23.6 An Asynchronous Array of Simple Processors for DSP Applications

Zhiyi Yu, Michael Meeuwsen, Ryan Apperson, Omar Sattari, Michael Lai, Jeremy Webb, Eric Work, Tinoosh Mohsenin, Mandeep Singh, Bevan Baas

University of California, Davis, CA

Modern applications increasingly require the computation of DSP workloads comprised of a variety of numerically-intensive DSP tasks. These workloads are found in communication, multimedia, embedded, and wireless applications, and often require very high levels of computation and high energy efficiency.

The Asynchronous Array of Simple Processors (AsAP) uses processor cores with small instruction and data memories to dramatically reduce area and power while increasing performance. Fig. 23.6.1 shows the architecture of an individual AsAP processor and the 6×6 array contained on the chip. Data enters the array through the upper left processor and exits from one of the right-column processors. Each processor has a simple RISC-style design with a nine-stage pipeline, detailed in Fig. 23.6.2, and contains two input ports, one output port, and a local clock oscillator. The 16b fixed-point datapath includes a 40b accumulator and executes 54 32b instructions, among which only the bit-reverse instruction is algorithm specific.

Memories in modern programmable processors typically consume large amounts of area and power, and frequently contain the critical path. Therefore, energy efficiency and speed can be dramatically increased with small memories. Two reasons make it possible to use very small memories in DSP processors. First, although it is possible to write code that requires large amounts of data and instruction memory, Fig. 23.6.3 shows that many common DSP tasks do not require large memory sizes. Second, a system with many processors can use individual processors or groups of processors to compute individual tasks and intermediate data can be transmitted between processors with a small memory requirement. Fig. 23.6.3 illustrates this fine-grain streaming processor approach. AsAP processors contain a 128×16b data memory and a 64×32b instruction memory, and easily contain many DSP tasks.

Each AsAP processor contains an independent clocking system including its own clock oscillator. The maximum clock tree span of less than 1mm in 0.18µm CMOS greatly simplifies the design process. The programmable ring oscillators are built from standard cells and change their frequency through parallel tri-state inverters in each stage, a 5 or 9-stage configuration, and a divider that ranges from 1 to 128, as shown in Fig. 23.6.4. Processors are clocked asynchronously with respect to each other and clocking circuits permit oscillators to operate at any frequency less than the maximum, as well as allowing arbitrary bursts and stalls.

The Globally Asynchronous Locally Synchronous (GALS) architecture and oscillator design enable each processor to completely power down (leakage only) if no work is available for nine clock cycles. When work becomes available, the clock is restored in less than one cycle. Clock stalling typically reduces power by more than 50% for complex applications. This clocking strategy also enables frequency adaptable operation to reduce power and to compensate for process and workload variations. The measured oscillator frequency covers a wide range from 1.66MHz to 702MHz. The maximum gap between useful samples (from 1.66MHz to 500MHz) is only 0.08MHz. To simplify the physical design, clock gating was not used in this prototype, and almost 2/3 of the power dissipation is due to the clocking system—so the future addition of clock gating is expected to greatly reduce power consumption.

Processors communicate only with adjacent processors to permit full-rate communication with low energy. Movement of one data word from one processor's memory to an adjacent processor's memory requires only 220pJ at 1.8V or 68pJ at 0.9V. Each processor can receive inputs from up to two adjacent processors, and can send output to any dynamically-configurable combination of its four neighbors through software.

Communication across clock boundaries is accomplished by 32-word dual-clock FIFOs. Fig. 23.6.4 shows the configurable synchronization registers and gray coding of read and write address information circuits that avoid metastability and permit reliable operation with completely arbitrary clock operation in both sending and receiving clock domains.

Figure 23.6.7 shows the first generation 36-processor standard-cell-based AsAP chip implemented using TSMC 0.18µm CMOS. Each processor occupies 0.66mm^2 and processors nearly directly abut each other. The chip is fully functional on first-pass silicon at 475MHz at 1.8V, and 116MHz at 0.9V near room temperature. Average typical power is 32mW at 1.8V and 475MHz while executing applications. The maximum power is 144mW per processor at 475MHz, resulting in a maximum energy dissipation of 0.3mW/MHz. At 0.9V and 116MHz, the maximum energy dissipation is 0.093mW/MHz. As shown in Fig. 23.6.5, AsAP is at least 20 times smaller and at least 14 times more energy efficient than other processors. A comparison of area utilization shows that AsAP uses a favorable 66% of its area for the core processor.

