Maximizing the Communication Parallelism for Wavelength-Routed Optical Networks-On-Chips

Mengchu Li†, Tsun-Ming Tseng†, Mahdi Tala◇, and Ulf Schlichtmann†
{mengchu.li, tsun-ming.tseng, ulf.schlichtmann}@tum.de, mahdi.tala@unife.it
†Technical University of Munich, Germany ◇University of Ferrara, Italy

***Abstract*—Enabled by recent development in silicon photonics, wavelength-routed optical networks-on-chips (WRONoCs) emerge as an appealing next-generation architecture for the communication in multiprocessor system-on-chip. WRONoCs apply a passive routing mechanism that statically reserves all data transmission paths at design time, and are thus able to avoid the latency and energy overhead for arbitration, compared to other ONoC architectures. Current research mostly assumes that in a WRONoC topology, each initiator node sends one bit at a time to a target node. However, the communication parallelism can be increased by assigning multiple wavelengths to each path, which requires a systematic analysis of the physical parameters of the silicon microring resonators and the wavelength usage among different paths. This work proposes a mathematical modeling method to maximize the communication parallelism of a given WRONoC topology, which provides a foundation for exploiting the bandwidth potential of WRONoCs. Experimental results show that the proposed method significantly outperforms the state-of-the-art approach, and is especially suitable for application-specific WRONoC topologies.**

I. Introduction

To deal with the ever-increasing on-chip communication in multiprocessor system-on-chip (MPSoC), optical networks-on-chips (ONoCs) emerge as an appealing next-generation architecture. The wavelength-division multiplexing (WDM) technology enables ONoCs to transmit multiple optical signals on different wavelengths through a single waveguide [1], and thus to offer much higher bandwidth compared to conventional NoC systems that use electrical interconnects. Besides, current research has also demonstrated the advantages of ONoCs in energy-efficiency [2], [3], since the power-consumption in ONoCs is relatively distance independent [4].

This work focuses on wavelength-routed optical networks-on-chips (WRONoCs), which belong to a promising family of ONoCs. As their name suggests, WRONoCs route optical signals based on their wavelengths. For each pair of communicating nodes in the network, a dedicated signal path is statically reserved at design time, so that the initiator node can transmit data to the target node at any time, regardless of other on-chip communications in progress. In this manner, WRONoCs avoid the latency and energy overhead for arbitration, and are thus especially suitable for applications that require time predictability and/or performance guarantees [5], [6].

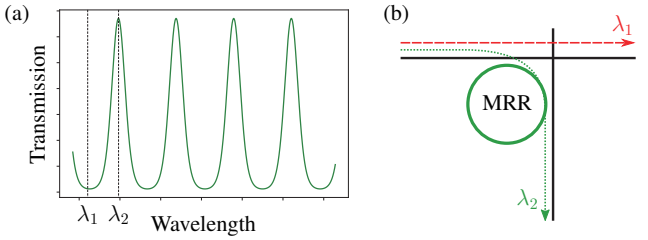

Figure 1: (a) The transmission spectrum of an MRR. (b) Signal routing mechanism supported by this MRR. Optical signals on different wavelengths are represented by dash lines in different colors.

The key network components that enable this passive routing mechanism are silicon microring resonators (MRR). An MRR consists of a looped waveguide, namely the microring, and a coupling mechanism to access the microring [7]. When an optical signal approaches an MRR, a resonance may or may not occur, dependent on the radius of the MRR and the wavelength of the signal. Specifically, if the optical path length of the MRR is an integer multiple of the wavelength of the signal[1], the signal will be "on-resonance" and coupled to the MRR; otherwise, the signal will be "off-resonance" and just ignore the MRR. For example, Figure 1(a) shows the periodic transmission spectrum of an MRR, where each of the four peaks represents a resonant wavelength. In particular, λ_2 is a resonant wavelength but λ_1 is not. Thus, as shown in Figure 1(b), when approaching this MRR through a nearby waveguide, a signal on wavelength λ_1 will just ignore the MRR, while a signal on wavelength λ_2 will be coupled to the MRR and switched to another direction.

With MRRs supporting different resonant wavelengths, collision-free signal paths can be established between every two communicating nodes in a WRONoC topology. For example, Figure 2(b) shows a GWOR topology [8] that applies MRRs of two different radii, the transmission spectra of which

[1]i.e., $\lambda_r = \frac{n_{\text{eff}} L}{m}$, where λ_r is the resonant wavelength, n_{eff} is a constant refractive index, L is the round trip length of the MRR, and m is a natural number.

