

PROPEL: Power and Area-Efficient Nanophotonic On-Chip Interconnect Architecture for Multicores

Avinash Kodi, *Member, IEEE*, Randy Morris, and Ahmed Louri, *Senior Member, IEEE*

Abstract—With technology scaling, growing wire delays and excess power dissipation of current metallic interconnects are predicted to significantly limit the performance of Network-on-Chips (NoCs) architectures. Recent research has focused on developing alternate solutions to current metallic interconnects. One potential solution is silicon photonics because of its higher bandwidth, reduced power dissipation, increased wiring simplification and its compatibility with (complementary-metal-oxide semiconductor) CMOS processing. In this paper, we propose PROPEL, a balanced power and area-efficient on-chip photonic interconnect for future multicores. PROPEL overcomes two fundamental issues facing NoCs architectures, namely power dissipation and area overhead, by a combination of multiplexing techniques (wavelength and space) and by exploiting the recent advances silicon photonics design space. Our results indicate that PROPEL is power, cost and area-efficient network when compared to the proposed on-chip optical topologies. Moreover, simulation results on synthetic traffic indicate that PROPEL outperforms both electrical and optical topologies for in-chip interconnects in terms of throughput and power.

I. INTRODUCTION

THE performance of future chip multiprocessors (CMPs) is expected to exponentially grow with technology scaling allowing more processing cores to fit within the same sized die. However, the increased wire delay problem combined with excess power dissipation in the sub-nanometer regime are expected to become fundamental bottlenecks for increased performance [1]. This has changed the design of on-chip wires which have moved from high-speed serial point-to-point ad-hoc wiring to more modular and regular network-on-chips (NoCs) paradigm [2]. Recent research has shown that the power consumption is a major issue facing NoCs [3], [4]. With technology scaling, increased repeater power combined with leakage power will further contribute to the increase in power dissipation. Moreover, electrical interconnect signaling problems, electromagnetic interference (EMI), crosstalk, and clock skew will cumulatively limit the performance and scalability of electrical interconnects [5], [6], [7].

One potential solution is to use optical technology to overcome the wire delay problem and power issues. Optical interconnects offer several well known advantages such as higher spatial and temporal bandwidths, lower cross-talk independent of data rates, higher interconnect densities, better

signal integrity at high frequencies, lower signal attenuation and lower power requirements at high bit rates; making it a solution of choice for long distance communication (LANs, WANs) and even short distances such as board-to-board and chip-to-chip communication [5], [6], [7]. However, the recent surge in photonic components and devices such as silicon-on-insulator (SOI) based micro-ring resonators compatible with complementary-metal-oxide semiconductor (CMOS) technology that offers extraordinary performance in terms of density (small footprint ($\sim 12\mu m$)), power efficiency ($\sim 0.1mW$) [21] and high bandwidth (~ 18 Gbps/channel) [8] characteristics are generating interest for even on-chip interconnects [9], [10], [11], [12], [13], [14], [15].

In this paper, we propose PROPEL - an on-chip nanophotonic interconnect architecture that addresses the power and bandwidth demands of future multicores with acceptable optical hardware complexity. PROPEL uses optical interconnects for long distance inter-router communication and electrical switching within the routers. This reduces the power dissipation on long inter-router links while electrical switching provides flow control to prevent buffer overflow. We leverage nanophotonic components/devices and exploit optical properties such as wavelength division multiplexing (WDM), space division multiplexing (SDM) and wavelength reuse to reduce power dissipation, increase the bandwidth density and reduce area requirements in an efficient manner. Moreover, we present a detailed optical implementation that includes power and area estimates and performance modeling using network simulation on synthetic traffic traces. Our results for 64 cores indicate the following: (1) PROPEL reduces the power consumption by 80% when compared to proposed on-chip electrical networks, (2) PROPEL is comparable and improves performance by more than 10% when compared to on-chip electrical and photonic networks with similar bisection bandwidths, and (3) PROPEL requires the least optical hardware (modulators, photodetectors, waveguides) and has the lowest area overhead as compared to proposed on-chip photonic networks.

Although there has been considerable work in off-chip optical interconnects, only few on-chip optical solutions have been proposed thus far. Collet et.al. [17] have concluded that for technology nodes ranging from $0.7\mu m$ to $0.05\mu m$, on-chip lasers will consume the bulk of the power, hindering the design of on-chip photonic networks. Shacham et.al. [12] have proposed circuit-switched photonic interconnects, where electronic set-up, photonic communication and tear-down have been proposed. A possible issue with this approach is the excess latency for path set-up which is performed using electrical interconnects. Kirman et.al. [9] have proposed an optical bus

Manuscript received May 1, 2009. This work was supported in part by the National Science Foundation grants CCR-0538945 and ECCS-0725765.

Avinash Karanth Kodi and Randy Morris are with the School of Electrical Engineering and Computer Science, Ohio University, Athens, OH 45701, USA e-mail: kodi@ohio.edu, randy.morris@eecs.ohio.edu.

Ahmed Louri is with the Department of Electrical and Computer Engineering, University of Arizona, Tucson, AZ 85721, USA e-mail: louri@ece.arizona.edu

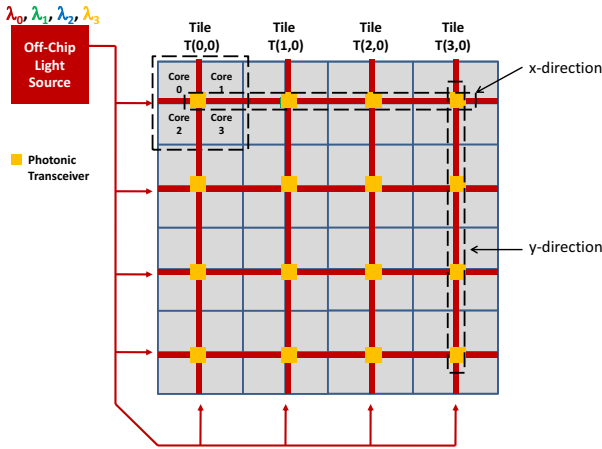


Fig. 1. Proposed layout of PROPEL architecture for 64 tile architecture. Four tiles are combined into a super-tile.

for intra-chip processor-L2 cache interconnect for 64 cores grouped into 4 sets. However, this design cannot be scaled and bus contention will increase with more cores unless more bandwidth (wavelengths) can be incorporated. Batten et.al. [10] have proposed a DRAM-processor interconnect using an opto-electronic global crossbar with electronic arbitration and photonic switching devices using double ring resonators. More recently, CORONA, a 3D-stacked, 256-core, on-chip fully connected optical crossbar with token based optical arbitration has been proposed [11]. This design scales as $O(N^2)$ which increases the cost and complexity of the network.

