SRing: A Sub-Ring Construction Method for Application-Specific Wavelength-Routed Optical NoCs

Zhidan Zheng, Meng Lian, Mengchu Li, Tsun-Ming Tseng, Ulf Schlichtmann
Chair of Electronic Design Automation
Technical University of Munich
Munich, Germany
{zhidan.zheng, m.lian, mengchu.li, tsun-ming.tseng, ulf.schlichtmann}@tum.de

***Abstract*—Wavelength-routed optical networks-on-chip (WRONoCs) attract ever-increasing attention for supporting high-speed communications with low power and latency. Among all WRONoC routers, optical ring routers attract much interest for their simple structures. However, current designs of ring routers have overlooked the customization problem: when adapting to applications that have specific communication requirements, current designs suffer high propagation loss caused by long worst-case signal paths and high splitter usage in power distribution networks (PDN). To address those problems, we propose a novel customization method to generate application-specific ring routers with multiple sub-rings, SRing. Instead of sequentially connecting all nodes in a large ring, we cluster the nodes and connect them with sub-ring waveguides to reduce the path length. Besides, we propose a mixed integer linear programming model for wavelength assignment to reduce the number of PDN splitters. We compare SRing to three state-of-the-art ring router design methods for six applications. Experimental results show that SRing can greatly reduce the length of the longest signal path, the worst-case insertion loss, and the number of splitters in the PDN, significantly improving the power efficiency.**

***Index Terms*—Wavelength-routed optical networks-on-chip, optical ring routers, sub-ring structure.**

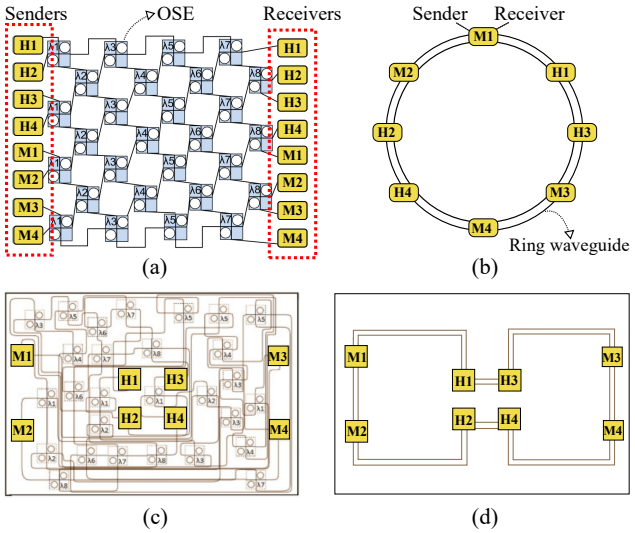

Fig. 1. (a) An 8-node λ-router topology, which places the senders and receivers of the same nodes at the opposite ends. (b) An 8-node ring topology, where the senders and the receivers of the same nodes are closely placed. (c) The layout results for the 8-node λ-router. (d) The layout results for the 8-node ring router.

I. Introduction

In the many-core era, traditional electrical networks-on-chip are confronted with two critical challenges: ever-increasing bandwidth requirements and significant communication energy [1], [2]. Optical networks-on-chip (ONoCs) have emerged as a disruptive technology solution to tackle these primary issues [3], [4]. Several critical advantages of silicon photonics have been demonstrated: bit-level parallelism with wavelength-division multiplexing (WDM), high-speed transmission in waveguides ($10.45\,\mathrm{ps/mm}$), and low energy consumption with microring resonators (MRRs) [1], [5], [6].

ONoCs can be classified into two categories based on their routing mechanisms: *active* and *passive* ONoCs [2]. On active ONoCs, before a node (sender) can transmit data to another node (receiver), a signal path needs to be established by a control module. On the other hand, passive ONoCs, also termed as *wavelength-routed* ONoCs (WRONoCs), reserve all required signal paths during design phase to prevent data collision, and thus save the time and energy for arbitration [2].

Current WRONoC routers can be categorized into two types based on their structures: *crossbar* [7]–[11] and *ring* [12]–[14]. Crossbar routers apply a matrix-like waveguide structure with nodes at the router boundaries and optical switching elements (OSEs) at the waveguide intersections to switch the optical signals between waveguides, as shown in Fig. 1(a). On the other hand, ring routers connect all nodes sequentially along two or more circular waveguides for data transmission in clockwise and counter-clockwise directions, as shown in Fig. 1(b). During physical implementation, crossbar routers usually require much placement & routing (P&R) effort, such as re-arranging the positions of OSEs, to adapt their matrix-like structure to actual layout constraints. This usually results in additional waveguide detours and crossings [15], as shown in Fig. 1(c). On the contrary, due to their simple structures, ring routers can easily be mapped to the physical plane without significant P&R effort, as shown in Fig. 1(d). Moreover, ring routers avoid the insertion loss and crosstalk noise caused by waveguide crossings [16]. Therefore, ring routers are considered an attractive option for WRONoC applications.

We notice that current ring router designs are based on the assumption that every node needs to communicate with every other node in the network, i.e. *full connectivity*, and thus connect all nodes sequentially. However, in practice, many applications do not ask for full connectivity [17]–[21]. For example, Fig. 2(a) shows the communication requirement of

