On-Chip Optical Technology in Future Bus-Based Multicore Designs

This work investigates the integration of CMOS-compatible optical technology to on-chip coherent buses for future CMPs. The analysis results in a hierarchical optoelectrical bus that exploits the advantages of optical technology while abiding by projected limitations. This bus achieves significant performance improvement for high-bandwidth applications relative to a state-of-the-art fully electrical bus.

Nevin Kırman
Meyrem Kırman
Rajeev K. Dokania
José F. Martínez
Alyssa B. Apsel
Matthew A. Watkins
David H. Albonesi
Cornell University

Current trends indicate that future chip multiprocessors (CMPs) may comprise tens or even hundreds of processing elements. Feeding data to so many on-chip cores, however, will be possible only if architecture and technology developments provide sufficient chip-to-chip and on-chip communication performance.

Optical technology and 3D integration are two potential solutions to *chip-to-chip* communication performance limitations. Still, *on-chip* communication faces considerable technological and architectural challenges. For example, global electrical interconnects do not scale well with technology.[1] Although delay-optimized repeater insertion and proper wire sizing can keep the delay nearly constant, this comes at the expense of power and active area, as well as a reduction in wire count (and thus bandwidth). Researchers have developed techniques for optimizing the power-delay product, but these techniques yield neither optimal power nor optimal latency. This and other technological issues—such as manufacturability, conductivity, crosstalk, and so on—constitute important roadblocks (*ITRS—International Technology Roadmap for Semiconductors,* http://public.itrs.net). As more cores are integrated, we expect the on-chip interconnect to take an increasingly larger fraction of chip area and power budgets.

Whereas 10 years ago electrical-to-optical translation costs and CMOS incompatibility appeared to be insurmountable barriers to the use of optics in on-chip communication, today the outlook is dramatically brighter. Because of rapid progress in the past five years in CMOS-compatible detectors, modulators, and even light sources, the latest *ITRS* considers on-chip optical interconnects as a potential replacement for global wires by 2013. In global-signaling applications, optical interconnects have the potential to fare favorably compared to their electrical counterparts, owing to their high speed, high bandwidth, low on-chip power,

Published by the IEEE Computer Society

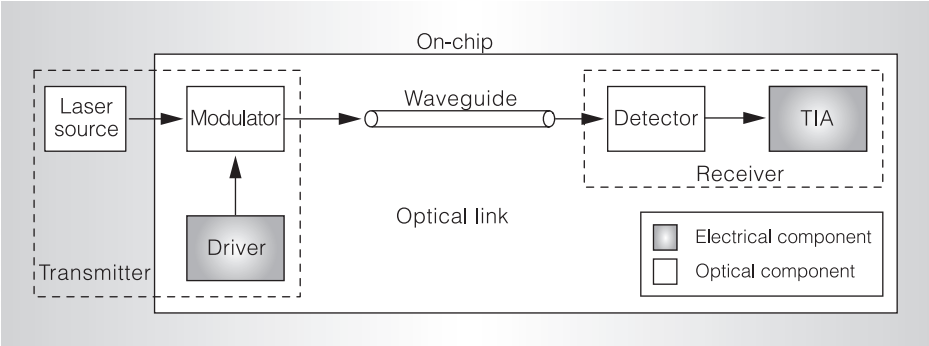

Figure 1. Simplified diagram showing the main components involved in modulator-based on-chip optical transmission.

good electrical isolation, low electromagnetic interference, and other characteristics.[2]

Although the technology is admittedly still in its early stages, there is enough data and understanding of on-chip, CMOS-compatible optical components to consider the broader architectural trade-offs in designing an on-chip optical network for future high-performance microprocessors. In this article, we investigate the potential of implementing a low-latency, high-bandwidth shared snoopy bus in future CMPs using optical technology. Through a carefully projected case study for a 32-nm CMP, we conduct the first exploration and evaluation of on-chip optical buses for this application, and provide insight into the potential advantages and limitations of the technology for catalyzing future interdisciplinary work.

Optical technology overview

As Figure 1 shows, on-chip, modulator-based optical transmission uses a transmitter, a waveguide, and a receiver.

Transmitter

Optical transmission requires a laser source, a modulator, and a modulator driver circuit. The laser source provides light to the modulator, which transduces electrical information (supplied by the modulator driver) into a modulated optical signal.

In this work, we opt for an off-chip laser source because it saves on-chip power, area, and cost. As the light enters the chip, optical splitters and waveguides (not shown in Figure 1) route it to different modulators for data transmission. (These distribution paths are a source of signal losses, as we will discuss later.)

We assume a recently proposed resonator-based modulator implementation that features low operating voltage and compact size.[3] Researchers have already proposed 10-µm-diameter ring-shaped modulators, and they are likely to shrink further.

Modulator performance depends in part on the *extinction ratio*—the on-to-off light intensity ratio. Higher extinction ratio is better for proper signal detection. Extinction ratio could limit the number of transmitters that can time-share the same wavelength on the same channel.

