THE FFT AND CROSS MULTIPLIER SYSTEMS OF THE VLBA CORRELATOR

by Ray Escoffier

Oct. 16, 1987

1.0) INTRODUCTION

This memo will describe the FFT and the cross multiplier systems of the VLBA correlator.

Much of the discussion in this memo will require some knowledge of the butterfly gate array chip being developed for the correlator. Therefore, the first topic to be covered is the ASIC (application specific integrated circuit) chip itself (also see VLBA correlator memo 87).

2.0 The Butterfly ASIC

Figure 1 gives an internal block diagram of the ASIC chip that will be used in the VLBA correlator.

2.1 Digital Signal Representation

The signal representations described below are generally expressed in complex floating point form. This signal representation is a non-standard form in which the real and imaginary components of a number have a common exponent bit field. Such numbers will be referred to in this memo in a short hand fashion, an example of which is 7,7,4 indicating 7-bit real and 7-bit imaginary mantissa fields and a common 4-bit exponent field. The mantissa fields are in sign-magnitude. Four basic precisions are used in the various applications of this ASIC:

1) Data points being Fourier transformed;

   7,7,4

2) Sin, Cos twiddle factors used in FFT butterfly;

   5,5,0

3) Points being cross multiplied;

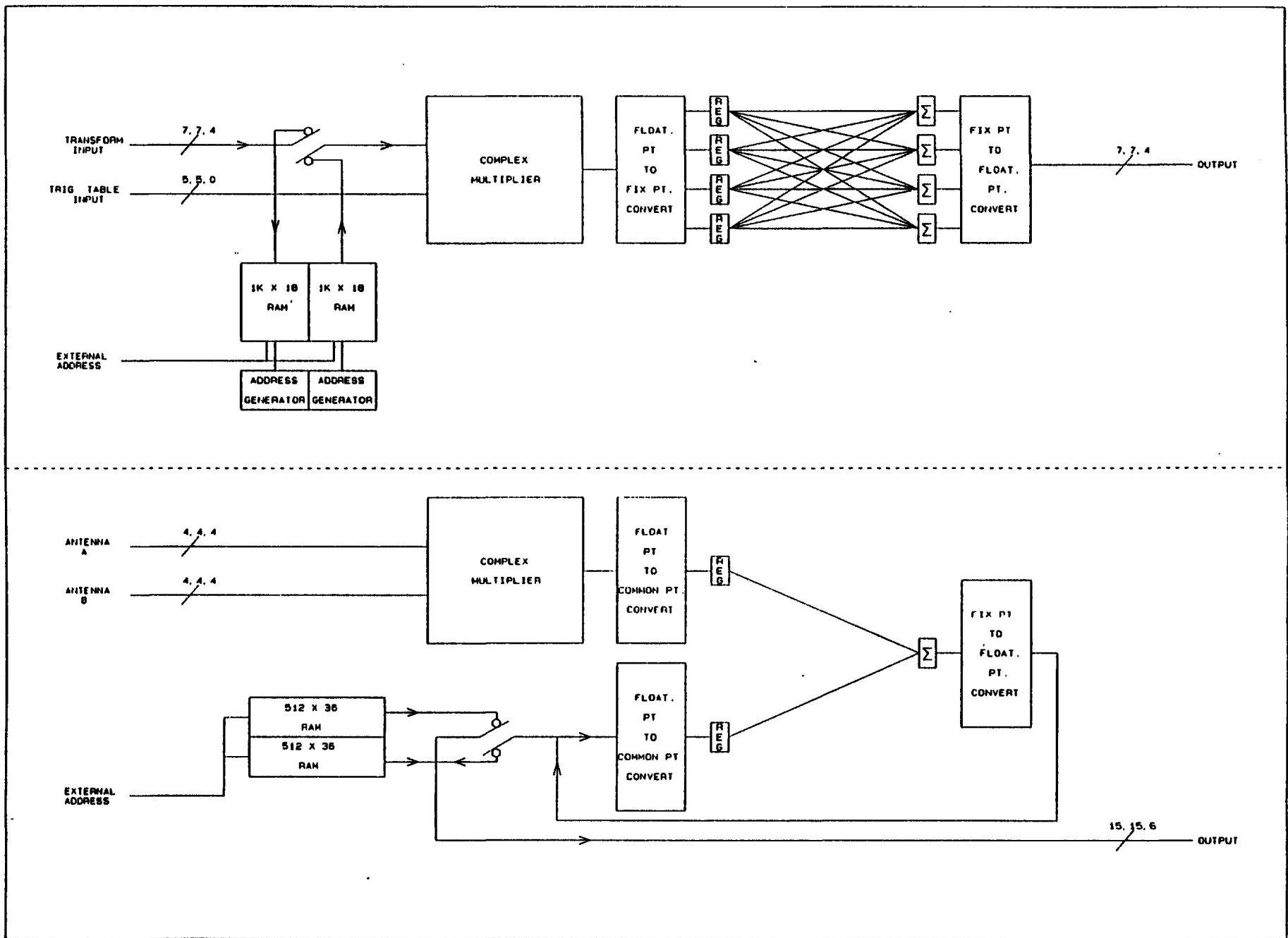   4,4,4

4) Accumulator precision;

   15,15,6

1

Figure 1

2.2 Functional Description

The ASIC chip to be used in the VLBA correlator is a multi-function chip that can be used in the following applications:

Radix 4 FFT butterfly;

The ASIC chip will perform radix 4 DIT FFT butterflys on four complex data points clocked serially into the chip on four consecutive clock cycles. Two input ports exist into the ASIC, one for the points being transformed and one for FFT twiddle factors (actually, a third port into the chip exists to assist in performing 2048 point transforms; this third port will be discussed below). One point and one twiddle factor enter the chip on the same clock edge. The ASIC is fully pipelined so that one complex data point enters and one complex output data point exits the chip every 32 MHz clock cycle. In order to avoid offset biases that would result from using such low precision two's complement arithmetic, the points into or out of the chip are expressed in sign-magnitude floating point. Within the chip, numbers are converted to a one's compliment fixed point format. RAM storage for the data points (not the twiddle factors) is inserted between the chip input port and the butterfly input circuitry. The RAM addresses for the FFT butterfly stages may be generated externally and applied to the chip via input pins or be obtained from an internal address generator. A RAM configuration of two independently addressable 1K X 18 banks allows double buffering of the points being transformed.

