Design of High Bandwidth Photonic NoC Architectures Using Optical Multilevel Signaling

Tzyy-Juin Kao
Department of Electrical and Computer Engineering
University of Arizona
Tucson, Arizona
Email: tjkao@email.arizona.edu

Ahmed Louri
Department of Electrical and Computer Engineering
George Washington University
Washington, DC
Email: louri@gwu.edu

Abstract—**Network-on-chip (NoC) is a key component for boosting the system performance of future chip multiprocessors. With the projected increase in the number of cores on the chip, the NoC is perceived to be the limiting component for performance and scaling. Photonic NoCs are under serious consideration for scaling future multicore architectures. In this paper, we propose two photonic NoC architectures based on an optical multilevel signaling technique that can double the transmission bandwidth and reduce the area requirements. Simulation studies show that the proposed methodology saves up to 53% of power and reduces the area overhead by as much as 81% compared with metallic-based NoCs.**

I. Introduction

Silicon photonic devices are compatible with standard CMOS technology and have brought light onto a chip [1]. The photonic link features a high data-transmission rate and low propagation loss, making it especially suitable for replacing long-distance wires. In the wavelength-division multiplexing (WDM) technique, several dozen wavelengths share a waveguide without interference and can be modulated and received individually [2]. Recent advancements in NoC design have leveraged the benefits of silicon photonics [3]–[5]; however, because designers put emphasis on improving both the speed and power, the chip area consumption was usually neglected. Many photonic NoC architectures require three-dimensional stacking technology [6] and an additional die to place excessive photonic devices, increasing the manufacturing costs.

Current photonic NoC architectures adopt micro-ring resonators to modulate wavelengths with an on-off keying (OOK) format, utilizing only the presence and absence of a wavelength to represent logical 1 and 0 [2]. To increase aggregate communication bandwidth without adding more links, many modulation formats used in fiber-optic communication are more advanced than OOK, such as optical multilevel signaling (OMLS). Because OMLS has multiple levels of amplitude to assign more bits of data at once, the capability of the bandwidth can be increased. Furthermore, OMLS offers a lower implementation complexity [7]; thus, it is more feasible to implement on a processor chip owing to the area constraint.

This paper presents a bottom-up approach to introduce OMLS-based NoC architectures. We briefly describe the structure of the OMLS link with a transmission bandwidth twice as large as the conventional photonic link [8]. Then, we employ a Clos network topology [4], [9] to develop a target system that can best harness the advantages of OMLS. Because of the multistage routing of Clos, packets can be distributed evenly across the middle routers, providing high load-balancing capability and high path diversity. Finally, we describe the implementation approach and propose two OMLS-based NoCs using off-chip and on-chip lasers for a 64-tile chip. Simulation results show that the proposed architectures consume a significantly lower chip area. This is achieved with only a slight or no increase in power consumption compared with the current photonic OOK-based NoCs.

II. OMLS-Based NoC Architectures

In this section, we illustrate the implementation approach of the on-chip OMLS link by using a simple 4-tile architecture as an example and propose two 64-tile OMLS-based NoC architectures. Readers could refer to [8] for more details about the on-chip OMLS link.

