

8 CORRELATOR

*Brent Carlson, Dave Fort
Martin Pokorny, Bruce Rowen
Last changed 2007-Aug-23*

Revision History:

2001-June-06: Initial release.

2001-July-17: Updates from M. Rupen, J. Romney comments. Add milestone table, table of “impacts and interfaces” to the rest of the system, and risk assessment table. Add clarification text to many sub-sections. Add sub-section on sub-band stitching. Upgrade the correlator layout diagram. Revise module costs to include refined pricing and circuit boards and cables for 32 stations, with racks for 40 stations.

2001-August-14: Minor revisions based on additional Rupen and Romney comments. First draft full release.

2002-February-1: Overhaul based on new Gigabit Ethernet output and Backend configuration. Add more M&C S/W and H/W details. Many refinements to many sections based on design refinements over the last several months. Due to cost increases, NRC will now only pay for and install a 32-station correlator (i.e. not racks for 40 stations).

2002-May-22: Add Correlator Backend discussion to Summary and Introduction. Insert Correlator Backend Requirements and Design sections **8.4** and **8.5**.

2003-Aug-15: General update includes: upgraded sharp-cutoff FIR capability; explicitly state that the correlator is capable of up to 4 million spectral channels per baseline with recirculation; recirculation on 4 streams is now a requirement; specification of stream statistics capabilities; one phase-cal extractor in every FIR chip; solidification of Correlator Chip capabilities; solidification of delay capabilities including narrow baseband very fine delay tracking; VSI_H I/O on the Station Board for VLBI-ready capability; the software operating system choice is Linux at all levels; updated risk assessment table.

2003-Aug-25: Update the section on RFI mitigation to include wideband and sub-band data valid blanking requirements.

2004-Nov-15: Major upgrade to take into account refined design information. No reduction in system functionality or performance, except the correlator chip uses a 3-level rather than 5-level phase rotator, reducing sensitivity by an additional 1.75%. Enhanced functionality may be possible with additional funds (e.g. “R2” recirculation on 8 streams).

2006-Mar-30: Table 8-2 (milestone schedule) and Table 8-3 (costs) updated with latest information. More references to new correlator documentation added. Phasing Board can do 4 sub-arrays, 4 sub-bands, 32 stations or 4 sub-arrays, 2 sub-bands, 64 stations. Granularity for phasing decreased to 1 antenna, allowing full antenna selection flexibility. Hardware to connect to correlator for phased-array auto-correlation or VLBI data recording included, with some restrictions. Figures 8-2 and 8-4 have been upgraded. Design includes “R2” recirculation, however enabling it requires more money than committed by NRC for production.

2007-Aug-23: Major overhaul to include changes resulting from “new connectivity scheme” (Carlson, NRC-EVLA Memo# 028). This allows more independent and logical configurations of sub-bands, and allows all bandwidth to be phased all of the time. With an additional GigE switched network, any sub-band can be selected for phased-array VLBI recording of the Mark 5C format. Start using “spectral magnification” instead of “recirculation” nomenclature. Update section on CBE to reflect new design.

Summary

The delivered system is a 32-station correlator, however the scaleable architecture supports up to 256 stations in 8-station increments. Each station is capable of handling a total bandwidth of 16 GHz, arranged as 8, 2 GHz **basebands**. The correlator contains dedicated hardware (lags) for 16,384 spectral channels per baseline at the widest bandwidths and uses “**spectral magnification**” (a.k.a. “**recirculation**”) to provide up to 4 million spectral channels per baseline at narrow(er)-bandwidths (or wide bandwidths with sensitivity losses). The system can flexibly use and deploy spectral channel resources within internally generated and user defined **digital sub-bands**. High performance pulsar processing capabilities are an integral part of the design. The system will be delivered with the capability of phasing all available bandwidth all of the time and, with an additional Gigabit Ethernet switch network, allows VLBI recording of dynamically selected sub-bands using the Mark 5C 10 GigE format. The **Correlator Backend (CBE)** consists of a parallel cluster of commodity computers with high-speed interconnects to the correlator and the image processing and archive system. The CBE is scalable in order to grow with increasing EVLA observational demands and correlator output data volumes, and flexible enough to handle all specified correlator operational modes. The CBE uses standard network communications hardware and software and won’t rely on specialized vendor-specific

implementations. In this chapter, correlator performance specifications are outlined and a reasonably complete design is presented that meets the specifications. The principal performance specifications for the correlator are shown in Table 8-1. Development milestones are shown in Table 8-2.

Table 8-1 EVLA Correlator Principal Performance Specifications.

No. of stations (antennas) (Sec. 8.2.1)	32 (architecture supports up to 256).
Max spectral channels/baseline @ max bandwidth of 2 x 8 GHz = 16 GHz (Sec. 8.2.2, 8.2.8)	16,384 (more with “wideband spectral magnification” and sensitivity losses).
Max spectral channels/cross-correlation with spectral magnification (Sec. 8.2.2, 8.2.8)	262,144 (total 4 million channels per baseline)
Polarization products (Sec. 8.2.3)	1, 2, or 4
No. of basebands/antenna (Sec. 8.2.4)	8 x 2 GHz each (more with narrower bandwidths)
Quantization (Sec. 8.2.10)	1, 2, 3, 4, or 8-bit initial quantization; 4 or 7-bit re-quantization after sub-band filter.
Correlator efficiency (Sec. 8.2.10)	~93.2% (4-bit initial quantization, 4-bit re-quantization, 3-level fringe rotation)
No. of sub-bands per baseband (Sec. 8.2.6)	16 (provision for up to 18 for “N+1” redundancy).
Sub-band bandwidth (Sec. 8.2.8)	128 MHz, 64 MHz, 32 MHz, ..., 31.25 kHz (multi-stage filter). Each sub-band’s width and position can be set independently of any other sub-band.
Sub-band tuning (Sec. 8.2.9)	Each sub-band should remain within an appropriate integer slot to minimize band-edge SNR loss. E.g. a 128 MHz sub-band should be within 1 of 16 equally spaced slots in a 2 GHz band. Greater tuning flexibility at narrower bandwidths is possible.
Spectral dynamic range (Sec. 8.2.10)	(Initial quantization) 3-bit: ~44dB; 4-bit: ~50dB; 8-bit: ~58dB. [Test: 2 “bunches” of 4 tones/bunch, each “bunch” contained within one sub-band (128 MHz); 99% tone (interference) power; ideal samplers; dynamic range measured outside sub-bands containing interference.] With 2 tones only, results are ~10dB, ~2dB, ~4dB (respectively) worse.
Auto-correlations (Sec. 8.2.20)	Wideband (4x2 GHz pairs): 4 products of 1024 spectral channels each, SNR loss =4. Sub-band: at least two auto-correlation products per antenna at a time, no SNR loss.
Pulsar processing (Sec. 8.2.14)	2 banks of 1000 time bins each/baseline. Up to 65,536 bins/baseline with software accumulation. Min. bin width: ~200 μsec (all spectral channels) ~15 μsec (64 spectral channels/sub-band/baseline). Also, pulsar gating with one timer+multi-gate generator per 2 GHz baseband.
Min. dump period (initial installed configuration) (Sec. 8.2.15)	100 milliseconds (all spectral channels). Faster with more CBE computers and/or fewer channels and/or baselines.
CBE input aggregate bandwidth	1.28 Gbytes/sec
CBE output aggregate bandwidth to archive	25 Mbytes/sec; this is the minimum requirement.
Max. dump period (Sec. 8.2.15)	Unlimited (within the CBE)
Maximum baseline (Sec. 8.2.16)	25,000 km with 0.5c FOTS transmission velocity (0.25 sec total delay buffer).
Sub-arrays (Sec. 8.2.18)	Cross-correlation: up to 8 sub-arrays, with each sub-array having a granularity of (any) 4 antennas. Phased-VLA: 2 sub-arrays per sub-band; no granularity restrictions.
Phased-VLA (Sec. 8.2.19)	Phase all bandwidth all of the time. Access to phased data, Mark 5C format via GigE, and internal HM Gbps simultaneously. Simultaneous operation with interferometer modes using same array phase-center.
VLBI (Sec. 8.2.21)	VLBI-ready. Requires additional software, VSI FPGAs on the Station Board, and VSI MDR-80 breakout connector/PCB or 10 GigE to LVDS interface card.
Interference mitigation (Sec. 8.2.23)	Post-corr. Temporal/spectral excision—narrowband interference modulation robust. Possibly provision for post-correlation interference cancellation. Fast RFI blanking: sub-band power over-range detection and data valid blanking with programmable dwell time.

Table 8-2 EVLA Correlator Development Milestones

Milestone	Approximate Date	Notes
Conceptual Design Review (CoDR)	November, 2001	Architecture/features review, specifications/design freeze.
Correlator Backend (CBE) 4 node clus	Q2, 2002	Minimal configuration for initial prototyping.
CBE 8+ node test cluster	Q3, 2002	Minimal configuration for functional testing.
CBE full functionality	Q4, 2003	Ready for system test
Preliminary Design Review (PDR)	Q2/Q3, 2005	Prototype design ready, review before proto. construction.
CBE earliest connect to corr h/w	Q4, 2006	First live testing in Penticton with prototype hardware.
Rack-to-rack cable installation	Q1, 2008	Rack-to-rack cables installed before racks.
Critical design review (CDR)	Q2, 2008	Review before on-the-sky tests and 1 st partial correlator delivery.
On-the-sky test at VLA; 1 st partial correlator delivery.	Q3, 2008	Provides 10 antenna, 1.5 GHz/pol'n capability.
Earliest possible "shared-risk" science	Q1, 2009	10 antenna, 1.5 GHz/pol'n.
Begin full installation at VLA	Q2, 2009	Begin installation of production boards.
Correlator commissioning	Q2, 2010	Full observational mode, turn off old correlator, intermittent NRC support required.
Project complete	Q4, 2010	Scheduled NRC support no longer required.

8.1 Introduction

The EVLA correlator design is based on the WIDAR concept (Carlson, IEE 2000) (Carlson, Memo# 001) (Carlson, Memo# 014) (Carlson, Memo# 024) (Carlson, A25290N0000) where wide (2 GHz) bands are sampled, split into smaller sub-bands with digital filters, and then correlated. A key anti-aliasing technique along with stable and calculable digital filter characteristics, allow the sub-bands to be seamlessly "stitched" together to yield the wideband cross-power spectrum. Using this technique it is possible to correlate data efficiently so that about an order-of-magnitude more spectral channels can be provided compared to what other time-domain parallelization techniques can yield. A design requirement for the EVLA is to provide 16,384 spectral channels per baseline in wideband modes, with more spectral channels available using "spectral magnification" (a.k.a. "recirculation"). Digital sub-banding has the additional benefit of increasing the flexibility of the correlator so that only those spectral regions of interest need use correlator resources. An 'XF' correlator has been chosen primarily to minimize the station hardware-to-baseline hardware bandwidth/cabling requirements—a significant consideration for a correlator system of this size.