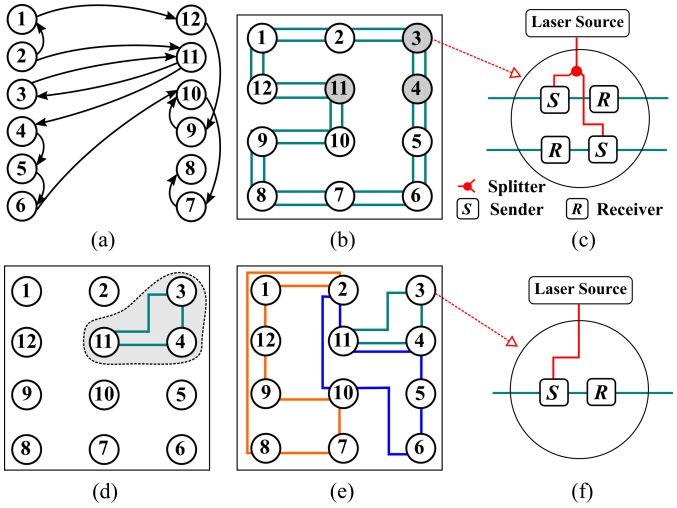

Fig. 2. (a) The communication requirements. (b) A classic 12-node ring router design. (c) Wavelengths are distributed from the laser source to both senders of node 3 via a PDN waveguide and a splitter. (d) A sub-ring connects nodes 3, 4 and 11. (e) A customized ring router design. (f) Node 3 now only has one sender supported by a single PDN waveguide.

a multi-window display (MWD) application [17], in which a node only communicates to a subset of the other nodes. For those specific applications, applying current ring router designs raises significant energy efficiency concerns for two reasons:

- Firstly, long signal paths between certain nodes result in high propagation loss, which increases power consumption. In the current ring structure, since all nodes are connected sequentially, each node is destined to be far away from some other nodes. For the MWD application, suppose that the nodes are arranged regularly on the chip. A classical ring router design is shown in Fig. 2(b), where the signal path between nodes 4 and 11 almost spans half the perimeter of the chip regardless of the transmission direction (clockwise or counter-clockwise), even though nodes 4 and 11 are physically close to each other.
- Secondly, redundant resources, such as extra senders and splitters, result in a significant power waste. In a ring router, it is conventional for a node to equip each ring waveguide with an individual sender to transmit data on that waveguide, and every sender is supplied with laser power from off-chip lasers via a power distribution network (PDN) [22]. Fig. 2(c) shows a typical PDN design for a node with two senders connected by a splitter that distributes the laser power of the same wavelengths to the senders. However, in specific applications, some nodes send signals to only a node, such as node 3 in the MWD application. For those nodes, it is unnecessary to equip them with senders at both ring waveguides because the redundant senders will not only increase power consumption but also require splitters to distribute the wavelengths, which causes additional splitter loss [22].

The above-mentioned problems can be conquered by customizing the waveguide connections in a ring router based on communication requirements. For example, for the MWD application, if we connect nodes 3, 4, and 11 with a sub-ring waveguide, as shown in Fig. 2(d), the signal sent from node 11 to node 4 can avoid the long detour. Fig. 2(e) shows

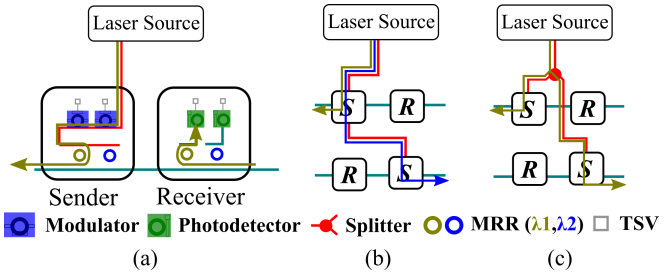

Fig. 3. (a) Working mechanism of a sender and a receiver. (b) If two senders are connected sequentially by the PDN (red lines), only wavelengths that do not resonate with the first sender can reach the second sender. (c) A PDN splitter is needed to provide the same wavelengths to two different senders.

a customized design consisting of three sub-ring waveguides, in which the signal paths are significantly shortened, and the redundant waveguides, as well as senders, can be removed, which also contributes to a simpler PDN, as shown in Fig. 2(f). That indicates the benefits of customizing the connections in a ring router to fit the required connectivity.

In this paper, we propose a customization method, **SRing**, to generate application-specific ring routers with multiple sub-rings. As the name suggests, we construct small rings to connect the nodes rather than a large ring. Specifically, we propose a clustering algorithm to cluster the nodes based on their physical locations and communication requirements and connect the nodes with sub-ring waveguides. To further reduce the power consumption, we propose a mixed integer linear programming (MILP) model to optimize the wavelength assignment in the network and thus minimizes the need for PDN splitters. We compare SRing with three state-of-the-art ring router design methods: ORNoC [12], CTORing [13], and XRing [14]. Results of the comparisons demonstrate the superiority of SRing over state-of-the-art methods in improving energy efficiency. For a benchmark of a realistic multimedia communication system [21], SRing reduces total laser power by more than 64% compared to the state-of-the-art design methods.

II. Background

A. Working Mechanism of Senders/Receivers

In WRONoC ring routers, a sender consists of an array of modulators and microring resonators (MRRs), and a receiver consists of an array of photodetectors and MRRs. Modulators and photodetectors perform the electronic-optical (E/O) and optical-electronic (O/E) conversion, respectively, and MRRs couple/decouple optical signals to/from the ring waveguide [22]. Fig. 3(a) illustrates the working mechanism of a sender and a receiver. Specifically, a PDN waveguide delivers wavelength λ_1 from the laser source to the sender, so that the sender can modulate an optical signal on λ_1 and couple it to the ring waveguide by the MRR that is on-resonance with λ_1. When the signal on λ_1 reaches a receiver with another on-resonance MRR, the signal will be coupled to the MRR from the ring waveguide and delivered to the photodetector.

When a node sends/receives signals through multiple waveguides, it requires a sender/receiver at each waveguide. In a typical ring router, each node sends/receives signals through two ring waveguides for clockwise and counter-clockwise transmission, respectively. If the senders at different waveguides are

connected by a single PDN waveguide, as shown in Fig. 3(b), all wavelengths that are on-resonance with the first sender will be coupled to the first ring waveguide, and only wavelengths that are not used by the first sender can be delivered to the other sender. Due to the limited available wavelengths, a ring router commonly uses the same wavelengths for data transmission at different ring waveguides. Thus, a splitter is needed to divide the power into multiple portions and deliver them to different senders, as shown in Fig. 3(c). When the laser power of the wavelengths gets split, it gets degraded by 3 dB (assuming a 50% splitting ratio) [23]. Thus, if we can reduce the number of senders that use the same wavelengths at the same node, we can reduce splitter usage and improve energy efficiency.

