Rauch, K.P. 2003, CARMA Memo 12.

Table 1. Spectral resolutions for COBRA-based correlator bands

Bandwidth (MHz)	Channels (per sideband)	δV[3 mm] (km/s)	V_{tot}[3 mm] (km/s)	δV[1 mm] (km/s)	V_{tot}[1 mm] (km/s)
500	17	94	1500	31	500
62	57	3.4	188	1.1	62.5
31	57	1.7	93.8	0.56	31.2
8	57	0.42	23.4	0.14	7.81
2	61	0.10	5.86	0.03	1.95

Table 2. Estimated spectral resolutions for option (2) [maximum resolution]

Bandwidth (MHz)	Channels (per sideband)	δV[3 mm] (km/s)	$V_{\rm tot}$[3 mm] (km/s)	δV[1 mm] (km/s)	$V_{\rm tot}$[1 mm] (km/s)
500	100	15	1500	5	500
250	175	4.3	750	1.4	250
125	250	1.5	375	0.5	125
62	300	0.63	188	0.21	62.5
31	350	0.27	93.8	0.09	31.2
8	400	0.06	23.4	0.02	7.81
2	400	0.015	5.86	0.005	1.95

Table 3. Estimated spectral resolutions for option (2) [maximum efficiency]

Bandwidth (MHz)	Channels (per sideband)	δV[3 mm] (km/s)	$V_{\rm tot}$[3 mm] (km/s)	δV[1 mm] (km/s)	$V_{\rm tot}$[1 mm] (km/s)
500[a]	100[a]	15[a]	1500	5[a]	500
250	75	10	750	3.3	250
125	125	3	375	1	125
62	200	0.94	188	0.31	62.5
31	225	0.42	93.8	0.14	31.2
8	250	0.09	23.4	0.03	7.81
2	250	0.023	5.86	0.008	1.95

[a]Ability to optimize efficiency not anticipated; see § 2. The values (and efficiency) in Table 2 apply.