LumiNOC: A Power-Efficient, High-Performance, Photonic Network-on-Chip

Cheng Li, *Student Member, IEEE,* Mark Browning, *Student Member, IEEE,* Paul V. Gratz, *Member, IEEE,* and Samuel Palermo, *Member, IEEE*

***Abstract*—To meet energy-efficient performance demands, the computing industry has moved to parallel computer architectures, such as chip multiprocessors (CMPs), internally interconnected via networks-on-chip (NoC) to meet growing communication needs. Achieving scaling performance as core counts increase to the hundreds in future CMPs, however, will require high performance, yet energy-efficient interconnects. Silicon nanophotonics is a promising replacement for electronic on-chip interconnect due to its high bandwidth and low latency, however, prior techniques have required high static power for the laser and ring thermal tuning. We propose a novel nano-photonic NoC (PNoC) architecture, LumiNOC, optimized for high performance and power-efficiency. This paper makes three primary contributions: a novel, nanophotonic architecture which partitions the network into subnets for better efficiency; a purely photonic, in-band, distributed arbitration scheme; and a channel sharing arrangement utilizing the same waveguides and wavelengths for arbitration as data transmission. In a 64-node NoC under synthetic traffic, LumiNOC enjoys 50% lower latency at low loads and ∼40% higher throughput per Watt on synthetic traffic, versus other reported PNoCs. LumiNOC reduces latencies ∼40% versus an electrical 2-D mesh NoCs on the PARSEC shared-memory, multithreaded benchmark suite.**

***Index Terms*—Low-power electronics, multiprocessor interconnection networks, nanophotonics, optical interconnects, ring resonator**

I. Introduction

PARALLEL architectures, such as single-chip multiprocessors (CMPs), have emerged to address power consumption and performance scaling issues in current and future VLSI process technology. Networks-on-chip (NoCs), have concurrently emerged to serve as a scalable alternative to traditional, bus-based interconnection between processor cores. Conventional NoCs in CMPs use wide, point-to-point electrical links to relay cache-lines between private mid-level and shared last-level processor caches [1]. Electrical on-chip interconnect, however, is severely limited by power, bandwidth, and latency constraints. These constraints are placing practical limits on the viability of future CMP scaling. For example, communication latency in a typical NoC connected multiprocessor system increases rapidly as the number of nodes increases [2]. Furthermore, power in electrical interconnects has been reported as high 12.1 W for a 48-core, 2-D mesh CMP at 2 GHz [1], a significant fraction of the system's power budget. Monolithic silicon photonics have been proposed as a scalable alternative to meet future many-core systems bandwidth demands, however, many current photonic NoC (PNoC) architectures suffer from high power demands and high latency, making them less attractive for many uses than their electrical counterparts. In this paper, we present a novel PNoC architecture which significantly reduces latencies and power consumption versus competing photonic and electrical NoC designs.

Recently, several NoC architectures leveraging the high bandwidth of silicon photonics have been proposed. These works can be categorized into two general types: 1) hybrid optical/electrical interconnect architecture [3]–[6], in which a photonic packet-switched network and an electronic circuit-switched control network are combined to respectively deliver large size data messages and short control messages and 2) crossbar or Clos architectures, in which the interconnect is fully photonic [7]–[15]. Although these designs provide high and scalable bandwidth, they either suffer from relatively high latency due to the electrical control circuits for photonic path setup, or significant power/hardware overhead due to significant over-provisioned photonic channels. In future latency and power constrained CMPs, these characteristics will hobble the utility of photonic interconnect.

We propose LumiNOC, a novel PNoC architecture which addresses power and resource overhead due to channel over-provisioning, while reducing latency and maintaining high bandwidth in CMPs. The LumiNOC architecture makes three contributions: first, instead of conventional, globally distributed, photonic channels, requiring high laser power, we propose a novel channel sharing arrangement composed of sub-sets of cores in photonic subnets; second, we propose a novel, purely photonic, distributed arbitration mechanism, dynamic channel scheduling, which achieves extremely low-latency without degrading throughput; third, our photonic network architecture leverages the same wavelengths for channel arbitration and parallel data transmission, allowing efficient utilization of the photonic resources, lowering static power consumption.

We show, in a 64-node implementation, LumiNOC enjoys 50% lower latency at low loads and ∼40% higher throughput per Watt on synthetic traffic, versus previous PNoCs. Furthermore, LumiNOC reduces latency ∼40% versus an

Manuscript received September 8, 2013; revised January 21, 2014 and March 25, 2014; accepted April 10, 2014. Date of current version 15 May 2014. This paper was recommended by Associate Editor R. O. Topaloglu.

The authors are with the Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX 77843 USA (e-mail: seulc@tamu.edu; mabrowning@tamu.edu; pgratz@gratz1.com; spalermo@ece.tamu.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCAD.2014.2320510

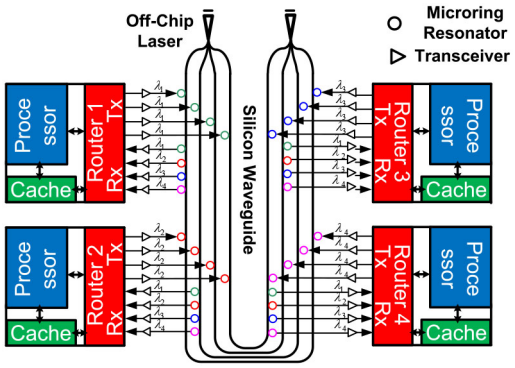

Fig. 1. Four-node fully connected photonic crossbar.

electrical 2-D mesh NoCs on PARSEC shared-memory, multithreaded benchmark workloads.

II. Background

PNoCs have emerged as a potential replacement for electrical NoCs due to the high bandwidth, low latency, and low power of nanophotonic channels. Fig. 1 shows a small CMP with four compute tiles interconnected by a PNoC. Each tile consists of a processor core, private caches, a fraction of the shared last-level cache, and a router connecting it to the photonic network. Fig. 1 also shows the details of an example PNoC, organized as a simple, fully connected crossbar interconnecting the four processors. The photonic channel connecting the nodes is shown as being composed of microring resonators (MRR) [16], [17], integrated photodetectors (PD) [18] (small circles) and silicon waveguides [19], [20] (black lines connecting the circles). Transceivers (small triangles) mark the boundary between the electrical and photonic domain. While the network shown is nonoptimal in terms of scalability, it is sufficient for introducing the components of a simple PNoC.

A. Microring Resonators (MRR)

MRRs can serve as either optical modulators for sending data or as filters for dropping and receiving data from on-chip photonic network. The basic configuration of an MRR consists of a silicon ring coupled with a straight waveguide. When the ring circumference equals an integer number of an optical wavelength, called resonance condition, most of the light from the straight waveguide circulates inside the ring and the light transmitted by the waveguide is suppressed. The resonance condition can be changed by applying electrical field over the ring, thus achieving electrical to optical modulation. MRRs resonance is sensitive to temperature variation, therefore, thermal trimming is required to tune the ring to resonate at the working wavelength.

B. Silicon Waveguides

In photonic on-chip networks, silicon waveguides are used to carry the optical signals. In order to achieve higher aggregated bandwidth, multiple wavelengths are placed into a single waveguide in a wavelength-division-multiplexing (WDM) fashion. As shown in Fig. 1, multiple wavelengths generated by an off-chip laser ($\lambda_1, \lambda_2, \lambda_3, \lambda_4$) are coupled into a silicon waveguide via an optical coupler. At the sender side, microring modulators insert data onto a specific wavelength through electro-optical modulation. The modulated wavelengths propagate through integrated silicon waveguide and arrive at the receiver side, where microring filters drop the corresponding wavelength and integrated PD convert the signals back to the electrical domain. In this paper, silicon nitride waveguides are assumed to be the primary transport strata. Similar to electrical wires, silicon nitride waveguides can be deployed into multiple strata to eliminate in-plane waveguide crossing, thus reducing the optical power loss [21].

C. 3-D Integration

In order to optimize system performance and efficiently utilize the chip area, 3-D integration (3-DI) is emerging for the integration of silicon nanophotonic devices with conventional CMOS electronics. In 3-DI, the silicon photonic on-chip networks are fabricated into a separate silicon-on-insulator (SOI) die or layer with a thick layer of buried oxide (BOX) that acts as bottom cladding to prevent light leakage into the substrate. This photonic layer stacks above the electrical layers containing the compute tiles.

In Fig. 1, the simple crossbar architecture is implemented by provisioning four send channels, each utilizing the same wavelength in four waveguides, and four receiving channels by monitoring four wavelengths in a single waveguide. Although this straightforward structure provides strictly nonblocking connectivity, it requires a large number of transceivers $O(r^2)$ and long waveguides crossing the chip, where r is the crossbar radix, thus this style of crossbar is not scalable to a significant number of nodes. Researchers have proposed a number of PNoC architectures more scalable than fully connected crossbars, as described below.