The top half of figure 1 shows the ASIC in its basic radix 4 butterfly configuration. The RAM buffer is shown on the input side of the chip double buffering the input data points. The complex multiplier block does the twiddle factor rotation of input points. After four points have been input, rotated, and converted to fixed point, the complex radix 4 summations between the four are done in 15-bit fixed point arithmetic. A final conversion back to floating point occurs at the ASIC output.

Radix 2 FFT butterfly;

The ASIC chip must also perform a radix 2 DIT FFT butterfly on two complex data points clocked serially into the chip on two consecutive clock cycles. Most of the details defined above for the radix 4 function are true for this application.

Radix 2 FFT butterfly (for a 2048 point FFT);

The ASIC chip will do a radix 2 butterfly as above except that two inputs data ports (in addition to the twiddle factor port) exist on the chip so that a point shuffle between

2

two FFT chains can occur at the final radix 2 butterfly stage. The internal RAM buffers onboard the ASIC chips are not large enough to hold all 2048 points of a 2048-point transform. When a 2048-point transform is required, two FFT engines located on the same FFT card will be used. Each will be loaded with one half of the 2048 points and each will perform 5 radix 4 butterflies on their collections of 1024 points. At the input to the 6th butterfly stages, a point shuffle between the two FFT chains will occur so that the final butterfly ASIC in each FFT engine will have the correct pairs of points in their respective RAMs to perform the final radix 2 butterfly stage and complete the 2048-point transform. In theory the point exchange could occur at any butterfly stage, but it is most easily done at the final stage.

Straight through function;

The ASIC chip can allow all points to flow straight through unaltered except for a possible rearrangement in time sequence.

Complex multiply/accumulator;

The complex multiplier portion of the ASIC will also be used to do the spectral cross multiplication. The very fast onboard RAM will then be used for short term accumulation storage. The number of spectral points at the FFT output will only be one half of the transform size since half of the points in the calculated spectrum cover an empty sideband. Hence, the most spectral points being output from a single FFT engine will be half of the 1024 RAM depth. Therefore, for the accumulation function, the two 1K X 18 RAM banks may be split a different way to yield two 512 X 36 RAM banks. A 36-bit RAM width allows the short term accumulation to occur with a quantization level of 15,15,6.