II. PROPEL: ARCHITECTURE AND IMPLEMENTATION

In this section, we describe the proposed architecture, PROPEL and its routing and wavelength assignment (RWA). We choose 22nm technology node for our work as prior research has shown that optics becomes advantageous at this feature size [9], [10], [11], [12], [13], [14], [15]. In PROPEL, we combine optical transceivers and electronic switches, as shown in Figure 1. The proposed off-chip broadband light source will generate W_N wavelengths, $\Lambda = \lambda_0, \lambda_1, \lambda_2, \lambda_3, \dots, \lambda_{W_N-1}$. By transmitting the continuous off-chip carrier signal in both x- and y-directions simultaneously, we modulate the signals at the optical transmitters. Figure 1 shows 4 cores and a shared L2 cache combined together to form a *tile*. This grouping reduces the cost of the interconnect as every core does not require lasers attached and more importantly, facilitates local communication through cheaper electronic switching. There are a total of M tiles in the x-direction and N tiles in the y-direction, for a total of $4 \times M \times N$ cores. Each tile is represented by $T(m, n)$ where $0 \leq m \leq M-1$ and $0 \leq n \leq N-1$. With $M = N = 4$, PROPEL can be designed for 64 cores and with $M = N = 8$, PROPEL can be designed for 256 cores. Each tile consists of dual-set (x and y) photonic transceivers and an electronic switch. It takes a maximum of 2 hops to for any tile to communicate with other tile in the network. This is a significant advantage over many electrical networks.

A. Intra-Tile Interconnect

Each tile consists of a set of modulators (transmitters) and a set of photodetectors (receivers) for both x and y directions. With a shared-L2 cache for the four cores, we will need a 7×7 crossbar switch for 64 cores; three for x-direction, three for y-direction and one for the shared-L2. With a private-L2 cache, the crossbar size increases to 10×10 crossbar. Research has shown that high-radix routers could monotonically reduce the overall cost of network (power, area and latency) [16]. In addition, these crossbar and buffer elements are designed in lower metal layers leading to lower power and area overhead with technology scaling. The packet, consisting of several flits, undergoes the usual router stages of RC (routing computation), VC (virtual channel) allocation, SA (switch allocation) and ST (switch traversal). We allow flit interleaving in the electrical domain (intra-tile) and packet interleaving in the optical domain to reduce the contention and processing overhead at the receiver as the optical link rate may not match the electrical router data rate. Flow control signaling is tied to packet flows and not individual flits. This requires additional buffering at the transmitter and receiver ports to hold entire packets and to overcome round-trip control flow information. We use on/off signaling implemented using electrical interconnects based on receiver buffer thresholds.

B. Inter-Tile Interconnect

We adopt dimension-order routing (DOR) for inter tile communication. The traffic first flows in the x-direction to an intermediate tile and then flows in the y-direction to reach the destination. We explain the routing in a single dimension (x) involving four tiles and similar design can be extended to y-dimension. Figure 2 shows tiles 0 to 3 arranged along the x-direction. Every tile modulates the same wavelength into a different waveguide. Each destination tile is associated with a waveguide called as the *home channel*. For example, tile $T(0,0)$ has four modulators (ring resonators), all of which are resonant with the wavelength λ_0 . Three λ_0 transmissions from tile $T(0,0)$ are used to communicate with the other three tiles $T(1,0)$, $T(2,0)$ and $T(3,0)$ on their home channel waveguides. The fourth resonant wavelength will be used to communicate with the memory bank. As shown in Figure 2, the home channel for tile $T(0,0)$ consists of four wavelengths, $\Lambda = \lambda_0 + \lambda_1 + \lambda_2 + \lambda_3$ transmitted by tile $T(0,0)$, $T(1,0)$, $T(2,0)$ and $T(3,0)$ respectively. The wavelength selective filters located at tile $T(0,0)$ will demultiplex all the wavelengths, except for λ_0 which originates from itself and is intended for the memory. Similarly, the wavelengths, λ_0 from tile $T(0,0)$, λ_1 from tile $T(1,0)$, λ_2 from tile $T(2,0)$ and λ_3 from tile $T(3,0)$ are combined and these are used to access the memory banks. These are also the same wavelength at which the above tiles will receive data from the memory module. Our goal is to provide a scalable bandwidth to the memory similar to inter tile communication. While we propose to access off-chip memory using photonic network, these functionalities have yet to be implemented in our simulations. Similar wavelength assignment is replicated even in the y-direction for inter-tile communication.

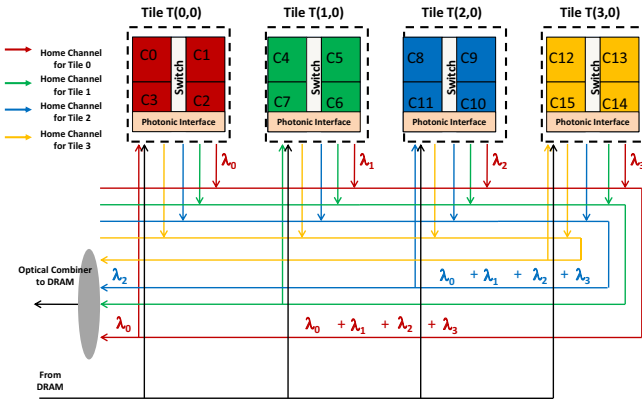


Fig. 2. The routing and wavelength assignment proposed for PROPEL for the x dimension.

