Tools and Methodologies for Designing Energy-Efficient Photonic Networks-on-Chip for High-Performance Chip Multiprocessors

Johnnie Chan, Gilbert Hendry, Aleksandr Biberman, and Keren Bergman
Department of Electrical Engineering
Columbia University
New York, NY 10027, USA
{johnnie, gilbert, biberman, bergman}@ee.columbia.edu

***Abstract*—Photonic interconnection networks have recently been proposed as a replacement to conventional electronic network-on-chip solutions in delivering the ever increasing communication requirements of future chip multiprocessors. While photonics offers superior bandwidth density, lower latencies, and improvements in energy efficiency over electronics, the photonic network designs that can leverage these benefits cannot be easily derived by simply mimicking electronic layouts. In fact, proper implementations of photonic interconnects will require the careful consideration of a variety new physical-layer metrics and design requirements that did not exist with electronics. Here, we review some of the currently proposed designs for chip-scale photonic interconnection networks, the design methodologies required to produce viable network topologies, and a simulation environment, called PhoenixSim, that we have developed to accurately model and study those metrics and designs.**

I. Introduction

Photonic interconnection networks have recently been proposed as a possible solution in alleviating many of the problems facing networks-on-chip (NoCs) for chip multiprocessors (CMPs). The core-to-core and core-to-memory communication requirements of modern CMPs are quickly outpacing the ability of electronic NoCs in supplying the necessary system bandwidth, motivating the need for innovative solutions such as photonics. In some high performance CMPs, as much as 50% of the total dynamic power is dissipated by the electronic interconnects [1]. The power dissipated by the interconnect is expected to continue to escalate with time, further necessitating the transition to a new chip-scale networking platform.

Photonics has the potential to provide better bandwidth, lower latency, and improved energy efficiency over electronics, however the two technology domains are fundamentally different in both operation and physical-layer characteristics. First, the photonic domain has the ability to use wavelength division multiplexing (WDM) which employs multiple optical signals tuned to different wavelengths on a single waveguide to create extremely high-bandwidth communication circuits. Secondly, the optical domain provides no practical means of enabling in-flight logical processing or packet buffering without utilizing an optical-electronic-optical (O-E-O) conversion stage. It is crucial to avoid such O-E-O conversion since it would significantly increase the power dissipated by the photonic network and diminish the suitability of the photonic topology. Additionally, optical amplification for signal regeneration is not easily feasible on a silicon substrate thereby limiting the distance an optical signal can travel. However, this lack of buffering, domain conversion, and regeneration results in a link energy dissipation that is practically distance independent, enabling greater energy efficiencies than electronics for chip- and board-scale communications. Lastly, photonic networks exhibit insertion loss and crosstalk that have no direct counterpart with electronics. For all these reasons, new methodologies and tools must be realized to accurately design these networks.

This paper presents an overview of how photonic interconnection networks are designed and can be used to create high-bandwidth and energy-efficient links for on- and off-chip communications. We also briefly describe PhoenixSim, a physical-layer photonic interconnection network simulator, which we have developed to accurately model photonic NoC designs and measure their performance characteristics [2]. Lastly, we discuss some metrics that need to be considered in properly designing photonic NoCs.

II. Photonic Interconnection Networks

We envision the integration of a dedicated photonic networking plane to provide low-power global communications through the three-dimensional integration (3DI) of future CMPs [3]. As shown in Fig. 1, 3DI enables the stacking of multiple layers that are each dedicated for a different purpose. The bottom plane would contain the many processing nodes of the CMP, while the next several layers would be dedicated to providing a large amount of local memory. Found at the top would be the photonic plane that globally connects the many cores and memory units.

Due to recent advancements in optical technologies, photonic NoCs have become attractive solutions for the problems facing traditional electronic implementations. All the

This work was supported in part by the Interconnect Focus Center, one of five research centers funded under the Focus Center Research Program, a Semiconductor Research Corporation and DARPA program.