The ASIC chip will input two floating point complex numbers, one on the butterfly data point input port and one through the twiddle factor input port, will perform a complex multiplication between the two, and add the complex result into a complex floating point accumulation obtained from the RAM. The accumulation result is stored in the RAM, as discussed above, across the entire 36-bit width of the RAM. This operation is seen in the lower half of figure 1.

Two operating modes for the ASIC are required in the multiply/accumulator application. In one mode, points to be multiplied will enter the chip in pairs, one pair every two (32-Mhz) clock periods. On average, the RAM must be read, a multiplication done, the result added to the accumulator and the new sum stored back in RAM in the two available clock ticks. The other mode (to process polarization observations) requires that point pairs be multiply/accumulated one pair every clock cycle.

In the latter mode, two consecutive point pairs entering the chip will be added into the same accumulation result (see section 3.0 for more detail). Thus, between reading the accumulator partial sum in RAM and writing the new sum back into the RAM, two cross multiplications will be made. In summary, the RAM will be read, two multiplications made, both results added via the accumulator to the same accumulation sum, and the new accumulation stored back in RAM every two clock ticks in this mode. In both of these modes, the RAM access requirement is one RAM read or one RAM write per clock cycle.

Number controlled oscillator;

The ASIC chip has a 10-bit slice of a number controlled oscillator (NCO) on board for external applications (this function is not represented in figure 1). Two secondary storage registers for loading an initial oscillator phase and oscillator rate are supplied. These secondary storage registers are loaded serially. The NCO adder carry lines between slices are pipelined so that any number of the bit slices may be tied together to make larger NCOs. The functions intended for these NCOs in the correlator are the short term fringe tracking and the fractional sample correction phase ramp generation.

3.0) THE FFT SYSTEM

The FFT portion of the VLBA correlator performs most of the signal processing functions required of the station logic in the VLBA correlator. Samples are collected serially from the playback interface and Fast Fourier Transforms are performed on them. This system of the VLBA correlator can service up to 8 channels from each of up to 20 stations.

3.1 THE FFT CARD

The next fundamental building block of the FFT section of the VLBA correlator after the butterfly ASIC is the FFT engine which performs the actual transforms. The system requires 160 of these engines to process all of the data from the playback system. A block diagram of an FFT engine is seen in figure 2. The operations performed in the circuitry represented by this figure include:

1. short term fringe phase generation
2. fringe rotation
3. windowing
4. FFT bit reversal
5. Fourier transformation
6. point shuffle for 2048-point transforms
7. fractional sample error correction phase ramp generation
8. fractional sample error correction
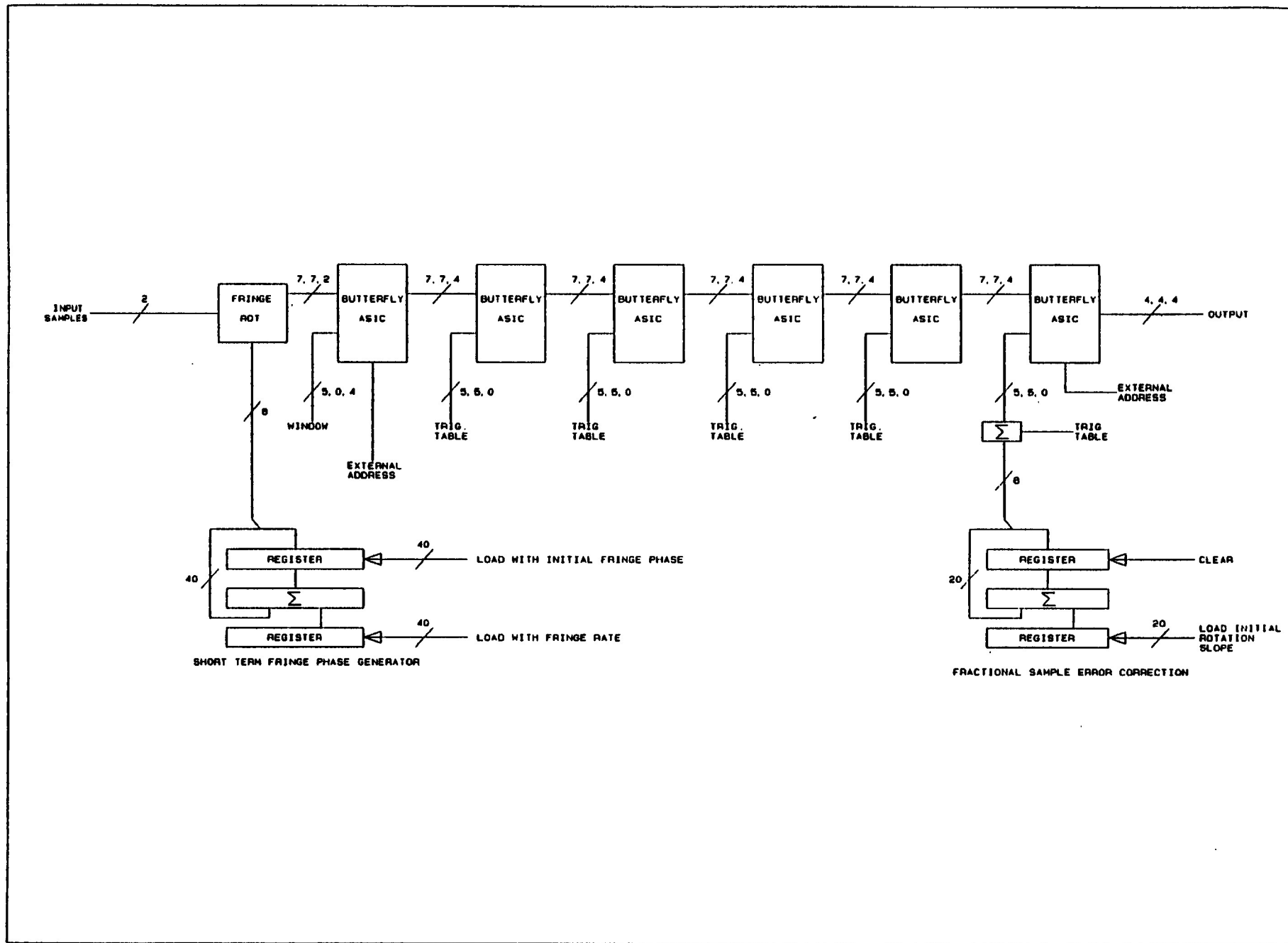9. drive into the cross-multiplier system

