System Design Description

80.04.00.00-002-D

Status: *Released*

2004-02-20

Prepared By:		
Name(s) and Signature(s)	**Organization**	**Date**
Larry R. D'Addario	NRAO	
Approved By:		
Name and Signature	**Organization**	**Date**
Xxx	Xxx	yyyy-mm-dd
Released By:		
Name and Signature	**Organization**	**Date**
xxx	xxx	yyyy-mm-dd

Change Record

Version	Date	Affected Section(s)	Change Request #	Reason/Initiation/Remarks
A	2003-09-10			
B	2003-12-15			
C	2004-01-30			
D	2004-02-20			

ALMA Document 80.04.00.00-002-D
System Design Description
Larry R. D'Addario
2004-02-20

Table of Contents

[Note: In this document, information preceded by '**' is a system level requirement or specification. Most of these are summarized in the Appendix.]

1 INTRODUCTION

This document provides an overview of the ALMA telescope design. It attempts to cover the complete signal path from the arrival of a wavefront at the antenna aperture to the archiving of measurements, along with the principal supporting subsystems. However, many things important to the telescope's operation are omitted or touched upon only briefly. These include the antennas and their configuration (except for the antenna optics in the signal path); local infrastructure, such as roads, buildings, power distribution, cabling, and communication; and computing that is not required during an observation. Only the basic 64-antenna array is covered here; the Atacama Compact Array is not considered because, to a large extent, it is currently expected to operate separately from the basic array. If necessary, interactions between the two arrays will be considered in a later version of this document.

This is considered a "system level" description. It shows the major design choices that have been made in order to meet the requirements and goals in an efficient manner, but it avoids discussing design details that might have been done differently while achieving the same performance. Documentation of such design details is the responsibility of the group doing the design. It is a project requirement that the detailed designs remain consistent with the system-level description given here. Changes to the system level design can be made after appropriate review of proposals; if changes are adopted, this document will be revised to include them.

It is assumed that the reader understands the purpose of the ALMA telescope and its main performance requirements, which are given elsewhere [1].

2 MAIN SIGNAL FLOW

In this section, the components in the main signal path are described in the order in which they are encountered by the astronomical signal. Many points are more easily understood by referring to the system block diagram, ALMA Document 80.04.01.00-004 [3].

2.1 Antennas

Each of the **64 antennas of ALMA is a **12 m diameter paraboloid on an elevation-over-azimuth mount, with an hyperboloidal subreflector producing a secondary (Cassegrain) focus 1.37 m behind the main reflector's vertex. The subreflector is on a computer-controlled hexapod mount allowing full focus tracking, and, on some antennas, it includes a nutation mechanism that allows a beam throw of up to 5 arcmin at a rate of up to 10 Hz. (The present plan and budget allow for **4 antennas to be equipped with nutators, but there is some uncertainty about whether this is necessary or sufficient. The answer depends on developing techniques for mitigating the effects of atmospheric transparency fluctuations on total power observations, and should be decided after tests on the prototype antennas.)

Front end assemblies for the various bands (see below) are distributed to fixed positions in the focal plane, and one of them is selected by an adjustment to the pointing of the main reflector. Thus, only one can be in use at a time, but rapid switching is possible.

The antennas will be arranged spatially in a variable way that is beyond the scope of this report. There are 216 fixed locations (foundations with power and signal cables) over which the 64 antennas can be distributed. The furthest station from the central Technical Building is approximately 12 km distant, but most stations are within a few km.

2.2 Front End (FE) Assembly

The required **frequency coverage (all atmospheric windows from 31.3 to 950 GHz) is achieved in 10 separate bands, as listed in Table 1. The band edge frequencies were chosen to provide good coverage of the windows while limiting the **edge frequency ratio for each band to < 1.35, considered the largest feasible value. The exact frequencies follow closely the recommendations of a scientific working group [14], with two exceptions: Band 1 is justified by the minimum-noise region around 35 GHz for continuum, by the lines at 43 and 44 GHz, and by overlap with the VLBA. It is technically advantageous to stay within the 33–50 GHz WR22 waveguide band, but the lower edge is extended to 31.3 GHz so as to include the 31.3 to 31.5 GHz ITU allocation to radio astronomy; performance below 33 GHz may be degraded. For Band 3, a lower limit of 89 GHz was recommended and is sufficient to provide continuous coverage with Band 2. But the edge was

extended to 84 GHz, covering the SiO line at 86 GHz, because of concern by some astronomers that Band 2 will never be implemented and because of (undocumented) claims that the extended range is technically feasible without loss of performance.

Table 1: Frequency Coverage By Band

Band No.	Frequency Range, GHz
1	31.3–45.0
2	67–90
3*	84–116
4	125–163
5	163–211
6*	211–275
7*	275–370
8	385–500
9*	602–720
10	787–950

*Only these bands will be implemented initially.

Cryogenically cooled components of the front end that are specific to one band are in separate assemblies known as "cartridges," but these are housed in a common vacuum chamber and cooled by a common cryocooler. Only four of the bands will be implemented initially, namely 84-116 GHz (band #3), 211-275 GHz (#6), 275-370 GHz (#7), and 602-720 GHz (#9). The vacuum chamber with cryocooler and cartridges is part of the overall FE Assembly, located in the center of the antenna's receiver cabin, a room behind the focal plane that moves with both axes of antenna pointing. To maintain good optical alignment, the FE Assembly is rigidly attached to the antenna structure very near the focal plane.

For the two lowest-frequency bands, the initial active element is an HFET amplifier. For all others, it is an SIS mixer. All bands have two channels receiving orthogonal polarizations. The HFET amplifiers, along with the filters and mixers that follow them, are cooled to a nominal temperature of 15K and the SIS mixers are cooled to a nominal temperature of 4.0K. A single 3-stage Gifford-McMahon cryocooler is used; the remaining stage is nominally at 80K, and it cools the outer radiation shield, infrared filters in the signal paths, and (for most bands) frequency multipliers for the local oscillator. The cryocooler is supported by a single-stage helium compressor that is mounted outdoors on a platform that rotates with the antenna in azimuth but not in elevation.

Each band has its own tertiary optics, consisting of a lens or a pair of ellipsoidal mirrors, to match the wave that arrives at the secondary focus to a corrugated feed horn. The signal is separated into the two polarizations, nominally linear and orthogonal, and delivered in waveguide to the amplifier or mixer. (The highest frequency band, not yet designed, might use quasi-optical coupling of the wavefront to the active element, rather than a feed horn and waveguide.) For bands 1 through 6 (to 275 GHz), polarization splitting is achieved in an orthomode waveguide junction just after the feed horn. For higher frequencies, it is achieved with a wire grid within the tertiary optics, and in these cases each channel has a separate feed horn. For further details of the optics design, see [2].

In all cases, the signal is converted to an **IF band of 4 to 12 GHz, either by the SIS mixer or by a Schottky diode mixer following the HFET amplifier. The mixers for both polarization channels are driven at the same LO frequency. For the HFET bands, a filter ahead of the mixer produces a single-sideband response (upper sideband for band 1 and lower sideband for band 2). The SIS bands are of two different types: dual sideband (2SB) and double sideband (DSB). The 2SB case supplies two IF outputs simultaneously, carrying opposite sidebands of the LO. The DSB case supplies a single IF output carrying a linear combination (nominally equal) of responses from both sidebands. For the 2SB cases, the **undesired sideband is suppressed at least 10 dB. Bands 3 through 7 use 2SB mixers, and bands 8 through 10 use

3

DSB mixers[1]. In order to utilize fully the signal transmission system and correlator, we require that each polarization channel **deliver at least 8 GHz of instantaneous bandwidth at IF. For 2SB channels only, it is possible to achieve this by utilizing only 4 GHz of the 4–12 GHz IF band for each signal. If a 2SB mixer achieves 8 GHz of IF bandwidth (as most are expected to do), the requirement would still be met if one of the sidebands is not used, and the requirement is exceeded if both sidebands are used. We choose always to use both sidebands of 2SB mixers. With these points in mind, the design choices are listed by band in Table 2.

Table 2: Reception characteristics by band

Band	Implementation	Response	IF band	Total bandwidth per polarization
			GHz	GHz
1	HFET amp	SSB-upper	4–12	8
2	HFET amp	SSB-lower	4–12	8
3	SIS 2SB	2SB	4–12	16
4	SIS 2SB	2SB	4–12	16
5	SIS 2SB	2SB	4–12	16
6	SIS 2SB	2SB	4–12	16
7	SIS 2SB	2SB	4–8	8
8	SIS DSB	DSB	4–12	8
9	SIS DSB	DSB	4–12	8
10	SIS DSB	DSB	4–12	8

The requirement for 8 GHz instantaneous bandwidth per polarization for all bands was set by its expected feasibility in SIS mixers. It exceeds the minimum scientific requirement, set by the width of spectral lines, by a factor of several. This expected feasibility then drove the requirements for signal transmission and correlation, which were made to match it.

Each IF output from a mixer is amplified in a cooled pre-amplifier closely associated with that mixer. Gains are near 35 dB in most cases, allowing the subsequent cable loss (especially the transition to room temperature) to have negligible effect on the receiver noise temperature.

The amplified IF signals are brought out of the vacuum chamber on coaxial cables and delivered to IF Processing assemblies. Each such assembly accepts one signal from each band and contains additional amplification, switching to select one band's signal for further processing, gain equalization, and variable attenuation to adjust the power to a specified level. There are 4 IF processing assemblies to support the 4 IF signals supplied by the 2SB bands; all are identical (for interchangeability), even though two of them have no inputs from bands 1, 2, 8, 9, or 10 (see Table 2).[2] Normally all IF Processing assemblies select the same band, and this is also the band that is selected by the antenna pointing.

The cooled amplifiers and mixers are supported by active d.c. bias circuits, and most SIS mixers also have electromagnets for d.c. magnetic field bias. These supporting circuits are remotely adjustable via interfaces to the Monitor/Control subsystem.

All of the components described in this section are parts of the Front End Assembly, which includes the vacuum chamber with cryocooler and an attached frame that houses the room temperature electronics. One FE Assembly is located in the center of the receiver cabin of the antenna, with its upper end rigidly attached to the antenna structure near the focal plane, which is near the ceiling of the receiver cabin.

The FE Assembly also contains various local oscillator components. These are described separately, later in this document (section 3.1.3).

---

[1] At this writing, 2SB mixers are under development for bands 3, 6, and 7 and DSB mixers are planned for band 9. Assuming that these choices hold, it is reasonable to assume that the undeveloped bands 4 and 5 will use 2SB mixers and 8 and 10 will use DSB mixers.

[2] If all of the 2SB bands become SSB, which is possible, then all bands will have only 2 IF signals and only 2 IF Processing assemblies would be needed. The Downconverters would also be simplified.

2.3 Analog signal processing

At each antenna, the signals from the 4 IF channels of the FE Assembly are delivered on coaxial cables to the Downconverter Assemblies in the Analog Rack, which is located on one side of the receiver cabin. In a Downconverter, an IF signal is mixed with a high-side LO (8–14 GHz) to convert it to a 2–4 GHz "baseband" channel. There are 4 such converters per polarization, each with an independently tunable LO, and a flexible switching arrangement allows an IF channel to be connected to 1, 2, or all 4 of the converters. For a clear understanding of this, see the system block diagram [3]. If a single IF channel is connected to all 4 converters, and if the LO frequencies are selected properly, the 4 baseband channels can span the whole 4–12 GHz IF channel. On the other hand, for bands with 2SB front ends, each sideband can be directed to two converters. Many other possibilities are available.

The same LO feeds one downconverter of each polarization, so the two baseband channels of nominally orthogonal polarization are always tuned to the same sky frequency.

The Downconverter provides, for each IF signal, amplification, gain equalization (to compensate for the attenuation slope of the cables from the FE Assembly to the Analog Rack), and programmable attenuation in 0.5 dB steps. The attenuation is adjusted to keep the signal level within the linear range of the detectors (see below) and downconversion mixers over a wide range of system temperatures. There is no local ALC loop. (An ALC loop could be created in software via the Monitor/Control system, but at present this is considered unnecessary and undesirable.)