A. Implementation Approach

Fig. 1(a) illustrates the structure of an OMLS link. A 2-way asymmetric splitter [10] is placed at the beginning of the transmitter, diverting two-thirds of the input laser power to the upper waveguide; the rest remains in the lower waveguide. Hence, the upper waveguide commands a larger amplitude power than the lower waveguide. Each waveguide on the transmitter side includes an identical series of rings that modulate the same set of wavelengths simultaneously, and both the upper and lower waveguides perform OOK modulation according to the electrical inputs. After the modulation, a combiner connecting both waveguides as inputs merges these two OOK signals back into one four-amplitude-level (4-ASK) signal owing to the constructive interference, assuming both waveguides present identical phase delays. Each amplitude level represents one of the four combinations of two bits (00, 01, 10, and 11). Finally, the modulated wavelength reaching the destination is coupled into a ring and converted back into two electrical signals. As a result, because there are two output bits in each time interval, the bandwidth of each wavelength is doubled compared with the conventional link.
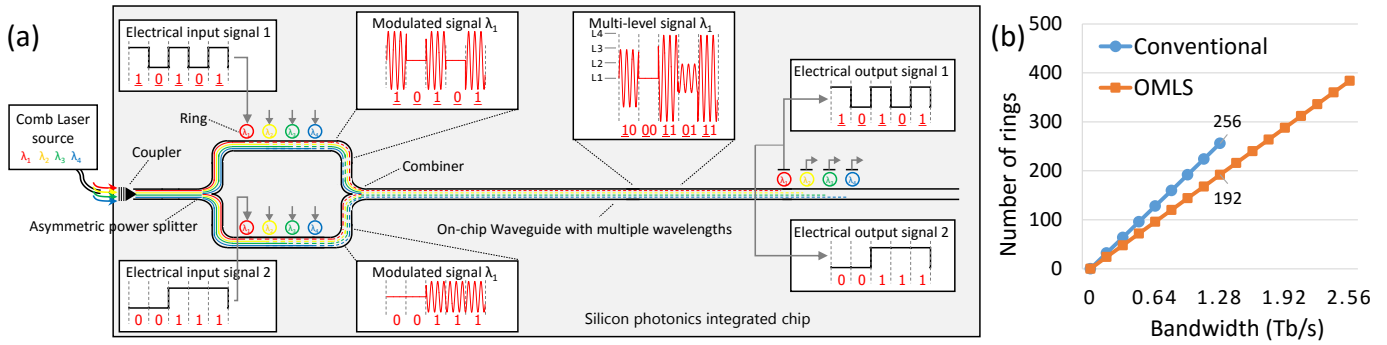
Fig. 1. (a) OMLS link: an example of encoding a four-amplitude-level signal into a single wavelength with two input binary signals. (b) Number of rings required for a singular conventional link and an OMLS link. The conventional link can only achieve half the bandwidth of the OMLS link.

Fig. 1(b) shows the overall numbers of rings required for the conventional and OMLS links, assuming that up to 128 wavelengths can be placed on a waveguide by transmitting them on opposite directions. A single OMLS link with rings modulated at 10 GHz can achieve a maximum bandwidth of 2.56 terabits per second, whereas a conventional link can achieve only half of that. Given the same transmission bandwidth, OMLS requires half as many wavelengths as OOK. As the number of wavelengths decreases, the number of rings required on both the transmitter and receiver sides also decreases. Therefore, although the OMLS link requires one additional ring per wavelength, the overall number of rings is less than that of the conventional link.

We use a small Clos topology for illustration purposes to simplify the design and verification. Fig. 2 shows the conventional and OMLS implementations of a 2-ary 3-stage Clos topology for a 4-tile chip. This Clos topology is composed of four tiles, six routers, and a total of sixteen links (Fig. 2(a)). In this example, because of the proximity of the tiles, two tiles are grouped into a cluster and assigned three routers—ingress, middle, and egress routers—which represent three different stages of the network. Tiles of the same cluster are electrically connected (gray lines) to their ingress and egress routers because of their adjacency. Routers in the same cluster are also electrically connected, whereas the inter-routers between different clusters are optically connected (colored lines).

