

## Research Highlights

# A Nanophotonic Interconnect for High-Performance Many-Core Computation

R. G. Beausoleil<sup>1</sup>, J. Ahn<sup>2</sup>, N. Binkert<sup>2</sup>, A. Davis<sup>2</sup>, D. Fattal<sup>1</sup>, M. Fiorentino<sup>1</sup>, N. P. Jouppi<sup>2</sup>, M. McLaren<sup>3</sup>, C. M. Santori<sup>1</sup>, R. S. Schreiber<sup>2</sup>, S. M. Spillane<sup>2</sup>, D. Vantrease<sup>2</sup>, and Q. Xu<sup>1</sup>

<sup>1</sup> INFORMATION AND QUANTUM SYSTEMS, HP LABORATORIES, 1501 PAGE MILL RD., MS 1123, PALO ALTO, CA 94304-1123

<sup>2</sup> EXASCALE COMPUTING, HP LABORATORIES,

1501 PAGE MILL RD., MS 1177, PALO ALTO, CA 94304-1177

<sup>3</sup> EXASCALE COMPUTING, HP LABORATORIES, FILTON ROAD, STOKE GIFFORD, BRISTOL BS34 8QZ, UK

### Abstract

Silicon nanophotonics holds the promise of revolutionizing computing by enabling parallel architectures that combine unprecedented performance and ease of use with affordable power consumption. Here we describe the results of a detailed multiyear design study of dense wavelength division multiplexing (DWDM) on-chip and off-chip interconnects and the device technologies that could improve computing performance by a factor of 20 above industry projections over the next decade.

### 1. Introduction

Moore's Law is still a fundamental technology driver for the information technology industry: the ITRS Semiconductor roadmap [1]

shows, in the next decade and a half, a continued reduction in feature sizes from 40 nm to the sub-10-nm regime. This growth in circuit density has brought us into the multi-core CPU era, and we are on the eve of a many-core (16 or more cores per socket) era. As transistor density increases, the number of transistors comprising a single computer core is not growing; rather, it has pulled back for reasons of power efficiency, allowing the number of cores packaged on the die to grow exponentially. In the many-core era, on-chip and off-chip communication are the critical issues for sustaining performance growth for the demanding, data-intensive applications for which these many-core chips are intended. Computational bandwidth scales linearly with the exponentially growing number of cores, but the rate at which data can be communicated across a chip using top-level metal wires is increasing very slowly. The rate at which data can be communicated through pins at the chip edge is also growing more slowly than the computational bandwidth, and the energy cost of cross-chip and off-chip (from processor socket to DRAM) communication significantly limits the achievable bandwidth. As a result, the field of computer architecture is now in a crisis. It is not clear that many-core processors will find widespread

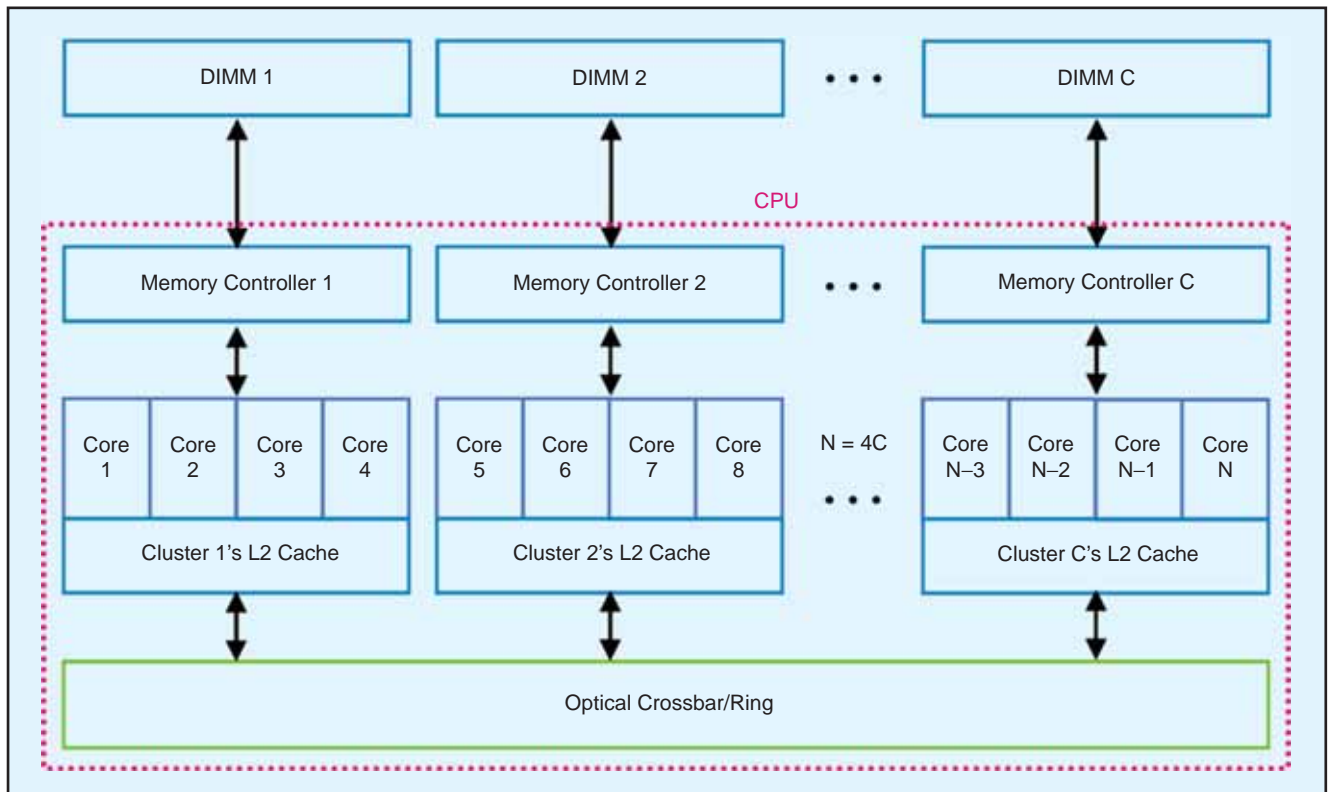


Figure 1: Schematic of the basic architecture for a 256-core processor. The cores are divided into clusters that share an L2 cache and control access to a particular unit of memory.