978-1-7281-4123-7/20/$31.00 ©2020 IEEE

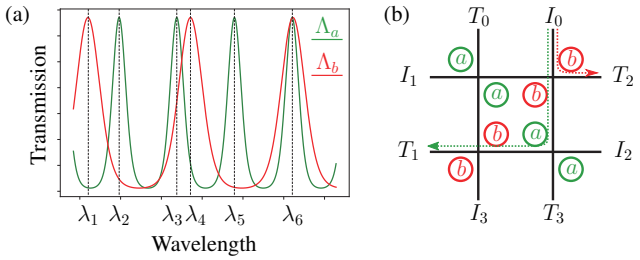

Figure 2: (a) The transmission spectra of two MRRs of different radii. (b) A GWOR topology. MRRs denoted by a have same spectrum as Λ_a, and MRRs denoted by b have the same spectrum as Λ_b. The signal paths (I_0,T_1) and (I_0,T_2) are represented by green and red dash lines, respectively.

are shown in Figure 2(a). In this topology, the signal paths from initiator I_0 to targets T_1 and T_2 can share the same output waveguide without data-collision, if the wavelength usage satisfies the following demands:

1) signals in path (I_0,T_1) are on-resonance with MRR a but off-resonance with MRR b; and
2) signals in path (I_0,T_2) are on-resonance with MRR b.

To this end, path (I_0,T_1) can use wavelengths λ_2 and λ_5, which are the resonant wavelengths of MRR a. Though λ_3 and λ_6 are also on-resonance with MRR a, they cannot be selected as they are (partially) on-resonance with MRR b. On the other hand, path (I_0,T_2) can use all the three resonant wavelengths λ_1, λ_4 and λ_6 of MRR b, since this path does not involve MRR a.

Currently, most WRONoC research assumes that each pair of initiator/target nodes communicates only with one bit at a time using a single wavelength. However, as we have seen, supported by the periodic transmission spectra of MRRs, multiple wavelengths can be available for a signal path. More specifically, in the topology shown in Figure 2(b), (I_0,T_1) can use two wavelength channels to transmit data simultaneously, while (I_0,T_2) can use three wavelength channels. Thus, the maximal achievable bit-level communication parallelism of (I_0,T_1) and (I_0,T_2) is 2 and 3, respectively.

Actually, we can expect higher communication parallelism in most state-of-the-art WRONoC topologies, since MRRs of 10-30μm radii can already support up to tens of resonant wavelengths in the typical band 1500nm–1600nm [9]. However, more resonant wavelengths do not necessarily lead to higher communication parallelism, because the available wavelengths in each signal path are also dependent on the degree of overlapping of the transmission spectra of the corresponding MRRs. Thus, to fully exploit the bandwidth potential of a WRONoC topology, a systematical optimization approach regarding the MRR usage in each signal path is necessary.

In this work, we formally model the passive routing behavior in WRONoCs, and propose a design automation approach to maximize the bit-level communication, by optimizing the selection of MRR-radii and the usage of wavelengths for data transmission in each signal path.

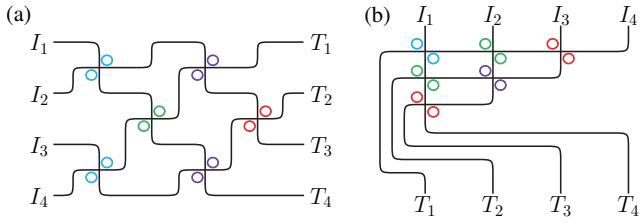

Figure 3: Representative WRONoC topologies that support full connectivity. MRRs of different radii are represented by circles in different colors.

II. Related Work

State-of-the-art WRONoC topologies can be classified into two categories:

1) General-purpose topologies that provide full connectivity, i.e., all initiators can send signals to all targets. One of the most recognized topologies in this category is the λ-router [10], where the MRR usage in each signal path is relatively balanced, as shown in Figure 3(a). Other representative topologies include GWOR [8][2], as shown in Figure 2(b), and Snake [11], as shown in Figure 3(b).
2) Application-specific topologies that provide customized connectivity, i.e., nodes that do not require data exchange are not connected. Such topologies can be reduced from general-purpose topologies [12] or synthesized with design automation approach [13], [14]. Due to the imbalance in the connectivity, the MRR usage varies a lot among different initiator/target pairs.