III. Related Work

Many PNoC architectures have been recently proposed which may be broadly categorized into four basic architectures: 1) electrical-photonic; 2) crossbar; 3) multistage; and 4) free-space designs.

A. Electrical-Photonic Designs

Shacham *et al.* [4] propose a hybrid electrical PNoC using electrical interconnect to coordinate and arbitrate a shared photonic medium [3]. These designs achieve very high photonic link utilization by effectively trading increased latency for higher bandwidth. While increased bandwidth without regard for latency is useful for some applications, it eschews a primary benefit of PNoCs over electrical NoCs, low latency. Recently, Hendry *et al.* [22] addressed this issue by introducing an all optical mesh network with photonic time division multiplexing (TDM) arbitration to set up communication path. However, the simulation results show that system still suffers from relatively high average latency.

B. Crossbar Designs

Other recent PNoC work attempts to address the latency issue by providing nonblocking point-to-point links between

nodes. In particular, several works propose crossbar topologies to improve the latency of multicore photonic interconnect. Fully connected crossbars [9] do not scale well, but researchers have examined channel sharing crossbar architectures, called single-write-multiple-read (SWMR) or multiple-write-single-read (MWSR), with various arbitration mechanisms for coordinating shared sending and/or receiving channels. Vantrease *et al.* [12], [13] proposed Corona, a MWSR crossbar, in which each node listens on the dedicated channel, but with the other nodes competing to send data on this channel. To implement arbitration at sender side, the author implemented a token channel [13] or token slot [12] approach similar to token rings used in early LAN network implementations. Alternately, Pan *et al.* [11] proposed Firefly, a SWMR crossbar design, with a dedicated sending channel for each node, but all the nodes in a crossbar listen on all the sending channels. Pan *et al.* [11] proposed broadcasting the flit-headers to specify a particular receiver.

In both SWMR and MWSR crossbar designs, over-provisioning of dedicated channels, either at the receiver (SWMR) or sender (MWSR), is required, leading to under utilization of link bandwidth and poor power efficiency. Pan *et al.* [10] also proposed a channel sharing architecture, FlexiShare, to improve the channel utilization and reduce channel over-provisioning. The reduced number of channels, however, limit the system throughput. In addition, FlexiShare requires separated dedicated arbitration channels for sender and receiver sides, incurring additional power, and hardware overhead.

Two recent designs propose to manage laser power consumption at runtime. Chen and Joshi [23] propose to switch off portions of the network at runtime dependent measured bandwidth requirements. Zhou and Kodi [24] proposed a method to predict future bandwidth needs and scale laser power appropriately.

C. Multistage Designs

Recently, Joshi *et al.* [7] proposed a photonic multistage Clos network with the motivation of reducing the photonic ring count, thus reducing the power for thermal ring trimming. Their design explores the use of a photonic network as a replacement for the middle stage of a three-stage Clos network. While this design achieves an efficient utilization of the photonic channels, it incurs substantial latency due to the multistage design.

Koka *et al.* [14] present an architecture consisting of a grid of nodes where all nodes in each row or column are fully connected by a crossbar. To maintain full-connectivity of the network, electrical routers are used to switch packets between rows and columns. In this design, photonic "grids" are very limited in size to maintain power efficiency, since fully connected crossbars grow at $O(n^2)$ for the number of nodes connected. Kodi and Morris [25] propose a 2-D mesh of optical MWSR crossbars to connect nodes in the x and y dimensions. In a follow-on work by the same authors Morris and Kodi [26] proposed a hybrid multistage design, in which grid rows (x-dir) are subnets fully connected with a photonic crossbar, but different rows (y-dir) are connected by a token-ring arbitrated shared photonic link. Bahirat and Pasricha [27] propose an adaptable hybrid design in which a 2-D mesh electrical network is overlaid with a set of photonic rings.

D. Free-Space Designs

Xue *et al.* [28] present a novel free-space optical interconnect for CMPs, in which optical free-space signals are bounced off of mirrors encapsulated in the chip's packaging. To avoid conflicts and contention, this design uses in-band arbitration combined with an acknowledgment based collision detection protocol.

Our proposed architecture, LumiNOC, attempts to address the issues found in competing designs. As in FlexiShare [10] and Clos [7], LumiNOC focuses on improving the channel utilization to achieve better efficiency and performance. Unlike these designs, however, LumiNOC leverages the same channels for arbitration, parallel data transmission, and flow control, efficiently utilizing the photonic resources. Similar to Clos [7], LumiNOC is also a multistage design, however, unlike Clos, the primary stage (our subnets) is photonic and the intermediate is electrical, leading to much lower photonic energy losses in the waveguide and less latency due to simplified intermediate node electronic routers. Similar to Xue *et al.* design [28], in-band arbitration with collision detection is used to coordinate channel usage; however, in LumiNOC, the sender itself detects the collision and may start the retransmit process immediately without waiting for an acknowledgment, which may increase latency due to timeouts and reduce channel bandwidth utilization. These traits give LumiNOC better performance in terms of latency, energy efficiency, and scalability.

IV. Power Efficiency in PNoCs

Power efficiency is an important motivation for photonic on-chip interconnect. In photonic interconnect, however, the static power consumption (due to off-chip laser, ring thermal tuning, etc.) dominates the overall power consumption, potentially leading to energy-inefficient photonic interconnects. In this section, we examine prior PNoCs in terms of static power efficiency. We use bandwidth per watt as the metric to evaluate power efficiency of photonic interconnect architectures, showing that it can be improved by optimizing the interconnect topology, arbitration scheme, and photonic device layout.

A. Channel Allocation

We first examine channel allocation in prior photonic interconnect designs. Several previous PNoC designs, from fully connected crossbars [9] to the blocking crossbar designs [8], [10]–[13], provision extra channels to facilitate safe arbitration between sender and receiver. Although conventional photonic crossbars achieve nearly uniform latency and high bandwidth, channels are dedicated to each node and cannot be flexibly shared by the others. Due to the unbalanced traffic distribution in realistic workloads [29], channel bandwidth cannot be fully utilized. This leads to inefficient energy usage, since the static power is constant regardless of traffic load. Over-provisioned channels also implies higher ring resonator counts,

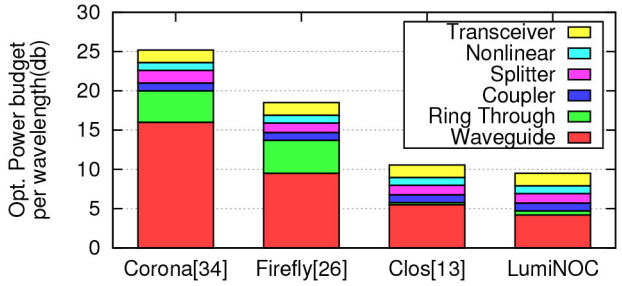

Fig. 2. Optical link budgets for the photonic data channels of various photonic NoCs.

which must be maintained at the appropriate trimming temperature, consuming on-chip power. Additionally, as the network size increases, the number of channels required may increase quadratically, complicating the waveguide layout and leading to extra optical loss. An efficient photonic interconnect must solve the problem of efficient channel allocation. Our approach leverages this observation to achieve lower power consumption than previous designs.

B. Topology and Layout

Topology and photonic device layout can also cause unnecessary optical loss in the photonic link, which in turn leads to greater laser power consumption. Many PNoCs globally route waveguides in a bundle, connecting all the tiles in the CMP [8], [11]–[13]. In these designs, due to the unidirectional propagation property of optical transmission, the waveguide must double back to reach each node twice, such that the signal being modulated by senders on the outbound path may be received by all possible receivers. The length of these double-back waveguides leads to significant laser power losses over the long distance.

Fig. 2 shows the optical link budgets for the photonic data channel of Corona [13], Firefly [11], Clos [7], and LumiNOC under same radix and chip area, based on our power model (described in Section VI-E). Flexishare [10] is not compared, since not enough information was provided in the paper to estimate the optical power budget at each wavelength. The figure shows that waveguide losses dominate power loss in all three designs. This is due to the long waveguides required to globally route all the tiles on a chip. For example, the waveguide length in Firefly and Clos network in a 400 mm^2 chip are estimated to be 9.5 and 5.5 cm, respectively. This corresponds to 9.5 and 5.5 dB loss in optical power, assuming the waveguide loss is 1 dB/cm [7]. Moreover, globally connected tiles imply a relatively higher number of rings on each waveguide, leading to higher ring through loss. Despite a single-run, bi-directional architecture, even the Clos design shows waveguide loss as the largest single component.

In contrast to other losses (e.g., coupler and splitter loss, filter drop loss, and photodetector loss) which are relatively independent of interconnect architecture, waveguide and ring through loss can be reduced through layout and topology optimization. We propose a network architecture which reduces optical loss by decreasing individual waveguide length as well as the number of rings along the waveguide.