To implement four optical connections (colored lines in Fig. 2(a)) between two clusters, we perform two different approaches as shown in Fig. 2(b). Either two conventional links or one OMLS link is required to implement this network and achieve the same aggregate communication bandwidth. We use four colors to correspond to each optical connection of the Clos topology. In the conventional optical approach (left side of Fig. 2(b)), the left waveguide is used for connecting the ingress and middle routers from cluster 1 to cluster 2 and vice versa. The right waveguide uses the same method to connect the middle and egress routers. The right side of Fig. 2(b) shows another approach that compresses all four connections into one OMLS link. Each electrical input signal is dissected and wired to two identically colored rings that modulate the same set of wavelengths to generate a double-bandwidth signal.
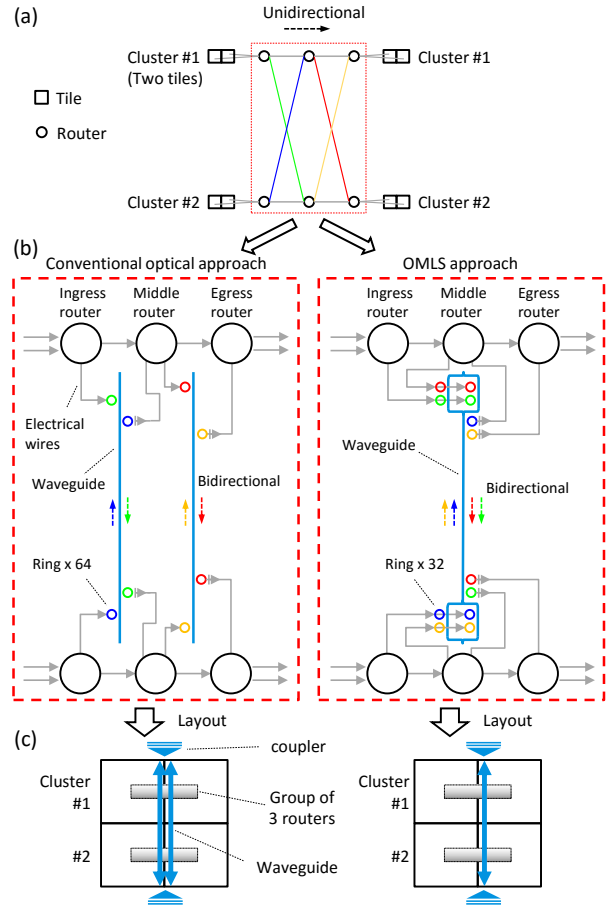

Fig. 2. (a) Basic 2-ary 3-stage Clos topology. The optical connections in the dashed box can be implemented with either two conventional photonic links or a single OMLS link (b). Each color of rings corresponds to the same color line in the topology. (c) Four-tile chip layouts for the links.

Two physical layouts on a 4-tile chip are shown in Fig. 2(c). The chip is divided into two clusters: upper and lower. Three routers of the same cluster are grouped together and placed in the center of every cluster. To achieve a bidirectional connection in the waveguides, two external laser sources with different sets of wavelengths are required, which are connected to both sides of the chip. The total number of rings required

(a) Clusters / Routers / Total 64 tiles

(b) Group of 3 routers / Tile / Cluster 1, Cluster 8, Cluster 2, Cluster 7, Cluster 3, Cluster 6, Cluster 4, Cluster 5

(c) ↓Laser sources↓ / Waveguide bundle

Nth link	Link between two clusters
1	1 ↔ 8
2	1 ↔ 2
3	1 ↔ 7
4	2 ↔ 7
5	2 ↔ 8
6	7 ↔ 8
7	1 ↔ 3
8	1 ↔ 6
9	2 ↔ 3
10	2 ↔ 6
11	3 ↔ 6
12	3 ↔ 7
13	3 ↔ 8
14	6 ↔ 7
15	6 ↔ 8

Nth link	Link between two clusters
16	1 ↔ 4
17	1 ↔ 5
18	2 ↔ 4
19	2 ↔ 5
20	3 ↔ 4
21	3 ↔ 5
22	4 ↔ 5
23	4 ↔ 6
24	4 ↔ 7
25	4 ↔ 8
26	5 ↔ 6
27	5 ↔ 7
28	5 ↔ 8

