Dynamic Ring-based Multicast with Wavelength Reuse for Optical Network on Chips

Feiyang Liu†, Haibo Zhang†, Yawen Chen†, Zhiyi Huang†, Huaxi Gu‡
†Department of Computer Science, University of Otago, Dunedin 9016, New Zealand
‡State Key Laboratory of ISN, Xidian University, Xi'an 710071, China
Email:{feiyang,haibo,yawen,hzy}@cs.otago.ac.nz, hxgu@xidian.edu.cn

***Abstract*—Multicast communication, which widely exists in multicore systems, can occupy a large quantity of network resources and lead to severe traffic congestions. Optical Network on Chip (ONoC) is considered as a promising interconnection technology for future multicore processors, due to its remarkable advantages of high bandwidth capacity and transmission speed. However, traditional multicast schemes cannot well utilize the limited number of wavelength channels in ONoC. In this paper, we propose a novel multicast scheme, DWRMR, based on Dynamically-established and Wavelength-Reusable Multicast Rings. In DWRMR, the multicast ring, which connects the source core with all destinations via a cyclic routing path, is dynamically constructed for each multicast group. Then multicast packets are transmitted in the manner of single-send-multi-receive using only single wavelength. The same wavelength can also be reused in link-disjoint multicast rings. Most importantly, in our scheme the established multicast ring can be shared among cores in the same multicast group for interactive multicast traffic via optical-token arbitration, which avoids setting up exclusive multicast routing paths for each core. We formulate the multicast ring routing and wavelength allocation problem as an integer linear programming problem, and propose a heuristic algorithm that is able to accommodate more multicast rings under the wavelength limitation. Simulation results indicate that DWRMR can reduce more than 50% of packet delay with slight hardware cost, or require only half number of wavelengths to achieve the same performance, compared with existing schemes.**

***Index Terms*—Optical Network on Chip, Multicast Scheme, Wavelength Reuse, Ring, Routing and Wavelength Assignment**

I. Introduction

Continuous advances in high-performance computing systems and semiconductor technology keep promoting the development of Chip Multi-Processors (CMPs). At present, up to hundreds of micro-cores can be integrated in a processor chip, *e.g.,* 72-core Tilera TILE-Gx, 80-core Intel Teraflops, and 256-core Kalray MPPA. With the rapid increase of network traffic among these high speed cores, the high bandwidth and energy efficient communication architecture becomes especially important. Multicast communication, in which packets from one source core need to be addressed simultaneously to multiple destinations, intensively exists in CMPs due to the demand on cooperative computing and cache coherence [1][2]. For a 64-core system running PARSEC benchmarks with MESI coherence protocol [3], the multicast traffic takes a large percentage in each application as shown in Fig. 1(a), *e.g.,* multicast packets hold about 39% in average and 88% at maximum for each core in the blackscholes application. Moreover, an

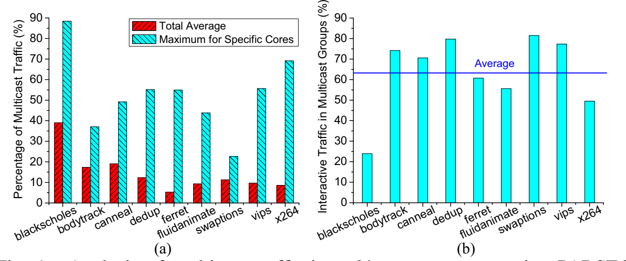


Fig. 1. Analysis of multicast traffic in a 64-core system running PARSEC benchmarks, (a) the percentage of multicast packets for each core; (b) the average percentage of interactive multicast packets in same multicast group.

important property of multicast traffic is that the cores in a *multicast group* (including the source and destination cores) have frequent *interactive multicast* communications within the multicast group itself. For instance, several cores share the same cache line while any core can change it and invoke the multicast for data update. As shown in Fig. 1(b), more than 60% of multicast packets are transmitted within the multicast group in average for different applications. If without an efficient multicast scheme, the source core needs to generate multiple copies and transmit them separately to each destination. That will lead to serious traffic congestion and low energy efficiency even with a small proportion of multicast (1%) [4]. Hence, the design of high performance and energy efficient interconnection architecture and routing protocol to support multicast communication plays an important role in CMPs.

Optical Network on Chip (ONoC), a chip-level optical interconnection technology, is promising for multicore processors [5][6]. ONoC has tremendous advantages in high performance inter-core communication, such as (*i*) low transmission delay and energy dissipation; (*ii*) extremely high bandwidth capacity by adopting Wavelength Division Multiplexing (WDM) which allows optical signals to be transmitted using different carrier wavelengths in the same waveguide (optical medium); (*iii*) very low signal decay and interference; (*iv*) area-compact and CMOS-compatible optical devices. However, compared with the electrical interconnects, it lacks optical processing logics and optical buffer, and supports only limited number of wavelength channels [7]. Generally, most existing ONoCs need to setup end-to-end optical routing path before communication by an electrical network [5], or use a global optical interconnection with the fixed wavelength assignment [8][9].

Currently, most multicast routing schemes can aggregate the copies of each multicast packet in some intermediate links,

978-1-5090-3531-1/16 $31.00 © 2016 IEEE
DOI 10.1109/MCSoC.2016.9

such as, tree-based [10] and path-based [11] schemes. They route multicast packets along the spanning trees/paths that originate from the source core and connect to all destinations. However, their performance and efficiency are constrained, since (i) the established multicast routing of trees/paths cannot be reused among the cores in the multicast group for interactive multicast traffic, since each tree/path is unidirectional from a specific source core to destinations; (ii) their performance closely depends on the distribution of destinations [11]. Moreover, these schemes cannot be used in ONoC directly, because they require multiple wavelengths or large fanout optical splitters to transmit different copies of multicast packet in parallel [12]. On the other hand, some ONoC architectures are inherently multicast enabled by transmitting multicast packets to each destination on a global wavelength-routed interconnect with fixed wavelength assignment [8][9]. However, they can only achieve non-blocking multicast routing for small scale ONoCs with tens of cores because of the constraint on the number of available wavelength channels.

