

ALMA CORRELATOR

*John Webber
Ray Escoffier
Chuck Broadwell
Joe Greenberg
Alain Baudry
Last revised 2001-02-07*

Revision History:

1998-09-18: Added chapter number to section numbers. Placed specifications in table format. Added milestone summary.

1999-04-09: Revised milestone dates and made date format conform to adopted standard. Revised tables and some text to reflect adoption of digital FIR filter. Changed text to reflect architectural change in delay line implementation. Revised block diagram.

2000-02-04: Changed maximum number of antennas to 64. Changed block diagram to correspond to current thinking about system architecture.

2000-03-31: Revised from MMA to ALMA baseline correlator. Incorporated changes resulting from correlator and systems PDRs.

2000-04-12: Clarified path of digitizer development and meaning of subarrays.

2000-12-05: Minor revisions to a few numbers. Addition of section describing possibilities for a future correlator.

2002-02-07: Minor revisions to text, mostly editorial. FX correlator reference added.

Summary

This section describes the ALMA correlator. The design described here is for a lag correlator with a system clock rate of 125 MHz. The goals of Phase 1 are to produce paper designs and some simulations of all major correlator elements, including the correlator chip, and to fabricate and test prototype hardware. The goals of Phase 2 are to produce a prototype minimally populated correlator, deliver such a prototype for use in the test interferometer, and deliver the complete correlator to the ALMA site.

Table 10.1 ALMA Correlator Specifications

Item	Specification
Number of antennas	64
Number of baseband inputs per antenna	8
Maximum sampling rate per baseband input	4 GHz
Digitizing format	3 bit, 8 level
Correlation format	2 bit, 4 level
Maximum baseline delay range	30 km
Hardware cross- correlators per baseline	1024 lags + 1024 leads
Autocorrelators per antenna	1024
Product pairs possible for polarization	HH, VV, HV, VH (for orthogonal H and V)

Although the specification is for 64 antennas, the design may be changed to accommodate a larger or smaller number of antennas with some impact on schedule. However, once the production units are begun, no change in the maximum number of antennas can be made without substantial redesign.

Table 10.2 Principal milestones for ALMA correlator

Preliminary Design Review	2000- 01- 20
Prototype correlator Critical Design Review	2001- 07- 31
Deliver prototype correlator to VLA site	2003- 05- 30
Deliver first quadrant to Chajnantor site	2004- 06- 18
Deliver last quadrant to Chajnantor site	2006- 10- 06

10.1 System Block Diagram

The system architecture described has been chosen as the best tradeoff to produce high reliability, robust operating margins, a minimum number of integrated circuits, and a minimum number of cable interconnects (see ALMA Memo 166). The performance of this architecture permits high versatility in correlator operation (see ALMA Memo 194). The adoption of a digital FIR filter eliminates many potential sources of systematic error (see ALMA Memo 204 and ALMA Memo 248).

The correlator system envisioned for ALMA includes digital filters, mode selection, a delay line and data format conversion stage, cross- and auto- correlators, long term accumulation, and initial digital computer processing.

A simplified functional block diagram for the ALMA correlator is given in Figure 10.1. This diagram presents a fairly conventional lag correlator except for the presence of the packetization stage.