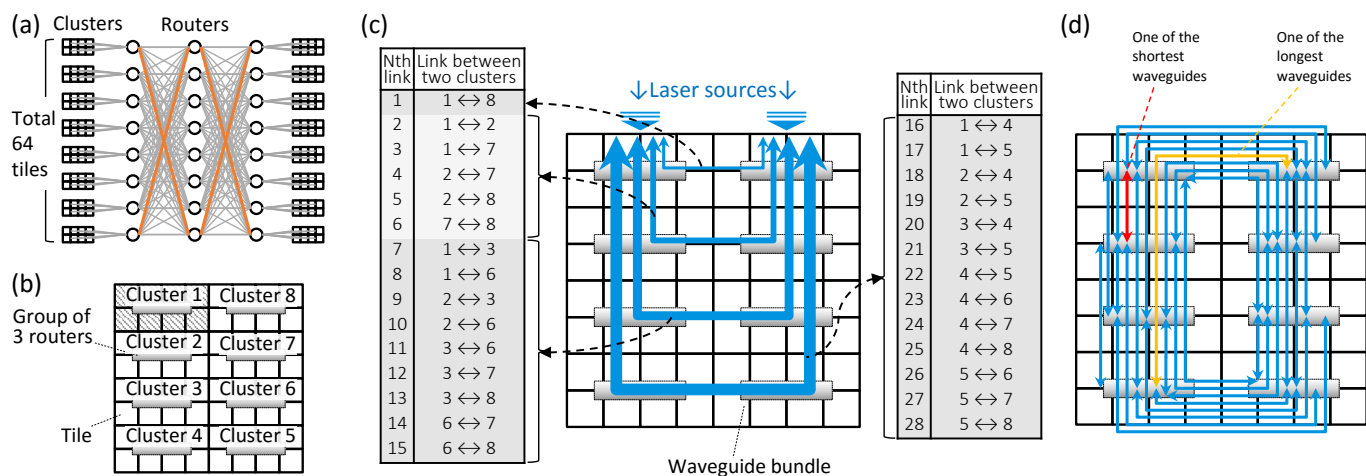**(d)** One of the shortest waveguides / One of the longest waveguides

Fig. 3. (a) 8-ary 3-stage Clos topology. The links are unidirectional, from left to right. The communication links between first and last clusters are highlighted. (b) The 64-tile chip layout is divided into eight clusters. Each cluster includes one group of three routers located at its center. For clarity, the links are not shown. (c) U-shaped OMLS-based architecture for this Clos topology on the 64-tile chip with OMLS links. The U-shaped arrows—having four different lengths—each represent a waveguide or a waveguide bundle. A thicker blue line indicates a larger number of waveguides (1, 5, 9, and 13). The tables list a total of 28 links and their bidirectional connections between any two clusters. (d) O-shaped OMLS-based architecture. The clusters have their own on-chip lasers to generate light individually. The double arrows represent bidirectional waveguides and form a layout with minimal lengths and no crossing point.

for the conventional optical and OMLS approaches is 512 and 384, respectively. Consequently, with the OMLS approach, the number of long-distance waveguides and rings are both reduced. This OMLS implementation approach is used to design 64-tile NoC architectures in the following subsections.

B. U-Shaped Photonic Clos Architecture

We propose a U-shaped photonic NoC architecture using the OMLS links. Fig. 3(a) shows an 8-ary 3-stage Clos topology for a 64-tile network, organized in eight tiles per cluster. Each cluster corresponds to three routers that are closely located, including one for each ingress, middle, and egress router. The network has the previously discussed implementation approach: inter-routers between different clusters are optically connected (total of 112 optical connections), and others are electrically connected. Because one OMLS link can be responsible for up to four optical connections, the number of OMLS links required is 28, regardless of the routing paths.

Fig. 3(b) shows a chip layout that distributes 64 tiles and 8 groups of routers symmetrically. To implement this 8-ary 3-stage Clos topology, 28 OMLS links illustrated by U-shaped double-sided arrows are placed, as shown in Fig. 3(c). The shortest arrow has only one waveguide, which is responsible for the communication between clusters 1 and 8. The other three arrows, having different thicknesses, indicate waveguide bundles, comprising 5, 9, and 13 waveguides from thin to thick. The corresponding communication links for each waveguide and waveguide bundle are presented in the tables in Fig. 3(c). To achieve bidirectional communication among clusters, two external laser sources are coupled to both sides of each waveguide and placed at the top of the chip. Therefore, the physical layout of the waveguides on a 64-tile chip forms a U-shaped structure.