The main objective of this paper is to design a high performance and efficient multicast scheme with the capability of reusing established multicast routing paths and wavelength channels. The contributions can be summarised as follows:

- We propose a multicast scheme, DWRMR, based on optical multicast ring and wavelength reuse. Each multicast ring is a dynamically constructed cyclic routing path to connect all the cores in a multicast group. Multicast traffic is transmitted in the manner of single-send-multi-receive.
- We design a flexible ONoC architecture which utilizes a control plane and a forwarding plane. The control plane manages multicast ring construction and wavelength assignment from a global perspective. The forwarding plane provides passive transmission of multicast packets.
- We devise an efficient multicast ring reuse scheme for interactive multicast traffic among cores in the same multicast group, based on the optical-token arbitration.
- We formulate the problem of multicast ring routing and wavelength allocation using Integer Linear Programming model, and design a heuristic algorithm to accommodate more multicast rings with limited wavelength channels.
- We evaluate DWRMR through extensive simulations using both synthetic traffics and realistic traces. Simulation results show that DWRMR can outperform existing schemes in communication performance and efficiency.

The rest of this paper is organized as follows. Section II introduces motivation and related work. Section III presents the network architecture and routing scheme. Problem formulation and a heuristic solution are given in Section IV. Section V evaluates the performance through simulations. Section VI concludes the paper and discusses the future work.

II. Related Work and Motivation

In the literature, most of the multicast routing schemes are *replication-based*. Generally, they can be classified into three categories: (i) *Unicast-Based*, the source core generates multiple packet copies and forwards them separately to destinations, as shown in Fig. 2(a); (ii) *Tree-Based* [10], it employs the spanning tree as routing path, in which the source core is the root and all destination cores are leaves, and copies of multicast packet are created in intermediate routers if encountering new branches, as shown in Fig. 2(b); (iii) *Path-Based* [4][11], the network is divided to several disjoint regions according to the distribution of destinations, and in each region a spanning path which locally connects to all destinations is formed. Then the source core generates a separate multicast copy for each region and transmits it along the predetermined path, as shown in Fig. 2(c). Exclusive wavelength assignment or optical splitters [13] need to be used to implement these routing schemes for parallel multicast transmission in ONoC. However, from the example given in Fig. 2, it can be seen that for the multicast communication among a source core {9} and destination cores {0,1,3,6,7,8,15}, replication-based multicast schemes need to establish multiple routing paths and transmit a separate optical packet on each path. Thus, they cannot effectively address the traffic conflicts during optical path setup, and low energy efficiency. *The most important drawback* is that an established routing is unidirectional and can merely be used by a fixed source. When other cores in the same multicast group have interactive multicast traffic, it needs to setup separate routing paths, *e.g.,* core {0} cannot reuse the routing paths of core {9}.

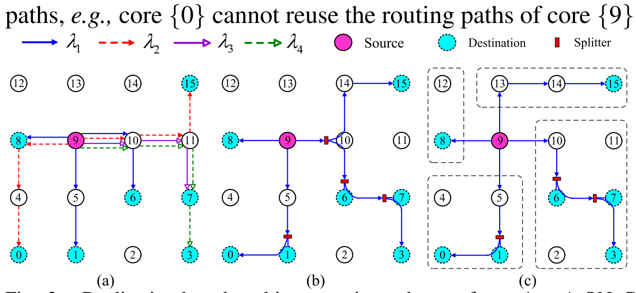


Fig. 2. Replication-based multicast routing schemes for a 4 × 4 ONoC, (a) unicast-based scheme with exclusive wavelengths assignment; (b) tree-based and (c) path-based schemes with optical splitters.

Some ONoC architectures employ a global interconnection, ring or crossbar, and fixed wavelength allocation to provide all-optical wavelength-routed communication [8][9]. These architectures are intrinsically multicast enabled by transmitting the multicast packet to every destination core via a different wavelength simultaneously. However, they are limited by the number of available wavelengths and only preferable for small scale ONoCs which connect tens of cores with sufficient wavelength channels. For larger scale ONoCs, the hierarchical interconnection and electrical-optical conversion can be used to extend these architectures, but that also introduces extra hardware cost and energy overhead. Moreover, with fixed wavelength assignment for all the connected cores, these architectures generally cannot achieve high wavelength efficiency.

The motivation of this paper is to design a high performance and efficient multicast scheme for ONoCs, in which the established multicast routing path and allocated wavelength can be reused among cores in the multicast group for interactive multicast, and multicast packets can be transmitted in the manner of single-send-multi-receive with single copy and wavelength. As shown in Fig. 3(a), we propose the ring-based multicast

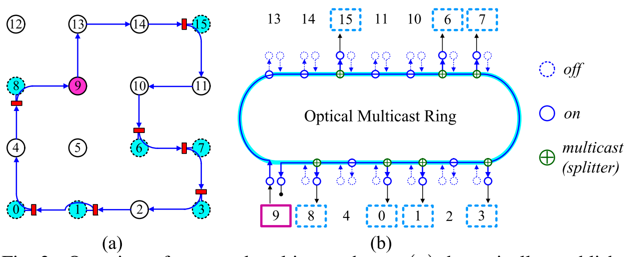


Fig. 3. Overview of proposed multicast scheme, (a) dynamically established multicast ring for routing; (b) single-send-multi-receive transmission.

scheme which connects all the cores in a multicast group using a dynamically-constructed cyclic routing path. The source core can transmit its multicast packets on the multicast ring with allocated wavelength, and every destination core can receive simultaneously by accessing the multicast ring as shown in Fig. 3(b). In this way, on an established multicast ring, each router that connects a destination core filters a small portion of optical signal and other intermediate routers transparently pass the optical signal. Benefited from the cyclic connection, every core in a multicast group can share the established multicast ring/wavelength, *e.g.,* after the multicast of core $\{9\}$, core $\{0\}$ can transmit its multicast packets by applying the multicast ring. The multicast ring reuse is conducted on the basis of an efficient optical token ring arbitration (details in Section III-C).