B. Performance Factors of WRONoC Ring Routers

Insertion loss and crosstalk noise are typical performance factors for WRONoCs [2]. In particular, crosstalk noise is much more significant in crossbar routers than in ring routers, because crosstalk signals are majorly generated at the MRRs and waveguide crossings. As ring routers do not rely on OSEs for signal routing, crosstalk noise is not a critical concern [22]. For ring routers, the insertion loss is an important performance factor as it defines the laser power [22].

The insertion loss of a signal in WRONoCs can be considered as the summation of the following losses [24]: *modulator loss* and *photodetector loss*, which are fixed for sending/receiving the signal; *drop loss* and *through loss*, which happen at MRRs; *splitter loss*, which depends on the number of PDN splitters; as well as *propagation loss*, *crossing loss*, and *bending loss*, which depend on waveguide structures. The worst-case insertion loss of a wavelength is the maximum insertion loss value over all signals on that wavelength, which defines the laser power of the wavelength [25]. The total laser power is the summation of the laser power of all wavelengths.

C. State-of-the-art WRONoC Ring Routers

ORNoC [12] and CTORing [13] are two ring router design methods that apply the same ring structure, i.e. connecting all network nodes sequentially with two or more parallel ring waveguides to transmit signals clockwise and counter-clockwise. CTORing differs from ORNoC in improving the way of assigning wavelengths to signal paths. As a result, CTORing reduces wavelength usage than ORNoC [13]. However, both ORNoC and CTORing rely on long ring waveguides and face high propagation loss [14]. Besides, they equip each node with an individual sender along each ring waveguide and connect every two senders with a splitter. which usually lead to many senders and splitters that increase laser power [22].

XRing [14] is the latest design approach, which proposes a new ring design: adding OSEs to ring routers to shorten the signal paths. Moreover, it removes redundant senders to reduce laser power and proposes a new PDN design. However, XRing applies the same assumption as previous works that connect every two senders with a splitter. Thus, XRing suffers from high splitter usage, introducing insertion loss and power penalties.

So far, no previous approach has properly modelled the excessive usage of splitters and the resulting costs. Besides, the exploration of different ring structures is still limited, which is, however, necessary to broaden the applicability and versatility of optical ring routers.

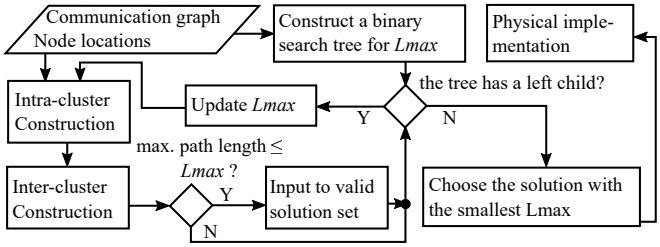

Fig. 4. The flow chart of the clustering algorithm.

III. Methodology

In this work, we propose a design method, SRing, to synthesize WRONoC ring routers with new ring structure. Instead of constructing a large ring to connect all nodes sequentially, we cluster the nodes and construct multiple small sub-rings for the required connectivity among the nodes. Then, we assign the wavelengths to the signal paths with an MILP model that minimizes the usage of wavelengths and splitters.

A. Sub-Ring Construction

For all nodes, we divide them into different clusters according to their *communication requirements* and *physical locations*, and construct a sub-ring for each cluster to support the *intra-cluster* communication. After that, we construct at most one additional sub-ring to interconnect all nodes that require *inter-cluster* communication. In other words, we ensure that each node has at most two senders dedicated to two ring waveguides. This setting limits the sender usage to avoid power waste.

To identify the best sub-ring connections that optimize the length of signal paths, we propose a clustering algorithm, as illustrated in Fig. 4. Firstly, considering that long signal paths lead to high propagation loss that increases energy penalties, we introduce a number L_{max} to denote the maximum permissible length of a signal path: the length of any signal path in any intra- or inter-cluster sub-rings must be smaller than or equal to L_{max}. To find the smallest possible L_{max}, we limit the range of L_{max} as $d_1 \leq L_{max} \leq d_2$ and construct a *balanced binary search tree*. We set d_1 to the maximum Manhattan distance[a] between any two communicating nodes and d_2 to the length of the longest signal paths if all nodes are connected sequentially as in a conventional ring router. By setting $\frac{d1+d2}{2}$ as the root of the search tree, we insert $2^h - 2$ equidistant real numbers within $[d_1, d_2]$ into the tree[b]. We start the clustering algorithm by setting L_{max} to the root. If the clustering solution satisfies the length limitation, we mark this solution as valid and update L_{max} as the left child of the root; otherwise, we discard this solution and update L_{max} as the right child of the root. If there is at least a left child of L_{max}, we repeat this process; otherwise, we stop it. At last, we choose the valid result with the smallest L_{max} to construct the sub-rings.

[a]The reason to use Manhattan distance is that the sub-ring waveguides will later be physically implemented either horizontally or vertically.

[b]h is a user-defined value that controls the height of the tree and the time complexity for searching.