The modulator driver comprises a series of inverter stages driving the modulator's capacitive load. A smaller capacitance improves the overall transmitter's power and latency, requiring fewer stages. We assume a modulator capacitance of 50 fF, even though it is expected to get smaller.

Waveguide

Waveguides are the paths through which light is routed. For on-chip applications, silicon and polymer are the most promising waveguide materials. Our work focuses on silicon waveguides.[4]

Receiver

An optical receiver performs the optical-to-electrical conversion of the light signal. It consists of a photodetector and a transimpedance amplifier (TIA) stage. In wavelength division multiplexing (WDM)—simultaneous transmissions at different wavelengths

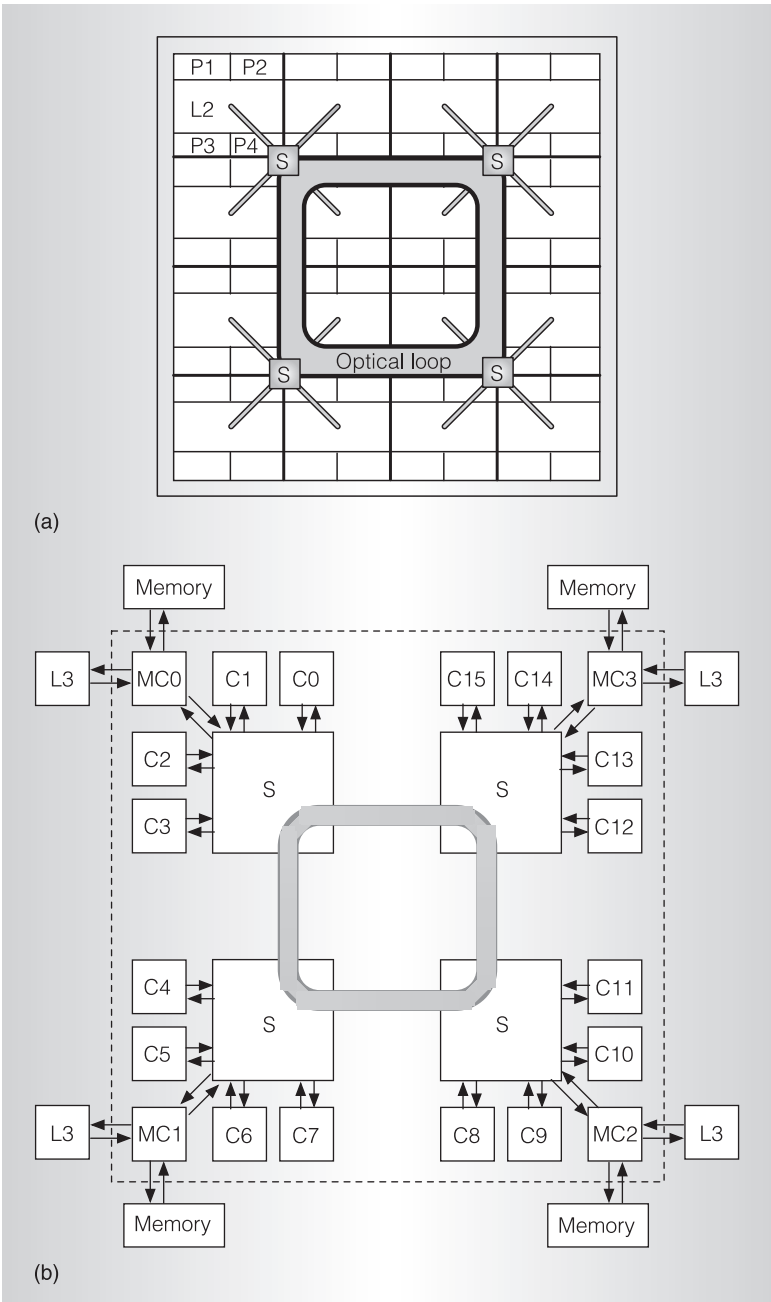

Figure 2. Simplified CMP floorplan diagram (a) and high-level system organization (b), showing the optical loop and the rest of the hierarchical bus. S: switch (separate switches for address/snoop and data buses); MC(0-3): memory controller; C(0-15): L2 cache.

in a waveguide—the receiver also requires a wavelength-selective filter.

Photodetector quantum efficiency is important: High quantum efficiency means lower losses when converting optical into electrical information. Detector size is also important for compactness and next-stage capacitance. Typically, the detector has a large base capacitance and poses a design challenge for subsequent high-speed gain stages. We assume a 100-fF detector capacitance in our study, which is achievable even with current technologies.

The TIA stage converts photodetector current to a voltage that subsequent stages threshold to digital levels. To achieve high-gain and high-speed detection, an analog supply voltage higher than the digital supply voltage might be necessary, thereby requiring higher power. We assume a TIA supply voltage 20 percent higher than the nominal supply.

Optoelectrical bus architecture

We target a 32-nm process technology and assume a 400-mm^2 die that accommodates 64 four-issue out-of-order cores, with enough additional space to allocate L2 caches, interconnect, and other system components; we also assume a 4-GHz core frequency.[4]

We opt for 16 L2 caches, each shared among four cores. We reasonably assume the availability of chip-to-chip optical technology, and set off-chip bandwidth to 256 Gbytes/s to L3 and 128 Gbytes/s to memory.

