First Experimental Demonstration of Optically-Connected SDRAM Across a Transparent Optical Network Test-Bed

Daniel Brunina, Caroline P. Lai, Ajay S. Garg, and Keren Bergman
Dept. of Electrical Engineering, Columbia University, New York, NY, USA; daniel@ee.columbia.edu

Abstract: **We experimentally demonstrate SDRAM-based memory nodes connected to an emulated microprocessor across a 4×4 optical interconnection network utilizing 4×2.5-Gb/s wavelength-striped WDM payloads. Error-free performance is verified by a series of SDRAM write and read transactions.**

Introduction

As the number of processor cores per die continues to grow, the need for sustained memory bandwidth has become a significant design challenge[1]. Memory devices remain limited to clock frequencies in the hundreds of megahertz, and must therefore meet growing bandwidth requirements by prefetching a larger amount of data and driving the memory bus at higher clock rates. Furthermore, to minimize the growing latency gap, system designers must place the processor, memory controller (MC), and main memory as physically close to each other as possible. The bandwidth, latency, and capacity requirements are inherently at odds in today's electronically interconnected memory systems, and continue to strain the limits of electronics in terms of energy density, pin count, wiring complexity, and skew. Overall, the electronic memory bus is becoming a limiting factor in the performance gains of future computing systems[2].

Optical interconnects have been proposed as a solution to these challenges due to their ability to provide high-bandwidth, energy-efficient, low-latency links over large distances[3]. This large bandwidth-distance product can be leveraged by future memory systems, enabling physically distant memory nodes (Fig. 1) that deliver large amounts of data over an optical network without sacrificing data-movement efficiency[4]. As a result, the design space of high-performance systems will be no longer tightly constrained by spatial locality. Future systems utilizing optically-connected SDRAM can be designed with larger memory capacities and bandwidths.

Here, 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 Micron MT47H64M16BT-37E synchronous dynamic random access memory (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.

Memory Access Protocol

Memory access granularity is dictated by the burst length, which specifies the number of data words accessed in each transaction. Every new memory access incurs tens of
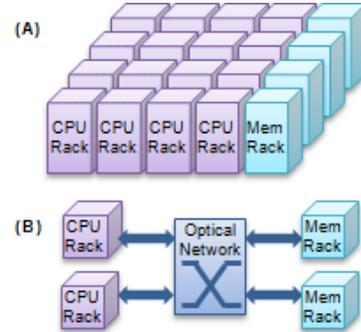
Fig. 1: (A) CPU racks with separate optically-connected memory racks; (B) a subset of the racks highlighting the optical interconnection network

nanoseconds of latency. Therefore, the burst length has increased in each new generation of SDRAM to its current value of eight words. The data words, typically 64 bits, are spread across several SDRAM chips, all sharing common command and address wires for parallel access. Data is transmitted between the MC and SDRAM by a wide, electronic double-pumped bus. The bus wires must be path-length-matched to minimize skew, which adds significant complexity to the system design. Ultimately, the demands placed on memory interconnect by greater parallelism and faster I/O transceivers will become too great for electronics and therefore necessitate an optical interconnect approach.

Our MC is modified to allow for longer, variable-size burst requests up to the size of a full SDRAM row, typically 1024 data words. By enabling longer continuous read or write bursts, we can exploit the inherent SDRAM parallelism to better utilize the available bandwidth offered by an optical solution. This functionality maps well to streaming and highly-parallel applications that operate on large blocks of data. The longer bursts also reduce latency by simplifying the MC design.

Optical Interconnection Network

In this work, the memory access operation is streamed through an experimentally implemented 4×4 optical switching network test-bed[5]. The test-bed is comprised of six 2×2 photonic switching nodes, each containing four semiconductor optical amplifiers (SOAs) that provide high-bandwidth capacity switching. The network leverages a wavelength-striped packet format where control bits are encoded on dedicated wavelengths and the wavelength-division multiplexed (WDM) payload information is modulated on the rest of the available band. Here, the 4×2.5-Gb/s payload channels representing the memory transactions are multiplexed with the appropriate header bits and transmitted through the test-bed. The control bits are recovered using wavelength filters and optical receivers

at each switching node, and are processed by a complex programmable logic device (CPLD). The CPLD then uses the corresponding SOAs to establish a circuit path through the network test-bed and to route the optical stream to the desired output port.

Experimental Setup and Results

The experimental setup, depicted in Fig. 2, consists of two circuit boards that establish an all-optical WDM lightpath over a transparent 4×4 optical circuit-switched network. Board A (Fig. 2a) and board B (Fig. 2c) are identical FPGA-based development boards, each containing an Altera Stratix II GX FPGA, four SDRAM chips, and a 4×2.5-Gb/s high-speed I/O transceiver. The data transceivers on one circuit board are connected to four 10-Gb/s optical modulators, which encode the memory access information optically using off-the-shelf distributed feedback (DFB) lasers. A wavelength-striped 4×2.5-Gb/s optical stream is established and injected in the network test-bed (Fig. 2b) with the appropriate control bits. The stream is then received at the output using 10-Gb/s p-i-n-TIA receivers and transmitted to the other circuit board.