The 4 baseband signals from each polarization are delivered to a Baseband Processor Assembly, adjacent to the corresponding Downconverter Assembly in the Analog Rack. Here each signal is treated identically. It is amplified, bandpass filtered to limit it to the 2–4 GHz range, gain equalized, and passed through another programmable attenuator. The attenuator is set to optimize the level delivered to the Digitizer. Again, ALC could be provided via software, but none is implemented locally. The attenuators are not necessarily phase invariant, so it is recommended that their settings be kept constant over an instrumental phase calibration cycle.

The Downconverter also provides accurate square law detectors covering the full 4–12 GHz band for each IF channel, and the Baseband Processor provides accurate square law detectors for each 2–4 GHz baseband channel. These are available for continuum total power astronomy. For that purpose, all the detector outputs are filtered and digitized to 16 bits every 2 msec. The results are delivered to the Antenna Bus Master (ABM) computer over a dedicated CAN bus (see further discussion under Monitor And Control, below).

2.4 Digitization

The 8 baseband signals are transmitted on coaxial cables to the Digital Rack, located on the opposite side of the receiver cabin from the Analog Rack. From this point onwards, the channels are organized as "polarization pairs," so that the two channels of opposite polarization and the same sky frequency are always kept together.

Each 2 GHz wide baseband signal is then digitized **to 3 bits resolution (8 levels) at 4 Gsamp/sec. The chosen 3-bit code is a Gray code and also a sign-magnitude code. The phase of the digitization clock is programmable, and this implements the fine part of the delay tracking required for interferometry (the coarse part, to 1 sample resolution, is implemented digitally in the correlator). The digitization clock is common to all 8 baseband channels. This is further discussed under Local Oscillator, below.

The 3 bitstreams of 4 Gb/sec from each digitizer are immediately demultiplexed by 16 to form 48 bitstreams of 250 Mb/sec each. This is a more convenient rate for interfacing to the Transmitter Assembly of the Digital Transmission System. Cancellation of the 180 deg phase switching introduced in an LO (see section 4.3 below) is achieved here by synchronously inverting the sign bit.

2.5 Signal transmission

Each polarization pair of digitizers thus generates an aggregate data rate of 24 Gb/sec, and all 4 pairs (8 channels) generate 96 Gb/sec. All of this data is transmitted on a single optical fiber to the central Technical Building for correlation. A total of 12 optical wavelengths is used, all in the vicinity of 1550 nm, with 3 wavelengths for each signal pair.

The data are encoded for transmission in an ALMA-specific manner [4] that is part of the detailed design (and thus not described here). Each optical carrier is modulated at 10 Gb/s and conveys 8 Gb/s of sample data; the excess is used by the protocol. The main requirements are **that the samples be recovered

in a transparent manner (3 bits per channel at 4 Gb/s) without loss, in spite of the fact that not all bits of a channel can be transmitted at the same optical wavelength, and dispersion causes variation in transmission delay with wavelength; that the transmission delay at the output of the receiver be repeatable for each channel; and that the decoded bit error rate be less than 10^{-6}. This requires that the data be buffered with variable delay at the receiving end so that the bit timing of the separate optical channels can be aligned. Although the total delay need not be known *a priori*, it must be repeatable after all or some of the DTS components have had a power cycle, and also after other losses of signal, such as a temporarily disconnected optical cable.

Each of the 216 antenna stations has a separate fiber connection to the Technical Building where it terminates at a patch panel. Manually installed optical patch cables then connect 64 of them to further processing. The loss budget is such that an optical amplifier is required for each of the 64 active fibers (although there is at present some uncertainty about this).

Each of the selected 64 fibers is then demultiplexed by wavelength to separate the 12 optical carriers. Each group of 3 wavelengths, conveying one polarization pair of channels, is routed to a receiver and decoder that is located within one quadrant of the correlator. Each such quadrant contains the receivers and decoders for one channel pair of all 64 antennas, and the other channel pairs are received in the other three quadrants. There are no interconnections among quadrants (see Correlation, below).

The decoders demultiplex the 10 Gb/s raw data streams to lower rates and remove the extra bits of the protocol so as to recover the original sample bits. These are delivered to the first stage of the correlator's station electronics at a clock rate of 125 MHz, so that each 4 GSa/s x 3b channel is represented as 96 parallel bitstreams. The decoder uses information in the protocol to align these bitstreams in time and to ensure that the total delay from the digitizer to this point is repeatable.

Under some circumstances, it is desired to have two or more quadrants of the correlator process data from the *same* channel pair. To allow this, a 3-pole optical crossbar switch is inserted after each wavelength demultiplexer. The three wavelengths of one channel pair can be routed to any one, two, or four of the correlator quadrants; in the latter cases, the signal is split in a passive optical power divider and one or more of the other optical wavelengths is terminated. Each optical receiver works over a broad wavelength range, so it can accept whichever wavelength set is routed to it. (At present, elimination of these switches for cost saving is under consideration. Even if this is done, space is being provided in the Technical Building so that they can be added later if necessary.)

2.6 Correlation

The correlator is designed to process all 16 GHz of signal bandwidth transmitted from each antenna, forming auto-correlation and cross-correlation functions of all channels. In its main modes, the architecture is basically of the "FXF" type, where each input channel is first analyzed coarsely into subchannels by frequency (F); correlations are computed for corresponding subchannels at multiple "lags" that are one sample time apart in delay (X); and then the correlation functions are transformed to the frequency domain (F).

The detailed design of the correlator is complicated, and there is a large variety of possible operating modes. Only a summary and some examples are provided here. The main performance parameters, covering frequently used modes but not all possible modes, are given in Table 3. For additional details, see [8][9][10][23].

The correlator hardware is organized by quadrants, each of which normally handles one channel pair for all antennas. If fewer than 4 channel pairs are being used, the optical crossbar switches can be used to route the same data to more than one quadrant, allowing finer spectral resolution to be achieved by having each quadrant compute different lags for the same signals. No interconnections are ever required between one quadrant and another. This architecture allows straightforward expansion of the total bandwidth by adding more "quadrants," but it forces the total number of antennas to be fixed once the structure of one quadrant is established.

Within a quadrant, the processing is separated into Station Electronics and Correlation Electronics.

Table 3: Correlator – Main Specifications
(Values apply in filterbank mode; corresponding bypass mode value in {...} if different.)

Number of antennas	64
Input (baseband) channels	8 (4 polarization pairs)
Quantization	8 levels into filters
	4 levels into cross-multipliers (16 at reduced bandwidth)
Bandwidth per BB channel	2 GHz maximum
	31.25 MHz minimum
Spectral resolution	All 4 baseband channel pairs in use:
(bypass) [filterbank] mode	4096 {128} points per channel w/o cross-pol
	2048 {64} points per channel with cross-pol
	2x above if 2 channel pairs used
	4x above if 1 channel pair used
Output rate	minimum integrating time 512 {16} msec
	(autocorrelations only: 4 {1} msec)
	maximum integrating time 65 sec
Subarrays	quadrants are independently configurable
	up to 6 subarrays: double buffered
	7 to 16 subarrays: single buffered

2.6.1 Station Electronics

The Station Electronics of one quadrant provides processing for one channel pair from each of 64 antennas. The signals for an antenna are received on three optical fibers and are demodulated, demultiplexed, and decoded by a Data Receiver module to recover the 3-bit sample streams from the two channels; all 6 bits are aligned in time and have a repeatable propagation delay (see Signal Transmission, above). The 4 GSa/sec streams are delivered at 125 Mb/sec in 32-sample words, or 96 parallel bit streams per channel.

The processing for one antenna is organized into a pair of Filter modules and a single Station Buffer module. Together they handle the delay tracking and provide many of the operating modes.

2.6.1.1 Filtering.

Each channel may then be processed by a digital filter bank, where it is analyzed into 32 subchannels [23]. Each subchannel includes digital downconversion to baseband followed by low-pass filtering, with the output sampling rate decimated by 32 to 125 MSa/sec. The filter bandwidths are programmable, but normally they are slightly less than the Nyquist bandwidth of 62.5 MHz and the center frequencies are chosen to produce slightly overlapping coverage of the channel. By discarding the edges of each subchannel in post correlation processing (after transforming to the frequency domain), this allows synthesizing continuous coverage of slightly less than the full 2 GHz channel with negligible aliasing among subchannels. The filter bank may also be bypassed, so that the full channel bandwidth is processed at once; this produces coarser spectral resolution but allows shorter integration times, as explained in section 2.6.3 below. We call the two cases "filter bank mode" and "bypass mode," respectively. In both modes, the samples are re-quantized to 4 levels (2b) or 16 levels (4b) for all further processing. With 4b samples, only half of the subchannels are transmitted downstream, either 16 subchannels from each channel or 32 subchannels from one channel only.

2.6.1.2 Buffering.

The filtered, decimated, and re-quantized sample streams are then buffered in a memory and re-organized. The buffer provides the variable coarse delay needed for interferometric tracking, and it also allows for a wide variety of processing modes. Data are always buffered in blocks of 4 MSa or 1 msec.

In bypass mode, the data at the buffer output are organized as 32 sample streams where each stream contains time-contiguous samples. The first stream contains the first 1/32 msec of the samples at 1/32 of the original sample rate, the second contains the next 1/32 msec, etc. During 1.0 msec, all 32 streams are processed in parallel so as to handle the original 1.0 msec of 4 GSa/sec data. Each stream goes to a separate "plane" of correlation electronics, so each plane processes a separate time slice of data. The correlation results from all time slices are later summed.

In filter bank mode, each subchannel's samples are effectively buffered separately and passed to one or more separate planes. If all 32 subchannels of both polarization channels are used, then each goes to

a different plane. If half or fewer subchannels are processed, then each can be sent to two or more planes and the delays can be set so that each plane computes a different range of lags; in this way, finer spectral resolution can be achieved. In addition, if a multiple of 4 planes is used to process one subchannel, then 4b×4b multiplication can be achieved. The latter feature requires that the filters be configured to provide 4b samples.

2.6.1.3 Delay Tracking.

Delay tracking is implemented separately for each channel (but in common for all subchannels in filter bank mode), in two stages. The bulk of the delay is produced by the Station Buffer, using the same memory that is used for re-ordering the sample streams. Data is always read from this memory in words of 64 samples, and the delay is adjusted by incrementing or decrementing the read address, so the resolution is 64 samples. In bypass mode, the sample interval is 250 psec (4 GSa/sec) so the resolution corresponds to 16 nsec. In filter bank mode (125 MSa/sec), the sample interval is 8 nsec so the resolution is 512 nsec.

In either mode, finer resolution is achieved at the input of the filter bank, where the sample interval is always 250 psec (4 GSa/sec). Here the data is processed in 32-sample words, so a delay of one clock provides a resolution of 32 samples or 8 nsec. A shift register provides a range of 0 to 63 clocks. In addition, a buffer with appropriate shift logic allows delay adjustment with a resolution of 1 sample and a range of 0 to 63 samples. Together these provide a range of 0 to 2047 samples or 0 to 511.75 nsec, sufficient to span the worst-case resolution of the bulk delay, with a resolution of 1 fast sample or 250 psec.

Fine and coarse delay tracking must be synchronized. When the fine delay reaches 0 or 2047 samples (in filter bank mode, 0 or 63 samples in bypass mode), the next 1-sample change is implemented by a change of 64 decimated samples (one address) in the coarse delay and 2047 samples (63 samples in bypass mode) in the opposite direction in the fine delay. (In single subchannel mode the situation is similar but the fine delay range is only 0 to 63 samples.) These changes should be applied at the same sample, which means that the fine change must occur earlier by the amount of the bulk delay plus the delay through the filter (although the filter delay is generally negligible). A step in coarse delay is allowed to occur only between the data blocks in the buffer, or every 1.0 msec. For sidereal rate tracking, this is at most every 5.5 sec (full bandwidth mode), and the fine delay steps occur at most every 86 msec. (As mentioned in Section 2.4 and further explained later in Section 3.1.5, sub-sample delay resolution is obtained by varying the phase of the sampling clock at the antenna. Synchronization of these is discussed in Section 4.5.)