Optical medium

Optical waveguides do not lend themselves gracefully to H-tree or highly angled structures common in electrical topologies, because turns and waveguide crossings can result in significant signal degradation. Thus, we propose building upon a simple loop-like structure, which is far better suited to the structural characteristics of optical waveguides. This loop-shaped bus consists of optical waveguides (residing on a dedicated silicon layer) that encircle a large portion of the chip, as Figure 2 shows. Multiple nodes connected to the bus, each issuing transactions for a processor or a set of processors, are equipped with necessary transmitters and receivers to interface with the optical medium.

We use WDM to realize a multibus. Multiplexing the multibus by address is achieved by partitioning the

available wavelengths among different address spaces. In contrast, in multiplex by node, each node has exclusive access to a subset of wavelengths. Multiplex by node has several advantages, including the following:

- Broadcasting on the optical bus does not require global arbitration.
- Placing an attenuator immediately before each modulator can remove the leftover light signal after going full circle.
- It needs fewer transmitters and receivers (which consume area and power).
- It's possible to optimize the required light power through individual coupling-ratio tuning at detectors during design to absorb just the right fraction of light power, because the relative position of each detector with respect to the (sole) transmitter is known for every wavelength.
- Extinction ratio constraints are less strict, because there is a single transmitter on each wavelength.

The downside of multiplex by node is that the number of nodes directly connected to the bus is, at best, limited to the number of wavelengths. In our study, we opt for a multiplex-by-node organization.

Bus organization

We propose an optoelectrical hierarchical bus, where the optical loop constitutes the top level of the hierarchy, and nodes deliver information to processors via electrical sublevels. Figure 2 depicts a possible four-node organization for our 64-processor CMP. We assume 64 bits for address and command, 72 bits for data, and 8 bits per snoop response. Node switches arbitrate among the incoming coherence requests from the L2 caches and broadcast the winner or winners on the optical address bus. Then, nodes arbitrate among the concurrent requests (all reaching the same outcome independently).

Next, the selected requests are delivered to all caches simultaneously, and the rest are retried later. Snoop responses from the caches are combined in the switches first locally and then globally after being broadcasted on the optical snoop bus; then the final snoop result propagates up to the caches. Eventually, if appropriate, data is sent down the optical data bus (undergoing arbitration), which the original requesting node collects and sends up the requesting L2 cache.

Topologies. We explore a range of 4 to 12 available wavelengths per waveguide,[4] and investigate several possible bus topologies, deriving for each of them area and power (listed in Table 1). In Table 1, H-$n\times k$AkD (H for hierarchical) designates a topology with n nodes on the optical bus and k address and data wavelengths per node, totaling nk wavelengths per waveguide in the address and data buses. We sweep through all possible configurations given the WDM projections: $k \in \{1, 2, 3\}$ for $n = 4$, and $k = 1$ for $n = 8$. In the cases of four nodes with $k > 1$, we also investigate topologies with more limited support for new address transactions per cycle—H-4×1AkD. Similarly, in the case of eight nodes, we explore reducing the electrical snoop bandwidth to 4 (H-8×1A1D(4S) where 4S stands for four snoops). This should generally result in area and power savings.

Frequency estimation. We estimate the bus operating frequency by calculating the time needed for the light to travel from any node to the farthest node on the (unidirectional) optical loop, so that the bus can transmit information to all nodes in one cycle. With the loop bus centered on the die (as Figure 2 shows), using the waveguide and optical-component delays provided in Table 2, and accounting for a 4 FO4 latching delay (based on *ITRS* data), we estimate that all buses can run safely at 2 GHz—half the cores' frequency.

Area estimation. For each organization, we estimate the required area on the active, optical, and metal layers. These area calculations, listed in Table 1, consider all address, snoop, and data buses. In the active area, we account for electrical switches in

Table 1. Area and power characterization of different optical bus topologies.

Optical bus topology	Snoop requests per bus cycle	Area (mm^2)				Power* (W)					
		Active silicon				Electrical level		Optical level		Total on-chip	
		Switch	Tx and Rx	Metal layer	Optical layer	Switch	Wiring	Tx and Rx	Optical	$\alpha = 1$	$\alpha = 0.5$
H-4×1A1D	4	1.71	0.39	15.21	33.68	1.75	12.82	0.60	0.79	15.56	9.04
H-4×2A2D	8	2.72	0.78	24.42	34.10	3.03	20.59	1.19	1.58	25.60	15.13
H-4×3A3D	12	4.00	1.17	33.64	34.51	4.64	28.36	1.79	2.37	35.98	21.49
H-4×1A2D	4	1.93	0.56	15.21	33.86	2.06	12.82	0.85	1.13	16.30	9.73
H-4×1A3D	4	2.13	0.72	15.21	34.04	2.37	12.82	1.11	1.47	17.03	10.41
H-8×1A1D	8	4.05	1.89	12.21	51.64	4.50	10.30	3.07	6.35	21.05	15.44
H-8×1A1D(4S)	4	3.08	1.59	7.60	51.30	3.25	6.41	2.58	5.33	14.91	11.34

