2010 Fourth ACM/IEEE International Symposium on Networks-on-Chip

Performance Evaluation of a Multicore System with Optically Connected Memory Modules

Paul Vincent Mejia*, Rajeevan Amirtharajah*,
Matthew K. Farrens†, and Venkatesh Akella*

* *Department of Electrical and Computer Engineering*
† *Department of Computer Science*
University of California, Davis
Davis, CA 95616

***Abstract*—Over the past years, there have been impressive advances in bandwidth, latency, and scalability of on-chip networks. However, unless the off-chip network bandwidth and latency are also improved, we might have unbalanced systems which will limit the improvements to overall system performance. In this paper, we show how dense wavelength-division multiplexing (DWDM) -based optical interconnects could be used to emulate multiple buses in a fully-buffered DIMM (FB-DIMM) -like memory system to improve both bandwidth and latency. We evaluate an optically connected memory using full-system simulations of an 8-core system running memory-intensive multithreaded workloads. We show that for the FFT benchmark, optically connected memory can reduce the average memory request latency by 29% compared to a single-channel DDR3 SDRAM system and provide an overall performance speedup of 1.20. We also show that at least two DDR3 memory channels are needed to match the performance of a single optical bus, which demonstrates the advantage of optical interconnects in terms of savings in the number of pins required.**

***Keywords*-DRAM, OC-DIMM, Optics, DWDM, Performance Evaluation.**

I. Introduction

There has been enormous interest in networks-on-chip and numerous papers about improving the bandwidth and latency of on-chip networks have been published [1]. However, there has been less attention paid to how to get data *into* and *out of* the chip, i.e. the challenges of off-chip interconnects. We believe that as the core count per chip increases to hundreds or thousands and networks-on-chip become more sophisticated and scalable, off-chip interconnects will carry a higher burden to sustaining high overall system performance. Therefore, it is important to study ways of improving the bandwidth and latency of off-chip interconnects, especially the interconnect between the CPU and DRAM.

Memory latency is improving much more slowly than memory bandwidth [2]. So, improving on-chip network bandwidth and latency without a corresponding improvement in the off-chip memory bandwidth and latency will result in *unbalanced* systems [3] and limit the scalability of future multicore processors and high performance SoCs. To improve off-chip bandwidth, one has to provide more memory channels which, in turn, increases the number of pins. Unfortunately, the growth of total package pin count is projected at only 10% per year [4]. This will limit how much bandwidth one can provide to a chip.

One solution to this problem is the use of 3D die stacking [5] which enables the use of optical interconnects with dense wavelength-division multiplexing (DWDM). With DWDM, multiple bits of data can be on the same waveguide simultaneously to realize an optical bus. A single-waveguide, 64-wavelength, 10 Gbps DWDM optical bus can provide a peak bandwidth of 80 GBps. As a comparison, a 64-bit DDR3-1333 memory channel has a peak bandwidth of 10.6 GBps. Thus, one optical fiber can provide the bandwidth of almost 8 DDR3-1333 channels without an increase in pin count. Though the higher bandwidth potential of optical interconnects is well-known, it is not clear whether optical interconnects can reduce the latency of a memory system and, if so, by how much.

The goal of this paper is to address this issue by leveraging DWDM-based optical buses in memory systems using off-the-shelf DRAM devices. The total memory access latency T_m is given by

$$T_m \approx T_q + T_d + T_a \tag{1}$$

where T_q is the queueing delay, T_d is the transport delay and T_a is the DRAM access time. With the increase in the number of cores per chip and the improvements to the on-chip networks, T_q will be the dominant factor in the overall memory latency. Single-channel memory systems suffer from high contention for the shared data bus, thus increasing the queuing delay of memory requests. Using electrical I/O would require doubling the pin count for memory interface to support two memory channels. The intuition behind our approach is as follows. We *emulate* multiple memory channels using multiple wavelengths on the optical bus, so that several memory transactions can be serviced simultaneously. This will reduce the queueing delay and, hence, reduce the overall memory latency.

This introduces some interesting research questions — what is the best memory access protocol to use in WDM

978-0-7695-4053-5/10 $26.00 © 2010 IEEE
DOI 10.1109/NOCS.2010.31

optical interconnects? How many buses should we emulate? What is the impact on memory latency? What is the impact on overall execution time of the program? This paper attempts to answer these questions.

The rest of this paper is organized as follows. Section 2 describes recent studies related to this paper. Section 3 describes the proposed optically connected memory system. Section 4 explains the experimental setup used in our study. Section 5 provides an assessment of the performance of the proposed memory system design compared to conventional DDR3 SDRAM systems. Lastly, section 6 concludes this paper.