2.6.1.4 Phased Array Output.

There is a **requirement to sum the corresponding samples of a selected subset of antennas into a single sample stream, and to provide this stream to a recording system for VLBI. Implementation of this feature is currently pending, with details TBD. Further work is not funded within the construction budget, nor is acquisition of any VLBI recording system. Further discussion is given in section 4.6.

2.6.2 Correlation Electronics

Sample streams from the Station Electronics are delivered to a separate set of racks containing the Correlation Electronics of one quadrant. Again, all processing is at a clock rate of 125 MHz and is organized into 32 "planes" that operate in parallel so as to handle an aggregate sample rate of 4 GSa/sec.

Each plane receives data streams from all 64 antennas, where each data stream carries 1/32 of the data of a channel pair. Depending on the filter mode and buffer mode, a plane might be processing a time slice or a frequency subchannel. An important feature of the design is that each data stream has only one source (an output of the buffer) and one destination (an input of one correlation plane), greatly simplifying the interconnection topology compared with alternative architectures. Each plane is a 64×64 matrix of correlator cells, and each data stream feeds both a row and a column of the matrix. Within each cell, four modes are possible, where x and y are the two channels of the polarization pair:

- 256-lag correlator for channel x only (xx);
- 256-lag correlator for channel y only (yy);
- two 128-lag correlators: xx and yy;
- four 64-lag correlators: xx, yy, xy, and yx.

All cells of all 32 planes are normally set to the same mode. The last mode is used if all polarization parameters of the signal are needed.

The correlation cells also support 2x oversampling. If the filters are set to a bandwidth of 31.25 MHz, then the (fixed) 125 MHz sampling rate is twice the Nyquist rate. In that case, the correlation cells can be

configured to compute lags spaced two samples apart; no aliasing occurs, and the frequency resolution is halved.

Along the matrix diagonal, autocorrelation functions are obtained for each antenna. The same antenna pair appears in both the above-diagonal and below-diagonal cells, where one computes cross-correlation functions for positive lags and the other for negative lags. In this way, each plane will support 256, 128, or 64 spectral frequency channels in the auto- or cross-power spectra that are obtained by Fourier transformation after correlation. In bypass mode, these spectral channels are spread across the full baseband channel (2 GHz). In filter bank mode, they are spread across each subchannel.

Larger numbers of spectral channels across each subchannel are obtained when not all subchannels are used, so that more than one plane can be used for each; then different planes are used to compute different lags using the same samples. In addition, multiples of 4 planes can be configured to provide 4b×4b multiplication. Discussion of all possibilities is beyond the scope of this report.

2.6.3 Long-Term Accumulator (LTA)

The correlation results from all planes are read into the LTA, the next stage of the Correlation Electronics, where they can be accumulated for longer times. If only the autocorrelation results are desired, they can be read and passed through the LTA every 1.0 msec. Otherwise, results are read every 16.0 msec and further integrated in the LTA. Each LTA result is a 32 bit value. To guarantee that overflow never occurs, the maximum integrating time is limited to 65 sec.

In bypass mode, where separate correlation planes are computing different time slices of the same sample streams, the LTA first sums those planes to obtain the correlation for all time slices.

Binning is supported by allowing the summation to be switched to any of 4 accumulation bins on any 16 msec boundary. Normally this is done synchronously with state changes in the upstream signal processing, such as LO phase switching or frequency switching; subreflector nutation; or on/off antenna pointing.

If the 64 antennas are partitioned into subarrays, then the antenna pairs of each subarray may have different modes in the LTA, including different integrating times and different bin-switching times. Up to 6 such subarrays are supported, and each quadrant may be partitioned independently. Results for inter-subarray antenna pairs are undefined.

2.6.4 Correlator Data Processor (CDP)

The LTA accumulations are transferred to an array of 16 PCs that perform the next stage of processing (plus one master PC, making 17 PCs in the whole CDP). Four of these computers are connected to the LTA of each of the four quadrants. Each computer is capable of handling up to 64 Mbyte/sec of data, or 256 Mbyte/sec/quadrant. However, at the minimum integration time of 16 msec the LTA produces 8.3885 Gbyte/sec/quadrant. To accommodate this limit, the user must select a combination of longer integrating time and coarser spectral resolution, or he may choose to transfer only a subset of the results, either not all lags or not all antenna pairs. In filter bank mode, all lags of each subchannel must be processed in order to resolve and eliminate aliased frequencies; but fewer than all 32 subchannels can be selected.

The CDP performs the following processing:

- time tagging
- quantization correction
- Fourier transformation from lag to frequency, including windowing
- in filter bank mode, assembling a continuous broadband spectrum from all available subchannels, discarding frequencies affected by aliasing
- correction of residual delay error (applying phase slope)
- applying complex gain corrections (e.g., from WVR measurements)
- additional integration
- formatting for transmission to archive

2.7 Archiving

Formatted data from the CDP is transmitted as UDP or TCP/IP packets on the 10 Gb/s ethernet link from the Technical Building at the ALMA Operations Site to the Operations Support Facility about 30 km away. There it is received by the Archive Processor and written to the archive disk array. Writing to the archive is possible at a maximum rate of approximately **40 Mbyte/sec, or about 4% of the CDP throughput or 0.12% of the maximum rate of generation of correlation products. The user must further restrict the integrating time or spectral resolution, or select a subset of the data for recording.

3 SUPPORT SYSTEMS

3.1 Local oscillators

The local oscillator subsystem is responsible for establishing all of the time synchronization in the array, on scales ranging from 48 msec (20.8 Hz) to $\ll$ 1 psec (1 THz). It is also responsible for generating the sinusoidal signals necessary for converting the received signals from RF to IF to baseband, and for tuning these as required to establish the desired sky frequency and interferometer phase (including fringe tracking, phase switching, and other interferometer-specific features). The latter are more properly known as "local oscillator" signals, but the subsystem must also supply various coherent references to other devices so as to achieve synchronization and accurate timing. These include digitizers, computers, and the correlator. It does this by distributing periodic reference signals derived from a common master oscillator.

The LO subsystem also forms part of the array master clock (discussed in Section 4.1 "Timing and clocks") in cooperation with a computer of the monitor-control subsystem. It does this by providing an interface to an external time scale (currently GPS) and by measuring the difference between external time and array time. Measures of time larger than 48 msec are obtained in the MC system by integration.

Table 4: Local Oscillators – Main Specifications

Item	*Specification*	*Goal (if different)*
Frequency Range		
1st LO	27.3 to 938 GHz	
2nd LO	8 to 14 GHz	
Output Power		
1st LO	Band dependent, see Table 4B	
2nd LO	+15dBm to each of 2 converters	
Sideband Noise, 1st LO	10 K/μW	3 K/μW
Amplitude Stability		
1st LO	.03% < 1 sec; 3% between adjustments	
2nd LO	.13% in < 1 sec	
Phase errors	see Table 4B	
Tuning step size, maximum		
On the sky	200 MHz	
SIS mixer 1st LO	500 MHz	
Indendently tuned subarrays	4	
Time for frequency change, maximum		
Within .03%	10 msec	1 msec
Otherwise	1.5 sec	0.1 sec
Repeatability	Phase-unambiguous synthesis	
	Stability specs apply across frequency changes	
General	All-electronic tuning (no mechanical adjustments)	

3.1.1 LO Specifications

The main specifications for the LO subsystem are summarized in Table 4. Many of the critical specifications are controversial and are under review at the time of this writing. Among these are the first LO power, the amplitude stability, and the phase variation. In all these cases, both the required value and the technically feasible value are uncertain.

3.1.1.1 First LO Power

It has always been difficult and expensive to develop significant amounts of sinusoidal power at millimeter and submillimeter wavelengths. Low noise, electrically tunable oscillators are not available above about 150 GHz, and those for 50 to 150 GHz (typically Gunn diode oscillators) have insufficient bandwidth for the ALMA application. Above 150 GHz, it is almost always necessary to use frequency multipliers driven by a lower-frequency oscillator; these typically have low efficiency, delivering only a few microwatts above 300 GHz. Fortunately, the power requirements of SIS mixers are small. This makes it possible, in many cases, to use weak coupling of the LO to the mixer, where 90 to 99% of the LO power is discarded.

The *required* power is strongly dependent on the mixer design, including whether weak or strong coupling is used, whether the mixer is simple (single-ended) or compound (balanced and/or sideband-separating), and the number of junctions and their impedance. (Strong coupling is possible by using frequency-selective couplers (e.g., Martin-Pupplet diplexers) or by using balanced mixers, but none of the present ALMA designs use these techniques.) The *available* power depends on the technology of oscillators, amplifiers, and multipliers, including design choices driven by considerations of cost and reliability. Thus the overall result is a compromise.

Taking these considerations into account results in the power requirements given in Table 4A. For further discussion, see [21].

Table 4A: First LO Power

Band	Freq Range	Junctions per mixer	Mixers per pol.	Minimum per pol.	Goal
	GHz			μW	μW
1	27.3–33	N/A	1	1000	5000
2	72–94	N/A	1	1000	5000
3	96–104	4	2	25	40
4	137–151	4	2	45	70
5	175–199	4	2	80	120
6	223–263	4	2	140	200
7	283–362	2	2	65	100
8	397–488	2	1	15	30
9	614–708	1	1	35	70
10	799–938	1	1	55	100

3.1.1.2 Phase Errors

We separate the overall phase accuracy requirements of the telescope by time scale. Variations of phase during one visibility measurement (integration) cause a reduction in signal and loss of sensitivity called "incoherence." Variations on longer time scales, especially the intervals on which corrections or calibrations are possible, cause systematic errors in determining the absolute visibility (beyond the random errors that are attributable to the system noise). It is tentatively assumed that the amount of incoherence can be known and thus that the signal reduction can be corrected, so that it causes no systematic error but just a reduction in signal-to-noise ratio.

Both the coherence and the systematic error may be dominated by the variable delay through the atmosphere rather than by the instrument. The magnitude of atmospheric effects is highly dependent on weather conditions, season, and time of day. They are to some extent correctable from sounding measurements based on auxiliary instruments (see Section 3.3) and on observations of astronomical calibrators. It is a goal of the instrument design that its contributions to the coherence and the systematic error should be well below those of the atmosphere under good observing conditions and after all available corrections have been applied.

To make this quantitative is difficult. Early in the ALMA project (2000-Feb), goals of achieving 90% coherence and 0.1 radian accuracy at 950 GHz were set [21] without serious consideration of their feasibility. The mean square errors were then allocated 50% to the atmosphere, 33% to the electronics, and the balance to the antenna structures. The electronics contribution was assumed to be mostly due to the LO.

Examination of site testing data and the use of soundings with water vapor radiometers (see also Section 3.3.4) at other telescopes allows more realistic requirements to be set, as shown in Table 4B. Here the requirements are expressed in time units, representing fluctuations or uncertainty in signal delay; most effects are nearly independent of observing frequency when expressed this way.