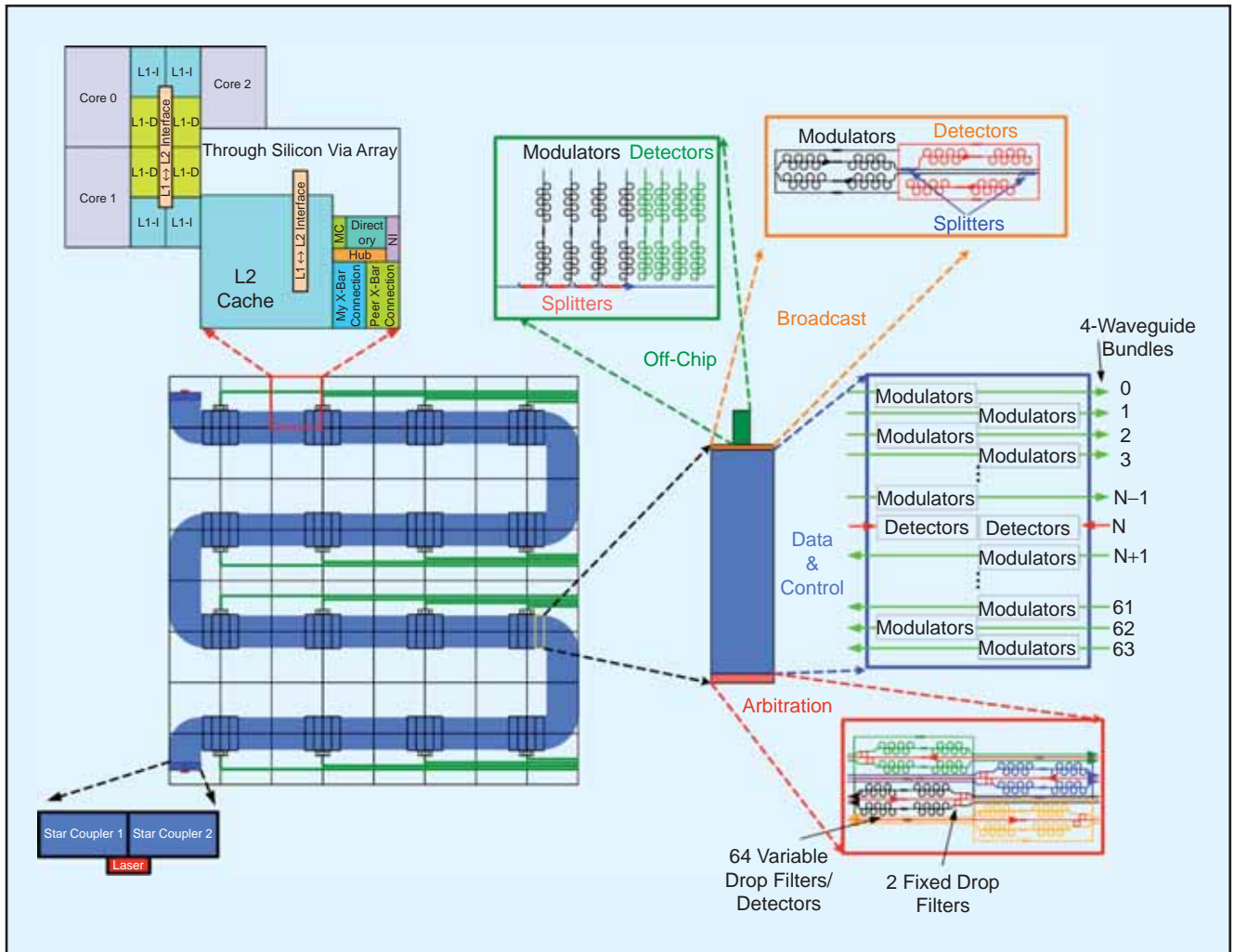


Figure 2: The nanophotonic interconnect die for a 256-core chip operating at 5 GHz. Subsystems illustrated here include: the 64 cluster tiles on the processor and memory dies; 270 identical parallel ridge waveguides allocated to cache line transfers and control messages (blue), and additional waveguides providing off-chip I/O (green); an expanded view of a photonic data/control block, an arbitration block, a broadcast block, and an off-chip communication hub; and laser power distribution.

use: program performance may be consistently disappointing due to limited communication bandwidths. In addition, exposing these limited bandwidths to the programmer makes the parallel programming task far more difficult.

If we strive to continue delivering exponential performance improvements over a broad range of computational applications during the next decade, then we are led inevitably to a symmetric architecture with the simplest possible crossbar interconnect that allows the programmer to exploit parallelism at a high level. But this simplicity requires that we rely on an interconnect technology that is no longer limited by the physics of copper wire. Here we argue that nanophotonics provides a feasible solution to this growing problem [2], and we present a simplified version of a novel photonic interconnect architecture to be presented at ISCA 2008 [3]. As we show in Section 6, the communication bandwidth per unit of dissipated power provided by this on-chip optical interconnect technology exceeds the maximum available from purely electrical interconnects by a factor of 20 for applications that heavily use the interconnect. Other proposals for on-chip nanophotonic interconnects have appeared recently, but most simply replace the long “global” wires with a bus [4] or a

circuit-switched network [5]. These approaches do not efficiently use the advantages of nanophotonics and, as a consequence, have limited bandwidth and large latency.