978-1-4244-5309-2/10/$26.00 ©2010 IEEE

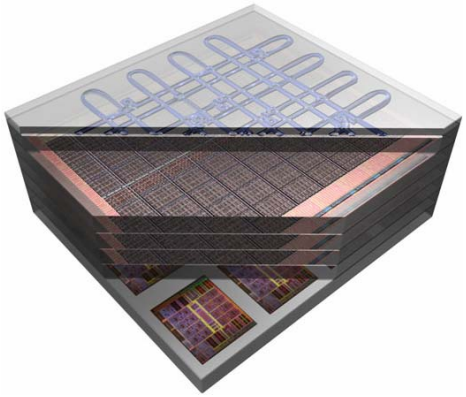

Figure 1. Example illustration of a 3DI chip, with planes dedicated to processing (bottom), memory (middle), and communications (top).

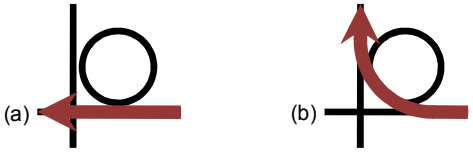

Figure 2. The ring resonator can be used to control the propagation of an optical signal by being configured to be in an (a) off-resonance or (b) on-resonance state.

optical devices that would be required to generate, route, and receive an optical signal have been demonstrated in silicon which renders them compatible with traditional complementary-metal-oxide-semiconductor (CMOS) fabrication technology. These devices includes waveguides [4], waveguide crossings [5], modulators [6], detectors [7], filters [8], and switches [9][10]. These devices can be combined to form a diverse range of topologies, giving chip designers a large design space to explore the possibilities.

A. *Methods for Routing Optical Signals*

The silicon ring resonator is a fundamental device in enabling the control of signal propagation on a photonic network. Ring resonators can be used off-resonance which allows the signals to pass by unobstructed (Fig. 2a), or in an on-resonance configuration which shifts the signal onto another waveguide (Fig. 2b). These two states of the ring resonator form the basis for all types of chip-scale photonic switching.

One method of controlling the flow of an optical signal is to treat the ring as a selective filter and controlling the wavelength of the incoming signal to be either on- or off-resonance (Fig. 3a). If a signal's wavelength is aligned with a resonant peak (i.e. resonant mode) of the ring, the signal will be on-resonance and drop into the ring, otherwise the signal will pass by. We call this *wavelength-selective switching* since the wavelength of the signal is used to determine the route that a signal will take. A network employing this type of switching has previously been proposed for use as an off-chip memory-access network architecture [11].

Alternatively, control can be enabled by holding the wavelength of the signal constant and actively tuning the ring's resonant profile through electro-optic means. This type of active control would result in a concurrent shift of the entire

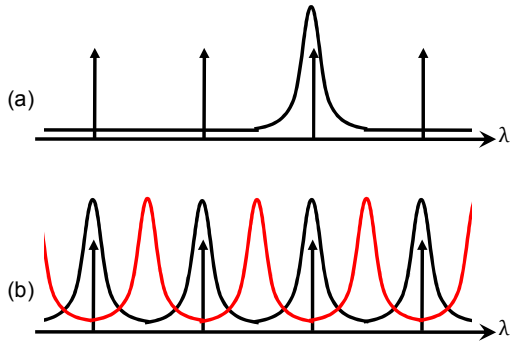

Figure 3. (a) Resonant profiles of ring-resonator designed as a filter. (b) The profile of a ring-resonator broadband switches which can be configured to be on-resonance (black) or off-resonance (red). Arrows represent the wavelengths of incoming optical signals.

resonant profile. Fig. 3b shows how shifting the profile from the black to red profile would cause all wavelengths to transition from on-resonance to off-resonance. We refer to ring resonators that use this method as broadband switches since the many wavelength channels of a WDM signal can each be aligned to a unique resonant mode and any shift in the resonant profile will simultaneously affect all wavelength channels. We label this form of control as *space switching* for its nature of shifting the entire WDM packet. A variety of on-chip spaced-switch networks have previously been proposed [12]–[13].

When comparing the two switching methods, spacing switching fully utilizes the optical spectrum to maximize link bandwidth by leveraging WDM, whereas wavelength-selective switching sacrifices some of the spectrum in order to use it as a controlling mechanism. Conversely, wavelength-selective switching can achieve much better latencies since it does not incur the delay overhead that is required by the circuit-switching protocol (discussed next) of the space-switching technique. Note that the two switching techniques described here can be considered distinctive techniques, but combinations of the two can also be used.