The requirements are based on the principle that the instrument should contribute less error than the atmosphere 95% of the time. We are thus interested in the atmospheric conditions at the 5th percentile. The atmospheric statistics at the ALMA site are well determined from more than six years of test data, as plotted in [22]. It is necessary to consider the joint distribution of delay fluctuations and water vapor content, since high frequency observing will not be done when the water vapor is high even if the delay fluctuations are

11

Table 4B: Phase Errors

			rms in fsec		
	Basis	Total	Atmos	Struct	Elec
Coherence					
Preliminary [21]	90% at 950 GHz	54.5	38.5	22.2	31.4
Uncorrected atmosphere	5th percentile, 45d		143		
	25th percentile, 45d		273		
Corrected via fast switching	5th percentile, 45d		87		
Corrected via WVRs	5th percentile, 45		75		
Current requirements† **	WVR resid. at 5th %ile	106	75	38	65
Accuracy					
Preliminary [21]	0.1 rad at 950 GHz	11.0	8.4	4.8	6.9
Atmos. with fast switching	5th percentile, 45d		25		
Antenna spec	non-repeatable			50	
Current requirements† **	f.s. resid. at 5th %ile	35	25	13	22

†Only the values for structure and electronics are requirements on the system performance. Those for the atmosphere and for the total error are estimates.

low. We find that a representative point on the 5th percentile curve has zenith values of 170 fsec rms and 0.69 mm for the delay fluctuation on a 300 m baseline (in a 10 min interval) and the condensed water vapor depth, respectively. These values are adjusted to 45d elevation (scaling by square root of air mass and by air mass, respectively) and $1/\sqrt{2}$ of the delay fluctuation is attributed to each antenna, giving 143 fsec and 0.96 mm as the conditions that correspond to the instrumental requirements.

After correction by fast switching, it is found via careful simulations [24] that the residual short-term fluctuations, affecting coherence, will be 123 fsec per baseline (86.9 fsec per antenna) and the residual calibration error will be 35.4 fsec per baseline (25.0 fsec per antenna). (See [24] for discussion of the simulation methods and assumptions. Different assumptions could lead to different estimates of the residuals, but the simulations are robust and quite insensitive to assumptions, so large variations from these values are unlikely.)

After correction using water vapor radiometer soundings, using the present estimates for WVR performance, a short-term residual of 75 fsec is predicted under the same conditions (see section 3.3.4). WVR soundings are not expected to be helpful with absolute calibration and visibility accuracy.

Considering all these results, we adopt 75 fsec as the instrumental short-term stability requirement (equal to the WVR residual), and 25 fsec as the instrumental long-term drift requirement. These apply to each antenna on the assumption that the antennas behave independently. The short-term requirement is the rms fluctuation over intervals of 10 sec or less, corresponding to the individual integrations on one source. The drift requirement applies to the phase change over a complete calibration cycle of at least 300 sec; it is desirable for it to be met for longer times, up to 1000 sec.

We allocate the instrumental errors between the structure and the electronics somewhat arbitrarily, with 25% of the squared error allowed for the structure and 75% for the electronics; these are the values shown in Table 4B. The structural allowance is then somewhat smaller than the 50 fsec allowed in the specification for the prototype antennas, but it is expected to be achievable.

Allocation of these phase error limits among subsystems within the electronics is currently pending, but the majority will be allocated to the first LO. For the purpose of establishing specifications, the limits should first be converted to phase at the observing frequency; then the coherence limit may be taken to refer to the rms phase deviation from its average value over any interval of 10 sec, and the accuracy limit may be taken to refer to the rms difference of 10-sec average phases measured 10 sec to 1000 sec apart.

3.1.1.3 Frequency ranges and tuning resolution

The first LO range for each band is immediately determined by the RF range (Table 1) and IF range (Table 2). At the RF band edges, it is acceptable to convert only to the IF band edges; and for the SIS bands (both 2SB and DSB cases) either sideband may be used to reach an RF band edge. Thus, the LO range for most SIS bands falls inside the RF range by 12 GHz on each end.
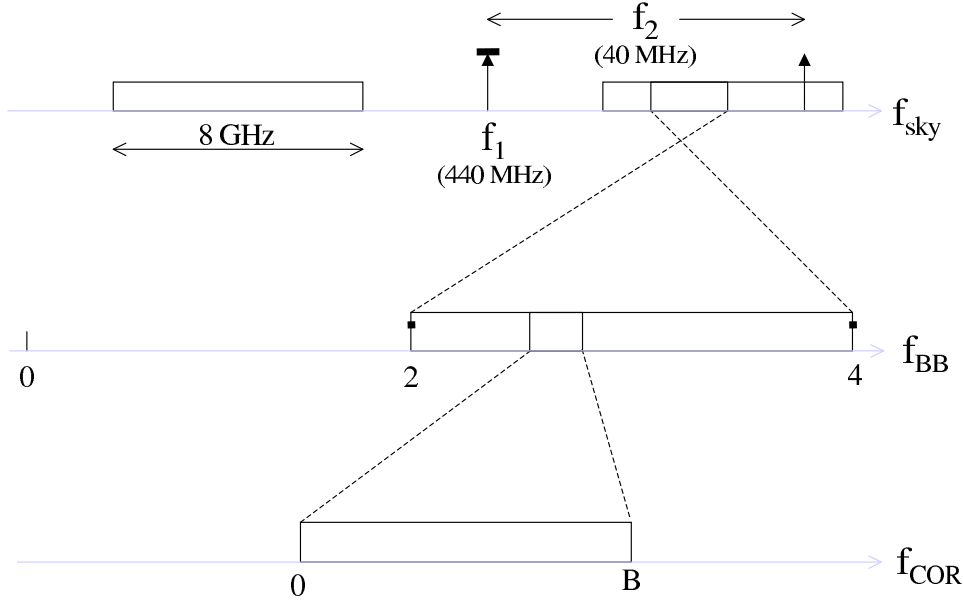
Figure 1: Tuning resolution achieved by the current LO design. From top to bottom, the selected observing band is shown as a function of sky frequency, baseband frequency, and correlator filter output frequency. The first LO frequency is f_1 and the second LO frequency is f_2. The maximum gap between available tunings for each LO is shown in parentheses, and is also depicted to scale by the small horizontal bars (above the arrow for f_1 and at 2 and 4 GHz on the f_{BB} plot for f_2). The correlator filter bandwidth B is freely tunable across the baseband channel via a digital LO [23].

The second LO range is similarly set by its need to convert from the 4–12 GHz IF band to the 2–4 GHz baseband, along with the design decision to operate on the high frequency side (LSB conversion).

The maximum tuning step is intended to ensure that **any sky frequency can be placed at the center of a baseband channel within $\pm5\%$ (±100 MHz) of the channel width. A separate requirement is placed on the first LO for the SIS mixer bands so as to ensure that **a line of interest can be placed near the middle of a selected sideband. The actual design substantially exceeds these requirements, as shown in Figure 1.

3.1.1.4 Phase switching and fringe tracking

These issues are covered in detail in [13] and further discussed under Methods and Algorithms, below. Here we mention only the implications for LO hardware design.

The first LO is **required to support 180d phase switching at intervals of 125 μsec so as to suppress spurious signals and d.c. offsets between the first mixer and the Digitizer. The 97% transition time (to 0.1 radian) for each phase change must be $< 2\,\mu$sec in order to avoid noticeable loss of SNR. It is desirable, but not required, that the second LO also support this feature. The present design allows either or both LOs to include such phase switching, except that the second LO transition time may exceed 2 μsec.

To allow sideband *separation* after correlation, 90d phase switching would be needed in the first LO. However, this feature is not required for ALMA[4]. Sideband *suppression* will be supported by offsetting the first and second LO frequencies by equal amounts from their nominal values (so that the offsets cancel for one sideband and add for the other), and by using a different offset at each antenna. These offsets should be multiples of $1/(.016\text{ sec}) = 62.5$ Hz so as to be effective in one 16 msec correlator integration cycle.

Fringe tracking can be provided in either LO. The present design provides hardware to support fringe tracking in both places, so the choice can be made in software. For bands using DSB mixers, some special considerations apply as discussed in [15].

---

[4] Phase switching in 90d steps could be provided, and the necessary synchronous switching of correlation products to separate integration bins could be provided at 16 msec intervals, resulting in a minimum cycle time of 1024 msec. Thus, sideband separation could be effective only over integration times that are integral multiples of 1024 msec.

All of these features are achieved through direct digital synthesizers that provide an offset reference to a phase locked loop.

3.1.1.5 Frequency changes, repeatability, and phase ambiguity

In order to support the "fast phase calibration" mode of observing, which is expected to allow removal of atmospheric phase fluctuations via observations of a nearby calibrator at a different frequency from the target source, ** the phase accuracy specifications for all LO signals apply across frequency changes. This means that after a change of observing frequency (including but not limited to switching to a different band) and subsequent return to the original frequency, the phases of all LOs applied to the signal path (first, second, and third [sampling clock]) must have the same phase (within the allowed phase noise and drift) as they would have had if the frequency change had not occurred.

Implementation of the above requirement is accomplished by **ensuring that all frequency synthesis is *unambiguous in phase.* This is more stringent than the requirement, but it has technical benefits. It means that phase is repeatable not only across a commanded frequency change, but also across a power cycle of one or more components, disconnection and reconnection of a cable, and various similar events that occur in practice. Generally, unambiguous synthesis requires that the synthesized frequency be an integer multiple of some reference frequency; i.e., the synthesis should not involve frequency division. This is not straightforward with nearly-continuous tuning, as is required for fringe tracking in the first or second LO and for delay tracking in the sampler clock. However, by making use of the properties of DDSs, along with the ability of the monitor/control system to deliver commands synchronously with the lowest frequency distributed reference (21 Hz), special techniques allow unambiguous synthesis even in this case (see Appendix A of [13]).

**Any frequency change must be accomplished in 1.5 sec, including phase settling to within the drift limit. This is also for support of the fast phase calibration mode, and is tied to the speed of antenna motion. Large frequency changes are also needed with no antenna motion, so as to determine the instrumental phase difference between two setups while observing one calibrator; for this reason, it is a goal (but not presently a requirement) to accomplish an arbitrary frequency change in substantially less than 1.5 sec. In addition, in order to support frequency-switched total-power spectroscopy, small frequency changes (within one band) must be accomplished much more rapidly. **For a change of .03% or less (maximum 285 MHz in band 10), the change must occur within 10 msec.

3.1.2 Reference generation and distribution

Nearly all time-dependent functions in the array must be coherent with a single master oscillator from which reference signals are derived and distributed. This is a hydrogen maser.

The master oscillator drives the Central Reference Generator (CRG), which synthesizes a set of fixed-frequency signals covering the range 21 Hz through 2.0 GHz (see the block diagram [3]). Some of these signals are used at the central building, and some are distributed to the antennas. See "Timing and Clocks," Section 4.1, for a discussion of the choices of reference frequencies.

A mm-wavelength reference is then synthesized for the first LO. This is the only variable-frequency signal that is distributed to the antennas. The process uses a microwave synthesizer to produce 8.62-11.08 GHz in 5 MHz steps (using primarily the 2 GHz and 5 MHz references from the CRG), followed by synthesis of 27-142 GHz as the difference between two laser-generated optical frequencies. The lasers are designated as "master" and "slave," with one master required for the array and one slave for each separately-tunable subarray. The two-frequency optical signal is sent to each antenna on a single-mode fiber. For each antenna separately, the optical signal passes through a line-length stabilizer based on two-way optical phase measurement of the master laser signal.

To support operation of more than one subarray with independent frequency selection, a separate laser synthesizer is provided for each, along with appropriate switching. The master laser may still be common to all antennas, but each subarray needs a microwave synthesizer, slave laser, and phase lock circuitry. **Four of these are planned and should be adequate. There is no technical barrier to providing more (only extra cost).

Meanwhile, the 21 Hz, 25 MHz, and 2 GHz references are transmitted together on a third laser carrier on the same fiber. The 21 Hz is transmitted by reversing the phase of the 25 MHz every 600,000 cycles (24 msec); this signal is added to the 2 GHz sinusoid, and the composite signal is used to intensity-modulate a

laser. This optical carrier is sufficiently near the master laser's wavelength that the line length stabilizer is effective for it, in spite of dispersion in the fiber.

At each antenna, the Reference Receiver assembly demodulates the 21 Hz, 25 MHz, and 2 GHz signals from their carrier; frequency-shifts the master laser carrier and transmits it back on the same fiber to the center for line length stabilization; produces an additional reference at 125 MHz; and distributes all the fixed references to various devices in the receiver cabin.

3.1.3 First LO

Table 5 lists various frequencies involved generating the first LO signals.