III. Architecture and Communication

A. Network Architecture

In this paper, we design a network architecture which can be logically divided into three planes: *core plane*, *control plane*, and *forwarding plane*, as shown in Fig. 4. This design exploits the idea of software defined network (SDN) in ONoC for the purpose of fast multicast ring construction. In this architecture, the core plane is the customer of multicast services, and it contains all micro-cores. The control plane is the kernel part. It collects multicast requests from the core plane, dynamically allocates multicast rings as routing paths, and configures the multicast rings in the forwarding plane. The control plane consists of a centralized Multicast Ring Allocator (MRA) and a cyclic arbitration channel. MRA allocates the multicast ring and carrier wavelength for each multicast group from the global perspective. The arbitration channel uses an optical ring to connect every core with MRA, and thus it can collect multicast requests and distribute configuration packets in a very fast manner. The forwarding plane provides passive transmission for multicast packets, and it is configured according to the allocation of the control plane. For physical implementation, the control plane and forwarding plane can be integrated in different layers and connected with the core plane using vertical links, *e.g.,* through silicon via (TSV). Thus, the architecture and communication in each plane can be designed and optimized separately, since their functions are independent.

1) Core Plane: Each core has a network interface (NI) for the coordination of communication, as shown in Fig. 5(a). Through the network interface, each core connects with an access point in the control plane and an optical router in the forwarding plane. Each NI consists of a local multicast table, a

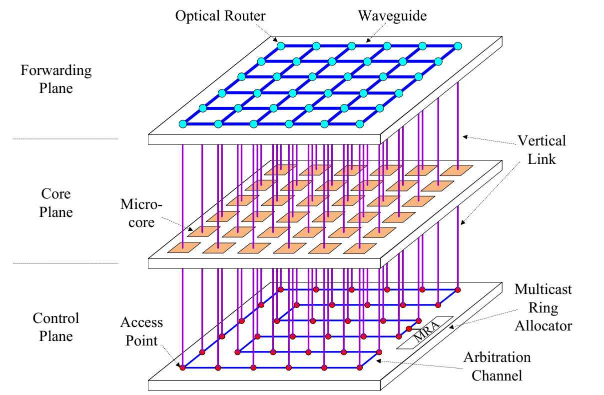


Fig. 4. Network architecture with three logical planes: core plane (micro-cores), control plane (centralized multicast ring/wavelength allocation), and forwarding plane (multicast packet delivery).

configuring unit, a laser power adjustment unit, and two pairs of electrical-to-optical (E/O) and optical-to-electrical (O/E) converters. The local multicast table records the information of established multicast rings which involve the current core, *i.e.,* addresses of the cores in each multicast group and the carrier wavelength. The configuring unit controls the connection state of corresponding optical router in the forwarding plane to establish/release the multicast ring. The laser power adjustment unit regulates the power intensity of optical signal according to the number of destinations before E/O conversion to improve communication reliability and power efficiency [14].

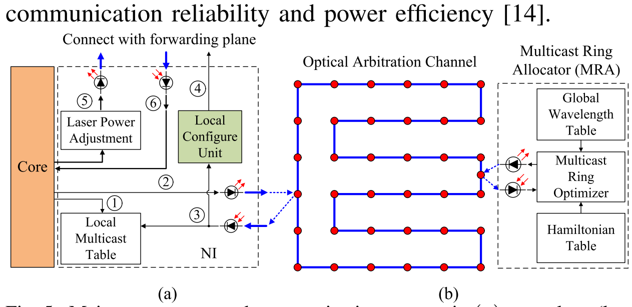
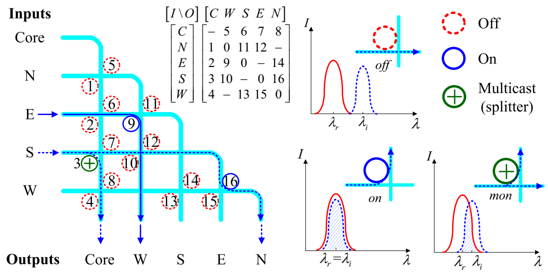


Fig. 5. Main components and communication process in (a) core plane (local configuration), and (b) control plane (centralized allocation).

2) Control Plane: It comprises an optical arbitration channel and a centralized multicast ring allocator, as shown in Fig. 5(b). The arbitration channel only transmits control packets, *e.g.,* multicast requests, multicast ring configuration and release packets. For fast multicast ring construction, the arbitration channel only needs to implement a N-to-1 and a 1-to-N optical buses. In N-to-1 optical bus, all the N cores use different wavelengths to send multicast requests to MRA in parallel; while in 1-to-N optical bus, MRA uses different wavelengths to simultaneously send configuration packets to the cores which are located on the allocated multicast ring.

In our design, MRA utilizes a heuristic algorithm to allocate the multicast ring and carrier wavelength. It consists of a global wavelength table, a Hamiltonian table, and a multicast ring optimizer. The global wavelength table stores the wavelength utilization of each optical link, while the Hamiltonian table records the connection of Hamiltonian cycle which is used as the primary multicast ring (details in Section IV-C). The multicast ring optimizer heuristically optimizes the mul-



(a) (b)

Fig. 6. Multicast-enabled router, (a) router principle; (b) switch status.

ticast ring both considering the minimization of peak wavelength utilization and the length of multicast ring, to accommodate as many multicast rings as possible with limited wavelengths. When a multicast ring is allocated, MRA updates the global wavelength table and sends out configuration packets.

3) Forwarding Plane: The forwarding plane is a passive network for multicast packets transmission. It consists of optical routers and waveguides in a $N\times N$ mesh. Each router has five pairs of input/output ports to connect with four neighbouring routers (north, east, south, west) and a core. As shown in Fig. 6, each router uses active microring resonators (MR) to implement configurable optical switching for dynamically established multicast rings. Each MR in the router has a unique resonant wavelength λ_r, which can be tuned by using thermal-optical or electrical-optical effects [6]. As shown in Fig. 6(b), if the resonant wavelength of MR is tuned to mismatch with input signal at *off-state*, $\lambda_i \neq \lambda_r$, the optical signal will transmit along the original waveguide; if the resonant wavelength is tuned to fully match with input signal at *on-state*, $\lambda_i = \lambda_r$, the optical signal will be coupled into MR and transmit along the other waveguide. We design the optical router for multicast by tuning the resonant wavelength to partially match with input signal at *multicast-state* (*mon*), $\lambda_i \cap \lambda_r$, so only a part of signal is coupled to MR and it can be output to both waveguides [14].