## 2. Model architecture

In the many-core era, the critical problems facing general purpose computer design are heat generation and communication bandwidth. Architectural compromises forced by these issues make programming difficult and impair application performance. The projected energy cost of cross-chip electrical communication (in the year 2017) is expected to be 5.8 pJ/b; although an extreme Rambus solution for off-chip signaling could achieve (in principle) 2.2 pJ/b, current practice reaches about 20 pJ/b [6]. At these costs, all of the available power will be consumed moving data unless a severe bandwidth choke is imposed. With optics, however, over the next decade we believe that we can reduce the energy required for cross-chip communication to less than 0.5 pJ/band to 0.1 pJ/b for off-chip communication. By exploiting the ability to move data affordably, wherever it is needed, we can reduce both power and total cost while improving bandwidth, performance, and programmability.

In Fig. 1, we show a block diagram of a many-core processor illustrating the basic idea behind the architecture considered here. The device is tiled with 64 identical four-core compute clusters, each of which has a memory controller that either interfaces to stacked memory or drives a photonic connection to off-chip memory to provide bandwidth that scales with processor performance. We interconnect the clusters with a photonic crossbar, offering enormous bandwidth, modest latency, and very low power consumption. This creates a symmetric architecture in which all memory is close to all processors; the programmer expresses parallelism at a high level, and is not burdened by issues of locality, thereby greatly reducing the difficulty of parallel program development. Furthermore, we can provide bandwidth of one byte per double-precision floating-point operation (flop) all the way to DRAM, eliminating the need to exploit very complex software techniques (multilevel tiling, for example) to guarantee locality of reference. The bandwidth provided from DRAM is large enough to eliminate an entire layer of cache (L3), with significant savings in power and cost, while also reducing latency and hardware complexity.

Energy efficiency is the primary motivation for many-core, and small, power-efficient cores and caches achieve the best possible performance per unit energy. We don't anticipate significant increases in CPU clock rates over the next decade, so we assume for our model that the cores use a 5 GHz clock. If we also assume that the cores are dual-issue, in-order, and multithreaded, and that they offer SIMD instructions allowing 4 multiply-accumulate and 4-word-wide load/store operations, then the compute bandwidth of the device is 10 teraflops (i.e., double-precision floating-point operations per second), requiring (at one byte per flop) 20 TB/s of bidirectional interconnect bandwidth. Our power models predict that this processor will dissipate approximately 200 W in the silicon itself.

### 3. The nanophotonic crossbar

Despite the remarkable progress made recently in broadband Mach-Zehnder electrooptic modulators [7] and quantum-well materials for electroabsorption modulators [8,9], we believe that — in the long term — the use of dense wavelength-division multiplexing (DWDM) in optical interconnects is inevitable. In order to minimize the need for buffering (and therefore eliminate the power dissipated by serialization and deserialization, or SERDES), we assume that we will be able to drive each physical channel at twice the clock frequency (i.e., read or write two bits per clock to/from digital registers), or 10 Gb/s. Therefore, we will need 16,000 physical channels in our interconnect to supply the required bidirectional bandwidth of 20 TB/s. However, in order to implement the crossbar with the performance specifications (particularly low latency) required by our architecture, we will need to avoid switching altogether and overprovision the physical chan-

nels significantly. For example, in order to provide an all-to-all single-hop interconnect topology, we will require total direct connections between the clusters. (In other words, at any given time, each cluster may be sending data to only one of the other clusters, but each cluster must always be connected physically to all 63 of the other clusters.) If we assume that we are using dies, and that each waveguide (or wire) occupies an area with dimensions  $2\text{ cm} \times 3\text{ }\mu\text{m}$  on average, we find that the interconnect will require over 200 layers on the die, which is clearly infeasible, regardless of whether we are using wires, direct optical modulation, or optical coarse wavelength division multiplexing (CWDM). As we shall see next, using DWDM in our design allows us to provide extraordinarily high interconnectivity in only a single layer.

As illustrated to scale in Fig. 2, one implementation of the nanophotonic crossbar employs 64 wavelengths multiplexed over 270 parallel waveguides, with 256 waveguides allocated to control and data, 2 to broadcast, and 12 to arbitration. Each cluster "listens" to a single dedicated bundle of 4 waveguides, at all wavelengths. These 256 individual bit-wide channels are grouped into logical channels for data and control messages. The timing of these logical channels is synchronous, each channel being used by one sender to send to its destination in any given 24 clock "epoch." A new, distributed, all-optical crossbar arbitration scheme described in Section 5 allocates the logical channels to sending clusters, allowing one sending cluster to transmit to a given destination cluster in a given epoch. Once granted permission by the arbiter to transmit, the sending cluster modulates all wavelengths on the allocated

# CORNING

## Corning Incorporated

is a diversified technology company with a rich history spanning more than 150 years. Corning combines our expertise in specialty glass, ceramic materials, polymers and the manipulation of properties of light with strong process and manufacturing capabilities to develop, engineer and commercialize significant innovative products for the flat panel display, telecommunications, environmental and life sciences markets.

We are seeking an applied scientist to conduct optical physics research and to provide expertise in fundamental understanding of laser interactions with materials of fundamental importance to Corning, including glasses and ceramics. This position will also have the opportunity to perform exploratory research to extend our understanding of laser processing.

### Responsibilities:

- Lead and/or support R&D teams to develop laser process technologies to enable next generation product or process features
- Direct special investigations of laser system capabilities related to

product or process performance including laser output and laser/material interaction

- Develop independent, exploratory research efforts to characterize and explore new laser/material interactions

### Required Skills:

- Expertise in high power lasers and material interaction (e.g. drilling, ablation, etc.)
- Experimental experience with high power lasers and optical beam delivery systems
- Ability to develop and lead independent research effort to support laser processing initiatives

### Education and Experience:

- Ph.D. in physics, electrical engineering, materials science or related field
- 5+ years experience with multiple types of laser systems (UV, CO<sub>2</sub>, diode and/or Nd:YAG) including beam delivery systems.
- 3+ years experience in field of laser processing of materials

### Submit your application online at:

<http://www.corning.com/careers/index.aspx> and search ref #160941

channel in order to send to the destination cluster. This on-chip, low-cost, high-bandwidth, low-power crossbar is able to handle as many as 64 inputs and outputs simultaneously, and is a revolutionary development that will be significant in many applications; it enables the kind of flattened, symmetric architecture we desire.