C. Arbitration Mechanism

The power and overhead introduced by the separated arbitration channels or networks in previous PNoCs can lead to further power efficiency losses. Corona, a MWSR crossbar design, requires a token channel or token slot arbitration at sender side [12], [13]. Alternatively, Firefly [11], a SWMR crossbar design, requires head-flit broadcasting for arbitration at receiver side, which is highly inefficient in PNoCs. FlexiShare [10] requires both token stream arbitration and head-flit broadcast. These arbitration mechanisms require significant overhead in form of dedicated channels and photonic resources, consuming extra optical laser power. For example, the radix-32 Flexishare [10] with 16 channels requires 416 extra wavelengths for arbitration, which accounts for 16% of the total wavelengths in addition to higher optical power for a multireceiver broadcast of head-flits. Arbitration mechanisms are a major overhead for these architectures, particularly as network radix scales.

There is a clear need for a PNoC architecture that is energy-efficient and scalable while maintaining low latency and high bandwidth. In the following sections, we propose the LumiNOC architecture which reduces the optical loss by partitioning the global network into multiple smaller sub-networks. Further, a novel arbitration scheme is proposed which leverages the same wavelengths for channel arbitration and parallel data transmission to efficiently utilize the channel bandwidth and photonic resources, without dedicated arbitration channels or networks which lower efficiency or add power overhead to the system.

V. LumiNOC Architecture

In our analysis of prior PNoC designs, we found a significant amount of laser power consumption was due to the waveguide length required for propagation of the photonic signal across the entire network. Based on this, the LumiNOC design breaks the network into several smaller networks (subnets), with shorter waveguides. Fig. 3 shows three example variants of the LumiNOC architecture with different subnet sizes, in an example 16-node CMP system: the one-row, two-rows, and four-rows designs (note: 16-nodes are shown to simplify explanation, in Section VI we evaluate a 64-node design). In the one-row design, a subnet of four tiles is interconnected by a photonic waveguide in the horizontal orientation. Thus, four nonoverlapping subnets are needed for the horizontal interconnection. Similarly, four subnets are required to vertically interconnect the 16 tiles. In the two-row design, a single subnet connects eight tiles while in the four-row design a single subnet touches all 16 tiles. In general, all tiles are interconnected by two different subnets, one horizontal and one vertical. If a sender and receiver do not reside in the same subnet, transmission requires a hop through an intermediate node's electrical router. In this case, transmission experiences longer delay due to the extra O/E-E/O conversions and router latency. To remove the overheads of photonic waveguide crossings required by the orthogonal set of horizontal and vertical subnets, the waveguides can be deposited into two layers with orthogonal routing [21].

Another observation from prior PNoC designs is that channel sharing and arbitration have a large impact on design power

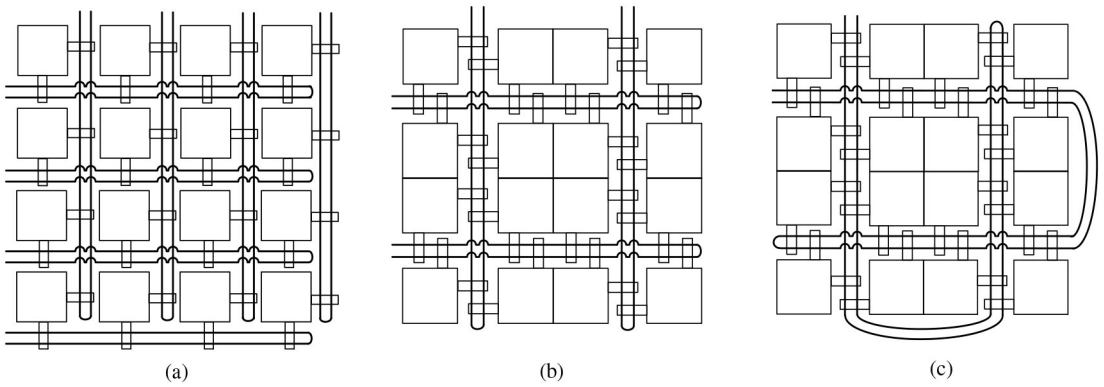

Fig. 3. LumiNOC interconnection of CMP with 16 tiles. (a) One-row (b) Two-rows and (c) Four-rows interconnection.

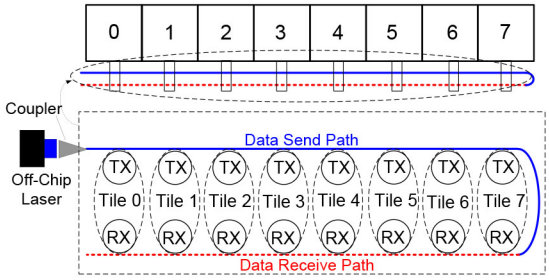

Fig. 4. One-row subnet of eight nodes. Circles (TX and RX) represent groups of rings; one dotted oval represents a tile.

efficiency. Efficient utilization of the photonic resources, such as wavelengths and ring resonators, is required to yield the best overall power efficiency. To this end, we leverage the same wavelengths in the waveguide for channel arbitration and parallel data transmission, avoiding the power and hardware overhead due to the separated arbitration channels or networks. Unlike the over-provisioned channels in conventional crossbar architectures, channel utilization in LumiNOC is improved by multiple tiles sharing a photonic channel.

A final observation from our analysis of prior PNoC design is that placing many wavelengths within each waveguide through deep WDM leads to high waveguide losses. This is because the number of rings that each individual wavelength encounters as it traverses the waveguide is proportional to the number of total wavelengths in the waveguide times the number of waveguide connected nodes, and each ring induces some photonic power losses. We propose to limit LumiNOC's waveguides to a few frequencies per waveguide and increase the count of waveguides per subnet, to improve power efficiency with no cost to latency or bandwidth, a technique we call "ring-splitting." Ring-splitting is ultimately limited by the tile size and optical power splitting loss. Assuming a reasonable waveguide pitch of 15 μm required for layout of microrings which have a diameter of 5 μm [30], this leaves 5 μm clearance to avoid optical signal interference between two neighbor rows of rings.

A. LumiNOC Subnet Design

Fig. 4 details the shared channel for a LumiNOC one-row subnet design. Each tile contains Λ modulating "Tx rings" and Λ receiving "Rx Rings," where Λ is the number of wavelengths multiplexed in the waveguide. Since the optical signal unidirectionally propagates in the waveguide from its source at off-chip laser, each node's Tx rings are connected in series on the "data send path," shown in a solid line from the laser, prior to connecting each node's Rx rings on the "data receive path," shown in a dashed line. In this "double-back" waveguide layout, modulation by any node can be received by any other node; furthermore, the node which modulates the signal may also receive its own modulated signal, a feature that is leveraged in our collision detection scheme in the arbitration phase. The same wavelengths are leveraged for arbitration and parallel data transmission.

During data transmission, only a single sender is modulating on all wavelengths and only a single receiver is tuned to all wavelengths. However, during arbitration (i.e., any time data transfer is not actively occurring) the Rx rings in each node are tuned to a specific, nonoverlapping set of wavelengths. Up to half of the wavelengths available in the channel are allocated to this arbitration procedure. with the other half available for credit packets as part of credit-based flow control. This particular channel division is designed to prevent optical broadcasting, the state when any single wavelength must drive more than one receiver, which if allowed would severely increase laser power [31]. Thus, at any given time a multiwavelength channel with N nodes may be in one of three states: idle—all wavelengths are un-modulated and the network is quiescent; arbitration—one more sender nodes are modulating N copies of the arbitration flags; one copy to each node in the subnet (including itself) with the aim to gain control of the channel; data transmission—once a particular sender has established ownership of the channel, it modulates all channel wavelengths in parallel with the data to be transmitted.

In the remainder of this section, we detail the following: Arbitration—the mechanism by which the photonic channel is granted to one sender, avoiding data corruption when multiple senders wish to transmit, including dynamic channel scheduling, the means of sender conflict resolution, and Data Transmission—the mechanism by which data is transmitted from sender to receiver. Credit return is also discussed.