4

Figure 2

All of the functions above are represented in figure 2 with the exception of the 2048-point shuffle.

The short term fringe phase generation circuit seen in figure 2 is loaded every 4 msec with the coefficients necessary for a two term Taylor series approximation of the fringe phase. These coefficients are the initial fringe phase and the rate of change of phase. Both of these coefficients have 40-bit resolution and the Taylor series approximation is updated every 32-Mhz clock cycle. The 8 most significant bits of the resulting 40-bit fringe phase approximation are used to do the actual fringe rotation. This 8-bit truncated fringe phase and the 1 or 2 bit samples both drive a 1K ROM which provides a look-up table complex multiplication. Given an 8-bit fringe phase and a 2-bit sample, the ROM outputs a 16-bit (7,7,2) complex result. The fringe rotation look-up table also has implicit in it the conversion from the sample format to a binary weighted code necessary for processing in the arithmetic of the FFT. The 2-bit exponent range of the rotated sample is sufficient to express the optimum weighting for two bit sampling (see correlator memo 75).

The window function is seen in figure 2 as applied at the twiddle factor input to the first butterfly stage of the FFT. In a decimation-in-time FFT, the twiddle factors for the first butterfly stage are all unity. This fact allows the use of the twiddle factor input to the butterfly ASIC chip, which would normally do a complex multiplication of incoming points, to do the real multiplication required by a window tapering function. Since the window weights are real, the quantization into the first stage butterfly has 5,0,4 resolution (actually, since the weights are always positive, the effective resolution comes out to be 4,0,4).

The bit reversal operation required in performing fast Fourier transforms is accomplished at the input to the first butterfly ASIC as the incoming points are stored in the input RAM buffer. An external address is supplied to the first stage input buffer and the bit reversal requirement is satisfied by the RAM address sequence.

All of the butterfly stages in the FFT engine past the first stage have 5,5,0 twiddle factors applied at the second input port of the butterfly ASIC as seen in figure 2. Simulations of fast Fourier transforms using such low resolution trig. table quantizations have shown poor performance and to reach the performance levels required by the VLBA correlator, the trig. table will consist of 32 actual tables that average to 10-bit resolution. These 32 tables will drive the FFT engines in time sequence (performing an integral number of passes through all 32 tables in one 100 msec short term integration) to improve performance.

