# ALMA Correlator Design Considerations and Questions

Steven T. Myers
National Radio Astronomy Observatory

Updated Version, December 16, 1999

**Abstract**

There are a number of outstanding issues and questions pursuant to the design of the ALMA correlator, LTA system, and real-time pipeline. These are posed here in summary form.

## 1 Introduction

The current baseline correlator and LTA design can be found at

*http://www.cv.nrao.edu/~jpisano/mma_corr.html*

while a technical summary of the capabilities and open questions related to this design are outlined in the draft memo at

*http://www.aoc.nrao.edu/~smyers/alma/*

(to which this memo is an addendum).

## 2 Issues and Open Questions

While the capabilities of the proposed ALMA correlator design are the same as those presented in Rupen & Escoffier (Memo 194). and the scientific requirements are still those most recently presented in the white paper by Rupen, Shepherd & Wright (1998), there are still some open questions and critical issues that need to be resolved regarding this correlator design. *The* *statements in italics are my own assessment of the particular issue in question.* At the end of this memo is a glossary of some of the more confusing terms used in this document.

1. In Memo 190, it was stated that the 256 lag chip would be able to provide a maximum of 1024 lags ( + 1024 leads ) per baseline, which translates to 2048 lags ( + 2048 leads ) with the 512 lag chip. This is done by ganging the four separate arrays (quadrants) of the correlator together to compute the extra lags for one baseband pair instead of correlating different baseband pairs. Because lags from more than a single array must be fetched to do a given baseline's FFT, this means that the computing system must be more complicated than what would be needed if all of a given baselines correlations occur in only one array.

   **For the ASAC:** Is the spectral resolution attained without crossing arrays (16384 channels over 62.5 MHz, 3.8 kHz per channel, or 38 m/s at 30 GHz and 10 m/s at 115 GHz) sufficient for the science drivers, or is the extra factor of 4 resolution worth the added complexity?

   **For the engineers:** Note that if full configurability of board inputs was available (see below), then the correlations could be adjusted so that all results for a given baseline were done in a given array, for example by only correlating a given quarter of the baselines in a quadrant of the correlator (effectively spreading the antennas over the 4 arrays, instead of the lags), and using the four boards of a given correlator plane to do the extra lags for each baseline. Is this possible and/or desirable?

   *I expect that the extra lags (and resolution) are not needed immediately, and as long as the capability could be gained through a reasonable upgrade of the computing and correlator interconnect later on, I would be inclined to keep the downstream computing as simple as possible at first. The option of reconfiguring the correlator inputs to put all lags of a given baseline on the same array, if possible, is even better. On the other hand, it may be a better design policy to keep the hardware simple, and hope that the ever improving computing and software can deal with the added complexity later on.*

2. In order that the first quadrant of the correlator be useful when delivered in 2004 with less than half the antennas available, it must be able to correlate all 4 baseband-pairs (which are normally mapped to

the different quadrants). Thus, there must be a mechanism to run the first quadrant to correlate 4 baseband-pairs for 32 antennas instead of 1 baseband-pair for 64 antennas (the standard mode when the correlator and antennas are complete) or even 2 baseband-pairs for 32 antennas (obtained by replacing the second 32 antenna inputs by a second baseband-pair and throwing away the products between different baseband pairs).

**For the engineers:** It is thought that some disposable hardware could be used to make the first quadrant be able to handle 4 baseband-pairs for 32 antennas. Should the capability to trade antennas for baseband pairs be part of the overall correlator design, and not just a fix for the first quadrant? It is possible that budgetary constraints will lead to having less than 64 antennas in the final US/European array, or that some subarray configurations might get by with less than the full correlator for some of the baselines. Thus, it might be desirable to be able to use some of the spare antenna inputs for extra things. Also, this might allow correlation of extra lags (to get higher spectral resolution) for a given set of baselines in the same quadrant, thus keeping the downstream processing simple.

**For the ASAC:** Is extra flexibility in trading antennas for baseband pairs (or lags) is sufficiently desirable to mandate this for the normal operation of the correlator, or can this be a throw-away kludge for the initial stages?