1) Arbitration: We propose an optical collision detecting and dynamic channel scheduling technique to coordinate

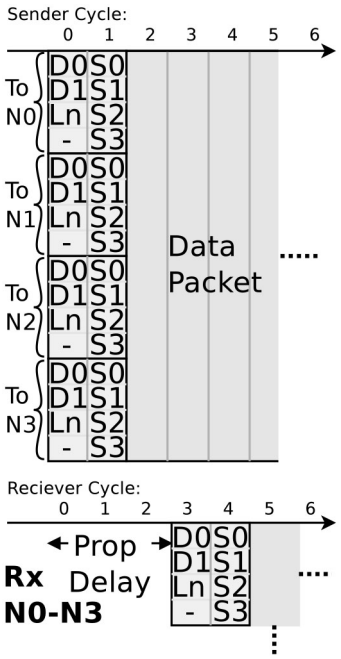

Fig. 5. Arbitration on a four-node subnet.

access of the shared photonic channel. This approach achieves efficient channel utilization without the latency of electrical arbitration schemes [3], [4], or the overhead of wavelengths and waveguides dedicated to standalone arbitration [10], [11], [13]. In this scheme, a sender works together with its own receiver to ensure message delivery in the presence of conflicts.

a) Receiver: Once any receiver detects an arbitration flag, it will take one of three actions: if the arbitration flag is uncorrupted (i.e., the sender flag has a 0 in only one location indicating single-sender) and the forthcoming message is destined for this receiver, it will enable all its Rx rings for the indicated duration of the message, capturing it. If the arbitration flags are uncorrupted, but the receiver is not the intended destination, it will detune all of its Rx rings for the indicated duration of the message to allow the recipient sole access. Finally, if a collision is detected, the receiver circuit will enter the dynamic channel scheduling phase (described below).

b) Sender: To send a packet, a node first waits for any on-going messages to complete. Then, it modulates a copy of the arbitration flags to the appropriate arbitration wavelengths for each of the N nodes. The arbitration flags for an example four-node subnet are depicted in Fig. 5. The arbitration flags are a t_{arb} cycle long header (2 in this example) made up of the destination node address (D0–D1), a bimodal packet size indicator (Ln) for the two supported payload lengths (64-bit and 576-bit), and a "1-hot" source address (S0–S3) which serves as a guard band or collision detection mechanism: since the subnet is operated synchronously, any time multiple nodes send overlapping arbitration flags, the "1-hot" precondition is violated and all nodes are aware of the collision. We leverage self-reception of the arbitration flag: right after sending, the node monitors the incoming arbitration flags. If they are uncorrupted, then the sender succeeded arbitrating the channel and the two nodes proceed to the data transmission phase. If the arbitration flags are corrupted (>1 is hot), then a conflict has occurred. Any data already sent is ignored and the conflicting senders enter the dynamic channel scheduling regime (described below).

The physical length of the waveguide incurs a propagation delay, t_{pd} (cycles), on the arbitration flags traversing the subnet. The "1-hot" collision detection mechanism will only function if the signals from all senders are temporally aligned, so if nodes are physically further apart than the light will travel in one cycle, they will be in different clocking domains to keep the packet aligned as it passes the final sending node. Furthermore, the arbitration flags only start on cycles that are an integer multiple of the $t_{pd} + 1$ to assure that no nodes started arbitration during the previous t_{slot} and that all possibly conflicting arbitration flags are aligned. This means that conflicts only occur on arbitration flags, not with data.

Note that a node will not know if it has successfully arbitrated the channel until after $t_{pd} + t_{arb}$ cycles, but will begin data transmission after t_{arb}. In the case of an uncontested link, the data will be captured by the receiver without delay. Upon conflict, senders cease sending (unusable) data.

As an example, say that the packet in Fig. 5 is destined for node 2 with no conflicts. At cycle 5, nodes 1, 3, and 4 would detune their receivers, but node 2 would enable them all and begin receiving the data flits.

If the subnet size were increased without proportionally increasing the available wavelengths per subnet, then the arbitration flags will take longer to serialize as more bits will be required to encode the source and destination address. If, however, additional wavelengths are provisioned to maintain the bandwidth/node, then the additional arbitration bits are sent in parallel. Thus, the general formula for $t_{arb} = ceil(1 + N + log_2(N)/\lambda)$ where N is the number of nodes and λ is the number of wavelengths per arbitration flag.

2) Dynamic Channel Scheduling: Upon sensing a conflicting source address, all nodes identify the conflicting senders and a dynamic, fair schedule for channel acquisition is determined using the sender node index and a global cycle count (synchronized at startup): senders transmit in $(n + cycle)$ mod N order. Before sending data in turn, each sender transmits an abbreviated version of the arbitration flags: the destination address and the packet size. All nodes tune in to receive this, immediately followed by the data transmission phase with a single sender and receiver for the duration of the packet. Immediately after the first sender sends its last data flit, next sender repeats this process, keeping the channel occupied until the last sender completes. After the dynamic schedule completes, the channel goes idle and any node may attempt a new arbitration to acquire the channel as previously described.

3) Data Transmission: In this phase, the sender transmits the data over the photonic channel to the receiving node. All wavelengths in the waveguide are used for bit-wise parallel data transmission, so higher throughput is expected when more wavelengths are multiplexed into the waveguide. Two packet payload lengths, 64-bit for simple requests and coherence traffic and 576-bit for cache line transfer, are supported.

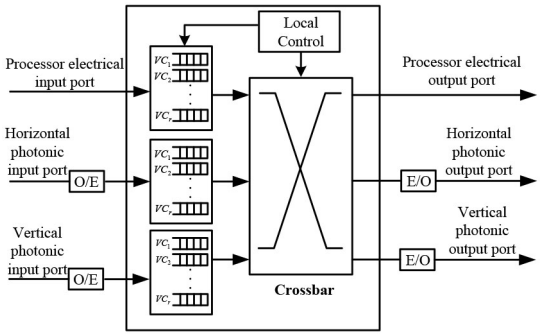

Fig. 6. Router microarchitecture.

4) Credit Return: At the beginning of any arbitration phase (assuming the channel is not in use for data transmission), 1/2 of the wavelengths of the channel are reserved for credit return from the credit return transmitter (i.e., the router which has credit to return) to the credit return receiver (i.e., the node which originally sent the data packet and now must be notified of credit availability). Similar to the arbitration flags, the wavelengths are split into *N* different sub-channels, each one dedicated to a particular credit return receiver. Any router which has credit to send back may then modulate its credit return flag onto the sub-channel to the appropriate credit return receiver. The credit return flag is encoded similarly to the arbitration flag. In the event of a collision between two credit return senders returning credit to the same receiver, no retransmission is needed as the sender part of the flag will uniquely identify all nodes sending credit back to this particular credit return receiver. Credit is returned on a whole-packet basis, rather than a flit basis to decrease overheads. The packet size bit *Ln* is not used in the credit return flag; credit return receivers must keep a history of the packet sizes transmitted so that the appropriate amount of credit is returned.

B. Router Microarchitecture

The electrical router architecture for LumiNOC is shown in Fig. 6. Each router serves both as an entry point to the network for a particular core, as well as an intermediate node interconnecting horizontal and vertical subnets. If a processor must send data to another node on the same vertical or horizontal subnet, packets are switched from the electrical input port to the vertical photonic output port with one E/O conversion. Packets which are destined for a different subnet must be first routed to an intermediate node via the horizontal subnet before being routed on the vertical subnet. Each input port is assigned with a particular virtual-channel (VCs) to hold the incoming flits for a particular sending node. The local control unit performs routing computation, VC allocation and switching allocation in crossbar. The LumiNOC router's complexity is similar to that of an electrical, bi-directional, 1-D ring network router, with the addition of the E/O-O/E logic.

VI. Evaluation

In this section, we describe a particular implementation of the LumiNOC architecture and analyze its performance and power efficiency.

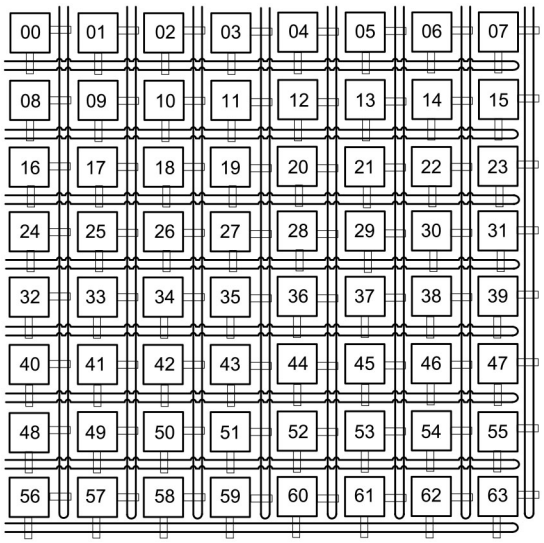

Fig. 7. One-row LumiNOC with 64 tiles.

A. 64-Core LumiNOC Implementation

Here, we develop a baseline physical implementation of the general LumiNOC architecture specified in Section V for the evaluation of LumiNOC against competing PNOC architectures. We assume a 400 mm^2 chip implemented in a 22 nm CMOS process and containing 64 square tiles that operate at 5 GHz, as shown in Fig. 7. A 64-node LumiNOC design point is chosen here as a reasonable network size which could be implemented in a 22 nm process technology. Each tile contains a processor core, private caches, a fraction of the shared last-level cache, and a router connecting it to one horizontal and one vertical photonic subnet. Each router input port contains seven VCs, each five flits deep. Credit-based flow control is implemented via the remainder of the photonic spectrum not used for arbitration during arbitration periods in the network.