The design of the **Correlator Backend (CBE)** is based on the requirement to not have to rely on a specialized high-speed interconnect fabric amongst CBE computers to perform required processing, and to be able to use commodity computers (PCs) and network hardware and software to meet performance goals. The data delivery network from the correlator to the CBE is designed so that each computer has all of the information (lag data) needed for processing one or more baselines—CBE computers do not have to exchange additional high-speed information. By designing the data delivery network in this fashion, processing by the CBE is highly parallelized, providing nearly linear scalability of performance with the addition of processing nodes. Use of multi-CPU processors in the nodes provides for sufficient compute power and flexibility to handle critical real-time input demands as well as data processing, formatting and internal monitor and control activities. Inter-node communications is limited to monitor and control messages that are handled by message passing middleware. *Intra*-node communications are handled by message passing middleware, and standard IPC (*i.e.*, inter-process communication) mechanisms.

8.2 Specifications

8.2.1 Number of Stations (Antennas)

The installation includes a full population of 32 stations. The architecture supports up to 256 stations in 8-station increments. Connectivity is optimized for 32 stations, at 16 GHz/station of bandwidth. Expansion beyond 32 stations can be done at full or partial bandwidth.

8.2.2 Spectral Channel Capability

Dedicated correlator resources (lags) for 16,384 spectral channels/baseline at the widest bandwidths are available. Spectral channels can be flexibly deployed to desired sub-bands/basebands. “Spectral magnification” (recirculation) provides a maximum of 262,144 spectral channels per *cross-correlation* for 1 polarization product, and 16,384 spectral channels per cross-correlation for 2 or 4 polarization products. Each sub-band can have a different bandwidth and magnification factor. Spectral magnification works by time-multiplexing the acquisition of correlator lags using synthesized lag delays in a memory buffer. The amount of time multiplexing is known as the **magnification (recirculation) factor**. In narrow(er)-band modes where the bandwidth reduction is the same as the magnification factor, no sensitivity degradation is realized in the cross-power spectrum. If the magnification factor is *greater* than the bandwidth reduction, there is a *root(magnification factor/bandwidth reduction factor)* decrease in sensitivity. Magnification can be used at maximum sub-band bandwidth (128 MHz) with the above indicated sensitivity reduction (referred to as **wideband spectral magnification**). When magnification is used, the correlator dump time and/or minimum phase-bin time is increased since it is necessary to obtain at least one pass of all lag data in each dump to produce a proper spectrum. The time increase factor is the same as the magnification factor.

8.2.3 Polarization

Basebands can be flexibly arranged as combinations of dual-polarization **baseband pairs** and single-polarization basebands (subject to antenna system flexibility). 1, 2, or 4 polarization products can be correlated and these are selectable on a baseband/sub-band basis.

8.2.4 Sampled Baseband Capacity

Each “station input” can handle 8, 2.048 GHz basebands sampled at 4.096 Gs/s. More sampled bands—up to 128 per station input—*could* be handled if they had less bandwidth each. This could be useful if it is desired to process more (narrower) sampled basebands (for example, to avoid regions of extreme RFI), but this is currently not an EVLA requirement. The correlator can flexibly handle various combinations and numbers of sampled bands provided sample rates are properly related.

8.2.5 Baseband Tuning

Basebands can be at any “sky” frequencies and any restrictions are governed entirely by antenna LO system flexibility.

8.2.6 Digital Sub-band Capability

The correlator has provision for up to 18 digital **FIR (Finite Impulse Response)** filters—implemented in a **Filter Chip**—for each 2 GHz baseband input. Typically, one of the sub-bands is used for receiver switching noise diode measurements (i.e. so it can use a sub-band of the baseband with no [time-variable] interference in it for system noise temperature calibration). The Filter Chip consists of up to 4 stages of filtering and, depending on configuration, can provide sharp-cutoff sub-bands as narrow as 31.25 kHz. Thus, “radar-mode” capability is effectively built into each Filter Chip. Refer to section 8.2.12 for more detailed information.

The *delivered* correlator is populated with 4 quadrants, and each of these quadrants processes all baselines for all sub-bands of one baseband pair. If one or more basebands (or sub-bands of basebands) are not used, the quadrants’ resources may be deployed to provide additional correlator resources for the basebands/sub-bands that are processed.

Each sub-band can independently be set for bandwidth and spectral magnification factor. Additionally, provision is made so that each *sub-band* could be on a different delay-center on the sky to support multi-beaming *within* a baseband. The maximum delay-center offset from the baseband's delay center is currently +/-16 μ sec.

8.2.7 Sub-band Stitching

Adjacent sub-bands can be seamlessly “stitched” together with a maximum sensitivity loss of a factor of 2 at the **sub-band boundary**. The rate of reduction in sensitivity away from the boundary depends on the “steepness” of the filter transition band. (Typically, with a flat passband -6 dB cutoff filter and 511 taps, the sensitivity loss is less than 20%, 2 MHz away from the sub-band boundary for a 128 MHz passband. This includes sensitivity loss effects from re-quantization and fringe rotation.) Stitching is performed by applying the total power measurements obtained in the Filter Chips *before re-quantization* and by applying calculated digital filter bandshape corrections (Carlson, Memo# 001) (Carlson, IEE 2000). Since the filter is applied with the LO offset in place, and this is removed in the cross-power spectrum result, baseline-based filter bandshape corrections should be applied that include the effective baseline LO offset as it affects the filter amplitudes. Depending on transition-band steepness, this special consideration is normally only required if the LO offset is greater than $\sim 1/10^{\text{th}}$ of the spectral-channel bin width. Each filter's total power measurement (before re-quantization) can only be used properly if the total power gain of each filter is known. This gain is calculable, but also depends on tap-weight scaling (i.e. the scaling of floating-point tap weights to integer bits used in the filter) that should (effectively) be relative to some common reference value for all filters on every Station Board. Depending on sub-band roll-off and narrowband signal strength in the proximity of the sub-band boundary, stitching may require the use of adjacent sub-bands' spectral points and careful windowing operations. Initial quantizer statistics and re-quantizer statistics are obtained in the correlator and are required for accurate data normalization.

8.2.8 Sub-band Bandwidth

Each of the 18 general-purpose digital filters can be configured for an output bandwidth starting at 128 MHz and decreasing in powers of 2 down to 31.25 kHz. Sharp cutoff filters are possible at all bandwidths by using some or all of the 4 stages of the Filter Chip. *All* filters are independently configurable in bandwidth and placement within the baseband. Refer to section 8.2.12 for a more detailed description of the Filter Chip.

8.2.9 Sub-band Tuning Flexibility

Digital pass-bands can be placed anywhere within integer “sub-slots” corresponding to the sub-band (slot) width. For example, if the (stage 1) filter has a 1/64 bandpass, then the filter can be placed in any of the 64 evenly spaced slots in the band. More tuning flexibility is provided by stage 2 of the Filter Chip, operating on up to 128 MHz of stage 1 output. Refer to section 8.2.12.

8.2.10 Sample Word Sizes and Correlator Efficiency

The initial baseband sampled word size can be any one of 8, 4, 3, 2, or 1 bits. Each sampled baseband in each antenna could have a different word size as long as the total digital transmission bandwidth does not exceed the fiber-optic transmission system bandwidth. The correlator supports 4-bit initial quantizer word sizes, but for cost reasons, the antennas deliver only 3 bits at 2 GHz baseband bandwidths. Refer to Table 8-1 and (Carlson, Memo# 009) for spectral dynamic range estimates. The correlator supports 8-bit initial sampling, but if used, only $\frac{1}{2}$ the baseband bandwidth is available since the sample word width has doubled. (Each baseband is independently configurable in sample word width.) (N.B. because of frequency shifting, it is possible to use time-interleaved samplers since spectral by-products generated from amplitude mismatches do not show up in the correlator cross-power spectrum.)

After digital FIR filtering, the correlator re-samples the data to 4 bits. Alternatively, in high SNR high dynamic range regions of the spectrum, the correlator can re-sample and correlate 7 bits (Carlson, Memo# 010). If 7-bit correlation is used, then $\frac{1}{2}$ the spectral channels and $\frac{1}{2}$ the sub-band bandwidth is available (because of internal correlator data-path routing limitations). Choice of re-sampling word size can be done on a per sub-band basis. Also, the re-sampling word size does not depend on the initial sampler word size (and vice versa).

Three-bit initial sampling and 4-bit re-sampling, along with 3-level correlator fringe rotation loss, results in a correlator efficiency of about 91% (Carlson, Memo# 011). (Four-bit sampling is ~98.5% efficient, 3-bit is ~96.5% efficient, and 3-level fringe rotation is ~96% efficient (Carlson, Memo# 002). Eight-bit sampling is very close to 100% efficient and thus has a negligible sensitivity loss.) For spectral dynamic range refer to Table 8-1.

8.2.11 Correlator Chip

The Correlator Chip contains 2048 complex-lags, arranged as 16, 128 complex-lag correlators. Adjacent internal complex-lag correlators can be concatenated together. There is no provision for directly concatenating Correlator Chips. Each accumulator is 23 bits long and is *not* truncated for high dynamic range correlation. 23-bit accumulators have a maximum integration time of 500 microseconds, however shorter readout times are needed to support spectral magnification and narrow phase binning. Lag-based, quantized-phase fringe stopping is performed with a 3-level fringe rotator (Carlson, 1999) (Carlson, Memo# 002). The maximum chip data rate is 256 Ms/s.

