

A 5.8 pJ/Op 115 Billion Ops/sec, to 1.78 Trillion Ops/sec 32nm 1000-Processor Array

Brent Bohnenstiehl, Aaron Stillmaker, Jon Pimentel, Timothy Andreas, Bin Liu, Anh Tran,
Emmanuel Adeagbo, Bevan Baas

University of California, Davis

Abstract

1000 programmable processors and 12 independent memory modules capable of simultaneously servicing both data and instruction requests are integrated onto a 32nm PD-SOI CMOS device. At 1.1 V, processors operate up to an average of 1.78 GHz yielding a maximum total chip computation rate of 1.78 trillion instructions/sec. At 0.84 V, 1000 cores execute 1 trillion instructions/sec while dissipating 13.1 W.

Introduction

Modern semiconductor fabrication technologies now enable the construction of integrated circuits containing over 1000 processors on a single chip [1]. However, for such systems to effectively compute workloads, new architectures are needed for the processors, the inter-processor interconnect, circuits that interact with larger memories, and the applications they execute [2–4]. KiloCore’s 1000 MIMD processors are arrayed in 32 columns and 31 rows with 8 processors and 768 KB inside 12 independent memories in a 32nd row (Fig. 1). While caches are generally extremely effective, they are problematic for resources such as chip I/O and cache-coherency circuits when the number of processors is scaled into the 100s or 1000s and they also dissipate significant power [5]. In contrast, KiloCore processors do not contain explicit caches and instead store data and instructions inside i) local memory, ii) an arbitrary number of nearby processors, iii) on-chip independent memory modules, or iv) off-chip memory.

KiloCore Architecture

Each processor issues one in-order instruction per cycle into its 7-stage pipeline from either its 128 x 40-bit local instruction memory or an independent memory module. None of the 72 supported instruction types are algorithm-specific. Processor data memories are implemented as two 128 x 16-bit banks to sustain a throughput of one instruction per cycle for common instructions that require two source operands (Fig. 2). However, profiled code of five disparate applications (AES encryption, low-density parity-check (LDPC) decoder, 100-byte database record sorting, 802.11a/g receiver, and software single-precision floating-point arithmetic implementations) showed that only 0.34% of operands could not be mapped to an address in only one bank and thus needed to be written to both banks redundantly to avoid conflicts during subsequent reads. Our scheme permits conflict-free addressing with optimal memory space maximization.

Communication on-chip is accomplished by a high-throughput circuit-switched network and a complementary very-small-area packet-switched network (Fig. 3). The source-synchronous circuit-switched network supports communication between adjacent and distant processors, as resources allow, with each link supporting a maximum rate of 28.5 Gbps. Execution of an instruction with an input operand transferred by the circuit-switched network from an adjacent processor dissipates 16% less energy compared to if it were transferred from the local data memory. If an input operand comes from a processor ten processors away, only an additional 22% energy is required compared to a local access. Routers utilize wormhole routing, operate autonomously, and

contain a 4 x 18-bit buffer on each of the five input ports—one for each cardinal direction and one for the host processor [6]. Maximum throughput is 45.5 Gbps per router and 9.1 Gbps per port at 1.1 V. At 0.9 V, maximum throughput is 27.1 Gbps at 3.36 mW and at 0.67 V, it is 8.1 Gbps at 429 μ W. Both network types contribute to an array bisection bandwidth of 4.2 Tbps.

Each of the 12 independent memory modules contains a 64KB SRAM, services two neighboring processors, and supports 28.4 Gbps of I/O bandwidth. When streaming instructions to a processor, a dedicated control module takes over program control and branch prediction control from the processor to more efficiently execute across branches (Fig. 4).

KiloCore’s 1000 processors, 1000 packet routers, and 12 independent memories are clocked by local and completely-unconstrained (below the maximum operating frequency) clock oscillators that do not use PLLs and may change frequency, halt within 1-5 clock periods, and restart in less than one clock period to reduce power dissipation. Processors, routers, and memory modules with no work to do dissipate exactly zero active power (leakage only). At 0.9 V, idle processors leak 1.1% of their typical operating power. Information reliably crosses unrelated clock domains through dual-clock FIFO buffers [7].

Programming is accomplished by a multi-step process including a mapping step that assigns programs to processors (Fig. 5). New mappings may be computed during runtime for purposes such as: simultaneous execution of unrelated workloads, optimizing mappings with consideration of PVT variations, avoiding faulty or partially-functional processors, or for self-healing for failures due to wear-out effects.

Chip Design and Measured Results

The chip was fabricated in a 32nm PD-SOI technology. The entire array is 7.94 mm by 7.82 mm and contains 621 million transistors. Except for the 64 KB SRAMs inside the independent memory modules, all memories are built from clock-gated flip-flops with synthesized interfacing logic which greatly simplifies the physical design and likely lowers the minimum operating voltage. Each processor contains 575,000 transistors and occupies 239 μ m by 232 μ m; therefore 18 processors occupy 1 mm².