II. Related Work

Tan et al. [6] demonstrated the feasibility of using an optical data bus for computer interconnections to avoid signal integrity issues inherent in using multi-drop electrical buses. Batten et al. [7] described a custom opto-electrical crossbar network that connects DRAM modules to processors. Corona [8] uses a pair of waveguides to connect the memory controllers to external 3D-stacked DRAMs. This paper describes a variation of the fully-buffered DIMM (FB-DIMM) network using off-the-shelf DRAM devices which we believe is less ambitious and, hence, more practical as a near term solution. The concept of optically connected DIMMs, or OC-DIMMs, was evaluated on synthetic traces primarily to illustrate the bandwidth and power advantages of the proposed design [9], [10]. Our work describes the OC-DIMM architecture in detail and involves full-system simulations with emphasis on memory request latency and overall system performance. Hendry et al. [11] described a system-level analysis of on-chip photonic networks, but detailed evaluation of the benefits of off-chip photonic networks was not considered. This paper can be viewed as a complementary to the work mentioned in the sense that it evaluates the system-level benefits of off-chip photonic interconnects.

III. Optically Connected Memory Modules

A. Architecture

The OC-DIMM system uses a serial packet-based protocol to communicate between the memory controller and the DIMMs. A similar protocol is used in FB-DIMM systems [12]. An FB-DIMM system uses a point-to-point interface between the memory controller and the nearest DIMM and between each other DIMM through the on-module advanced memory buffers (AMB). On the other hand, an OC-DIMM system uses an optical broadcast bus that is interfaced to all the DIMMs through the on-module optical memory controller (OMC).

Figure 1 shows the OC-DIMM memory system topology interfaced with a 3D-stacked multicore system which includes the optical components on a separate layer. Using DWDM, multiple wavelengths from a single waveguide

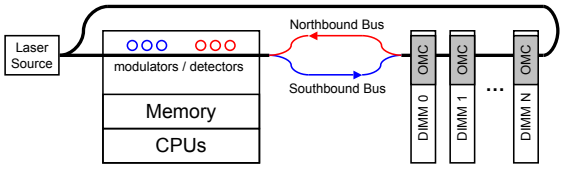

Figure 1. OC-DIMM system topology.

are grouped together to form the southbound (SB) and the northbound (NB) buses of the OC-DIMM system. The SB bus carries commands or WRITE data packets to the DIMMs. The NB bus carries READ data packets from the DIMMs to the memory controller. This architecture enables simultaneous READ and WRITE cycles on different DIMMs.

The memory controller broadcasts commands and WRITE data packets on the SB bus. If a command is addressed to the DRAM devices in a particular DIMM, the OMC on that DIMM decodes the command and controls the DRAM devices. If WRITE data is sent on the SB bus, the OMC of the designated DIMM stores the data in its FIFO buffer and subsequently stores them to the DRAM devices. In most cases, multiple SB packets are needed to send an entire block of WRITE data. When a READ command is sent on the SB bus, READ data is returned through the NB bus. Similarly, multiple NB packets are needed to return a block of READ data.

B. Access Protocol

The proposed OC-DIMM access protocol is adopted from the protocol used in FB-DIMM systems. Modifications were made so that the OC-DIMM system can benefit from the flexibility provided by the DWDM optical interconnects. The main difference is the size of a packet. FB-DIMM systems use fixed-sized packets: 168 bits for a NB packet and 120 bits for a SB packet. An OC-DIMM packet can be made as small or as large depending on the number of wavelengths available and the organization of these wavelengths. Consequently, the bus bandwidth of the SB and NB buses can vary.

Data that is sent in one DRAM clock cycle constitutes a packet. Assuming an OC-DIMM system that is built based on PC3-10600 DIMMs (comprised of 666 MHz DDR3-1333 chips) and a per wavelength bandwidth of 10 Gbps, then a single wavelength can transmit 15 bits of data per packet. Each packet may contain commands, READ data, or WRITE data.

The number of wavelengths needed in a SB bus is given by

$$\lambda_{SB} = \left\lceil \frac{CRC + PacketID + (CMD \times Slots)}{15} \right\rceil \quad (2)$$

where CRC is the number of checksum bits, $PacketID$ is a 2-bit identifier that specifies the packet type, CMD is

Table I
WAVELENGTH ASSIGNMENTS.

	Notation			
	2NB	4NB	8NB	16NB
NB Bus	24×2	12×4	7×8	3×16
SB Bus	15	15	7	15
Misc.	1	1	1	1
Total Wavelengths	64	64	64	64

a 24-bit command, and *Slots* is the number of command and/or WRITE data slots per SB packet. A SB packet can be either a *command* or a *command-with-data* packet. Independent commands to separate DRAM ranks can be sent in one packet, making it possible to process multiple memory transactions to different ranks simultaneously. For instance, a SB packet that can carry 4 slots can initiate up to 4 memory requests concurrently. In this example, the SB bus needs to be 8 wavelengths wide, based on equation (2) and assuming a 22-bit CRC.

Each slot in a SB packet can also be used to transmit a 16-bit WRITE data along with a 2-bit ECC and a 6-bit rank ID that identifies the destination rank. For example, a SB bus that can carry 5 slots is capable of transmitting one command and a 64-bit WRITE data. Multiple such command-with-data packets may be needed to construct a block of WRITE data. In this example, 8 command-with-data packets are needed to write a 64-byte cache block into a given rank of DRAM devices. The command slots of a command-with-data packet may be the accompanying RAS or CAS command of the WRITE transaction, READ commands to a different rank, or simply NOPs.