Current WRONoC research mostly neglects the communication parallelism and assumes that each initiator/target pair uses exactly one wavelength. [9] lays the groundwork for fault-free and maximally-parallel WRONoC design, which proposes the first formal methodology for selecting MRRs of proper radii to avoid data-collision and to increase the communication parallelism. The limitation of this method is that it overlooks the differences in MRR usage among the signal paths, and assumes that all MRRs are used equally. [15] points out that this assumption is overly pessimistic and results in an inaccurate estimation of the communication parallelism. To take full advantage of the bandwidth merits of WRONoCs, our work proposes a new design automation approach to maximize the communication parallelism considering the routing behavior in each signal path.

III. Problem Formulation

As shown in the above figures, state-of-the-art WRONoC topologies denote MRR-types as symbols, e.g. a, b, or color-

[2]GWOR topologies exclude self-connectivity, i.e., a node is not connected to itself.

ful circles. This abstraction says nothing about the physical features such as the radii of the MRRs and the exact value of wavelengths for data transmission in each signal path. These physical features, however, are essential parameters that characterize the communication parallelism of the topologies. Our target problem is to determine these physical features for a given symbolic WRONoC topology to achieve the maximized communication parallelism.

A. Input

- A WRONoC topology that defines a signal path for every pair of initiator/target nodes that require data transmission. In particular, the on- and off-resonance MRR-types in each signal path should be specified.
- A set of MRR-radii options. The radius of an MRR defines its resonant wavelengths, which can be calculated applying the methods proposed in [9] and [16].

B. Output

- The radius of each MRR type in the topology.
- The exact values of wavelengths for data transmission in each signal path.

C. Alternative Maximization Objectives

- The worst-case parallelism in a single path.
- The aggregate parallelism of all paths.

IV. Method

We propose an integer-linear-programming approach to explore this optimization problem in a comprehensive manner.

A. Input Abstraction

For a given input topology, we introduce a set $\mathcal{P}$ to denote the signal paths, and a set $\mathcal{M}$ to denote the various types of MRRs being used, where MRRs of the same type share the same radius. Based on this denotation, for each signal path $p\in\mathcal{P}$, we introduce two sets $\mathcal{ON}_p\subseteq\mathcal{M}$ and $\mathcal{OFF}_p\subseteq\mathcal{M}$ to denote the on- and off-resonance MRRs in this path.

Example 1:

The GWOR topology shown in Figure 2(b) can be denoted as the follows:

$$\mathcal{P}=\{(I_0,T_1),(I_0,T_2),(I_0,T_3),(I_1,T_0),(I_1,T_2),(I_1,T_3),\cdots\};$$
$$\mathcal{M}=\{a,b\};$$
$$\mathcal{ON}_{(I_0,T_1)}=\{a\},\quad \mathcal{OFF}_{(I_0,T_1)}=\{b\};$$
$$\mathcal{ON}_{(I_0,T_2)}=\{b\},\quad \mathcal{OFF}_{(I_0,T_2)}=\emptyset;\quad \cdots$$

Besides, we denote the input set of MRR-radii options as $\mathcal{R}$. For each radius $r\in\mathcal{R}$, we construct a sequence Ω_r to record its corresponding resonant wavelengths, where $\Omega_r^{(i)}$ denotes the i-th resonant wavelength supported by radius r.

Example 2:

If we consider the transmission spectra shown in Figure 2(a) as a set of MRR-radii options $\mathcal{R}=\{r_a,r_b\}$, the resonant wavelengths can be denoted as the follows:

$$\Omega_{r_a}=\{\lambda_2,\lambda_3,\lambda_5,\lambda_6\};\ \textit{where}\ \Omega_{r_a}^{(1)}=\lambda_2,\ \Omega_{r_a}^{(2)}=\lambda_3,\ \cdots$$
$$\Omega_{r_b}=\{\lambda_1,\lambda_4,\lambda_6\},\ \textit{where}\ \Omega_{r_b}^{(1)}=\lambda_1,\ \Omega_{r_b}^{(2)}=\lambda_4,\ \Omega_{r_b}^{(3)}=\lambda_6.$$