Tx: transmitter; Rx: receiver; α: switching activity factor
* Total on-chip power is the sum of switch, wiring, Tx and Rx, and half of the optical power components (because of a 3-dB coupling loss, only half of the optical power is actually consumed on chip). All dynamic power components in switching, wiring, and Tx-Rx columns assume $\alpha = 1$. For $\alpha = 0.5$, only the total sum is provided.

each node, as well as transmitters and receivers on the optical bus. We use Orion power-performance simulator to estimate switch areas. We estimate the active area taken up by transmitters and receivers by using their total counts in the system and conservatively assuming that the modulator occupies 80 μm^2; the modulator driver, 50 μm^2; the photodetector, 100 μm^2; and the TIA, 50 μm^2. We calculate the area occupied in the optical layer as the sum of waveguide, modulator, detector, and wavelength-selective filter areas (80 μm^2). We assume the component areas just specified and a 5.5-μm silicon waveguide pitch. The resulting active area is relatively modest, and the required optical-layer area easily fits within 400 mm^2.

Finally, we estimate the metal wiring area required for the electrical subinterconnects. We assume a global wire pitch of 400 nm, and wire lengths of 4.5 mm for four-node configurations and 2.25 mm for eight-node configurations (estimated according to the floorplan in Figure 2). From each cache to its node, the links include single address and data paths and as many snoop-response paths as needed in each topology (number of snoop requests per cycle in Table 1). From each node to a cache, the links include a single data path and as many snoop-request and snoop-result paths as indicated in Table 1.

Power estimation. Table 1 also shows a detailed breakdown of power consumption in the electrical sublevels (switches and wiring) and in the optical bus for all topologies under consideration. We estimate the static and dynamic power consumed by the switches again using the Orion power-performance simulator. We use the meth-

Table 2. Delays of various optical components at different technology nodes.[5]

Component	Delay (ps)		
	45-nm	32-nm	22-nm
Modulator driver	25.8	16.3	9.5
Modulator	30.4	20.0	14.3
Detector	0.6	0.5	0.4
Amplifier	10.4	6.9	4.0
Silicon waveguide	10.45 per mm	10.45 per mm	10.45 per mm

odology described by Ho, Mai, and Horowitz[1] for power-delay optimized repeater insertion, and the methodology described by Ho[6] for wire sizing when estimating the static and dynamic power consumed by wires.

Two main power components are attributable to the optical loop: electrical and optical power. Power consumed by the modulator drivers (117 μW per driver) and TIAs (257 μW per TIA) contribute to the electrical on-chip power. We use *ITRS* device projections and standard circuit procedures to estimate per-component power values.

Optical power is the off-chip power that the modulator requires to modulate and transmit the information optically from one node to the others. In our analysis, we first calculate the minimum optical threshold power required to detect a signal correctly, which is based on the receiver's output voltage swing and signal-to-noise ratio requirement, as suggested by O'Connor.[7] In our case, the minimum detector current requirement comes to 30 μA. It is possible to design the detectors to tap only the minimum amount of power adequate for signal detection, resulting in minimum overall optical power. Beginning with the minimum power required at the farthest receiver in the optical loop, we calculate the input power required at the transmitter's modulator by visiting nodes in reverse order up to the transmitter, and accumulating at each step the tapping power at detectors and the power losses incurred. Each modulator requires this amount of optical power, and because we assume a continuous laser source, that power will be always consumed, regardless of whether data is being transmitted. Table 3 lists the major power losses. We assume the photodetector to have a quantum efficiency of 0.8.[7]

We account for the remaining losses in the optical system—such as those due to on-chip coupling, splitters, and so on; Table 1 reports the minimum required total optical power for each configuration. Note that only half of this optical power contributes to the total on-chip power consumption; the other half is lost during light coupling into the chip.

Table 3. Major power losses incurred by an on-chip optical transmission system.

Source	Loss (dB)
On-chip coupling*	3.0
Silicon waveguide*	1.3 per cm
Splitter*	0.2
Modulator insertion**	1.0
Interlayer coupling	1.0
Bending*	0.5

* O'Connor.[7]
** Almeida et al.[8]

Discussion. The preferable topologies in terms of area and power are H-4×1A{1,2,3}D and H-8×1A1D(4S), although we observed empirically that the data bandwidth of H-4×1A1D is too low. In comparison, all the other configurations have excessive power and area expenses, owing to a variety of factors: higher snoop bandwidth, greater number of receivers and transceivers, larger switch crossbars and arbitration logic, and so on. In the four-node configuration, the power consumption of the optical components is relatively low compared to that of the electrical subnetwork.

Among the preferred organizations, H-4×1A2D and H-4×1A3D require lower laser power and are more flexible, because they can dynamically allocate the wavelengths for requests from every four L2 caches.