The number of wavelengths needed in a NB bus is given by

$$\lambda_{NB} = \left\lceil \frac{CRC + READ + ECC}{15} \right\rceil. \qquad (3)$$

For example, to transmit a 128-bit READ data, a NB packet must carry the 128-bit data, a 16-bit ECC, and a 22-bit CRC. Therefore, a NB bus needs to be allocated 12 wavelengths to satisfy this bandwidth requirement. This wavelength assignment matches the DIMM bandwidth of a PC3-10600, which is 10.6 GBps.

C. Wavelength Assignment

1) Multiple Data Channel Configuration: In this study, we assumed a photonic technology that can multiplex up to 64 wavelengths per waveguide, which is the expected capability in future 22 or 16 nm process [7], [8]. Also, we used a single waveguide for performance comparison with a single electrical memory channel. Table I shows different ways of how wavelengths from a single waveguide can be assigned to form the SB and the NB buses of an OC-DIMM system. The notation given describes the number of NB buses in the system.

A 4NB OC-DIMM system can process 4 READ requests concurrently, and each 12-wavelength wide NB bus

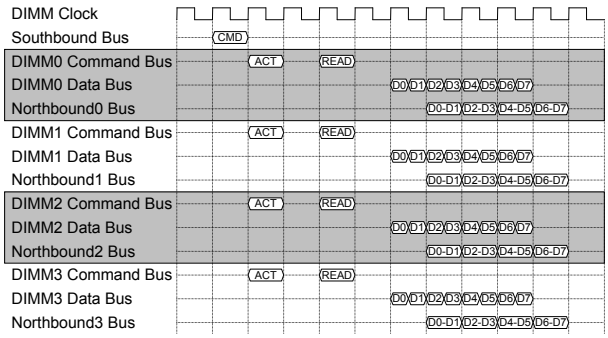

Figure 2. READ transaction in a 4NB OC-DIMM system.

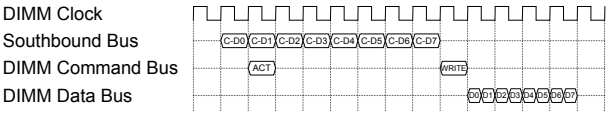

Figure 3. WRITE transaction in a 4NB OC-DIMM system.

is capable of carrying 128 bits of READ data per packet. This system's 15-wavelength wide SB bus is capable of transferring 64 bits of WRITE data, half of what a single NB bus can deliver. Therefore, a 4NB system can provide a total peak bandwidth of $(10.6 \times 4) + 5.3 = 48$ GBps.

Figure 2 illustrates how READ requests are processed concurrently in a 4NB OC-DIMM system. Assuming that the transaction queue has at least 4 pending READ requests to different DIMMs, DRAM commands to each of these DIMMs are sent in a single SB packet (CMD). Commands are received and decoded by each of the OMCs. The OMCs relay DRAM commands (ACT for row activation and READ for column read) to their respective DRAM ranks to perform READ transactions concurrently. The DRAM devices burst data back to the OMCs where they are formatted into packets before they are sent to the memory controller through the NB buses. Standard DIMMs have a burst width of 64 bits, so it takes a burst length of 8 beats, or 4 DRAM clock cycles, to move an entire 64-byte cache block out of the DRAM devices. Figure 2 is not drawn to scale for easy illustration. Nevertheless, the memory controller must schedule commands accordingly so that DRAM timing constraints (t_{RAS}, t_{CAS}, etc.) are satisfied.

Figure 3 illustrates how a WRITE transaction is executed in a 4NB OC-DIMM system. The memory controller sends command-with-data packets (C-Dx) to the designated DIMM through the SB bus. Since a 15-wavelength wide SB bus can transmit 64 bits of WRITE data per packet, the memory controller needs to send 8 command-with-data packets to write an entire 64-byte cache block. The first SB packet (C-D0) consists of the row-activate command (ACT) and the first sub-block of WRITE data. The last SB packet (C-D7) of a WRITE transaction consists of the column-write command (WRITE) and the last sub-block of the WRITE data. The SB packets in between carry the rest

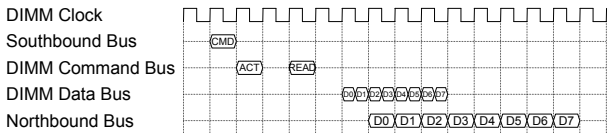

Figure 4. READ data transfer on a 7-wavelength NB bus.

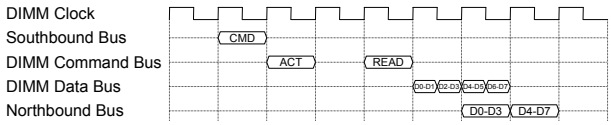

Figure 5. READ data transfer on a 24-wavelength NB bus.

of the WRITE data sub-blocks and may be accompanied by READ commands to other DIMMs, if available. Otherwise, NOPs are inserted.