Some complex applications have been programmed and measured on the AsAP chip. The programs are lightly optimized and unscheduled. A nine-processor JPEG encoder core is shown in Fig. 23.6.6 and dissipates 224mW and has a throughput of 215k 8×8 pixel blocks per second at a clock rate of 300MHz. This throughput is more than nine times faster than implementations built on traditional architectures [1]. A 22-processor 802.11a/802.11g transmitter is fully-compliant with the IEEE standard over all eight rates [2] and contains additional upsampling, filtering, and synchronization functions not required by the standard. The transmitter dissipates 407mW and has a throughput 30% of the 54Mb/s requirement at a clock rate of 300MHz. This result compares well with a reported throughput of 1.7Mb/s for a 24Mb/s rate on a TMS320C6201 [3]. A configuration test bit allows clock oscillator stalling to be disabled, which shows that oscillator stalling reduces JPEG power by 53% and 802.11a/g power by 65%.

Acknowledgments:
The authors acknowledge support from Intel, UC Micro, NSF Grant No. 0430090, MOSIS, and a UCD Faculty Research Grant; and thank R. Krishnamurthy, M. Anders, S. Mathew, S. Muroor, W. Li, C. Chen, and D. Truong.

References:
[1] K. Sakiyama et al., "Finding the Best System Design Flow for a High-Speed JPEG Encoder," *Proc. DAC*, pp. 577-578, Jan.,2003.
[2] M. Meeuwsen et al., "A Full-Rate Software Implementation of an IEEE 802.11a Compliant Digital Baseband Transmitter," *Proc. SIPS*, pp. 124-129, Oct., 2004.
[3] M. Tariq et al., "Development of an OFDM based high speed wireless LAN platform using the TI C6x DSP," *Proc. ICC*, pp. 522-526, Apr., 2002.
[4] A. Bright et al., "Creating the BlueGene/L Supercomputer from Low-Power SoC ASICs," *ISSCC Dig. Tech. Papers*, pp. 188-189, Feb., 2005.
[5] S. Agarwala et al., "A 600MHz VLIW DSP," *ISSCC Dig. Tech. Papers*, pp. 56-57, Feb., 2002.
[6] B. Flachs et al., "A Streaming Processing Unit for a CELL Processor," *ISSCC Dig. Tech. Papers*, pp. 134-135, Feb., 2005.
[7] M. Taylor et al., "Evaluation of the RAW Microprocessor: An Exposed-Wire-Delay Architecture for ILP and Streams," *Proc. ISCA*, pp. 2-13, June, 2004.
[8] R. Witek et al., "StrongARM: A High-Performance ARM Processor," *Proc. CMPCON*, pp. 188-191, Feb., 1996.

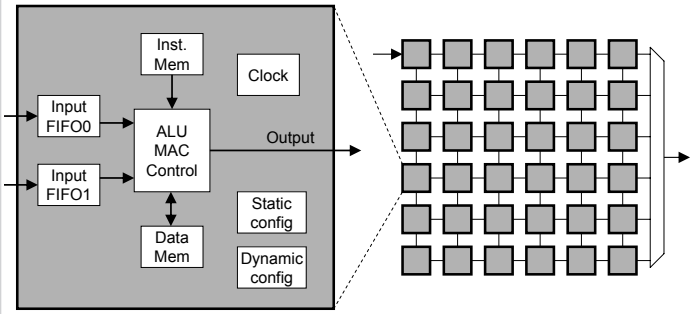


Figure 23.6.1: AsAP block diagram.

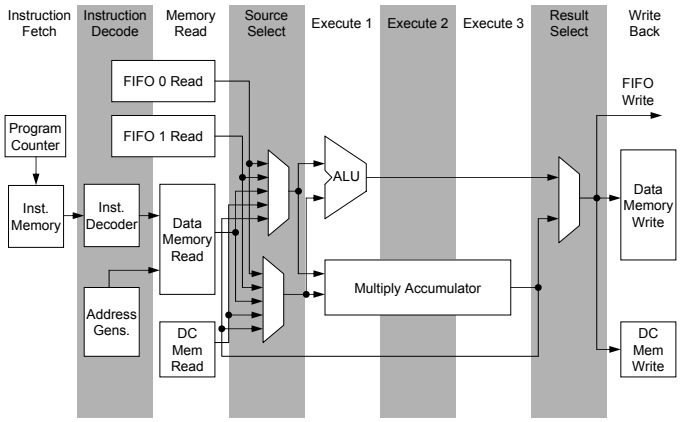


Figure 23.6.2: AsAP 9-stage pipeline.

Task	Inst. Mem requirement (words)	Data Mem requirement (words)
N-pt FIR	6	2*N*
8-pt DCT	40	16
8x8 2-D DCT	154	72
Conv. coding (*k* = 7)	29	14
Huffman DC encode	107	24
Huffman AC encode	96	310
N-pt convolution	29	2*N*
64-pt complex FFT	97	192
Bubble sort	20	1
N merge sort	50	*N*
Square root	62	15
Exponential	108	32

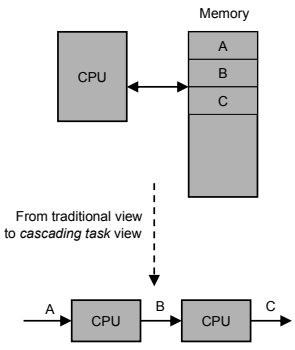


Figure 23.6.3: Memory requirements for common DSP tasks and cascading task data flow.