*My gut feeling is that the flexibility will be important down the line, so if it can be done, or later added easily as an upgrade, this is worth the trouble.*

3. The previous LTA spec was that one result could be read out per output stream (FPDP bus) every 80ns. Note that in auto-correlation mode (where sixteen 1 ms products are stored and output each 16 ms accumulation time), in each quadrant of the correlator we get

$$32 \text{ ant} \times 16 \text{ ms products} \times 512 \text{ lags} = 262144 \text{ results}$$

per 16 ms accumulation time, which is half the cross-correlation rate of

$$32 \text{ ant} \times 32 \text{ ant} \times 512 \text{ lags} = 524288 \text{ results.}$$

Thus, 2 output streams per quadrant at 80ns (or 4 streams at 160 ns for a VME bus) per result can clear the auto-products in 10.5 ms.

Thus, the OTF specs presented in Rupen, Shepherd & Wright could be easily satisfied. However the initial one stream at 80ns is a factor of two too slow. Therefore, a minimum output stream rate of

$$2 \text{ streams} \times (80 \text{ ns})^{-1} \times 4 \text{ bytes} \;=\; 100 \text{ MB/s}$$

for the LTA to FFT data bus to handle for each quadrant (400 MB/s minimum for the entire correlator). There is provision in the updated LTA design to be able to output results every 20 ns, where a single stream per quadrant could clear the auto-products every 5.2 ms and cross-products in 10.5 ms (but at 200 MB/s)!

**For the engineers and ASAC:** Given that the minimal bus rate of 100 MB/s is pushing our computer design as-is, is it worthwhile to build in the capability to read the LTA every 20ns now? If so, do we increase the bus speed to deal with this in a single stream, or increase the number of streams, or both? Do we cut the initial bus rate to half this or less now, and plan on eventual upgrade?

**For the ASAC:** do you foresee important science drivers to eventually push even higher rates (eg. OTF cross-correlation mode, fast solar flare monitoring, pulsars)?

*The science driver for data dump rate is OTF total-power, which probably needs the 100 MB/s rate, and thus that should be our minimal goal. Note that this is inline with previous science specifications. It would be nice if there were easy upgrade paths also.*

**Further questions:** Is the limitation of the dump time fundamental enough to warrant desiring to access the correlation products at their maximum rate of 1 ms (bypassing the LTAs)? Are further speed increases necessary above that? For example, faster scanning would be allowed for cross-correlation if this were the case.

4. The ability to flexibly subarray is important to the ALMA design (as was stressed in previous documents), as it impacts the way that we plan to do baseline determinations, do total-power observations along with full interferometry, be able to monitor time-variable phenomena, possibly incorporate VLBI into ALMA operations, and other considerations.

   **For the engineers:** What are the implications of being able to subarray on hardware and software? In particular, are there restrictions built in to how the subarrays must be allocated? For example, it would seem to me that the minimum block that can be allocated to a subarray is that which feeds a given adder tree / FPDP stream. Also, if the arrays for a certain subarray are planned to be ganged together for extra resolution, then the correlator has to know to take the signals from array 1 to feed consecutively to the other arrays. Is this a problem? Is the ability to subarray totally under control of the software interface to the correlator and real time system?

**For the ASAC:** If there are software/hardware restrictions, how strict can these be before they impact the proposed operation? In particular, are there any hard minimal specifications?

*Because of the size and potential of ALMA, I feel that it would be a crime to design out any flexibility in subarraying at this point. My gut feeling is that the added complexity in hardware would be small, and the difficulty in software manageable. There are too many open possibilities to restrict subarray allocation (eg. adding a number of small antennas to the array, continuous reconfiguration, etc.).*

5. It appears that this correlator design meets all the desired specs (except for planetary radar, see below) and our current limitation is the handing of the LTA output for FFT and post-processing. The only things the correlator and related real-time system seem to restrict are the fundamental sampling rate (125 MHz), minimum dump time (1 ms auto-products, 16 ms cross-products), and the number of lag products per intersection (512).

   **For the Engineers:** Are there reasonable (easy and cheap) upgrade paths from this correlator design? For example, one could imagine adding 32 more planes to each array (for extra lags and increased resolution)?

   **For the ASAC:** Is the spectral resolution of this correlator design too limiting? If so, then adding more planes at the current sampling rate would increase the resolution (as increasing the sampling rate would have widespread changes throughout the system). Going to an FX design might also build in more flexibility, but likely at maximal cost.

   *My feeling, after looking over all the correlator specs, is that the proposed correlator design can fulfill all the foreseen science requirements, and that simple incremental upgrade paths can be designed in to cope with most likely future needs.*

**For the ASAC:** Given the above, should ALMA be planning (and more to the point, budgeting) now on an advanced correlator, or given the likely budget constraints, should the project focus on improving other things such as receivers?