Processor maximum operating frequencies range from 1.70 GHz to 1.87 GHz with an average of 1.78 GHz at 1.10 V. The KiloCore chip is flip-chip mounted by 564 C4 solder bumps inside a stock BGA package that delivers full power to only the approximately 160 central processors; therefore, a maximum execution rate of 1.78 trillion MIMD operations per second per chip (Fig. 6) is achievable only with a custom-designed package. At a supply voltage of 0.56 V, processors dissipate 5.8 pJ per operation at 115 MHz, which enables a chip to process 115 billion operations per second while dissipating only 1.3 W (Table 1). Processors achieve their optimal energy times time of 11.1 (pJ x ns)/op at a voltage of 0.9 V. Independent memories operate from 1.77 GHz at 1.1 V down to 675 MHz at 760 mV. Routers operate from 1.49 GHz at 1.1 V down to 262 MHz at 665 mV.

Acknowledgments

This work was supported by DoD and ARL/ARO Grant W911NF-13-1-0090; NSF Grants 0903549, 1018972, 1321163, and CAREER Award 0546907; SRC GRC Grants 1971 and 2321, and CSR Grant 1659; and C2S2 Grant 2047.

References

- [1] S. Borkar, *DAC*, 2007, pp. 746–749.
- [2] S. Vangal et al., *ISSCC*, 2007, pp.98–589.
- [3] D. Truong et al., *JSSC*, vol. 44, no. 4, pp. 1130–1144, 2009.
- [4] M. Butts, *IEEE Micro*, vol.27, no.5, pp. 32–40, 2007.
- [5] M. Horowitz, *ISSCC*, 2014, pp. 10–14.
- [6] A. Tran, B. Baas, *TVLSI*, vol. 22, no. 6, pp. 1391–1403, 2013.
- [7] R. Apperson, et al., *TVLSI*, vol. 15, no. 10, pp. 1125–1134, 2007.

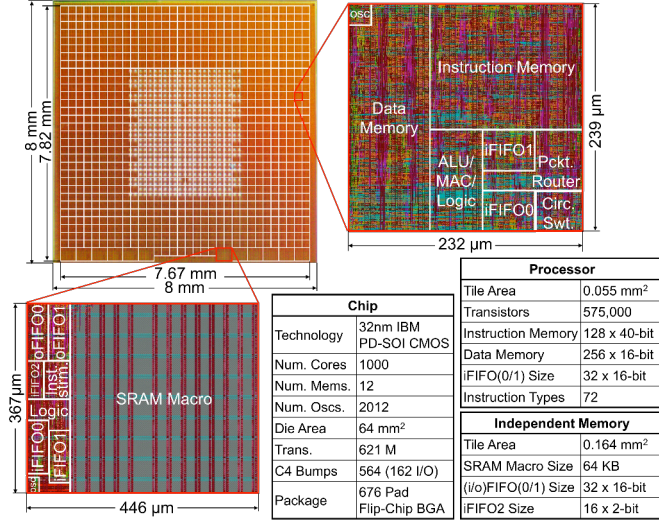


Fig. 1: Die photo of the KiloCore array, and annotated layout plots of a single processor tile and a single independent memory tile.

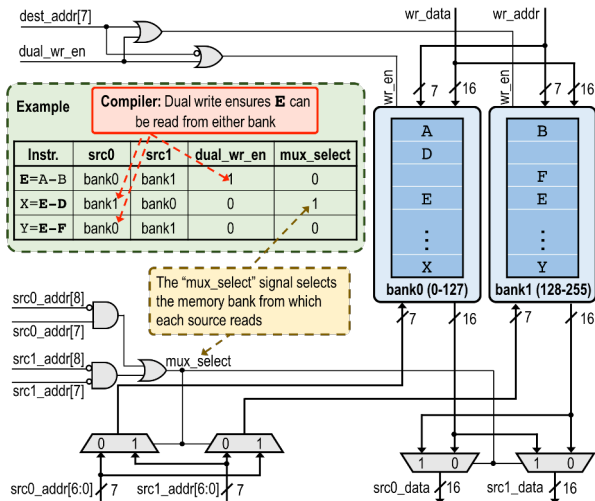


Fig. 2: Circuit diagram demonstrating dual memory bank reads and write backs.

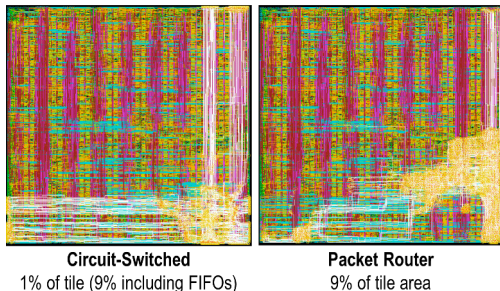


Fig. 3: Processor layout plots with inter-processor communication network circuits highlighted.