An OC-DIMM system can also be configured to have more than 4 NB buses, thereby increasing the number of concurrent READ requests that it can handle. However, due to a limited number of wavelengths per waveguide, this approach come at the expense of increasing the latency of individual READ transactions. An 8NB system has twice the number of NB buses but half the bus bandwidth. With 7 wavelengths, each NB bus is capable of transferring 64 bits of READ data per packet for a peak bandwidth of 5.3 GBps, which is only half the bandwidth of a PC3-10600 DIMM. Figure 4 illustrates how an OC-DIMM system with a 7-wavelength wide NB bus returns a READ transaction. Even if the burst length required to move a block of data out of the DRAM devices is 8 beats, transmission of data back to the memory controller will still take more cycles because of limited NB bus bandwidth. The figure shows that a 7-wavelength wide NB bus requires twice the number of NB packets to send a 64-byte block of READ data. Therefore, an 8NB OC-DIMM system increases the amount of concurrency in terms of the number of NB data buses at the expense of increased data transmission time.

Using a 64-wavelength waveguide, an OC-DIMM system can have up to a maximum of 16 NB buses. For the same reasons explained previously, the increase in the number of NB buses limits the available bandwidth per NB bus. In a 16NB system, a 3-wavelength NB bus provides only one-eighth of a PC3-10600 DIMM's bandwidth, which can further increase data transmission time.

2) High-Bandwidth Bus Configuration: An OC-DIMM system can also have fewer, yet higher-bandwidth NB buses by allocating more wavelengths to them. In a 2NB OC-DIMM system, each NB bus has a bandwidth of 21.33 GBps. To take advantage of this, a NB bus must be interfaced to a matching pair of DIMMs that operates in lockstep and sends 128 bits of data per DRAM burst. This setup is similar to the dual-channel configuration supported by the Intel 875P chipsets [13]. The DRAM devices now require a burst length of 4 beats, instead of 8, to read a 64-byte cache block. Figure 5 illustrates how an OC-DIMM system with a 24-wavelength wide NB bus returns a READ request in less transmission time.

Ideally, more wavelengths can be assigned to a NB bus and more DIMMs can be matched to operate in lockstep to further reduce the burst length to 2, or even 1. However, current DDR2 and DDR3 SDRAM devices support burst lengths of 4 and 8 only.

Table II
SYSTEM CONFIGURATION.

Processors	8 out-of-order cores, 4 GHz clock, 32-entry Instruction Window, 64-entry Re-order Buffer, 1 KB YAGS branch predictor
L1 I/D Cache	128 KB, 4-way associative, 64-byte line
L2 Cache	4 MB, 8 banks, 4-way associative, 64-byte line
Main Memory	8 GB DIMMs, 64-bit DIMM data channel, DDR3-1333 SDRAM devices, 10.6 GBps DIMM bandwidth, t_{RCD}-t_{RP}-t_{CL} of 9-9-9

IV. EVALUATION

A. Full-System Simulator

For our experiments, we used the Simics full system simulation platform [14]. We used Simics as a functional simulator to emulate a Sun server running an unmodified Solaris operating system. We leveraged this platform by using the timing simulator modules, Opal and Ruby, from the GEMS toolset [15]. Opal models an out-of-order SPARCv9 processor while Ruby models a multiprocessor memory system. These tools are used to simulate an 8-core CMP that implements a directory-based MOESI cache coherence protocol. While Ruby provides a detailed model of a multiprocessor memory hierarchy, it does not accurately model the DRAM behavior. In order to simulate a detailed DRAM system, we integrated the DRAMsim memory system simulator [16] with Ruby. Table II shows a summary of the simulated system parameters.

We modified DRAMsim to model our proposed OC-DIMM system. We also configured DRAMsim to model a conventional DDR3 memory module for performance comparisons. We used an 8 GB PC3-10600-like dual-rank DIMMs to model both OC-DIMMs and standard DDR3 DIMMs. The device parameters were based on Micron's 4Gb TwinDie DDR3 SDRAM components [17]. Additonally, we configured DRAMsim to model a transaction queue size of 256.

B. Workload Characteristics

We used multithreaded benchmarks from the SPLASH-2 [18] and PARSEC-2.1 [19] suites. We ran full-system

Table III
WORKLOADS AND CONFIGURATIONS.

Program	Problem Size	# DRAM Requests
FFT	1 M data points	5.24 M
Radix	2 M integers	2.37 M
Canneal	100 K elements	2.99 M

Table IV
OC-DIMM CHANNEL LATENCY BREAKDOWN.

Latency Component	Value
Controller-to-DIMM flight	10.45 ps/mm
SB DIMM-to-DIMM flight	10.45 ps/mm
Optical receiver delay	7.4 ps
Command check and decode	3000 ps
DRAM access	27000 ps
Data Serialization	3000 ps
Optical transmitter delay	36.3 ps
NB DIMM-to-DIMM flight	10.45 ps/mm
DIMM-to-Controller flight	10.45 ps/mm
Frame-into-Controller	1500 ps

simulations of programs from both benchmark suites using the setup described in section IV-A and a single-channel DDR3 memory controller populated with a single dual-rank 8 GB DDR3 memory module. From these experiments, we picked those programs that require high off-chip memory bandwidth demand and reasonable simulation time, namely FFT, Radix, and Canneal. Table III shows a summary of our selected workloads.

