Optically-Connected Memory for Energy-Efficient Computing Systems

Daniel Brunina*, Caroline P. Lai, Ajay S. Garg, and Keren Bergman
Columbia University, Department of Electrical Engineering
**corresponding author: daniel@ee.columbia.edu*

Abstract: *We demonstrate an advanced, scalable optically-interconnected memory system, delivering a high-performance photonic transmission link between dynamic random access memory (DRAM) and a microprocessor. Error-free operation is achieved between SDRAM and an emulated microprocessor using wavelength-striped 4×2.5-Gb/s streams across a 4×4 optical test-bed.*

Introduction

The scaling of computational performance of servers and data centers is vastly outpacing the performance gains of memory systems, and has resulted in the creation of a memory bottleneck in system performance. Power dissipation acts as a significant cause for this performance gap by limiting the scale of memory systems, and as server power consumption continues to increase [1] the need for energy-efficient, high-performance memory systems will become more important. Conventional electronic interconnects have a poor energy-distance relationship due to transmission line effects, which becomes a significant design challenge as physical wiring distance and data rates scale up with increased memory capacity and bandwidth. Electronic interconnects will be unable to meet the increasing demands of performance and energy efficiency, and therefore new advances in memory interconnect technology are required to enable continued scaling of computing systems.

Optical interconnects have been demonstrated as a solution to the memory interconnect problem through their ability to provide high-bandwidth, energy-efficient, low-latency links over large distances [2]. The large bandwidth-distance product of optics can be leveraged by memory systems to enable physically distant memory nodes (Fig. 1) that deliver large amounts of data over an optical link with time-of-flight latencies and without sacrificing data-movement efficiency [3]. Overall, by reducing the power demands of the memory interconnect, high-performance computing systems can dramatically improve performance while simultaneously reducing power consumption.

In this work, we experimentally demonstrate multiple memory nodes connected to a host processor over a 3-stage all-optical network test-bed. A memory node consists of a circuit board containing four chips of Micron MT47H64M16BT-37E SDRAM, a high-speed Altera field-programmable gate array (FPGA), and a 4×2.5-Gb/s transceiver providing an aggregate bandwidth of 10 Gb/s. The processor is modeled on the FPGA of a second identical circuit board. The transceivers modulate four wavelength channels, which propagate through an implemented network test-bed with a wavelength-striped format. Switching functionality of the optical network test-bed is utilized to demonstrate the processor connected to multiple remote memory nodes.

Memory Access Protocol

The proposed memory access protocol utilizes circuit-switching to provide high sustained memory bandwidth and guaranteed end-to-end transmission of memory accesses over the optical network. Here,

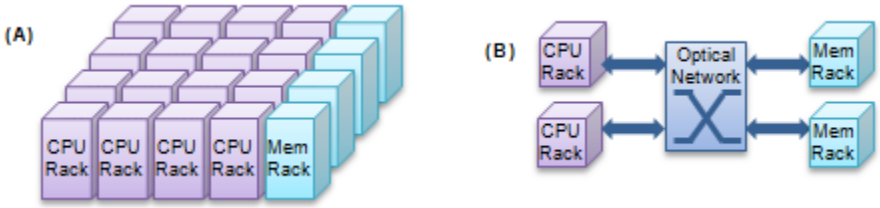

Fig. 1: (A) CPU racks with separate optically-connected memory racks; (B) a subset of the racks highlighting the optical interconnection network

the memory controller acts as a network arbiter by granting a processing core access to a remote memory node only after the memory controller has created the necessary circuit path. The circuit path setup latency can be amortized by allowing for longer continuous memory accesses, or *bursts*, due to the fact that each processor maintains exclusive access to a memory node for the duration of each memory access.

Experimental Setup and Results

The experimental setup presented here (Fig. 2) consists of two identical FPGA development boards that leverage an all-optical, wavelength-striped communication link through an implemented 4×4 all-optical network test-bed [4]. One development board acts as the microprocessor and the other as a main memory network node. Each development board contains an FPGA, four SDRAM modules, and transceivers banks capable of up to 4×2.5-Gb/s bidirectional communication. The transceivers interface with optical modulators to create a 4×2.5-Gb/s wavelength-striped optical stream. The stream propagates through the test-bed and is received using p-i-n-TIA receivers, which transmit the electronic stream to the second development board. The resulting configuration is one where all communication between the microprocessor and memory is performed all-optically.

The microprocessor performs a series of memory write and read operations over the full memory address space to validate complete memory system functionality, as well as overall network stability and reliability. Each test involves first writing to all memory addresses with a different bit pattern, and then reading back from all locations for verification. The bit patterns are: all 1's, all 0's, PRBS, and the bit sequence corresponding to each memory address being accessed. The read data is compared against expected bit patterns as it streams in from memory over the optical test-bed. A counter in the microprocessor tracks the amount of verified data in real-time, and signals that error-free performance has been achieved after one gigabit of data is verified.

Conclusion

We experimentally demonstrate an emulated microprocessor communicating with SDRAM network nodes over an implemented optical network test-bed. Error-free operation is achieved by successfully writing and reading one gigabit of data without bit mismatches (effective memory bit- error rate less than 10^{-9}). This work represents a significant step toward the integration of optical links into memory architectures, which enables computing systems with significantly greater memory capacity and bandwidth than would ever be possible with electronic interconnects alone.

References

[1] G. Astfalk, Applied Physics A **95,** 933–940 (2009).
[2] D. Brunina *et al.*, Photonics in Switching, PMC5 (2010).
[3] Y.Katayama et al., HPCA 46-50 (2007).
[4] A.Shacham et al., JLT **27** (7) (2009).

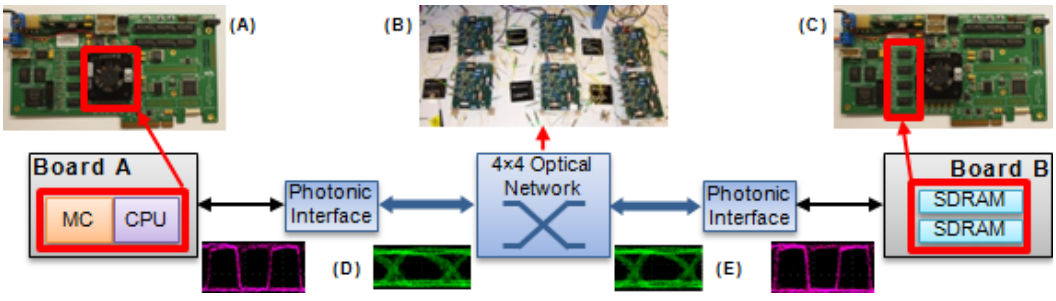

Fig. 2: Experimental setup: (A) Board A implements a CPU with on-chip MC on highlighted FPGA; (B) 4×4 optical network test-bed; (C) Board B with highlighted SDRAM; (D) Optical and received electronic 2.5-Gb/s eyes of read data from memory; (E) Optical and received electronic 2.5-Gb/s eyes of write data from CPU