The RWA algorithm designed for inter tile communication involves selective merging of same wavelengths from source tiles into separate home channels for destination tiles. This design maximizes the bandwidth via WDM and re-uses the same wavelengths on different waveguides via SDM. The electronic switching performs localized arbitration for the output optical transmitters within each tile. As the wavelength for the destination tile is fixed, there is no more contention once the local electronic switching is completed. The effective bandwidth of the nanophotonic interconnect, $B = W_N \times Wg_N \times B_R$, where W_N is the number of wavelengths, Wg_N is the number of waveguides and B_R is the effective bit rate of the channel. With $W_N = 1$, $Wg_N = 1$ and $B_R = 10$ Gbps, we obtain 10 Gbps of inter-tile communication. It could be possible to increase the bit rates beyond 10 Gbps as has been reported [25], however we may need additional equalization circuits that could consume substantial area on the chip. Another approach to increase the bandwidth is to increase the wavelengths or the waveguides. Increasing the waveguides increases the area occupied by the channels and the transmitter/receiver circuitry, where as increasing the wavelengths increases only the transmitter and receiver circuitry. As prior work has shown the feasibility of using 64 wavelengths, we assume similar number of wavelengths for our approach [11], [10]. As we have four tiles, we divide 64 wavelengths among four tiles to provide 160 Gbps of inter tile communication bandwidth.

Figure 3 shows a possible implementation of PROPEL. The off-chip broadband signal is split for x -dimension communication, y -dimension communication and DRAM memory banks. Every tile uses the same set of wavelength to communicate with row/column tiles. Figure 4 shows the implementation of one row of tiles (0 to 3) along the x dimension. The top half of the figure shows the transmitters which are color coded (purple: λ_{0-15} , orange: λ_{16-31} , green: λ_{32-47} and blue: λ_{48-63}). The bottom half shows the filters (ring resonators) that are used to detect the signal. For example, Tile 0 will use wavelengths λ_{0-15} to communicate with tiles 1, 2 and 3 on different waveguides which correspond to different home

channels. Similarly, Tiles 1, 2 and 3 will use wavelengths λ_{16-31} , λ_{32-47} and λ_{48-63} to communicate with other tiles respectively. Tile 0 will use wavelengths λ_{0-15} , tile 1 will use λ_{16-31} , tile 2 will use λ_{32-47} and tile 3 will use λ_{48-63} to communicate with memory. Therefore, the wavelengths associated with a tile (For example, Tile 0 is associated with λ_{0-15} , Tile 1 is associated with λ_{16-31}) are modulated at the transmitter, however they are not detected at the tile's receiver. All the home channel waveguides are combined and the multiplexed signal is transmitted to the memory banks where the individual signals are detected by the corresponding memory banks.

III. SILICON NANOPHOTONIC DEVICES AND COMPONENTS

In this section, we briefly describe the silicon nanophotonic interconnects and components required to design PROPEL. Nanophotonic interconnect will require (i) lasers to generate the carrier signal, (ii) modulators and drivers to encode the data, (iii) medium (waveguides, fibers, free space) for signal propagation, (iv) photodetectors to detect light and (v) back-end signal processing (transimpedance amplifiers (TIA), voltage amplifiers, clock and data recovery) to recover the transmitted bit.

Transmitters: Direct modulation of the carrier signal is possible using vertical-cavity surface emitting lasers (VCSELs), however these devices are not suitable for on-chip interconnects [19], [18]. Indirect modulation with an external laser will need an on-chip modulator. This approach will have, the advantage that the power of the laser is not included in the total power budget of the chip since it will be external to the chip. Two CMOS-compatible modulators proposed recently are Mach-Zehnder (MZ) modulator and the Microring resonator. Microring resonator will couple light through it only if it satisfies the relation: $\lambda_0 \times m = n_{eff} \times 2\pi R$, where R is the radius of the microring resonator, n_{eff} is the effective refractive index, m an integer and λ_0 is the resonant wavelength [20], [21]. Recent designs have shown ring resonators can modulate at 12.5 Gbps with a modulator delay of 80 ps [21]. Due to smaller footprint ($\sim 10\mu m$), lower power (0.1 mW) and high-speed modulation, ring resonators are more preferred over MZ modulators [23]. Although ring resonators show great promise, they are very sensitive to temperature due to the thermo-optic coefficient (TOC) of silicon ($\Delta n/\Delta T = 1.86 \times 10^{-4} K^{-1}$, where Δn is the change in the refractive index and ΔT is the change in temperature) which leads to a resonance shift of $\Delta\lambda_0 \sim 0.11$ nm/K [22]. Recent work has shown that temperature variations can be mitigated by adjusting the bias current (upto 15K) [22]. The pre-driver electrical circuit is a chain of tapered inverters used to drive the modulator's capacitive load ($\sim 60fF$).

Waveguides: CMOS compatible waveguides can be made of high index cores such as Si (3.5) or low index cores such as polymers (1.5). High index core offers a smaller waveguide pitch where as low index core offers a lower propagation delay. Silicon waveguides, which have a smaller pitch of 5.5 μm , a lower propagation time of 10.45 ps/mm and a signal

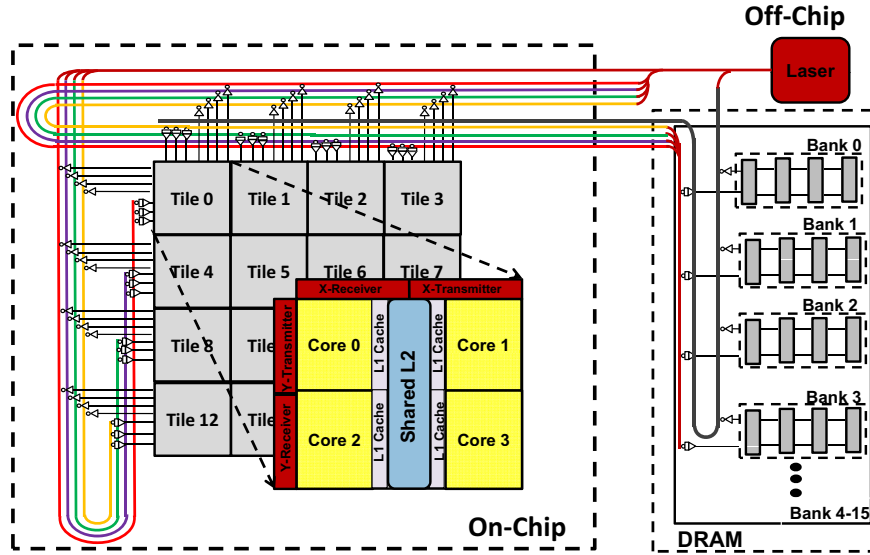


Fig. 3. The proposed PROPEL implementation consisting of x and y dimension connectivity. The off-chip broadband signal is split among the x , y and DRAM modules.

attenuation of 1.3 dB/cm is chosen due to ease of integration with other on-chip photonic components [9], [15]. Recent research into estimating the number of wavelengths that can be multiplexed onto the same waveguide have shown that with single-ring filters and 2.7 GHz free spectral range (FSR), we may have upto 12 wavelengths, which provides around 200 GHz channel spacing between adjacent wavelengths. With double-rings which improves the filtering, it maybe possible to pack 64 wavelengths with tighter 60 GHz spacing [10]. Therefore, in our design we utilize 64 wavelengths and extensively reuse these wavelengths to achieve scalable bandwidth.