B. *Routing Algorithm for Space-Switched Networks*

Correctly routing optical signals through a space switched network requires electronic control of the electro-optic ring resonators [12]. This task of manipulating and provisioning optical resources is delegated to a control plane. The control plane is designed as a separate low-power low-bandwidth electronic network, with a layout that mirrors the photonic plane. This mirroring is done to ensure that control messages can mimic the propagation of an optical signal and alter the state of ring resonators as needed.

The lifetime of an optical path begins with the arrival of a transmission request at a core, and ends after the transmission of data has been accomplished. When a request for network resources occurs at a core, a *path-setup* packet will be transmitted into the network, duplicating the route the photonic message would take. This is done in order to reserve the necessary broadband ring switches to form a valid optical path between the source and destination cores. Note that all reserved rings remain in their rest state (no applied bias) until

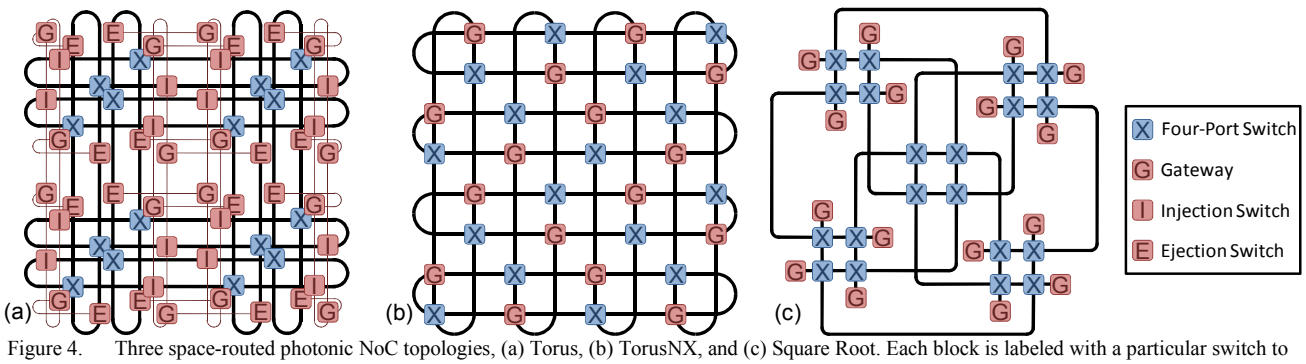

Figure 4. Three space-routed photonic NoC topologies, (a) Torus, (b) TorusNX, and (c) Square Root. Each block is labeled with a particular switch to represents different ring resonator and waveguide arrangements that are used to support differing functionalities.

after the full path has successfully been provisioned to avoid unnecessary energy dissipation from prematurely controlling the rings. If a needed ring along the way is already reserved or being used, the path is considered blocked and a *path-blocked* packet is returned, following the same route back so that the rings that have been reserved can be released. If the path-setup packet successfully provisions all the needed rings and reaches the destination, a *path-ack* packet is returned to the source along the same path, while signaling each reserved ring along the way to enter the on-resonance or off-resonance state as necessary. As soon as the path-ack returns to the source, the full path has been provisioned and the two optically connected cores can now utilize the high-speed communication channel. Once the last bit of data leave the source on the photonic plane, a *path-breakdown* packet is transmitted along the control plane to de-allocate all reservations and return each ring to their original state.

C. Photonic Network Topologies

Three different space-routed networks that we have designed are illustrated in Fig. 4a, 4b, and 4c in 4×4 configurations (X×Y expresses the number of nodes in the X and Y dimension). The blocks labeled 'G' denote gateways, locations on each node where an attached processing core can initiate or receive data transmissions. The thicker lines represent two waveguides used for bi-directional data transmission and the blocks marked 'X' represent four-port non-blocking photonic switches which are composed of eight ring-resonators and are used to effectively route data through the network [14]. Together, they form the main network through which data is routed.

The Torus (Fig. 4a) topology was the first proposed space-routed network [12]. It requires an additional access network, represented by thinner lines (additional waveguides) and the blocks denoted by 'I' (injection) and 'E' (ejection) to facilitate entering and exiting the main network. TorusNX (Fig. 4b) improves the Torus topology by introducing new a switch design that eliminates the need for the access network and directly integrates the gateway into the main topology [13]. Square Root (Fig. 4c) is an alternative hierarchical topology optimized to reduce the required number of waveguide crossings and switching points [13]. Due to the recursive nature of constructing the Square Root, the number of nodes along the X and Y dimension of the topology must be equal and a positive integer power of two (i.e. 2, 4, 8, 16, …). TorusNX and Square Root were both designed in response to preliminary physical-layer shortcomings of the Torus, since insertion losses due to waveguide crossings and the large number of switches had a dramatic impact on system performance.