Evaluation

We evaluate the performance impact of the proposed optical bus by comparing it with a state-of-the-art electrical bus. We use cycle-accurate, execution-driven simulations to model a 64-core CMP featuring dynamic superscalar cores and a snoopy-coherent memory subsystem. Each four-way, out-of-order core runs at 4 GHz, and each has access to a four-issue, private, write-through L1 data cache. Every four cores share an eight-way, banked, write-back L2 cache through a crossbar. The snoopy, fully pipelined bus that we are studying connects all 16 L2 caches. We use a MESI protocol that permits cache-to-cache clean block transfers. A banked, shared L3 cache resides

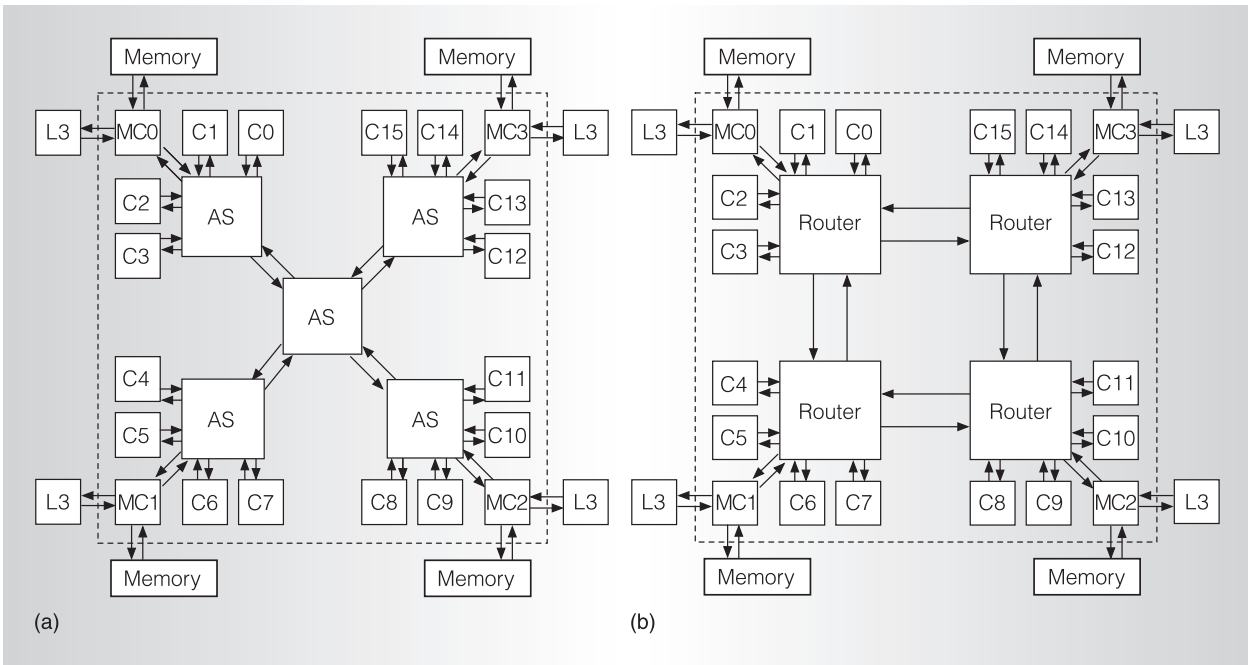

Figure 3. Modeled electrical baseline address network (a) and data network (b). AS: address switch; MC(0-3): memory controller; C(0-15): L2 cache. (Not to scale.)

off chip, but with tags on chip. L3 is accessed in parallel with main memory, and it is exclusive of L2 caches. All caches have 64-byte cache block size. We model four on-chip L3 and memory controllers, each connecting to one-fourth of L3 and memory via 64-Gbyte/s and 32-Gbytes/s links, respectively. Memory round-trip latency from the memory controllers is 320 processor cycles.

We use 11 applications from the Splash-2 suite, all with default input data sizes except for the cholesky application, for which we use the tk29.O input file. Following common practice for Splash-2 applications, we use reduced cache sizes to compensate for the applications' reduced working sets: 64 × 8 Kbytes for L1; 16 × 256 Kbytes for L2; and 1 × 16 Mbytes for L3. We use MIPS binaries compiled with -O3 optimization level. We fast-forward the initialization part of the applications and run them to completion.

Electrical bus

For the comparison, we establish a state-of-the-art electrical baseline with power and active and metal area characteristics similar to those of the competing optoelectrical buses.

In our electrical baseline, the address bus is a hierarchical tree organization (similar to a single snooping coherence domain in the Sun Fireplane system interconnect[9]) that yields low latency and competitive bandwidth relative to other alternatives for our configuration. As Figure 3a shows, in the system four L2 caches and a memory controller (which manages one-fourth of the off-chip L3 and memory) connect to an address switch (AS), and four such address switches connect to a top-level address switch—all through point-to-point links. Requests issued by L2 caches are arbitrated in the switches at each level of the tree, until they reach the top level and are selected. From that point on, broadcasting a snoop request down to all caches, combining snoop responses up at the top-level switch, and again broadcasting the final snoop result down to the caches takes a fixed number of cycles. We implement a multibus by selecting multiple snoop requests at the top-level address switch and employing as many snoop request and response buses as needed.