For multicast communication, the MRs in the optical router can be configured according to the matrix shown in Fig. 6(a). For example, if the optical signal inputs from *east* port and outputs to *west* port, MR $\{9\}$ is tuned to *on-state*; if optical signal inputs from *south* port and outputs to *core* and *north* port, MR $\{3\}$ is tuned to *multicast-state* and MR $\{16\}$ is tuned to *on-state*. It can be seen that only MRs $\{1,2,3,4\}$ may be tuned to *multicast-state*, thus this multicast-enabled router only introduces slight hardware cost for tuning MRs. Some communications, *e.g.,* from *north* to *west*, do not need to tune any MR, so they are labelled to $\{0\}$ in the matrix.

B. Communication Scheme

In our design, when a source core has multicast packets to transmit, firstly it checks in the local multicast table whether there already exists an established multicast ring connecting this multicast group, as shown in Fig. 5(a). If the multicast ring has been established, *i.e.,* the addresses of source and destination cores match with an existing multicast group in the local multicast table, the source core can apply for it according to the multicast ring reuse scheme (details in Section III-C); otherwise, the source core needs to send a request packet, which contains the addresses of the source and destination cores, to the control plane to establish a new multicast ring.

In control plane, the multicast request is transmitted from the corresponding access point to MRA in optical arbitration channel with a specific wavelength, as shown in Fig. 5(b). According to the distribution of destination cores and the global wavelength utilization, MRA calculates an optimized multicast ring as routing path and allocates an optical wavelength by using the heuristic algorithm. In the multicast ring, the source core connects to all the destinations with some intermediate cores. Then, MRA reserves the allocated wavelength in the global wavelength table, and sends the configuration packets to all the cores that are located in the new multicast ring (including intermediate cores) in parallel via different wavelengths in the arbitration channel to construct the multicast ring. The configuration packet contains the addresses of source and destination cores, and the carrier wavelength.

When the configuration packet is received by the core that on the multicast ring, as shown in Fig. 5(a), the configuring unit changes the connection state of the connected optical router to establish the multicast ring. According to the matrix in Fig. 6(a), there are three cases for optical router configuration: (i) in the source router, one MR is tuned to *on-state* to inject optical signal to multicast ring and another MR is tuned to *on-state* to collect the residual signal; (ii) in routers connecting to a destination core, one MR is tuned to *multicast-state* for the receiving of local core and another MR is tuned to *on-state* for the next hop; (iii) in intermediate routers, only one MR is tuned to *on-state* to an output for passing by the optical signal directly to the next hop. Moreover, if the core is the source or a destination core, the local multicast table will record the newly allocated multicast ring and wavelength.

After multicast ring construction and laser adjustment, the multicast packet is modulated to optical signal and transparently transmitted in forwarding plane. At the destination side, each NI receives the multicast packet and sends it to the destination core directly after O/E conversion. Finally, when the multicast traffic is over for all the cores in the multicast group, *e.g.,* if a multicast ring is idle for more than p cycles, the current source which owns the multicast ring sends a teardown packet to MRA. MRA frees the reserved wavelength in the global wavelength table and sends a release packet to the cores on the multicast ring. Then, the local multicast table in NIs and configured optical routers in forwarding plane are released.

C. Multicast Ring Reuse

In our scheme, multicast ring reuse is based on the optical token ring arbitration, and it is also implemented in the established multicast ring. The optical token represents the temporary ownership of the multicast ring. As shown in Fig. 7, the multicast ring has two states, *routing* and *arbitration*, which can interchange by tuning the MRs. For example in Fig. 7, a multicast ring connects 10 cores with the source core $\{1\}$ and 6 destination cores $\{3,5,6,7,9,10\}$. In Fig. 7(a), when the multicast ring is at the *routing* state, the router which connects to a destination core is tuned with one MR *on* and another

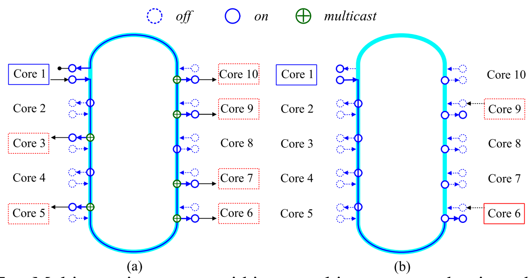


Fig. 7. Multicast ring reuse within a multicast group, by interchanging between two states: (a) multicast routing, and (b) optical token arbitration.

multicast, while intermediate routers are tuned with only one MR *on* to pass by the optical signal. In this way, the multicast packet can be received by every destination. If core $\{1\}$ has finished its multicast traffic, it multicasts a notification packet to inform the optical token arbitration will begin. Then, the multicast ring shifts to the *arbitration* state. As shown in Fig. 7(b), NIs in the destination cores configure their optical routers by (i) only tuning the MR which connects next-hop to *on* for passing by the token if it has no multicast traffic, *e.g.,* core $\{3,5,7,10\}$; (ii) only tuning the MR which connects local core to *on* if it has multicast traffic, *e.g.,* core $\{6,9\}$. Hence, when the original source core $\{1\}$ injects the optical token, the first encountered candidate, core $\{6\}$, can win the token. Then in the next clock cycle, core $\{6\}$ becomes the new source core to transmit its multicast packet and core $\{1\}$ shifts to a destination. Otherwise, if no core competes for the token, it is transmitted back to the original source, and the source core waits for q cycles to multicast a new notification for another round of token hunting until the source core has new multicast packet or one core wins the token. Due to the high speed of optical signal, the token arbitration only lasts for d_c cycles, which is determined by the length of multicast ring and optical propagation speed. No matter finally the optical token is received by a destination core or the original source, the multicast ring will automatically shift to the *routing* state after d_c cycles. Moreover, to prevent a multicast ring being occupied by one core in too long time, the optical token will be released after holding for a fixed time interval. Hence, every core in the multicast group has equal chance to use the multicast ring.

IV. Multicast Ring Construction Algorithm

The kernel problem in DWRMR is multicast ring construction and wavelength allocation. We formulate it using integer linear programming, and our optimization target is to enhance the wavelength efficiency by accommodating as many multicast rings as possible under the wavelength constraint. Without loss of generality, we focus on mesh-based ONoC in this paper.