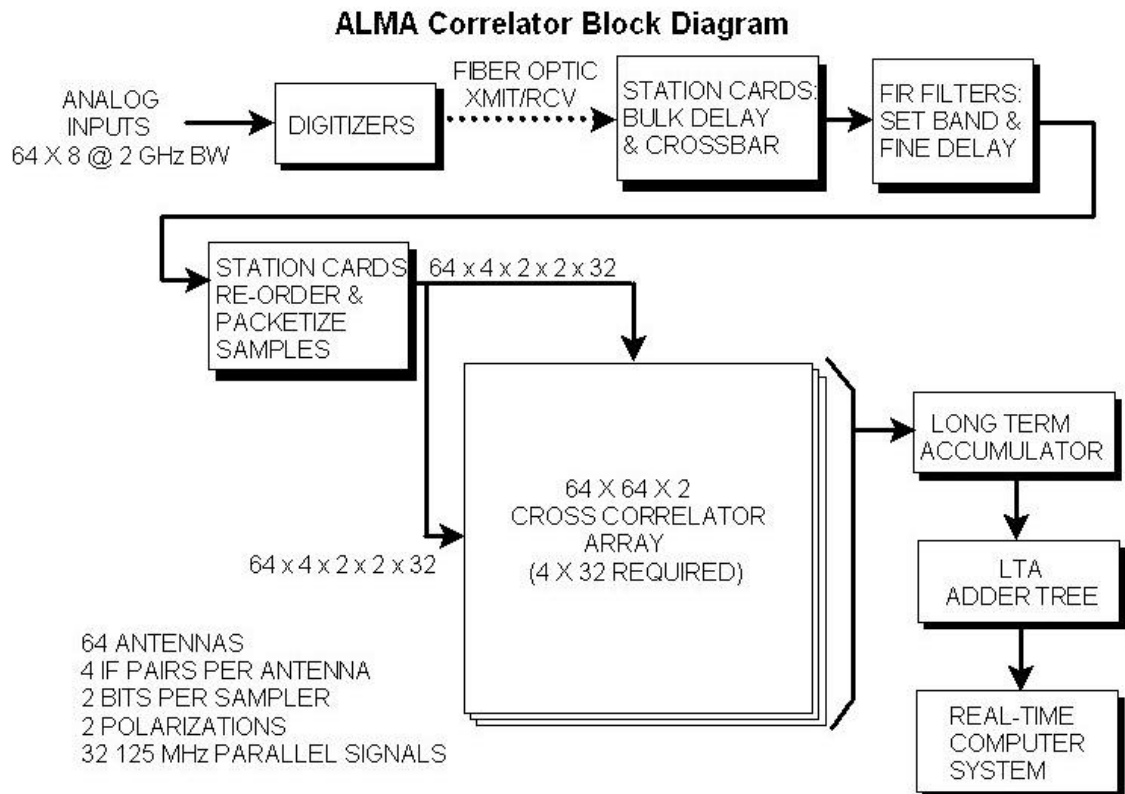


Figure 10.1: Simplified correlator functional block diagram. The digitizers, fiber optic IF transmission system, and real-time computer system are not part of the correlator but are shown for clarity. Computers which control the correlator are omitted.

The analog outputs of the baseband system drive digitizer inputs where 3-bit, 8-level digitization is performed at 4 GSamp/second. For details, see section 9.2. The samples are transmitted over fiber optic cables at a rate of 96 Gbit/s from each antenna to the correlator.

When less than 2 GHz bandwidth is desired, the samples are used as the input to the digital filter. The use of 3-bit quantization at the FIR filter input results in a small (~4%) loss of SNR compared to perfect analog filtering; the output re-quantization to 2 bits provides suitable input to the correlator. State counters are provided on both the input and output stages to permit setting analog levels to the digitizers correctly and for use in total power calibration.

Logic in the station cards routes outputs from the digital filters into the station cards which perform the packetization function. When fewer than 8 digitizers per antenna are being used, this stage will assure high system efficiency by replicating active digitizer outputs into unused memory areas and hence into otherwise unused correlators where additional lags can be generated. In this way, maximum performance will be obtained for the observational mode desired.

The digital filter stage will also do the sample decimation for observations in which sample rates less than 4 GS/second are needed. A 32-sample delay is required just before the digital filter in order to perform the finest resolution delay adjustment.

Adjusting the signals to the appropriate timing by means of a bulk delay is provided on the station cards which precede the FIR filters, in very efficient high density RAMs. For a 30 KM delay range, 524,288 RAM bits per digitizer output bit are required.