A 64-node LumiNOC may be organized into three different architectures: the one-row, two-row, and four-row designs (shown in Fig. 3), which represent a trade-off between interconnect power, system throughput, and transmission latency. For example, power decreases as row number increases from one-row to two-row, since the single waveguide is roughly with the same length, but fewer waveguides are required. The low-load latency is also reduced due to more nodes residing in the same subnet, reducing the need for intermediate hops via an electrical router. The two-row subnet design, however, significantly reduces throughput due to the reduced number of transmission channels. As a result, we choose the "one-row" subnet architecture of Fig. 3(a), with 64-tiles arranged as shown in Fig. 7 for the remainder of this section. In both the horizontal and vertical axes, there are eight subnets which are formed by eight tiles that share a photonic channel, resulting in all tiles being redundantly interconnected by two subnets. As discussed in Section II, 3-DI is assumed, placing orthogonal waveguides into different photonic layers, eliminating in-plane waveguide crossings [21].

As a general trend, multirow designs tend to decrease power consumption in the router as fewer router hops are required

to cover more of the network. Because of the diminishing returns in terms of throughput as channel width increases, however, congestion increases and the bandwidth efficiency drops. Further, the laser power grows substantially for a chip as large as the one described here. For smaller floorplans, however, multirow LumiNOC would be an interesting design point.

We assume a 10 GHz network modulation rate, while the routers and cores are clocked at 5 GHz. Muxes are placed on input and output registers such that on even network cycles, the photonic ports will interface with the lower half of a given flit and on odd, the upper half. With a 400 mm^2 chip, the effective waveguide length is 4.0 cm, yielding a propagation delay of $t_{pd} = 2.7$, 10 GHz, network cycles.

When sender and receiver reside in the same subnet, data transmission is accomplished with a single hop, i.e., without a stop in an intermediate electrical router. Two hops are required if sender and receiver reside in different subnets, resulting in a longer delay due to the extra O/E-E/O conversion and router latency. The "one-row" subnet-based network implies that for any given node 15 of the 63 possible destinations reside within one hop, the remaining 48 destinations require two hops.

1) Link Width Versus Packet Size: Considering the link width, or the number of wavelengths per logical subnet, if the number of wavelengths and thus channel width is increased, it should raise ideal throughput and theoretically reduce latency due to serialization delay. We are constrained, however, by the 2.7 network cycle propagation delay of the link, and the small packet size of single cache line transfers in typical CMPs. There is no advantage to sending the arbitration flags all at once in parallel when additional photonic channels are available; the existing bits would need to be replaced with more guard bits to provide collision detection. Thus, the arbitration flags would represent an increasing overhead. Alternately, if the link were narrower, the 2.7 cycle window would be too short to send all the arbitration bits and a node would waste time broadcasting arbitration bits to all nodes after it effectively "owns" the channel. Thus, the optimal link width is 64 wavelengths under our assumptions for clock frequency and waveguide length.

If additional spectrum or waveguides are available, then we propose to implement multiple parallel, independent network layers. Instead of one network with a 128-bit data path, there will be two parallel 64-bit networks. This allows us to exploit the optimal link width while still providing higher bandwidth. When a node injects into the network, it round-robins through the available input ports for each layer, dividing the traffic amongst the layers evenly.

2) Ring-Splitting: Given a 400 mm^2 64-tile PNoC system, each tile is physically able to contain 80 double-back waveguides. However, the ring-splitting factor is limited to four (32 wavelengths per waveguide) in this design to avoid the unnecessary optical splitting loss due to the current technology. This implies a trade off of waveguide area for lower power. The splitting loss has been included in the power model in Section VI-E.

3) Scaling to Larger Networks: We note, it is likely that increasing cores connected in a given subnet will yield increased contention well. A power-efficient means to cover the increase in bandwidth demand due to more nodes would be to increase the number of layers. We find the degree of subnet partitioning is more dependent upon the physical chip dimensions than the number of nodes connected, as the size of the chip determines the latency and frequency of arbitration phases. For this reason, our base implementation assumes a large, 400 mm^2 die. Increasing nodes while retaining the same physical dimensions will cause a sub-linear increase in arbitration flag size with nodes-per-subnet (the source ID would increase linearly, the Destination ID would increase as $log(n)$), and hence more overhead than in a smaller sub-net design.

B. Experiment Methodology

To evaluate this implementation's performance, we use a cycle-accurate, microarchitectural-level network simulator, ocin _tsim [32]. The network was simulated under both synthetic and realistic workloads. LumiNOC designs with 1, 2, and 4 network layers are simulated to show results for different bandwidth design points.

1) Photonic Networks: The baseline, 64-node LumiNOC system, as described in Section VI, was simulated for all evaluation results. Synthetic benchmark results for the Clos LTBw network are presented for comparison against the LumiNOC design. We chose the Clos LTBw design as the most competitive in terms of efficiency and bandwidth as discussed in Section VI. Clos LTBw data points were extracted from the paper by Joshi *et al.* [7].

2) Baseline Electrical Network: In the results that follow, our design is compared to a electrical 2-D mesh network. Traversing the dimension order network consumes three cycles per hop; one cycle for link delay and two within each router. The routers have two virtual channels per port, each 10 flits deep, and implement wormhole flow control.

3) Workloads: Both synthetic and realistic workloads were simulated. The traditional synthetic traffic patterns, uniform random and bit-complement represent nominal and worst-case traffic for this design. These patterns were augmented with the P8D pattern, proposed by Joshi *et al.* [7], designed as a best-case for staged or hierarchical networks where traffic is localized to individual regions. In P8D, nodes are assigned to one of eight groups, made up of topologically adjacent nodes and nodes only send random traffic within the group. In these synthetic workloads, all packets contain data payloads of 512-bits, representing four flits of data in the baseline electrical NoC.

Realistic workload traces were captured for a 64-core CMP running PARSEC benchmarks with the sim-large input set [33]. The Netrace trace dependency tracking infrastructure was used to ensure realistic packet interdependencies are expressed as in a true, full-system CMP system [34]. The traces were captured from a CMP composed of 64 in-order cores with 32-KB, private L1I and L1D caches and a shared 16MB LLC. Coherence among the L1 caches was maintained via a MESI protocol. A 150 million cycle segment of the PARSEC benchmark "region of interest" was simulated. Packet sizes for realistic workloads vary bimodally between

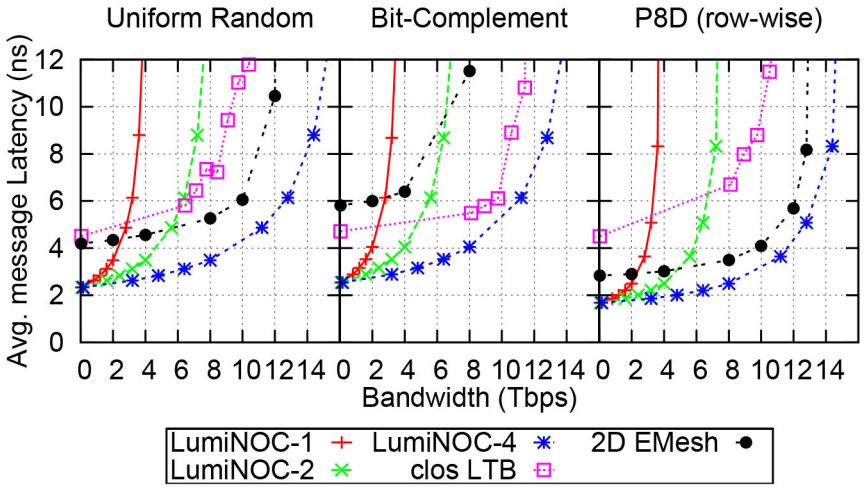

Fig. 8. Synthetic workloads showing LumiNOC versus Clos LTBw and electrical network. LumiNOC-1 refers to the one-layer LumiNOC design, LumiNOC-2 the two-layer, and LumiNOC-4 the four-layer.

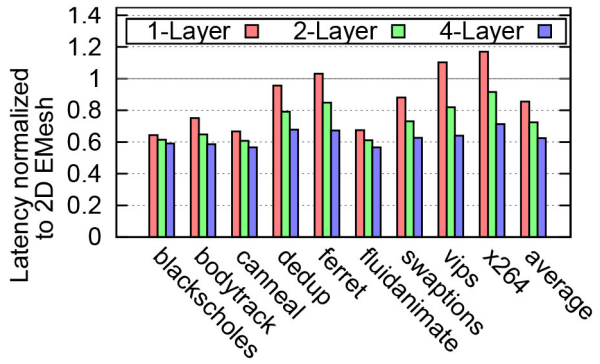

Fig. 9. Message latency in PARSEC benchmarks for LumiNOC compared to electrical network.

64 and 576 bits for miss request/coherence traffic and cache line transfers.

C. Synthetic Workload Results

In Fig. 8, the LumiNOC design is compared against the electrical and Clos networks under uniform random, bit complement, and P8D. The figure shows the low-load latencies of the LumiNOC design are much lower than the competing designs. This is due primarily to the lower diameter of the LumiNOC topology, destinations within one subnet are one "hop" away while those in a second subnet are two. The one-layer network saturates at 4 Tbps realistic throughput as determined by analyzing the offered versus accepted rate.