8.2.12 Digital Filter Chip

The digital filter chip consists of the following functional blocks (operating sequentially on the data from input to output):

1. **Sub-band delay** line of up to 32 (+/-16) microseconds. This delay line may be used for sub-band multi-beaming (Carlson, EVAL-NRC Memo# 014) or it can be (effectively) by-passed.
2. **Stage 1 filter**. This is a 512-tap, 16-phase poly-phase FIR filter with 32 taps per phase, and an integrated 16x16, 4-bit wide cross-bar switch. It operates on 4-bit or 8-bit input data (but only 256 taps on 8-bit data). The conversion of 1, 2 or 3-bit samples to 4-bit samples is done in another chip. For (VLBI-mode) very fine ($\pm 1/32^{\text{nd}}$ of a sample) narrowband delay tracking, each phase of the filter is loaded with sub-sample delay coefficients, and the delay tracker selects the appropriate phase in real time with no blanking as delay changes. The maximum output bandwidth from this filter is 128 MHz, and the practical minimum output bandwidth is ~16 MHz. The output of this filter is 16 bits, and the output can go to stage 2, or the final re-quantizer. Sub-band filtering using this stage should stay within integer slots to minimize the SNR degradation region at the edges of the filter.
3. **Stage 2 filter** with 64, 128, 256 or 512 taps, depending on the decimation factor of 2, 4, 8 or 16 respectively, so that the same *relative* filter cutoff steepness is maintained independent of bandwidth. The maximum input bandwidth of this filter is 128 MHz. This filter contains an integrated digital single-sideband mixer so that output sub-bands are finely tunable (with a 32-bit frequency synthesizer and high dynamic range mixer) anywhere within the input bandwidth. This mixer may be used or by-passed and it is also used for very fine delay tracking when the final output bandwidth of the filter is very narrow. The output of this filter is 16 bits and may go to stage 3 or the final re-quantizer.
4. **Stage 3 filter** with 64 to 512 taps operating the same as stage 2 except that there is no single-sideband mixer. The maximum input bandwidth of this stage is 8 MHz. The output of this filter is 16 bits and may go to stage 4 or the final re-quantizer.
5. **Stage 4 filter** with 512 taps at decimation factors of 2, 4, 8, or 16. The maximum input bandwidth of this stage is 0.5 MHz. The minimum output bandwidth from this is 31.25 kHz in order that there are an integral number of output samples in 10 milliseconds.

Note that since stages 2, 3, and 4 operate on 16 bits, even when all stages are in use, there is only one re-quantization loss in the Filter Chip—that of the final re-quantizer.

6. **Pre-re-quantizer power meter**, with two accumulation bins for noise diode calibration. This power meter operates on the 16 bits out of one of the filter stages before re-quantization.
7. **Fast RFI detector/blanker**. This block operates on the 16-bit sub-band output data before re-quantization. It has a programmable threshold that when triggered blanks the output data valid for a programmable dwell time.
8. **Final re-quantizer**. Sixteen bits from one of the four filter stages is selected to be re-quantized to 4, 5, 6, 7 or 8 bits by this block. Note that the Baseline Board can only accept 4 or 7-bit samples.
9. **Sideband flipper**. This flips the sign of every other sample to change the frequency sense of the output sub-band. This block can be enabled or disabled.

-
10. **Phase-cal extractor.** This is a tone extractor with a 32-bit frequency synthesizer that operates on the re-quantized data.
 11. **State counters and power meter.** This block acquires re-quantizer output statistics. Refer to section 8.2.20.

The final output of the Filter Chip is 4 bits wide, and contains the 4-bit or 7-bit (time-multiplexed) re-quantized data.

8.2.13 Radar Mode

Each Filter Chip has the ability to output sharp-cutoff, narrow sub-bands required by radar mode and so “radar mode” is effectively no longer a separate mode of the correlator. However, radar processing requires 1 Hz resolution on ~30 kHz bandwidth while simultaneously obtaining reasonable spectral resolution on the wide 2 GHz baseband that the 30 kHz is part of. One way this can be accomplished is by using two, 128-lag correlator cells within one sub-band of one quadrant of the correlator and spectral magnification $\times 256$ to get 1 Hz resolution on the 31.25 kHz, while using another quadrant of the correlator to correlate all of the required sub-bands that make up the 2 GHz baseband. Additionally, it is possible for the CMIB to capture and readout filtered and re-quantized data for software processing. The maximum filter output bandwidth for which all data can be captured is 4 Msample/sec, subject to CMIB and network performance. The minimum specification is to be able to capture all data (8-bit or 4-bit samples) at a 31.25 kHz bandwidth for one sub-band of each baseband. Currently, two 31.25 kHz 8-bit sub-bands can be captured and wider bandwidths with correspondingly fewer bits are also possible.

8.2.14 Pulsar Processing

There are 2 banks of **1000** time bins each per baseline. One bank is active while the other bank is being downloaded to CBE (Correlator BackEnd) computers. Alternatively, 1 bank of 2000 time bins can be used if correlator dead time while downloading data is acceptable. If all spectral channels are dumped, then the minimum bin width is ~200 μsec ; if only 64 spectral channels/sub-band/baseline are dumped, then the time bin can be as narrow as ~15 μsec . Up to 65,536 bins/baseline can be accommodated with CBE computer software accumulation. Pulsar gating with one timer and multi-gate generator per 2 GHz baseband is available. The multi-gate generator can produce 16 pulsar gates with configurable delays relative to the timer epoch so that each sub-band can be gated “on” at different times to track different pulse arrival times at different frequencies.

8.2.15 Real-Time Data Output Performance

The real-time data output performance is governed by several factors. The correlator hardware itself has a very wideband data output pipeline so it is most likely that any performance limitations are determined by the performance and configuration of the CBE computers. The minimum dump period for all spectral channels if the extreme, highest-performance output pipeline is used (4, 1 Gbit/sec links) is ~2.6 msec. This is a dump rate of ~7 Gvis/sec in a 32-station correlator. The delivered system has a pipeline—*out of the Baseline Boards*—capable of dumping all spectral channels every ~11 milliseconds. With the *planned* number of CBE computers all spectral channels should be able to be dumped about every 100 milliseconds (~167 Mvis/sec in a 32-station correlator). If fewer spectral channels are dumped, then shorter dump times could be obtained. The maximum correlator hardware **LTA (Long Term Accumulator)** integration time is signal-characteristic dependent but is about 16 seconds for low SNR cross-correlations and 2 seconds for auto-correlations. CBE computers can integrate data for an arbitrarily long period of time.

8.2.16 Delay

The delivered correlator contains enough delay buffering for 0.262144 seconds of delay and this translates into ~25,000 km baselines if there is a $0.5c$ FOTS data transmission velocity over the same distance. The delay may be increased in the future by replacing the Delay Module mezzanine card (on all Station Boards), if desired. The delay rate that the correlator can handle is limited only by the delay synthesizer update rate of 64 MHz. Each baseband can have its own independent delay model and hence independent delay center on the sky. Precision, fully digital $\pm 1/32^{\text{nd}}$ of a sample delay tracking on 2 GHz basebands is a feature of the WIDAR architecture (Carlson, Memo# 007). There is no associated data blanking as the correlator tracks delay. WIDAR sub-sample delay tracking eliminates the need

and uncertainty associated with sampler clock phase modification. In addition, the Filter Chip provides the ability to finely track delay on narrower basebands (≤ 128 MHz) by loading each “phase” of the stage 1 poly-phase FIR filter with delay-interpolation coefficients, and seamlessly selecting the correct phase of the filter to implement the correct sub-sample delay in real time. This method eliminates the need to perform special delay functions on a baseline basis in the Correlator Chip, and provides $\pm 1/16^{\text{th}}$ sample of baseline delay tracking with virtually no restrictions on delay rate.

8.2.17 Doppler/Frequency Shift

The Correlator Chip contains digital complex phase-rotators with effectively no limitations in Doppler phase rate or artificial frequency shift. The rotators are driven by linear digital frequency synthesizers that operate at 64 MHz and whose coefficient can be updated every 10 msec. The fundamental limitation is the sub-band bandwidth, but it is suggested that the maximum phase rate not exceed $\frac{1}{2}$ the widest sub-band so that phase does not contribute to Correlator Chip heating (through fast toggling of CMOS transistors). Digital filter anti-aliasing requires offsetting each antenna’s Local Oscillator by a small amount. It is suggested that this be about 10 kHz, but tunable in 100 Hz steps for narrowband radar mode (Carlson, Memo# 005). There should be an adequate frequency shift between signals/antennas being correlated so that digital mixer edge effects are not apparent and so that sufficient anti-aliasing attenuation occurs. A minimum of 100 cycles of differential phase rotation within an *incoherent* integration period is recommended. If desired, frequency shifts could be dynamic so that anti-aliasing occurs even over arbitrarily long *coherent* integration times.

8.2.18 Sub-arrays

A maximum of 8 sub-arrays are allowed; each sub-array has a granularity of 4 antennas where any 4 antennas may be assigned to a sub-array. Less than a multiple of 4 antennas can be used in a sub-array, with the unused correlator resources not usable in other sub-arrays. Additionally, separate sub-arrays can have mutual antennas as long as the configuration within a sub-array is consistent within the constraints of correlator data routing (and as long as the configuration software is capable of doing this!). Phased-VLA sub-arrays are completely flexible, with 2 sub-arrays per sub-band available.

8.2.19 Phased-VLA

All bandwidth may be phased all of the time, with 2 sub-arrays per sub-band allowed (provided enough 1 GigEthernet connections are available). If so equipped, a separate GigEthernet switch network could allow dynamic selection of sub-bands recorded for VLBI. Phased output on GigEthernet is the Mark 5C format, with each 1 GigEthernet output containing a single phased stream; conversion of multiple 1 GigEthernet packet streams to 10 GigEthernet for Mark 5C data recording is accomplished using a commercial switch.

All phased bandwidth may be auto-correlated all of the time (with 2048 channels for each sub-band) if less than 32 antennas are cross-correlated. In addition, all phased bandwidth is available all of the time via the correlator’s internal “raw” HM Gbps format, and this output may be used for any reason at any time by specialized hardware that is not part of correlator delivery.

8.2.20 Auto-correlations, Data Statistics, and Phase-Cal

Four wideband auto-correlation products are provided for every baseband pair. Each product has 1024 (up to a maximum 32,768) spectral channels, but with a factor of 4 sensitivity loss (sensitivity losses are greater for more than 1024 channels) over an ideal auto-correlation. This loss of sensitivity comes from acquiring the auto-correlations in 64-lag chunks every 10 milliseconds due to hardware limitations. *Sub-band* auto-correlations are acquired with cross-correlator hardware, although only 2 products at a time per antenna may be acquired. Sub-band auto-correlation results may contain transition-band aliasing so it is not possible to seamlessly stitch sub-band auto-correlation spectra together (except where a “cross-auto-correlation” is performed—if the antenna LO system is sufficiently flexible). Sixteen wideband state counters are provided per baseband (64-bit data highway) that, in ≤ 4 -bit mode are time-multiplexed across the input data streams. Time-multiplexing is under CMIB control and parameters can be modified

every 10 milliseconds. In 8-bit input mode, there is one accumulator and it can be set to count occurrences of any of the 256 states in a similar time-multiplexed fashion. After filtering and re-quantization, one accumulator (state counter) is used to time-multiplex the acquisition of state counts across 16 (4-bit) or 128 (7-bit) possible states. Also, full sensitivity total power accumulators are provided for both 4-bit and 7-bit re-quantized data. Finally, each Filter Chip contains a dedicated phase-cal tone extractor with a linear frequency synthesizer and full delay-tracking compensation that operates on the filtered and re-quantized data stream.