The packetization block seen in figure 10.1 will take the 32 parallel outputs of each digitizer and, using RAMs, both generate lags and re-sort the samples. In this block, the 32 parallel outputs of a high speed digitizer would be converted from each carrying every 32nd sample to each carrying short (about 1 msec) bursts of contiguous samples. If the N-wide parallel (2-bit) output of a high speed digitizer (each output carrying every Nth sample) were to drive the correlators using a conventional architecture, an N-by-N matrix of correlators would be required to insure every sample is correlated with every other sample. For N = 32, this would mean a matrix of 1024 small correlators to correlate the output of every baseband input of every baseline.

By using the format conversion scheme, the 32-wide parallel output from a high speed digitizer will be transformed into 32 parallel signals each carrying 1 millisecond data packets of contiguous samples that need drive only an N-by-1 array of correlators. This simplification in the correlator circuit requirements is obtained at the cost of an inefficiency of about 0.2% which results because the end bits in adjacent 1 msec time segments of samples will not get correlated with each other.

(Note that the conversion from a conventional N-by-N architecture to an N-by-1 architecture does not improve the spectral resolution performance of the correlator. The performance is set by the number of hardware correlators in the system. The conversion does, however, greatly simplifying the system wiring in that all N-by-N signals from two

antennas do not have to be wired to closely spaced electronics, thus simplifying the wiring matrix driving the cross correlators as well as reducing the number of I/O pins required by logic cards and integrated circuits.)

An additional benefit of the format conversion strategy is that it allows the system the same advantage as a recirculating correlator: when the bandwidth being processed is reduced by a factor of 2, the number of lags the system is capable of generating goes up by a factor of 2. This results in a factor of 4 increase in frequency resolution for a factor of 2 decrease in bandwidth.

Still another advantage of the format conversion (by far the most important in the ALMA correlator) is that it allows a minimum cable interconnect complex between the station electronics and the correlators. It also eliminates any requirement to interconnect correlator arrays in low bandwidth modes. Since the number of data interfaces between these two stages in the ALMA correlator surpasses that of any other astronomical correlator system by a factor of almost 100, this aspect of the system architecture is most important.

The cross correlator matrix of figure 10.1 is used to correlate the digitizer outputs of every antenna with those of every other antenna. At the intersection of any antenna X and another antenna Y in this matrix, there will be a correlator chip. This correlator will compute lag products for the XY baseline while the antenna Y and antenna X intersection of the matrix computes the baseline lead products. Auto correlation products for each antenna are obtained from correlators on the matrix diagonal.

In order to minimize further the station electronics to cross- multiplier cable interconnect, a very compact cross correlator matrix is essential. The design for the ALMA correlator places an entire 64 X 64 cross correlator matrix on 4 adjacent printed circuit cards, constituting a correlator *plane*. Each plane handles a 1/32 slice of the 4 GHz decimated digital data stream, at a 125 MHz data rate.

The distribution of signals from the station electronics to interface boards ("paddleboards") on the rear of the correlator planes assures that no signal drives more than one load. Two versions of the paddleboard will be produced: the first version will distribute the signals such that all IF signals from up to 32 antennas can be fully processed using only one quadrant of the correlator; the second version will distribute the signals such that all IF signals from the full 64 antennas can be processed using the entire correlator. This allows for interim operation of up to half of the array with one quadrant of the correlator while the remainder of the correlator (and antennas!) is built.

One disadvantage of this architecture is that once the number of antennas for the array has been set, future expansion of the correlator beyond this number is not practical.

The custom lag correlator chip has a dual 4- by- 4 array of correlators, each handling 2 polarizations. The chip can be programmed via a microprocessor supplied program word for its position in the matrix, which will select one of three correlator configurations:

1. four short correlators to compute the lags of all 4 polarization products (HH, VV, HV, and VH).
2. two longer correlators to compute just the lags for the two polarization components (HH and VV).
3. a single long correlator to compute lags for only one of the two baseband

inputs.

The estimated size of this custom correlator chip is approximately 2,000,000 gates. It will run on a 125 MHz clock. The chip package will be an industry standard surface mount package with 240 pins.