Also seen in figure 2 is the fractional sample error correction logic. The fractional sample error will be tracked in the delay model generator. Every 1024 clock cycles, the fractional sample error will be used to deliver a 20-bit rotation slope to the NCO seen in figure 2. As the spectral points are input to the final butterfly stage, each is rotated by a phasor generated by the NCO. This phasor starts at zero for the first spectral point and ramps up linearly through the spectrum. Thus, the first point of a spectrum is rotated through zero degrees, the next through one times the phase slope, and the Nth point of the spectrum through N times the phase slope.

The actual FFT card has 4 FFT engines on it and 40 FFT cards will be required to build the correlator. Figure 3 gives a block diagram of an FFT card and its 4 FFT engines with the twiddle factor inputs, fringe rotation, and fractional sample correction logic omitted for clarity. The cross-coupling interconnection just before the last butterfly ASIC does the above mentioned 2048-point shuffle. This 2048-point FFT requirement and the system interconnection requirements into the cross-multiplier system specify that the four FFT engines on a card be driven by two adjacent channels of two adjacent stations (for example channels 4 and 5 of stations 16 and 17).

3.0 THE CROSS-MULTIPLIER SYSTEM

The cross multiplier system will be driven by the output of the FFT system. As data points proceed down the FFT engine, they are represented in a 7,7,4 format. Once the FFT has been performed, however, the need for computational precision is lessened. The samples entered the FFT at the 2-bit quantization level (at most) and information is not gained via the FFT, the information in the samples is just rearranged by the FFT. Hence, once the possibility of errors introduced by low precision in the highly systematic FFT computation is over, cross multiplication may proceed at a lower resolution. Thus, in order to save wires in the FFT system to cross multiplication system interconnect, the data resolution will be scaled back to 4,4,4 at the cross multiplier input.

The cross multiplier system is conceptually very simple, consisting of 8 triangular arrays of multipliers. Each array is a 20 by 20 half matrix (including the diagonal) in which the spectral output of one channel from each of the 20 stations will drive one column and one row. In such a half matrix, each station spectral point will be multiplied (complex conjugate multiplied) by the same spectral point from every other station (auto-multiplication occurs on the diagonal of the half matrix).
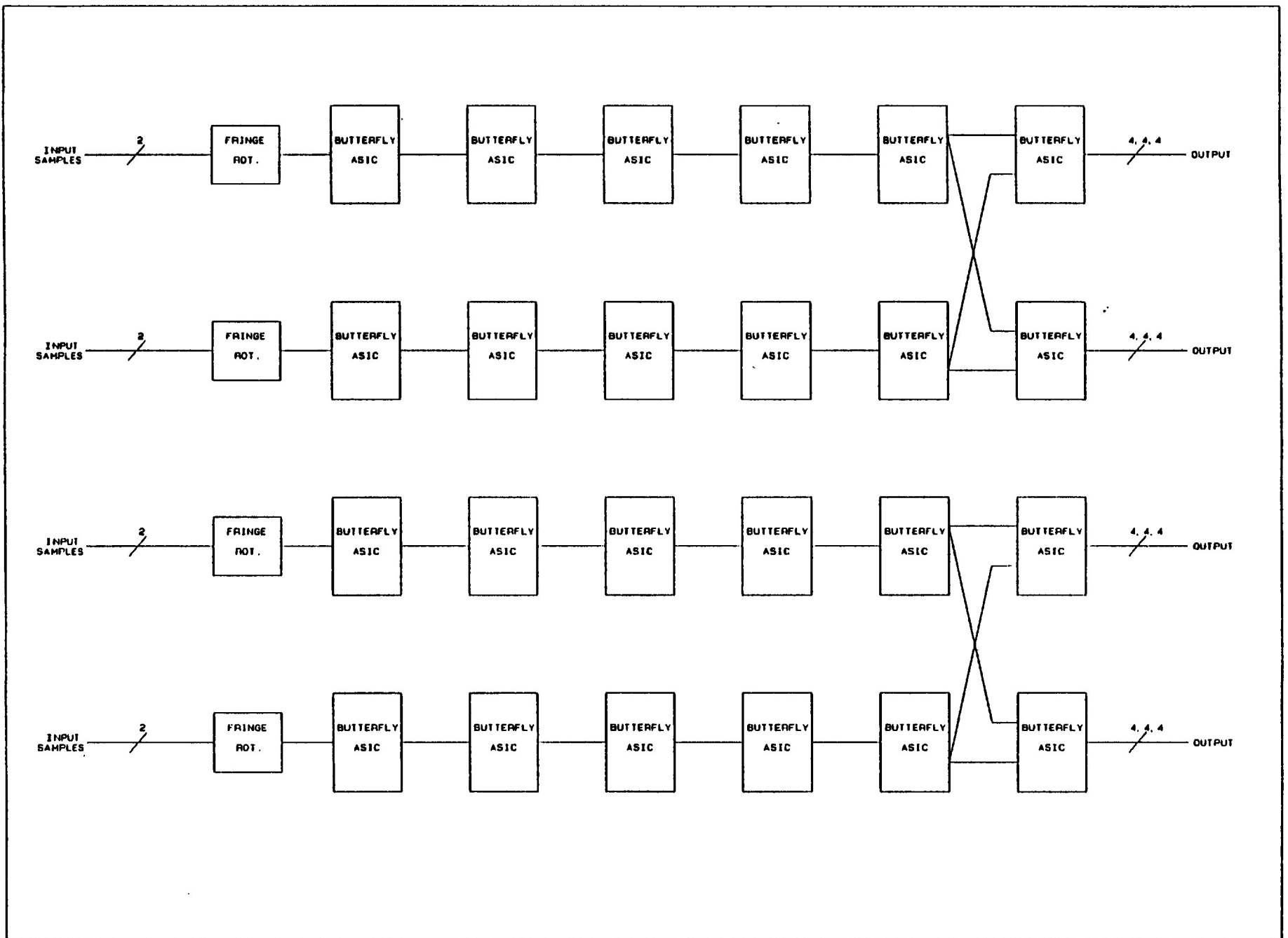
6