B. Radii Assignment

For each MRR-type $m\in\mathcal{M}$ and each radius option $r\in\mathcal{R}$, we introduce a binary variable b_r^m to represent whether MRRs of type m have radius r. To ensure that exactly one radius is assigned to each MRR type, we introduce the following constraint:

$$\forall m\in\mathcal{M}\quad \sum_{r\in\mathcal{R}} b_r^m=1. \tag{1}$$

Since MRRs of different types must have different radii, we introduce the following constraint to ensure that each radius is selected at most once:

$$\forall r\in\mathcal{R}\quad \sum_{m\in\mathcal{M}} b_r^m\leq 1. \tag{2}$$

C. Wavelength Assignment

The wavelength usage in a signal path is constrained by the resonant wavelengths of the MRRs in this path. Specifically, a wavelength ω can be used for data transmission in a path $p\in\mathcal{P}$, if it satisfies the following demands:

$$\begin{aligned} &\forall m\in\mathcal{ON}_p && \omega\ \textit{is on-resonance with}\ m;\\ \&\quad &\forall m'\in\mathcal{OFF}_p && \omega\ \textit{is off-resonance with}\ m'. \end{aligned}$$

Thus, if a path consists of at least one on-resonance MRR, the number of wavelengths assigned to this path cannot surpass the number of the resonant wavelengths supported by this on-resonance MRR. We record the maximum possible number of the resonant wavelengths supported by any MRR-radius option as a constant number n, i.e., $n:=\max_{r\in\mathcal{R}}|\Omega_r|$.

Example 3:

If we consider the MRR-radii options in Example 2, we will have:

$$n:=\max\{|\Omega_{r_a}|,|\Omega_{r_b}|\}=\max\{3,4\}=4;$$

Therefore, for each path $p\in\mathcal{P}$ with $\mathcal{ON}_p\neq\emptyset$, we introduce n continuous variables $(\omega_k^p)_{1\leq k\leq n}$ as wavelength placeholders for data transmission in path p. These variables can take any value between 0 and an upper bound defined by the given wavelength band. E.g. for the typical wavelength band 1500nm–1600nm, the upper bound of ω_k^p will be defined as 1600. If ω_k^p takes a non-zero value, it represents the exact value of a wavelength assigned to path p, and if ω_k^p takes 0, it represents an empty placeholder.

Example 4:

If we consider the topology in Example 1 and the MRR-radii options in Example 2, for each path, we will introduce 4 continuous variables:

For path (I_0,T_1): $\omega_1^{(I_0,T_1)},\omega_2^{(I_0,T_1)},\omega_3^{(I_0,T_1)},\omega_4^{(I_0,T_1)}$

For path (I_0,T_2): $\omega_1^{(I_0,T_2)},\omega_2^{(I_0,T_2)},\omega_3^{(I_0,T_2)},\omega_4^{(I_0,T_2)}$

…

We introduce the following constraints to ensure that if an on-resonance MRR m in path p selects a radius r, i.e. $\sum_{m\in\mathcal{ON}_p} b_r^m=1$, the wavelengths ω_k^p used in path p must be chosen from the resonant wavelengths specified by Ω_r:

$$\forall p\in\mathcal{P}\ \forall r\in\mathcal{R}\ \forall 1\le k\le|\Omega_r|:$$

$$\omega_k^p+\sum_{m\in\mathcal{ON}_p} b_r^m\cdot M\le\Omega_r^{(k)}+q_1\cdot M+M, \quad (3)$$

$$\omega_k^p-\sum_{m\in\mathcal{ON}_p} b_r^m\cdot M\ge\Omega_r^{(k)}-q_1\cdot M-M, \quad (4)$$

$$\omega_k^p+\sum_{m\in\mathcal{ON}_p} b_r^m\cdot M\le q_2\cdot M+M, \quad (5)$$

$$\omega_k^p-\sum_{m\in\mathcal{ON}_p} b_r^m\cdot M\ge -q_2\cdot M-M, \quad (6)$$

$$q_1+q_2=1, \quad (7)$$

where $M\gg\omega_k^p$ is an extremely large auxiliary constant, and q_1 and q_2 are two auxiliary binary variables. When none of the on-resonance MRR in path p selects the radius r, i.e., $\sum_{m\in\mathcal{ON}_p} b_r^m=0$, the inequations become tautology regardless of the values of q_1 and q_2. Thus, the above constraints only come into effect when the radius r is assigned to an on-resonance MRR in path p, i.e., $\sum_{m\in\mathcal{ON}_p} b_r^m=1$. In this case, constraints (3)–(6) can be simplified as the follows:

$$\omega_k^p\le\Omega_r^{(k)}+q_1\cdot M,\quad (3) \qquad \omega_k^p\le q_2\cdot M,\quad (5)$$

$$\omega_k^p\ge\Omega_r^{(k)}-q_1\cdot M,\quad (4) \qquad \omega_k^p\ge -q_2\cdot M.\quad (6)$$

According to constraint (7), either q_1 or q_2 must be 0 while the other must be 1. Thus, w_k^p takes either $\Omega_r^{(k)}$ or 0 as its value. Specifically, $\omega_k^p=\begin{cases}\Omega_r^{(k)} & \text{if } q_1=0\ \wedge\ q_2=1,\\ 0 & \text{if } q_1=1\ \wedge\ q_2=0.\end{cases}$

Furthermore, we need to ensure that if ω_k^p selects $\Omega_r^{(k)}$, $\Omega_r^{(k)}$ is off-resonance with all $m'\in\mathcal{OFF}_p$. To this end, we keep a safe distance Δd between $\Omega_r^{(k)}$ and the resonant wavelengths of the off-resonance MRRs. Specifically, we define a radius r' to be "conflicting" with $\Omega_r^{(k)}$, if there exists at least one $\omega_{r'}\in\Omega_{r'}$, so that $\Omega_r^{(k)}-\Delta d<\omega_{r'}<\Omega_r^{(k)}+\Delta d$.

a) Example 5

Consider the topology in Example 1 and the MRR-radii options in Example 2. And consider $p=(I_0,T_1)$, $r=r_a$, *and* $k=2$. *Thus, we are interested in whether variable* $\omega_2^{(I_0,T_1)}$ *can take value* $\Omega_{r_a}^{(2)}=\lambda_3$. *If* Δd *is set properly, we should find out that* $\lambda_4<\lambda_3+\Delta d$. *Since* $\lambda_4\in\Omega_{r_b}$, *we say* r_b *conflicts with* $\Omega_{r_a}^{(2)}$.

We collect all the MRR-radii options that are conflict with $\Omega_r^{(k)}$ into a conflict set $\mathcal{R}^{\text{conflict}}_{\Omega_r^{(k)}}$, and introduce the following constraint:

$$\forall m'\in\mathcal{OFF}_p \quad \sum_{r'\in\mathcal{R}^{\text{conflict}}_{\Omega_r^{(k)}}} b_{r'}^{m'}+q_2\le 1. \quad (8)$$

Thus, if an off-resonance MRR of type m' in path p is of a radius $r'\in\mathcal{R}^{\text{conflict}}_{\Omega_r^{(k)}}$, i.e., $\sum_{r'\in\mathcal{R}^{\text{conflict}}_{\Omega_r^{(k)}}} b_{r'}^{m'}=1$, q_2 will be forced to take 0, which will then prevent ω_k^p from taking the value $\Omega_r^{(k)}$, according to constraints (3)–(7).

Besides, since there are $n:=\max_{r\in\mathcal{R}}|\Omega_r|$ continuous variables $(\omega_k^p)_{1\le k\le n}$ as wavelength placeholders for each path p, where $n\ge|\Omega_r|$ for any radius $r\in\mathcal{R}$, we introduce the following constraints to set the redundant placeholders to 0:

$$\forall p\in\mathcal{P}\ \forall r\in\mathcal{R}\ \forall|\Omega_r|+1\le k\le n:$$

$$\omega_k^p+\sum_{m\in\mathcal{ON}_p} b_r^m\cdot M\le M, \quad (9)$$

$$\omega_k^p-\sum_{m\in\mathcal{ON}_p} b_r^m\cdot M\ge -M. \quad (10)$$

When r is not assigned to any on-resonance MRR in path p, i.e. $\sum_{m\in\mathcal{ON}_p} b_r^m=0$, the inequations (9) and (10) become tautology; otherwise, the redundant wavelength placeholders $(w_k^p)_{|\Omega_r|+1\le k\le n}$ will be forced to take 0.