The different synthetic traffic patterns bring out interesting relationships. On the P8D pattern, which is engineered to have lower hop counts, all designs have universally lower latency than on other patterns. However, while both the electrical and LumiNOC network have around 25% lower low-load latency than uniform random, Clos only benefits by a few percent from this optimal traffic pattern. At the other extreme, the electrical network experiences a 50% increase in no-load latency under the bit-complement pattern compared to uniform random while both Clos and the LumiNOC network are only marginally affected. This is due to the LumiNOC having a worst-case hop count of two and not all routes go through the central nodes as in the electrical network. Instead, the intermediate nodes are well distributed through the network under this traffic pattern. However, as the best-case hop count is also two with this pattern, the LumiNOC network experiences more contention and the saturation bandwidth is decreased as a result.

TABLE I
COMPONENTS OF OPTICAL LOSS

Loss Component	Value	Loss Component	Value
Coupler	1 dB	Waveguide	1 dB/cm
Splitter	0.2 dB	Waveguide Crossing	0.05 dB
Non-linearity	1 dB	Ring Through	0.001 dB
Modulator Insertion	0.001 dB	Filter Drop	1.5 dB
Photodetector	0.1 dB		

D. Realistic Workload Results

Fig. 9 shows the performance of the LumiNOC network in one-layer, two-layers, and four-layers, normalized against the performance of the baseline electrical NoC. Even with one-layer, the average message latency is about 10% lower than the electrical network. With additional network layers, LumiNOC has approximately 40% lower average latency. These results are explained by examining the bandwidth-latency curves in Fig. 8. The average offered rates for the PARSEC benchmarks are of the order of 0.5 Tbps, so these applications benefit from LumiNOC's low latency while being well under even the one-layer, LumiNOC throughput.

E. Power Model

In this section, we describe our power model and compare the baseline LumiNOC design against prior work PNoC architectures. In order for a fair comparison versus other reported PNoC architectures, we refer to the photonic loss of various photonic devices reported by Joshi *et al.* [7] and Pan *et al.* [10],

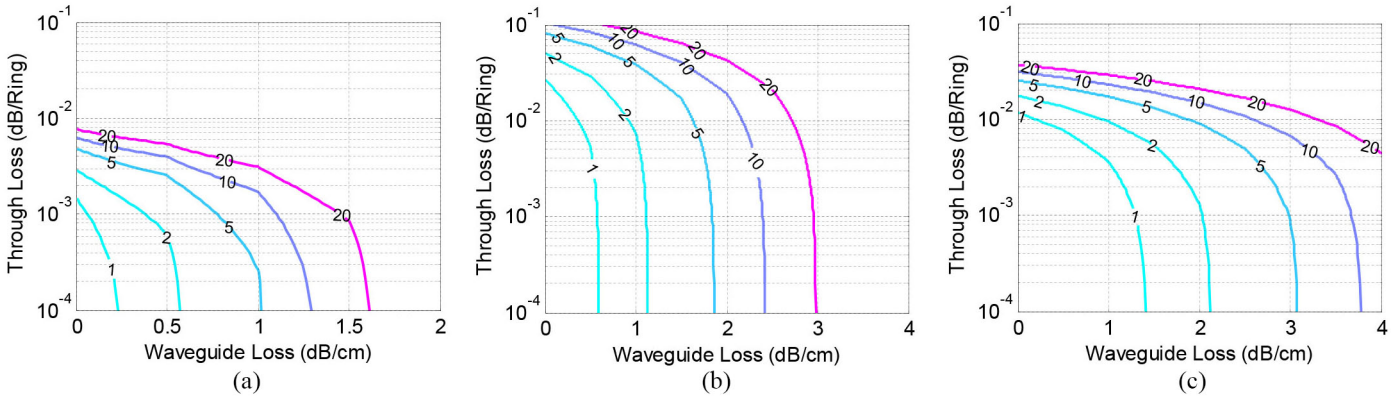

Fig. 10. Contour plots of the electrical laser power (ELP) in watts for networks with the same aggregate throughput. Each line represents a constant power level (watts) at a given ring through loss and waveguide loss combination (assuming 30% efficient electrical to optical power conversion). (a) Crossbar. (b) Clos. (c) LumiNOC.

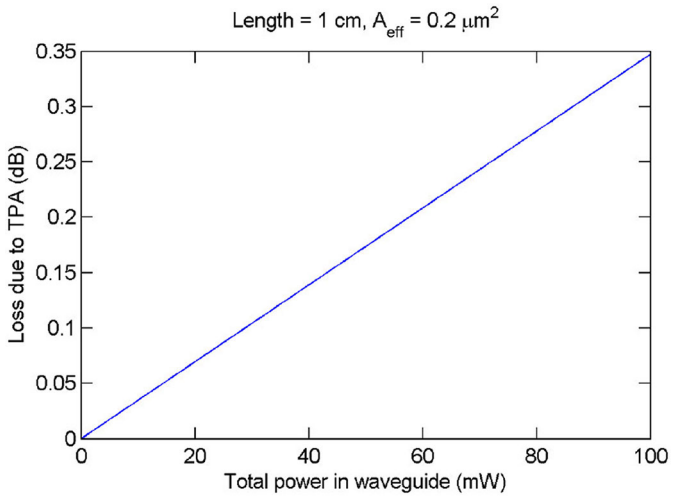

Fig. 11. Nonlinear optical loss in the silicon waveguide versus optical power in waveguide; waveguide length equals 1 cm with effective area of 0.2 um^2. Figure produced by Jason Pelc of HP labs with permission.

shown in Table I. Equation 1 shows the major components of our total power model

$$TP = ELP + TTP + ERP + EO/OE. \quad (1)$$

TP = total power; ELP = electrical laser power; TTP = thermal tuning power; ERP = electrical router power; and EO/OE = electrical to optical/optical to electrical conversion power. Each components is described below.

1) ELP: Electrical laser power is converted from the calculated optical power. Assuming a 10 μW receiver sensitivity, the minimum static optical power required at each wavelength to activate the farthest detector in the PNoC system is estimated based on 2. This optical power is then converted to electrical laser power using 30% efficiency

$$P_{optical} = N_{wg} \cdot N_{wv} \cdot P_{th} \cdot K \cdot 10^{\left(\frac{1}{10} \cdot l_{channel} \cdot P_{WG\_loss}\right)} \cdot 10^{\left(\frac{1}{10} \cdot N_{ring} \cdot P_{t\_loss}\right)}. \quad (2)$$

In 2, N_{wg} is the number of waveguide in the PNoC system, N_{wv} is the number of wavelength per waveguide, P_{th} is receiver sensitivity power, $l_{channel}$ is waveguide length, P_{wg_loss} is optical signal propagation loss in waveguide (dB / cm), N_{ring} is the number of rings attached on each waveguide, P_{t_loss} is modulator insertion and filter ring through loss (dB / ring) (assume they are equal), K accounts for the other loss components in the optical path including P_c, coupling loss between the laser source and optical waveguide, P_b, waveguide bending loss, and $P_{splitter}$, optical splitter loss. Fig. 10 shows electrical laser power contour plot, derived from 2, showing the photonic device power requirements at a given electrical laser power, for a SWMR photonic crossbar (Corona) [13], Clos [7], and LumiNOC with equivalent throughput (20Tbps), network radix and chip area. In the figure, x and y-axis represent two major optical loss components, waveguide propagation loss and ring through loss, respectively. A larger x- and y-intercept implies relaxed requirements for the photonic devices. As shown, given a relatively low 1 W laser power budget, the two-layer LumiNOC can operate with a maximum 0.012 dB ring through loss and waveguide loss of 1.5 dB/cm.

We note that optical nonlinear loss also effects the optical interconnect power. At telecom wavelengths, two-photon absorption (TPA) in the silicon leads to a propagation loss that increases linearly with the power sent down the waveguide. TPA is a nonlinear optical process and is several orders of magnitude weaker than linear absorption. This nonlinear loss, however, also has significant impact on the silicon-photonic

TABLE II
CONFIGURATION COMPARISON OF VARIOUS PHOTONIC NOC ARCHITECTURES— N_{core} : NUMBER OF CORES IN THE CMP; N_{node} : NUMBER OF NODES IN THE NOC; N_{rt} : TOTAL NUMBER OF ROUTERS; N_{wg} : TOTAL NUMBER OF WAVEGUIDES; N_{wv} : TOTAL NUMBER OF WAVELENGTHS; N_{ring} : TOTAL NUMBER OF RINGS

Literature		N_{core}	N_{node}	N_{rt}	N_{wg}	N_{wv}	N_{ring}
EMesh [1]		128	64	64	NA	NA	NA
Corona [13]		256	64	64	388	24832	1056K
FlexiShare [10]		64	32	32	NA	2464	550K
Clos [7]		64	8	24	56	3584	14K
LumiNOC	1-L	64	64	64	32	1024	16K
	2-L	64	64	64	64	2048	32K
	4-L	64	64	64	128	4096	65K