Table 5
ALMA FIRST LO MULTIPLICATION SCHEME

Band	RF min	RF max	IF min	IF max	LO min	LO max	Ncold	Ref min	Ref max	Nos max	LO resolution	N2	N1	Driver min	Driver max	VCO min	VCO max
	GHz	GHz	GHz	GHz	GHz	GHz		GHz	GHz		MHz			GHz	GHz	GHz	GHz
1	31.3	45.0	4.0	12.0	27.3	33.0	1	27.3	33.0	3	15.0	1	1	27.3	33.0	27.3	33.0
2	67.0	90.0	4.0	12.0	79.0	94.0	1	79.0	94.0	9	45.0	3	2	79.0	94.0	13.2	15.7
3*	84.0	116.0	4.0	8.0	92.0	108.0	1	92.0	108.0	11	55.0	3	2	92.0	108.0	15.3	18.0
4	125.0	163.0	4.0	12.0	137.0	151.0	2	68.5	75.5	7	70.0	3	1	68.5	75.5	22.8	25.2
5	163.0	211.0	4.0	12.0	175.0	199.0	2	87.5	99.5	11	110.0	3	2	87.5	99.5	14.6	16.6
6	211.0	275.0	4.0	12.0	223.0	263.0	3	74.3	87.7	9	135.0	3	2	74.3	87.7	12.4	14.6
7	275.0	370.0	4.0	8.0	283.0	362.0	3	94.3	120.7	11	165.0	3	2	94.3	120.7	15.7	20.1
8	385.0	500.0	4.0	12.0	397.0	488.0	5	79.4	97.6	9	225.0	3	2	79.4	97.6	13.2	16.3
9	602.0	720.0	4.0	12.0	614.0	708.0	5	122.8	141.6	13	325.0	3	3	122.8	141.6	13.6	15.7
10	787.0	950.0	4.0	12.0	799.0	938.0	9	88.8	104.2	11	495.0	3	2	88.8	104.2	14.8	17.4

Definitions

RF	signal frequency on the sky
LO	LO frequency at first mixer
Ncold	cold multiplication factor
N1,N2	warm multiplication factors
Nos	optical synthesizer multiplication factor
Ref	frequency of reference signal sent from Technical Building to antenna
VCO	frequency of YIG tuned oscillator used as VCO of antenna PLL

"Resolution" is step size at final LO frequency and assumes that drive to optical synthesizer has 5 MHz resolution.
* Band 3 LO range supports 4-8 GHz IF as a contingency, although the planned IF range is 4-12 GHz.

Generation of the mm/sub-mm wavelength first LO is based on phase-locking a local VCO to the highest-feasible reference frequency that can be distributed from the center. The two-laser reference signal is recovered by photomixing, and photodetector technology now limits this to the frequencies< 200 GHz. The phase-locked signal is then multiplied to higher frequencies, as needed, in diode multiplier assemblies; these are operated at cryogenic temperatures ($\sim$80K) to maximize efficiency. As shown in Table 5, reference frequencies in the range 68.5 to 141.3 GHz are needed to cover bands 2 through 10 with cold multiplication factors of 1 (no multiplier), 2, 3, 5, or 9. The band 1 reference and VCO are at 27.3 to 33 GHz.

The local VCO is implemented as a YIG-tuned microwave oscillator and amplifier/multiplier chain. The YTO frequencies and hence the required multiplication factor are chosen for engineering reasons, but the minimum oscillator frequency is kept above 12 GHz to avoid the possibility of spurious sidetones of the LO falling in the RF band that extends $\pm$12 GHz from the LO. Details of the frequencies involved are listed in Table 5.

The locked VCO is offset from the reference by approximately 30 MHz from a direct digital synthesizer [DDS, part of the Fine Tuning Synthesizer Assembly (FTS)] through which phase tracking (fringe rotation) and phase switching are implemented. To allow phase switching at intervals as short as 125 μsec, the settling time should be minimized (see section 3.1.1.4) so that the PLL bandwidth should be relatively large.

3.1.4 Second LO

The second conversion (IF to baseband) requires LOs at 8–14 GHz. Four second LO synthesizers are provided to allow independent tuning of each polarization pair of 2 GHz baseband channels.

The design locks a YTO to harmonics of 125 MHz, offset by a DDS that is tunable from approximately 20 to 43 MHz. This provides a very flexible arrangement with gaps in frequency no larger than 40 MHz, far exceeding the sky tunability requirement of 200 MHz. Any sky frequency can be placed within $\pm$20 MHz of a selected baseband frequency, so that it is possible to avoid the Nyquist band edges even at the narrowest

correlator bandwidth (31.25 MHz). **The DDS is required to be part of an FTS Assembly identical to the one used in the First LO.

3.1.5 Digitization clock (Third LO)

As explained in connection with the main signal flow, each 2–4 GHz baseband channel is digitized at a rate of 4 Gsamples/sec. The sampling clock may be regarded as a third LO whose nominal frequency is fixed. The process is complicated by the fact that the phase of the sampling clock is used to implement fine adjustment of the signal delay.

Most of the delay tracking for interferometry is implemented digitally in the correlator, but only to a resolution of one sample (250 psec). Finer resolution is needed to minimize loss of SNR due to delay tracking error [12]. If the resolution is 1/8 sample, the worst-case loss is 5% at the 4 GHz edge of the channel. **The sampling clock phase is therefore required to be adjustable to an accuracy of 1/8 sample or less. If this is achieved with discrete steps of $1/L$ sample ($L \geq 8$), then steps occur at a minimum interval of $(86 \text{ msec})/L$ for sidereal tracking. Every L such steps, the correlator inserts a full-sample delay change and the phase of the sampling clock must simultaneously change by nearly a full sample. Here "simultaneously" means that both delay changes (correlator and sampler) occur at the same data sample, so they are actually separated in time by the propagation delay, including the (variable) bulk delay within the correlator. **It is sufficient that synchronization of these events be accurate to 10 μsec or less; then fewer than .01% of the samples will have slightly incorrect delay during the change. For a more detailed discussion, see [12].

Before cross-correlation, the variable-phase sample streams from the different antennas must be re-captured onto a common clock. Due to the high sampling rate, this is accomplished after demultiplexing to a clock rate of 250 MHz, where the demultiplexing clock is also of variable phase. Blocks of 16 samples are then re-captured onto a fixed-phase 250 MHz clock; this is done at the antenna before transmission to the Technical Building. Details of the implementation are given in [11].

We have 8 separate signal channels, each with its own digitizer, from each antenna. The total delays for these channels are nearly the same, but there can be small yet significant differences because of different cable lengths, analog electronic components, and dispersion in the optical fiber. Nevertheless, **it is sufficient to drive all 8 digitizers with the same variable-phase clock. The delay differences can be compensated digitally with a resolution of 1 sample, and the fine delay can track the average of all channels. In this way, the maximum delay error for any one channel is 0.5 sample. It is important that this error is *constant* in time. It can then be corrected in post processing by adding a phase correction to each spectral channel of $2\pi f \, d\tau$, where f is the spectral channel's offset at baseband from the 4 GHz sampling frequency and $d\tau$ is the (constant) delay error. The spectral resolution should always be less than about 100 MHz to ensure negligible loss of SNR.

3.2 Control and monitoring

The Monitor/Control (MC) subsystem consists of a network of real-time computers and their connections to the numerous electronic and electro-mechanical devices that make up the ALMA telescope, along with the software that runs on them. Most devices are connected to the computer network by a specialization of the industry-standard CAN bus [16] known as the ALMA Monitor/Control Bus (AMB) [18].

The computers are in various locations and they communicate with each other via high speed Ethernet links [17]. They are:

- Antenna Bus Master (ABM), one at each antenna. Each ABM provides communication to and from every device at the antenna over an AMB.
- Array Real Time Machine (ARTM), at the central Technical Building. This machine provides communication with all devices in the Technical Building except the correlator, mostly over an AMB, but in special cases over other standard ports. It also implements the array master clock by counting cycles of the 48-msec-period (21 Hz) reference from the Central Reference Generator and it can, when necessary, command the CRG to adjust the 48 msec phase to coincide with external time from a GPS receiver.
- Correlator Control Computer (CCC), at the Technical Building. This machine communicates with each processing unit of the correlator, including each board of the Station Electronics and of the Correlation Electronics. It does so over an internal multi-drop CAN bus that is similar to, but not identical to, the AMB.

- Array Control Computer (ACC), at the Operations Support Facility (OSF), about 30 km (46 km by road) from the Technical Building. This machine is the master controller of the entire telescope, with the others operating as slaves to it. It also has other duties, such as scheduling and calibration, that need not run in real time and are not logically part of the MC subsystem.
- Archive Processor, at the OSF. This machine was discussed earlier as part of the main signal path, because it writes the correlation results to the archive. But it also writes to the archive monitor data gathered from devices by the ABMs, ARTM, and CCC, and in this role is part of the MC subsystem.

All of these machines (or at least the MC tasks of those that also have other functions) must operate in real time and their tasks must meet strict timing deadlines. However, only the ABMs, ARTM, and CCC use real-time operating systems, and the Ethernet communication among them is not deterministic. The design assumes that the necessary real-time performance will be achieved by having sufficient excess capacity in processing speed and communication speed (e.g., Gb/sec Ethernet links for Mb/sec average data rates), along with extensive buffering. Quantitative specifications have not been prepared.

The ABMs, ARTC, and CCC are all located at the high-altitude telescope site. To enhance reliability, they are all diskless; their software is loaded over the network from the lower-altitude OSF. Large-capacity disks for software and data (especially the archive) are located at the OSF.

The AMB is a deterministic, multi-drop, serial communication bus that uses twisted pair cabling at a raw signaling rate of 1 Mb/sec. The maximum cable length is 40 m, and it can support up to 128 nodes. Electrical, timing, and message format specifications are those of CAN [16], so all controllers and line drivers are off-the-shelf parts. In the ALMA application, additional protocol rules are imposed. In particular, all communications must be initiated by one node, the master; all others are slaves. The master (ABM or ARTC) sends commands and monitor requests to the (slave) devices; for monitor requests, the device must respond with the requested monitor data within a 150 μsec timeout. Each message contains a 29 bit ID, where the first 11 bits are used as the slave node address and the remaining 18 bits identify the message's meaning to that node. Details are given in [18].

At each antenna, a separate AMB is dedicated to communicating with the Total Power Digitizer assemblies within the Downconverter Monitor/Control modules. This allows sufficient bandwidth to read all 12 total power detectors every 2 msec. All other devices at the antenna can share a single AMB, but several separate busses are implemented for engineering convenience.

Each device connected to the AMB must include a Standard Interface (AMBSI), which is a small printed circuit board designed for ALMA. It handles the AMB protocol and provides signals and data to/from the device when messages directed to it are received. Two versions of AMBSI have been developed, allowing a variety of connection methods to the device. Details are given in [19].

Within devices, the monitor and control functions are expected to adhere to the following principles. These principles help to keep the device designs simple, which allows them to be stable and thus much more maintainable.

- Control at the lowest possible level must be provided. That is, the simplest internal functions of the device must be accessible to remote control via MC. All possible states should be accessible, even those not required for normal operation (e.g., invalid combinations of switch positions), unless safety would be compromised.
- Monitoring as far as possible from a control point must be provided. That is, the monitor data should show the *effect* of the control signal inside the device, not merely that the control signal was received. It should be possible to determine the complete internal state of the device from monitor data.
- Implementation of "high level" local control is rarely necessary, and is discouraged.
  – However, it may be justified if an excessive command rate across the AMB would otherwise be needed.
  – High level functionality should be implemented on the MC side of the AMB.
- Implementation of automatic local control is undesirable. Decisions should be made at higher levels in order to ensure consistency. However, if necessary to protect equipment or personnel, safety interlocks and automatic shutdowns should be implemented locally.

The monitor/control software is organized so as to take high-level observing commands from the scheduling system and to turn these into detailed control of the array, finally producing science data products and monitor data archives. In brief, the scope of the control software is "scheduler to data."