Receivers: The optical receiver is composed of light detection (photodiodes), amplification (TIA, voltage amplifier) and clock and data recovery. With the need to absorb light and convert into electrical pulses, Germanium is being used for two reasons: It has significant photo-absorption between 1.1 μm and 1.5 μm and is already being used in CMOS processes [11]. Recent receiver design have used Ge-on-silicon-on-insulator (Ge-on-SOI) photodiodes along with Si CMOS amplifiers to operate at 1.1 V consuming only 1.1 $mW/Gbps$ of power with an area of $175 \times 150 \mu m^2$ with a delay of 40 ps. The SiGe photodiode is CMOS compatible, has a high responsivity (0.56 A/W), detects signal with a bit error rate (BER) of 10^{-12} and is sensitive to optical wavelengths at (850nm, 1350nm, 1550nm)[24]. This receiver power of 1.1 $mW/Gbps$ adopted by this work is more than what has been proposed for CORONA (0.078 $mW/Gbps$) [11], but is closer to estimates from [12] of 0.8 pJ/bit (equal to 0.8 $mW/Gbps$). The receiver power from [11] is lower as they consider a receiverless design where germanium based wavelength-selective detector (ring resonator) is chosen with low detector capacitance of 1 fF. It should also be noted that if micron scale photodetectors are considered that generate few fF of load capacitance, then receiverless approach could become feasible [5]. However, in this paper, we adopt an existing receiver design and as the

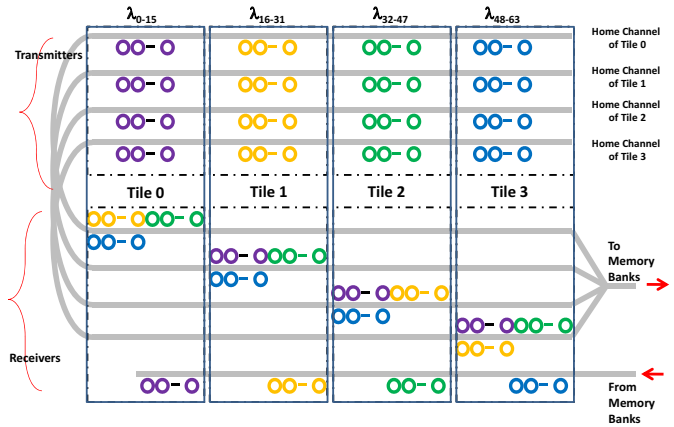


Fig. 4. The top half shows the transmitters (modulators) and the bottom half shows the receivers (filters). Every tile is associated with similar numbered wavelengths for transmission. For example, Tile 0 is associated with λ_{0-15} which is used to transmit to tiles 1, 2 and 3. Moreover, this wavelength set is used for memory communication.

receiver involves analog components, their scaling rules are slower than VLSI technologies [18] and therefore, we consider a higher power consumption for optical interconnects.

IV. PERFORMANCE EVALUATION

In this section, we compare the area and optical hardware complexity of PROPEL to proposed photonic interconnects such as the Shared-Bus from Cornell [9], Processor-DRAM interconnect from MIT [10], CORONA from HP [11] and the Circuit-Switch interconnect from Columbia[12]. We further model and simulate PROPEL and compare to both electrical networks such as the Mesh, Concentrated Mesh (CMesh) and

Flattened-Butterfly (FB) [16] and optical networks ([11], [12]) for synthetic traffic traces. The Processor-DRAM [10] was not chosen for performance comparison as they are designed for core-memory interconnect, where as PROPEL is designed for inter-core communication. In what follows, we briefly provide power and area estimates for NoC link and router. Then we compare 64 and 256 core versions of PROPEL with competing electrical and optical networks based on optical hardware required and provide simulation results.

A. Electrical Power and Area Estimations

For electrical interconnects, we consider wires implemented in semi-global metal layers for inter-router links. The wire capacitances, resistances and device parameters were obtained from International Roadmap for Semiconductors and Berkeley Predictive Technology Models. The power per segment of a repeater-inserted wire is given by $P_{segment} = P_{dynamic} + P_{leakage} + P_{short-ckt}$ where $P_{dynamic}$ is the switching power, $P_{leakage}$ is the power due to the subthreshold leakage current and $P_{short-ckt}$ is the power due to the short-circuit current [26], [27]. At 90nm technology node, we obtain a link power of 10.27 mW for 1 GHz clock and a V_{dd} of 1.2 V for a flit width of 128 bits [26] by considering a power-optimal repeater insertion. At 22nm, ITRS projects the clock to be 9 GHz. For a flit size of 128 bits, the power dissipation will be 198 mW. To reduce the power dissipation at future technology nodes, we reduce the network frequency to 1 GHz and reduce the power consumption to 22 mW. The area consumed by the wires is determined as $Area_{wires} = N_W \times p_w$ where N_W is the number of wires per link (the bit-width of the link) and p_w is the wire pitch at the given technology ($\sim 0.0422 \text{ mm}^2$) [26].

For on-chip (SRAM cell-array) buffers, the dynamic power consumed is the sum of the power expended in writing a flit into the buffer and the power consumed to read out the flit from the buffer [28]. At the 90 nm technology considered, an SRAM cell has an estimated width of $1.16 \mu\text{m}$ and a height of $0.87 \mu\text{m}$ [29], giving an area of $1.0092 \mu\text{m}^2$. Results from Intel's 90 nm technology [30] also indicate an area of $1 \mu\text{m}^2$ for the SRAM cell. We then determined the power and area values across technologies by scaling the parameters. The power dissipated by a SRAM buffer for 128 bits at 1 GHz clock is 15.8 mW [26]. At 22 nm, we estimate the buffer power to be 4.03 mW and occupies an area of $185 \mu\text{m}^2$. A 5×5 matrix crossbar with tri-state buffer connectors [28] is considered for the regular NoC design. The area of the crossbar is estimated by the number of input/output signals that it should accommodate. For 90nm, the power consumed by 5×5 crossbar is 3.6 mW and this scales to 0.65mW at 22nm.

