

The Case for Low-Power Photonic Networks on Chip

Assaf Shacham
Columbia University
Dept. of Electrical Engineering
assaf@ee.columbia.edu

Keren Bergman
Columbia University
Dept. of Electrical Engineering
bergman@ee.columbia.edu

Luca P. Carloni
Columbia University
Dept. of Computer Science
luca@cs.columbia.edu

ABSTRACT

Packet-switched networks on chip (NoC) have been advocated as a natural communication mechanism among the processing cores in future chip multiprocessors (CMP). However, electronic NoCs do not directly address the power budget problem that limits the design of high-performance chips in nanometer technologies. We make the case for a hybrid approach to NoC design that combines a photonic transmission layer with an electronic control layer. A comparative power analysis with a fully-electronic NoC shows that large bandwidths can be exchanged at dramatically lower power consumption.

Categories and Subject Descriptors

C.1.2 [Processor Architectures]: Multiple Data Stream Architectures (Multiprocessors)—*Interconnection architectures.*

General Terms

Design, Performance.

Keywords

Network-on-Chip, Optical Communication.

1. INTRODUCTION

The quest for both high performance and low power has lead to a new emerging trend in high-performance microprocessors design with the arrival of the first commercial chips hosting multiple processing cores like the SUN Niagara, the IBM CELL, and the Intel Duo. It is reasonable to expect that the number of these cores will continue to grow, leading to various generations of chip multiprocessors (CMP). Packet-switched micro-networks based on regular scalable structures such as meshes or tori have been proposed to implement on-chip global communication in multicore processors [1, 2, 11]. These networks-on-chip (NoC) are made of carefully-engineered links and represent a shared medium that can provide enough bandwidth to replace many traditional bus-based and/or point-to-point links. A prototype chip with a NoC connecting 80 cores was recently presented [15]. While NoCs potentially dissipate less power than a set of equivalent point-to-point communication links [1, 2], a growing fraction of the on-chip power dissipation is due to on-chip communications [6, 10, 15].

Leveraging the unique advantages of optical communication to construct *photonic NoCs* offers a potentially disruptive

technology solution that can provide ultra-high throughput, minimal access latencies, and low power dissipation that remains independent of capacity. An optical interconnection network that can capitalize on the capacity, transparency, and fundamentally low power consumption of silicon photonics could deliver *performance-per-watt* that is simply not possible with all-electronic interconnects. Photonic channels can support large amounts of data traffic across longer distances in a bandwidth-oriented design of a network connecting processing cores and memories. Besides the *power wall*, a photonic networks can address also the *memory wall* by allowing seamless delivery of off-chip communication bandwidth with minimal additional power consumption. Electronic technology can complement the photonic network in overcoming some of the limitations inherent to photonics, namely processing and buffering.

The photonics opportunity is made possible now by recent advances in nanoscale silicon photonics. High speed optical modulators at data rates exceeding 12.5Gb/s [18] have been reported. The integration of silicon photonic devices with CMOS integrated circuits for chip-to-chip communication recently became commercially available [4]. These remarkable achievements lead us to envision the integration of a fully functional photonic NoC on a single die. In this paper we present a novel architecture for a photonic NoC that uses silicon photonic technologies to provide a low-power solution for on-chip communication, and thus offer unparalleled advantages in terms of *performance-per-watt*.

2. ARCHITECTURE OVERVIEW

Photonic technology offers unique advantages in terms of energy and bandwidth but lacks two necessary functions for packet switching: buffering and processing are very difficult to implement. Electronic NoCs, conversely, have many advantages in flexibility, abundant functionality and ample buffering space, but their transmission bandwidth per line is limited and their energy requirements are higher.