C. Memory Request Latency Breakdown

1) OC-DIMM Channel Latency: Table IV summarizes the latency breakdown of an OC-DIMM system channel populated with DDR3-1333 components. The delay values of the optical components are predicted values for a 32 nm process technology [20], [21]. The OMC component delays were derived from the AMB delays of an FB-DIMM system based on DDR2-667 components [22]. These delays were adjusted for a DDR3-based memory module. We used these latency numbers in our OC-DIMM model for our performance studies.

2) Queuing Delay and Data Transmission Time: The previous subsection described the OMC-related overheads experienced by an OC-DIMM and the amount of flight time required on an optical medium. This subsection describes two other components of the total latency of a memory request.

Queuing delay is the amount of time that a memory request has to wait in queue before it is serviced. In systems running memory-intensive workloads, the number of pending memory requests in the memory controller's transaction queue can be huge. In situations like this, the queuing delay can become a significant part of the total memory request latency. Therefore, a good memory design must be able to minimize the queuing delay incurred by a memory request. Data transmission time, on the other hand, is the time for a packet to pass through the bus, not including time of flight, and is equal to the size of the packet divided by the bus' bandwidth [23].

We are particularly interested in addressing queuing delay and data transmission time by using the OC-DIMM memory system. Creating multiple data buses by organizing groups of wavelengths allows multiple requests to be serviced concurrently, thus addressing queuing delay. Furthermore, DWDM optical buses provide high bandwidth density which can help reduce data transmission time.

Table V
OC-DIMM SYSTEM ORGANIZATION AND PEAK BANDWIDTH.

Total DIMMs	Config Notation	BW Per NB Bus (GBps)	BW Per SB Bus (GBps)	Total BW (GBps)
1	1NB-1D	10.67	21.33	32.00
2	1NB-2D	21.33	21.33	42.67
	2NB-1D	10.67	21.33	42.67
4	1NB-4D	21.33	21.33	42.67
	2NB-2D	21.33	5.33	48.00
	4NB-1D	10.67	5.33	48.00
8	1NB-8D	21.33	21.33	42.67
	2NB-4D	21.33	5.33	48.00
	4NB-2D	10.67	5.33	48.00
	8NB-1D	5.33	2.67	45.33
16	1NB-16D	21.33	21.33	42.67
	2NB-8D	21.33	5.33	48.00
	4NB-4D	10.67	5.33	48.00
	8NB-2D	5.33	2.67	45.33
	16NB-1D	1.33	5.33	26.67
32	1NB-32D	21.33	21.33	42.67
	2NB-16D	21.33	5.33	48.00
	4NB-8D	10.67	5.33	48.00
	8NB-4D	5.33	2.67	45.33
	16NB-2D	1.33	5.33	26.67

D. OC-DIMM System Organization

This study encompasses all the possible organizations of an OC-DIMM system with 1 to 32 dual-rank DIMMs, organized into 1 to 16 NB buses. Table V summarizes the different OC-DIMM organizations possible using a single 64-wavelength waveguide. The organizations are grouped according to the total number of DIMMs in the system. The notation used in the table describes the number of NB buses and the number of DIMMs that share a NB bus. For example, the notation 8NB-2D means an OC-DIMM organization of 8 NB buses with two DIMMs sharing a NB bus.

Each DIMM provides a peak bandwidth of 10.6 GBps. A simple OC-DIMM organization have a matching NB bus bandwidth and 1 or more DIMMs that share a NB bus. OC-DIMM systems with a NB bus bandwidth of 21.3 GBps and at least 2 DIMMs per NB bus are configured such that a matching pair of DIMMs operates in lockstep to match the NB bus bandwidth, as described in section III-C2. OC-DIMM systems whose NB bus bandwidth is less than the DIMM's bandwidth are those that stretch the number of NB buses to 8 or 16 at the expense of lower per-NB-bus bandwidth.

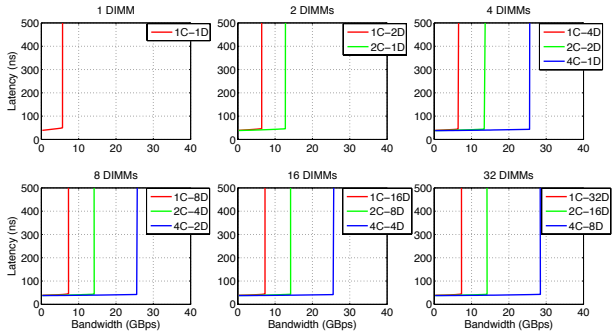

Figure 6. DDR3 SDRAM latency-bandwidth curves.

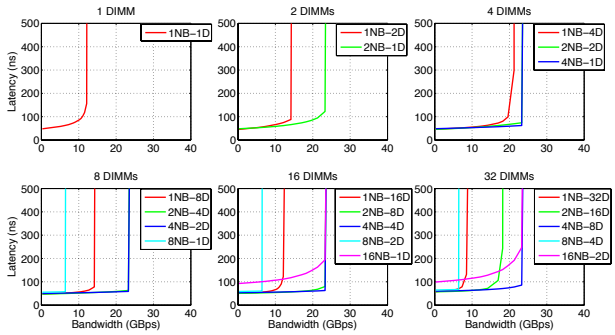

Figure 7. OC-DIMM latency-bandwidth curves.