Logically, the software is partitioned so that control flows in a master-slave fashion from a central executive who controls high-level ("composite") software devices that in turn control their constituent parts. The lowest level software devices are referred to as device controllers, and represent a proxy for the actual hardware — that is, they communicate with the hardware. Monitor data as well as astronomical data are collected from the devices by a process in the real-time computer attached to the hardware, and buffered there for distribution via a publish/subscribe mechanism to consumers of the data, which include the processes that format and archive the data. The software is distributed among computers so that only the device controllers and software directly concerned with low-level device activities are on the local real-time computer. All higher-level software entities are concentrated on the ACC.

3.3 Built-in calibration and sounding

In general, calibration of the ALMA instrument will rely on observations of known astronomical sources. To supplement this and to compensate for the shortage of appropriate natural sources, some devices will be built into each antenna for the purpose of generating calibration signals. In addition to calibrating the instrument, it is necessary to determine and correct for the attenuation and delay of the atmosphere, which is generally variable in space (across the array) and in time (faster than a full astronomical calibration). Built-in devices for such sounding of the atmosphere are also planned. Some concepts include devices which provide a combination of information about the instrument and the atmosphere.

Instrumental calibration is a matter of determining the overall gain of each signal path to a detector. For interferometry and for (single dish) spectroscopic radiometry, the detector is the correlator. For continuum radiometry, either the correlator or the analog square law detectors in the Downconverter can be used. Interferometry requires knowing the complex gain of each channel, including phase. Accurate measurement of polarized flux requires knowledge of the actual polarization state of each channel (usually stable), as well as the phase difference between the gains of the two channels of a pair (typically variable).

Details of the calibration and sounding procedures are beyond the scope of this report. At this writing they are still under discussion and there is considerable uncertainty about how they will actually be done. Here we discuss only the built-in instrumentation.

3.3.1 Magnitude of the instrumental complex gain

Several different schemes have been suggested for built-in signal sources, including a 2-temperature thermal source in the subreflector; chopping between the sky and an ambient load; chopping between the sky and a semi-transparent vane; and chopping among sky, ambient load, ambient semi-transparent vane, and hot semi-transparent vane. All except the subreflector-mounted device would be placed just in front of the focal plane and would have to be large enough to cover all bands. An assembly with moving parts, mounted above the Front End Assembly, is envisioned.

At this writing, it is not clear which, if any, of these concepts will be adopted and very little detailed design has been accomplished.

3.3.2 Phase of the instrumental complex gain

There is no known method to determine the absolute interferometric phase using built-in sources. However, a narrow-band source in the subreflector, where the signal is generated photonically and is tunable over the whole ALMA frequency range, was proposed as a method of measuring phase *fluctuations*, provided that it could be sufficiently stable. Development of this idea was dropped for cost reduction.

It is feasible to supply a source that enables measuring the phase difference of the gains of the two channels of a polarization pair on the same antenna. This is necessary for polarimetry and is difficult to do astronomically. The photonic reference might have accomplished this. At present no such source is planned.

3.3.3 Atmospheric attenuation

No special instrumentation is planned for measurement of atmospheric attenuation. However, it can be inferred from atmospheric emission measurements using the main receivers as radiometers via "tipping" curves; one or a few antennas can do this occasionally. To detect rapid fluctuations in attenuation, chopping against an ambient temperature load can be used (if implemented for gain calibration) to provide approximate results. Again, details of these techniques are beyond our scope here.

A dedicated tipping radiometer, covering the whole ALMA frequency range, is under consideration but not currently planned. It could consist of a small reflector antenna with either a Fourier transform spectrometer or a full ALMA front end assembly.

3.3.4 Atmospheric delay

Variation of the delay through the atmosphere can have a severe effect on interferometry. On short baselines the delay is nearly the same for both antennas, but for baselines of 1 km or more the delays become less correlated so that the path to each antenna must be determined and corrected separately. Even on shorter baselines the difference is sometimes significant.

Two approaches to removing these delay variations are planned. The first is the "fast phase calibration" mode of observing, discussed earlier. Here the observing time is split between the target source and a nearby calibrator during cycles of 10 sec or more. The calibrator is typically observed at a lower frequency band, where it is likely to be stronger. This technique involves no special equipment, but it does impose strong requirements on the antenna drive and servo to provide fast and accurate movement and on the LO and signal processing systems to provide fast frequency switching with stable and repeatable phase. The second approach involves the assumption that the unpredictable part of the atmospheric delay is almost entirely due to water vapor, and it depends on the installation of devices on each antenna to measure the column density of water vapor along the signal path.

The water vapor column density is measured by special radiometers (WVRs) built into each antenna. These measure the brightness temperature of the sky at several frequencies around the 183 GHz water line. Assuming that the sky brightness at these frequencies is dominated by emission from the water vapor, the spectral shape helps determine the altitude, temperature, and degree of saturation of the line. In principle, this allows the column density of water vapor to be deduced, and from it the signal delay at the observing frequency.

The accuracy achievable with WVR corrections is not known, but a project **requirement is that the contribution of radiometer noise to the estimate of path length fluctuations be less than $.01w + 10\,\mu$m, where w is the condensed water depth along the line of sight (column density expressed as depth of condensed water). An additional error occurs because of uncertainties in atmospheric modeling; this is currently estimated to be $.02\,\delta L$, where δL is the uncorrected rms fluctuation in path length. Furthermore, some path length fluctuation occurs because of density variations in the dry air, and these are invisible to the WVRs. The effect of these density variations is very uncertain, but they may be significant under good observing conditions at Chajnantor, when the other errors are small. As a guess, our budget currently allows an rms uncertainty of $10\,\mu$m at zenith. The noise, scale error, and density fluctuation errors should be independent, so the net rms residual after WVR corrections is taken to be [25]

$$\sqrt{(.02\,\delta L)^2 + (.01w + 10\,\mu\text{m})^2 + (10\,\mu\text{m})^2}.$$

Under common conditions of $\delta L = 500\,\mu$m, and $w = 1800\,\mu$m, the residual should be less than $31.3\,\mu$m (105 fsec), and under good conditions of $\delta L = 143\,\mu$m and $w = 1000\,\mu$m it should be less than $22.5\,\mu$m (75 fsec). In tests on existing telescopes, the smallest residual fluctuation ever achieved has been $25\,\mu$m (83 fsec); however, the ALMA instruments are expected to be more precise. The specification applies only to the *fluctuations* in delay over intervals of 5 minutes or less, and only at constant air mass. The total path length, changes over longer intervals, and changes with air mass may be limited by WVR calibration, WVR gain drift, and atmosphere modeling errors. Additional **requirements on the WVR are that it achieve the specified accuracy at a sampling interval of 1 sec or less, and that its beam on the sky be aligned with that of the observing receiver within 10 arcmin. (The WVR is coupled to the sky through the subreflector and main reflector just like the observing receiver, and therefore the beams must have a non-zero offset. A pickoff mirror in the center of the focal plane is used for the WVR.)

3.3.5 Antenna Pointing

Each antenna has mechanical and structural provisions for mounting an optical telescope to the main reflector backup structure so that it can view the sky through a hole in the main reflector in a direction very nearly aligned with the antenna beam. The optical telescope contains a video camera whose analog output signal is modulated onto an optical carrier and can be brought out of the antenna on optical fiber. The optical telescope will be installed and used to verify pointing and tracking accuracy during the initial outfitting and commissioning of each antenna. It is not intended to have optical pointing telescopes on any operational antennas; if this should later prove to be desirable, hardware for digitizing and processing and/or transmitting the images, as well as the optical telescope itself, would need to be added to each antenna.

3.4 Auxiliary optics

The package of movable vanes and loads that is intended to provide for calibration of the power gain will also contain some devices that can be inserted into the input beams of one or more bands. At present, two such devices are planned. Each can be inserted or removed by a mechanical actuator upon command.

3.4.1 Solar filter

To allow observations of the Sun, an attenuator of **13 dB can be inserted into the path for any band.

3.4.2 Quarter wave plate

To convert the input polarization from nominally linear to nominally circular, **a quarter wave plate can be inserted into the input path for Band 7 only. The plate has a design center frequency of 345 GHz. Clearly the bandwidth should be maximized and the insertion loss should be minimized, but there are no specific requirements.

This feature is being included because it was requested by some astronomers through the ASAC [5]. All polarization parameters can be measured without it.

4 METHODS AND ALGORITHMS

4.1 Timing and clocks

Throughout the ALMA telescope, including the Technical Building and each of the antennas, a set of fixed-frequency, periodic reference signals is available to any device. These signals are coherent across the array, and are all derived from a single master oscillator as described earlier. **The nominal frequencies are:

f1 = 2.0 GHz, the R2G signal;

f2 = 25.0 MHz, the R25M signal; and

f3 = 20+5/6 Hz (20.833.. Hz, 1/.048 Hz exactly), the R21 signal.

The actual frequencies, with respect to the SI second, are kept very close to these values. Note that ratios of the nominal frequencies are integers; these are exact. The value of f3 is intended to avoid small-harmonic relationships to the power line frequencies.

An important concept is that of "array time," which is the continuous measure of time on which the array operates internally. It is determined, up to an ambiguity interval of 48 msec, by the phases of the three reference signals at the outputs of the Central Reference Generator in the Technical Building. At other locations (especially at the antennas), the distributed versions of the reference signals represent the array time plus the propagation delay of the distribution system; if true array time is needed to high accuracy, then the user must calibrate the propagation time and subtract it. The f1 reference is the most accurate measure of time, but it has an ambiguity interval of 0.5 nsec. Therefore, at all distribution points f2 is required to be accurate to $\ll$ 1 cycle of f1, so as to extend the ambiguity interval to 40 microsec. Similarly, f3 must resolve a cycle of f2, extending the ambiguity to 48 msec. For larger time intervals, array time is determined by the monitor control subsystem in the Array Real Time Computer (see Monitor and Control, above), which integrates the phase of f3 by counting cycles. Time within the computer must be accurate enough to resolve a cycle of f3, thus extending the ambiguity interval by any desired amount. The MC system is responsible for distributing this time to all devices that need it (including other computers) while maintaining the ability to resolve a cycle of f3. It does this by sending a message to the device and guaranteeing its delivery time to be within a known cycle of f3.

It is intended to keep array time very close to International Atomic Time. Nevertheless, all devices in the ALMA telescope are synchronized to array time, and only indirectly (and less accurately) to any external measure of time.

Many of the considerations affecting timing and synchronization in the ALMA telescope are covered in ALMA Memo 298 [6]. Detailed discussion of the implementation of the master clock is given in Specification 09001NX0003 [7]. Some of the main principles are listed below.

- **Reference frequencies are maintained within 1 part in 10^{11} of their nominal values with respect to the SI second. (This is feasible if the master oscillator is a hydrogen maser that is regularly checked against distributed international timing signals.)
- R2G and R25M are sinusoidal and are delivered to users on coaxial cables. R21 is a logic waveform delivered on a bus conforming to RS485, on the same cable as the serial bus for monitor/control messages (see Monitor and Control, earlier in this document). It spends a minimum of 1 microsecond and a

maximum of 25% in the logic-1 state during each cycle. The rising edge of R21 marks "timing events" that are used for synchronization.