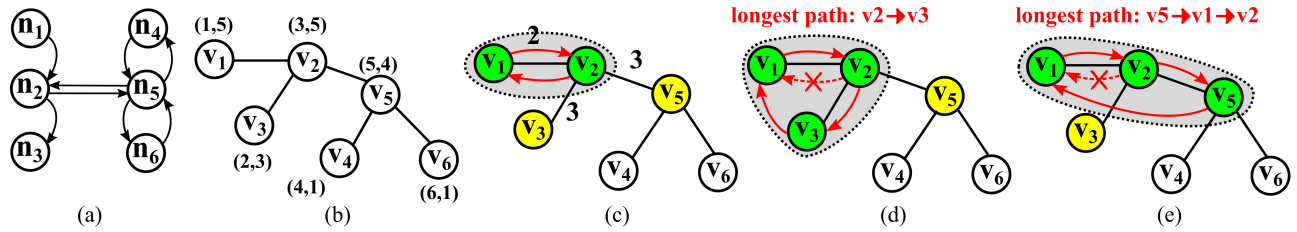

Fig. 5. (a) The communication graph. (b) Graph $G = (V, E)$. (c) v_1 is closer to v_2 and thus is selected to form the initial cluster with v_2. The red arrows denote the intra-cluster sub-ring and the corresponding data transmission directions. (d) v_3 is absorbed into the sub-ring by replacing the waveguide segment (v_2, v_1) with two waveguide segments (v_2, v_3) and (v_3, v_1). (e) v_5 is absorbed into the sub-ring by replacing the waveguide segment (v_2, v_1) with two waveguide segments (v_2, v_5) and (v_5, v_1).

To represent the input communication requirements and the physical locations of a network consisting of n nodes ($n \in \mathbb{N}$) we create a graph $G = (V, E)$, where $V = \{v_1, v_2, ...v_n\}$ is the set of vertices representing the nodes and $E = \{e(v_i, v_j) | v_i, v_j \in V$, node i and node j has communication$\}$ is a set of undirected edges. For example, Fig. 5(a) shows the communication graph of a DSP network [17], based on which we construct a graph with six vertices shown in Fig. 5(b). The coordinates next to each vertex denote the locations of the corresponding nodes on the layout plane.

1) Intra-Cluster Sub-Ring Construction: For each vertex, we consider it to be the *initial vertex* as the start to construct intra-cluster sub-ring and select the one that minimizes the length of signal paths. Among all adjacent vertices of the initial vertex v_i, we pick the one with the smallest Manhattan distance to v_i to construct the initial cluster C_1 together with v_i and connect them with a sub-ring waveguide. We set the signal transmission direction in the sub-ring to be the direction that minimizes the length of the longest intra-cluster signal path. For example, in Fig. 5(c), suppose that we arbitrarily choose v_2 to begin with, we select v_1, the closest vertex to v_2 among the adjacent vertices of v_2, to form the initial cluster and connect them with a sub-ring waveguide. Since the length of the path from v_1 to v_2 stays the same regardless of the transmission direction, we arbitrarily set the direction as clockwise.

For each vertex in the initial cluster, we mark its unvisited adjacent vertices as *candidates* to expand the sub-ring. For each candidate, we apply an *absorption* method, which will be described in details later, to add it into the sub-ring. We then calculate the length of each signal path in this sub-ring. If the length of the longest signal path in the resulting sub-ring exceeds L_{max}, we mark the candidate as *invalid*. Among all valid candidates, we *absorb* the one with the least increment in the length of the longest signal paths into the sub-ring.

Specifically, our *absorption* method is a modified version of a method proposed in [26], which considers the impact on the longest path length during absorption. To *absorb* a new vertex v_x into the sub-ring, we check every existing waveguide segment (v_y, v_z) in the sub-ring by replacing it with two waveguide segments (v_y, v_x) and (v_x, v_z). We then calculate the length of the signal paths in the resulting sub-ring. If the length of any signal path exceeds L_{max}, we mark the sub-ring as *invalid*. Among all valid absorption options, we absorb the one that minimizes the longest length of signal paths.

For example, as shown in Fig. 5(c), v_3 and v_5 are candidates for the initial cluster. Thus, we perform absorption for each of them. To absorb v_3, we replace one of the waveguide segments between v_1 and v_2 with two waveguide segments (v_2, v_3) and (v_3, v_1), as shown in Fig. 5(d). The longest signal path is the path from v_2 to v_3 with length 3. On the other hand, absorbing v_5 will lead to a signal path with length 7, as shown in Fig. 5(e). Suppose that L_{max} is 8, both absorption options are valid. But since 3 is smaller than 7, we choose v_3 to expand the cluster.

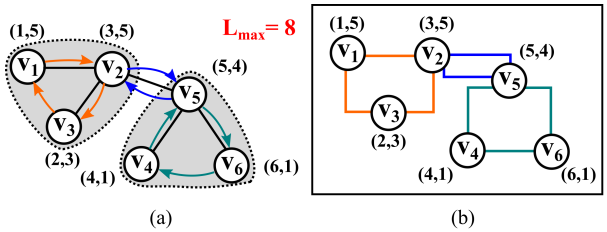

Fig. 6. (a) The valid clustering solution with the smallest L_{max}, where three sub-rings are represented by orange, blue, and turquoise arrows, respectively. (b) The corresponding implementation of the clustering solution.

For a newly absorbed vertex, we mark its unvisited adjacent vertices as candidates and continue the construction of this intra sub-ring until no other vertices can be absorbed due to the violation to L_{max}. If there are unclustered vertices, we consider each of them as an initial vertex and repeat this step to construct the intra-cluster sub-rings until all vertices are clustered.

2) Inter-Cluster Sub-Ring Construction: After constructing the intra-cluster sub-rings, we create a set V_{inter} for all nodes that require inter-cluster communications. For each $v \in V_{inter}$, we consider it as the initial vertex to construct the inter-cluster sub-ring using the same method for intra-cluster sub-rings. If the longest signal path in the inter-cluster sub-ring exceeds L_{max}, we mark this initial vertex as invalid and check the next initial vertex. If we fail to find any valid initial vertex, we mark the whole clustering solution as invalid, increase L_{max} to its right child, and return to the intra-cluster construction, as shown in Fig. 4; otherwise, we choose the initial vertex that minimizes the longest signal path for the inter-cluster sub-ring.