Each individual lag of the correlator chip consists of a 2-bit, 4-level, times 2-bit, 4-level, multiplier whose output is integrated in a 25-bit accumulator. Each accumulator has secondary storage for the 16 most significant bits. Each chip contains $256 * 16 = 4096$ of these lags.

There are eight Xilinx FPGAs on each correlator card for reading the accumulated results from the correlator chips and transmitting the results to the Long Term Accumulator Cards.

Each correlator card also has five Xilinx FPGAs to perform the Analog Sum function, which is required for forming a single beam from a set of participating antennas for Very Long Baseline Interferometry.

For observations in which fewer than 8 baseband inputs are being used, more lags can be produced by dedicating more than one correlator array to process the outputs of active baseband inputs. In this case, cards in the data format conversion stage will be used to form a virtual connection, the effect of which is to link two or more correlator arrays in series. The delayed input to the correlator chips that are to compute the higher level lags will be displaced in time the appropriate number of bits by offset RAM addressing in the data format conversion cards.

The long term accumulation block seen in figure 10.1 integrates the correlator outputs for the desired duration. The correlator chips will produce a total of 1,048,576 lag results in one plane. The long term accumulation block must provide double buffered integration storage for every result (since in some modes, every plane has distinct sets of lag results) in a total of four separate bins. This is a requirement of 1,073,741,842 storage locations, spread over 64 long term accumulator cards, or 16,777,216 results per card.

The adoption of a digital FIR filter has a potential system-wide consequence: it makes more attractive the baseline plan of performing the digitization at the antenna and transmitting the data to the correlator over a digital rather than an analog fiber optic link. This is due to the fact that, with analog filters, sampling at the antenna implies placing the analog filters at the antenna, with resulting stringent specifications on filter temperature stability which could be difficult to meet. The advantage of digitizing at the antenna is that the limited SNR and gain instability of an analog fiber optic link are eliminated. The disadvantages are possible shielding difficulties for the sampling clock and the (at present) high cost of digital transmission compared to the cost of two 8 GHz wide analog channels.

10.2 Performance

This section gives performance parameters for some typical operating modes of the ALMA correlator. The ALMA correlator will be programmable on a baseband by baseband basis and, hence, some baseband inputs may be processed in one mode

while other baseband inputs are processed in other modes.

Bandwidths per baseband input range from a maximum of 2 GHz down in factor of 2 steps to 31.25 MHz. For 8 baseband inputs per antenna, this yields a maximum bandwidth per antenna of 16 GHz.

Sub- arrays will also be possible using the ALMA correlator. The maximum number of correlator sub- arrays for ALMA is limited to 16 by the adoption of automatic transfer of results from the Long Term Accumulator to the VME computers which will receive the data.

There are 8 digitizers per antenna. The baseband inputs driving the digitizers will consist of 4 dual polarization pairs; for each pair, 4 polarization cross- products may be computed. Each digitizer is assumed to digitize at 4 GHz and hence to be driven by RF signals at most 2 GHz in bandwidth. The maximum bandwidth processed is thus 16 GHz split into 2 GHz pieces. Note that the analog baseband constraints of the planned ALMA baseband processing system will impose limits as well.

The smallest division of lags in the projected correlator chip is 64 lags. Because of the architecture, this will produce 64 lead and 64 lag channels and hence 64 spectral points per product. This smallest correlator division means that in the full- up configuration, all baseband inputs active at maximum bandwidth and all 4 polarization products being computed, 64 spectral points will be produced for every baseline, every spectrum. This gives a frequency resolution per spectral channel of 31.25 MHz.