- **At any one location, the phase stability of R21 on all time scales shall be less than 0.1 cycle of R25M (4 nsec); and the phase stability of R25M shall be less than 0.1 cycle of R2G (50 psec). The phase stability of R2G shall be better than 1 psec with respect to the master oscillator. (This requires stabilization of the transmission path length to each antenna.) However, there is no specification for the absolute phase of any signal nor for the relative phases among them, which may vary from one location to another. This means that, at any point in the telescope, it is possible to have a stable measure of time to an accuracy of 1 psec and an ambiguity interval of $1/f3 = 48$ msec by means of these signals alone (i.e., without maintaining any local clock).
- At one location in the central building, the phase of R21 will be adjusted so that it has a known relationship to external measures of time. **In particular, the 0-to-1 transition of each 125th cycle of R21 (every 6 seconds) shall coincide with the UTC second within 10 microseconds. Knowledge of the difference between those transitions and the UTC second shall be maintained to < 100 nsec.
- Any device that requires synchronization of its timing to the rest of the telescope shall accomplish that synchronization using one or more of these signals. A device should use only those reference signals needed to achieve its accuracy and ambiguity requirements.
- Any frequency that is a multiple of f3 may be synthesized with unambiguous phase from the references. Devices may synthesize such frequencies for internal use. If any device-internal signal requires synchronization to other parts of the telescope, then it must be synthesized in this way; thus, it must have a frequency that is a multiple of f3.
- If a frequency not in the given set is needed by multiple devices within one room, then that frequency may be synthesized once and distributed locally. Care must be taken to maintain a phase-stable distribution network. In particular, 125 MHz will be synthesized and distributed at each antenna.
- The array master clock is driven by the same reference signals, and it maintains knowledge of the complete time by counting (see [7]). The reference signals provide knowledge of the array time with an ambiguity interval of $1/f3=48$ msec. Any device which requires synchronization on longer intervals must do so using commands received via the Monitor-Control (MC) system [6].

4.2 Fringe rotation

Phase tracking of the LO, or fringe rotation, is implemented via direct digital synthesizers that are the main components of the FTS Assemblies in the First and Second LOs. As discussed under Synchronization below, the MC subsystem determines the desired phase (relative to the nominal LO frequency and hence the nominal FTS frequency) as a function of time, where time is measured from the local timing events. Since a DDS operates at high speed (125 MHz clock in our design), it is not practical to control its phase directly, but its frequency can be adjusted with very fine resolution and the adjustments can be made to occur on a known clock cycle. Whereas the MC system knows the exact value of all frequency commands, and whereas the internal arithmetic of the DDS uses only integers, the MC computer can calculate exactly the change in phase between any two clock cycles. Practical DDS devices allow the phase to be set to zero on a known clock cycle, and therefore the MC computer can know the phase at any time thereafter. In this way, it is possible for MC to cause the phase to follow the desired function of time by providing only frequency commands. The necessary algorithm is described more fully in Appendix A of [13].

4.3 Phase switching

By systematically switching the phase of the first LO by 180 degrees in such a way that it spends half of its time in one state and half in the other, and then cancelling the effect of this by simultaneously changing the sign of the signal immediately after digitization, it is possible to suppress the response to undesired signals that may be coupled into the signal path between the first mixer and the digitizer. Such signals will not be affected by the LO phase switching, but they will be modulated by the digital sign reversals and will retain this modulation when cross-correlated against signals from other antennas. By selecting the switching patterns so that they are mutually orthogonal among all antennas over some time interval, cross-correlations computed over the same interval will have zero net contribution from the undesired signals. This does not require that those signals be the same at all antennas, but it does require that they be constant over the integration time. This suppression of undesired signals occurs only in cross correlation and is not effective for autocorrelation nor for total power detection. It could be effective when cross-correlating oppositely

polarized channels of the same antenna if different orthogonal switching waveforms could be used, but this is not possible in the ALMA design because the first and second LOs are each common to the channels of a polarization pair.

The LO phase switching is implemented in the DDS of the FTS assembly, and its cancellation is implemented in the DTS Transmitter assembly. For the first LO, the PLL for which this DDS is providing an offset reference is not necessarily at the final LO frequency; for band 4 and above, the PLL output is multiplied in frequency by a factor of 2, 3, 5, or 9 (see Table 5). Therefore, the phase offset applied at the DDS must be 90, 60, 36, or 20 degrees, respectively.

We require a set of $N = 64$ mutually orthogonal binary sequences. The only known practical such set is that of Walsh functions [20]. To achieve orthogonality over an interval T, a sign change must occur in one or more functions every $T/2N$. To achieve effective suppression over the correlator's minimum integrating time of $T = 16$ msec then requires sign changes every $125\,\mu$sec. Furthermore, these functions are orthogonal only if they are properly aligned in time. This alignment is required when the signals are cross-correlated, not when the sign changes are applied at the antennas; in between, they suffer different amounts of delay and this delay varies as a source is tracked. Achieving this synchronization is discussed in Section 4.5 below.

It is also possible to apply the sign switching at the second LO, but then any undesired signals that enter the IF path between the first and second mixers will not be suppressed. Since both LOs use the same FTS Assembly, this option remains open as a function of software. However, the first LO's PLL is required to have a fast response in order to support the $125\,\mu$sec switching interval and we do not impose such a requirement on the second LO. Therefore, sign switching in the second LO might need to be slower.

If 90 deg phase switching were implemented in the first LO in addition to the 180 deg phase switching just described, it would be possible to separate the upper and lower sidebands of the first mixer after correlation. This is not required in ALMA, so it will not be discussed further here. (See [13] and Sideband Suppression, below.)

4.4 Sideband Suppression

Generally, the IF contains responses from both sidebands of the first mixer. For 2SB mixers, one sideband is suppressed by at least 10 dB, but often much greater suppression is desired. Because the ALMA signal processing contains two LOs, it is possible to arrange that signals in one sideband have zero cross correlation. Let each LO be offset in frequency from its nominal value by δf_i for antenna i. Then for one sideband the offsets cancel and for the other they add. The cross product of antennas i and j is then unaffected for the first sideband, but for the second it is multiplied by a sinusoid at $2\delta f_i - 2\delta f_j$. If the integrating time is an integral number of cycles of this modulation frequency, then any signal in the second sideband correlates to zero. This is easily achieved for all antenna pairs by setting $\delta f_i = i\,\delta f_0$, $i = 1, ..., 64$, where $\delta f_0 \geq 1/T$ for integrating time T. For example, the choice $\delta f_0 = 1\,$kHz is convenient.

For bands with DSB first mixers, continuum sensitivity can in principle be better by a factor of $\sqrt{2}$ if no suppression is done and the sideband responses are equal. But this would require extraordinarily accurate delay tracking (see [15]) along with 90 deg phase switching. With unequal sideband responses, the improvement is less. For a spectral line source in one sideband, sideband suppression does not affect sensitivity but it greatly simplifies calibration. Therefore, sideband suppression will almost always be used for interferometry in the DSB bands. In total power (single dish) continuum, no suppression is possible.

4.5 Synchronization

As illustrated in Figure 2, there are three functions that require synchronization in time over substantial distances, either among antennas or between each antenna and the Technical Building. These are delay tracking, phase tracking (fringe rotation), and phase switching. The issues involved as discussed briefly here, and more fully in [6].

As explained earlier (Sections 2.6.1 and 3.1.5), delay tracking is implemented in three stages, with most of it in the correlator's station buffer at a resolution of 64 (possibly decimated) samples. The next stage is in the correlator's input circuitry; it provides 250 psec (1 sample) resolution and 2047 sample range. And the final stage is in the digitizer (at the antenna), which provides sub-sample resolution with a range of nearly 1 sample by varying its clock's phase. The fine delay steps inserted at the antenna must be synchronized to the coarser steps inserted at the correlator. A correlator delay step occurs just before a sample that appears at the output of the correlator's delay line at times that are synchronized with the R21 signal in the Technical Building. For a fine delay step to occur at the same sample, it must have been inserted by the
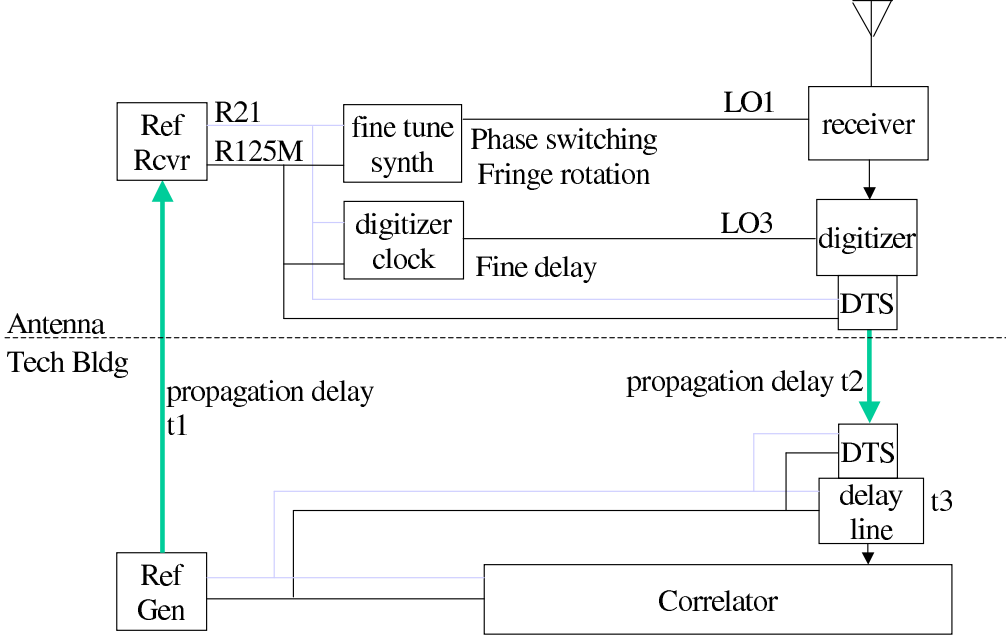
Figure 2: Signal and LO paths, illustrating the delays that must be taken into account for synchronization of delay tracking, phase tracking, and phase switching.

digitizer clock exactly $t_2 + t_3$ earlier, where t_2 is the propagation time from sampling to the correlator, and t_3 is the the the delay within the correlator, including its (variable) delay line (see Fig. 2). However, time at the antenna is determined relative to the timing events of its local R21 signal, and these occur later than those at the Technical Building by t_1, the outgoing propagation delay. For example, as illustrated in Figure 3, to cause the fine delay step to occur at the sample that exits the correlator's delay line at the same time as the timing event there ("sample1"), the digitizer clock must insert the step $t_1 + t_2 + t_3$ prior to the antenna's timing event. The MC subsystem must know this time interval and take it into account when commanding the delay steps. This is actually simpler than it might seem, because the variable delay t_3 must already include compensation for the propagation delay t_2, in order that the signals from the same wavefront of all antennas arrive together at the correlator's delay line output. That is, we can write

$$t_2 + t_3 = t_g + t_0$$

where t_g is the geometrical delay, which can be accurately computed from the antenna location and source position alone, without any knowledge of delays within the electronics; and t_0 is a common offset that is the same for all antennas. Thus only the value of t_1 needs to be determined for each antenna, along with the single common offset t_0; these should be quite stable. They must be measured whenever an antenna is moved or a relevant cable is changed.

For proper phase tracking, the phase offset applied to the LOs at any one time must correspond to the wavefront phase for a signal from the tracking direction in the sky. This can be calculated for each antenna as a function of array time, relative to a reference location near the center of the array where the phase offset is taken to be zero. However, the FTS Assembly that determines the LO phase has precise time information only with respect to R21, which is offset from array time by the outgoing propagation delay t_1. Therefore, the value of t_1 must be determined for each antenna and used by the MC subsystem in calculating the commanded phase. Specifically, the design provides that the MC system must transmit the desired phase at specified antenna timing events. This means that the synchronization is implemented in software and does not depend on any additional hardware. (MC must also transmit enough additional information, such as derivatives of the desired phase, to allow the FTS to interpolate accurately between commands. For sidereal rate tracking, providing the phase and its first derivative at each timing event should be sufficient; but providing commands more often or less often, with fewer or more parameters, is possible if necessary.
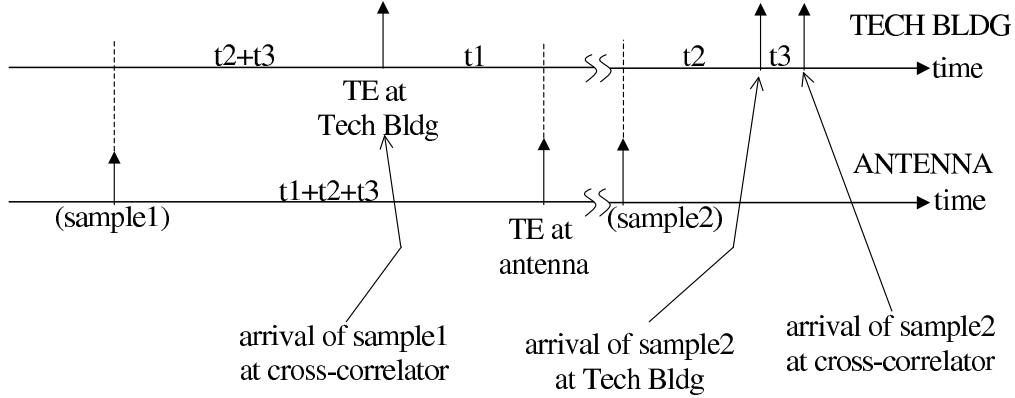
Figure 3: Timing diagram illustrating the synchronization of sampling at an antenna with timing events (TEs) at the antenna and at the Technical Building. See Fig. 2 for notation. 'Sample1' happens to arrive at the cross correlator simultanously with the timing event there; it must have been captured at the antenna somewhat earlier, as shown.