We compare mesh and PROPEL in terms of power consumed as both networks have buffer read/write, crossbar traversal and link traversal. Buffer power is consumed for both networks when a flit is read and written (4.03 mW). For crossbar traversal, the power dissipation in mesh and PROPEL are different due to the difference in crossbar sizes. In a mesh network, the crossbar power is 0.65 mW; for PROPEL

network, the crossbar power is 0.8 mW. The total power dissipation for a flit traversing across one link and router in a mesh network is estimated to be 26.68 mW, and the power dissipation for PROPEL is 6.13 mW. This results in a substantial (5X) reduction in power consumption.

B. Area and Optical Hardware Complexity Analysis

In this subsection, we analytically compare the optical hardware complexity in terms of wavelengths, optical components (splitters/couplers, ring resonators), total optical power budget, opto-electronic power dissipation and opto-electronic area requirements. For all networks, we assume an off-chip laser source and the following losses consistent across all networks [9], [10], [11], [12], [13]: a star splitter loss (L_S) of $-3 \times (\log_2 N)$ where N is the number of times the waveguide is split, a splitter/coupler loss (L_C) of -3 dB (50% loss of signal), off-chip laser-to-fiber coupling loss (L_{LF}) of -0.5 dB, off-chip to on-chip fiber-to-waveguide coupling loss (L_{FW}) of -2 dB, waveguide loss (L_W) of -1.3 dB/cm, bending loss (L_B) of -1 dB, a modulator traversal loss (L_M) of -1 dB, a waveguide crossover loss (L_X) of -0.05 dB and a waveguide-to-receiver loss (L_{WR}) of -0.5 dB.

PROPEL: The total area required for implementing PROPEL is split into two layers, an optical layer consisting of the modulator, waveguides, and photodetectors; and an electrical layer consisting of the pre-drivers, routers, and receiver circuitry. For the optical layer, we need a total of 3,072 ring resonators (192 per tile, 96 each for x- and y-directions), 32 silicon waveguides (16 each for x and y-directions) and 1,536 photodetectors (96 per tile). For each set of modulators and detectors on a tile, the total area overhead is approximately 0.0145 mm^2 per direction, giving a total area overhead of 0.029 mm^2 per tiles. Additional area results from the 32 optical waveguides that need to circulate around the chip which is approximated to be 5 cms. This gives PROPEL an approximate optical area overhead of 17 mm^2 . For the electrical layer, the pre-driver design at 22nm yields an area of $0.3 \mu^2\text{m}$. The receiver is composed of transimpedance amplifier, voltage amplifier and clock and data recovery circuit, which will occupy 0.026250 mm^2 at 90nm. We conservatively assume similar area requirements at 22nm, though this value will decrease. For a 7×7 crossbar proposed in 22nm, this will occupy 0.18 mm^2 . We assume a flit size of 128 bits, 4 virtual channels and 4 flit buffers per virtual channel for the buffer design and this will occupy 0.022 mm^2 per router. Therefore, the total electrical overhead is estimated to be 50 mm^2 . Next, we estimate the overall optical power loss in the network. The overall optical power loss is given by $L_S + L_{LF} + L_{FW} + 2 \times L_M + L_{WR} + 4 \times L_B + 32 \times L_X + L_W$, where L_S will be -15 dB ($= -3 \times \log_2 32$) and L_W will be -6.5 dB. This makes the total optical loss for PROPEL to be -32.1 dB.

CORONA [11]: As CORONA is a crossbar, we scale to 64 nodes and compare with PROPEL in Table 1. For the optical layer, there are a total of 72,192 ring resonators, 99 waveguides, and 7,424 photodetectors giving Corona an approximate area overhead of 64.6 mm^2 . The electrical area consists of simply the receiver circuit. The electrical area

TABLE I
OPTICAL HARDWARE COMPLEXITY COMPARISON BETWEEN VARIOUS
ON-CHIP PHOTONIC ARCHITECTURES FOR 64 CORES

| | Circuit Switch [12] | Shared Bus [9] | CORONA [11] | PROPEL |
|---------------------------|------------------------|-------------------|----------------|--------|
| Wavelengths | 24 | 4 | 64 | 64 |
| Waveguides | 64 | 168 | 99 | 32 |
| Ring Resonators | 16,576 | 2,688 | 72,192 | 3,072 |
| Power Loss (dB) | 37 | 39.2 | 49.2 | 32.1 |
| Optical Area mm^2 | 16 | 46 | 64.6 | 17 |
| Electrical Area mm^2 | 60 | 55 | 195 | 50 |
| Photodetectors | 1,536 | 2,016 | 7,424 | 1,536 |

overhead for 7,424 receivers is approximately 195 mm^2 . The overall power loss is calculated in a similar manner as PROPEL and is determined to be -49.6 dB

Circuit-Switched [12]: Circuit-switch is an all optical network that uses a high speed electrical network to setup the optical path. For the optical layer, there are a total of 16,576 ring resonators, 64 waveguides, and 1,536 photodetectors which approximately occupy 16 mm^2 . Assuming a flit size of 32-bits for circuit setup, we assume a 1,088 32-bit 4×4 crossbar. The area overhead for the electrical setup is estimated to be 22 mm^2 , which leads to a total electrical area overhead to be 60 mm^2 including receiver and transmitter circuitry. In calculating the power loss in the circuit switch, the worst path length needs to be considered. The overall power loss is calculated in a similar manner as PROPEL and is approximately 37 dB.

Shared Bus [9]: For the optical layer, there are a total of 2,688 ring resonator, 168 waveguides, and 2,016 photodetectors that occupy an approximate area of 46 mm^2 . The electrical area consists of the receiver circuitry and the 5×5 electrical crossbar. The electrical area overhead for 2,016 receivers and 16 5×5 electrical crossbar is approximately 55 mm^2 . The overall power loss is determined to be -39.2 dB.