8.2.21 VLBI

The correlator is fundamentally a VLBI correlator and the system will be delivered with all of the “hooks” in place for VLBI. Each Station Board has two VSI (VLBI Standard Interface) inputs and two VSI outputs—one of each per baseband (data highway) to allow data to be piped into the Filter Chips from some source (such as a VSI playback device) and out of the Filter Chips to some destination (such as a VSI recording device). Each VSI input or output can handle sixteen 2-bit sampled data streams at rates up to 256 Msamples/sec. VSI signals on the Station Board are broken-out to two connectors, each of which contains a VSI input and a VSI output. These connectors plug into Common Backplanes. To convert from the backplane connectors to standard VSI-H MDR-80 connectors requires special purpose breakout modules, which are also not part of the delivered system. Also, to use these interfaces, VSI FPGAs need to be installed, and this is not currently planned for the EVLA. Refer to Figure 8-3.

8.2.22 Maintenance

All (semi-conductor-populated) modules and module-to-module communications are designed for hot-swap capability. Additionally, the design is such that swapping out one module has the minimum possible impact on other modules and their data products. The estimated MTTR is about 10 minutes (with maintenance personnel on-site). The **total system** MTBF is currently estimated at 77 hours (Carlson, A25010N0003) at the 90% confidence level—meaning that there is a 90% probability that the MTBF is greater than 77 hours. State-of-the-art commercial devices, design, and production techniques are employed for maximum benefit, and an Environmental Stress Screen (ESS) program will be employed in an effort to reduce infant mortality defects from the system. Regular semiconductor failures are not anticipated. All hardware modules have active (via computer) and dead-man (thermal switch) temperature monitoring and shutdown. Separate cooling fan monitors are employed so that fan failures can be detected immediately, rather than waiting for components to heat up. It is possible to remotely power-cycle individual modules using a power control computer that is not part of normal correlator processing (for increased reliability). While the correlator is on-line, embedded synchronization codes allow for constant monitoring of module health and module-to-module communication integrity. When off-line (for example, when slewing antennas) it is possible to enable internal test vectors for complete correlator system testing. The intent is that a test is treated like a normal observation except that, instead of processing data “from the sky”, test vectors are processed instead. *The degree to which antenna and antenna transmission systems are included in this kind of testing is currently undefined.*

8.2.23 Interference Mitigation

The correlator contains some special real-time burst interference nulling hardware. This includes accepting data valid flagging from the antenna and not correlating when it is flagged bad, and detecting saturation before the re-quantizer in the Filter Chip to flag and not correlate invalid sub-band data. Additionally, high-speed dumping (with scaleable performance CBE computing), and high spectral dynamic range provided with many-bit samples enable post-correlation, temporal/spectral excision. The WIDAR design strongly attenuates the modulating effects of time-variable narrowband interference on normalized correlation coefficients (Carlson, Memo# 009), so post-correlation excision of non-saturating burst-like interference should be quite effective. Post-correlation interference cancellation, should it be found to be effective, can easily be handled since the interference detection antenna is just another antenna to the correlator. The correlator also has the capability of processing 8-bit sampled data for high-spectral dynamic range even in the presence of overwhelmingly powerful narrowband interference.

8.2.24 System Timing

All actions in the correlator are synchronized to distributed “TIMECODES” and a 128 MHz clock. An external

TIMECODE and 128 MHz clock is split, provided to the correlator, and distributed internally such that correlator operation is not susceptible to single card or single rack failures. TIMECODE operates on a 1 PPS (pulse per second) basis, and all operation in the correlator is synchronized at chip and board level to this signal. While synchronous clocks are distributed in the correlator, final synchronization is only required and achieved within each chip (FPGA or ASIC) where processing occurs.

8.2.25 Computing and Data Highways

The Correlator installation includes three classes of computers. The top-level monitor and control (MCCC) and power control (CPCC) computers are N+1 redundant units, either CompactPCI or rack-mount PC based systems. Each Station Board and Baseline Board is directly controlled by an embedded PC/104+ format computer called a CMIB (Correlator Module Interface Board) installed as a mezzanine card. All CMIBs are networked via 10/100/1000 COTS Ethernet switches to the MCCC. CMIBs obtain their configuration and real-time control information from the MCCC, and provide status back to the MCCC. The CPCC allows for 1+1 redundant remote power monitor and control independent of the MCCC/CMIB network. CBE data processing computers are COTS (Commercial Off-The-Shelf) “blade” PCs connected to correlator data generators (primarily Baseline Boards) via a monolithic commercial wire-speed 1 GigEthernet switch. This allows data to be routed to any blade PC for load sharing or fault-tolerant operation. The number of blade PCs is TBD, depending on performance, and the requirement to handle 100 msec dump times.

One possible correlator network topology is shown in Figure 8-4—other topologies for monitor and control are possible. All monitor and control network switches are nominally 24-port devices. A standard Linux OS distribution is used in all these computers with the CMIBs running a pre-emptable kernel version to address the tighter timing requirements of this computing system layer. All inter-system communications rely on standard Unix/Linux facilities to minimize development effort and maximize scalability and compatibility with external systems. CMIB hardware control is abstracted behind standard Linux device drivers to allow higher level control processes to be developed with a greater amount of portability should computing platforms be changed/upgraded. All monitor and control communications are message-based to best isolate failing systems from impacting overall system performance and to provide for a highly modular distributed system. Messages are XML (eXtensible Markup Language) based. Use of XML provides correlator system designers and users an industry standard protocol that is both human readable and includes a wealth of industry standard applications and tools to manipulate XML messages. The software package provides the system engineer easy access to correlator hardware internals via sophisticated GUIs as well as provides system users a polished orthogonal interface where the tedious details of hardware setup are hidden to avoid confusion, but still available if desired.

8.2.26 Environment

The correlator is designed for a “benign office environment” with an ambient temperature of $\sim 25^{\circ}\text{C}$ at the altitude of the VLA (Webber, Carlson, A25012N0000). Board and rack design is such that the operating temperature range is 0°C to $+35^{\circ}\text{C}$. However, for reliability the ambient temperature should be kept at about $+15^{\circ}\text{C}$. The (32-station) correlator, including the CBE, requires an estimated 170 kW of power. All correlator modules and racks (except the CBE, MCCC, and CPCC computers) operate from a centralized -48 VDC N+1 redundant, on-line serviceable power plant with 5 minutes of full-power backup, which has now been installed in the correlator room. The CBE, MCCC, and CPCC computers operate from standard battery-backed 110 VAC.

There are a total of 8 station racks and 8 baseline racks, each rack fully loaded with 16 Station and Baseline Boards each respectively. Racks are arranged such that the station racks are in the center, and the baseline racks are arranged in a ring around them; all racks are arranged on a regular grid with 2.5' clearance between racks. These are 24" racks that measure 2' W x 3' W x 7' H. The MCCC and CPCC computers (along with boot server computers) are contained within two standard 19" racks. The CBE occupies a maximum of 3 19" racks and consumes a maximum 15 kW of power. The station and baseline racks contain integrated “rack-as-a-duct” cooling, obtain their cold air supply from the floor, and exhaust warm air out the top. The CBE, MCCC, and CPCC racks obtain their cold air

supply from the room, and exhaust it to the room, likely in a front-to-back airflow configuration. Refer to (Webber, Carlson, A25012N0000) for detailed correlator room requirements and specifications.

All correlator boards are designed and tested to meet FCC Part 15 Subpart J Class ‘B’ conducted EMI levels. This is in an effort to minimize high-frequency ground-loop currents that can disrupt or reduce the reliability of rack-to-rack high-speed data flow. The –48 VDC power plant meets FCC Part 15 Class ‘A’ conducted EMI levels, a requirement for industrial equipment. The COTS computers (CBE, MCCC, and CPCC) meet Class ‘A’ conducted EMI levels, and possibly Class ‘B’ levels (Class ‘B’ is approximately 10 dB more stringent than Class ‘A’). Coupling of COTS computer conducted EMI into high-speed rack-to-rack signal lines is less likely due to the isolation provided by Ethernet networks. A standby mode is planned that, in the event of a mains AC power failure, maintains correlator temperature and power for as long as possible to avoid system power and temperature cycles.

The EVLA correlator system is designed and specified for a 20-year lifetime. Final sign-off of the system (the “Project Complete” milestone of Table 8-2) occurs when it is fully integrated and tested with the EVLA system, operates seamlessly and reliably with the EVLA telescope as a whole, and requires no further NRC personnel support. This includes all of the software needed to operate and maintain the correlator system, to meet the requirements defined in this chapter.

8.3 Correlator Architecture

8.3.1 System Overview

Perhaps the easiest way of understanding the correlator system is to first understand what it can do in very basic terms. The correlator consists of station-based digital filter banks and baseline-based quadrant correlators, operating on a total of 8, 2 GHz bandwidth inputs (from each antenna). A hierarchy of distributed cross-bar switches allows any filter’s outputs (i.e. a sub-band) filtered from any 2 GHz input to use any quadrant’s cross-correlator resources (with minor restrictions noted in following examples). This is illustrated in Figure 8-1.