The same high bidirectional bandwidth to off-stack memory is a requirement for our target performance. Even with the number of pins growing by 40 percent every generation, electrically-connected memory will in 2017 provide only one fifth of a byte per flop to the processor. To fully exploit the advantages of optics, we reorganize memory into stacks with a photonic interface below the layers of DRAM, connected by fiber to the memory ports of the many-core processor stack. This approach provides adequate bandwidth of about one byte per flop (and eventually even more if necessary), solving the chip-edge bandwidth problem. It does so while saving considerable power (up to 200 W if communicating at 10 TB/s electrically at 2.2 pJ/b). Optical interfaces enable a single DRAM chip to source an entire cache line, making better use of the large internal DRAM memory bandwidth and reducing power even further. This will provide the bandwidth and power benefits of a processor-in-memory (PIM) architecture while keeping the programmability of a classical symmetric multiprocessor.

#### 4. Nanophotonic components

This ultra-high-bandwidth DWDM network is enabled by a number of CMOS-compatible nanophotonic devices that are (at most) only a few years beyond the current state of the art:

1. Low-loss silicon-on-insulator (SOI) waveguides have measured losses as low as 0.2 dB/cm [10], and will not need much improvement. However, the commercially available SOI wafers used for this purpose today are custom-made to satisfy light confinement requirements by increasing the thickness of the traditional oxide buffer layer, and are therefore relatively expensive. In addition, this thick oxide layer confines heat within the thin top silicon layer, which translates into a large temperature build-up for sensitive resonant devices such as drop filters, modulators, and ring lasers. Therefore, in the long term, it will be important to develop means for creating nanophotonic components using pure silicon wafers, which provide high thermal conductivity at low cost.

2. Resonant receiver-less Ge detectors. The smaller bandgap of Ge than Si (0.7 eV vs. 1.1 eV) allows detection of optical signals in the 1300 nm wavelength range, and with properly designed material can likely extend to 1550 nm. The potential ability to fabricate Ge layers on Si offers the possibility of integrating optical components with Si integrated circuits for efficient electric-optical transduction. Ge can be embedded into resonant detectors to allow a single wavelength in a waveguide to be detected. This also enables low-capacitance detectors that eliminate the need for power-hungry amplifiers and clock recovery to build a receiver-less detection scheme [11].
3. Resonant modulators. As described below, we propose to use ring resonators that selectively modulate a single wavelength on a given waveguide and can be moved to an "OFF" state where they are transparent to the data flux in the waveguide. The modulators work by changing the index of silicon rings using charge injection. Published results indicate that the target modulation rate of 10 Gb/s can be achieved with current technology [12]. These rings will need to be kept resonant with the chosen wavelength by thermal tuning of the refractive index.
4. Multiwavelength lasers with precisely controlled frequency intervals are ideal for low-cost DWDM systems. If only one of the frequency channels is servo-locked to an on-chip standard cavity, then all of the other frequency modes will track the controlled mode. One of the possible approaches is the Fabry-Perot comb laser based on quantum dots [13], which has already been used to demonstrate a bit-error-rate of at 10 Gb/s over ten longitudinal modes [14,15]. Another possible approach is the mode-locked hybrid Si/III-V evanescent laser [16], which uses a silicon-waveguide laser cavity wafer-bonded to a III-V gain region. In this case, any ambient temperature change in the environment will cause approximately the same refractive index shift in the laser cavity and the silicon waveguides and resonators that form the DWDM network. Our design study shows that the laser need only provide 1–2 W of total optical power to supply a 20 TB/s network if the detector capacitance is low enough that only 30,000 photons are needed to drive a 1 V swing at the detector's output terminal.

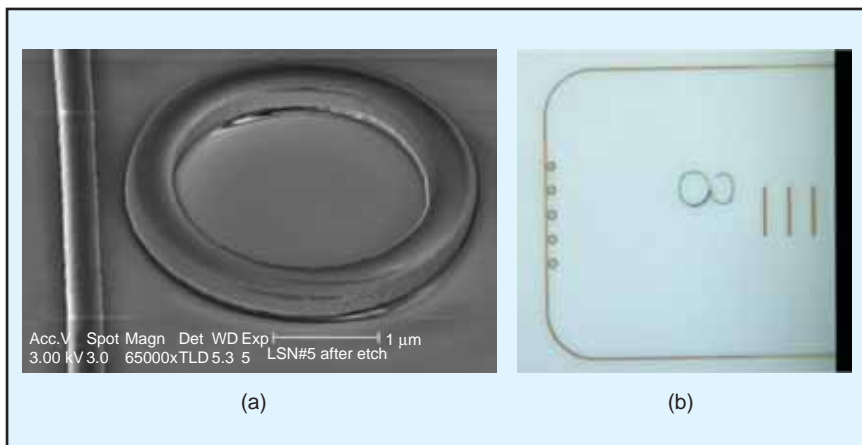


Figure 3: (a) An SEM picture with 40°-tilted view of a microring resonator with a 1.5-μm radius coupled to a waveguide with an optimized (reduced) width. (b) A microscope picture of cascaded microring resonators coupled to a U-shaped waveguide at the edge of the chip.