Table 4. Area and power characterization of two possible topologies for the baseline electrical bus.

Snoop requests per bus cycle	Area (mm^2)		Power (W)		Total on-chip power* (W)	
	Switches and routers	Wiring	Switches and routers	Wiring	$\alpha = 1$	$\alpha = 0.5$
2	1.47	15.9	1.42	13.40	14.82	8.08
4	1.66	22.81	1.68	19.23	20.91	11.29

* Total on-chip power is the sum of all electrical power components. Dynamic power components in switching and wiring columns assume $\alpha = 1$. For $\alpha = 0.5$, only the total sum is provided.

We assume an H-tree layout, and by using power-delay-optimized, repeatered wires, we can accommodate a 2-GHz bus clock frequency—half the cores' speed. Under no contention, a request's address phase spends a total of 13 bus cycles on the bus: 4 bus cycles for request arbitration, 3 for snoop request, and 6 for snoop-response combining and result broadcasting (excluding time spent in the caches).

The data network, shown in Figure 3b, consists of a four-node bidirectional ring. As with the address switches, each data router serves requests from and to four local caches and a memory controller connected to it through point-to-point links. Routing is deterministic and balanced. Transfers within a node use a bypass path within the local router. In the absence of contention, it takes 14 bus cycles to transfer a cache line on the data network to a cache in the farthest node.

Following the estimation methodology we use for the optical buses, we obtain the area and power characteristics for the electrical bus, listed in Table 4. When compared to buses H-4×1A{1,2,3}D, an electrical bus with support for an equal number of snoop requests per bus cycle (four) exhibits comparable power consumption and active device area, but a 50 percent increase in metal area overhead. On the other hand, an electrical baseline with support for half as many snoop requests per bus cycle has area and power characteristics similar to its optoelectrical counterparts. Thus, for our electrical baseline, we choose the configuration supporting half as many snoop requests per bus cycle.

Optoelectrical bus

For the optoelectrical bus, we model configuration H-4×1A3D (described earlier). The uncontended latencies are 10 bus cycles for arbitration plus snoop-request and snoop-response phases, and 12 bus cycles to transfer cache line data on the bus across bus nodes.

Results

Figure 4a shows speedups of the optoelectrical system relative to the electrical baseline. In general, the optoelectrical configuration provides high data bandwidth via WDM, achieving significant speedups: a geometric mean of 1.13, and a peak of 1.71. The performance improvements closely correlate to the global L2 miss rates, also provided in Figure 4a.

To further clarify the sources of performance improvement, Figure 4b shows the average latency breakdown (in bus cycles) of bus transactions in the two configurations. (In the plots, the data transfer category excludes memory or cache access times.)

We observe latency advantages for the optoelectrical configurations in both the address-snoop and data networks. In the address-snoop network, moving from electrical to electro-optical technology reduces effective latency by 22 percent on average (34 to 28 bus cycles). Recall that, even in the absence of contention, the optoelectrical buses have a latency advantage over our electrical baseline. Moreover, the optoelectrical buses can support twice as much snoop-request and snoop-response bandwidth as the electrical baseline using similar power and area. Some applications (barnes, radiosity, raytrace, and water-spatial) have significant contention in the arbitration phase. Our simulations show that this is caused mostly by the serialization of conflicting requests to the same cache line (in our bus protocol, conflicting requests to