A. Preliminary Definition

In a multicore system, we assume it has N^2 cores in $N\times N$ mesh topology. We can model it as a directed graph $G(V,E)$, in which each node v_i in set V corresponds to the combination of core $\{i\}$ and the connected router, and each link e_{ij} in set E is the unidirectional optical interconnect from node v_i to an adjacent node v_j. We label node v_i with a two-dimensional coordinates (x_i,y_i) where $0\le x_i,y_i\le N-1$. Then, the connectivity from node v_i to node v_j, denoted by χ_{ij}, can be represented by the existence of link e_{ij}, *i.e.,* $\chi_{ij}=1$ if $\exists e_{ij}\in E$, and $\chi_{ij}=0$ otherwise. Thus, in mesh we have

$$\chi_{ij}=\begin{cases}1, & |x_i\pm 1|=x_j,\ y_i=y_j \text{ or } x_i=x_j,\ |y_i\pm 1|=y_j\\ 0, & \text{otherwise.}\end{cases} \quad (1)$$

Since each multicast packet has a source core and multiple destinations, and the interactive multicast packets can be transmitted from any of these cores to the others, we define the *multicast group*, denoted by M_f, to be an arbitrary number of cores which have multicast communications. In our scheme, a multicast group is the basic object for the multicast ring routing and wavelength allocation. The multicast group M_f is represented as $M_f=\{S_f,D_f(k)\}$, where S_f and $D_f(k)$ are the source core and k destination cores respectively.

In ONoC architecture, we suppose each optical link can support a maximum of $W_{\max}$ wavelengths, and the bandwidth of each wavelength channel (>10 *Gbps*) is sufficient for multicast traffic. The set of wavelengths $\Lambda=\{\lambda_1,\lambda_2,...,\lambda_{W_{\max}}\}$ can be used in every node $v_i\in V$ and link $e_{ij}\in E$. Let $W_{ij}=\{\omega_{ijn}\}$ indicate the status of each wavelength in link e_{ij}, *i.e.,* if the n^{th} wavelength λ_n is available then $\omega_{ijn}=1$, otherwise $\omega_{ijn}=0$. Thus, the wavelength utilization of link e_{ij}, denoted by u_{ij}, is

$$u_{ij}=1-\frac{1}{W_{\max}}\sum_{n=1}^{W_{\max}}\omega_{ijn}. \quad (2)$$

B. Multicast Ring Model

In our scheme, the multicast ring is modelled as an unidirectional cyclic routing path in the form of an ordered node permutation, denoted by $P_f=\langle p_0,p_1,...,p_{l-1},p_l\rangle$, where $p_i\in P_f$ is the i^{th} node from the source, and $p_0=p_l=S_f$. Each multicast ring should satisfy the following constraints: (i) *spanning cycle*: the multicast ring should connect all the nodes of multicast group M_f with a limited number of intermediate nodes; (ii) *single wavelength*: all the links in a multicast ring should use the same wavelength λ_f; (iii) *no link overlapping*: each multicast ring can pass an optical link at most once.

In order to construct as many multicast rings as possible with a limited number of wavelengths in optical interconnects, our scheme is mainly concerned with two perspectives: (i) minimizing the peak wavelength utilization in each optical link; (ii) minimizing the length of each multicast ring. Thus, for each multicast ring P_f, we define an integer variable X_i to indicate the number of times it passes node v_i; and an integer variable Y_{ijf} to indicate the number of times it passes link e_{ij} using λ_f. Hence, in our scheme the problem of multicast ring routing and wavelength allocation for each multicast group M_f can be formulated using integer linear programming model as:

$$\textbf{Minimize} \sum_{e_{ij}\in P_f} u_{ij}, \quad (3)$$

subject to

$$X_i\ge\begin{cases}0, & \text{if } v_i\notin\{S_f,D_f(k)\};\\ 1, & \text{if } v_i\in D_f(k);\\ 2, & \text{if } v_i=S_f;\end{cases} \quad (4a)$$

$$l=\sum_{v_i\in V} X_i-1\le N^2; \quad (4b)$$

$$\chi_{ij} = 1, \quad \text{if } \ v_i = p_n, \ v_j = p_{n+1}; \tag{4c}$$

$$\omega_{ijf} = 1, \quad \text{if } \ v_i = p_n, \ v_j = p_{n+1}, \ \lambda_f \in \Lambda; \tag{4d}$$

$$0 \leq Y_{ijf} \leq 1, \quad \text{if } \ v_i = p_n, \ v_j = p_{n+1}, \ \lambda_f \in \Lambda. \tag{4e}$$

Constraints Eq. (4a) and Eq. (4b) state each multicast ring should start/end from/to the source node and connect with all destination nodes with a limited length l. Eq. (4c) indicates any adjacent nodes in the multicast ring should be physically connected. Eq. (4d) is the wavelength limitation that λ_f is available for all the links in a multicast ring. Eq. (4e) states the multicast ring cannot pass an optical link more than once.

Note that our model aims to minimize the summation of wavelength utilization of optical links in each multicast ring, instead of minimizing the multicast ring's length or peak wavelength utilization, thus it can prevent early wavelength exhaustion in some popular links. For example, $M_f = \{v_a, v_b, v_c\}$, if a cyclic path $P_1 = \langle v_a, v_b, v_c, v_d, v_a \rangle$ has the minimal length while path $P_2 = \langle v_a, v_b, v_e, v_c, v_f, v_a \rangle$ has the minimal summation of wavelength utilization, *i.e.*, $u_{ab}+u_{bc}+u_{cd}+u_{da} > u_{ab}+u_{be}+u_{ec}+u_{cf}+u_{fa}$, and wavelengths λ_1 and λ_2 are both available for P_1 and P_2, then our model selects P_2 as multicast ring and λ_2 as carrier wavelength. With the limited number of wavelengths, our scheme can hold more multicast rings.

C. Heuristic Allocation Scheme

We propose a heuristic algorithm for multicast ring routing and wavelength allocation considering computation efficiency, because the ILP model for similar problem is NP-complete in general [12]. In our algorithm, we use the precomputed Hamiltonian cycle as the primary multicast ring, then map the multicast group M_f on it, and heuristically optimize the connections between any adjacent nodes by replacing them with paths having the minimal overall wavelength utilization.