Figure 3

In practice, the only complication to the simplicity stated above is the requirement to form polarization cross products. This requirement precipitates two responses in the hardware of the correlator. First, the 8 half matrix multiplier arrays described above will be divided into pairs that are located in close proximity to each other. This action is taken to lessen the cable interconnect requirements. The arrays pairs will service channels of the VLBA that, for polarization observations, will be opposite polarizations of the same frequency band and hence all the spectral points that must be multiplied together (cross hand and parallel hand multiplications) will be near each other.

The second response to the polarization requirement is the action described in the section on the ASIC cross multiplier application (section 2.2). When polarization cross products are required, the total number of multiplications per second in the cross multiplier section doubles. This doubling in not a problem for the complex multiplier itself since normally each cross multiplier chip need only perform a complex multiplication every two clock cycles. The problem comes in accessing the RAM accumulator. For polarization observations, there must be two RAM operations per clock cycle, one read and one write which is beyond the performance level of the ASIC RAMs. The solution to this problem is to set the maximum transform size for polarization observations at 512 points. At 512 points, the ASIC chip buffers can hold the points for two complete transforms. With two complete spectra available at the same time, the requirements for access to the accumulator RAMs can be diluted by cross multiplying point X of spectra # 1, adding the result into the accumulated partial sum and then cross multiplying point X of spectra # 2 and adding this result into the same accumulator partial product before storing the new partial result. Thus the need for RAM access is reduced by half.

3.1 THE CROSS MULTIPLIER CARD

The basic component of the cross multiplication system is the cross multiplier card.

