ATA Memo 24

Correlators: General Design Considerations

L. D'Addario

2001 March 13

1. Cost Equations

If cost is measured by the total number of multiplications and additions per second, then it is well known that the costs of XF and FX correlators are, respectively,

$$C_{\mathrm{XF}} = aBMN^2 \tag{1}$$

and

$$C_{\mathrm{FX}} = bB \log M + cBN^2, \tag{2}$$

for some coefficients a, b, c that depend on the technology and details of the implementation but are otherwise constant. Here B is the total bandwidth processed, N is the number of antennas, and M is the number of spectral channels into which B is analyzed. These equations can be misleading because they obscure some important issues, as we shall see shortly. Nevertheless, to the extent that they are valid they provide a few insights.

First, all cost terms are proportional to total bandwidth. Second, at sufficiently large N both architectures have a cost proportional to N^2; that is, the "X" term dominates. In present practice, it seems that $N > 30$ is enough for this to be true. The cost ratio of the two architectures is then aM/c, so the XF architecture is more expensive at sufficiently large M. Bunton [1] estimates that the break-even point is at about $M = 5$. Since scientifically interesting values of M are almost always much larger than this, it appears that the FX architecture is almost always cheaper.

The problem with (1) and (2) is that their basic premise is flawed. The cost cannot be measured by counting the number of multiplies and adds. For one thing, not all multiplies are equal. An FX correlator requires two real multiplies per baseline per spectral channel, but an XF correlator requires a complex multiply (which consists of 4 real multiplies and 2 adds). For broadband astronomical signals (noise) the quantization can be very coarse, so that only 2 bit (or maybe 3-level) multiplies are needed for the XF correlator; but the dynamic range is much larger after the F part of an FX correlator, requiring finer quantization and more expensive multiplies and adds. (This is true for astronomical signals alone, considering astrophysical masers and other strong lines; when interference is considered,even more dynamic range is needed.) However, these facts can be handled in (1) and (2) by making $c > a$.

A more fundamental flaw is that the expressions neglect the cost of accumulation. For each multiplier, the results must be summed over a significant time before being written to long term (slow) storage. As M increases, an FX correlator's multiplication rate is constant because the M frequencies are each processed $1/M$ of the time. But the number of accumulation registers into which the results must be added is proportional to MN^2, just as it is for the XF correlator. Both (1) and (2) should contain an additional term for the cost of this storage, and at sufficiently large M that cost will dominate the FX case.

A second fundamental flaw of (1) and (2), and one that we will dwell upon at length later, is that they fail to include the cost of interconnections. For any architecture, if we have an array of N spatially distributed antennas, then the signals will arrive in N separate bundles. Some processing (in particular the F part of an FX correlator) can be done on each antenna-bundle separately, without interaction with the others, but eventually all must be pairwise interconnected. If (as is usually required) the correlations of all baselines are to be computed, then every antenna must be connected to every other. It is thus unavoidable that there must be at least N cables of some sort to bring the antenna signals to the cross-correlation unit(s). The cost of each connection depends on bandwidth B. In practice, many times this number of interconnections is usually necessary; as we'll show later, once N and B are sufficiently large, the number of interconnections grows as BN^3. Thus, the large-N case is said to be "copper dominated" (or "glass dominated" if optical fibers are used) rather than "silicon dominated."

2. Cost Minimization Strategies

We confine ourselves primarily to the case of many antennas ($N > 30$), or an array at least as large as the VLA. If we also consider large bandwidth ($B > 1\,\mathrm{GHz}$), then in present technology the necessary signal processing requires that the electronic hardware occupy several (perhaps many) equipment racks and consume large amounts of power (many tens, and perhaps hundreds, of kW). Let us call the correlator "large" if its implementation requires several racks and/or high power. A narrower-bandwidth correlator (say, 100 MHz) can also be large if the number of antennas is sufficient (say, 100).

A large correlator must be partitioned among racks, and further among circuit boards ("cards") and still further among integrated circuits ("chips"). This is straightforward for the antenna-based parts of the circuitry (including the "F" part of an FX correlator) because each signal can be kept separate and processed independently until it reaches the cross-correlation circuitry. Therefore, we concentrate on partitioning the cross correlators. This must be done in such a way that the separate units do not interact with each other; that is, each should receive input signals only from the antennas and not from another correlation unit, otherwise the number of interconnections becomes unmanageable very quickly. There are at least three different ways to do this: by frequency-slicing, by time-slicing, and by baseline. That is, respectively:

a. For each antenna, bandwidth B can be analyzed into K subbands, with each assigned to separate hardware.

b. For each antenna, the signal can be segmented into blocks of time-contiguous samples, with successive blocks in each group of K blocks assigned to separate hardware. All K hardware units operate in parallel, and each can be K times slower than a single element that would process all the data.

c. For the $N(N+1)/2$ baselines, assign $1/K$ of them to each of K separate hardware units.