D. Parallelism Maximization

The bit-level communication parallelism in each signal path $p\in\mathcal{P}$ can be regarded as the number of wavelengths assigned to this path. Thus, maximizing the communication parallelism in path p is equivalent to maximizing the number of continuous-variables ω_k^p with $1\le k\le n$ that take non-zero values.

We introduce a binary variable b_k^p for each wavelength placeholder ω_k^p to represent whether it takes a non-zero value. This is modeled with the following constraint:

$$\forall p\in\mathcal{P}\ \forall 1\le k\le n:\quad b_k^p\le\omega_k^p. \quad (11)$$

Thus, if ω_k^p takes 0, b_k^p will be forced to also take 0; and if ω_k^p takes a non-zero value, i.e. a value equal to a wavelength which is much bigger than 1, b_k^p will be free to take either 0 or 1. As we will see later, b_k^p will take 1 in the latter case, constrained by the optimization objective.

Based on the above constraints, we introduce two alternative optimization objectives:

- **Worst-Case Single-Path Parallelism** We denote the minimal number of wavelengths assigned to a single path

in the given topology as v_{worst}, which is modeled with the following constraint:

$$\forall p \in \mathcal{P}: \quad v_{\text{worst}} \leq \sum_{1 \leq k \leq n} b_k^p. \tag{12}$$

This constraint ensures that v_{worst} can only take a value smaller than or equal to the worst-case single-path parallelism, and we will see later that this is sufficient for modeling v_{worst}.

- **Aggregate Parallelism** We denote the sum of the wavelengths assigned to each path in the given topology as v_{total}, which is equivalent to $\sum_{p \in \mathcal{P}} \sum_{1 \leq k \leq n} b_k^p$.

Thus, the optimization model can be formulated as:

$$\begin{aligned} &\text{Maximize:} \quad \alpha \cdot v_{\text{worst}} + \beta \cdot v_{\text{total}}, \\ &\text{Subject to:} \quad (1) - (12), \end{aligned}$$

where α and β are adjustable weight coefficients. This optimization objective ensures that v_{worst} and b_k^p will take the largest possible value as long as α and β are not both set to 0. Thus, it is sufficient for constraints (11) and (12) to only specify the upper bounds.

V. Experimental Results

A. Inputs and Parameter Settings

We implement our approach in C++, and solve the integer-linear-programming model with an optimization solver called Gurobi [17]. We apply this approach to six WRONoC topologies of different scales. All of the topologies have been proposed in published research. Specifically, there are four general-purpose topologies (GWORs [8], λ-router [10], and Snake [11]) where the MRR usage among different signal paths is relatively balanced, and two customized topologies [13] where the MRR usage among different signal paths varies relatively more.

We provide MRR-radii options range between 5μm and 30μm, with an incremental step of 0.25μm, as suggested in [15].

We set the available wavelength band to be 1500nm – 1600nm, and we keep a minimal spacing between any two wavelengths for data transmission as 0.8nm, i.e., $\Delta d = 0.8$nm, since this is the minimum requirement to achieve < -20 dB of crosstalk for quality factor $\mathcal{Q} = 10^4$ [18].

B. Results Comparison

For a fair comparison between our approach and the state-of-the-art, we re-implement the radii-selection method proposed in [9] using integer-linear-programming methods. We run both approaches on a computer with two 2.6GHz CPUs.

Table I[3] shows the detailed input features and the optimization results. We apply our methods under two different settings by adjusting α and β in the maximization targets, as shown in Section IV-D. The worst-case parallelism oriented setting applies $\alpha = 1$ and $\beta = 0$, and the aggregate parallelism oriented setting applies $\alpha = 0$ and $\beta = 1$.

TABLE I: Input features and optimization results

Topology	$\lvert\mathcal{P}\rvert$	$\lvert\mathcal{M}\rvert$	Method	v_{worst}	v_{total}	$\#\omega$	Time
4×4 GWOR	8	2	state-of-the-art	16	136	34	1s
			v_{worst}-oriented	16	139	37	134s
			v_{total}-oriented	14	148	40	18s
4×4 Snake	12	4	state-of-the-art	10	120	40	15s
			v_{worst}-oriented	11	134	61	1116s
			v_{total}-oriented	10	172	62	2586s
4×4 λ-router	12	4	state-of-the-art	10	129	43	8s
			v_{worst}-oriented	10	120	46	1999s
			v_{total}-oriented	10	160	52	3410s
5×5 GWOR	16	4	state-of-the-art	10	168	42	23s
			v_{worst}-oriented	10	175	55	367s
			v_{total}-oriented	7	200	57	2215s
4×4 Custom	6	4	state-of-the-art	10	64	42	25s
			v_{worst}-oriented	16	103	34	79s
			v_{total}-oriented	13	106	33	270s
8×8 Custom	36	6	state-of-the-art	6	224	37	10s
			v_{worst}-oriented	6	244	48	2923s
			v_{total}-oriented	2	288	69	4192s

