A Low-Power Correlator ASIC For Arrays With Many Antennas

Larry R. D'Addario and Douglas Wang
Jet Propulsion Laboratory, California Institute of Technology
Pasadena, California, USA
email: ldaddario@jpl.nasa.gov

***Abstract*—We report the design of a new application-specific integrated circuit (ASIC) for use in radio telescope correlators. It supports the construction of correlators for an arbitrarily large number of signals. The ASIC uses an intrinsically low-power architecture along with design techniques and a process that together result in unprecedentedly low power consumption. The design is flexible in that it can support telescopes with almost any number of antennas N. It is intended for use in an "FX" correlator, where a uniform filter bank breaks each signal into separate frequency channels prior to correlation.**

I. Introduction

We have developed a new application-specific integrated circuit (ASIC) that allows the construction of correlators with far lower power consumption than previously and that supports correlation of almost any number of signals. It is based on the low-power architecture described in [1]. The chip implements one X-unit of an FX correlator, processing all $2N$ signals from N dual-polarization antennas for a portion of the bandwidth. The basic concept is shown in Figure 1. The chip contains

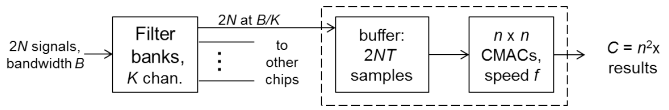

Fig. 1. Power-optimized correlator architecture. The ASIC is shown in the dashed box. Each chip processes all $2N$ signals using an array of n^2 CMACs, but it does so for only a fraction of the bandwidth. A filter bank (not part of the correlator chip) breaks each input signal into frequency channels narrow enough so that one chip can process all signals.

an array of $n^2 = 4096$ complex multiply-accumulate units (CMACs). This is sufficient for all correlations of 64 signals ($N = 32$) to be computed in parallel. When N is larger, input data are buffered in an on-chip memory and the CMACs are reused as many times as necessary to compute all correlations. Each such computation of 4096 correlations is called a sub-integration (SI). Since $2N^2$ correlations (including self-correlations) are needed, $x = 2(N/n)^2$ SIs are required. If the CMACs run at clock rate f, the chip can process bandwidth $b = f/x$. The memory must hold T samples of each signal in order to compute an integration of length T.

The design has been synthesized and simulated so as to obtain accurate estimates of the chip's size and power consumption. It is intended for fabrication in a 32 nm silicon-on-insulator process, where it will require less than 12 mm2 of silicon area and achieve an energy figure of merit (FoM) of 1.76 to 3.5 pJ per CMAC operation, depending on the number of antennas. Operation has been analyzed in detail up to $N = 4096$. The system-level energy efficiency, including board-level I/O, power supplies, and controls, is expected to be 5 to 7 pJ. Existing correlators for the JVLA ($N = 32$) [2] and ALMA ($N = 64$) [3] telescopes achieve about 5000 pJ and 1000 pJ respectively using ASICs in older technologies. To our knowledge, the largest-N existing correlator is LEDA at $N = 256$ [4]; it uses GPUs built in 28 nm technology and achieves about 1000 pJ. Correlators being designed for the SKA telescopes (N=128 and N=512) using FPGAs in 14 nm or 16 nm technologies are predicted to achieve about 100 pJ.

We believe that this analysis provides reliable estimates of the performance, but physical design has not yet been completed and devices have not yet been fabricated.

II. Performance

Figure 2 summarizes the performance for $N = 32$ and all mulitiples of 64 up to 4096. Results are computed for

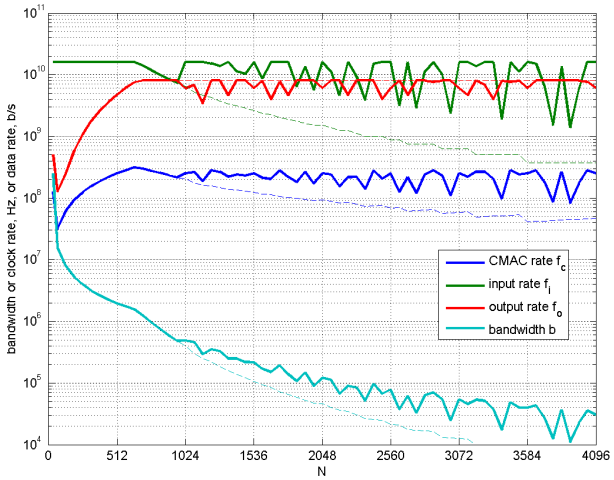

Fig. 2. Performance of the chip as a function of number of antennas N. Bandwidth, I/O rates, and CMAC rate are plotted. For $N \geq 1024$, performance is improved by breaking the processing into partial integrations. Dashed lines show the performance that can be obtained if all signals are processed at once.

$N = 32$ and each multiple of 64 up to 4096. It is possible to maintain nearly the maximum throughput (as measured by the CMAC clock rate) over the entire reange. There are some unfavorable values of N where the device cannot be used as efficiently, such as when $N/64$ is a prime number, but even

then the performance remains good. For an ideal device (where the processing rate remains constant and there is negligible overhead), the bandwidth decreases as $1/N^2$. Fig. 2 shows that the upper envelope of bandwidth vs. N approaches this limit. Figure 3 shows the corresponding energy FoM vs. N. For example, at $N = 512$ the chip is able to process a bandwidth of $b = 1.953\,\text{MHz}$ while dissipating $P = 1.824\,\text{W}$, so the FoM is $2N^2b/P = 1.78\,\text{pJ}$. To process a total bandwidth of 500 MHz, 256 of these chips would be needed.