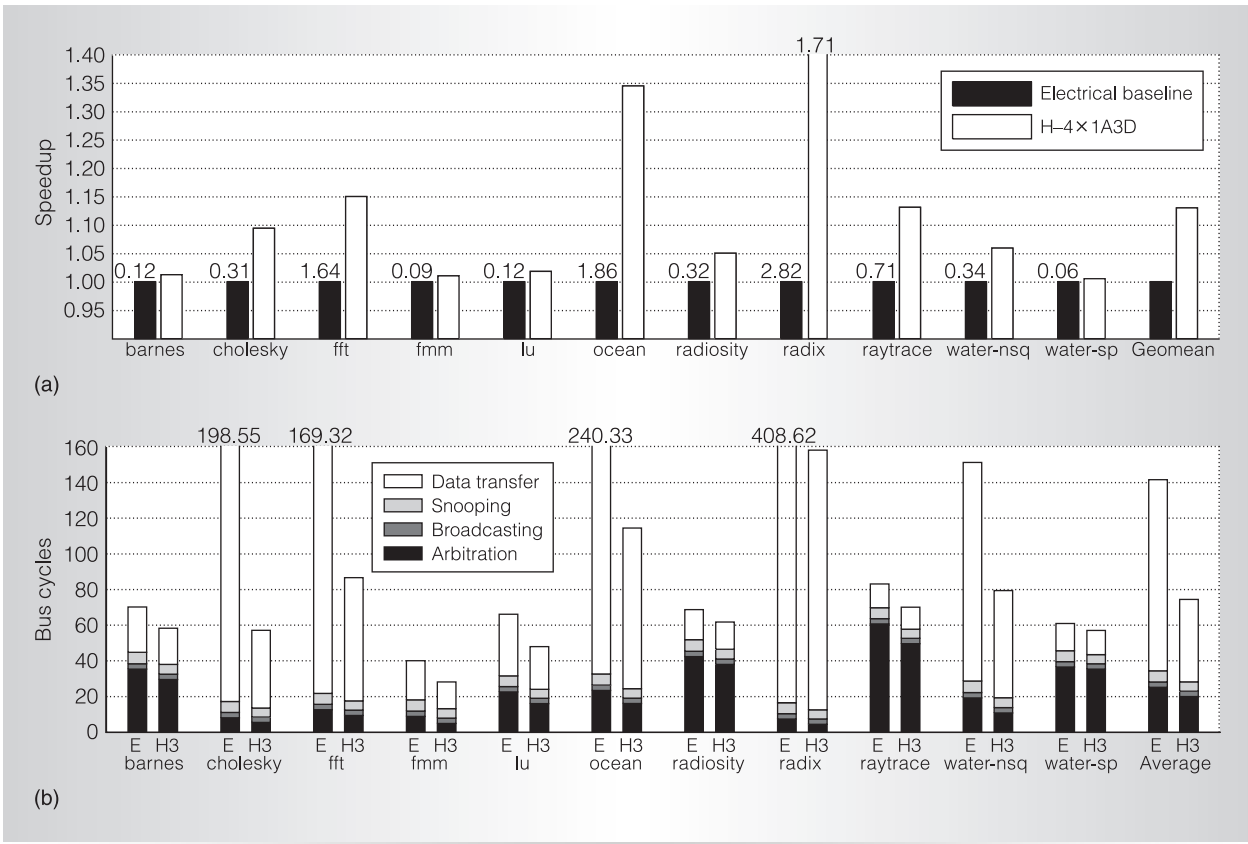

Figure 4. Speedups achieved by H-4×1A3D relative to the electrical baseline, for the Splash-2 suite (a) and average latency breakdown (in bus cycles) of bus transactions in baseline electrical (E) and H-4×1A3D (H3) buses (b). In (a), values on top of the black bars indicate the average global L2 miss rates (percent). In (b), data transfer excludes cache or memory access times.

a cache line with an outstanding request are deferred). This serialization can be optimized at the protocol level. Indeed, for the configurations under study, the main overall benefit comes from reduced contention (and thus effective latency) for data transfers. The data network struggles to supply the bandwidth needed to satisfy these requests. It is in the data network that the availability of extra wavelengths through WDM yields the largest performance improvements. Still, some applications suffer from significant contention in the data network even for H-4×1A3D, leaving room for further improvement. We identify the main cause to be contention at the L2 cache input and output ports. Notice that the bandwidth to and from the caches (and memory controller) is unchanged in all configurations despite the increased data bandwidth on the optical loop. Also, those higher-contention applications would benefit from additional wavelengths.

Our evaluation shows that incorporating optical technology in bus-based CMPs can benefit performance, and that WDM support could be a critical ingredient to this benefit in both address-snoop and data networks. The fact that WDM comes at very small additional area and power is encouraging. In the particular design points that we evaluated, the data network's contribution to performance turned out to be dominant.

Future CMPs with tens to hundreds of cores will demand very-high-performance, power-efficient on-chip interconnects, for which CMOS-compatible optics technology offers a promising solution. A

hybrid optoelectrical CMP interconnect can provide significant speedups for high-bandwidth applications while consuming reasonable on-chip power. Additional interdisciplinary research involving microarchitects and optical device researchers will be required for the technology to become a practical CMP interconnect solution.

An area of interest involves the best use of optical waveguides and WDM for a particular CMP design. On the one hand, our results indicate that increasing the number of WDM channels alone can significantly impact performance for bus-intensive applications in the CMP configuration studied. On the other hand, other CMP organizations (for example, more nodes on the interconnect) might require other organizations that present very different power-performance and complexity challenges—number of modulators, power and bandwidth requirements at the nodes, and so on.

The performance and system-power improvements obtainable using optics is limited by how far the technology penetrates into the bus design. Our hierarchical approach, for instance, addresses a fraction of the bus latency via optics, but a large fraction remains entirely electrical. This shortcoming poses an interesting opportunity for joint research at the bus design and optics component fronts. Another interesting area for exploration is the temperature management of the optical components, which is a critical system-level issue because the optical modules are very sensitive to temperature variations.