3) Physical Implementation: After obtaining the solution with the smallest L_{max}, we route waveguides on the physical plane to implement the sub-rings. Specifically, we connect two nodes by routing a waveguide either horizontally or vertically. We manually optimize the routing results to minimize the number of waveguide crossings and bends. Thanks to the simple ring structure, the physical implementation of the sub-rings does not require much effort. For example, Fig. 6(a) shows a clustering solution and Fig. 6(b) shows its layout.

B. Wavelength Assignment

From the layout results, we obtain the length of sub-ring waveguides and signal paths. To assign the wavelengths, we propose an MILP optimization model that, for the first time, optimizes the impact of splitter usage on the the number of wavelength usage and the worst-case insertion loss.

First, we introduce a set S for signal paths and another set Λ for wavelengths. For each $s \in S$ and $\lambda \in \Lambda$, we introduce a binary variable $b_{s,\lambda}$ to indicate if signal path s is assigned with wavelength λ. To ensure that each signal path is assigned with exactly one wavelength, we introduce the following constraints:

$$\forall s \in S: \quad \sum_{\lambda \in \Lambda} b_{s,\lambda} = 1 \tag{1}$$

If multiple signal paths overlap at any waveguide segment, those signal paths need to be assigned with different wavelengths to avoid data-collision. For each $s \in S$, we create a set $S^s_{conflict}$ to store all signal paths that overlap with s and introduce the following constraints to prevent the same wavelength from being assigned to overlapping signal paths:

$$\forall s \in S,\ \forall \lambda \in \Lambda: \ b_{s,\lambda} + \sum_{s' \in S^s_{conflict}} b_{s',\lambda} \leq 1 \tag{2}$$

To calculate the total wavelength usage, we introduce an integer number i_{wl} and model it with the following constraints:

$$i_{wl} = \sum_{\lambda \in \Lambda} min(\sum_{s \in S} b_{s,\lambda}, 1) \tag{3}$$

For each $\lambda \in \Lambda$, if it is assigned to more than one signal path, i.e. $\sum_{s \in S} b_{s,\lambda} \geq 1$, this wavelength counted in i_{wl}.

In SRing, a node has at most two senders for intra- and inter-cluster communications. If the two senders share any wavelengths, they must be connected to a splitter; otherwise, a splitter is not required. We introduce a set N for all nodes and a binary variable $b^{n_i}_{sp}$ for each $n_i \in N$ to represent whether node n_i requires a splitter. For each $n_i \in N$, we store its signal paths for intra- and inter-cluster communication in the sets $S^{n_i}_{intra}$ and $S^{n_i}_{inter}$, respectively. We introduce the following constraints:

$$\forall n_i \in N,\ \forall \lambda \in \Lambda: \sum_{s \in S^{n_i}_{intra}} b_{s,\lambda} + \sum_{s' \in S^{n_i}_{inter}} b_{s',\lambda} \leq 1 + b^{n_i}_{sp} \tag{4}$$

For each signal path, we denote its insertion loss resulting from the physical layout excluding the PDN as L_s, which is calculated as the sum of all losses introduced in Sec. II-B except the loss caused by splitters. Note that the unit of all losses is dB. Then, we introduce a constant L_{sp} to denote the loss caused by a PDN splitter, which is the sum of the splitter loss and a splitting ratio. The loss caused by a splitter is valid for a signal path only when $b^{n_i}_{sp} = 1$. We introduce a continuous variable il_s to indicate the insertion loss of a signal path s at a node $n_i \in N$ and model it as follows:

$$\forall s \in S^{n_i}_{intra} \cup S^{n_i}_{inter}: il_s \geq b^{n_i}_{sp} * L_{sp} + L_s \tag{5}$$

We introduce a continuous variable il^{Smax} to indicate the worst-case insertion loss over all signal paths:

$$\forall s \in S: \quad il^{Smax} \geq il_s \tag{6}$$

For the worst-case insertion loss of each used wavelength, we introduce a continuous variable il^{max}_{λ} and model it as follows:

$$\forall \lambda \in \Lambda,\ \forall s \in S: \quad il^{max}_{\lambda} \geq il_s - (1 - b_{s,\lambda}) * \Xi \tag{7}$$

where Ξ is a very large auxiliary number. Only when a wavelength is used by at least one signal, the worst-case insertion loss of that wavelength is valid.

At last, we set the wavelength usage, the worst-case insertion loss over all signal paths, and the summation of the worst-case insertion loss of each wavelength as our optimization objectives:

$$Minimize: \quad \alpha * i_{wl} + \beta * il^{Smax} + \gamma * \sum_{\lambda \in \Lambda} il^{max}_{\lambda} \tag{8}$$

where α, β, and γ are user-defined weight coefficients that control the optimization preference.

IV. Results

We implemented SRing in C++ and solved the MILP model using Gurobi [27]. All weight coefficients in Eq. 8 were set to 1[c]. In SRing, we constructed the PDNs using the approach proposed in [22]. We compare SRing to three state-of-the-art design methods for ring routers: ORNoC [12], CTORing [13], and XRing [14], for seven typical benchmarks in WRONoCs [18], [28]. For a fair comparison, we implemented the three methods in C++, adopted the ring settings in CTORing[d], and constructed signal paths only for the required communications in the applications[e]. All experiments in this paper were conducted on a computer with an Intel 8-node $3.4\,\mathrm{GHz}$ CPU. For all tests, we applied the technology parameters from [22].

A. General Comparison

We tested SRing for seven applications: four large-scale, low-communication-density multimedia systems (MWD [17], VOPD [19], MPEG [29], and D26 [21]) and three small-scale, high-communication-density processor-memory networks (8PM-24 [30], 8PM-32 [12], and 8PM-44 [18]). Details and results are in Table I and Fig. 7.

In general, SRing improves energy efficiency significantly compared to the other three design methods. For example, for D26, the largest-scale networks, SRing decreases the total laser power by more than 64% compared to the others. In every case, SRing has the minimum laser power consumption among the four ring routers, as shown in Fig. 7.