$|\mathcal{P}|$: the number of signal paths in the topology; $|\mathcal{M}|$: the number of MRR types in the topology; v_{worst}: the minimum single-path communication parallelism; v_{total}: the aggregate communication parallelism; $\#\omega$: the number of distinct wavelengths used for data transmission; Time: time for solving the ILP problem.

We can observe the following trends from the experimental results:

- The worst-case communication parallelism in a single path negatively correlates with the number of MRR-types. As the number of MRR types increases, the resonant wavelengths of the on- and off-resonance MRRs in a path is more likely to overlap, which then leads to the reduction of the available wavelengths for data transmission.
- Our approach outperforms the state-of-the-art approach in both optimization settings. Though the state-of-the-art approach can achieve the optimal results in worst-case single-path parallelism in some topologies (where the worst-case path consists of MRRs of all different types), it cannot fully exploit the aggregate parallelism. As visualized in Figure 4, our approach significantly improves the aggregate communication parallelism by $8\%-65\%$. In particular, based on our path-level routing model, our approach provides good support to the customized topologies where the MRR usage is relatively imbalanced among the signal paths.
- When applying our approach, the wavelength usage does not necessarily increase with the communication parallelism. For example, in the 4×4 customized topology, compared with the state-of-the-art approach, our approach

[3]In all the statistics in this section, we refer to a "signal path" as a path that consists of at least one on-resonance MRR. The wavelength usage in paths that do not consist of on-resonance MRRs can be freely selected from a rich set of options beyond the resonant wavelengths, and is thus not considered for the evaluation of the communication parallelism.

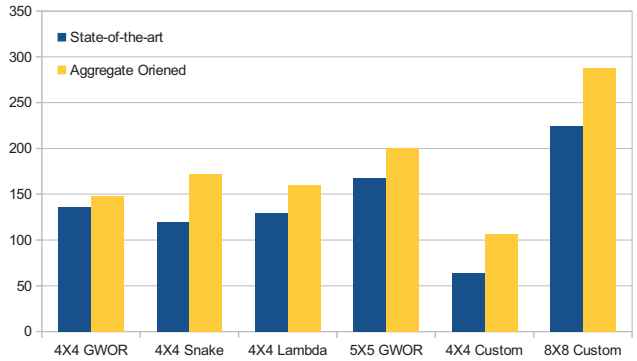

Figure 4: The comparison results of the aggregate communication parallelism between the state-of-the-art approach and our approach.

increases the aggregated parallelism from 64 to 106, and simultaneously reduces the wavelength usage from 42 to 33. This is because that we do not prohibit the usage of wavelengths that are on-resonance with multiple MRR-types, as long as these MRR-types do not exist in the same signal paths.

- The program run time of our approach positively correlates with the number of paths and MRR-types in a topology. Though current WRONoC topologies are mostly of small scales, it is essential to develop speed-up methods to prepare for future challenges.

VI. Discussion

The increase of the communication parallelism comes at the cost of using more wavelengths, which indicates more laser power consumption and more crosstalk noise. To alleviate these concerns, an intuitive approach is to increase the spacing between every two wavelengths, i.e., the Δd. However, a more systematic approach should be to consider the communication parallelism in topology design. As shown in the experiments, for application-specific topologies, it is possible to achieve high communication parallelism with modest wavelength usage. This direction calls for further research efforts.

VII. Conclusion

In this work we propose a mathematical modeling approach for the maximization of the communication parallelism of a given WRONoC topology. Compared with state-of-the-art approaches, we remove the assumption that all MRRs are used equally in all signal paths, and performs a comprehensive analysis of the routing behavior. As a result, our approach enables significant improvement on the bit-level communication parallelism, and is especially suitable for application specific WRONoC topologies.