We propose a photonic NoC architecture that employs a hybrid design, where an optical interconnection network is used for bulk message transmission, and an electronic network is used for distributed control and short message exchanges. Both networks use the same 2-D planar topology that maps well on CMP planar layout. Each core in the CMP is equipped with a network interface, a *gateway*, whose goal is to perform the necessary E/O and O/E conversions, communicate with the control network and execute several other related tasks like synchronization. Every photonic message transmitted is preceded by an electronic control packet (a *path-setup* packet) that is routed on the electronic network, acquiring and setting-up a photonic path for the message. Buffering of messages, which is impossible in the photonic network, only takes place for the electronic pack-

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

DAC 2007, June 4–8, 2007, San Diego, California, USA.

Copyright 2007 ACM 978-1-59593-627-1/07/0006 ...\$5.00.

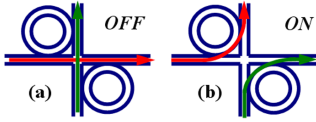


Figure 1: Photonic switching element (PSE): (a) *OFF* state: a passive waveguide crossover. (b) *ON* state: light is coupled into rings and forced to turn.

ets during the path-setup phase. The photonic messages are transmitted without buffering once the path has been acquired. This approach has many similarities with optical circuit switching, a technique used to establish long-lasting connections between nodes in the optical Internet core [12].

The main advantage of using photonic paths relies on a property of the photonic medium known as *bit-rate transparency* [12]: unlike routers based on CMOS technology that must switch with every bit of transmitted data, leading to a dynamic power dissipation that scales quadratically with the bit rate [10], photonic switches switch on and off once per message, and their energy dissipation does not depend on the bit rate. This property facilitates the transmission of very high bandwidth messages while avoiding the power cost that is typically associated with them in traditional electronic networks. Another attractive feature of optical communications results from the *low loss in optical waveguides*: at the chip scale, the power dissipated on a photonic link is completely independent of the transmission distance. Energy dissipation remains essentially the same whether a message travels between two cores that are $2mm$ or $2cm$ apart. Furthermore, the employment of photonic hardware for intrachip communication enables seamless integration of optical interconnects for *off-chip communications*.

The rest of this section summarizes the main issues in the architecture and the design of our hybrid photonic NoC. For a more detailed presentation, the reader is referred to [14].

Building Blocks. The fundamental building block of the photonic network is a broadband *photonic switching element* (PSE), based on a microring-resonator structure. The switch is essentially a waveguide intersection, positioned between two ring resonators (Fig. 1). The rings have a certain resonance frequency, derived from material and structural properties. In the *OFF* state, when the resonant frequency of the rings is different from the wavelength (or wavelengths) on which the optical data stream is modulated, the light passes through the waveguide intersection uninterrupted, as if it is a passive waveguide crossover (Fig. 1a). When the switch is turned *ON*, by the injection of electrical current into *p-n* contacts surrounding the rings, the resonance of the rings shifts such that the transmitted light, now in resonance, is coupled into the rings making a right angle turn (Fig. 1b), thus creating a switching action. PSEs and modulators based on the aforementioned effect have been realized in silicon and a switching time of $30ps$ has been experimentally demonstrated [18]. Their merit lies mainly in their extremely small footprint, approximately $12\mu m$ ring diameter, and their low power consumption: less than $0.5mW$, when *ON* [18]. When the switches are *OFF*, they act as passive devices and consume nearly no power. Ring-resonator based switches exhibit good crosstalk properties ($> 20dB$), and a low insertion loss, approximately $1.5dB$ [17]. These switches are typically narrow-band, but advanced research efforts are now undergoing to fabricate wideband structures capable of

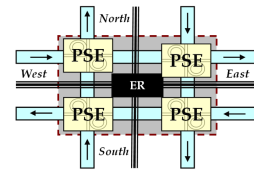


Figure 2: In a 4×4 switch an electronic router controls 4 PSEs.