Finally, we believe that other interesting optical network topologies can be explored—possibly including flat switch-type networks. MICRO

References

1. R. Ho, W. Mai, and M.-A. Horowitz, ''The Future of Wires,'' *Proc. IEEE*, vol. 89, no. 4, Apr. 2001, pp. 490-504.
2. D.A. Miller, ''Rationale and Challenges for Optical Interconnects to Electronic Chips,'' *Proc. IEEE*, vol. 88, no. 6, June 2000, pp. 728-749.
3. C.A. Barrios, V.R. de Almeida, and M. Lipson, ''Low-Power-Consumption Short-Length and High-Modulation-Depth Silicon Electrooptic Modulator,'' *J. Lightwave Tech*, vol. 21, no. 4, Apr. 2003, pp. 1089-1098.
4. N. Kırman et al., ''Leveraging Optical Technology in Future Bus-Based Chip Multiprocessors,'' *Proc. Int'l Symp. Microarchitecture* (Micro 06), IEEE CS Press, 2006, pp. 492-503.
5. G. Chen et al., ''Predictions of CMOS Compatible On-Chip Optical Interconnect,'' *Proc. Int'l Workshop System-Level Interconnect Prediction*, ACM Press, 2005, pp. 13-20.
6. R. Ho. *On-Chip Wires: Scaling and Efficiency*, doctoral dissertation, Dept. of Electrical Engineering, Stanford University, 2003.
7. I. O'Connor, ''Optical Solutions for System-Level Interconnect,'' *Proc. Int'l Workshop System-Level Interconnect Prediction*, ACM Press, 2004, pp. 79-88.
8. V.R. Almeida et al., ''All-Optical Switching on a Silicon Chip,'' *Optics Letters*, vol. 29, no. 24, Dec. 2004, pp. 2867-2869.
9. A. Charlesworth, ''The Sun Fireplane System Interconnect,'' *Proc. ACM/IEEE Conf. Supercomputing*, ACM Press, 2001, pp. 1-14.

Nevin Kırman is a PhD student in electrical and computer engineering at Cornell University. Her research interests include high-performance multicore architectures, interconnects, and memory systems. She received BS degrees in electronics and communication engineering and in computer engineering from Istanbul Technical University, Turkey. She is the recipient of an HPCA Best Paper Award and an Intel fellowship (both in 2005). She is a member of the IEEE and the ACM.

Meyrem Kırman is a PhD student in electrical and computer engineering at Cornell University. Her research focuses on high-performance multicore architectures, interconnects, and memory systems. She received BS degrees in control and computer engineering, and in electronics and communication engineering, both from Istanbul Technical University, Turkey. She is the recipient of an HPCA Best Paper Award (2005), a McMullen fellowship

(2003), and an Intel fellowship (2006). She is a member of the IEEE and the ACM.

Rajeev K. Dokania is pursuing MS and PhD degrees as a member of the Optoelectronic VLSI group in the School of Electrical and Computer Engineering at Cornell. Previously, he worked for two years at Intel in DFx on 10-Gb Ethernet router and multicore server processor designs. His research interests include circuit modeling, low-voltage process-invariant analog circuit design, optical interconnects, DFX architecture, and high-speed I/O designs for optical communication. He received his BTech degree in electrical engineering from the Indian Institute of Technology, Kharagpur, India.

José F. Martínez is an assistant professor of electrical and computer engineering at Cornell University and a member of the Computer Systems Laboratory. His research interests include novel technologies for microprocessor design as well as reconfigurable and self-optimizing architectures. Martínez is the recipient of an NSF Career Award, an IBM Faculty Award, an HPCA '05 Best Paper Award, and a Kenneth A. Goldman '71 Teaching Excellence Award. His work on speculative synchronization appeared in IEEE Micro Top Picks in 2003. He received a PhD in computer science from the University of Illinois at Urbana-Champaign. He is a member of the IEEE, the ACM, the AAAS, and the ASEE.

Alyssa B. Apsel is the Clare Luce Boothe Assistant Professor of Electrical and Computer Engineering at Cornell University, where she heads the Optoelectronic VLSI Laboratory. Her current research interests include optoelectronic CMOS systems and architectures, short-distance interconnects, low-power RF systems, and high-speed CMOS interfaces. She received a PhD from the Johns Hopkins University. She holds two patents, received a best paper award at the 2000 Midwest Symposium on Circuits and Systems, a best paper award at ASYNC 2006, and an NSF Career Award in 2004, and was named to *MIT Technology Review*'s top 100 young innovators for 2004. She is also an associate editor of *IEEE Transactions on Circuits and Systems I,* a member of the Science and Engineering Council of the OSA, and a member of the technical program committees for SPIE Optics East, Photonics West, IEEE LEOS meeting, and IEEE ISCAS.

Matthew A. Watkins is pursuing MS and PhD degrees in electrical and computer engineering at Cornell University. His current research addresses on-chip optical interconnects and their architectural implications for future microprocessors. Watkins received BS degrees in computer engineering and electrical engineering from the University at Buffalo. He is a member of the IEEE.

David H. Albonesi is an associate professor of electrical and computer engineering at Cornell University and a member of the Computer Systems Laboratory. His research interests include adaptive, power-efficient, and reliability-aware computer architectures; multithreaded and multicore microprocessors; and microarchitectures exploiting new technologies. He received his PhD in 1996 from the University of Massachusetts, Amherst. He has received the NSF Career Award and IBM Faculty Partnership awards. He is a senior member of the IEEE and a member of the ACM. He is also *IEEE Micro*'s new Editor in Chief.

Direct questions about this article to Nevin Kırman, Computer Systems Laboratory, Cornell University, Ithaca, NY 14853; nkirman@csl.cornell.edu.