Hamiltonian cycle is a spanning cycle that passes all the nodes in a given network exactly once. In a $N \times N$ ONoC, two adjacent nodes are connected by two unidirectional optical links. Thus, when N is even, there are two link-disjoint Hamiltonian cycles as shown in Fig. 8(a); while when N is odd, there is only one link-disjoint spanning cycle which needs to pass at least one node twice. That is because in a cyclic channel the number of links in positive directions ($x+$ or $y+$) should equal to the number in negative directions ($x-$ or $y-$). In this paper, we use the Hamiltonian/spanning cycle as the primary multicast ring in our algorithm. Without loss of generality, we focus on the architecture with an even number N.

As shown in Fig. 8(a), two link-disjoint Hamiltonian cycles which in counter-clockwise and clockwise directions are denoted by H^- and H^+ respectively. Specifically, in the counter-clockwise Hamiltonian cycle H^-, we have the relationship between the address of any node (x_i, y_i) and its relative position t counting from source node $(x_s, y_s) = (0, 0)$,

$$t = \begin{cases} (N^2 - y_i)\%(N^2), & x_i = 0; \\ y_i \times (N-1) + x_i, & 0 < x_i < N, \ y_i\%2 = 0; \\ y_i \times (N-1) + N - x_i, & 0 < x_i < N, \ y_i\%2 = 1. \end{cases} \tag{5}$$

For example, in Fig. 8(a), $t = 2, 6, 9, 14$ when $(x_i, y_i) = (2,0)$, $(1,1)$, $(3,2)$, $(0,2)$. From Eq. (5), it can be seen that the relationship between the address of any node and its relative position in Hamiltonian cycle is uniquely determined. The relationship in the clockwise Hamiltonian cycle H^+ or counting from any source node (x_s, y_s) can be achieved by reversing or shifting operation, such as the Hamiltonian cycle H^+ starting from source node $(1, 2)$ as shown in Fig. 8(b).

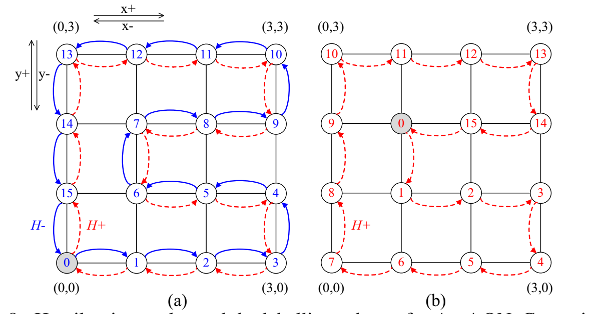


Fig. 8. Hamiltonian cycles and the labelling scheme for 4×4 ONoC, starting from the source core of (a) $(x_s, y_s) = (0, 0)$ in H^-, and (b) $(x_s, y_s) = (1, 2)$ in H^+, respectively.

The proposed heuristic algorithm operates in three steps. (i) *Mapping and dividing*. For a new multicast group M_f, the multicast ring P_f is initialized as the Hamiltonian cycle H that starts from the source node S_f, while H^- and H^+ are employed alternatively for load-balance. The destination nodes $D_f(k)$ are mapped on H according to their relative positions as in Eq. (5). Then P_f is naturally divided by the source node and k destination nodes into $k+1$ segments, $P_f = \{P_0, ..., P_k\}$. (ii) *Segment optimization*. For each segment path P_i in P_f, we search an alternative P^* with the minimal overall wavelength utilization u^*. Firstly, all alternative paths for P_i (not longer than P_i) are calculated into $\{P_m^*\} = \{P_1^*, P_2^*, ...,\}$, and each P_m^* is compared with the current optimal P^*. When an alternative P_m^* has lower wavelength utilization than the current optimal alternative, *i.e.*, $u_m^* < u^*$, then P_m^* replaces P^*. Note that if multiple alternatives have the minimal wavelength utilization u^*, the one with the minimal length l_m^* is selected as P^*. Then, the set of available wavelengths W_i^* is achieved by checking the common free wavelengths in $\forall e_{ij} \in P^*$. (iii) *Assembling*. The final multicast ring P_f is assembled from all the local optimizations, and the carrier wavelength λ_f is randomly chosen from the intersection of available wavelengths sets W_i^* of all segments. If there is no available wavelength λ_f for a multicast ring P_f, the calculated connection of P_f will be stored and wait for an available wavelength to be released.

V. Performance Evaluation

A. Simulation Setup

We evaluate DWRMR through extensive simulations using both synthetic traffics and realistic traces, and compare it with existing multicast schemes, including unicast-based routing (UM), tree-based routing (TM), and path-based routing (PM). We develop a simulation platform based on Noxim [15] to implement DWRMR and other schemes. Unicast communication, which can be done using XY routing [5] and with a separate set of wavelengths, is not simulated in this paper. To make fair comparison, we use the same parameter settings for all the schemes, as well as similar processes in optical path construction for other schemes, *i.e.,* centralized routing and wavelength allocation instead of hop-by-hop reservation. The parameter settings are summarized in Table I. The network size is set to 8×8 in mesh topology. In each optical link, the number of available wavelengths is set from 16 to 64. The channel bandwidth of all the optical devices, including E/O&O/E converters and optical routers, is set to 10 *Gbps/wavelength*. The system clock works at the frequency of 5 *GHz*. With this clock frequency, the propagation speed of optical signal is 8 *routers/cycle* [16]. The processing delay for centralized routing and wavelength allocation takes 2 *cycles* in all the schemes. E/O and O/E conversions both take 1 *cycle*. Each multicast packet has a constant size of 64 *bits*.

TABLE I
SIMULATION SETTINGS

Optical		Electrical	
Maximal wavelengths	16, 32, 64	Clock frequency	5 *GHz*
Channel bandwidth	10 *Gbps*/λ	Processing delay	2 *cycles*
Transmission speed	8 *routers/cycle*	E-O conversion delay	1 *cycle*

In the simulation, we evaluate the performance of different multicast schemes in terms of average packet delay and network throughput. Multicast packet delay is defined as the time interval from the source core generating a new multicast packet until the packet being transmitted in the forwarding plane to all of destinations. Network throughput is defined as the average volume of multicast packets received in each core over a time period. In synthetic simulations, we assume multicast traffic is subjected to following distribution: (i) each core generates multicast packets independently with a data rate of θ *packets/cycle*, which follows Poisson distribution, $\theta \in [0,1]$[4]; (ii) the number of destinations in each multicast packet follows Normal distribution, $k \sim N(\mu_k, \sigma_k)$, μ_k and σ_k are the expected value and standard deviation; (iii) in each multicast packet, the source and destinations are distributed uniformly; (iv) the probability of interactive multicast, *i.e.,* packets transmitted in the multicast group from different cores, denoted by τ, also follows Poison distribution, $\tau \in [0,1]$. In trace based simulation, we extract multicast traffic from realistic traces for a 64-core system running PARSEC benchmark [3].

