

Dynamic Home Circuit Construction for Datacenter Networks Using LOBS-HC Ring

Wan Tang^{1*}, Bo Yi¹, Ximi Yang¹, Jingcong Li²

¹ College of Computer Science, South-Central University for Nationalities
Wuhan 430074, China

[e-mail: tangwan@scuec.edu.cn; yibobooscar@gmail.com; yangximin@scuec.edu.cn]

² College of Information Engineering, Minzu University of China
Beijing 100081, China

[e-mail: jcli@pku.org.cn]

*Corresponding author: Wan Tang

Received December 6, 2014; revised March 22, 2015; accepted April 19, 2015; published May 31, 2015

Abstract

Optical switching will be applied in datacenter networks because electronic switching is costly and power-consuming. In this paper, considering the ring-based interconnection using optical switching in the core of a datacenter, we study the home circuit (HC) construction for the labeled optical burst switching with home circuit (LOBS-HC), a new paradigm trying to share wavelengths among the HCs from the same source. In particular, aiming to construct HCs dynamically and properly, a scheme named optimal path matching and symmetric HC matching (OPM-SHM) is proposed. The main idea of OPM-SHM is to dynamically construct HCs by sharing wavelength(s) not only among the same-source HCs but also with symmetric HCs which have different sources other than the original LOBS-HC features. The simulation results demonstrate that OPM-SHM achieves better performance than some other methods in terms of burst loss rate and wavelength utilization of physical links. More specially, it maintains good load balancing for the datacenter network using an LOBS-HC ring. In addition, due to the symmetric feature of SHM, the proposed scheme can decrease the upper bound of the average hop count of the routing paths to half of the ring size.

Keywords: Datacenter, optical switching, labeled optical burst switching with home circuit (LOBS-HC), ring networks, wavelength utilization

A preliminary version of this paper appeared in ChinaCom, August 14-16, 2014, Maoming, China. This version improves the idea of the proposed scheme OPM-SHM significantly, and evaluates it on an intra-datacenter LOBS-HC-based network simulation platform under dynamic network traffic. This work is supported by the National Natural Science Foundation of China (NSFC) (No. 61103248) and the State Key Laboratory of Advanced Optical Communication Systems & Networks, China.

1. Introduction

Cloud computing and big-data processing are driving the emergence of large datacenters [1, 2]. A huge datacenter consists of tens to hundreds of thousands of pods, each of which contains thousands of servers. The dynamic communication traffic inside one datacenter is huge and burst. So the transmission of huge data urgently needs fast switching technologies. However, electronic switching consumes a significant amount of power and also requires a large number of transmitters. Optical switching can lower the high cost and power consumption greatly [3]. But the mature optical switching, i.e. optical circuit switching (OCS), has not been able to handle burst data well and utilize bandwidth effectively so far [4].

Fortunately, a novel paradigm called labeled optical burst switching with home circuit (LOBS-HC) was proposed in 2010 [5]. LOBS-HC aims to provide the same quality of service (QoS) to traffic as in OCS while maintaining as fast a switching speed as in optical burst switching (OBS) [6, 7] by building the home circuit (HC). One HC can be used by not only its own traffic flow but also other traffic flows, and HCs with the same source may share one or more wavelengths. By good management of the data cache, and synchronization and transmission in the source nodes, LOBS-HC can ensure that the in-profile bursts from the same source will not collide with each other, and the overall bandwidth requirement will not exceed the capacity of the links.

Since bidirectional ring topology is common in datacenter networking [8], an all-optical ring network based on LOBS-HC without wavelength conversion is considered in this paper. Additionally, shortest path first (SPF), a simple but effective way verified in [9, 10], is applied in bidirectional ring routing. Furthermore, considering the grouping of the same-source HCs, we extend the routing and wavelength-assignment (RWA) issue, which is common in both the OCS and OBS networks [11], to a grouping and wavelength-assignment (GWA) issue in the LOBS-HC networks.