switching several wavelengths simultaneously, each modulated at tens of Gb/s. It is also reasonable to assume that the loss figures can be improved with advances in fabrication techniques. We use groups of four PSEs controlled by an *electronic router* (ER) to form a 4×4 switch (Fig. 2). The 4×4 switches are interconnected by the inter-PSE waveguides, carrying the photonic data signals, and by metal lines connecting the ERs. Control packets (e.g. path-setup) are received in the electronic router, processed and sent to their next hop, while the PSEs are switched *ON* and *OFF* accordingly. Once a path-setup packet completes its journey through a sequence of electronic routers, a chain of PSEs is ready to route the optical message. Owing to the small footprint of the PSEs and the simplicity of the logic design of the ER, which only handles small control packets, the 4×4 switch can have a very small area. We estimate it at $70\mu m \times 70\mu m$ based on the size of the microring resonator devices [18]. The 4×4 switch is internally blocking, a fact that may affect the NoC performance. Methods to address this issues are discussed in [14].

Topology. We propose a photonic NoC for a chip multiprocessor (CMP) where a number of homogeneous processing cores are integrated as tiles on a single die. The communication requirements of such a system is best served by a 2-D regular topology such as a mesh or a torus [11]. The distributed control scheme of 2-D topologies also contribute to improved scalability - an important property, considering the expectations that the core-count in CMPs will continue to increase in the near future. In this work we assume to serve a 36-core system by a 6×6 2D mesh.

Network contention is a major source of latency in the path-setup procedure. The photonic NoC can be augmented with additional paths so that the probability of contention is lowered and the path-setup latency is reduced. Owing to the small footprint of the switches, the simplicity of the routers, and the fact that the PSEs only consume power when they cause messages to turn, the power and area cost of adding parallel paths is not large. Hence, *path multiplicity* can be used as a cost-effective method of improving performance and reducing contention in the absence of traditional means for contention resolution such as buffers.

Routing and Flow Control. Dimension order routing is a simple routing algorithm for mesh and torus networks, requiring minimal logic in the routers. We use XY dimension order routing on the photonic NoC, with a slight modification required to accommodate the injection/ejection rules of the optical messages [14]. The flow control technique in our photonic NoC greatly differs from common NoC flow control methods due to the fundamental differences between electronic and photonic technologies. In particular memory elements (registers, SRAM...) cannot be used to buffer messages or even to delay them while processing is done. Electronic control packets are thus exchanged to acquire photonic paths, and the data are only transmitted, with

very high bandwidth, once the path has been acquired. The path-acquisition procedure requires the path-setup packet to travel a number of electronic routers and undergo some processing in each hop. Contention may delay the packet, leading to a path-setup latency on the order of tens of nanoseconds. Once a path is acquired, the transmission latency of the optical data is very short, depending only on the group velocity of light in a silicon waveguide: approximately 6.6×10^7 m/s, or $300ps$ for a $2 - cm$ path crossing a chip [7]. The exchange of packets carrying small chunks of data can be done on the electronic control network which is, in essence, a low-bandwidth electronic NoC. These messages are not expected to create congestion because of their small size,

Network Interfaces. Electronic/Optical (E/O) and Optical/Electronic (O/E) conversions are necessary in our photonic NoC. Each node includes a photonic network interface: a *gateway*. Small footprint microring-resonator-based silicon optical modulators with data rates up to $12.5Gb/s$ [18] as well as SiGe photodetectors [5] have been reported recently and become commercially available [4], to be used as photonic chip-to-chip interconnects. The laser sources can be located off chip, externally coupled, as is typically the case in off-chip optical communication systems [4]. The network gateways also include the circuitry necessary for clock synchronization and recovery and serialization/deserialization.

Since electronic signals are fundamentally limited in their bandwidth to a few GHz, larger data capacity is typically provided by increasing the number of parallel wires. The optical equivalent of this wire parallelism can be obtained with a large number of simultaneously modulated wavelengths using wavelength division multiplexing (WDM) at the gateways. The translating device, which can be implemented using microring resonator modulators, converts directly between space-parallel electronics and wavelength-parallel photonics in a manner that conserves chip space as the translator scales to very large data capacities [9]. Optical time division multiplexing (OTDM) can additionally be used to multiplex the modulated data stream at each wavelength and achieve even higher transmission capacity [8]. The energy dissipated in these large parallel structures is not small, but it is still smaller than the energy consumed by the wide busses and buffers currently used in NoCs: the E/O and O/E conversions in the gateway interfaces occur once per node in the photonic NoC, compared to multiple ports at each router in electronic equivalent NoCs [13].