Table 1 shows various optical components and losses of various photonic interconnects for scaled versions of 64 cores. Shared-bus was originally designed with four wavelengths and increasing the number of wavelengths will change most parameters. As will be explained later, shared-bus architecture is limited by the crossbar throughput and therefore, any increase in the wavelength will not change the performance. As can be seen, PROPEL reduces the optical hardware complexity while requiring the least number of ring resonators and has the lowest optical power loss. Moreover, PROPEL can be designed with minimum optical and electrical area overhead. PROPEL requires $3.8 \times$ less optical area than CORONA and $2.7 \times$ less optical area than Shared-Bus and is comparable to circuit-switch architecture. PROPEL requires $3.8 \times$ lesser electrical area than CORONA with the assumption of the specific electrical receiver circuitry adopted for this design [24].

C. Throughput, Latency and Power

In this subsection, we first describe our simulation methodology and present our results on synthetic traffic traces. We simulated PROPEL on several traces including Uniform Random, and permutation patterns, such as Bit-Reversal, Butterfly, Matrix Transpose, Complement and Perfect Shuffle. A cycle accurate simulator was used to evaluate the performance of PROPEL and the above mentioned networks. We assumed a packet size of 4 flits with the flit size of 128 bits. Identical bisectional bandwidth and buffering for each electrical network was considered. For FB, we assumed delays of 1, 2 and 3 cycles to communicate over 1, 2 and 3 routers respectively to account for longer links in a single dimension. For CORONA and PROPEL, we simulate L2 caches with a crossbar connecting the cores to the optical transmitters to improve performance. CORONA provides a channel bandwidth of 2.56 Tbps and bisection bandwidth of 40.96 Tbps. PROPEL provides a channel bandwidth of 160 Gbps and a bisection bandwidth of 5.12 Tbps. To maintain similar bandwidths for circuit-switch, we assumed 240 Gbps of optical channel rate.

Synthetic Traffic: Figure 4 shows the throughput and average network latency per packet for uniform traffic. From Figure 4(a), we can see that CORONA outperforms all network due to its enormous channel bandwidth of 2.56 Tbps as compared to PROPEL which provides only 0.16 Tbps, a 16X reduction. However, PROPEL offers only 15% lower throughput than CORONA with a significant reduction in optical hardware and network cost. PROPEL outperforms Mesh and FB by 33% and 10% respectively with identical bisection bandwidths. While PROPEL outperforms electrical network, the real advantage of PROPEL can be seen in terms of power dissipation as shown in the next plot. PROPEL is better than circuit-switch traffic by over 50% for uniform traffic. The two-fold reason is that we are considering short packets and the traffic is random. This creates more contention in circuit-switch network and it will not be able to amortize the cost of setting-up the circuit. Shared-bus network saturates early due to the traffic build-up at the two overlapped switches at the entry and exit points from the optical network. There are four cores connected to the first switch, and these four sets of four cores connected to the second switch before entering the optical network. All 16 cores will contend to enter into the shared bus using the two level switches. The network load is significant to saturate the bus even at very low loads. Figure 4(b) shows the average network latency for 64 cores.

The throughput for all traffic traces for various networks are shown in Figure 5(a). In the figure, the results are normalized relative to the mesh network, showing the increase in throughput of each network relative to the mesh. PROPEL's performance is comparable and even better than most electrical networks and is slightly lower than CORONA. Circuit-switch performs better for Butterfly and Perfect Shuffle traffic traces as these communication traces involve less contention. Shared-bus also improves performance with synthetic traffic traces as select source cores communicate with select destination cores which reduces the random nature of uniform traffic traces. As PROPEL reduces the cost of the network, it trades-

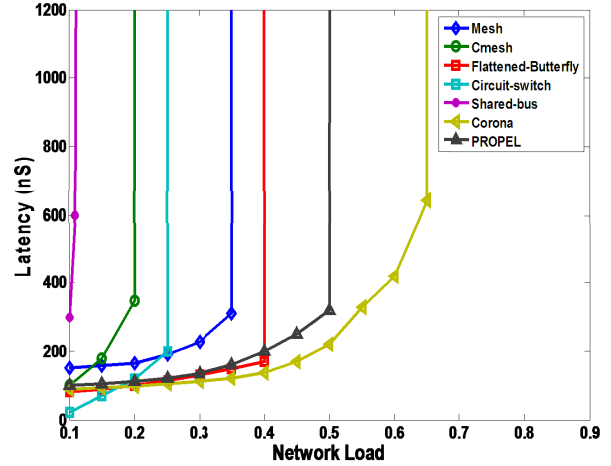
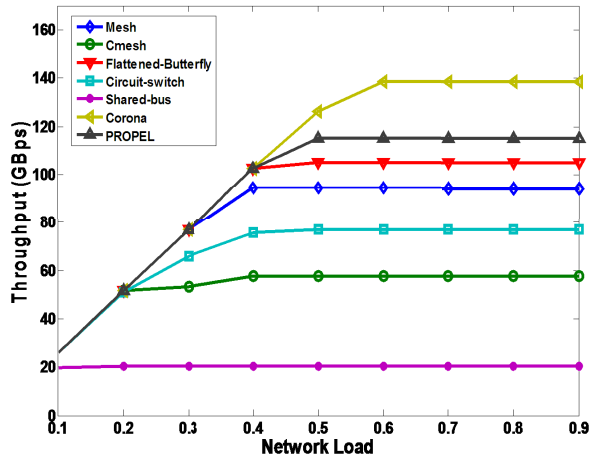


Fig. 5. Simulation results showing (a) throughput and (b) network latency for uniform traffic for various nanophotonic and electrical interconnects with 64 cores.