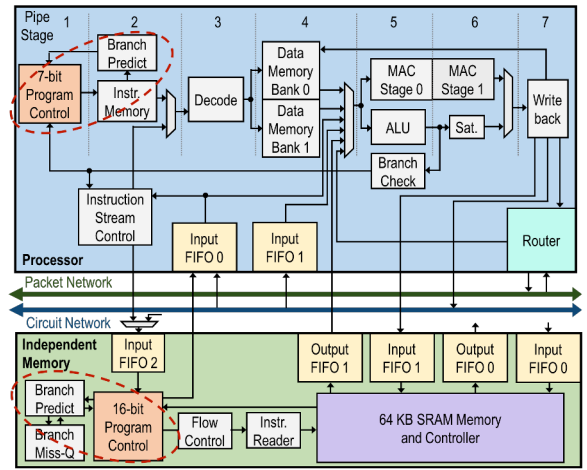


Fig. 4: Pipeline diagram of a KiloCore processor and an adjacent independent memory with program control highlighted in each block.

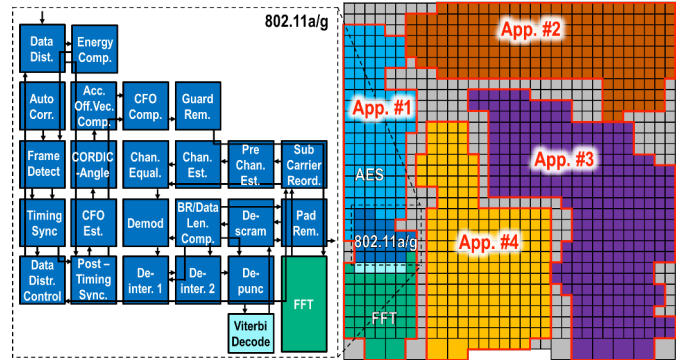


Fig. 5: Example multi-application mapping.

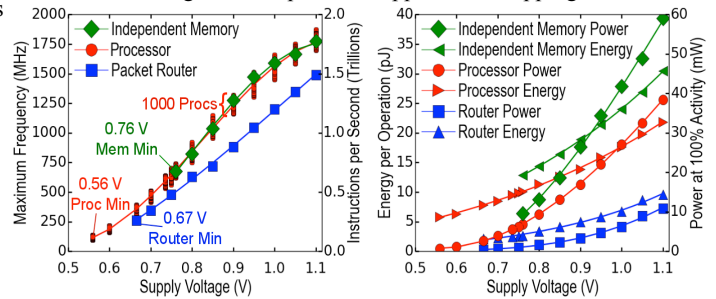


Fig. 6: Measured data at various supply voltages.

Table 1: Measured data and comparisons.

| | Proc. Count | Clock Freq (MHz) | Supply Voltage (V) | Per Proc. Power (mW) | Energy /Op (pJ) | ExT (pJ x ns) | Max. Bisect. BW (Tb/s) |
|--------------------------------|-------------|----------------------------|----------------------------|----------------------------|----------------------------|-------------------------------------|------------------------|
| Sleepwalker BoI 65nm [JSSC'13] | 1 | 25 23.6 | 0.4 0.375 | 0.065* 0.052* | 2.6 2.2 | 104 93.2 | N/A |
| TeraFlops 65nm [2] | 80 | 4000 3130 | 1.2 1.0 | 2260** 1230** | 70.6† 49.1† | 17.7 15.7 | 2.65 |
| AsAP2 65nm [3] | 167 | 1070 66 | 1.2 0.675 | 47.5 0.61 | 44 9.2 | 41.1 139 | 0.998 |
| Am2045 130nm [4] | 336 | 300 | - | 23.8* | 79.4 | 265 | 0.713 |
| KiloCore 32nm This work | 1000 | 1782 1237 638 115 | 1.1 0.9 0.75 0.56 | 39.6 17.7 6.9 1.3 | 21.9 13.8 9.9 5.8 | 12.2 11.1 15.4 50.3 | 4.24 |

* Per processor power calculated by: energy per operation x clock frequency

** Per processor power calculated by: total power / number of cores

† Energy/Op calculated by: power / clock frequency / IPC (conservatively assumes processor is fully active every clock cycle)

Note:

A single processor dissipates 5.8 pJ per instruction including instruction execution, data reads/writes, and network accesses, at 0.56 V. When fully active at 115 MHz and including leakage current, the processor has a total power consumption of **0.69 mW**. Instruction energy is calculated using a weighted average based on the code profile of a 327-processor Fast Fourier Transform (FFT) application.

The chip power value of 1.3 mW reported in Table 1 includes power dissipated by the chip-level I/O circuits which operate at 1.8 V.