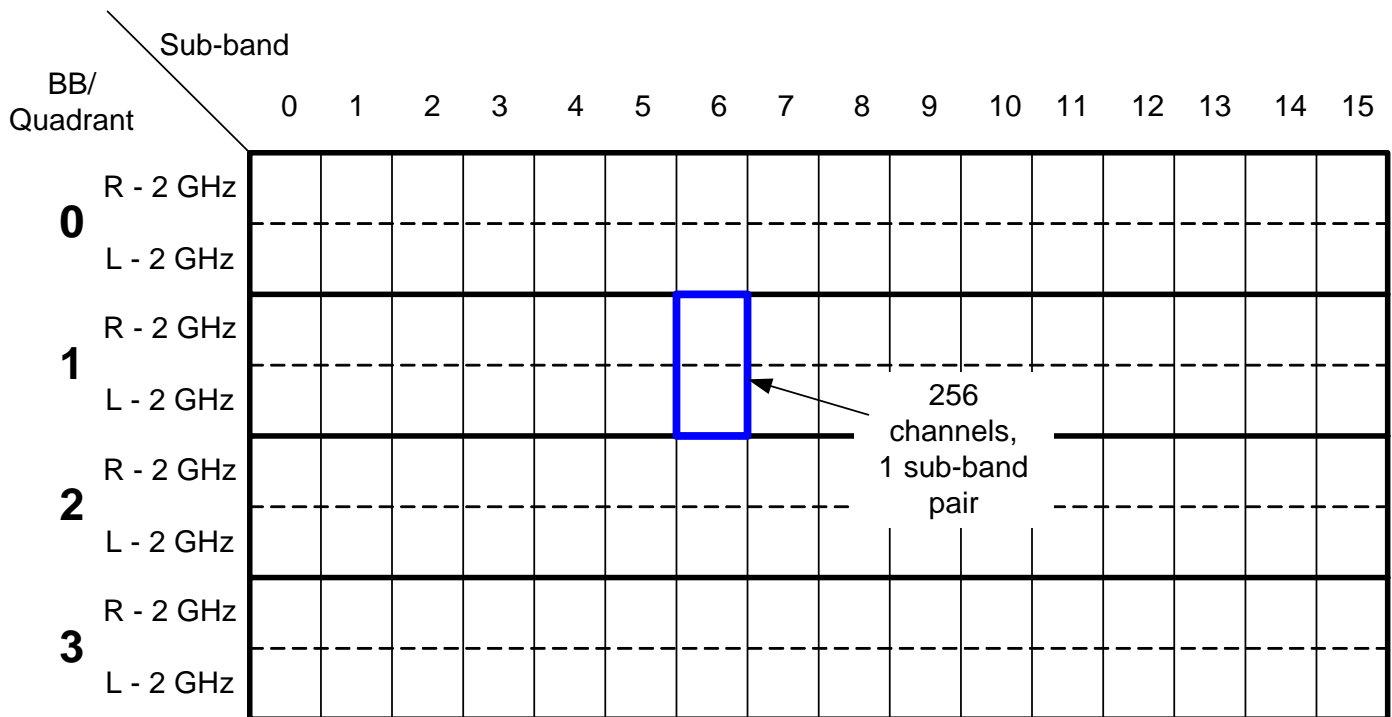


Figure 8-1 Simplified correlator capabilities diagram.

Each small square in Figure 8-1 represents both a filter output (sub-band), and a cross-correlation resource for every baseline. Each cross-correlation resource, shown as the smallest square in the figure, provides 128 spectral channels

across a 128 MHz sub-band for one polarization product, or 2 x 64 spectral channels across a 128 MHz sub-band for 2 polarization products. (Or, viewed another way, each of 2 small-adjacent-dotted-line-separated squares provide 256 channels total, which can be allocated to 1, 2, or 4 polarization products, as shown in the **blue highlighted rectangle**). More spectral channels can be provided using “spectral magnification” (a.k.a. “recirculation”) for concomitant decreases in sub-band bandwidths, however, the physical correlation resources remain the same. Alternatively, or in addition, if more spectral channels are required for a particular sub-band, then sub-band cross-correlation resources within the same quadrant, or within a different quadrant can be allocated to that sub-band.

An example allocation of sub-band filter bank and cross-correlation resources is shown in Figure 8-2. Each full product cross-correlation and the resources it consumes are represented with a different colour, and the sub-band filter bank (R+L polarizations and from which 2 GHz BB) that is being processed is represented with a large black dot. For example, sub-band 0 from BB 0 uses 16 x 128 channel = 2048 channel cross-correlation resources, presumably because higher spectral resolution is required. Sub-band 14 from BB 1 uses 8 x 128 channel resources, etc. The **general rule when allocating resources in this table**, is that first select the sub-band of the BB desired (i.e. place the black dot), then if more spectral resources are required, sweep across (either right, left, or both) if sub-band resources within the same BB can be used, and then sweep or select downwards (or upwards) for every sub-band resource to use additional quadrant resources, if desired. *Generally this results in square or rectangular blocks of resource usage, although not necessarily contiguous as noted in Figure 8-2.* The “horizontal sweeping” uses the Station Board cross-bar switch, and the vertical sweeping uses the quadrant cross-bar switch (**X-bar Board** in Figure 8-3)

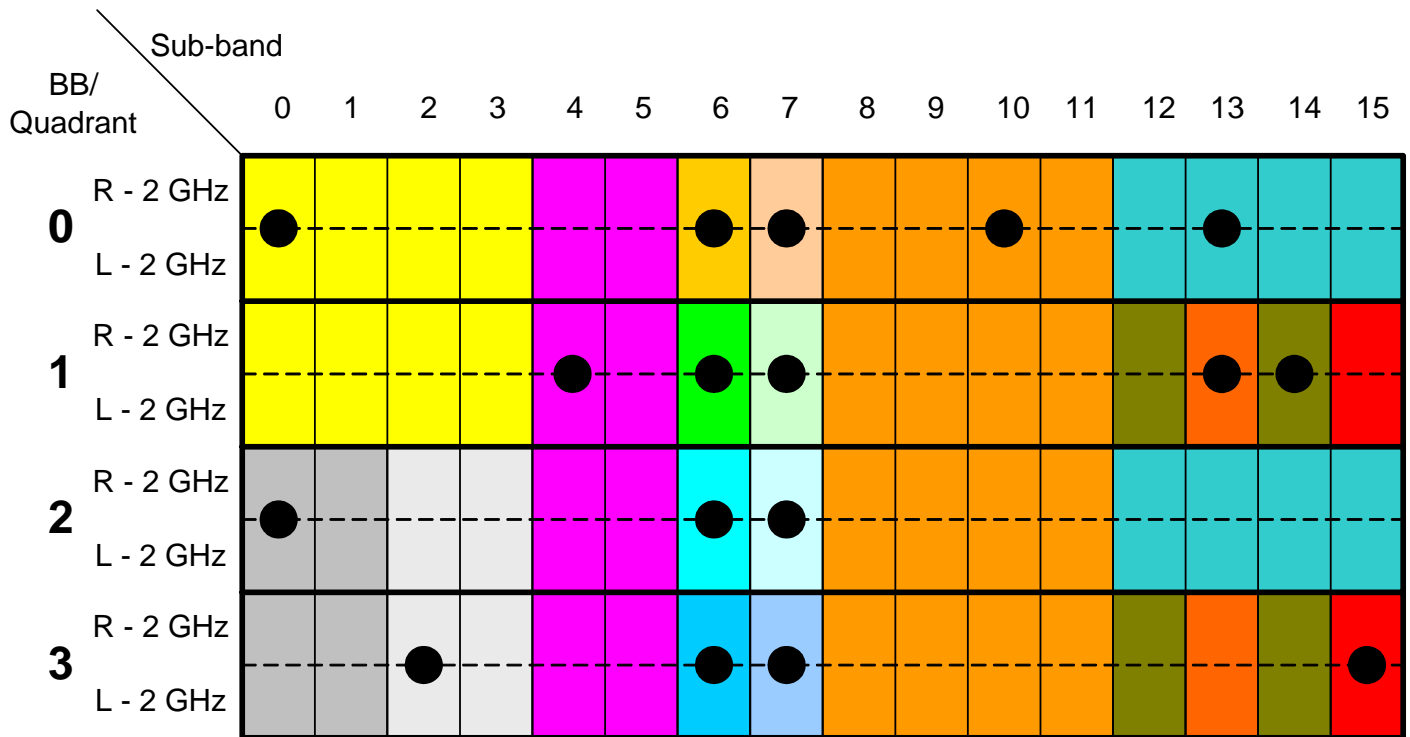


Figure 8-2 Example filter and cross-correlation resource allocation.

The width and placement of a sub-band is *entirely* governed by the Filter Chip tap weights and chosen decimation factor (filter output sample rate). Each sub-band is independent of other sub-bands in terms of bandwidth, number of spectral channels, and integration parameters. As illustrated in the above figures, unused sub-band correlation resources can be allocated to other sub-bands in flexible ways.

Similar but slightly different diagrams apply when 8-bit initial sampling and/or 7-bit correlation is used. With 8-bit initial sampling, there are 1/2 as many BBs, at 1/2 the bandwidth each (although the same number of squares), and with

7-bit correlation, each polarization product requires two small-dotted-line-separated squares (the **blue square** in Figure 8-1) and the maximum sub-band bandwidth is 64 MHz.

8.3.2 System Module Connectivity

Figure 8-3 is a simplified system module connectivity diagram, showing one of each major module. Each Station Board accepts 2, 2 GHz sampled data streams via fiber from an antenna. Each of the 16 outputs of a Station Board contains a sub-band pair from the filter bank, selected via a cross-bar switch, and these are connected to an X-bar Board input. The X-bar Board allows any input to be connected to any output, facilitating the resource allocation flexibility shown in Figures 8-1 and 8-2. The Baseline Board accepts 32 sub-band pairs and performs phasing and cross-correlation functions. Refer to (Carlson, NRC-EVLA Memo# 028, April 13, 2007) for more details.

Two adjacent Baseline Boards are required to perform all 32-station cross-correlations, all polarization products, for a sub-band pair, and the design of the board allows for the required distribution to the next adjacent board, as well as expansion beyond 32 stations, as shown. Correlation “lag frames” are transmitted from the LTAs on the Board on GigEthernet packets. The Baseline Board contains logic to perform phased-array functionality and, once data is phased, it is sent out in Mark 5C Ethernet format on one of up to 3 dedicated 1 GigEthernet links. Phased data can also be routed to the correlator matrix for auto-correlation, or out the input connector for additional specialized processing (not shown).

8.3.3 System Network Topology

The correlator system is designed for scaleable performance: there are virtually no bottlenecks to output data flow and the system’s real-time data handling performance is largely governed by CBE COTS computing performance. The proposed network configuration for the correlator is shown in Figure 8-4. M&C network switches shown in the figure are 24-port COTS switches, while the CBE switch, depending on the number of CBE blades, requires approximately 200 ports. In the figure the **MCCC** is the Main Correlator Control Computer and the **CPCC** is the Correlator Power Control Computer. Each Station and Baseline Board has an embedded CMIB. Baseline Board data are transmitted to CBE computers on Gigabit Ethernet through a monolithic wire-speed Gigabit Ethernet switch—required so that all data required for a particular FFT arrives at one computer. This eliminates the need for an additional wideband network fabric that would be required if a distributed FFT is performed. More straw-man details of network topology and CBE processing are in (Rowen, 2001), (Morgan, 2003), and section 8.4. Not shown are the network for the FORM boards, the Ethernet switch network for routing phased data packets to VLBI recorders, nor CPCC to module connections for remote power monitor and control.

8.3.4 System Installation

The correlator is a large system. For cost and performance reasons, it is desirable to minimize the correlator installation footprint. A smaller footprint requires shorter and less expensive cables and results in better signal performance—particularly at the clock and data rates under consideration. The floor plan for a 32-station correlator is shown in Figure 8-5. ‘S’ racks are Station Racks, and ‘B’ racks are Baseline Racks.