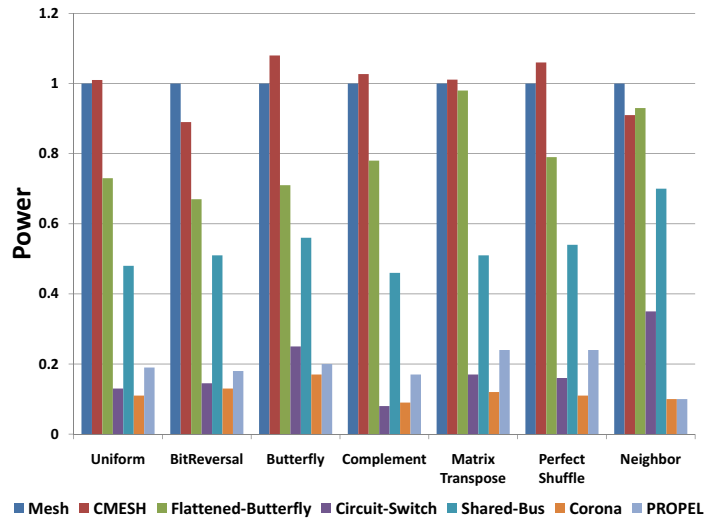
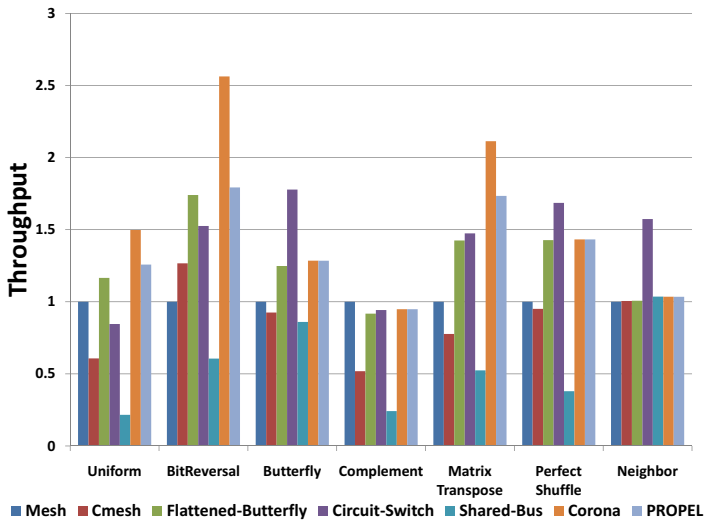


Fig. 6. Simulation results showing (a) saturation throughput and (b) power dissipation for different synthetic traffic patterns for 64 cores.

off performance with network cost and area. Figure 5(b) shows the normalized power dissipation. In the figure, the results are normalized relative to the mesh. As seen, PROPEL reduces the power by 5X when compared to mesh network. In fact, all nanophotonic networks reduce the power dissipated when compared to electrical networks with reduced frequency. Increasing the frequency will increase the power dissipation for electrical networks and opto-electronic networks such as PROPEL and Shared-Bus. While CORONA and circuit-switch have least power consumption, we do not take into account the buffering required at the end-points. As these are fully optical networks, buffers will be required at the end-points for receiving and transmitting the packets. This is accounted in PROPEL as backpressure from the channel allows more packets to be in the network. Circuit-switch has a higher power consumption due to electrical setup which increases with contention.

V. CONCLUSION

In this paper, we tackled the problem of scalable opto-electronic on-chip interconnects to solve the bandwidth and power dissipation problems of future NoCs. We proposed PROPEL architecture for 22nm technology node. The optical complexity analysis clearly showed that PROPEL is significantly cost-efficient than previously proposed on-chip photonic interconnects while delivering comparable performance at reduced power dissipation. Moreover, this architecture has the desirable scalable features identical to a mesh architecture which can be scaled in two dimensions, and provides fault-tolerance due to multi-path connectivity.

REFERENCES

- [1] R. Ho, K. W. Mai, and M. A. Horowitz, "The future of wires," *Proceedings of the IEEE*, vol. 89, pp. 490–504, April 2001.
- [2] L. Benini and G. D. Micheli, "Networks on chips: A new soc paradigm," *IEEE Computer*, vol. 35, pp. 70–78, 2002.