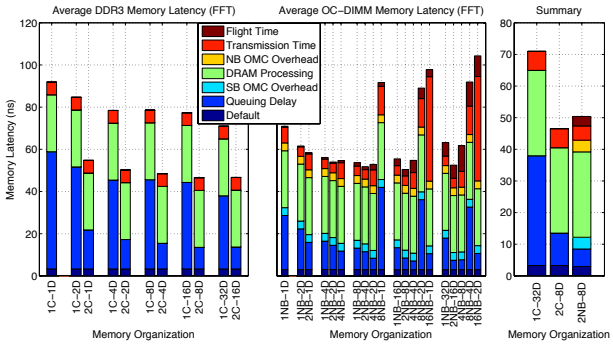

Figure 8. FFT Memory Request Latency Breakdown. Average memory request latency is reduced by 29% in OC-DIMM compared to a single-channel DDR3 SDRAM. Two channels of DDR3 memory is needed to yield a latency reduction of 7.66% over OC-DIMM.

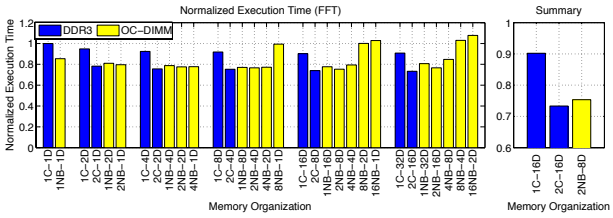

Figure 9. FFT Execution Time. OC-DIMM provided a performance speedup of 1.20 compared to a single-channel DDR3 SDRAM. Two channels of DDR3 memory is needed to yield a speedup of 1.03 over OC-DIMM.

V. Results

A. Latency-Bandwidth Characteristics

Figure 6 shows the latency-bandwidth characteristics of a DDR3 SDRAM system. These latency and bandwidth values were measured for a random address trace with varying memory request rate. We investigated a configuration space of 1 to 32 dual-rank DIMMs organized into 1 to 4 channels. A similar notation is used to denote a DDR3 SDRAM system's organization. The notation 2C-4D means an organization of 2 channels with 4 DIMMs sharing a channel. Each subplot in the figure is organized according to the total number of DIMMs in the system.

The first observation is that the sustained bandwidth is significantly increased as the number of channels is increased. Having multiple independent channels reduces the amount of contention due to shared resources, like the address, command, and data buses, and allows for multiple memory transactions to be serviced concurrently. This greatly reduces the queuing delay for memory requests. Another observation is that an increase in the number of DIMMs (or ranks) per channel slightly increases the sustained bandwidth. Having multiple ranks in a channel provides more memory banks to distribute requests to and more opportunities for scheduling and pipelining requests. Overall, a 4-channel DDR3 SDRAM system can sustain a bandwidth of up to 28.4 GBps. Having 2 channels can sustain up to 14.2 GBps, while having a single channel can sustain up to 7.3 GBps.

Figure 7 shows the latency-bandwidth characteristics of an OC-DIMM system. The sustained bandwidth is significantly increased as the number of NB buses is increased from 1 to 2. Ignoring 8NB OC-DIMM systems for a moment, 1NB systems have the lowest sustainable bandwidth because it suffers from high contentions for the shared NB bus. On the other hand, the poor bandwidth performance of 8NB systems is due to the very limited bandwidth of its SB bus. Since commands are transmitted through the SB bus, 8NB systems suffer highly from the inability to schedule as much commands as possible to take advantage of the multiple NB buses available. In fact, 8NB systems have the lowest SB bus bandwidth among all the OC-DIMM organizations. A 16NB system benefits from having twice as much SB bandwidth and yields one of the highest sustained bandwidth, although it experiences very high latency due to long transport delays, as explained in section III-C1.

The latency-bandwidth characteristics of 2NB and 4NB systems are very much alike. A 4NB system experiences lower queuing delays by being able to service more requests concurrently. On the other hand, a 2NB system has higher-bandwidth NB buses that help reduce data transport delay by about half. As the plots show, these two OC-DIMM organizations' benefits equally outweigh each other. The only marked difference occurs when the OC-DIMM system has 32 DIMMs. In this case, a 2NB-16D system has a deeper NB channel than a 4NB-8D system. A deeper NB channel

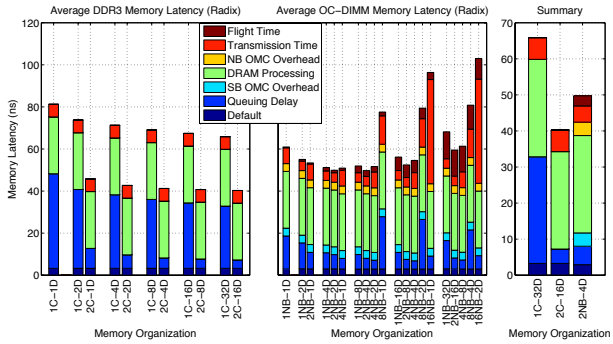

Figure 10. Radix Memory Request Latency Breakdown. Average memory request latency is reduced by 24.48% in OC-DIMM compared to a single-channel DDR3 SDRAM. Two channels of DDR3 memory is needed to yield a latency reduction of 19.03% over OC-DIMM.