C. O-Shaped Photonic Clos Architecture

We propose a second OMLS-based NoC architecture, wherein off-chip lasers are replaced by state-of-the-art on-chip lasers to further reduce the waveguide lengths and optimize the power efficiency. To design this architecture, we use the same 8-ary 3-stage Clos network topology, implementation approach, and 64-tile chip as the previously described architecture. On-chip lasers provide the benefits of layout flexibility, the capability of switching on and off sources, and the elimination of coupling power losses [11]. Because lasers can be placed anywhere on the chip, both sides of the waveguides can be located inside the chip and do not need to gather at the edge. The groups of routers are distributed across the chip, so the laser sources can be similarly distributed. Each cluster can have its on-chip lasers at the center; thus, bidirectional waveguides can start at one cluster and end at another. Consequently, all OMLS links are reallocated back-to-back without overlapping with the minimum-length requirements, forming an O-shaped architecture, as shown in Fig. 3(d). One of the shortest waveguides, highlighted in red, communicates between clusters 1 and 2; one of the longest waveguides, highlighted in yellow, communicates between clusters 4 and 8.

We estimate the required waveguide lengths by assuming that each waveguide begins and ends at the center of its cluster and that the unit is the length of a tile. The results indicate that the O-shaped architecture improves the waveguide lengths significantly because of the on-chip lasers. The shortest waveguides are 2 tile units in length and can be used for connecting two adjacent clusters. The longest waveguides in the U-shaped and O-shaped architectures are 18 and 10 tile units long, respectively. Although the total number of waveguides for both architectures is 28, the overall waveguide length is 62% lower for the O-shaped architecture.
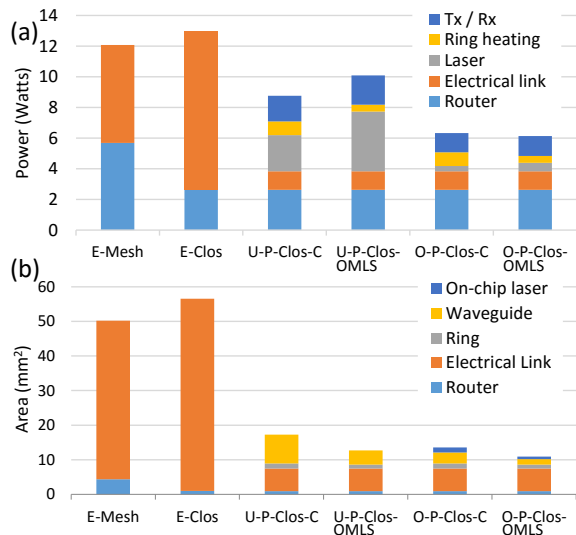
Fig. 4. (a) Power breakdown and (b) area overhead for different architectures.

III. EVALUATION

This section presents an architectural analysis of the power and area for NoCs operating at 5 GHz on a 400-mm^2 chip. P-Clos architecture offers a uniformly low latency, leveraging the benefits of both Clos topology (low hop counts) and photonic technology (fast transmission) [4]. To evaluate the power consumption, we modify the photonic models of the DSENT simulator [12] to include OMLS links and use its default parameters. All the devices are based on a 22-nm CMOS technology. Our results include two electrical architectures for baseline comparison: mesh (E-Mesh) and Clos (E-Clos). The proposed photonic Clos (P-Clos) architectures can be categorized into the following four configurations: 1) U-P-Clos-C: U-shaped photonic Clos architecture using conventional photonic links. 2) U-P-Clos-OMLS: U-shaped photonic Clos architecture using OMLS links. 3) O-P-Clos-C: O-shaped photonic Clos architecture using conventional photonic links. 4) O-P-Clos-OMLS: O-shaped photonic Clos architecture using OMLS links.