- [3] J. D. Owens, W. J. Dally, R. Ho, D. N. Jayasimha, S. W. Keckler, and L. S. Peh, "Research challenges for on-chip interconnection networks," *IEEE Micro*, vol. 27, no. 5, pp. 96–108, September–October 2007.
- [4] Yatin Hoskote, Sriram Vangal, Arvind Singh, Nitin Borkar, and Shekhar Borkar, "A 5-ghz mesh interconnect for a teraflops processor," *IEEE Micro*, pp. 51–61, September/October 2007.
- [5] David A. B. Miller, "Device requirements for optical interconnects to silicon chips," to appear in *Proceedings of the IEEE, Special Issue on Silicon Photonics*, 2009.
- [6] A. F. Benner, M. Ignatowski, J. A. Kash, D. M. Kuchta, and M. B. Ritter, "Exploitation of optical interconnects in future server architectures," *IBM Journal of Research and Development*, vol. 49, no. 4/5, pp. 755–775, September 2005.
- [7] Raymond G. Beausoleil, Philip J. Kuekes, Gregory S. Snider, Shih-Yuan Wang, and R. Stanley Williams, "Nanoelectronic and nanophotonic interconnect," *Proceedings of the IEEE*, vol. 96, no. 2, pp. 230–247, February 2008.
- [8] S. Manipatruni, Q. Xu, B.S. Schmidt, J. Shakya, and M. Lipson, "High speed carrier injection 18 gb/s silicon micro-ring electro-optic modulator," in *IEEE/LEOS 2007*, 21–25 Oct 2007.
- [9] N. Kirman, M. Kirman, R.K. Dokania, J. Martnez, A.B. Apsel, M.A. Watkins, and D.H. Albonesi, "Leveraging optical technology in future bus-based chip multiprocessors," in *Proceedings of the 39th International Symposium on Microarchitecture*, December 2006.
- [10] Christopher Batten and et.al., "Building manycore processor-to-dram networks with monolithic silicon photonics," in *Proceedings of the 16th Annual Symposium on High-Performance Interconnects*, August 27–28 2008.
- [11] Dana Vantrease and et.al., "Corona: System implications of emerging nanophotonic technology," in *Proceedings of the 35th International Symposium on Computer Architecture*, June 2008.
- [12] Assaf Shacham, Keren Bergman, and Luca P. Carloni, "Photonic networks-on-chip for future generations of chip multiprocessors," in *IEEE Transactions on Computers*, September 2008, pp. 1246–1260.
- [13] Ian O'Connor, "Optical solutions for system level interconnects," in *Proc. 2004 International Workshop on System Level Interconnect Prediction (SLIP'04)*, Paris, France, February 2004, pp. 290–294.
- [14] Cary Gunn, "Cmos photonics for high speed interconnects," *IEEE Photonics Technology Letters*, vol. 26, pp. 58–66, 2006.
- [15] Guoqing Chen, Hui Chen, Mikhail Haurylau, Nicholas Nelson, David Albonesi, Philippe Fauchet, and Eby Friedman, "On-chip optical interconnect roadmap: Challenges and critical roadmap," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 12, no. 6, pp. 1699–1705, November/December 2006.
- [16] John Kim, William J. Dally, and Dennis Abts, "Flattened butterfly: Cost-efficient topology for high-radix networks," in *Proceedings of 34th Annual International Symposium on Computer Architecture (ISCA)*, June 2007, pp. 126 – 137.
- [17] J.H.Collet, F.Caignet, F.Sellaye, and D.Litaize, "Performance constraints for on-chip optical interconnects," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 9, pp. 425–432, March/April 2003.
- [18] Samuel Palermo, Azita Emami-Neyestanak, and Mark Horowitz, "A 90 nm cmos 16 gb/s transceiver for optical interconnects," *IEEE Journal of Solid-State Circuits*, vol. 43, pp. 1235–1246, May 2008.
- [19] Christian Kromer, Gion Sialm, Christoph Berger, Thomas Morf, Martins L. Schmartz, Frank Ellinger, Daniel Erni, Gian-Luca Bona, and Heinz Jackel, "A 100-mw 4x10gb/s transceiver in 80-nm cmos for high-density optical interconnects," in *2005 IEEE International Solid State Circuits Conference*, February 2005, vol. 1, pp. 334–602.
- [20] Q. Xu, B. Schmidt, S. Pradhan, and M. Lipson, "Micrometre-scale silicon electro-optic modulator," *Nature Letters*, vol. 435, pp. 325–327, 2005.
- [21] Qianfan Xu, Sasikanth Manipatruni, Brad Schmidt, Jagat Shakya, and Michal Lipson, "12.5 gbit/s carrier-injection-based silicon micro-ring silicon modulators," *Optics Express: The International Electronic Journal of Optics*, vol. 15, January 2007.
- [22] S. Manipatruni, R. Dokania, B. Schmidt, N. Droz, C. Poitras, A. Apsel, and M. Lipson, "Wide temperature range operation of micron-scale silicon electro-optic modulators," *Optics Letters*, vol. 33, no. 19, September–October 2008.
- [23] William M. J. Green, Michael J. Rooks, Lidija Sekarie, and Yuri Vlasov, "Ultra-compact, low rf power, 10 gb/s silicon mach-zehnder modulator," *Optics Express*, pp. 17106–17113, December 2007.
- [24] Steven J. Koester, Clint L. Schow, Laurent Schares, and Gabriel Dehlinger, "Ge-on-soi-detector/si-cmos-amplifier receivers for high-performance optical-communication applications," *Journal of Lightwave Technology*, vol. 25, pp. 46–57, January 2007.
- [25] N. Sherwood-Droz, H. Wang, L. Chen, B.G. Lee, A. Biberman, K. Bergman, and M. Lipson, "Optical 4x4 hitless silicon router for optical networks-on-chip (noc)," *Optics Express*, vol. 16, no. 20, Sept 2007.
- [26] Ashwini Sarathy Avinash K Kodi and Ahmed Louri, "ideal: Inter-router dual-function energy- and area-efficient links for network-on-chip (noc)," in *Proceedings of the 35th International Symposium on Computer Architecture (ISCA'08)*, Beijing, China, June 2008, pp. 241–250.
- [27] K. Banerjee and A. Mehrotra, "A power-optimal repeater insertion methodology for global interconnects in nanometer designs," *IEEE Transactions on Electron Devices*, vol. 49, no. 11, pp. 2001–2007, November 2002.
- [28] H. S. Wang, X. Zhu, L. S. Peh, and S. Malik, "Orion: A power-performance simulator for interconnection networks," in *Proceedings of the 35th Annual ACM/IEEE International Symposium on Microarchitecture*, Istanbul, Turkey, November 18–22 2002, pp. 294–305.
- [29] A. Y. Zeng, K. Rose, and R. J. Gutmann, "Cache array architecture optimization at deep submicron technologies," in *Proceedings of the IEEE International Conference on Computer Design: VLSI in Computers and Processors*, San Jose, CA, USA, October 11–13, 2004, pp. 320–325.
- [30] S. Thompson and et.al., "A 90 nm logic technology featuring 50 nm strained silicon channel transistors, 7 layers of cu interconnects, low-k ild, and 1 μm^2 sram cell," in *International Electron Devices Meeting*, San Francisco, CA, USA, December 9–11, 2002, pp. 61–64.



Avinash Karanth Kodi received the Ph.D. and M.S. degrees in Electrical and Computer Engineering from the University of Arizona, Tucson in 2006 and 2003 respectively. He is currently an Assistant Professor of Electrical Engineering and Computer Science at Ohio University, Athens. His research interests include computer architecture, optical interconnects, chip multiprocessors (CMPs) and network-on-chips (NoCs). Dr. Kodi is a member of IEEE.



Randy Morris received his B.S. and M.S. degrees in Electrical Engineering and Computer Science from Ohio University, Athens in 2007 and 2009 respectively. He is currently pursuing his PhD degree at Ohio University. His research interests include optical interconnects, network-on-chips (NoCs) and computer architecture.



Ahmed Louri received the PhD degree in Computer Engineering in 1988 from the University of Southern California (USC), Los Angeles. He is currently a full professor of Electrical and Computer Engineering at the University of Arizona, Tucson and the director of the High Performance Computing Architectures and Technologies (HPCAT) Laboratory (www.ece.arizona.edu/~ocppl). His research interests include computer architecture, network-on-chips (NoCs), parallel processing, power-aware parallel architectures and optical interconnection networks.

Dr. Louri has served as the general chair of the 2007 IEEE International Symposium on High-Performance Computer Architecture (HPCA), Phoenix, Arizona. He has also served as a member of the technical program committee of several conferences including the ACM/IEEE Symposium on Architectures for Networking and Communications Systems (ANCS), the OSA/IEEE Conference on Massively Parallel Processors using Optical Interconnects, among others. Dr. Louri is a senior member of the IEEE and a member of the OSA.