The improvement is mainly driven by reducing the worst-case insertion loss. On one hand, we optimize the propagation loss with our clustering algorithm, which minimizes the length of sub-ring waveguides and signal paths. For example, for MWD, SRing decreases the length of the longest signal paths by 78%, 71%, and 33% compared to ORNoC, CTORing, and XRing, respectively. On the other hand, SRing optimizes the loss of splitters by minimizing the splitter usage. We calculate the number of splitters passed by every signal path in each ring router and denote the largest value as $\#\mathrm{sp}_w$. As shown in Table I, SRing has the least $\#\mathrm{sp}_w$ among all design methods. Therefore, SRing reduces the worst-case insertion loss with the losses in PDNs, denoted as il^{all}_w, by 14% – 26% compared

[c]Since $\sum_{\lambda \in \Lambda} il^{max}_{\lambda}$ is typically larger than $\sum_{\lambda \in \Lambda} b_{\lambda}$ and il^{Smax}, setting all coefficients to 1 focuses on optimizing $\sum_{\lambda \in \Lambda} il^{max}_{\lambda}$.

[d]i.e. using two ring waveguides to minimize power according to [13].

[e]For the PDNs of ORNoC and CTORing, we apply our PDN design. For XRing, we construct the PDNs using its proposed approach.

TABLE I
COMPARISON RESULTS OF ORNOC, CTORING, AND SRING FOR SEVEN TEST CASES.

	MWD #N = 12, #M = 13				VOPD #N = 16, #M = 21				MPEG #N = 12, #M = 26				D26 #N = 26, #M = 68				8PM-24 #N = 8, #M = 24				8PM-32 #N = 8, #M = 32				8PM-44 #N = 8, #M = 44			
	L	il_w	$\#sp_w$	il_w^{all}	L	il_w	$\#sp_w$	il_w^{all}	L	il_w	$\#sp_w$	il_w^{all}	L	il_w	$\#sp_w$	il_w^{all}	L	il_w	$\#sp_w$	il_w^{all}	L	il_w	$\#sp_w$	il_w^{all}	L	il_w	$\#sp_w$	il_w^{all}
ORNoC	1.8	5.2	5	21.7	3.0	6.0	5	22.7	2.2	5.5	5	21.7	5.0	7.9	6	29.2	1.2	4.8	4	17.6	1.4	4.9	4	18.2	1.8	5.2	4	18.4
CTORing	1.4	4.4		21.0	1.4	4.9		21.5	1.1	4.7		21.0	2.4	5.8		26.7	0.7	4.2		17.9	0.9	4.2		18.0	0.8	4.5		18.4
XRing	0.7	4.2	5	20.3	1.4	4.4	6	23.9	1.0	4.4	6	23.6	2.4	4.9	7	28.4	0.6	4.2	5	20.0	1.4	4.5	5	20.1	0.8	4.3	6	23.7
SRing	0.4	4.1	4	17.5	1.4	4.4	4	17.7	1.0	4.4	4	17.6	2.4	4.9	5	21.7	0.6	4.2	3	14.2	1.4	4.6	3	14.5	1.4	4.7	3	14.7

#N: the number of nodes; #M: the number of messages; L: the length of the longest signal path denoted in mm; il_w: the worst-case insertion loss **without** including the losses in PDNs denoted in dB; $\#sp_w$: the largest number of splitters passed by signal paths; il_w^{all}: the worst-case insertion loss of a wavelength **including** the losses in PDNs denoted in dB.

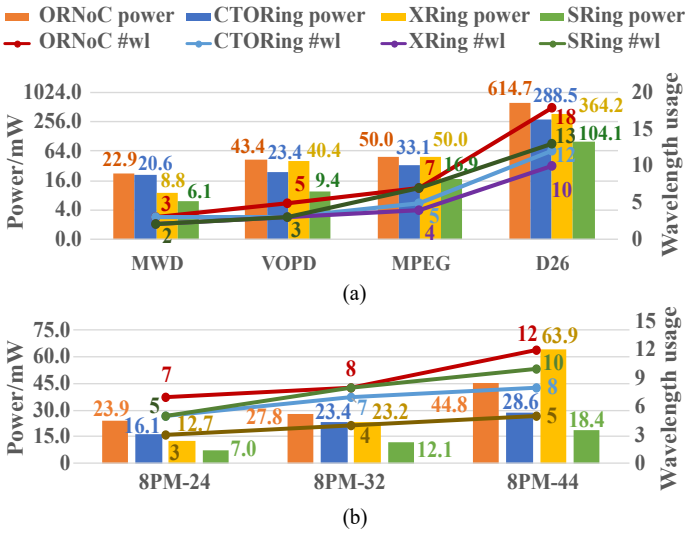

Fig. 7. Total laser power and the wavelength usage (#wl) of ORNoC, CTORing, XRing, and SRing for (a) the four multimedia communication systems and (b) the 8-node processor-memory networks.

to the other three methods. That contributes to a significant reduction in power consumption.

For the wavelength usage, ORNoC has the most wavelengths, and XRing has the fewest wavelengths. The wavelength usage of SRing depends on the communication density, i.e. the ratio of #M to #N. When the communication density is low, such as MWD and VOPD, SRing minimizes the wavelength usage using our MILP model. When the communication density is medium, such as 8PM-24 and 8PM-32, SRing and CTORing exhibit similar wavelength usage, which is slightly more than XRing's. When the communication density is high, like 8PM-44, or a node needs to talk to almost all other nodes, such as MPEG, SRing requires more wavelengths than CTORing. That is because our MILP model minimizes the number of splitters to optimize power consumption rather than sharing the same wavelengths among the senders. Thanks to the minimum splitter usage, SRing enhances energy efficiency significantly, even when the communication density is high. In contrast, XRing requires many splitters to reduce the number of wavelengths and thus consumes more laser power among the four methods.