Generally, there are two GWA categories, static GWA (SGWA) and dynamic GWA (DGWA). In SGWA, the traffic is static and the entire connection requests are known in advance. Based on this, SGWA can be divided into two independent modules: grouping and wavelength assignment. The grouping module takes charge of allocating and placing HCs into groups, and the wavelength assignment module is responsible for assigning wavelengths to the groups in a global fashion. On the other hand, in the DGWA, the traffic arrives at random. An HC is set up for the connection request from a dynamically arriving flow, and it will be released as long as all bursts of the flow have arrived at the destination. Unlike in SGWA, grouping and wavelength assignment cannot be implemented separately in DGWA. Currently, several proposals regarding the SGWA of LOBS-HC have been studied [12-15]. However, to our knowledge, most of the existing research does not focus on the DGWA.

In this paper, we propose a DGWA scheme named optimal path matching and symmetric HC matching (OPM-SHM). For each arrival flow, OPM-SHM tries to construct an HC in two stages. First, it builds a new HC that shares wavelength(s) with other HCs depending on the OPM. If it fails, it applies SHM to re-assign an unused wavelength to construct the HC as the second stage. In particular, when a new flow arrives, matching degrees used to measure the coverage between the route of the flow and a same-source group whose spare bandwidth is enough for the flow, will be calculated first. The wavelength occupied by a group having the maximum matching degree will be chosen to construct an HC. If none of the groups is

appropriate, an unused wavelength should be used to construct an HC for the flow. In this situation, SHM takes into consideration the wavelengths occupied by *symmetric groups* first. The sources of the symmetric group and the data transmission direction are the same as that of the flow. Once a wavelength has been successfully selected to construct the HC, the corresponding wavelength sequences should be updated.

In this paper, we extend the path matching which is originally proposed in [16]. Besides, the HC of a flow may be able to share a wavelength with different-source HCs via different strategies. Finally, compared with other HC-sharing approaches, such as Random, Least-Used and Most-Used, the proposed OPM-SHM can keep the wavelength utilization more balanced.

The rest of the paper is structured as follows. Section 2 introduces the related work on this subject. Section 3 formulates the DGWA problem in LOBS-HC ring networks. The main idea of OPM-SHM is described in detail in Section 4, and then, in Section 5, the simulation results are given with analysis. Finally, Section 6 draws the conclusion from the work.

2. Motivation and Related Work

Below, we first present some challenges facing the development of optical datacenter networks. We then introduce the main features of LOBS-HC and the related work addressing the GWA issue in the ring network based on LOBS-HC.

2.1 Challenges for Optical Switching in Datacenter Networks

Datacenters today are essential for Internet enterprises. How to effectively transport big data to other devices or datacenters becomes very important and the transportation speed will impact on the performance and efficiency of the datacenter network.

Currently, there is still no effective way to enlarge the bandwidth capacity of datacenters. With the tremendous growth of mobile devices, such as smartphones and laptops, the Internet infrastructure of enterprises needs more and more capacity to support these devices. In order to meet user requirements, datacenter networks are expected to grow on a massive scale. A modern datacenter consists of thousands of servers and large datacenters with hundreds of thousands of servers are becoming common, which means that the traffic generated within datacenters is growing significantly on an exponential level.

High energy cost is another challenge for datacenters. Generally, maintainers satisfy the continuous increasing bandwidth requirement by adding more network devices, which not only occupies extra space but also increases the energy consumption of electricity and cooling systems. Electronic switching consumes 10~20 percent of the total power of datacenters, and it is expected to increase in the near future.

Furthermore, the current datacenters which are mainly built from commodity Ethernet switches can no longer remain attractive [3]. The high oversubscription ratio in electronic networks can lead directly to high connection complexity and power consumption, and also results in traffic congestion.

All these challenges drive the need for high-performance technologies and architectures for intra-datacenter networks. Therefore, optical switching is seen as a promising solution to address the issues raised by the above challenges due to its attractive features, such as high bandwidth and reliability, and low power consumption and cabling complexity [4]. People tend to focus on electronic-optical hybrid switching technology [17, 18, 19], and finally try to realize all-optical switching [20]. At present, the primary optical switching technologies, such as OCS and OBS, can support high throughput in networks and reduce the communication cost and

latency among inter-datacenters or intra-datacenters. However, they cannot make use of resources effectively, while still providing a guaranteed bandwidth. Given current datacenter architectures, a better way to provide required bandwidth between dynamically changing sets of nodes is to build a non-blocking switch fabric for datacenter networks, where LOBS-HC is one of the promising paradigms to address the issue [14].

