ALMA Correlator Study
WP2.4: Corner-turn platform

Jack Hickish

July 2016

1 Introduction

The ALMA Correlator Study assumes an upgraded ALMA correlator will be of an FX architecture, widely employed for digital correlators in radio astronomy. Hence, the correlator will require a *corner-turn*, or data transpose, between the F and X stages. Such a transpose allows data processing to be parallelised on a per-antenna basis in the F-stage and a per-frequency basis in the X stage.

In this document we consider the physical hardware used to implement the corner-turn, possible candidates being

- Ethernet switches
- Custom back-plane
- Custom switching system

We begin by detailing the top-level specifications which determine the requirements of the corner-turning system. Specifications are taken from Rupen et al.

2 Interconnect Specifications

The specifications relevant to the corner-turning system discussed in this document are:

Rupen spec.	Description	Value	Symbol
2	Number of antennas	up to 80	N
5	BBC Bandwidth	14 GHz	B
4	Number of BBCs per polarization	1	n
7	Cross-correlation input bitwidth	4+4 bit	w

We make the following further assumptions:

1. The number of polarizations processed, p, is 2.

2. Correlators for each BBC are effectively independent. That is, the complete correlator may be constructed from n distinct correlators.

Considering a correlator for a single BBC, we may consider the inputs of the corner turner to be pN F-Engines. Each F-Engine generates data at a rate $B \times w$ bits/s. For the assumed ALMA correlator, this is 112 Gb/s.

In the interests of simplicity and platform agnosticism we consider the pN corner-turner inputs to be physically separate. However, each of these 112 Gb/s streams may be split over parallel interfaces in order to be practically feasible. For example, a 112Gb/s stream may be implemented using 3 parallel 40 Gb/s Ethernet interfaces, or via 12 10 Gb/s serial links.

Note that assuming that all pN correlator inputs are physically separate increases the complexity of the corner turner, but minimally constraints the F-Engines. In practice it may be possible to combine multiple antennas into a single F-processor, reducing the number of inputs to the corner-turner (with a corresponding increase in bandwidth of each input).

The output data-rate of the corner-turner is the same as the input, i.e., $pN \times 112$ Gb/s.

3 F-Engine interface

In this document we assume an F-Engine processes data from a single BBC for a single polarization of a single antenna. The total ALMA correlator has $N_f = pN$ F-Engines. We assume nothing about the F-Engine interface except that it is capable of outputting a total of 112 Gb/s over f independent links. Where $f > 1$, we assume that the mulitple outputs contain sub-bands of the total processed bandwidth. In essence, the downstream corner-turner and correlator is then f clones of a smaller correlator processing a bandwidth $\frac{B}{f}$.

4 X-Engine interface

We assume that a complete correlator system comprises N_x X-Engines, each processing a subset of the total correlator bandwidth. The number of X-Engines is determined by the computational performance of a single unit, and need not be related to the number of F-Engines, pN. There is no assumed requirement on the X-engine interface, other than a single X-Engine is capable of sinking its fraction of the total system bandwidth: $\frac{pNBw}{N_x}$. This may be achieved via a single wide-band link, or via x multiple parallel links.

Where the connection between the N_f F-Engines and N_x X-Engines is not direct (because, for example, it is mediated by an Ethernet switch) there is no requirement that the protocol of the F- and X-Engine interfaces should be the same.

5 Custom Backplane/Switch

The present ALMA correlator uses 16384 LVDS twisted-pair cables operating at 250 MHz, representing "the greatest design challenge in the system" [1], to connect the *station cards* (effectively equivalent to F-engines) to the *correlator cards* (similarly equivalent to the X-engines). With the increased specifications of the next-generation ALMA correlator a corner-turn implementation using the same technology would see the total number of cables increase roughly five-fold, mainly driven by the doubling in bandwidth and sample bitwidth. Increasing the per-lane speed by a factor of two or even four may reduce the total number of cables but for the purposes of simplicity we will consider other options.

PCI-Express is a common standard for connecting many processing boards via a backplane type configuration supporting transfer speeds up to 125 Gbps per endpoint (for Gen3 with 16 lanes). Additionally the standard allows data transfers between slaves by using bus mastering. All PCI-Express endpoints, however, must connect to either a root complex or a switch and given these devices with 16 or more endpoints are rare or non-existent we will not consider this technology for ngALMA (which will require upwards of 64 endpoints).

Another popular standard is RapidIO commonly used in high-performance computing and data centers. RapidIO supports both backplane- and switch-based systems with sub-microsecond latency transfers at speeds of up to160 Gbps per port. However the availability of switches with greater than 16 ports is limited and are generally in the form of chips that would need to be designed into a PCB backplane. Additionally the limited maximum payload size of 256 bytes may present challenges for a correlator corner-turn design compared to the jumbo packets supported by Ethernet.

Generally, we believe the cost and NRE of designing a custom corner-turn platform would greatly outweigh the benefits offered by technologies such as RapidIO or PCI-Express. This is especially important considering the off-the-shelf availability of very fast (∼100 Gbps) and very large (>64 ports) Gigabit Ethernet switches.

6 Ethernet Switch

An Ethernet switch has a variety of attractive attributes:

- No hardware NRE.
- Industry-standard interface, widely supported by commodity hardware (FPGA boards, CPU/GPU platforms).
- Extremely tolerant to changes in F- and X-Engine implementations or changes to number of antennas.
- Trivially supports hardware testing via CPU-driven test-vector injection.
- May supports conversion between protocols – 10/40/100 GbE.

Though the available Ethernet technology at the time of deployment of a new correlator is uncertain, we can demonstrate the feasibility of an Ethernet corner-turn solution based on 2016 technology and assume that future solutions will be cheaper and denser.

We hypothesize the following system:

- $N_f = pN = 160$ F-Engines
- F-Engine output protocol is 100 GbE, with $f = 2$ independent interfaces, each carrying $\frac{112}{2} = 56$ Gb/s.
- $N_x = 256$ X-Engines, each processing 89.6 Gb/s, over a single 100 GbE link.

Such a system requires, for each of the two BBCs, two duplicates of a corner-turner with 112 F-Engine inputs, and 128 X-Engine outputs. Thus, the corner-turner is a switch with at least $128 + 112 = 240$ 100 GbE ports. Such switches are available off-the-shelf today. For example, the Arista 7508R with up to 288 100 GbE ports.

Alternatively, one can construct a larger switch from smaller modules, which can have more predictable performance for the all-to-all corner-turn systems required by correlators. In this case, the same system can be achieved with 8 individual 64-port 100 GbE switches, such as the 7260CX-64.

For other interfacing standards, such as 40 GbE using F-Engines with $f = 3$ or $f = 4$ output ports, similar systems can be constructed using 40 GbE switches.

References

[1] Escoffier, R. P., Comoretto, G., Webber, J. C., Baudry, A., Broadwell, C. M., Greenberg, J. H., Treacy, R. R., Cais, P., Quertier, B., Camino, P., Bos, A., and t, A. W. Gun. The ALMA correlator. *A&A*, 462(2):801–810, 2007.