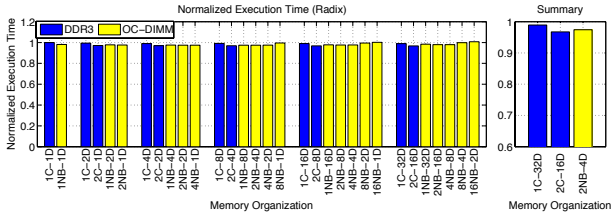

Figure 11. Radix Execution Time. OC-DIMM provided a performance speedup of 1.02 compared to a single-channel DDR3 SDRAM. Two channels of DDR3 memory is needed to yield a speedup of 1.01 over OC-DIMM.

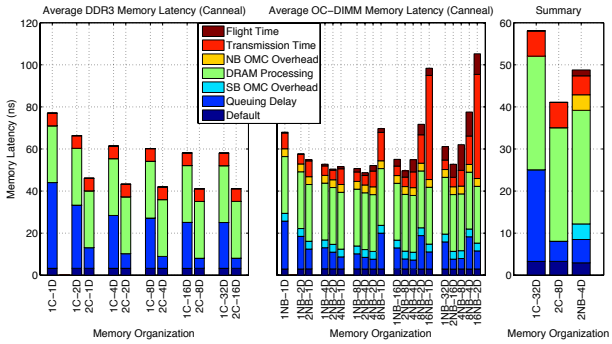

Figure 12. Canneal Memory Request Latency Breakdown. Average memory request latency is reduced by 15.99% in OC-DIMM compared to a single-channel DDR3 SDRAM. Two channels of DDR3 memory is needed to yield a latency reduction of 15.81% over OC-DIMM.

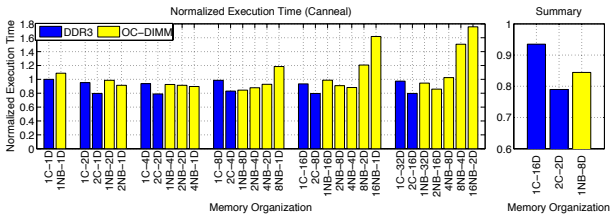

Figure 13. Canneal Execution Time. OC-DIMM provided a performance speedup of 1.11 compared to a single-channel DDR3 SDRAM. Two channels of DDR3 memory is needed to yield a speedup of 1.07 over OC-DIMM.

means more in-flight memory transactions that try to contend for the NB bus, thus increasing the queuing delay.

Overall, the maximum sustainable bandwidth of an OC-DIMM system is about 23.3 GBps, which is attainable by using any of the several optimum OC-DIMM organizations — 2NB-1D, 2NB-2D, 2NB-4D, 2NB-8D, 4NB-1D, 4NB-2D, 4NB-4D, and 4NB-8D. Although a 16NB system can sustain as much bandwidth, its high latency makes it less attractive compared to the others. Compared to a DDR3 SDRAM system, the OC-DIMM system's sustained bandwidth is 3.2× the sustained bandwidth of a single-channel DDR3 SDRAM system. A DDR3 SDRAM system would need up to 4 channels of memory in order to attain this bandwidth.

B. Latency and Execution Time

This section compares the performance of the OC-DIMM memory system against conventional 1-channel and 2-channel DDR3 SDRAM systems for multithreaded workloads. Figure 8 shows the average memory request latency breakdown of FFT for both DDR3 SDRAM and OC-DIMM systems. A 2-channel DDR3 SDRAM system yielded a lower average memory latency compared to a 1-channel DDR3 SDRAM. The significant decrease in queuing delay is due to the fact that a multi-channel memory system can handle more requests concurrently. Adding more DIMMs to a channel resulted in modest decrease in latency.

As expected, 2NB and 4NB OC-DIMM systems performed better than the other OC-DIMM organizations. 16NB systems experienced long transmission times which have become a significant portion of the total memory request latency. 8NB systems' significant queuing delay is due to the fact that they are limited by the SB bus bandwidth. Memory requests are not being scheduled immediately because of contention for the SB bus, forcing them to wait longer in the queue.

The summary plot of figure 8 compares the best-case average memory request latency of a 1- and 2-channel DDR3 SDRAM with OC-DIMM. OC-DIMM reduced the average memory request latency by 29% compared to a 1-channel DDR3 SDRAM system. A 2-channel DDR3 SDRAM system can reduce the average memory latency by 7.66% at the cost of doubling the pin count for the extra memory channel.

Figure 9 shows the execution time of the FFT program across the different memory configurations. A performance speedup of 1.20 is gained by using OC-DIMM compared to a 1-channel DDR3 SDRAM system. A DDR3 SDRAM system has to utilize 2 memory channels in order to perform better than the OC-DIMM system. A 2-channel DDR3 SDRAM system can provide a program speedup of 1.03% over OC-DIMM.