B. Discussion: Solution Quality

We illustrate the computational complexity and the solution quality of our clustering algorithm and MILP model by randomly generating a hundred thousand solutions[f] and compare

[f]Specifically, we randomly cluster nodes, sequentially connect the nodes in each cluster to form sub-rings, and randomly assign wavelengths to signal paths. If the same wavelengths are assigned to the signal paths that overlap at the same waveguide sections, we denote a solution as *infeasible*.

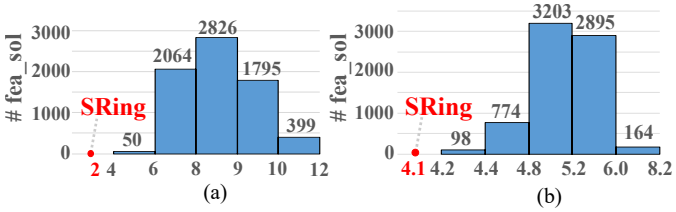

Fig. 8. The histogram of (a) #wl and (b) il_w for MWD. The bar located between a and b represents the number of feasible solutions, denoted as #fea_sol, with results in the range $(a, b]$.

TABLE II
PROGRAM RUNTIME OF SRING DENOTED IN SECONDS.

	MWD	VOPD	MPEG	D26_media	8PM-24	8PM-32	8PM-44
SRing	0.12	0.22	0.36	6.32	0.27	0.52	2.40

the feasible ones to the results of SRing.

Among all cases, we can only obtain feasible solutions for the MWD and VOPD applications. The number of feasible solutions dramatically drops when the number of nodes or communications increases. For MWD, 7134 (7%) out of 100,000 solutions are feasible, while only 19 ($\leq$ 1%) feasible solutions for VOPD can be found. On the contrary, SRing can easily obtain feasible solutions for each case and does not suffer a high computational burden, as illustrated in Table II. SRing can finish the design in a few seconds. Fig. 8(a) and (b) show the histogram of the wavelength usage and the worst-case insertion loss in ring routers for all feasible solutions found in MWD, respectively. In both graphs, we use a red circle to represent the solution generated by SRing. In both metrics, SRing has better results versus all feasible solutions, which indicates the good quality of SRing's solution.

V. CONCLUSION

In this work, we propose SRing, a method to construct multiple sub-rings for application-specific WRONoC ring routers. The sub-ring structure, a key component of SRing, is designed to better fit the customized communication requirement. This structure is formed by clustering nodes based on their physical proximity and communication demand, thereby shortening the signal paths. Then, we propose an MILP model that simultaneously optimizes wavelength- and splitter usage. The experimental results show that SRing outperforms the state-of-the-art WRONoC ring routers in terms of both energy and computational efficiency. By shortening the signal path and removing redundant splitters with sub-rings, SRing can reduce insertion loss and total laser power consumption.

ACKNOWLEDGMENT

This work is supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project Number 496766278.

REFERENCES

[1] Scott Van Winkle, Avinash Karanth Kodi, Razvan Bunescu, and Ahmed Louri. Extending the Power-Efficiency and Performance of Photonic Interconnects for Heterogeneous Multicores with Machine Learning. In *2018 IEEE International Symposium on High Performance Computer Architecture (HPCA)*, pages 480–491, 2018.

[2] Tsun-Ming Tseng, Alexandre Truppel, Mengchu Li, Mahdi Nikdast, and Ulf Schlichtmann. Wavelength-Routed Optical NoCs: Design and EDA — State of the Art and Future Directions: Invited Paper. In *2019 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, pages 1–6, 2019.

[3] Mahdi Tala, Marco Castellari, Marco Balboni, and Davide Bertozzi. Populating and Exploring the Design Space of Wavelength-Routed Optical Network-on-Chip Topologies by Leveraging the Add-Drop Filtering Primitive. In *IEEE/ACM International Symposium on Networks-on-Chip (NoCS)*, pages 1–8, 2016.

[4] Aditya Narayan, Yvain Thonnart, Pascal Vivet, César Fuguet Tortolero, and Ayse K. Coskun. WAVES: Wavelength Selection for Power-Efficient 2.5D-Integrated Photonic NoCs. In *2019 Design, Automation & Test in Europe Conference & Exhibition (DATE)*, pages 516–521, 2019.

[5] W. Bogaerts, P. De Heyn, T. Van Vaerenbergh, K. De Vos, S. Kumar Selvaraja, T. Claes, P. Dumon, P. Bienstman, D. Van Thourhout, and R. Baets. Silicon microring resonators. *Laser & Photonics Reviews*, 6(1):47–73, 2012.

[6] Sudip Shekhar, Wim Bogaerts, Lukas Chrostowski, John E. Bowers, Michael Hochberg, Richard Soref, and Bhavin J. Shastri. Roadmapping the next generation of silicon photonics. *Nature Communications*, 15(1):751, Jan 2024.

[7] Zhidan Zheng, Mengchu Li, Tsun-Ming Tseng, and Ulf Schlichtmann. Light: A Scalable and Efficient Wavelength-Routed Optical Networks-On-Chip Topology. In *2021 Asia and South Pacific Design Automation Conference (ASP-DAC)*, page 568–573, 2021.

[8] M. Briere, B. Girodias, Y. Bouchebaba, G. Nicolescu, F. Mieyeville, F. Gaffiot, and I. O'Connor. System Level Assessment of an Optical NoC in an MPSoC Platform. In *2007 Design, Automation Test in Europe Conference Exhibition (DATE)*, pages 1–6, 2007.

[9] Xianfang Tan, Mei Yang, Lei Zhang, Yingtao Jiang, and Jianyi Yang. On a Scalable, Non-Blocking Optical Router for Photonic Networks-on-Chip Designs. In *2011 Symposium on Photonics and Optoelectronics (SOPO)*, pages 1–4, 2011.

[10] Zhidan Zheng, Mengchu Li, Tsun-Ming Tseng, and Ulf Schlichtmann. LightR: A Fault-Tolerant Wavelength-Routed Optical Networks-on-Chip Topology. *Applied Sciences*, 13(15), 2023.