In all cases, we take the number of separate cross correlatoin hardware units to be K, which is chosen to make the unit have a practical size in the available technology. The size may be limited by internal connections (topology), by power dissipation, or by external connections. It is also possible to do the partitioning in cascaded stages, using a different one of the three methods at each stage.

For example, the ALMA correlator now under construction [2] partitions its $B = 16\,\mathrm{GHz}$ bandwidth into 4 parts at each antenna (method a), allowing the correlator to be constructed as four independent "quadrants." Each quadrant requires 8 racks of electronics. Within a quadrant, each 1 msec of signal duration is further partitioned into 32 time slices (method b); these are then processed by 32 independent correlation "planes," each of which requires 4 cards. Each plane handles all baselines of the $N = 64$ antenna array.

The choice of partitioning method can be crucial to the minimization of cost.

First, notice that method c (partitioning by baseline) is generally most expensive for interconnections. This is because all antenna signals must go to all K correlator units, so there are NK interconnections each carrying the full bandwidth B. In the other methods, there are also NK interconnections, but each carries only $1/K$ of the signals. There could still be cases (especially for small N and large B or M) where method c produces the lowest overall cost, but not for the large-N situations on which we are focusing.

Next, consider that partitioning by frequency-slicing (method a) is natural to the FX architecture and will almost always be chosen in that case. However, it is neither necessary nor natural for the partitioning factor K to be the same as the final number of frequency channels M. If the application requires large bandwidth B (forcing relatively large K) but low resolution M, it could happen that $M < K$. Then further partitioning by time-slicing (method b) might be reasonable. Conversely, if high resolution is required at relatively small bandwidth, then it would be possible to force $K = M$ by making more partitions than necessary; but doing so requires more interconnections, so it is better to keep K small and let each correlation unit handle several of the M spectral channels.

Finally, consider that partitioning by time-slicing costs no arithmetic operations at all. It may require re-ordering the data and/or changing the number of parallel paths on which it appears (and thus changing the clock rate), but this can be done by memory elements alone; each sample remains unchanged. This is usually quite inexpensive. It is therefore the method of choice for the XF architecture as well as for situations where no further frequency analysis is needed (as in the $M < K$ situation mentioned above).

These general ideas must be applied to the case at hand, which depends on parameters N, B, and M and on the available technologies. The first decision is on the main type of architecture, which might be XF or FX or some hybrid (e.g., FXF [3]). (I ignore here the very different approach where the *spatial* Fourier transform is done before detection [4]. This appears to be useful only for filled arrays, which are of little interest for astronomical imaging.) The choice is not obvious, since it depends on many issues as discussed in section 1. The total cost also depends on things that do not scale easily, such as non-recurring engineering: designs which are very regular and thus minimize the number of different assemblies can be much cheaper because NRE is better amortized. Secondary considerations like operating flexibility (e.g., can M be traded against B, or can some baselines be sacrificed for more bandwidth or resolution?); maintenance cost; and future expandibility may be critical to some applications. For these reasons, it is usually worthwhile to do straw-man preliminary designs in two or more architectures so as to find the total costs; these should include the interconnections, packaging, and NRE.

3. Interconnections

For each basic architecture being considered for a large correlator, the next decision should be the partitioning. As argued above, partitioning by baseline is not efficient and so it will not be further considered. Then the total bandwidth to be transmitted from the antennas to the cross-correlation units is always NB, and the number of separate signals involved must be at least NB/b, where b is the capacity of one signal in the selected transmission technology. This could require a large number of wires (or fibers), but it is irreducible regardless of architecture. The actual number of signals could be much larger, because there are NK topologically separate paths required. The minimum possible number is then $\max(NB/b, NK) = N \max(B/b, K)$.

From this it follows that it is best to partition the cross correlator into as few units as possible, so that each unit is as large as possible. This means that each unit should handle the largest possible fraction of the bandwidth or of a time segment. Obviously a single unit ($K = 1$) is optimum, but we are considering a "large" correlator, so we now assume that a single unit is not practical.