Figure 4 gives a block diagram of the cross multiplier card (with examples of antenna numbers and IF channels provided in the input capture registers of the card). This card is a dual 4 X 4 matrix of ASIC chips. The output spectra of four FFT engines will drive the four column inputs of one of the matrices and the output of four other FFT engines (on the card array diagonal they will be the same) will drive the four row inputs. Fifteen of these cards are required to make up two complete 20 by 20 arrays of cross multipliers. Cards located on the diagonal of the resulting triangular array can be partially populated with ASIC chips to save money.
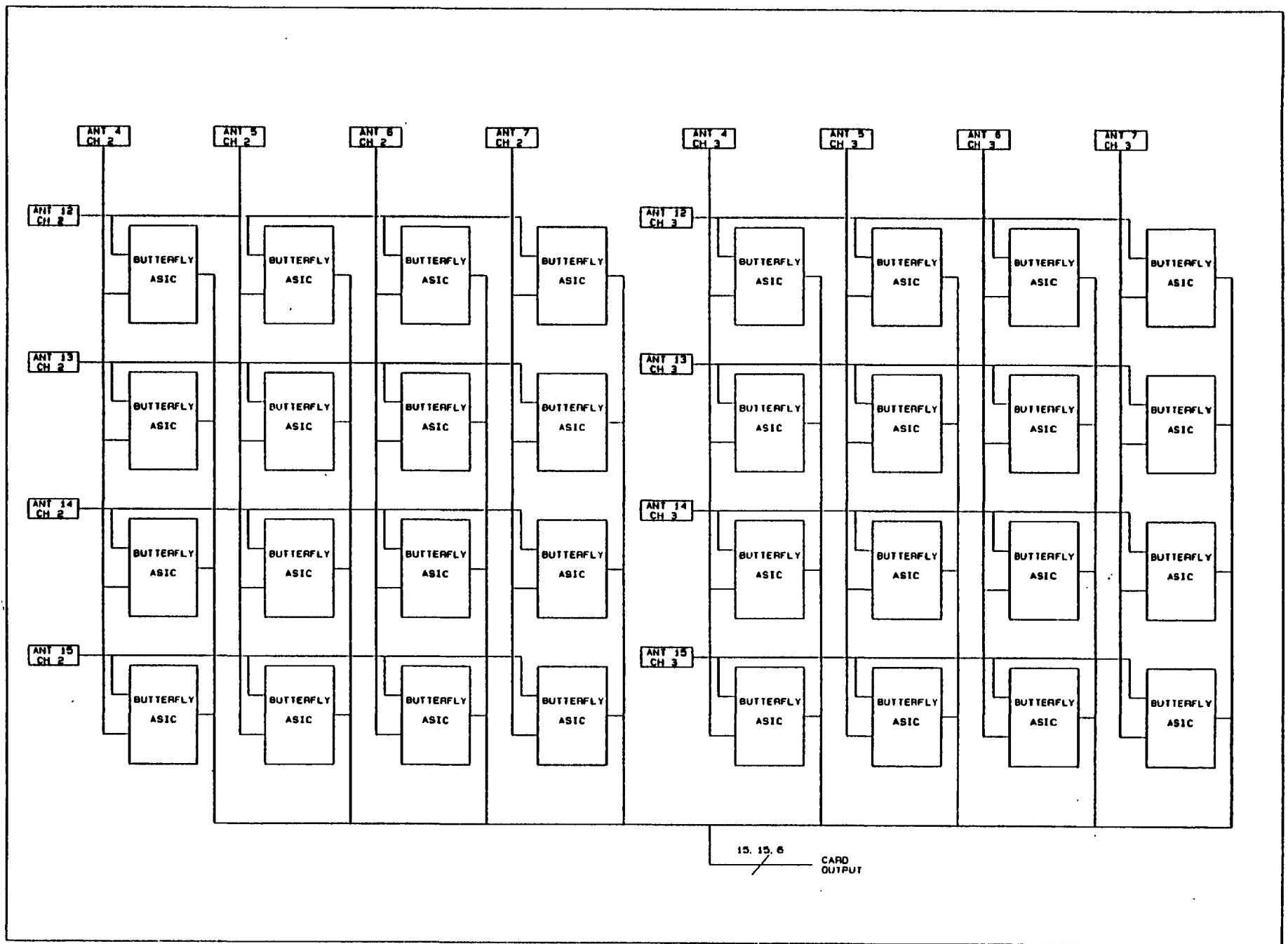
7

Figure 4

The complete VLBA correlator system requires 8 20 by 20 multiplier arrays, and hence 60 of the cards seen in figure 4 are required in the final system.

Figure 4 is actually a simplified diagram that shows how the card is normally used. Figure 5 illustrates how the same card of figure 4 is in fact connected. The row inputs, instead of being segregated by channel between the two 4 by 4 matrices of ASICs as shown in figure 4, interconnect between the two arrays. By interconnecting the two 4 by 4 arrays, polarization observation requirements are provided for. Every two clock cycles of a polarization observation, one ASIC will clock in a point from the L channel of station A and points from both the L and R channel of station B. A second ASIC (that could be on a different card) will clock in a point from the R channel of station A and a point from both the R and L channels of station B. On one clock cycle the two chips perform the RR and LL multiplications respectively and on the next clock the RL and LR multiplications (recall that the spectral points exit the FFT engine and enter the cross multiplier at one point every two clock cycles).