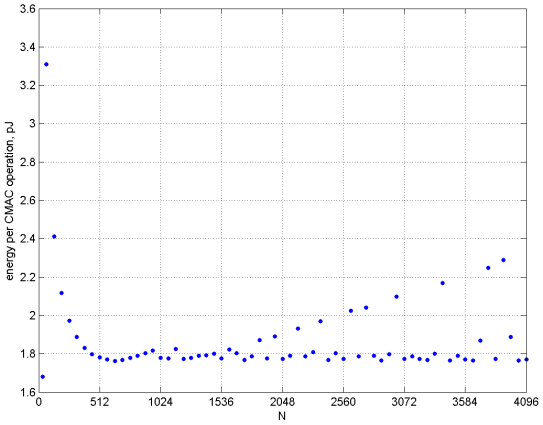

Fig. 3. Energy per CMAC operation (figure of merit for power) vs. N. This number remains near its best value of 1.76 pJ for most N. The result here includes only the ASIC power; additonal power is used by other devices that are needed for integrating the chips into a system.

Performance is limited by input/output bandwidth rather than by internal speeds. For cost reasons, the input and output rates are limited to 16 Gb/s and 8 Gb/s respectively. This can be seen in Fig. 2. At $N \leq 640$, the chip is input rate limited, otherwise it is output rate limited. Higher bandwidth per chip could be achieved if these rates were increased; that would reduce the number of chips needed in a large system, but it would have little effect on the system power.

III. Selected Design Details

We mention only a few important points in this summary paper. A more complete paper is in preparation.

Input data consist of complex signal samples represented as 4 b+4 b twos-complement numbers (real+imaginary) and the results are delivered as 16 b+16 b complex numbers.

Figure 4 is a block diagram of the chip, showing its major modules and all input/output ports. Input data are received

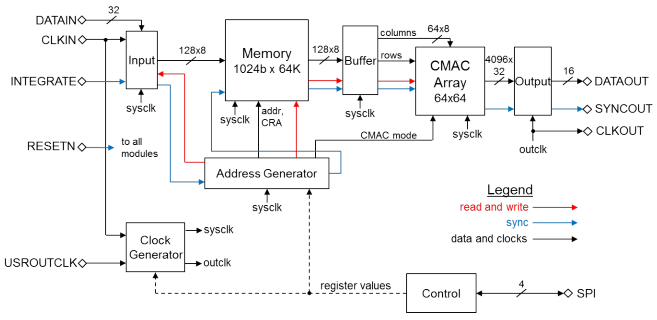

Fig. 4. Block diagram of the correlator ASIC.

as 32 bits in parallel on each rising edge of CLKIN, and consist of four 8 b samples of different signals at the same sampling time. Assertion of INTEGRATE marks the first data of a new integration. Successive input words supply samples from other signals and other sampling times in a particular order until all $2NT$ samples of one integration have been received and written to the memory. The memory stores data as 1024 b words or 128 samples. The address generator module determines the sequencing of write and read cycles and provides the appropriate addresses for each. It also provides synchronization signals that organize the chip's computations into sub-integrations. On memory read cycles, the RAM delivers 1024 b words to a small buffer which organizes the data into 64×8-bit words corresponding to the rows and columns of the CMAC array. The CMAC array module performs the correlations. At the end of an SI, 4096 complex results are available in the CMAC readout registers. These are delivered sequentially by the output module to 16 parallel output pins synchronous with 8192 cycles of CLKOUT. SYNCOUT is asserted during the first CLKOUT cycle of each SI. The frequency of CLKOUT must be large enough to read out all CMACs before the next SI is finished. The performance results assume that the maximum rate at each input or output pin is 500 MHz. Simulations show that this is easily achieved.

The memory and CMAC array operate on an internal clock (sysclk), synthesized from INCLK by the clock generator module using a phase locked loop (PLL). OUTCLK can also be synthesized by the clock generator, or it can optionally be supplied at input USROUTCLK. Each of these has a minimum required frequency, but operation of the chip is driven entirely by INCLK.

A set of twelve 20 b control registers can be written and read via a serial peripheral interface (SPI) slave port. These set up the device for a particular application by specifying the values of N and T and the PLL parameters that determine the clock frequencies. For each SI, three registers must be written with the starting memory addresses for reading row data, reading column data, and writing new data during the next SI. This arrangement gives the chip the flexibility to work in a wide variety of applications.

Acknowledgment

This work was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration and funded through the internal Research and Technology Development program.

References

[1] L. D'Addario, "Low-Power Correlator Architecture For the Mid-Frequency SKA," SKA Memo 133, 2011 March 21. Available at http://www.skatelescope.org/memos.

[2] M. McKinnen, private communication (email of 2010 Feb 11, "EVLA correlator power").

[3] A. Wooten and A. R. Thompson, "The Atacama Large Millimeter/Submillimeter Array." *Proc. IEEE*, **97**, 1463–1471, 2009.

[4] J. Kocz, L. J. Greenhill, B. R. Barsdell, *et al.*, "Digital Signal Processing Using Stream High Performance Computing: A 512-Input Broadband Correlator for Radio Astronomy." *J. of Astron. Instr.*, **4**, 1550003, 2015.