[11] Luca Ramini, Paolo Grani, Sandro Bartolini, and Davide Bertozzi. Contrasting wavelength-routed optical NoC topologies for power-efficient 3d-stacked multicore processors using physical-layer analysis. In *2013 Design, Automation Test in Europe Conference Exhibition (DATE)*, pages 1589–1594, 2013.

[12] Sébastien Le Beux, Jelena Trajkovic, Ian O'Connor, Gabriela Nicolescu, Guy Bois, and Pierre Paulin. Optical Ring Network-on-Chip (ORNoC): Architecture and design methodology. In *2011 Design, Automation and Test in Europe*, pages 1–6, 2011.

[13] Marta Ortín-Obón, Luca Ramini, Víctor Viñals-Yúfera, and Davide Bertozzi. A tool for synthesizing power-efficient and custom-tailored wavelength-routed optical rings. In *2017 Asia and South Pacific Design Automation Conference (ASP-DAC)*, pages 300–305, 2017.

[14] Zhidan Zheng, Mengchu Li, Tsun-Ming Tseng, and Ulf Schlichtmann. XRing: A Crosstalk-Aware Synthesis Method for Wavelength-Routed Optical Ring Routers. In *2023 Design, Automation & Test in Europe Conference & Exhibition (DATE)*, pages 1–6, 2023.

[15] Luca Ramini, Davide Bertozzi, and Luca Carloni. Engineering a Bandwidth-Scalable Optical Layer for a 3D Multi-core Processor with Awareness of Layout Constraints. In *2012 IEEE/ACM Sixth International Symposium on Networks-on-Chip*, pages 185–192, 05 2012.

[16] Sébastien Le Beux, Hui Li, Gabriela Nicolescu, Jelena Trajkovic, and Ian O'Connor. Optical crossbars on chip, a comparative study based on worst-case losses. *Concurrency and Computation: Practice and Experience*, 26(15):2492–2503, 2014.

[17] Shijun Lin, Li Su, Haibo Su, Depeng Jin, and Lieguang Zeng. Hierarchical cluster-based irregular topology customization for networks-on-chip. In *2008 IEEE/IFIP International Conference on Embedded and Ubiquitous Computing*, pages 373–377, 2008.

[18] Mengchu Li, Tsun-Ming Tseng, Davide Bertozzi, Mahdi Tala, and Ulf Schlichtmann. CustomTopo: A Topology Generation Method for Application-Specific Wavelength-Routed Optical NoCs. In *2018 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, page 1–8, New York, NY, USA, 2018.

[19] Pinar Kullu, Yilmaz Ar, Suleyman Tosun, and Suat Ozdemir. Mapping application-specific topology to mesh topology with reconfigurable switches. *IET Computers & Digital Techniques*, 14(1):9–16, 2020.

[20] Krishnan Srinivasan, K.S. Chatha, and G. Konjevod. Linear programming based techniques for synthesis of Network-on-Chip architectures. *Proceedings - IEEE International Conference on Computer Design: VLSI in Computers and Processors*, pages 422–429, 01 2004.

[21] Ciprian Seiculescu, Srinivasan Murali, Luca Benini, and Giovanni Micheli. SunFloor 3D: A Tool for Networks on Chip Topology Synthesis for 3-D Systems on Chips. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 29:9–14, 04 2009.

[22] Marta Ortín-Obón, Mahdi Tala, Luca Ramini, Víctor Viñals-Yufera, and Davide Bertozzi. Contrasting Laser Power Requirements of Wavelength-Routed Optical NoC Topologies Subject to the Floorplanning, Placement, and Routing Constraints of a 3-D-Stacked System. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, PP:1–14, 03 2017.

[23] Alexandre Truppel, Tsun-Ming Tseng, and Ulf Schlichtmann. PSION 2: Optimizing Physical Layout of Wavelength-Routed ONoCs for Laser Power Reduction. In *Proceedings of the 39th International Conference on Computer-Aided Design (ICCAD)*. Association for Computing Machinery, 2020.

[24] Alexandre Truppel, Tsun-Ming Tseng, and Ulf Schlichtmann. Accurate Infinite-Order Crosstalk Calculation for Optical Networks-on-Chip. *Journal of Lightwave Technology*, 41(1):4–16, 2023.

[25] Anja Von Beuningen, Luca Ramini, Davide Bertozzi, and Ulf Schlichtmann. PROTON+: A Placement and Routing Tool for 3D Optical Networks-on-Chip with a Single Optical Layer. *ACM Journal on Emerging Technologies in Computing Systems*, 12(4), 12 2015.

[26] Tsun-Ming Tseng, Mango C.-T. Chao, Chien-Pang Lu, and Chen-Hsing Lo. Power-Switch Routing for Coarse-Grain MTCMOS Technologies. In *2009 International Conference on Computer-Aided Design (ICCAD)*, page 39–46, New York, NY, USA, 2009.

[27] Gurobi Optimization, Inc. *Gurobi Optimizer Reference Manual*. http://www.gurobi.com, 2024.

[28] Zhidan Zheng, Mengchu Li, Tsun-Ming Tseng, and Ulf Schlichtmann. ToPro: A Topology Projector and Waveguide Router for Wavelength-Routed Optical Networks-on-Chip. In *2021 IEEE/ACM International Conference On Computer Aided Design (ICCAD)*, pages 1–9, 2021.

[29] D. Bertozzi, A. Jalabert, Srinivasan Murali, R. Tamhankar, S. Stergiou, L. Benini, and G. De Micheli. Noc synthesis flow for customized domain specific multiprocessor systems-on-chip. *IEEE Transactions on Parallel and Distributed Systems*, 16(2):113–129, 2005.

[30] Sébastien Le Beux, Ian O'Connor, Gabriela Nicolescu, Guy Bois, and Pierre Paulin. Reduction methods for adapting optical network on chip topologies to 3D architectures. *Microprocessors and Microsystems: Embedded Hardware Design*, 37(1):87–98, 2013.