The short term accumulator contains two banks of 512 X 36 bits to double buffer accumulation results. Every 100 mseconds, a bank switch is performed and in the 100-mseconds that follow, data from the double buffered bank is output into the long term accumulator. Figure 4 shows the outputs of all ASIC chips tri-stated together to allow RAM readout (figure 5 omits the output connections for clarity). In actually, the outputs from the two 4 by 4 matrices may be kept separate and each card may have two outputs.
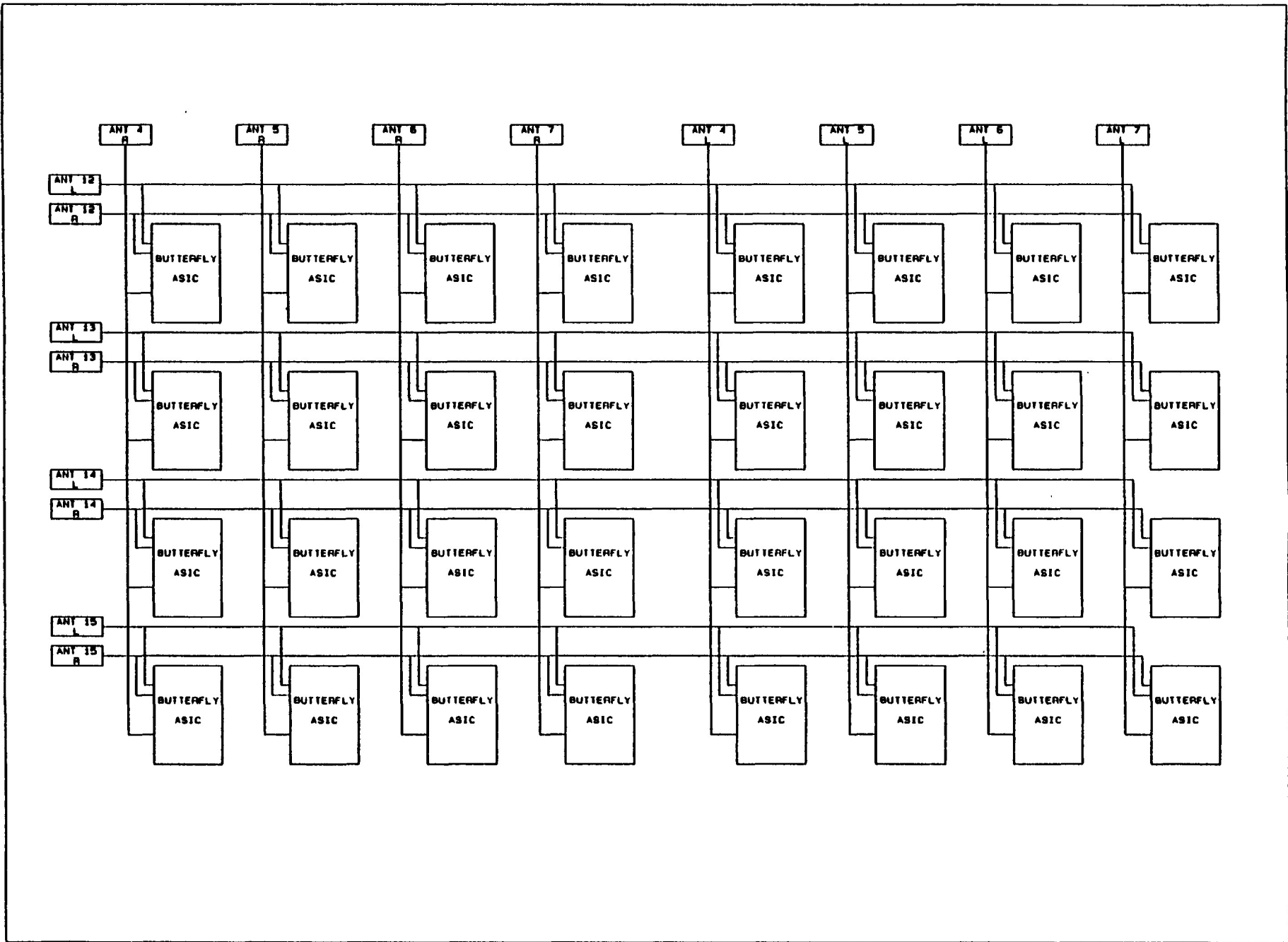
Figure 5