The next figures show the average memory request latencies and execution times for Radix (figures 10 and 11) and Canneal (figures 12 and 13). These figures also show over-

all performance speedup in using the OC-DIMM memory system compared to a standard 1-channel DDR3 memory.

VI. Conclusion

Multiwavelength optical buses can provide high bandwidth density than electrical interconnects and can scale well with the increasing memory bandwidth demand of future multicore systems. Our OC-DIMM memory system leverages this technology for processor-DRAM communication. Organizing multiple wavelengths into groups to form multiple independent data channels helps reduce the queuing delay of memory requests because of increased concurrency. Full-system simulations show that, for memory-intensive multithreaded workloads on an 8-core CMP, a single-waveguide OC-DIMM system yields a memory request latency reduction of up to 29% and a performance speedup of up to 1.20 compared to conventional single-channel DDR3 SDRAM systems. In the future, we plan to extend this work to include a detailed system-level power analysis of the memory subsystem.

Acknowledgment

This work was supported in part by the CITRIS seed grant #3-34PICOB.

References

[1] N. E. Jerger and L.-S. Peh, "On-chip networks," *Synthesis Lectures on Computer Architecture*, vol. 4, no. 1, pp. 1–141, 2009.

[2] D. A. Patterson, "Latency lags bandwith," *Commun. ACM*, vol. 47, no. 10, pp. 71–75, 2004.

[3] G. M. Amdahl, "Architectural trends in large systems," *SIGMICRO Newsl.*, vol. 2, no. 4, pp. 17–29, 1972.

[4] ITRS, "Int'l. technology roadmap for semiconductors, 2007 edition, assembly and packaging," 2007.

[5] B. Black *et al.*, "Die stacking (3D) microarchitecture," in *Int'l. Symp. on Microarchitecture*, 2006, pp. 469–479.

[6] M. Tan *et al.*, "A high-speed optical multi-drop bus for computer interconnections," in *Symp. on High Performance Interconnects*, 2008, pp. 3–10.

[7] C. Batten *et al.*, "Building manycore processor-to-DRAM networks with monolithic silicon photonics," in *Symp. on High Performance Interconnects*, 2008, pp. 21–30.

[8] D. Vantrease *et al.*, "Corona: System implications of emerging nanophotonic technology," in *Int'l. Symp. on Computer Architecture*, 2008, pp. 153–164.

[9] A. Hadke *et al.*, "Design and evaluation of an optical CPU-DRAM interconnect," Oct. 2008, pp. 492–497.

[10] ——, "OCDIMM: Scaling the DRAM memory wall using WDM based optical interconnects," in *Symp. on High Performance Interconnects*, 2008, pp. 57–63.

[11] G. Hendry *et al.*, "Analysis of photonic networks for a chip multiprocessor using scientific applications," in *Int'l. Symp. on Networks-on-Chip*, 2009, pp. 104–113.

[12] J. Haas and P. Vogt, "Fully-buffered DIMM technology moves enterprise platforms to the next level," *Technology@Intel Magazine*, 2005.

[13] Intel, "Intel 875P chipset datasheet," Feb. 2004.

[14] P. Magnusson *et al.*, "Simics: A full system simulation platform," *Computer*, vol. 35, no. 2, pp. 50–58, Feb 2002.

[15] M. M. K. Martin *et al.*, "Multifacet's general execution-driven multiprocessor simulator (GEMS) toolset," *SIGARCH Comput. Archit. News*, vol. 33, no. 4, pp. 92–99, 2005.

[16] D. Wang *et al.*, "DRAMsim: A memory system simulator," *SIGARCH Comput. Archit. News*, vol. 33, no. 4, pp. 100–107, 2005.

[17] Micron, "4Gb: x4, x8 TwinDie DDR3 SDRAM functionality," 2008.

[18] S. C. Woo *et al.*, "The SPLASH-2 programs: Characterization and methodological considerations," in *Int'l. Symp. on Computer Architecture*, 1995, pp. 24–36.

[19] C. Bienia *et al.*, "The PARSEC benchmark suite: Characterization and architectural implications," in *Int'l. Conf. on Parallel Architectures and Compilation Techniques*, 2008, pp. 72–81.

[20] N. Kirman *et al.*, "Leveraging optical technology in future bus-based chip multiprocessors," in *Int'l. Symp. on Microarchitecture*, 2006, pp. 492–503.

[21] G. Chen *et al.*, "Predictions of CMOS compatible on-chip optical interconnect," in *Int'l. Workshop on System Level Interconnect Prediction*, 2005, pp. 13–20.

[22] B. Jacob, S. W. Ng, and D. T. Wang, *Memory Systems: Cache, DRAM, Disk*. Morgan Kaufmann, 2007.

[23] T. Pinkston and J. Duato, "Interconnection networks," in *Computer Architecture, Fourth Edition: A Quantitative Approach*, J. L. Hennessy and D. A. Patterson, Eds. Morgan Kaufmann Publishers Inc., 2006.