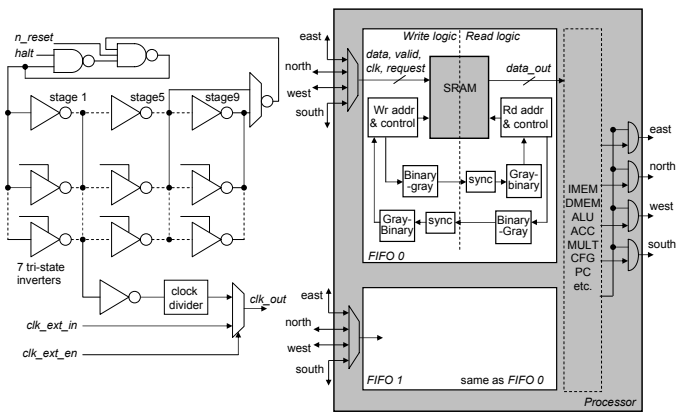


Figure 23.6.4: Programmable clock oscillator and inter-processor communication diagram.

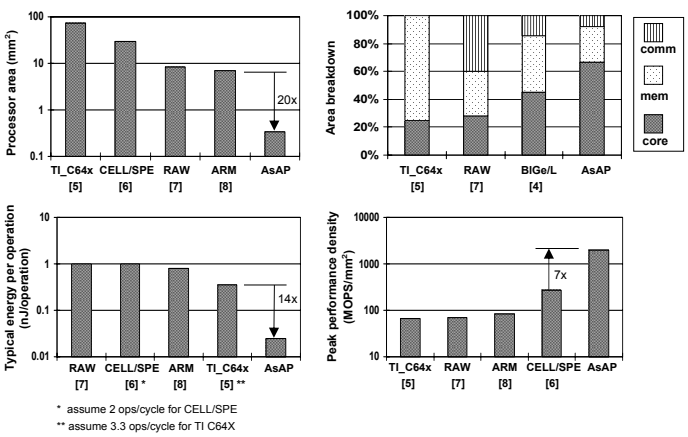


Figure 23.6.5: Area, power and performance of several processors scaled to 0.13µm.

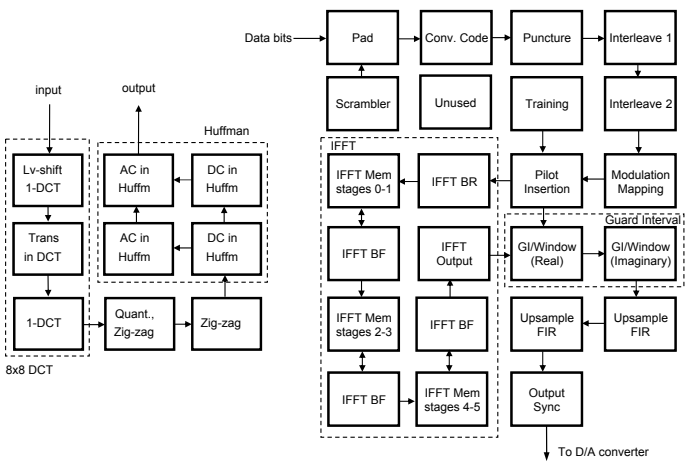
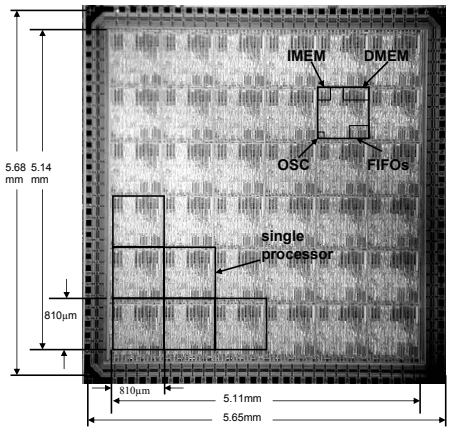


Figure 23.6.6: JPEG and 802.11a/802.11g transmitter implementations.

Continued on Page 663



Technology	TSMC 0.18µm
Transistors	8.5 Million
Area	5.68mm x 5.65mm = 32.1mm²
Area, 1 proc	0.66mm²
Max speed	475MHz @ 1.8V 17.1GMAC/s
Typical power	32mW @ 1.8V @ 475MHz (1 proc)
Max power	144mW @ 1.8V @ 475MHz (1 proc)
Max energy/op	300pJ/Op = 0.3mW/MHz @ 1.8V 93pJ/Op = 0.093mW/MHz @0.9V (1 proc)
Package	PGA121M

Figure 23.6.7: Chip micrograph of the 6×6 AsAP array.