A. Power

For the ring-modulation power, the depletion-mode ring, which can be tuned electrically, is adopted. Theoretically, OMLS requires 4.8 dB more laser power at the receiver circuit to differentiate 4-ASK signals and obtain the same bit error rate [7]. The power breakdowns for different network architectures are shown in Fig. 4(a). Particularly, because Clos networks are composed of many long distance links, they appear to be well-suited for optical interconnects. Therefore, when implementing conventional photonic links, U-P-Clos-C eliminates 41% of the wire power and reduces 32% of the total power. By using OMLS links, U-P-Clos-OMLS reduces the overall power by 22% compared with E-Clos. For further improving the power efficiency by using the O-shaped architecture, O-P-Clos-OMLS can save up to 53% of the power compared with E-Clos.

B. Area

The electrical links are the dominating factor, as shown in Fig. 4(b); however, their effect can be minimized by using photonic technology. For P-Clos architectures, the remaining electrical links enable short-distance communications between tiles and routers within each cluster. After removing most of the long distance electrical links in E-Clos, U-PClos-C can save up to 70% of the area overhead. Substituting OMLS links in U-P-Clos-OMLS and O-P-Clos-OMLS further reduces the area of the waveguides by half because the waveguide bandwidth is doubled. U-P-Clos-OMLS reduces the area of photonic components to less than 4 mm^2, decreasing the area overhead by as much as 81% and 37% compared with E-Clos and U-P-Clos-C, respectively. This not only yields a similar power consumption to O-P-Clos-C but also reduces the number of waveguides and rings required

IV. CONCLUSION

We propose a compact photonic OMLS structure and its application to NoC design. The OMLS link doubles the transmission bandwidth of each waveguide by transmitting data into a 4-ASK signal. It exhibits great potential for improving bandwidth, area, and cost of optical interconnects, and for NOCs in particular. To highlight the advantages of OMLS for NoCs, we propose detailed implementation of OMLS-based photonic architectures, and the simulation results indicate: 1) low area requirement; 2) high feasibility of monolithic integration; and 3) only a slight or no increase in power. Consequently, the use of OMLS for NoC design is a promising approach to satisfy communication demands of future multicore architectures.

REFERENCES

[1] R. Meade *et al.*, "Integration of silicon photonics in bulk CMOS," in *Symp. VLSI Technology: Dig. Tech. Papers*, Honolulu, HI, 2014, pp. 1–2.

[2] M. Georgas *et al.*, "Addressing link-level design tradeoffs for integrated photonic interconnects," in *CICC*, San Jose, CA, 2011, pp. 1–8.

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

[4] A. Joshi *et al.*, "Silicon-photonic clos networks for global on-chip communication," in *NoCS*, San Diego, CA, 2009, pp. 124–133.

[5] R. Morris, A. Kodi, A. Louri, and R. Whaley, "Three-dimensional stacked nanophotonic network-on-chip architecture with minimal reconfiguration," *IEEE Trans. Comput.*, vol. 63, no. 1, pp. 243–255, Jan. 2014.

[6] B. Black *et al.*, "Die stacking (3D) microarchitecture," in *MICRO*, Orlando, FL, 2006, pp. 469–479.

[7] M. Atef and H. Zimmermann, *Optical Communication Over Plastic Optical Fibers: Integrated Optical Receiver Technology*. Springer, 2012.

[8] T. J. Kao and A. Louri, "Optical multilevel signaling for high bandwidth and power-efficient on-chip interconnects," *IEEE Photonics Technol. Lett.*, vol. 27, no. 19, pp. 2051–2054, Oct. 2015.

[9] C. Clos, "A study of non-blocking switching networks," *The Bell System Technical Journal*, vol. 32, no. 2, pp. 406–424, Mar. 1953.

[10] R. Thapliya, T. Kikuchi, and S. Nakamura, "Tunable power splitter based on an electro-optic multimode interference device," *Appl. Opt.*, vol. 46, no. 19, p. 4155–4161, 2007.

[11] M. J. R. Heck and J. E. Bowers, "Energy efficient and energy proportional optical interconnects for multi-core processors," *IEEE J. Sel. Topics Quantum Electron.*, vol. 20, no. 4, pp. 332–343, 2014.

[12] C. Sun *et al.*, "DSENT - a tool connecting emerging photonics with electronics for opto-electronic networks-on-chip modeling," in *NoCS*, Copenhagen, 2012, pp. 201–210.