As N or B increases, more units will be required in a given technology. In the large-N limit, we find from (1) and (2) that K must be proportional to BN^2. Then the total number of connections NK is proportional to BN^3, so interconnections rapidly become the dominant design problem as N gets large.

To see what would be a practical "unit" correlator, we should look at its internal structure. Since it must compute all baselines, a natural topology is as a two-dimensional matrix of size $N \times N$, where each cell handles one baseline [2][5]. Each antenna's signal is distributed to one entire row and also to one entire column. Self-correlations are computed along the diagonal, and these are usually needed. The cells on one side of the diagonal are in principle redundant and need not be constructed; but it is possible to maintain the square symmetry by designing each cell to compute half of the desired results (say, positive lags on one side and negative ones on the other, for an XF architecture; or real part of the correlation on one side and imaginary on the other, for an FX architecture) without requiring any additional signal paths.

Consider implementing a unit correlator as a single chip. If N is large, using a separate input pin for each antenna may exceed the practical pin limit of a cost-effective package. This can be fixed by making K large enough that B/K is small, so that signals from several antennas can be time-multiplexed onto one pin. Still, the processing capacity within the chip may not be enough unless the data rate is sufficiently slow; this also requires making K large enough. So either the processing capacity or the I/O capacity will set a lower limit for K in the chip-per-unit implementation.

But it is not necessary to put each unit on one chip. Consider making a unit correlator as one circuit board. The board can contain an $L \times L$ array of identical chips, maintaining the 2-D matrix topology and multiplying the processing capacity by L^2. Each chip handles only some of the baselines. No additional interconnections are needed; there are just N to each board. The board will also have a pin limit; if N exceeds this, then the time-multiplexing used at the chip level can be used here as well.

Nor is it necessary to limit a unit to a single board, but expansion to multiple boards makes it harder to maintain the matrix topology. It can be done if the number of boards is kept square ($= P^2$), but every antenna signal must be distributed to $2P - 1$ of them via a backplane or similar mechanism. The backplane interconnections are not quite so regular, so this is beginning to get messy. If there are so many boards that they do not fit in one chassis (serviceable by one backplane) then it is messier still. Nevertheless, multiple-board units are practical if the number of boards is small; $P = 2 \ldots 4$ (4 to 16 boards) may be acceptable at moderate N, since this implies $3N$ to $7N$ logical backplane signals.

So finally we have a practical limit on the size of one correlation unit: a few circuit boards (or, at very large N, perhaps only one circuit board). Such a unit should be designed to correlate all antennas for as much of the bandwidth/time-segment as possible, thereby minimizing the number of units K. This makes the number of topologically separate interconnections NK as small as possible.

Having done this, each of the NK connections has a fixed path, from antenna n to correlator k where $n = 1 \ldots N$ and $k = 1 \ldots K$. There is rarely a need to switch any of them.

The signal format on each connection is a matter for detailed design, and the best choice (least expensive, most reliable) depends on available technologies. It is possible either to time multiplex all sample bits onto a single wire/fiber or to transmit all bits of several samples simultaneously on many wires.

4. Future Upgrading

Attempts are sometimes made to include in the design provisions for future upgrading, and this may alter some of the conclusions reached above. For example, if it is assumed that in a future technology each correlator unit will be able to provide the same processing at a higher speed, then all antenna electronics and interconnections might be built for larger bandwidth and initially operated at lower clock rates; when the correlator units are eventually replaced, the clock rates could be increased and we would immediately have larger bandwidth. Expansion in other dimensions, such as the number of antennas, could also be considered.

It might also happen that the initial implementation is limited only by funding, in which case provisions could be made for adding more correlator units or more antennas of exactly the same design when money later becomes available.

Each of these upgrade strategies is likely to fail. In the first case (assuming technology advances), it is expensive and risky to overbuild one part of the system so as to support an anticipated upgrade of another. There would have to be strong reasons to believe that the technology of the first part will advance to a much smaller extent than that of the second part. Partitioning the system to match such a prediction is difficult. In the second case (expansion of the original design), if several years elapse before the expansion then it is likely that building additional copies of the original units will be unattractive (because better parts are available) or impossible (because the original parts are no longer available).

It makes more sense to match all parts to present technologies, and to replace all of them when and if sufficient improvements become feasible.

5. Example

Consider a telescope with the following parameters, which are representative of the ATA: $N = 350$, $B = 200\,\mathrm{MHz}$, $M = 2000$. (The bandwidth includes two channels of opposite polarization, and each is to be analyzed into $M_1 = M_2 = 1000$ channels. But this has little effect on the discussion here.) Suppose that we have available an interconnection technology that allows up to 10 Gb/s to be transmitted on one connection.