3. HIGH LEVEL POWER ANALYSIS

The main motivation for the design of a photonic NoC is the potential dramatic reduction in the power dissipated on high-bandwidth intrachip communications. To evaluate this power reduction we perform a comparative high-level power analysis between two equivalent on-chip interconnection networks: a photonic NoC and a reference electronic NoC. They are equivalent in the sense that must provide the same bandwidth to the same number of cores. For our case study, we assume a CMP with 36 processing cores, each requiring a peak bandwidth of $800 Gb/s$ and an average bandwidth of $512Gb/s$. These numbers match widely accepted predictions on future on-chip bandwidth requirements in high-performance CMPs. We will see that in this high-bandwidth realm, photonic technologies can offer a dramatic reduction in the interconnect power. We assume a uniform traffic model, a mesh topology and XY dimension

	65 nm	45 nm	32 nm
Clock Frequency [GHz]	3.2	4	5
E_{LINK} [pJ/mm/bit]	0.58	0.46	0.34
E_{BUFFER} [pJ/bit]	0.16	0.13	0.12
$E_{CROSSBAR}$ [pJ/bit]	0.93	0.63	0.36
E_{STATIC} [pJ/bit]	0.06	0.11	0.35

Table 1: Predictions for future technology nodes.

order routing. Of course, different conditions can be used, but as our goal is to provide an equal comparison plane, this choice provides a simple “apples-to-apples” comparison.

Reference Electronic NoC. The reference electronic network is a 6×6 mesh, where each router is integrated in one processor tile and is connected to four tiles. A router micro-architecture that has been widely proposed in the NoC literature [2,11] is based on an input-queued crossbar with a 4-flit buffer on each input port. The router has five I/O ports: one for the local processor and four for the network connections with the neighbour tiles (N, S, E, & W). We estimate the power expended in an electronic NoC under a given load using the method developed by Easley and Peh in [3]: this assumes that whenever a flit traverses a link and the subsequent router, five operations are performed: (1) reading from a buffer; (2) traversing the routers’ internal crossbar; (3) transmission across the inter-router link (4) writing to a buffer in the subsequent router, and (5) triggering an arbitration decision. The energy required for a single hop through a link and a router ($E_{FLIT-HOP}$) is the sum of the energies spent in these operations.

Table 1 reports the values of the energy spent in these operations (buffer reading and writing energies are combined, arbiter energy is neglected) that were obtained with the ORION NoC simulator [16]. ORION account for the static energy dissipated in the router and converts it to a per-bit scale. $E_{FLIT-HOP}$, the energy expended to transmit one flit across a link and a subsequent router, is computed based on the energy estimates in Table 1 as well as the link length and flit-width which vary for different technology nodes. The total energy expended in a clock cycle can be computed as

$$E_{NETWORK-CYCLE} = \sum_{j=1}^{N_L} U_{L_j} \cdot E_{FLIT-HOP}$$

where U_{L_j} is the average number of flits traversing link j per clock cycle, an estimate on the utilization of link j . Then, the power dissipated in the network is equal to

$$P_N = E_{NETWORK-CYCLE} \cdot f$$

where f is the clock frequency. For the a 6×6 mesh under uniform traffic using XY routing and an injection rate of $\alpha = 0.625$ the global average link utilization is $\bar{U} = 0.75$. Hence, the energy expended in a clock cycle in the reference electronic NoC (which has 120 links) is:

$$E_{NETWORK-CYCLE} = 0.75 \cdot 120 \cdot E_{FLIT-HOP}$$

and the total power dissipated is estimated as:

$$P_{E-NOC} = E_{NETWORK-CYCLE} \cdot f$$

The results appear in Table 2. The main conclusion that can be drawn from this analysis is that when a truly high communication bandwidth is required for on-chip data exchange, even a dedicated, carefully designed NoC may not be able to provide it within reasonable power constraints. Since the electronic transmission is limited in bandwidth to a few GHz at most, high transmission capacity require the use of