B. Simulation with Synthetic Traffics

1) Comparison with Different Multicast Schemes: In this set of simulations, the number of available wavelengths is 64. The number of destinations for each multicast packet follows Normal distribution $N(16,5)$. In the multicast traffic, there is $\tau = 0.1$ percent of interactive multicast in the same multicast group. It can be seen from Fig. 9 that our scheme outperforms other multicast routing schemes in both average packet delay and throughput at any data rate. In Fig. 9(a), when the data rate is low, *e.g.,* 1 *Gbps/core*, the average packet delay for DWRMR is only 30.0 *ns*, whereas it can reach 44.0, 39.8, and 36.3 *ns* for other three schemes respectively. As the data rate increases, the average packet delay using all these schemes increases as well, since there are more conflicts on routing paths and wavelengths with more multicast traffic on

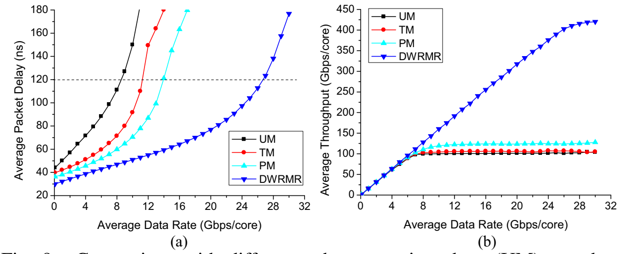


Fig. 9. Comparison with different schemes, unicast-bast (UM), tree-based (TM), and path-based (PM), in average (a) packet delay and (b) throughput.

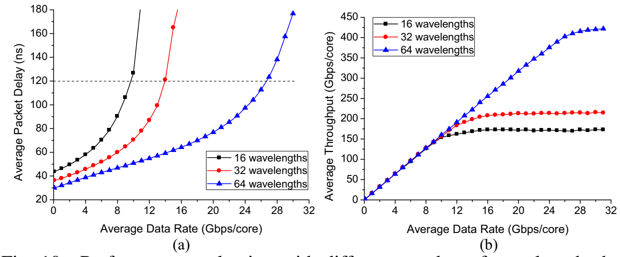


Fig. 10. Performance evaluation with different number of wavelength channels for DWRMR, in average (a) packet delay and (b) throughput.

the network. However, our scheme has the lowest increasing speed. If we define the maximal acceptable multicast delay to be 120 *ns*, then the maximum data rates that can be tolerated in each scheme are about 8.5, 11.5, 13.5, and 27.5 *Gbps/core* respectively. It indicates DWRMR can achieve more than twice multicast capacity of other multicast schemes with the same delay requirement. The reasons for much lower packet delay and higher maximal data rate in DWRMR are (i) the established multicast ring and allocated wavelength can be reused among cores in the same multicast group, while in other schemes the established routing path cannot be reused, and it needs to establish a new multicast routing for each core and each multicast frequently; (ii) the multicast ring routing and wavelength allocation in DWRMR minimize both the length and accumulated wavelength utilization, while other schemes greedily route multicast packets along the shortest path from the source core to the destinations, which can lead to unbalanced wavelength utilization, *i.e.,* no available wavelength in some links. In Fig. 9(b), the average throughput tends to linearly increase before the network get saturated due to wavelength limitation. Similarly, DWRMR can achieve much higher throughput than other schemes.

2) Influence of the Number of Wavelengths: As shown in Fig. 10, we evaluate the influence on different number of wavelengths for DWRMR. In the simulation, we set the maximal number of available wavelengths $W_{\max}$ to be 64, 32, and 16 respectively. It is obvious that DWRMR can achieve much lower packet delay and higher throughput with more available wavelength channels. The most important discovery is that by comparing with the simulation results in Fig. 9, DWRMR can even has better performance with 32 wavelengths than other multicast schemes with 64 wavelengths. That indicates DWRMR has the ability to efficiently utilize the limited wavelengths to accommodate more multicast communications. Moreover, if the multicore system has a fixed performance requirement on multicast communication, DWRMR can use only

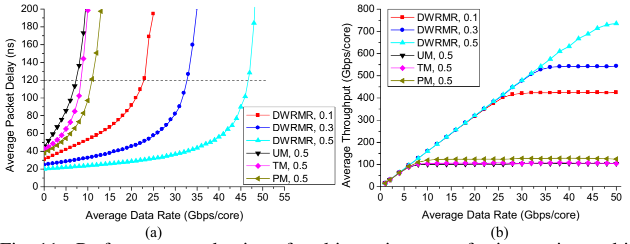


Fig. 11. Performance evaluation of multicast ring reuse for interactive multicast traffic, $\tau = 0.1, 0.3, 0.5$, in average (a) packet delay and (b) throughput.

half number of wavelengths than other schemes. It can greatly reduce the demand of on-chip lasers, E/O&O/E converters, and microring resonators. Due to the complexity of optical devices, this reduction is more considerable compared to the hardware cost on recording multicast rings and wavelengths.

3) Multicast Ring Reuse: DWRMR has the advantage of reusing established multicast rings and wavelengths. In this set of simulations, we evaluate the performance with different percentage of interactive multicast traffic in the same multicast group, $\tau = 0.1, 0.3, 0.5$. It can be seen from Fig. 11 that (i) for DWRMR scheme, since multicast ring and wavelength can be reused for interactive multicast in the multicast group, the average delay for multicast ring allocation and configuration is significantly reduced, and the probability of conflicts on wavelength allocation can also be decreased, thus the higher percentage of interactive multicast traffic is, the more significant improvement on performance can be achieved; (ii) for other schemes, since the established multicast routing path is released just after the source core finishes its multicast traffic, while other cores cannot reuse it, thus interactive multicast traffic has little influence on them. Note that when $\tau = 0.1$, 0.3, 0.5, the maximal achievable data rates of DWRMR with a delay constraint of 120 *ns* are more than 2.2, 3.2, and 4.6 times higher than other schemes, while the maximal throughput are 3.4, 4.4, and 5.9 times higher than them respectively.