In this plan, maximum 7 m-length cables are required for data distribution from the Station Racks to the Baseline Racks. Signal arrival time mismatch at the Baseline Boards is completely compensated for by buffers in the input cross-bar switch of the board, and the Correlator Chip. Racks need only front and rear access and can be installed side-by-side. Each rack holds 16 boards plus a 6U sub-rack; in the Station Rack, the 6U sub-rack holds 8 X-bar Boards. Provided there is floor space, more racks can easily be added at a later date without requiring replacement of existing cabling. Each rack is 7.5 feet high. All high-speed cabling is within the racks and under the (raised) floor. Any other cabling (e.g. network cables shown in Figure 8-4) is run in overhead cable trays. The MCCC, CPCC, and CBE computers are located in the correlator room as shown in Figure 8-5

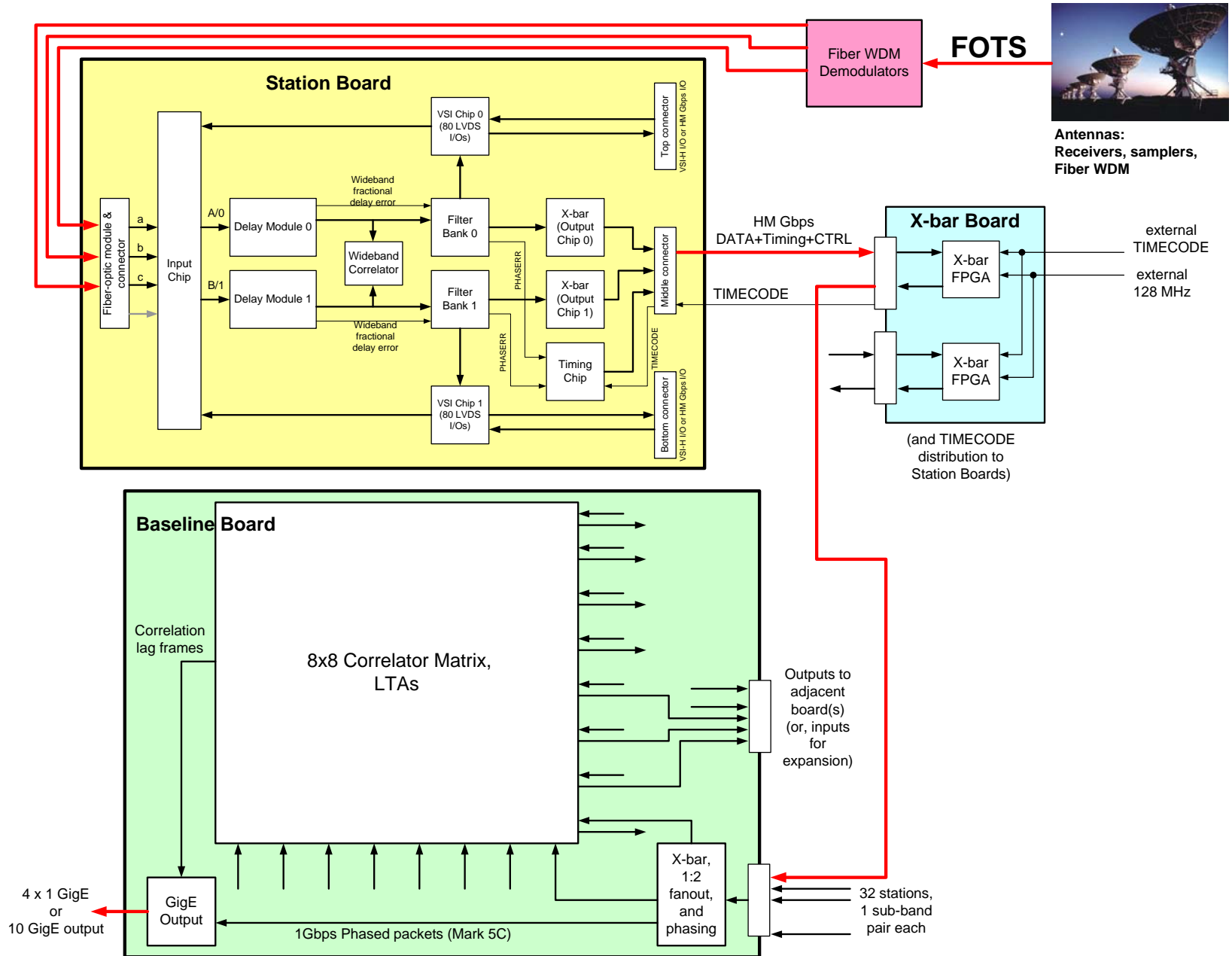


Figure 8-3 Correlator module connectivity diagram.

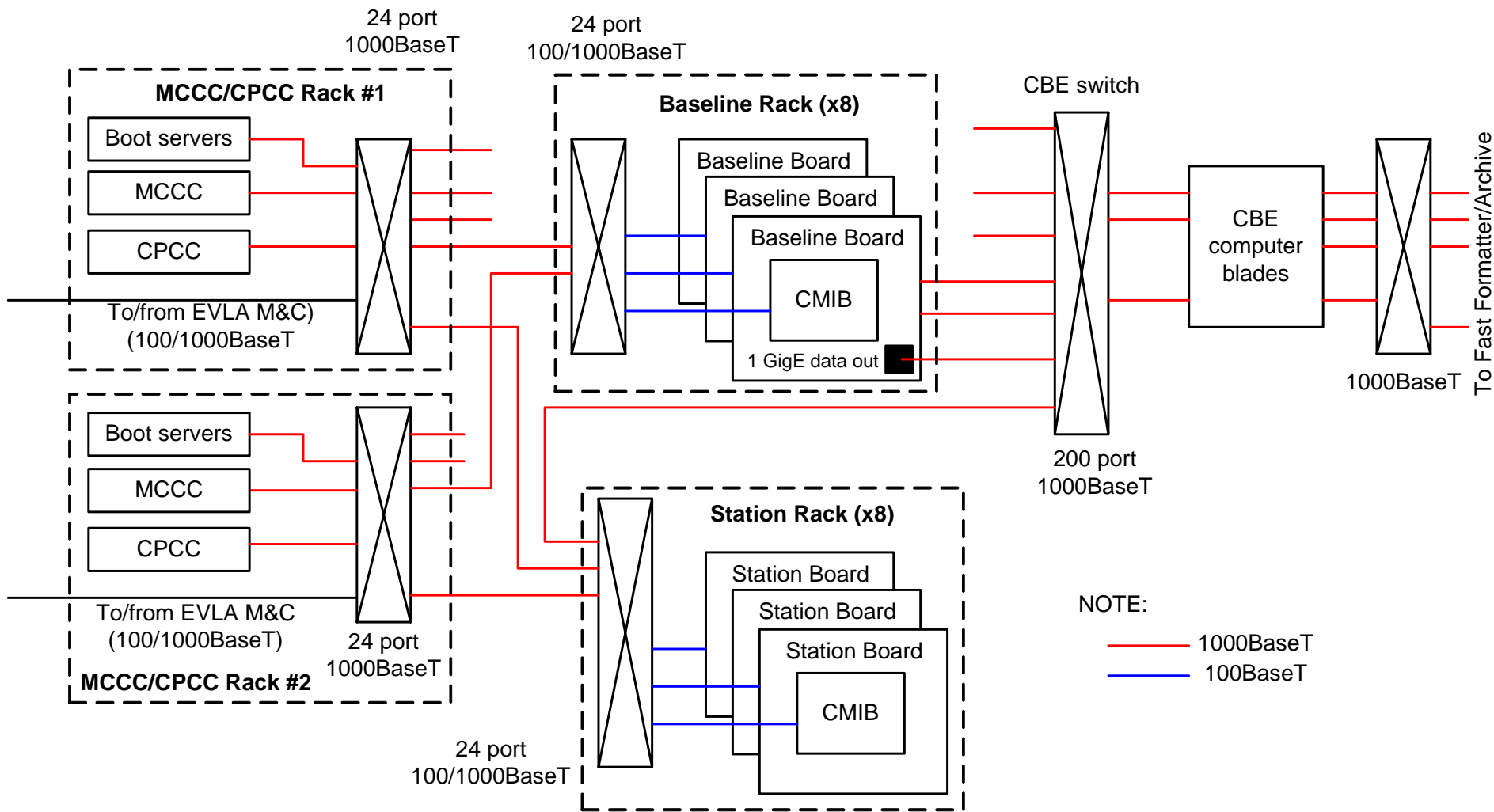


Figure 8-4 Simplified correlator network topology.