The resulting configuration is one where the processor and MC communicate electronically, while all communication between the MC and main memory is performed transparently and all-optically through the optical network test-bed. Two network control addresses are verified to demonstrate processor accesses to multiple optically-connected memory nodes.

The FPGA on board A implements the functionality of a microprocessor with an on-chip MC, as well as an optical packet format (OPF) module and a central arbiter for the optical network. Here, the processor first generates and then later requests data for main memory accesses. The MC interfaces with the central arbiter to control circuit-path setup and teardown for write and read data, processor-to-memory and memory-to-processor respectively, which simplifies the design of the remote memory node and reduces memory access latency.

Board B implements the SDRAM-based memory node. The FPGA on board B implements a version of the OPF for transmitting or receiving data over the optical network, along with SDRAM driver circuitry necessary to double-pump the memory data bus. The memory node does not contain its own control logic and merely responds to memory access requests generated by the processor.

To experimentally validate the optically-connected memory system, our emulated microprocessor runs a sequence of four tests that generate true memory transactions across the optical test-bed. The tests are repeated indefinitely while the microprocessor calculates a real-time effective memory-bit-error rate (EMBER). Each test first writes a specific bit pattern to all memory locations, and then reads back all memory locations while verifying the received data. The bit patterns for the four tests are: all zeroes, 2^{32}-1 PRBS, all ones, and bits representing the destination memory address.

Each memory access is of a burst length equal to an SDRAM row size, 2^{10} columns, and a sufficient number of accesses are issued to test the full address space. With an aggregate bandwidth of 10 Gb/s, each burst transmits on the circuit path for over 3.2 μs. As the processor receives read data from the network, it verifies each bit and updates a counter to quantitatively track the amount of data that has been streamed from the network and correctly verified. Any mismatch between expected and received data signals an error. Error-free (EMBERs<10^{-9}) operation is achieved once the data counter reaches one gigabit, thus verifying full memory functionality and addressability, network stability, and repeatability. Optical and received electrical 2.5-Gb/s eye diagrams are shown in Fig. 2d and Fig. 2e.

Conclusions

We have experimentally demonstrated the error-free functionality of optically connecting a microprocessor, as emulated on a high-speed FPGA, to multiple off-the-shelf SDRAM-based memory nodes across an implemented optical network test-bed. This work represents a significant step toward future system designs that can leverage the high bandwidth-distance product and energy scalability of optics to bridge the performance gap between computing cores and memory.

References

1 R. Yavatkar, 12th IEEE Symposium on High-Performance Computer Architecture (HPCA) (2006).
2 Wm. A. Wulf et al., ACM SIGARCH **23** (1) (1995).
3 B. J. Offrein et al., ECOC, 3.1.3 (2009).
4 Y. Katayama et al., HPCA 46-50 (2007).
5 A. Shacham et al., JLT **27** (7) (2009).
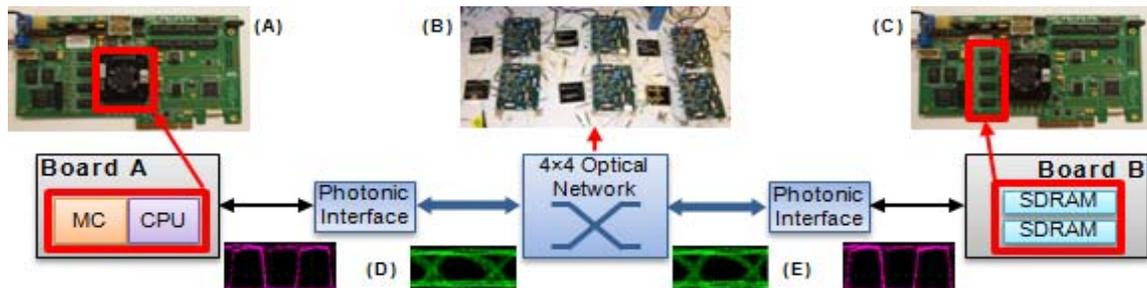
Fig. 2: Experimental setup: (A) Circuit board A realizing a CPU with on-chip MC on highlighted FPGA; (B) 4×4 optical network test-bed; (C) Circuit board B with highlighted SDRAM; (D) Optical and received electronic 2.5-Gb/s eyes of read data from memory; (E) Optical and received electronic 2.5-Gb/s eyes of write data from CPU