In our design of a single-layer DWDM nanophotonic interconnect, we have chosen the silicon microring resonator as our foundational component because it has small size, high quality factor Q, transparency to off-resonance light, and no intrinsic reflections. Using injected charge, the refractive index of the microring can be changed, shifting the fundamental frequency of the cavity either into or out of resonance with an incident light field. The microring can act as an optical filter [17,18], and it can be made into electrooptic modulators [12,19,20], lasers [21] and detectors when carrier injection, optical gain, or optical absorption mechanisms are incorporated. In the past, the key characteristic of a silicon microring resonator yet to be

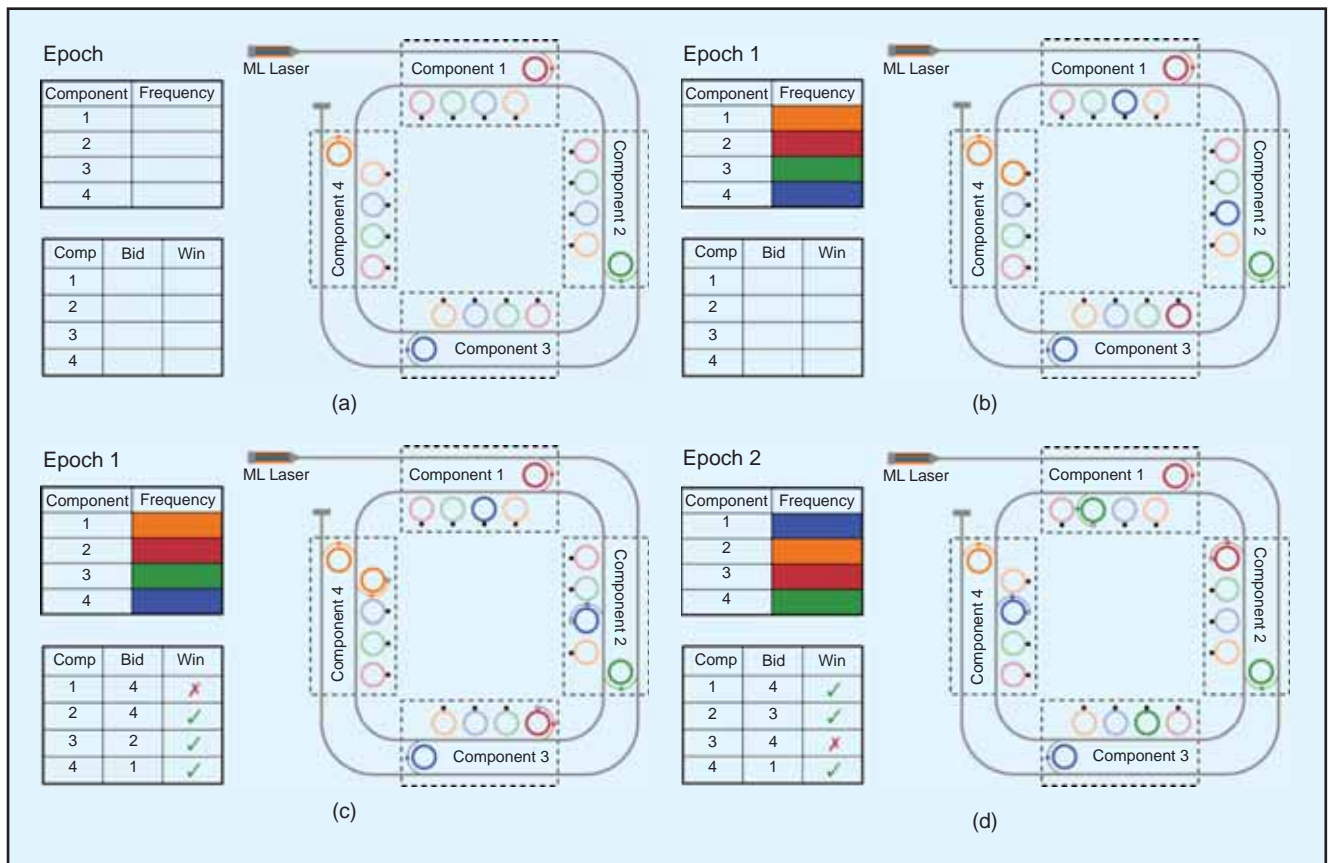


Figure 4: A simplified schematic example of an optical arbitration network used to provision four system resources. This scheme can run in parallel to (and independently from) the network used to transmit data between the cores. For example, each transmitter can determine whether a particular target receiver is available without communicating with the receiver.

demonstrated experimentally is a radius approaching the minimum possible value that allows an intrinsic  $Q$  of 20,000, which is about  $1.4 \mu\text{m}$ . As shown in Fig. 3(a), we have fabricated Si microrings with radii of  $1.5 \mu\text{m}$  [22] with intrinsic  $Q$ s of 18,000 and effective mode volumes around  $1.0 \mu\text{m}$ . When coupled to an optimally-designed silicon strip waveguide that minimizes spurious light scattering and increases the critical dimensions of the geometry (easing fabrication requirements), the coupled  $Q$  approaches the theoretical maximum possible value for a ring of that size (9,000 out of 10,000). In Fig. 3(b), we show cascaded silicon microring resonators that can be used as a modulator or filter bank in a nanophotonic network.

In the case where we use the microring as a modulator, a small size is critical for several reasons. First, a smaller size means that more modulators can be fit into a given area, therefore providing higher integration density. Second and more importantly, the power consumption of the modulator, which is a key performance factor for electrooptical modulators, is directly proportional to the circumference and inversely proportional to the  $Q$  of the resonator. Reducing the size of the ring without sacrificing the  $Q$  is critical for low-power operation. Third, the total bandwidth of a microring-based DWDM modulation system [20] is limited by the free spectral

range (FSR) of the microring resonator, which is inversely proportional to the circumference of the ring. A smaller microring modulator has a larger FSR, which can therefore accommodate more wavelength channels and have higher aggregate data bandwidth. In our case, the choice of a  $1.5 \mu\text{m}$  radius and the demonstration of the near-maximum-possible coupled  $Q$  of 9,000 provides a FSR of about 8 THz, and a filter bandwidth of about 20 GHz, which is nearly ideal for our interconnect architecture.