	65 nm	45 nm	32 nm
Flit width	256	208	168
Link length [mm]	3.33	2.33	1.67
$E_{FLIT-HOP}$ [pJ]	788	406	235
P_{E-NoC} [W]	227	146	106

Table 2: Power consumption of electronic NoC.

many parallel lines [2], which lead to high power dissipation for transmission and buffering. Admittedly the above analysis is based on a simple circuit implementation, but even if aggressive electronic circuit techniques such as low-swing current mode signaling are employed, the overall NoC power consumption that is necessary to meet the communication bandwidth requirements in future CMPs will likely be too high to manage within reasonable packaging constraints.

Photonic NoC. Since our photonic NoC is based on an hybrid design (Sec 2), its power dissipation can be estimated as the sum of three components: the photonic network, the electronic control network, and the O/E E/O interfaces.

1. *Transmission Network.* Path multiplicity is a low-power cost-effective solution to compensate for the lack of buffers in the photonic network. In this design we assume a path multiplicity factor of 2, meaning a 12×12 photonic mesh, comprised of 576 PSEs (144 4×4 switches), serves the 6×6 CMP. The power analysis of a photonic NoC is fundamentally different from the electronic network analysis since it mainly depends on the state of the PSEs: in the *ON* state, when the message is forced to turn, the power dissipated is less than $0.5mW$ [18], while in the *OFF* state, when a message proceeds undisturbed or when no message is forwarded, there is no dissipation. Hence, the total power consumption in the network depends on the number of switches in *ON* state, which can be estimated based on network statistics and traffic dynamics. We assume that in the photonic NoC each message makes, at most, 4 turns. Assuming a peak bandwidth of $960Gb/s$ and an injection rate of 0.6, the average bandwidth is $576Gb/s$. The average number of messages in the network at any given time is calculated as $36 \times 0.6 = 21.6$. The average number of PSEs in the *ON* state is about 86 in a 576-PSE NoC. Hence, the total power consumption can be estimated as:

$$P_{P-NoC,transmission} = 86 \cdot 0.5mW = 43mW$$

dramatically lower than anything that can be approached by an electronic NoC.

2. *Control Network.* The power analysis of the electronic control network is based on the fact that this is essentially an electronic NoC, i.e, similar to our reference electronic NoC except for the larger dimensions (12×12 compared to 6×6). We assume that each photonic message is accompanied by two 32-bit control packets and the typical size of a message is 2 KBytes. Then, the total power consumed by the electronic control network can be approximated as:

$$P_{P-NoC,control} = P_{E-NoC} \cdot 2 \cdot \frac{32}{16384} \cdot 2 = 0.82W$$

3. *Network Interfaces.* To generate the $960Gb/s$ peak bandwidth we assume a modulation rate of $10Gb/s$. The modulated data streams are then grouped using $\times 12$ OTDM to $\times 8$ WDM to form $960Gb/s$ messages. The OTDM and WDM multiplexers are passive elements, so power is dissipated mainly in the 96 modulators and 96 receiver circuits in each gateway. Since there is presently no equivalent to the ITRS for the photonic technology, predictions on the power consumption of photonic elements vary greatly. A

reasonable estimate for the energy dissipated by a modulator/detector pair, at $10Gb/s$ is $1.1pJ/bit$ today. We estimate that using silicon ring-resonator modulators and SiGe detectors, the energy will decrease to about $0.11pJ/bit$ in 8-10 years. Consequently, the total power dissipated by 36 interfaces under the conditions described above is:

$$P_{P-NoC,gateways} = 0.11pJ/bit \times 36 \times 576Gb/s = 2.3W$$

Hence, the estimated power consumed by the photonic NoC to exchange data between 36 cores at an average bandwidth of $576Gb/s$ is the sum of the three components: **3.2W**.