References

[1] D. Vantrease, R. Schreiber, M. Monchiero, M. McLaren, N. P. Jouppi, M. Fiorentino, A. Davis, N. Binkert, R. G. Beausoleil, and J. H. Ahn, "Corona: System implications of emerging nanophotonic technology," *ACM SIGARCH Computer Architecture News*, vol. 36, no. 3, pp. 153–164, 2008.

[2] A. Shacham, K. Bergman, and L. P. Carloni, "Maximizing gflops-per-watt: High-bandwidth, low power photonic on-chip networks," in *P=ac2 Conference*, 2006, pp. 12–21.

[3] Z. Li, A. Qouneh, M. Joshi, W. Zhang, X. Fu, and T. Li, "Aurora: A cross-layer solution for thermally resilient photonic network-on-chip," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 23, no. 1, pp. 170–183, 2015.

[4] S. Werner, J. Navaridas, and M. Luján, "A survey on optical network-on-chip architectures," *ACM Computing Surveys (CSUR)*, vol. 50, no. 6, p. 89, 2018.

[5] M. Ortín-Obón, L. Ramini, V. Viñals-Yúfera, and D. Bertozzi, "A tool for synthesizing power-efficient and custom-tailored wavelength-routed optical rings," in *Proc. Asia and South Pacific Des. Autom. Conf.*, 2017, pp. 300–305.

[6] G. Nicolescu, S. Le Beux, M. Nikdast, and J. Xu, *Photonic Interconnects for Computing Systems: Understanding and Pushing Design Challenges*. River Publishers, 2017.

[7] 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*, vol. 6, no. 1, pp. 47–73, 2012.

[8] X. Tan, M. Yang, L. Zhang, Y. Jiang, and J. Yang, "On a scalable, non-blocking optical router for photonic networks-on-chip designs," in *Symp. Photonics and Optoelectronics (SOPO)*, 2011.

[9] A. Peano, L. Ramini, M. Gavanelli, M. Nonato, and D. Bertozzi, "Design technology for fault-free and maximally-parallel wavelength-routed optical networks-on-chip," in *Proc. Int. Conf. Comput.-Aided Des.*, 2016, pp. 3:1–3:8.

[10] M. Brière, 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 *Proc. Design, Automation, and Test Europe Conf.*, 2007, pp. 1084–1089.

[11] L. Ramini, P. Grani, S. Bartolini, and D. Bertozzi, "Contrasting wavelength-routed optical noc topologies for power-efficient 3d-stacked multicore processors using physical-layer analysis," in *Proc. Design, Automation, and Test Europe Conf.*, 2013, pp. 1589–1594.

[12] S. L. Beux, I. O'Connor, G. Nicolescu, G. Bois, and P. G. Paulin, "Reduction methods for adapting optical network on chip topologies to 3d architectures," *Microprocessors and Microsystems: Embedded Hardware Design*, vol. 37, no. 1, pp. 87–98, 2013.

[13] M. Li, T.-M. Tseng, D. Bertozzi, M. Tala, and U. Schlichtmann, "CustomTopo: A topology generation method for application-specific wavelength-routed optical nocs," in *Proc. Int. Conf. Comput.-Aided Des.*, 2018, p. 100.

[14] A. Truppel, T.-M. Tseng, D. Bertozzi, J. C. Alves, and U. Schlichtmann, "PSION: Combining logical topology and physical layout optimization for Wavelength-Routed ONoCs," in *Proc. Int. Symp. Phy. Des.*, 2019.

[15] M. Tala and D. Bertozzi, "Understanding the design space of wavelength-routed optical noc topologies for power-performance optimization," in *2018 IFIP/IEEE International Conference on Very Large Scale Integration (VLSI-SoC)*. IEEE, 2018, pp. 255–260.

[16] A. Parini, L. Ramini, G. Bellanca, and D. Bertozzi, "Abstract modelling of switching elements for optical networks-on-chip with technology platform awareness," in *Proceedings of the Fifth International Workshop on Interconnection Network Architecture: On-Chip, Multi-Chip*. ACM, 2011, pp. 31–34.

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

[18] K. Preston, N. Scherwood-Droz, J. S. Levy, and M. Lipson, "Performance guidelines for WDM interconnects based on silicon microring resonators," *CLEO: Science and Innovations*, 2011.