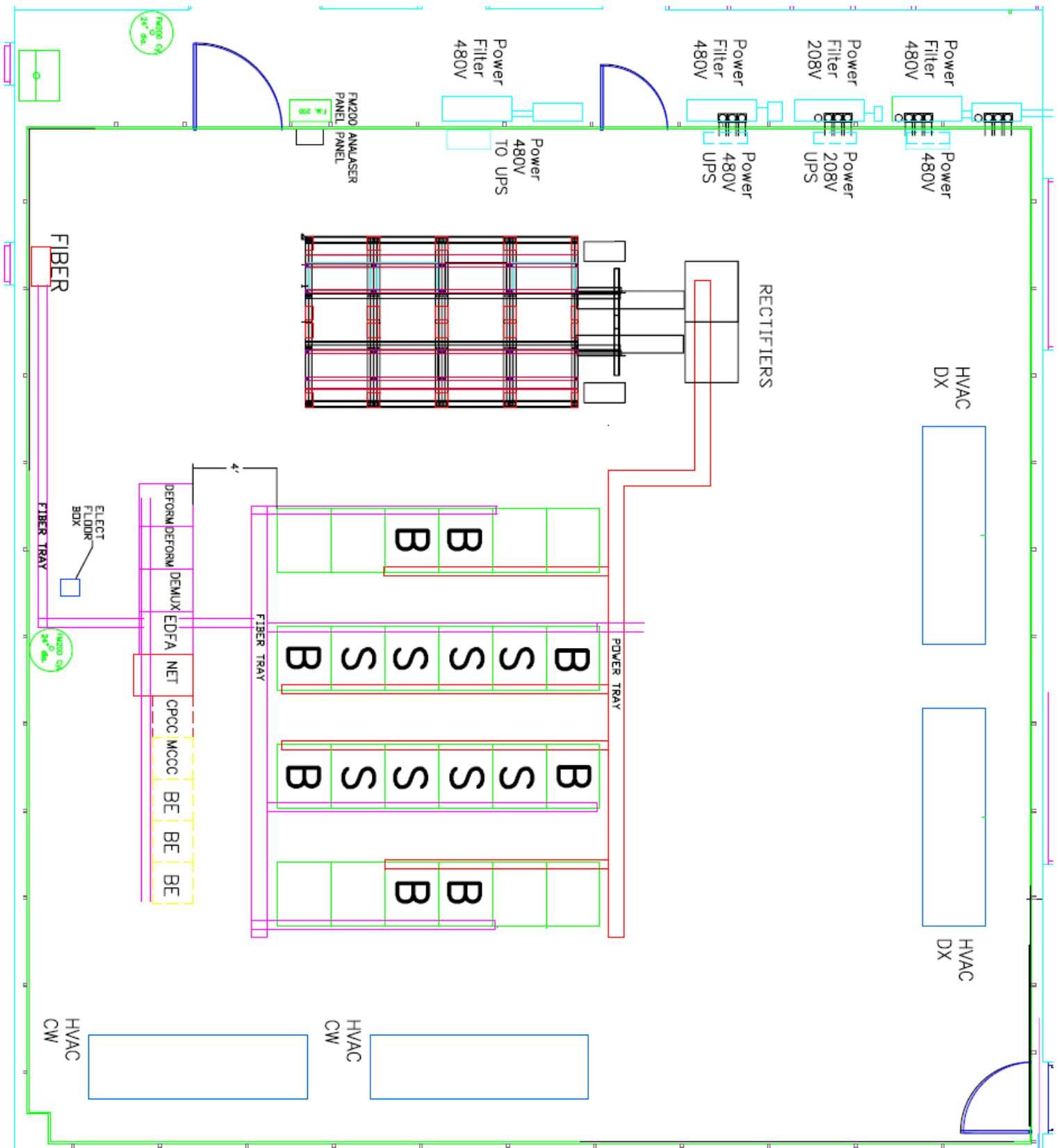


Figure 8-5 EVLA 32-station correlator system floor-plan.

8.4 Correlator Backend (CBE) Requirements

A complete requirements specification is given in: “System Requirements Specification: EVLA Correlator Backend“, project document A25251N0000, revision 2.0, May 10, 2002. It can be found on the Computing Working Documents web page at <http://www.aoc.nrao.edu/evla/techdocs/computer/workdocs/index.shtml>. The following is an overview of the key requirements.

8.4.1 Assumptions

Packetization of lag frames (a lag frame consists of one lag section of up to 128 complex lags and identifier information from one Correlator Chip (section 8.2.11)) including setting of the correct CBE node destination IP address (section 8.3.3) for the given baseline is handled by the correlator. The CBE provides a mapping of baselines to node IP addresses for the current correlator mode. The lag frame packets don't necessarily arrive in a set order and their delivery is a one-time event. Resends of missed or bad packets won't be possible due to Baseline Board hardware limitations and performance requirements.

Indirect (i.e., non-correlator lag frame) data arrive in a timely fashion. That is, with no significant delay prior to its being needed at a particular point in backend data manipulation, processing, or formatting.

The archive subsystem is designed to handle output rates and volumes delivered by the backend during times of peak production. The CBE provides results ready to be ingested by the archive.

8.4.2 CBE Input

The backend shall be able to receive the following correlator outputs: lag frames, quantizer power measurements, Filter Chip parameters (power measurements etc), frequency shift parameters, windowing parameters, and quantizer and re-quantizer state counts. It shall also be able to receive observational mode, meta-data (“sky frequencies”, polarizations etc.), status requests, and other EVLA data from the Monitor and Control System.

8.4.3 CBE Output

The CBE shall be able to deliver formatted observational output to the archive, and status, warning, error, and system component failure and recovery reports to M&C.

8.4.4 Correlator Interface

All lag frame data shall be sent directly across the correlator to backend interface using Gbit Ethernet and UDP/IP. All backend cluster nodes shall have a direct path to each correlator output point (the baseline boards) through a single Ethernet switch. The interface shall have sufficient bandwidth to meet the initial maximum aggregate data transfer rate of 1.6 Gbytes/sec. All lag frames from the same baseline (that could be distributed across baseline boards) shall be routed to the same backend cluster node.

8.4.5 M&C Interface

All non-lag frame correlator data along with other EVLA data, M&C requests, backend responses and backend-generated reports shall pass to and from the M&C System via the Virtual Correlator Interface (VCI). If sufficient bandwidth is not available to handle all traffic, critical auxiliary correlator data may have to be routed directly from the main correlator control computer.

8.4.6 Archive Interface

All final, formatted astronomical results shall be sent directly to the archive across this interface. It shall have sufficient bandwidth to meet the initial maximum aggregate data transfer rate of 25 Mbytes/sec. All backend cluster nodes shall have a path to the archive system.

8.4.7 User Interface

The backend shall be capable of presenting a command line interface on any and all nodes for use by software development and test personnel. It shall also have selectable internal diagnostic modes that produce printed values for key variables at critical locations in the code. A system-level GUI shall be available for high-level monitoring of the backend status, including, for example, node status, processing load, and network status.

8.4.8 Data Processing

Backend cluster nodes shall be able to perform the following data manipulation and processing tasks: lag set assembly (a “lag set” is required for an FFT), data valid normalization, coarse quantization correction, time stamp adjustment, residual phase rotation correction, FFT, interference removal/reduction, windowing, integration, and output formatting. These are capabilities that shall be available, although only a subset is normally used on any given data stream. The backend *may* also be required to perform sub-band stitching operations.

8.4.9 Internal Monitor and Control

The backend shall be self-monitoring. Input, output and data processing rates shall be measured and error and warning statistics shall be maintained in order to continually monitor system health and anticipate problems. All internally generated reports and status information shall be passed to the M&C for presentation to the outside world.

8.4.10 Reliability

The backend shall be capable of attempting recovery from a number of failure modes. The failure of a single node, including the node running the Backend monitor and control functions, shall not affect any other node. The loss of an external network connection shall not affect internal operations until all on-board storage resources are filled, in the case of loss of the archive connection; or until necessary auxiliary data is needed, in the case of loss of the M&C connection. The system shall be able to kill and restart corrupted processes, and reboot failed processors and network connections. It shall report all problems, recovery attempts and outcomes to the M&C. The goal is to avoid total system reboots for a period of time greater than or equal to the normal EVLA maintenance interval.

8.4.11 Scalability

The CBE system shall be scalable to higher rates of input, output and data processing, with an ultimate objective of meeting the full data generating capability of the correlator (16 Gbytes/sec). Hardware shall be extensible in a manner that is transparent to software and vice versa. Upgrades shall meet seamlessly with unchanged components.

8.5 Correlator Backend Design

The CBE is a distributed cluster based system with the nodes logically linked via message passing middleware. High-speed external switched networks are used to connect to the correlator and archive systems. The node hardware is multi-CPU Intel or Intel-clone processors configured with large amounts of memory and disk storage. The current operating system of choice is Linux, and the message passing middleware is either PVM or MPI. All are open source, and based on widely accepted industry standards.

There are two main software subsystems running on the nodes. One node (known as the *head node*), with one or more shadow nodes, runs a subsystem consisting of the Backend internal monitor and control functions. The remaining nodes (known as *compute nodes*) run the processing pipeline subsystem that consists of input, sorting, data processing and output functions. Several of the compute nodes will be available as standby nodes to provide failover capability in the event of compute node failures.

8.5.1 Backend Control Function

Backend Control is the gateway to the CBE. All inbound and outbound non-correlator frame data pass through it. Backend Control also maintains a statistical model of the CBE system state. It incorporates measurement, error and warning data from the processing nodes along with periodic status checks performed by the Monitor Function. There are three classes of messages: the most basic are messages that are simply routed to another destination; the second

class is messages that are also routed, but in the process data for the statistical model are extracted; and the third type has the Control Function as a destination and is also used to update the statistical model. Backend Control generates messages, based on the state of the statistical model, to request check and repair, and offload services from the Monitor Function.

8.5.2 Backend Monitor Function

The Monitor Function performs system component monitor and recovery operations based on directions received from the Backend Control Function. It performs status checks of networks, processors and processes, it attempts network and processor restarts, and is able to kill and restart damaged processes. It also performs off-loads of data processing from malfunctioning to standby nodes.

8.5.3 Compute node functionality

Compute node functionality is separated into two stages: the input stage and the lag processing stage. The function of the input stage is to assemble complete “lag frame sets” from the raw lag frames arriving from the baseline boards, as well as order the frames according to their timestamps. The function of the lag processing stage is to operate on the stream of lag frame sets delivered by the input stage, for example, normalizing the lags and applying a Fourier transformation.

8.5.3.1 Input stage

The input stage is a pipeline process that accepts lag frames from the correlator/Backend network interface as input, and creates complete, time-ordered lag frame sets as output. It incorporates a timeout functionality to maintain the stream of lag frame sets in the absence (or late arrival) of expected frames. Input is discarded under a limited set of conditions: the arrival of faulty input frames (*i.e.*, invalid frame checksum value or other invalid frame header values), the arrival of late input frames, the presence of unexpected input frames (*i.e.* frames not expected under the current configuration), and an out-of-memory condition in the input stage process.

8.5.3.2 Lag processing stage

The lag processing stage comprises a set of pipeline processes that accept a time-ordered stream of lag frame sets as input, and apply a series of transformations and side-effect producing operations on the sets in the stream. This stage applies transformations such as data valid normalization and FFT, and creates output records for the archive subsystem. Each pipeline in this stage consists of a sequence of selected processing elements that may be individually configured. Processing exceptions occurring in the processing elements, such as floating point arithmetic exceptions, are trapped and flagged. The pipeline incorporates buffers that allow for the arrival of required auxiliary data before forwarding a data element to the following processing element in the pipeline.

Creation of archive records also occurs in the lag processing stage. The creation of output records is a distributed process that occurs over many backend compute nodes concurrently. The distributed nature of archive record creation avoids the need to send data from the compute nodes to an intermediary process that would assemble the data from multiple compute nodes into a single archive record. (The backend control function writes record-wide metadata in the archive records.) Archive records are created on a distributed filesystem hosted either by the compute nodes themselves using local disk for storage, or by other computers providing the disk storage. Complete archive records are sent to the archive, and are made available on the same filesystem on which they were created to other EVLA computing subsystems (TelCal, DCAF) for a limited time.

8.6 Deliverables

Table 8-3 summarizes the modules that are under development and will be delivered by the NRC. This table includes items for a 32-station correlator. 5% module spares are not included in this table but will be provided.