C. Simulation with Realistic Traces

Fig. 12 gives the simulation results for different multicast schemes with realistic traces. It can be seen that the average packet delay is significantly reduced in all applications using DWRMR. In average, the packet delays achieved by using other multicast schemes are 108.2 *ns*, 48.9 *ns*, and 35.8 *ns* respectively, while DWRMR can reduce it to only 14.9 *ns*, which is 50% less than even the best of other schemes. That is because, as demonstrated in Fig. 1(b), a large proportion of multicast communication ($>60\%$) in the multicore system is interactive multicast transmitted within the same multicast group, thus the advantage of multicast ring reuse of DWRMR can be well exploited in these applications. Other multicast schemes have no such ability and need to spent more time on establishing multicast routing paths. Thus, DWRMR can achieve much better performance as well.

VI. Conclusions

This paper proposes a novel multicast scheme, DWRMR, for Optical Network on Chips. It facilitates multicast communication by combining the wavelength reusable routing with

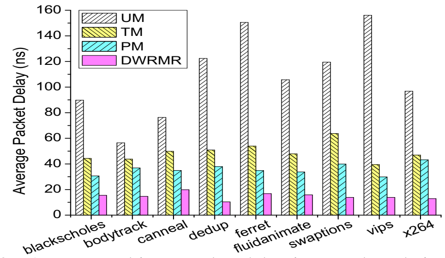


Fig. 12. Average multicast packet delay in trace-based simulations.

dynamically established optical multicast ring. Compared with existing multicast schemes, our scheme has following advantages: (*i*) established multicast ring can be reused within the same multicast group, thus it avoids the overhead of setting up exclusive routing paths for each multicast; (*ii*) each multicast ring is dynamically constructed, so that the wavelength assignment is more flexible and the same wavelength can be reused in link-disjoint multicast rings; (*iii*) multicast packets are transmitted in the multicast ring using only single copy and wavelength. Our future work is to further analyse the hardware cost and power consumption of the proposed multicast scheme, explore the implementation for dynamically-varied multicast groups and the implementation in other topologies and larger scale networks, and investigate the optimal multicast ring allocation scheme.

References

[1] S. Rodrigo, J. Flich, J. Duato, and M. Hummel, "Efficient unicast and multicast support for CMPs," in *IEEE/ACM Proc. MICRO*, 2008, pp. 364–375.

[2] A. Karkar, T. Mak, K. Tong, and A. Yakovlev, "A survey of emerging interconnects for on-chip efficient multicast and broadcast in many-cores," *IEEE Circuits Syst. Mag.*, vol. 16, no. 1, pp. 58–72, 2016.

[3] J. Hestness, B. Grot, and S. Keckler, "Netrace: dependency drive trace-based network-on-chip simulation," in *IEEE Proc. NoCArc*, 2010.

[4] S. Ma, N. Jerger, and Z. Wang, "Supporting efficient collective communication in NoCs," in *IEEE Proc. HPCA*, 2012, pp. 1–12.

[5] A. Shacham, K. Bergman, and L. Carloni, "Photonic Networks on Chip for future generations of chip multiprocessors," *IEEE Trans. Comput.*, vol. 57, pp. 1246–1260, Sep. 2008.

[6] C. Batten, A. Joshi, V. Stojanovic, and K. Asanovic, "Designing chip-level nanophotonic interconnection networks," *IEEE Trans. Emerg. Sel. Topics Circuits Syst.*, vol. 2, no. 2, pp. 137–153, 2012.

[7] K. Preston, N. Droz, J. Levy, and M. Lipson, "Performance guidelines for WDM interconnects based on silicon microring resonators," in *IEEE Proc. CLEO*, May 2011, pp. 137–153.

[8] I. O. Connor, "Optical solutions for system-level interconnect," in *ACM Proc. SLIP*, 2004, pp. 79–88.

[9] S. L. Beux, J. Trajkovic, I. O'Connor, G. Nicolescu, G. Bois, and P. Paulin, "Optical ring Network-on-Chip (ORNoC): architecture and design methodology," in *IEEE Proc. DATE*, 2011, pp. 1–6.

[10] F. Samman, T. Hollstein, and M. Glesner, "Adaptive and deadlock-free tree-based multicast routing for Networks-on-Chip," *IEEE Trans. Very Large Scale Integr. Syst.*, vol. 18, no. 7, pp. 1067–1080, 2010.

[11] M. Ebrahimi, M. Daneshtalab, P. Liljeberg, J. Plosila, J. Flich, and H. Tenhunen, "Path-based partitioning methods for 3D Networks-on-Chip with minimal adaptive routing," *IEEE Trans. Comput.*, vol. 63, no. 3, pp. 718–733, Mar. 2014.

[12] L. Gong, X. Zhou, X. Liu, W. Zhao, W. Lu, and Z. Zhu, "Efficient resource allocation for all-optical multicasting over spectrum-sliced elastic optical networks," *IEEE/OSA J. Opt. Commun. Netw.*, vol. 5, no. 8, pp. 836–847, Aug. 2013.

[13] R. Morris, E. Jolley, and A. Kodi, "Extending the performance and energy-efficiency of shared memory multicores with nanophotonic technology," *IEEE Trans. Parallel Distrib. Syst.*, vol. 25, no. 1, pp. 83–92, Jan. 2014.

[14] B. Fu, Y. Han, H. Li, and X. Li, "Accelerating lightpath setup via broadcasting in binary-tree waveguide in optical NoCs," in *IEEE Proc. DATE*, 2010, pp. 933–936.

[15] F. Fazzino, M. Palesi, and D. Patti, "Noxim: the Network-on-Chip simulator," in *https://github.com/davidepatti/noxim*, 2014.

[16] J. Liu, J. Yang, and R. Melhem, "Gasolin: global arbitration for streams of data in optical links," in *IEEE Proc. IPDPS*, 2015, pp. 93–102.