Concluding Remarks. Although the power analysis used here is rather simplistic and uses many assumptions to ease the calculation and work around missing data, its broader conclusion is unmistakable. The potential power difference between photonics-based NoCs and their electronic counterparts is immense. Even when one accounts for inaccuracies in our analysis and considers predicted future trends the advantages offered by photonics represent a clear leap in terms of *bandwidth-per-watt* performance.

Acknowledgments

AS and KB acknowledge the support of the NSF under Grant CCF-0523771 and the U.S. Dept. of Defense under subcontract B-12-664. LC acknowledges the support of the NSF under Grant No. 0541278.

4. REFERENCES

- [1] L. Benini and G. D. Micheli. Networks on chip: A new SoC paradigm. *IEEE Computer*, 49(2/3):70–71, Jan. 2002.
- [2] W. J. Dally and B. Towles. Route packets, not wires: On-chip interconnection networks. In *Design Automation Conf.*, pages 684–689, June 2001.
- [3] N. Easley and L.-S. Peh. High-level power analysis for on-chip networks. In *Intl Conf. on Compilers, Architecture, and Synthesis for Embedded Systems*, Sept. 2004.
- [4] C. Gunn. CMOS photonics for high-speed interconnects. *IEEE Micro*, 26(2):58–66, Mar./Apr. 2006.
- [5] A. Gupta et al. High-speed optoelectronics receivers in SiGe. In *17th Intl. Conf. on VLSI Design*, pages 957–960, Jan. 2004.
- [6] J. L. Hennessy and D. A. Patterson. *Computer Architecture: A Quantitative Approach*. Morgan Kaufmann, 2006.
- [7] I.-W. Hsieh et al. Ultrafast-pulse self-phase modulation and third-order dispersion in si photonic wire-waveguides. *Optics Express*, 14(25):12380–12387, Dec. 2006.
- [8] S. Kawanishi et al. 3 Tbit/s (160 Gbit/s $\times 19$ channel) optical TDM and WDM transmission experiment. *Electronic Letters*, 35(10):826–827, May 1999.
- [9] B. G. Lee et al. Demonstrated 4×4 Gbps silicon photonic integrated parallel electronic to WDM interface. In *Optical Fiber Communications Conf. (OFC)*, Mar. 2007.
- [10] T. Mudge. Power: A first-class architectural design constraint. *IEEE Computer*, 34(4):52–58, 2001.
- [11] T. M. Pinkston and J. Shin. Trends toward on-chip networked microsystems. *Intl. J. High Performance Computing and Networking*, 3(1):3–18, 2001.
- [12] R. Ramaswami and K. N. Sivarajan. *Optical Networks: A Practical Perspective*. Morgan Kaufmann, 2002.
- [13] A. Shacham, K. Bergman, and L. P. Carloni. Maximizing GFLOPS-per-Watt: High-bandwidth, low power photonic on-chip networks. In *IBM P=ac² Conf.*, Oct. 2006.
- [14] A. Shacham, K. Bergman, and L. P. Carloni. On the design of a photonic network on chip. In *The 1st IEEE Intl Symp. on Networks-on-Chips (NOCS)*, May 2007.
- [15] S. Vangal et al. An 80-tile 1.28 TFLOPS network-on-chip in 65 nm CMOS. In *Intl. Solid State Circuits Conf.*, Feb. 2007.
- [16] H.-S. Wang et al. Orion: A power-performance simulator for interconnection networks. In *IEEE/ACM Intl. Symp. on Microarchitecture (MICRO-35)*, Nov. 2002.
- [17] F. Xia, L. Sekaric, and Y. A. Vlasov. Ultracompact optical buffers on a silicon chip. *Nature Photonics*, 1:65–71, Jan. 2007.
- [18] Q. Xu et al. 12.5 Gbit/s carrier-injection-based silicon microring silicon modulators. *Optics Express*, 15(2):430–436, 22 Jan. 2007.