6. One science application that does require extremely high spectral resolution is that of planetary radar, where resolutions of near 1 Hz are required. There are radars under development at 7–10 mm and 3mm wavelength, so this should not be discounted for ALMA. Note that 1 Hz at 30 GHz corresponds to 0.1 m/s velocity, which is why this is so critical.

   Nominally, the best we can do is 3.8kHz (see above) for one array, and 954 Hz for all four arrays ganged together. An additional factor of nearly 1000 is thus needed. It is true that if the fundamental sample rate of 125 MHz (which with the $32\times$ factor becomes 4 GHz which is required for 2 GHz bandwidth) is somehow switchable to 125 kHz, then the resolution would be correspondingly increased. However, this does not appear to be in the current design.

   **For the ASAC:** Is this of sufficient importance to include in the correlator design? Could (or should) this be deferred for a second generation correlator? If it were not to be in the design, but add-on (possibly complex) hardware could be designed to do this, should this be made a WBS item?

   **For the Engineers:** Does the current design support ultra-high (Hz) resolution? If so — how? If not — is there an upgrade path, or some reasonable add-on hardware to deal with this special case. Given that 1-Hz resolution is a science specification for the VLA Upgrade correlator, could (or should) a joint solution (for ALMA and the VLA) be pursued?

   *My guess is that the current design does not easily acommodate 1 Hz spectral resolution, and that if it is deemed important to have available this capability at the start of ALMA, then some extra hardware may be required.*

## 3 Glossary

**amplifier chain** Actually, the chain of amplifiers and mixer (in case of SIS) which can deal with a single polarization in the RF band which includes both sidebands. There are 2 of these (one for each polarization) in a given *LO/IF chain*.

**array** One quarter or quadrant of the correlator consisting of 32 planes each with 4 correlator boards that can handle the cross-correlations for 64 antennas with 512 lags times 32 planes.

**FPDP** The Front Panel Data Port is the hardware envisioned to read the LTA data onto the bus so that the FFT computers can acquire the lag products needed

**intersection** The part of a single correlator *plane* where the paired digital streams (one for each polarization E and H) from two antennas X and Y get multiplied for a given number of lag products. For example, in the upgraded chip design, a total 512 lag channels can be produced at each intersection (512 in a single polarization product, 256 for two polarizations, and 128 for the full four polarization products). Note also that for every pair of antennas XY, there is a lag intersection XY and lead intersection YX which must eventually combined to form the complex visibilities for the desired frequency channels.

**LO/IF chain** One fully-tunable (over the bandpass of the receiver) set of the four outputs of a receiver consisting of both upper and lower sideband in each of two orthogonal polarizations. There is a single (tunable) LO input and two polarizations (from separate amplifier chains) which are sideband separated into 2 sidebands times 2 polarizations each from the same 2 GHz wide RF band.

**LTA** The Long Term Accumulator sits after the correlator, and accumulates (coadds) the 1 ms output from the correlator if needed for multiples of the basic 16 ms accumulation period. The fundamental dump rate of the correlator is given by the integer number of accumulator periods (1 or more) that elapse during the time that it takes to read the full correlator output. Note that the LTA itself is accumulating the next set of data while the "current" set of outputs are read to the bus.

**plane** One layer of an array of the correlator consisting of 4 boards. One plane can handle cross-correlation of 64 antennas with 512 lags each.

**polarization** In most of the planned receivers, there will be two orthogonal polarizations (for each LO/IF chain) which get processed through separate amplifier chains and form independent correlator products. These I have designated as E and H plane, for their waveguide counterparts, since the actual sky polarization (whether linear EH or circular RL) depends upon the actual receiver design for a given band.

**quad** (or **quadrant**) See *array*.

**stream** A time-series of digital samples, from the antenna digitizer (4 GHz sampling), the multiplexed digital signal downlinked from the antenna to the correlator station, the inputs to the correlator, or the output from the LTAs on the bus to the FFT computers, depending upon the context. For example, there are 8 digital streams into the downlink MUX. Or 2 or 4 streams from a single quadrant LTA assembly.

## 4 Update History

1. *First Draft, 6 Dec 1999, stm*
2. *Updated, 16 Dec 1999, stm, added planetary radar issue*