One of the most important requirements that must be met by a nanophotonic interconnect is that its total thermal dissipation

Technology Node (nm)	40 nm	28 nm	17 nm	14 nm
Clusters/chip	4	16	64	64
Cores/cluster	4	4	4	16
Computational performance (Tflops)	0.64	2.56	10.24	40.96
On-chip interconnect				
Bi-directional bandwidth (TB/s)	1.28	5.12	20.48	81.92
Power (W)	3.4	18.0	38.4	118.4
Energy/bit (fJ)	332	439	235	181
Off-chip interconnect				
Bi-directional bandwidth (TB/s)	1.28	5.12	20.48	81.92
Power (W)	1.8	4.3	8.9	27.5
Energy/bit (fJ)	177	105	54	42

TABLE 1: Projected DWDM interconnect performance at various technology nodes. The power consumption of the interconnect includes lasers and modulators, as well as the power needed to keep resonant detectors and modulators locked.

remain below 25% that of the underlying silicon transistors, or less than 50 W for our target processor in 2017. As discussed above, we expect that the laser source will contribute about 5 W to this total, but the on-chip microrings will contribute a much larger quantity of heat. There are three possible modes for electrical power dissipation in rings: fabrication error trimming, resonance frequency biasing, and direct data modulation. Because of fabrication imperfections, each ring will have a resonance frequency that is slightly different from the design goal, and must be “trimmed” into the correct spectral location. We can rely on two schemes to fine-tune the rings: we can use carrier injection to blue-shift the resonance, and thermal heating to red-shift it. In the worst-case scenario we need 185  $\mu\text{W}/\text{nm}$  to red-shift a 3- $\mu\text{m}$  ring through heating and 125  $\mu\text{W}/\text{nm}$  to blue-shift it through current injection. The approach that combines heating and current injection, however, is only viable if the critical dimension control of the fabrication process is better than 1 nm. If this condition is not met, then thermal control alone needs to be used at the expense of a much larger power consumption.

The electric current required to modulate a ring is given by  $C/\text{pulse}$  at 5 GHz, or 30 fJ at 1 V, assuming that we detune the ring 40 GHz from resonance to obtain an extinction ratio of 10 dB or greater. This corresponds to a raw dissipated power of 150  $\mu\text{W}$  per online ring. However, the modulator voltage driver circuit will necessarily dissipate electrical power, since the modulator acts as a capacitive load with a 10  $\mu\text{A}$  leakage current in the “on” state, and has a peak current of 1 mA during the transition. Recently, this problem has been solved on-die by manufacturers of packaged CMOS photonic devices using AD-DA conversion drivers, but the power dissipated in these drivers has been much greater than the power expended in the modulators themselves. Therefore, in the long term we believe that it will be important to develop purely analog CMOS drivers to reduce the electronic overhead by a factor of 30–100 over the current state-of-the-art. We believe that the efficiency of these drivers will scale only slowly with circuit feature size, resulting in a modulation power below 0.5 mW at the 17-nm technology node.

Generally, in the reconfigurable network we will trim all rings (i.e., both modulators and detectors) away from resonance, and then use current injection to bring the necessary rings online once arbitration is complete. Given the silicon parameters mentioned above, and that an active ring will be online during the entire epoch, the total power requirement per online ring is about 30  $\mu\text{W}$ , or about 0.1 mW including analog driver overhead at the 17-nm technology node. Therefore, in 2017 we expect the power dissipated by all on-chip rings to be approximately 40 W. We have

modeled the performance of the on-chip and off-chip interconnects shown in Fig. 2 at several technology nodes in Table 1. The total power consumption is the sum of the on-chip and off-chip estimates, and includes all of the laser, modulation, and trimming contributions outlined above.

## 5. All-optical arbitration

One of the most significant contributions to the interconnect latency is the time required to determine the availability of system resources, arbitrate collisions between requests, and then grant access to the resource requestors. For example, in an all-to-all multihop switched interconnect architecture (e.g., a torus), electrical signals representing requests must be sent to an arbitration processor, and wait for the outcome of a computation and then receipt of another electrical signal before transmission can commence. However, a key advantage of our solution is that — at the cost of overprovisioning the nanophotonic components — the transmitters themselves can reconfigure the crossbar in a few clock cycles, thus avoiding the need for hand-shaking procedures that increase latency. Nevertheless, we still require an arbitration system to handle collisions. The intrinsic parallelism of optical signals allows us to propose a novel, all-optical, low-latency arbitration system that does not require digital electronic computation or communication between the transmitter and receiver, using a protocol that can be run completely independently of the on-chip data network. Our analysis shows that this protocol provides nearly the best possible throughput under light and moderate loads, and about 90% of the best possible throughput under heavy loads.

A schematic diagram of a simple version of optical arbitration is shown in Fig. 4 for the case where there are four system resources (e.g., L2 caches in cores or clusters of cores) to be allocated. A four-wavelength (e.g., mode-locked) laser provides optical power to each component in a single distribution waveguide, and each wavelength is dropped onto the arbitration waveguide at a specific location in the ring. For example, the “red” wavelength (which in fact may belong to a particular channel near 1310 nm) is always dropped onto the arbitration waveguide by a microring resonator near component 1 that is always tuned to that wavelength. At the beginning of the first 24-clock-cycle “epoch,” each resource is assigned a unique wavelength using a predetermined algorithm known to each component, and each component prepares a “bid” for a particular resource (in this case, the right to transmit data to another component). In an electrical arbitration system, this bid would be an electrical signal sent to a dedicated subprocessor, but here the bid is made locally by tuning an adjacent drop filter to the wavelength assigned to the desired target resource. For example,

here component 1 bids for resource 4 by activating (i.e., tuning into precise resonance) a local microring resonator that is designed to drop the wavelength currently assigned to resource 4 (i.e., “blue” during this epoch) onto an integrated photodetector. If the optical power sensed by the photodetector rises above a designated threshold, then component 1 has won the right to transmit to resource 4. However, in this epoch, component 2 also bids for access to resource 4; since component 2 is “upstream” from component 1, and the

Benchmark	Optical Performance	Electrical Performance	Scaled Optical/Electrical
PTRANS (GB/S)	9102	459.0	22
STREAM (GB/S)	10240	605.0	19
GUPS	40	2.4	19
DGEMM (Gflops)	37	37.0	1
FFT (Gflops)	1734	879.0	2
MPI (GB/s)	20	1.2	19

TABLE 2: HPCS benchmark performance for the proposed architecture at the 17-nm technology node.