TABLE III
POWER EFFICIENCY COMPARISON OF DIFFERENT PHOTONIC NOC ARCHITECTURES— ELP : ELECTRICAL LASER POWER; TTP : THERMAL TUNING POWER; ERP : ELECTRICAL ROUTER POWER; EO/OE : ELECTRICAL TO OPTICAL/OPTICAL TO ELECTRICAL CONVERSION POWER; ITP : IDEAL THROUGHPUT; RTP : REALISTIC THROUGHPUT; TP : TOTAL POWER

Literature		ELP (W)	TTP (W)	ERP (W)	EO/OE (W)	ITP (Tbps)	RTP (Tpbs)	TP (W)	**RTP/W (Tbps/W)**
EMesh [1]		NA	NA	NA	NA	10	3.0	26.7	**0.1**
Corona [13]		26.0	21.00	0.52	4.92	160	73.6	52.4	**1.4**
FlexiShare [10]		5.80	11.00	0.13	0.60	20	9.0	17.5	**0.5**
Clos [7]		3.30	0.14	0.10	0.54	18	10.0	4.1	**2.4**
LumiNOC	1-Layer	0.35	0.33	0.13	0.30	10	4.0	1.1	**3.6**
	2-Layers	0.73	0.65	0.26	0.61	20	8.0	2.3	**3.4**
	4-Layers	1.54	1.31	0.52	1.22	40	16.0	4.6	**3.4**

link power budget, if a high level of optical power (∼Watt) is injected into silicon waveguide. Fig. 11 shows the computed nonlinear loss of a 1 cm waveguide versus the optical power in the waveguide. It shows a nonlinear loss of ∼0.35 dB for up to ∼100 mW waveguide optical power. In LumiNoC, the nonlinear effect has been included in the optical power calculation.

2) TTP: Thermal tuning is required to maintain microring resonant at the work wavelength. In the calculation, a ring thermal tuning power of 20 μW is assumed for a 20 K temperature tuning range [7], [10]. In a PNoC, total TTP is proportional to ring count.

3) ERP: The baseline electrical router power is estimated by the power model reported by Kim *et al*. [35]. We synthesized the router using TSMC 45 nm library. Power is measured via synopsis power compiler, using simulated traffic from a PARSEC [33] workload to estimate its dynamic component. Results are analytically scaled to 22 nm (dynamic power scaled according to the CMOS dynamic power equation and static power linearly with voltage).

4) EO/OE: The power for conversion between the EO/OE is based on the model reported by Joshi *et al.* [7], which assumes a total transceiver energy of 40 fJ/bit data-traffic dependent energy and 10 fJ/bit static energy. Since previous PNoCs consider different traffic loads, it is unfair to compare the EO/OE power by directly using their reported figures. Therefore, we compare the worst-case power consumption when each node was arbitrated to get a full access on each individual channel. For example, Corona is a MWSR 64×64 crossbar architecture. At worst-case, 64 nodes are simultaneously writing on 64 different channels. This is combined with a per-bit activity factor of 0.5 to represent random data in the channel.

While this approach may not be 100% equitable for all designs, we note that EO/OE power does not dominate in any of the designs (see Table III). Even if EO/OE power is removed entirely from the analysis, the results would not change significantly. Further, LumiNOC experiences more EO/OE dynamic power than the other designs due hops through the middle routers.

F. Power Comparison

Table II lists the photonic resource configurations for various PNoC architectures, including one-layer, two-layer, and four-layer configurations of the LumiNOC. While the crossbar architecture of Corona has a high ideal throughput, the excessive number of rings and waveguides results in degraded power efficiency. In order to support equal 20 Tbps aggregate throughput, LumiNOC requires less than $\frac{1}{10}$ the number of rings of FlexiShare and almost the same number of wavelengths. Relative to the Clos architecture, LumiNOC requires around $\frac{4}{7}$ wavelengths, though approximately double number of rings.

The power and efficiency of the network designs is compared in Table III. Where available/applicable, power and throughput numbers for competing PNoC designs are taken from the original papers, otherwise they are calculated as described in Section VI-E. ITP is the ideal throughput of the design, while RTP is the maximum throughput of the design under a uniform random workload as shown in Fig. 8. A 6 × 4 2GHz electrical 2-D mesh [1] was scaled to 8 × 8 nodes operating at 5 GHz, in a 22 nm CMOS process (dynamic power scaled according to the CMOS dynamic power equation and static power linearly with voltage), to compare against the photonic networks.

The table shows that LumiNOC has the highest power efficiency of all designs compared in RTP/Watt, increasing efficiency by ∼40% versus the nearest competitor, Clos [7]. By reducing wavelength multiplexing density, utilizing shorter waveguides, and leveraging the data channels for arbitration, LumiNOC consumes the least ELP among all the compared architectures. A four-layer LumiNOC consumes ∼1/4th the ELP of a competitive Clos architecture, of nearly the same throughput. Corona [13] contains 256 cores with four cores sharing an electrical router, leading to a 64-node photonic crossbar architecture; however, in order to achieve throughput of 160 Gb/s, each channel in Corona consists of 256 wavelengths, 4X the wavelengths in a one-layer LumiNOC. In order to support the highest ideal throughput, Corona consumes the highest electrical router power in the compared PNoCs.

Although FlexiShare attempts to save laser power with its double-round waveguide, which reduces the overall nonresonance ring through-loss (and it is substantially more efficient than Corona), its RTP/W remains somewhat low for several reasons. First, similar to other PNoC architectures, FlexiShare employs a global, long waveguide bus instead of multiple short waveguides for the optical interconnects. The global long waveguides cause relatively large optical loss and overburden the laser. Second, FlexiShare is particularly

impacted by the high number of ring resonators ($N_{ring} = 550K$—Table II), each of these rings need to be heated to maintain its proper frequency response and the power consumption of this heating dominates its RTP/W. Third, the dedicated physical arbitration channel in FlexiShare costs extra optical power. Finally, similar to an SWMR crossbar network (e.g., Firefly [11]), FlexiShare broadcasts to all the other receivers for receiver-side arbitration. Although the authors state that, by only broadcasting the head flit, the cost of broadcast in laser power is avoided, we would argue this would be impractical in practice. Since the turn-around time for changing off-die laser power is so high, a constant laser power is needed to support the worst-case power consumption.

VII. Conclusion

PNoCs are a promising replacement for electrical NoCs in future many-core processors. In this paper, we analyze prior PNoCs, with an eye toward efficient system power utilization and low-latency. The analysis of prior PNoCs reveals that power inefficiencies are mainly caused by channel overprovisioning, unnecessary optical loss due to topology and photonic device layout and power overhead from the separated arbitration channels and networks. LumiNOC addresses these issues by adopting a shared-channel, photonic on-chip network with a novel, in-band arbitration mechanism to efficiently utilize power, achieving a high performance, and scalable interconnect with extremely low latency. Simulations show under synthetic traffic, LumiNOC enjoys 50% lower latency at low loads and ∼40% higher throughput per Watt on synthetic traffic, versus other reported PNoCs. LumiNOC also reduces latencies ∼40% versus an electrical 2-D mesh NoCs on the PARSEC shared-memory, multithreaded benchmark suite.

References

[1] J. Howard *et al.*, "A 48-core IA-32 processor in 45 nm CMOS using on-die message-passing and DVFS for performance and power scaling," *IEEE J. Solid-State Circuits*, vol. 46, no. 1, pp. 173–183, Oct. 2011.

[2] J. Kim, D. Park, T. Theocharides, N. Vijaykrishnan, and C. R. Das, "A low latency router supporting adaptivity for on-chip interconnects," in *Proc. DAC*, Anaheim, CA, USA, pp. 559–564, Jun. 2005.

[3] G. Hendry *et al.*, "Analysis of photonic networks for a chip multiprocessor using scientific applications," in *Proc. 3rd ACM/IEEE NoCS*, San Diego, CA, USA, 2009, pp. 104–113.

[4] A. Shacham, K. Bergman, and L. P. Carloni, "On the design of a photonic network-on-chip," in *Proc. 1st NoCS*, Princeton, NJ, USA, 2007, pp. 53–64.

[5] A. Shacham, K. Bergman, and L. P. Carloni, "Photonic NoC for DMA communications in chip multiprocessors," in *Proc. 15th Annu. IEEE HOTI*, Stanford, CA, USA, 2007, pp. 29–38.

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

[7] A. Joshi *et al.*, "Silicon-photonic Clos networks for global on-chip communication," in *Proc. 3rd ACM/IEEE NoCS*, San Diego, CA, USA, 2009, pp. 124–133.

[8] N. Kirman *et al.*, "Leveraging optical technology in future bus-based chip multiprocessors," in *Proc. 39th Annu. IEEE/ACM MICRO*, Orlando, FL, USA, 2006, pp. 492–503.

[9] A. Krishnamoorthy *et al.*, "Computer systems based on silicon photonic interconnects," *Proc. IEEE*, vol. 97, no. 7, pp. 1337–1361, Jul. 2009.

[10] Y. Pan, J. Kim, and G. Memik, "FlexiShare: Channel sharing for an energy-efficient nanophotonic crossbar," in *Proc. 16th IEEE HPCA*, Bangalore, India, 2010, pp. 1–12.