Table 8-3 NRC-supplied correlator deliverables

Qty	Item/Description
128	Station Board (c/w mezzanine cards)
64	X-bar Board
768	Common Backplane
128	Baseline Board (c/w mezzanine cards)
1	High-speed cables for 32 stations
1	Sub-racks and racks for 32-station correlator
1	COTS computers (MCCC, CPCC, CBE (50), Ethernet switches))
1	48VDC, 4000A plant including batteries, shipping, but not power cables or installation. Can be field-upgraded to 6000A. AC-AC UPS for Backend COTS PCs (est. 30 kVA)
n/a	Correlator software: monitor/control interface, and mapper software.

Table 8-4 summarizes additional modules and components that NRC does not develop or supply as part of the correlator installation. Quantities are for a 32-station correlator.

Table 8-4 Additional (NRAO-supplied) correlator deliverables.

Qty	Item/Description
128	Dual-input (2 x 2 GHz bandwidth) Fiber-Optic Receiver Module (FORM) card to plug into the Station Board. C/w test vector receivers and test vector transmitters (to Station Board receivers). Each baseband output is 4 Gs/s arranged as 16 de-multiplexed streams, 4 (or 3) bits/stream, @ 250/256 Mbps each.
1	1 GigE switch for selection of phased data for VLBI, and conversion to 10 GigE for the Mark 5C recorder.
n/a	Correlator embedded real-time software, system/board/chip-level GUIs, CBE software

8.7 Interfaces and Impacts on Other Systems

Table 8-5 summarizes correlator interfaces and/or associations to the external world, and a description of possible impacts on other parts of the system.

Table 8-5 Table of correlator interfaces and potential impacts on other systems.

Interface/Location	Description	Impacts
Fiber-Optic Receiver Module (FORM). Sec. 8.2.4, 8.2.10	3 fibers into each module. Two, 48-bit data highways out of each module. Each data highway contains a 2 GHz BB, 3-bit samples at 4 Gs/s. Provision for 4-bit samples. Provision for one 8-bit sampled stream at 2 Gs/s.	BERT transmitter in the antenna's fiber-optic transmitter and in the (correlator) receiver module allows transmission system and interface testing. Real-time, non-invasive CRC checks allow on-line connectivity testing between the FORM and the Station Board. Supports 1, 2, 3, 4, or 8-bit sampling with flexible baseband widths. 8-bit sampling reduces the sampled bandwidth by a factor of 2.
Delay Module (Station Board). Sec. 8.2.16	This module inserts wavefront delay in the station data path. The depth of this delay determines the maximum baseline.	Design is for a total of 0.25 seconds of delay with +/-122 ps resolution. Delay can be increased with new module. Designed to tradeoff bandwidth for number of BBs.
Correlator clock/timing interface (TIMECODE) Sec. 8.2.24	Reference clock (128 MHz), and reference time tick (1 PPS). Required for correlator TIMECODE generation.	Requires clock and time epoch (1 PPS) from array maser/timing master.
LO system (antenna). Sec. 8.2.6, 8.2.7, 8.2.12, 8.2.13, 8.2.17	LO offsets for anti-aliasing, sub-sample delay tracking, and narrowband harmonic/inter-modulation product reduction.	Requires 100 Hz LO tuning resolution for LO offset capability. An antenna can have the same LO offset in every one of its basebands. Optionally, different LO offsets in the same antenna allow sub-band "cross auto-correlation". System control should ensure that minimum acceptable net phase rotation rate is ensured on all baselines. Time-variant LO offsets could be employed for more aliasing attenuation on long <i>coherent</i> integration times. LO offsets could be turned off, and correlator would lose sub-sample delay tracking and anti-aliasing capability.
Noise diode switching (antenna). Sec. 8.2.6	Noise diode switching in the antenna receivers for system noise calibrations. A reference Filter Chip synchronously switches with the noise diode to acquire power data with the diode on and with the diode off.	Switching/binning in the correlator is synchronized to switching in the antenna using a timer and a priori knowledge of the switching period and phase. It is not yet defined what this switching rate is.
VLBI recorder interface. Sec. 8.2.19	1 GigE switch with one or more 10 GigE ports for Mark 5C data recording.	NRAO-purchased COTS item.
VSI-H I/O (Sec. 8.2.21)	Single antenna VLBI I/O to/from Station Boards	With the advent of 10 GigE for VLBI recorders, this is likely obsolete. A 10 GigE to HM LVDS module could be developed to use this interface.
Internal correlator monitor and control bus (Station, Baseline Boards). Sec. 8.3.2, 8.3.3	Interface to Station, Baseline, and Phasing boards. 100 Mbps Ethernet and embedded PC/104+ "CMIB".	Station Board data products output from this interface (auto-corr, sub-band power, quantizer statistics, phase-cal) internally, but through the CBE computers externally.
Internal data output interface (Baseline Board). Sec. 8.2.25, 8.3.2	Baseline Board data output pipeline on 1000BaseT Ethernet.	Wideband output with delivered output data rate of ~100 Mbytes/sec from each Baseline Board. Potential for upgrade to ~800 Mbytes/sec (10 Gigabit Ethernet) output capacity from each Baseline Board.
Correlator system monitor and control interface. Sec. 8.3.3	Network interface to higher-level control computers. 100/1000BaseT Ethernet. Refer to (Vrcic, A25201N0000)	Virtual correlator interface to allow high-level configuration, control, and monitoring.
Correlator system data output interface. Sec. 8.3.3	1000BaseT through switch to CBE COTS computers.	CBE computers perform FFTs, excise interference, and perform longer-term integration.

8.8 Risk Assessment

Table 8-6 Areas of risk, and planned risk mitigation strategies in descending order of importance.

Risk	Risk Mitigation
Personnel	All engineers have been hired and appear to be working well on development. Hardware engineering will soon be ramping down as we move into more prototype and production fabrication.
Speed (256 MHz clock rates)	Well in hand and proven technology with initial prototype successes.
Correlator Chip	Prototypes fully tested and qualified; signed off for production end of June 2007.
Filter Chip	The design successfully fits in a Xilinx Virtex-4 SX35 device, which is cheap enough and cool enough for production. The FPGA eliminates design risk associated with an ASIC, leaving only the reliability of the SX35 as a risk item.
Personnel turn-over	Define and enforce documentation standards to minimize single person dependencies. No personnel turn over yet in the project.
Disruptive ground loop noise	Use differential signaling from rack-to-rack. Ensure all modules meet FCC Part 15 Subpart J Class B conducted EMI levels. Use large, low-impedance shunts between racks. Use signal-cable common-mode chokes if necessary.
Major supplier insolvent	This has not happened. Purchase of production quantity FPGAs, memories, and cables is in progress.

8.9 References

- Carlson, B.R., Dewdney, P.E., Efficient wideband digital correlation, Electronics Letters, IEE, Vol. 36 No. 11, p987, 25 May, 2000.
- Carlson, Brent, A Proposed WIDAR Correlator for the Expansion Very Large Array Project: Discussion of Capabilities, Implementation, and Signal Processing, NRC-EVLA **Memo# 001**, May 18, 2000.
- Carlson, Brent, WIDAR Correlator Sensitivity Losses, NRC-EVLA **Memo# 011**, January 30, 2001.
- Carlson, Brent, A Closer Look at 2-Stage Digital Filtering in the Proposed WIDAR Correlator for the EVLA, NRC-EVLA **Memo# 003**, June 29, 2000.
- Carlson, Brent, Simulation Tests to Quantify the Spectral Dynamic Range and Narrowband Interference Robustness of the WIDAR Correlator for the EVLA, NRC-EVLA **Memo# 009**, Nov. 1, 2000.
- Carlson, Brent, Refined WIDAR EVLA Correlator Architecture, NRC-EVLA **Memo #014**, October 2, 2001.
- Carlson, Brent, EVLA 'WIDAR' Correlator Description for the Preliminary Design Review, NRC-EVLA **Memo# 024**, June 17, 2005.
- Carlson, B., USER MANUAL: Programmer's Guide to EVLA Correlator System Timing, Synchronization, Data Products, and Operation, **A25290N0000**, Revision DRAFT2, March 17, 2006.
- Carlson, Brent, Simulation Tests of Phasing Subsystem Signal Processing in the WIDAR Correlator for the EVLA, NRC-EVLA **Memo# 008**, Nov. 7, 2000.
- Carlson, B.R., Dewdney, P.E., Burgess, T.A., Casorso, R.V., Petrachenko, W.T., Cannon, W.H., The S2 VLBI Correlator: A Correlator for Space VLBI and Geodetic Signal Processing, Publications of the Astronomical Society of the Pacific, 1999, 111, 1025-1047.
- Carlson, Brent, An Analysis of the Effects of Phase Dithering in a Lag-based Fringe-Stopping XF Correlator, NRC-EVLA **Memo# 002**, May 26, 2000.
- Carlson, Brent, Requirements for 8-bit Processing in the Proposed WIDAR Correlator for the EVLA, NRC-EVLA **Memo# 010**, January 29, 2001.
- Carlson, Brent, Simulation Tests of Sub-Sample Delay Tracking in the Proposed WIDAR Correlator for the

Expanded Very Large Array, NRC-EVLA **Memo# 007**, October 3, 2000.

Carlson, Brent, Summary of Discussions Held During the July 10-14, 2000 Workweek in Socorro Regarding the EVLA-WIDAR Correlator, NRC-EVLA **Memo# 005**, August 22, 2000.

Carlson, Brent, An Optimized Connectivity Scheme for the EVLA Correlator, NRC-EVLA **Memo# 028**, April 13, 2007.

Carlson, Brent, REQUIREMENTS AND FUNCTIONAL SPECIFICATION: EVLA Correlator Phasing Board, **A25110N0000**, Revision DRAFT, November 8, 2005.

Carlson, Brent, TEST AND VERIFICATION REPORT: EVLA Correlator Reliability Report #1, **A25010N0003**, Revision: Report #1, June 6, 2005.

Webber, Ralph, Carlson, Brent, REQUIREMENTS AND FUNCTIONAL SPECIFICATION: EVLA Correlator Room, **A25012N0000**, Revision DRAFT3, October 25, 2005.

Crochiere, R.E., and Rabiner, T.R., Multirate Digital Signal Processing, Prentice-Hall, New Jersey, 1983.

Morgan, Thomas R., Requirements and Functional Specification, EVLA Correlator Backend, **A25252N0000**, Rev. 1.0, September 23, 2003.

Rowen, B., WIDAR Correlator Backend Processing Options, NRAO, Socorro, 2001, November 19.

Vrcic, S., Protocol Specification, EVLA Correlator Monitor and Control—Virtual Correlator Interface (VCI), **A25201N0000**, October 28, 2004.