III. Photonic Metrics and Simulation Methodology

Design metrics are useful in determining the usefulness and feasibility of a NoC design. Electronics and photonics share many system-level metrics such as throughput and latency as well as some physical-layer metrics such as power efficiency. However, in order to fully understand the capabilities and limitations of a photonic NoC, a designer must also consider additional metrics unique to photonics, namely insertion loss and crosstalk. In fact, some of these new metrics directly impact others metrics, clearly making them a critical consideration for creating a high-performance energy-efficient design [13].

Since photonics is physically different from electronics, it renders typical network simulators incapable of clearly determining accurate performance characteristics of such networks. Also, current fabrication technology is incapable of producing entire photonic network topologies which further stresses the need for physically-accurate network simulation. The PhoenixSim simulation environment was developed with the goal of incorporating detail models of photonic components while also maintaining a level of customizability to allow a variety of topology designs to be explored. The toolset is utilized in a hierarchical manner, enabling the creation of a library of highly-parameterizable device models to be used as building blocks to form larger networking components as well as entire topologies.

A. Insertion Loss vs. Throughput and Scalability

Due to the lack of signal regeneration in silicon and the avoidance of O-E-O conversion, care must be taken in determining the amount of insertion loss that the photonic network should sustain. Insertion loss is a measure of the optical power attenuation that a photonic signal incurs as it travels through a photonic network. Major sources of loss come from waveguides, waveguide crossings, and ring resonators. Incurring too high of an insertion loss will result in a signal with insufficient optical power to be properly received. Furthermore, increasing the number of nodes connected to the network would introduce additional

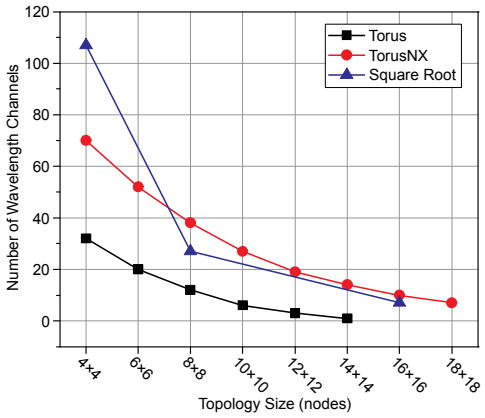

Figure 5. Maximum possible number of wavelengths for each topology.

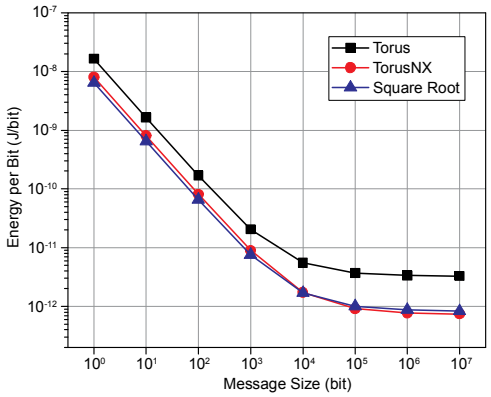

Figure 6. Transmission efficiency of the photonic topologies.

waveguides and photonic devices which will invariably increase the networks insertion loss. This shows that the maximum insertion loss of a network directly affects how scalable and complex the network can be.

The parallelism that can be achieved with using WDM is also bounded by insertion loss. The wavelength channels in a WDM message effectively behave like many distinct messages which cause the total laser power to be equally divided up among them. This is done due to the existence of physical limitations in the photonic devices that are evaluated based on aggregate power and not the power of the individual channel. This actually reduces the amount of insertion that each wavelength channel can work with and further reduces how complex the network can become. On the other hand, it may be desirable to limit the topology size and support a large number of wavelength channels for its throughput benefits.

Insertion loss has a clear relationship to how well the network can scale and the throughput that can be achieved. This creates a parameter space that requires consideration based on the needs of the system that the NoC will be integrated into. This type of analysis is incorporated into PhoenixSim, with an example output plotted in Fig. 5.