Given the full- up performance as defined above, the number of lags that the correlator can produce for a given experiment results from the following considerations:

1. If polarization cross- products are not required, a factor of 2 more lags (finer resolution) can be obtained. The particular configuration can be selected on a baseband pair by baseband pair basis.
2. If fewer than 8 baseband inputs are required, lags go up as 1 over the fraction of baseband inputs used ($1/2$ the baseband inputs, 2 times the lags).
3. If a lower bandwidth than 2 GHz per baseband input is required, lags go up as 1 over the fraction of maximum bandwidth ($1/4$ the maximum bandwidth, 4 times the lags) until a factor of 32 is reached. After that, the number of lags stays constant. The particular configuration can be selected on a baseband by baseband basis.

Note that item 3 implies the characteristic described above that for each reduction by a factor of 2 in bandwidth, an increase of a factor of 4 in resolution is obtained (up to the factor of 32 limit after which the resolution improves by only 2 for each factor of 2 reduction in bandwidth).

Table 10.3 below illustrates some of the possible modes. The first four columns relate to the correlator proper. The columns relating to velocity range and resolution assume 90% of the analog bandwidth will be usable. (See ALMA Memo 194 for additional illustration of the ALMA correlator performance.)

Table 10.3 Selected correlator modes

# of Digitizers	Bandwidth/ Digitizer	Cross-pol Products?	Channels/ Product	At 230 GHz, in velocity space:	
				Range	Resolution km/s
8	2 GHz	Yes	64	9391	40.8
8	2 GHz	No	128	18783	20.4
8	1 GHz	No	256	9391	5.1
8	500 MHz	Yes	256	2348	2.5
8	250 MHz	No	1024	2348	0.32
4	2 GHz	Yes	128	4696	20.4
4	1 GHz	No	512	4696	2.5
4	500 MHz	Yes	512	1174	1.3
4	250 MHz	No	2048	1174	0.16
2	2 GHz	Yes	256	2348	10.2
2	1 GHz	No	1024	2348	1.3
2	500 MHz	Yes	1024	587	0.64
2	250 MHz	No	4096	587	0.08

Two natural time intervals associated with the correlator are 1 msec and 16 msec. These are the two short term integration cycles available in the correlator chips. The 1 msec short term integration cycle is available only on the array diagonal (auto-correlation results only). The 16 msec cycle is available both on and off the array diagonal (auto- and cross- correlation results). The Long Term Accumulator (LTA) provides longer term accumulation for the 16 msec results, in integer multiples of 16 msec, up to approximately 65 seconds maximum. The LTA does not provide longer term accumulation for the 1 msec results. It does provide buffers for 16 consecutive sets of 1 msec results, so access to the results is on 16 msec boundaries instead of 1 msec boundaries.

The function of the adder tree block seen in figure 10.1 varies with correlator mode. At

full bandwidth (2 GHz), the lag results from all 32 planes must be summed together in the adder tree, while at minimum bandwidth, distinct sets of lags are produced in each plane and must be passed through the adder tree block. Intermediate bandwidths require intermediate sets of planes to be summed.

In cross- correlation mode, using 16 msec integrations at minimum bandwidth, a 32 GByte/sec output rate would be required, if all results were transmitted out of the correlator. The correlator output capacity is specified as 1 GByte/sec (64 MByte/sec on each of 16 streams). Alternatives of longer integration times, restricted numbers of lags, or fewer baselines are provided, allowing lower output rates.

The current functional specification of the LTA is given in ALMA Memo 294.

10.3 Size and Power Requirement Estimate

Table 10.4 Preliminary ALMA correlator module & printed circuit card requirements

Item	# required	Size	Power
FIR filter card	512	6U euro card	80 w
Station card	512	6U euro card	20 w
Correlator card	512	9U euro card	160 w
Control card	160	6U euro card	40 w
Long term accumulator	64	9U euro card	60 w
TOTALS	1760		143,360 w

It is estimated that the station- dependent part of the system (digitizer, filter, mode, and memory) will require 1/4 a rack per antenna, or 16 racks for 64 antennas. The remainder of the system, proportional to the number of antennas squared (correlator, control, and accumulator) will occupy 16 racks for 64 antennas. The grand total of racks is therefore about 32.