It has been estimated [6] that a cross-correlation board can be built to handle 1 MHz of bandwidth for all baselines (0.5 MHz for each polarization) by using 4 large FPGA chips (Xilinx Vertex-II series). This assumes an FX architecture, so that the signal has already been analyzed into the required 0.1 MHz channels and the samples are represented as complex numbers. If one of these boards is taken to be a correlation unit, then we need $K = 200$ of them and $NK = 70000$ interconnections, each of which transmits about 16 Mb/s (assuming 16 b per sample, complex). This is a poor use of the available 10 Gb/s technology.

One solution is to build an intermediate device called a "corner turner" to re-order the data. We could then time-multiplex all data from one antenna (or from several antennas) onto a single connection for transmission to the corner turner, requiring 1.6 Gb/s per antenna (at 16 b/sample); after a massive re-organization, the data is time-multiplexed so that 1 MHz of the bandwidth (10 of the 2000 spectral channels) from all antennas is on each of $K = 200$ connections for transmission to the cross correlators, requiring 5.6 Gb/s per connection. The number of interconnections is then reduced from NK to $N + K = 550$, at the cost of the corner turner. The latter can be implemented, for example, with a large and fast RAM. If $N = K$, there is also a memoryless implementation [7].

Another approach is to reduce K by aggregating a set of the cross correlation boards into one correlation unit. For example, a bank of 20 boards would handle 20 MHz of bandwidth. This would be transmitted on a 320 Mb/s serial link from each antenna, which could be used to drive all boards in parallel via the back plane with each board ignoring all but its 5% of the data. There are now $K = 10$ such correlation units, requiring a total of 3500 interconnections. This is a large but manageable number. Each fails to make good use of the 10 Gb/s technology, so a less expensive technology would be used. The corner turner is completely avoided, so its cost and complexity are saved. It is likely (although I have not proved it here) that this approach is cheaper than the one that uses a corner turner.

There may be additional approaches that permit reducing K much further. The board described above [6] needs to compute $(B/K)N(N + 1)/2 = 6.14 \times 10^{10}$ complex multiply-accumulate (MAC) operations per second. But one correlation board of the ALMA correlator [2][8] computes 3.28×10^{13} real MAC/sec. Both boards are the same size (9U). Allowing a factor of 4 for the complex vs. real calculations, the ALMA board is 133 times more powerful. The difference is explained by several factors: (a) The ALMA board operates on 2-bit numbers, but the FPGA board uses (8,8)-bit complex numbers. This has little effect inside the chips, where gate count is dominated by accumulators, but it affects pin count and intra-board wiring. (b) The architectures are different: ALMA uses a lag (XF) correlator, so each input sample is used for many MACs per baseline (64 or 128, depending on mode); the FPGA design is based on an FX approach, where each sample is used only once per baseline. This too has little affect on computational requirements, but it affects I/O. (c) The ALMA board uses full-custom chips, rather than general-purpose FPGAs. One result of (a) and (b) is that the intra-board wiring limits the one board to carrying only 4 FPGAs, while the ALMA board carries 64 of the custom chips. Overall, there appears to be a potential for increasing the capacity of the correlator boards for the case at hand by a factor of at least 16 and possibly as much as 64. That would allow a $K = 1$ system to have 4 to 16 boards with only 350 interconnections, or a one-board-per-unit-correlator system to have $K = 4$ to 16, with 1400 to 5600 interconnections.

REFERENCES

[1] J. Bunton, "An improved FX correlator," ALMA Memo 342, 2000-Dec-19. `http://www.alma.nrao.edu/memoseries`

[2] R. Escoffier, "The ALMA correlator." MMA Memo 166, 1997-Apr-01. `http://www.alma.nrao.edu/memoseries`

[3] B. Carlstrom and P. Dewdney, "Efficient wideband digital correlation." *Electronics Letters*, 2000.

[4] A. Rogers, "Array processing estimates for the ATA." ATA Memo 12, 2000-Nov-04.

[5] L. Urry, "A new correlator architecture eases wiring problem." BIMA Technical Memo #71, 1998-Nov-13.

[6] D. Werthimer, private communication (email of 2001-Mar-08).

[7] L. Urry, "A corner turner architecture." ATA Memo 10, 2000-Oct-30.

[8] J. Webber et al., "ALMA correlator." Chapter 10 of ALMA Project Book, revised 2001-Feb-07. `http://www.alma.nrao.edu/projectbk/construction`