B. *Power Dissipation*

The network-level power dissipation is a major component in limiting performance scaling of chip-scale systems. Photonic on-chip networks have been shown to drastically outperform electronic networks in both performance and energy, especially in the case of traffic patterns that require large data transmissions [15]. Power and energy analysis has also been integrated into PhoenixSim (Fig. 6).

IV. CONCLUSIONS

Photonic technology offers a unique solution to mitigating many performance limitations facing today's electronic NoCs. Because of the fundamentally different nature of photonics, new strategies for designing photonic topologies must be devised to ensure that such a design can fully leverage what photonics has to offer. Novel tools that understand the unique characteristics of photonics, like PhoenixSim, are needed so that researchers can accurately model, analyze, and create optimized designs for them.

REFERENCES

[1] N. Magen, A. Kolodny, U. Weiser, and N. Shamir, "Interconnect-power dissipation in a microprocessor," in *Proc. of the 2004 International Workshop on System Level Interconnect Prediction*, pp. 7–13, 2004.

[2] J. Chan, G. Hendry, A. Biberman, L. P. Carloni, K. Bergman, "PhoenixSim: a simulator for physical-layer analysis of chip-scale photonic interconnection networks," submitted to *Design, Automation and Test in Europe*, 2010.

[3] A. W. Topol, et al., "Three-dimensional integrated circuits," *IBM Journal of Research and Development*, vol. 50, no. 4/5, pp. 491–506, Jul/Sep 2006.

[4] F. Xia, L. Sekaric, and Y. Vlasov, "Ultracompact optical buffers on a silicon chip," *Nature Photonics*, vol. 1, pp. 65–71, 2006.

[5] W. Bogaerts, P. Dumon, D. V. Thourhout, and R. Baets, "Low-loss, low-cross-talk crossings for silicon-on-insulator nanophotonic waveguides," *Opt. Lett.*, vol. 32, no. 19, pp. 2801–2803, 2007.

[6] M. R. Watts, D. C. Trotter, R. W. Young, and A. L. Lentine, "Ultralow power silicon microdisk modulators and switches," in *Proc. IEEE International Conference on Group IV Photonics*, Sorrento, Italy, paper WA2, Sept. 2008.

[7] S. J. Koester, et al., "Ge-on-SOI-detector/Si-CMOS-amplifier receivers for high-performance optical-communication applications," *IEEE/OSA Journal Lightwave Technology*, vol. 25, no. 1, pp. 46–57, Jan. 2007.

[8] B. E. Little, et al., "Ultra-compact Si-SiO_2 microring resonator optical channel dropping filters," *IEEE Photon. Technol. Lett.*, vol. 10, no. 4, Apr. 1998.

[9] B. G. Lee, A. Biberman, P. Dong, M. Lipson, and K. Bergman, "All-optical comb switch for multiwavelength message routing in silicon photonic networks," *IEEE Photon. Technol. Lett.*, vol. 20, pp. 767–769, 2008.

[10] B. G. Lee, et al., "High-speed 2×2 switch for multi-wavelength silicon photonic networks on-chip," *IEEE/OSA Journal Lightwave Technology*, vol. 27, no. 14, pp. 2900–2907, Jul. 2009.

[11] C. Batten, et al., "Building many-core processor-to-DRAM networks with monolithic CMOS silicon photonics," *IEEE Micro*, vol. 29, no. 4, pp. 8–21, Jul.–Aug. 2009.

[12] A. Shacham, K. Bergman, and L. P. Carloni, "Photonic networks-on-chip for future generations of chip multiprocessors," *IEEE Transactions on Computers*, vol. 57, no. 9, pp. 1246–1260, Sept. 2008.

[13] J. Chan, G. Hendry, A. Biberman, K. Bergman, "Architectural Design Exploration of Chip-Scale Photonic Interconnection Networks Using Physical-Layer Analysis," submitted to *Optical Fiber Communication Conference*, 2010.

[14] B. G. Lee, A. Biberman, K. Bergman, N. Sherwood-Droz, M. Lipson, "Multi-wavelength message routing in a non-blocking four-port bidirectional switch fabric for silicon photonic networks-on-chip," in *Proc. Optical Fiber Communications Conference*, OMJ4, Mar. 2009.

[15] G. Hendry, et al., "Analysis of photonic networks for a chip multi-processor using scientific applications," in *International Symposium on Networks-on-Chip*, May 2009.