These details are part of the MC design.) In principle, for this purpose the total values of t_1 need not be known for all antennas, since only their differences are important. But the total values are needed for the delay tracking discussed above, so they can be used here as well.

Finally, regarding phase switching, it is necessary to arrange that all antennas begin a phase switching cycle at samples that will arrive simultaneously at the correlator's delay line output. This situation is nearly the same as for delay tracking, and can be handled in the same way. Here only the differences in delays are important, so the common component of the signal delay t_0 need not be known. As discussed in section 4.3, phase switching is implemented in the FTS assembly of either the first or second LO, and it must be simultaneously cancelled by a sign reversal of the samples just after the digitizer.

In summary, we have two kinds of time-critical events at an antenna whose time of occurrence varies with respect to the local R21. These are phase switching sign changes and fine delay steps. (For fringe tracking, commands are always tied to R21 and timing compensation is done in software.) It is possible that additional, similar time-critical events will be defined by future modes or new hardware. For all such cases, we have chosen an implementation in which a signal with the required timing is generated locally in each hardware device by using the distributed R21 and R125M references. No variable-timing signal is transmitted from one device to another. The desired time offset with respect to R21 is sent to the device as a command from MC, and a new command is sent whenever the value changes significantly. The following devices in the current design are affected by this: 1st LO FTS; 2nd LO FTS (four units); Digitizer Clock; and DTS Transmitter (which includes the Digitizer). For phase switching (FTS assemblies and DTS Transmitter), it is only the *starting time* of the switching cycle that needs to be specified; the appropriate Walsh function must then be generated locally. An accuracy of about 1 μsec is necessary. For delay tracking, an accuracy of about 10 μsec is sufficient, as noted in section 4.1.5. It should be easy to achieve much better accuracy in the local signal synthesis of digital signals (the R125M reference allows a resolution of 8 nsec), but the overall accuracy is limited by the response speed of the associated phase locked loops.

4.6 VLBI support

The ALMA telescope is **required to be usable as a single station for VLBI experiments along with other telescopes. To achieve high sensitivity, large collecting area is desired, and therefore we **require the ability to combine any subset of the antennas (including all 64) into a phased array. This means that signals from the selected antennas are summed to produce one sample stream for each polarization, and that the LO phases are adjusted in real time so that the signals add coherently.

As mentioned earlier (section 2.6.1), implementation of the summation feature is pending and its parameters are not yet determined. This includes the bandwidth(s) of the summation, whether the result is in analog or digital form, and the output format. This is partly because VLBI signal transmission technology is evolving, and it is uncertain what methods will be current when ALMA is in operation. To date, most VLBI has involved recording digital signals on magnetic tapes, shipping the tapes to a correlation center,

and processing in non-real-time. A new generation of recording equipment based on magnetic disks is now under development. Eventually, real-time transmission of signals is envisioned, perhaps via satellite or via optical fiber networks. Consequently, no VLBI recording equipment is included in the ALMA budget.

The requirement to provide real-time absolute phase tracking in the LOs is unique to VLBI. When ALMA is operated as a stand-alone synthesis telescope, phase calibrations and corrections can be applied to the visibilities in post-correlation processing and need not be known in real time. A large class of antenna-based phase errors can be removed in the imaging process by "self calibration." Atmospheric corrections, partly from rapid calibrator observations and partly from WVR data, can also be applied. In the phased-array mode for VLBI, none of these possibilities is available. It is necessary to observe a calibrator, solve for the antenna-based instrumental and atmospheric phases, and apply the appropriate adjustments to the LO phases, all prior to observing the target source. While this imposes no new requirements on the hardware, it does require complex and efficient software. Due to variation of the atmospheric delay with time, across the array, and between calibrator and target, maintaining accurate phasing is likely to be successful only for compact configurations of the antennas and only at the lower frequency bands. Therefore, support of this mode is **required only for bands 1 through 6 (to 270 GHz) and only for antennas within 1 km of the array center. It may eventually be possible to extend the technique to higher frequencies and larger configurations. The usefulness of doing so depends on the availability of other telescopes suitable for VLBI at the same frequencies.

Assuming that the relative phases are adjusted so that the signals can be coherently summed, the phase stability of the composite signal must then be considered. It is affected by phase variations that are common to all antennas, including both the atmosphere and the instrument; these have no effect on visibilities measured by the stand-alone array. This imposes a **requirement on the common-mode LO stability that would not otherwise be needed. It is for this reason that ALMA's master oscillator will be a hydrogen maser; for the stand-alone array, a high quality crystal oscillator would be adequate. The rest of the common-mode LO system should not degrade the H maser stability on time scales of 0.1 to 100 seconds.

REFERENCES

[1] S. Guilloteau, A. Wootten, E. van Dishoeck, B. Butler, "ALMA Scientific Requirements." In preparation.
[2] J. Lamb et al., "ALMA receiver optics design." ALMA Memo 362, April 2001.
[3] System Block Diagram, ALMA Document 80.04.01.00-004. Rev G, dated 2003-05-30, was current in August 2003.
[4] R. Freund, "Digital transmission system signalling protocol." ALMA Memo 420, September 2002.
[5] R. Crutcher, "Polarization observations with ALMA." Appendix C of *Report of the ALMA Scientific Advisory Committee March 2000 Meeting*, pp 11–16.
    http://www.alma.nrao.edu/committees/ASAC/asacreport.sep00.pdf
[6] L. D'Addario, "Timing and synchronization." ALMA Memo 298, March 2000.
[7] L. D'Addario, "ALMA Master Clock Implementation." Specification 09001NX0003, 2000-10-24; revised 2003-08-06 as ALMA Document 80.04.00.00-001.
[8] ALMA Project, *ALMA Construction Project Book,* version 4.0, 2001-Feb-08. Chapter 10, "Correlator" by J. Webber *et al.*
[9] C. Broadwell and R. Escoffier, "The ALMA correlator long term accumulator sub-system." ALMA Memo 294, March 2000. (Some details are now obsolete.)
[10] R. Escoffier and J. Webber, "Enhancing the performance of the baseline ALMA correlator." ALMA Memo 441, November 2002.
[11] M. Torres, "Sampling clock and fine delay system for ALMA." Informal notes dated 2003-Feb-13. Available from the author, torres@iram.fr. (Not to be confused with obsolete version dated 2002-Dec-18.)
[12] L. D'Addario, "Notes on delay tracking for ALMA: resolution and tolerance." Memo dated 2003-02-08, ALMA document 50.00.00.00-002. Available from the author, ldaddario@nrao.edu, or via ALMAEDM.
[13] L. D'Addario, "Fringe tracking, sideband separation, and phase switching in the ALMA telescope." ALMA Memo 287, February 2000.
[14] A. Wootten et al., "Frequency band considerations and recommendations." MMA Memo 213, 1998-May-28.

[15] A. Thompson and L. D'Addario, "Relative sensitivity of double- and single-sideband systems for both total power and interferometry." ALMA Memo 304, April 2000.

[16] W. Lawrenz, *CAN System Engineering.* Springer-Verlag, New York, 1997.

[17] A. Ferris, "Physical architecture of the ALMA system." Undated technical note, with block diagram.

[18] M. Brooks and L. D'Addario, "ALMA monitor and control bus interface specification." ALMA document 70.35.70.03-001, version H, 2001-09-07.

[19] M. Brooks, "AMBSI1 Standard Interface Design Description," ALMA document 70.35.10.01-001, version B dated 2002-03-05; and W. Koski, "ALMA Monitor and Control Bus AMBSI2 Standard Interface Design Description," ALMA document 70.35.10.02-001, version A dated 2001-05-03.

[20] L. D'Addario, "Orthogonal functions for phase switching and a correction to ALMA Memo 287." ALMA Memo 385, 2001-Aug-09.

[21] ALMA Project, *ALMA Construction Project Book,* version 4.0, 2001-Feb-08. Chapter 7, "Local Oscillators," L. D'Addario, ed.

[22] L. D'Addario and M. Holdaway, "Joint distribution of atmospheric transparency and phase fluctuations at Chatnantor." LAMA Memo 801, 2003-Nov-11.

[23] B. Quertier, G. Comoretto, A. Baudry, A. Gunst, A. Bos, "Enhancing the ALMA correlator performances with a second generation correlator digital filter system." ALMA Memo 476, 2003-Aug-27.

[24] M. Holdaway and L. D'Addario, "Simulation of atmospheric phase correction combined with instrumental phase calibration using fast switching." LAMA Memo 803, 2004-Jan-23.

[25] R. Hills, private communication.

APPENDIX: Summary of Technical Requirements and Specifications

Parameter	Value
Number of antennas	64
Antenna diameter	12m
Antennas with nutators	4
Frequency coverage	31.3–950 GHz, all atmospheric windows
Edge frequency ratio of each band	< 1.35
IF range	4–12 GHz
1st mixer sideband ratio	> 10 dB, SSB and 2SB
	< 3 dB, DSB
Instantaneous bandwidth	> 8 GHz in each of two orthogonal polarizations
Digital transmission	no loss of samples
	all samples in order
	repeatable latency
	bit error rate $< 10^{-6}$
Archive writing rate	> 40 MB/s for correlation products; > 60 MB/s total
Tuning resolution: any sky frequency	at desired baseband frequency within $\pm5\%$ of baseband width
	at desired IF within $\pm5\%$ of IF bandwidth
Phase switching	180d in 1st LO and digitizer
	Walsh functions, 64 orthogonal by antenna
	Cycle time .016 sec
	Synchronization within 1 μsec
Phase accuracy	requirements apply across frequency changes
electronics	24 fsec drift, 65 fsec rms jitter
structure	14 fsec repeatability, 38 fsec rms vibration
Phase ambiguity	all frequency synthesis unambiguous
Frequency change time	< 1.5 sec (< 0.1 sec desired)
	< 10 msec for .03% or less
Independently tunable subarrays	≥ 4
FTS assemblies	identical for 1st and 2nd LOs
Sampling clock phase control	1/8 sample accuracy
Delay synchronization	$< 10\mu$sec (antenna to correlator)
Sampling clock	common to all channels at antenna
Water vapor radiometer	
path error due to noise	$.01w + 10\,\mu$m
sampling rate	1 Hz
beam offset	< 10 arcmin
Timing reference frequencies	2 GHz (R2G)
	25 MHz (R25M)
	20+5/6 Hz (R21)
Absolute accuracy of ref. frequencies	1×10^{-11}
Phase stability of references	(at destination devices)
R2G	±1 psec rms (relative to Master)
R25M	±50 psec rms (relative to R2G)
R21	±4 nsec rms (relative to R25M)
Array time accuracy	R21 at REFGEN out $\equiv$ UTC within 10 μsec
	UTC known to 100 nsec
VLBI support	usable as a single station (phased up)
	sum output for available for any subset of antennas
	phasing up required only for bands 1–6
	common mode stability for 0.1–100 sec should not
	degrade H maser stability