“blue” wavelength is dropped onto the arbitration waveguide near component 3, it is component 2 that wins the arbitration round and access to resource 4. Since the optical power at the “blue” wavelength is removed from the waveguide by component 2, the “blue” photodetector sees a low optical intensity, and must wait until the next epoch to try again to transmit data to resource 4. However, during the next epoch, the wavelengths representing the available resources are reassigned, and now component 1 wins the right to transmit to resource 4 even though component 3 tries to bid for the same resource. A more sophisticated token-ring optical arbitration scheme [3] eliminates the need to synchronize execution during epochs, and allows latency to be reduced even further.

## 6. System performance models

We have modeled the performance of the nanophotonic architecture described in the previous section (as well as an idealized electrical equivalent) for the HPC Challenge benchmarks [23], which typify high-performance data access patterns. In our model, we calculated performance limits due to CPU, interconnect, and memory bandwidths. We have assumed that the benchmarks have been implemented as multi-threaded shared-memory programs with data imperfectly placed, requiring communication through the on-chip interconnect. We compare our nanophotonic architecture to a many-core electrically-interconnected alternative system, for which we assume an on-chip mesh network, power-limited to 50 W, and an electrical connection to memory with bandwidth limited by the pin count and pin bandwidth anticipated by ITRS in 2017. Our estimated results for the 17 nm technology node are shown in Table 2. The final column lists the modeled ratio of optical performance to electrical performance per unit of dissipated heat. Note that four of the benchmarks show a factor-of-20 improvement for nanophotonics over wires. The other two benchmarks do not show significant improvements because they are not bandwidth constrained.

## 7. Conclusion

The many-core architecture presented here — with the cores divided into silicon compute clusters, connected to each other and to off-chip memory using nanophotonic technology — will continue to evolve [3] as we further explore the implications of a highly parallel interconnect for the programmer. We believe that the use of DWDM in integrated-circuit interconnects is inevitable, and that the optical components that we describe here are essential elements of that approach. The potentially high bandwidth of an optical interconnect in general-purpose many-core processors will be significantly compromised if an electronically-reconfigurable circuit-switched mesh or torus architecture [5] (essentially a photonic implementation of today’s copper-wire global interconnects) is employed. Instead, we propose to

build an all-optical arbitration system that relies on the same nanophotonic building blocks as the data-transmission network, allowing the transmitters themselves to determine whether a receiver is available, and to begin sending in only a few clock cycles. We have modeled the performance of this system using the HPCC benchmarks, and we have found that a performance increase of 20 over a purely electronic interconnect — essentially 4–5 “extra” Moore’s Law generations — is possible for applications that heavily use the interconnect. This extraordinary performance boost is a critical goal for those of us advocating such a radical departure from current semiconductor engineering practice: the transition to this new interconnect technology will be so painful for the IT industry that only an order-of-magnitude improvement in compute bandwidth will make the risk and effort worthwhile.

## References

- [1] <http://www.itrs.net/>.
- [2] R. G. Beausoleil, P. J. Kuekes, G. S. Snider, S.-Y. Wang, and R. S. Williams, “Nanoelectronic and Nanophotonic Interconnect (Invited Paper),” Proc. IEEE 96, 230–247 (2008).
- [3] D. Vantrease, R. Schreiber, M. Monchiero, M. McLaren, N. P. Jouppi, M. Fiorentino, A. Davis, N. Binkert, R. G. Beausoleil, and J. Ahn, “Corona: System Implications of Emerging Nanophotonic Technology,” in Proceedings of the 35 International Symposium on Computer Architecture (ISCA 2008) (Beijing, China, 2008). To appear.

## Engineered Custom Test Equipment Short Manufacturing Run Products

Fiber Optic • Microwave/RF • Advanced Technologies  
Consulting • Engineering • Manufacturing • Support  
Commercial • Industrial • Defense • Emerging Industry

**Fiber Optic Translators** - Transmitters, receivers, transceivers, regenerators, & wavelength converters. Single mode and multi-mode. Analog and/or digital NRZ, RZ, CRZ, etc. modulation. Direct, EA, or Lithium Niobate modulators. Fixed or tunable OTX and/or ORX. One or more wavelengths and/or channels. WDM, CWDM, DWDM. 850, 1310, 1550 nm, etc. Data rates to ~43 Gb/s. Extended temperature range and/or military ruggedizing. VOA, O or E splitters, power monitors, SBS suppression, microwave breakouts, many other options.

**Clock Regenerators** - Specialized fiber optic and/or microwave receivers with digital clock or clock-data recovery, analog pass-thru available. Binary or multi-level data versions. 850, 1310, 1550 nm, etc. SE/Diff 50-75-100-150 ohms. Data rates to ~13 Gb/s. Used for sampling oscilloscope eye pattern triggers and data-only jitter tests.

**Fiber Optic Spans** - Programmable communication “superhighway in a box”. Real lengths to 1,000+ KM. Used for fiber optic transmitter, receiver, amplifier, etc. dispersion, compensation, and regeneration tolerance tests.