[11] Y. Pan *et al.*, "Firefly: Illuminating future network-on-chip with nanophotonics," in *Proc. 36th ISCA*, Austin, TX, USA, 2009.

[12] D. Vantrease, N. Binkert, R. Schreiber, and M. H. Lipasti, "Light speed arbitration and flow control for nanophotonic interconnects," in *Proc. 42nd Annu. IEEE/ACM MICRO*, New York, NY, USA, 2009, pp. 304–315.

[13] D. Vantrease *et al.*, "Corona: System implications of emerging nanophotonic technology," in *Proc. 35th ISCA*, Beijing, China, 2008, pp. 153–164.

[14] P. Koka *et al.*, "Silicon-photonic network architectures for scalable, power-efficient multi-chip systems," in *Proc. 37th ISCA*, Saint-Malo, France, 2010, pp. 117–128.

[15] Y. H. Kao and H. J. Chao, "BLOCON: A bufferless photonic Clos network-on-chip architecture," in *Proc. 5th ACM/IEEE NoCS*, Pittsburgh, PA, USA, May 2011, pp. 81–88.

[16] Q. Xu, S. Manipatruni, B. Schmidt, J. Shakya, and M. Lipson, "12.5 Gbit/s carrier-injection-based silicon microring silicon modulators," in *Proc. CLEO*, Baltimore, MD, USA, 2007, pp. 1–2.

[17] I. Young *et al.*, "Optical I/O technology for tera-scale computing," in *Proc. IEEE Int. Solid-State Circuits Conf.*, San Francisco, CA, USA, 2009, pp. 468–469.

[18] M. Reshotko, B. Block, B. Jin, and P. Chang, "Waveguide coupled Ge-on-oxide photodetectors for integrated optical links," in *Proc. 5th IEEE Int. Conf. Group IV Photon.*, Cardiff, U.K., 2008, pp. 182–184.

[19] C. Holzwarth *et al.*, "Localized substrate removal technique enabling strong-confinement microphotonics in bulk Si CMOS processes," in *Proc. CLEO/QELS*, San Jose, CA, USA, 2008, pp. 1–2.

[20] L. C. Kimerling *et al.*, "Electronic-photonic integrated circuits on the CMOS platform," in *Proc. Silicon Photon.*, San Jose, CA, USA, 2006, pp. 6–15.

[21] A. Biberman *et al.*, "Photonic network-on-chip architectures using multilayer deposited silicon materials for high-performance chip multiprocessors," *ACM J. Emerg. Tech. Comput. Syst.*, vol. 7, no. 2, pp. 1305–1315, 2011.

[22] G. Hendry *et al.*, "Time-division-multiplexed arbitration in silicon nanophotonic networks-on-chip for high-performance chip multiprocessors," *J. Parallel Distrib. Comput.*, vol. 71, pp. 641–650, May 2011.

[23] C. Chen and A. Joshi, "Runtime management of laser power in silicon-photonic multibus NoC architecture," *IEEE J. Sel. Topics Quantum Electron.*, vol. 19, no. 2, Article 3700713, Mar.–Apr. 2013.

[24] L. Zhou and A. Kodi, "PROBE: Prediction-based optical bandwidth scaling for energy-efficient NoCs," in *Proc. 7th IEEE/ACM NoCS*, Tempe, AZ, USA, 2013, pp. 1–8.

[25] A. Kodi and R. Morris, "Design of a scalable nanophotonic interconnect for future multicores," in *Proc. 5th ACM/IEEE ANCS*, Princeton, NJ, USA, 2009, pp. 113–122.

[26] R. W. Morris and A. K. Kodi, "Power-efficient and high-performance multi-level hybrid nanophotonic interconnect for multicores," in *Proc. 4th ACM/IEEE NoCS*, Grenoble, France, May 2010, pp. 207–214.

[27] S. Bahirat and S. Pasricha, "UC-PHOTON: A novel hybrid photonic network-on-chip for multiple use-case applications," in *Proc. 11th ISQED*, San Jose, CA, USA, 2010, pp. 721–729.

[28] J. Xue *et al.*, "An intra-chip free-space optical interconnect," in *Proc. 37th ISCA*, New York, NY, USA, 2010, pp. 94–105.

[29] P. Gratz and S. W. Keckler, "Realistic workload characterization and analysis for networks-on-chip design," in *Proc. 4th CMP-MSI*, 2010.

[30] C. Li *et al.*, "A ring-resonator-based silicon photonics transceiver with bias-based wavelength stabilization and adaptive-power-sensitivity receiver," in *Proc. IEEE ISSCC*, San Francisco, CA, USA, Feb. 2013, pp. 124–125.

[31] M. R. T. Tan *et al.*, "Photonic interconnects for computer applications," in *Proc. ACP*, Shanghai, China, 2009, pp. 1–2.

[32] S. Prabhu, B. Grot, P. Gratz, and J. Hu, "Ocin_tsim-DVFS aware simulator for NoCs," in *Proc. SAW*, 2010.

[33] C. Bienia, S. Kumar, J. P. Singh, and K. Li, "The PARSEC benchmark suite: Characterization and architectural implications," in *Proc. 17th PACT*, Toronto, ON, Canada, Oct. 2008.

[34] J. Hestness and S. Keckler. (2010). *Netrace: Dependency-tracking traces for efficient network-on-chip experimentation*. Dept. Comput. Sci., Univ. Texas at Austin, Austin, TX, USA. Tech. Rep. TR-10-11 [Online]. Available: http://www.cs.utexas.edu/~netrace

[35] H. Kim, P. Ghoshal, B. Grot, P. V. Gratz, and D. A. Jimenez, "Reducing network-on-chip energy consumption through spatial locality speculation," in *Proc. 5th ACM/IEEE NoCS*, Pittsburgh, PA, USA, 2011, pp. 233–240.

Cheng Li (S'13) received the M.S. degree in computer engineering from the Illinois Institute of Technology, Chicago, IL, USA, and the Ph.D. degree in electrical engineering from Texas A&M University, College Station, TX, USA, in 2009 and 2013, respectively.

From 2001 to 2007, he was at Alcatel, Shanghai, China, where he was involved in the design of broadband access system. He is currently a Research Scientist with HP Laboratories, Palo Alto, CA, USA. His current research interests include design of high-speed transceiver circuits for optical/photonic links and photonic interconnects for network-on-chip systems.

Mark Browning (S'13) received the B.S. degree in nuclear engineering and the M.Eng. degree in computer engineering from Texas A&M University, College Station, TX, USA, in 2009 and 2013, respectively.

He is currently with Capsher Technology, College Station, writing physics software for extended reach and horizontal drilling engineers. His current research interests include high performance computing and photonic network-on-chip systems.

Paul V. Gratz (S'04–M'09) received the B.S. and M.S. degrees in electrical engineering from the University of Florida, Gainesville, FL, USA, and the Ph.D. degree in electrical and computer engineering from the University of Texas at Austin, Austin, TX, USA, in 1994, 1997, and 2008, respectively.

He is an Assistant Professor with the Department of Electrical and Computer Engineering at Texas A&M University, College Station, TX, USA. From 1997 to 2002, he was a Design Engineer with Intel Corporation, Santa Clara, CA, USA. His current research interests include energy efficient and reliable design in the context of high performance computer architecture, processor memory systems, and on-chip interconnection networks.

Dr. Gratz's paper "B-Fetch: Branch Prediction Directed Prefetching for In-Order Processors" was selected as one of the four Best Papers from IEEE Computer Architecture Letters in 2011. At ASPLOS '09, he co-authored *An Evaluation of the TRIPS Computer System*, which received the Best Paper Award. In 2010, he received the Teaching Excellence Award–Top 5% Award from the Texas A&M University System for a graduate-level advanced computer architecture course he developed.

Samuel Palermo (S'98–M'07) received the B.S. and M.S. degree in electrical engineering from Texas A&M University, College Station, TX, USA, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, USA, in 1997, 1999, and 2007, respectively.

From 1999 to 2000, he was with Texas Instruments, Dallas, TX, USA, where he was involved in the design of mixed-signal integrated circuits for high-speed serial data communication. From 2006 to 2008, he was with Intel Corporation, Hillsboro, OR, USA, where he was involved in high-speed optical and electrical I/O architectures. In 2009, he joined the Electrical and Computer Engineering Department of Texas A&M University, where he is currently an Assistant Professor. His current research interests include high-speed electrical and optical links, high performance clocking circuits, and integrated sensor systems.

Dr. Palermo was a recipient of the 2013 NSF-CAREER Award. He is a member of Eta Kappa Nu. He currently serves as an Associate Editor for IEEE TRANSACTIONS ON CIRCUITS AND SYSTEM–II and has served on the IEEE Circuits and Systems Society Board of Governors from 2011 to 2012. He has co-authored the Jack Raper Award for Outstanding Technology-Directions Paper at the 2009 International Solid-State Circuits Conference.