The power estimates given in Table 10.4 above are based on the experience gained in the development of the GBT spectrometer. The biggest unknown at this time is the dissipation to be expected in the custom correlator chip, 32,768 of which will be required in the system. The GBT correlator chip dissipates about 5 watts with a clock rate of 125 MHz. Such a high chip dissipation in the ALMA correlator would mean both high system power requirements and lower reliability because of the difficulty in removing the heat from the system at the high altitude site.

By using low voltage chip technology it is predicted that the custom correlator chip described in this document can be built with about a 1.5 watt power requirement, 2 watt maximum. The chip represents about a factor 2 increase in the level of integration when compared to the GBT correlator chip (twice the number of transistors). By using a more

modern process, with finer component features and low voltage technology, a smaller chip with lower power requirements should be possible. The smaller silicon size should also mean a higher yield in the manufacturing process.

10.4 Second Generation Future Correlator

The ALMA Baseline Correlator described above will provide a correlator which meets the baseline science requirements. However, it is clear that advances in semiconductor technology and correlator architecture will eventually make that correlator obsolete, and that planning for a future correlator is necessary.

10.4.1 Introduction and Correlator Nomenclature

While the NRAO team designs the ALMA Baseline Correlator with delivery of the first quadrant scheduled by mid-2004, European teams in the ALMA project are working on a preliminary design study and sub-systems prototyping for a second generation Future Correlator. The goal is to offer higher correlation efficiency, more spectral channels and perhaps more flexibility than with the Baseline Correlator while benefiting from the technical advances one can anticipate before completion of the full ALMA Correlator. At the same time, the Japanese are working on an FX architecture design (c.f. ALMA Memo 342 at <http://www.alma.nrao.edu/memos/html-memos/abstracts/abs342.html>.) The current nomenclature is then, Baseline Correlator, Future Correlator and FX Correlator for the baseline ALMA project and the European and Japanese team projects, respectively.

With Japan joining the ALMA project the Enhanced ALMA will require an Enhanced Correlator or a change to the baseline correlator design to accommodate any additional 12-m antennas beyond the 64 served by the current baseline correlator design. The 10 to 16 smaller antennas of the ALMA Compact Array may be served along with the 12-m antennas by a single correlator, or by a separate, dedicated Compact Array correlator. This Enhanced Correlator should be defined by the joint efforts of the NRAO-European-Japanese correlator teams.

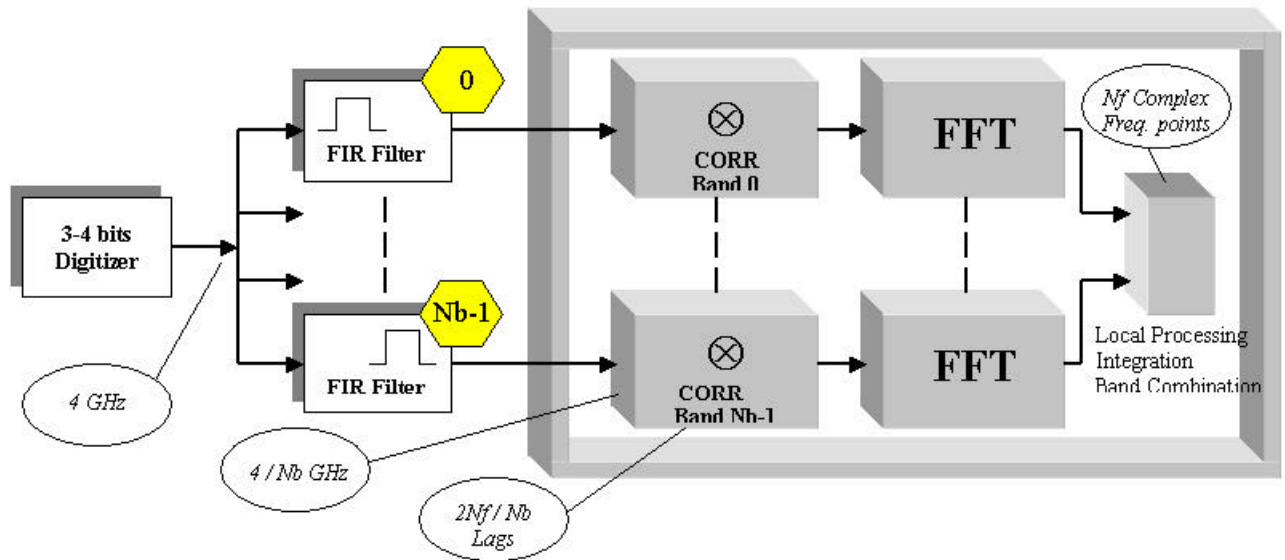
We concentrate here on the main features of the European Future Correlator. Europe proposes to prototype and build a digital hybrid correlator whose architecture is somewhat intermediate between the XF and FX architecture designs and incorporates features common to both designs. The basic concept was first discussed in a European kick-off meeting held in April 2000. Further presentation was made in September 2000 during the ASAC meeting and during an informal meeting held with the Japanese team. (See minutes of the latter meeting in http://www.eso.org/projects/alma/committees/backwg/minutes10sept_correl.doc.)