**FEC Translators** - Transmitters, receivers, and transceivers with forward error correction between SONET, SDH, 10GE, G.709, G.975, Super-FEC, etc. up to 13 Gb/s.

**Many Other Choices** - Electronic Translators. Precision Functional Test or Assembly Fixtures. Communication Switch Matrices. Critical Process Equipment. High Voltage Equipment. Specialized Research Equipment. Test Bed Development.

**Example Products** - **Fiber Optic Transceiver**, NRZ to NRZ-RZ-CRZ, 10-13 Gb/s, C+L, SBS enabled laser, manual/LAN operated, 1U x 22” rackmount, 19 operating modes • **Fibre Channel Transceiver**, wide temperature range, military ruggedized, 38999 connectors • **Programmable Fiber Optic Span**, 0 to 118.75 km in 6.25 km steps, with DCF and variable gain DWDM EDFA, manual/LAN operated, 7U x 22” rackmount

Multi-functional units and quantities from 1 to 10+ our specialty.  
Why make it when you can buy it? Download a catalog today!



**Third Millennium Engineering**  
www.tmeplano.com • 972-491-1132

- [4] N. Kirman, M. Kirman, R. K. Dokania, J. Martínez, A. B. Apsel, M. A. Watkins, and D. H. Albonese, "Optical Technology in Future Bus-based Multicore Designs: Opportunities and Challenges," *IEEE Micro* 27, 56–66 (2007).
- [5] K. Bergman and L. Carloni, "Power efficient photonic networks on-chip," in *Proc. Soc. Photo-Opt. Instrum. Eng.*, vol. 6898, p. 689813 (2008).
- [6] R. Ho, *On Chip Wires: Scaling and Efficiency*, Ph.D. thesis (2003).
- [7] W. M. Green, M. J. Rooks, L. Sekaric, and Y. A. Vlasov, "Ultra-compact, low RF power, 10 Gb/s silicon Mach-Zehnder modulator," *Opt. Express* 15, 17106–17113 (2007).
- [8] Y.-H. Kuo, Y.-K. Lee, Y. Ge, S. Ren, J. E. Roth, T. I. Kamins, D. A. B. Miller, and J. S. Harris, "Strong quantum-confined Stark effect in germanium quantum-well structures on silicon," *Nature* 437, 1334–1336 (2005).
- [9] J. E. Roth, O. Fidaner, R. K. Schaevitz, Y.-H. Kuo, T. I. Kamins, J. S. Harris, and D. A. B. Miller, "Optical modulator on silicon employing germanium quantum wells," *Opt. Express* 15, 5851–5859 (2007).
- [10] A. Liu, R. Jones, L. Liao, D. Samara-Rubio, D. Rubin, O. Cohen, R. Nicolaescu, and M. J. Paniccia, "A high-speed silicon optical modulator based on a metal-oxide-semiconductor capacitor," *Nature* 427, 615–618 (2004).
- [11] A. Bhatnagar, C. Debaes, H. Thienpont, and D. A. B. Miller, "Receiverless detection schemes for optical clock distribution," in *Proc. Soc. Photo-Opt. Instrum. Eng.*, vol. 5359, pp. 352–359 (2004).
- [12] Q. Xu, B. Schmidt, S. Pradhan, and M. Lipson, "Micrometre-scale silicon electro-optic modulator," *Nature* 435, 325–327 (2005).
- [13] A. Kovsh, I. Krestnikov, D. Livshits, S. Mikhlin, J. Weimert, and A. Zhukov, "Quantum dot laser with 75nm broad spectrum of emission," *Opt. Lett.* 32, 793–795 (2007).
- [14] A. Gubenko, I. Krestnikov, D. Livshits, S. Mikhlin, A. Kovsh, L. West, C. Bornholdt, N. Grote, and A. Zhukov, "Error-free 10 Gbit/s transmission using individual Fabry-Perot modes of low-noise quantum-dot laser," *Electron. Lett.* 43, 1430–1431 (2007).
- [15] <http://www.innolume.com/>.
- [16] B. R. Koch, A. W. Fang, O. Cohen, and J. E. Bowers, "Mode-locked silicon evanescent lasers," *Opt. Express* 15, 11225–11233 (2007).
- [17] S. Xiao, M. H. Khan, H. Shen, and M. Qi, "A highly compact third-order silicon microring add-drop filter with a very large free spectral range, a flat pass-band and a low delay dispersion," *Opt. Express* 15, 14765–14771 (2007).
- [18] M. S. Nawrocka, T. Liu, X. Wang, and R. R. Panepucci, "Tunable silicon microring resonator with wide free spectral range," *Appl. Phys. Lett.* 89, 071110 (2006).
- [19] Q. Xu, S. Manipatruni, B. Schmidt, J. Shakya, and M. Lipson, "12.5 Gbit/s carrier-injection-based silicon microring silicon modulators," *Opt. Express* 15, 430–436 (2006).
- [20] Q. Xu, B. Schmidt, J. Shakya, and M. Lipson, "Cascaded silicon micro-ring modulators for WDM optical interconnection," *Opt. Express* 14, 9430–9435 (2006).
- [21] A. W. Fang, R. Jones, H. Park, O. Cohen, O. Raday, M. J. Paniccia, and J. E. Bowers, "Integrated AlGaInAs-silicon evanescent race track laser and photodetector," *Opt. Express* 15, 2315–2322 (2007).
- [22] Q. Xu, D. Fattal, and R. G. Beausoleil, "Silicon microring resonators with 1.5- $\mu\text{m}$  radius," *Opt. Express* 16, 4309–4315 (2008).
- [23] [http://icl.cs.utk.edu/hpcc/hpcc\\_results.cgi](http://icl.cs.utk.edu/hpcc/hpcc_results.cgi).

**Visit the LEOS web site for  
more information:  
[www.i-LEOS.org](http://www.i-LEOS.org)**