The European Enhanced Correlator concept and the WIDAR correlator concept proposed by the Canadian group for the VLA are closely related.

10.4.2 Future Correlator Overview

One key idea behind the digital hybrid correlator concept is that for a given spectral resolution the cross-correlation requirements diminish with the number of sub-bands sub-dividing each digitized 2 GHz input band. The current concept considers 16 partially overlapping sub-bands in order to match a correlator clock frequency of order 250 to 300 MHz. In this frequency-division

demultiplexing scheme there are several important issues under investigation: flexible interconnection of correlator cells, digital filtering and requantization, digital total power calibration for sub-band concatenation and FFT, etc. A simple block diagram is shown in Fig. 10.2. 3- or 4- bit correlation is a major goal of the Future Correlator study. The advantage is significant because it is beneficial to all types of observations compared to 2-bit correlation. Going from full 2-bit to full 3-bit operation diminishes the quantization losses and is equivalent to adding about 10% more collecting area.



ALMA Future Correlator

Digital Hybrid Concept
 (Frequency division multiplexing)
 One 2-4 GHz Sub-band

Figure 2. Future Correlator simplified block diagram

The number of sub-bands per 2 GHz total band is an a priori free parameter in the digital hybrid correlator design. This parameter has an impact on both the total cost (more FIRs means lower clock rates) and power consumption (depending not only on the adopted correlator chip design but also on the performances of the FIR filter ASIC design).

In contrast with the Baseline Correlator in which the FIR filters are used to narrow the input bandwidth, in the Future Correlator FIRs are used for both band narrowing and frequency demultiplexing of the 2 GHz input band. It is interesting to note that FIR filtering is similar to the FFT stage in the FX correlator design. In the Future Correlator the total number of spectral channels is driven by several issues and the number of lags per baseline associated with each 2 GHz sub-band is not yet adopted. The maximum spectral resolution depends on flexible allocation of lag resources using cross-bar switches to interconnect the correlator cells. The present goals are 8000 spectral channels across 8 GHz bandwidth and about 5 kHz maximum

resolution.

An initial document describing the signal routing and the hybrid system architecture based on generic correlator modules, cell partitioning and interconnections of an NxN array of cells fed by N correlator input signals has been prepared by A. Bos for discussion in the European team. Other on-going tasks include fast backplane technology studies, signal processing implementation platforms and VHDL simulations, and design of FIR filter ASICs operating above 100 MHz. In addition, there is good progress in the construction of an end-to-end model to simulate the Future Correlator signal flow and the effects of requantizing after the FIRs. An operational mock-up comprising both hardware and software sub-systems will be prepared in 2001 to test the Future Correlator concept and identify the most critical components/sub-systems. Results will be given in the ALMA Phase 1 Design Study document.