2.2 Labeled Optical Burst Switching with Home Circuit (LOBS-HC)

LOBS-HC is based on the labeled OBS framework and aims to provide lossless data transmission with a limited delay as in OCS, but in a more efficient manner [6]. In LOBS-HC, HCs are not wavelength routed or switched. The traffic on the HC is switched in burst granularity. For the same-source HCs, there are two ways of sharing wavelengths. One, multiple same-source HCs can share wavelengths through common links; the other, one HC may be able to share the wavelengths having been assigned to other different-source HCs with low priority. Hence, whenever a traffic request arrives, a corresponding HC will be built with a certain bandwidth. Being assigned logically, the bandwidth resource is not monopolized by the corresponding traffic. It is also worth noticing that traffic is still switched burst-by-burst in core LOBS-HC nodes.

Assuming that there is an N -node network, and the traffic demand of the flow associated with the source-destination node pair (s, d) is given, where s is the source and d is the destination, the given traffic demand is between the average and the peak traffic of the flow and sufficient for providing the required QoS. The wavelength(s) allocated in LOBS-HC is adequate for the flow to transmit through the network without resource contention. The HC established for the node pair (s, d) is not a physical but a logical connection. The bursts carried by the HC are still switched on the burst-by-burst basis at the core nodes; therefore, the wavelength resource assigned to a given HC can be reused by other flows during the burst-send interval time. In particular, two HCs from the same source (perhaps with different destinations) can share wavelength(s) [6]. It is different from an OCS circuit, which is wavelength-switched and used solely by only one flow. In addition, LOBS-HC can still provide the guaranteed bandwidth to the flow. Because the source node synchronizes their O/E/O(optical-electrical-optical) conversion and transmission, the bursts with the same source will not collide with each other and are called in-profile bursts. The following example illustrates the main feature of LOBS-HC.

A five-node topology is shown in Fig. 1. Denoting an HC from source s to destination d by $HC_{s,d}$ and its required bandwidth by $h_{s,d}$, there are four HCs, i.e., $HC_{0,1}$, $HC_{0,2}$, $HC_{0,3}$ and $HC_{0,4}$. The normalized bandwidths of the HCs are $h_{0,1} = 0.3$, $h_{0,2} = 0.4$, $h_{0,3} = 0.5$, and $h_{0,4} = 0.6$, respectively. If a wavelength can be used by only one HC in each link, four wavelengths will

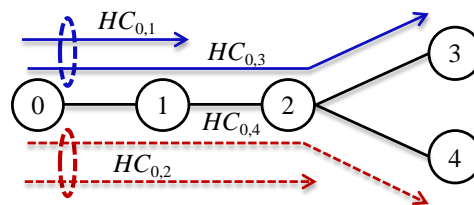


Fig. 1. Four HCs with the same-source node

be occupied on link (0, 1). However, in LOBS-HC, the four HCs from node 0 can share wavelengths. Because $h_{0,1} + h_{0,3} < 1.0$ and $h_{0,2} + h_{0,4} < 1.0$, we can group $HC_{0,1}$ and $HC_{0,3}$, and group $HC_{0,2}$ and $HC_{0,4}$, respectively. Finally, only two wavelengths are used in link (0, 1).

2.3 GWA in Ring-based LOBS-HC Networks

The main issue in dynamic LOBS-HC networks is to construct HCs for traffic flows. The construction process consists of routing (routing can be ignored under the ring topology), grouping, and wavelength assignment (i.e. GWA). The goal of grouping is to group multiple HCs optimally from a same-source node. These grouped HCs require sub-wavelength bandwidths and share a minimal number of wavelengths after being grouped together. According to the results of grouping, there are two ways to assign proper wavelengths to HCs. In other words, HCs with the same source can share wavelengths through their common links, otherwise, one HC can occupy the remainder wavelength bandwidth from other HCs having different sources. By the two wavelength sharing methods, the burst loss probability will be reduced.

Several mathematical formulations and heuristic algorithms have already been proposed to address the GWA issue. In [14], ring intra-datacenter networks using LOBS-HC were studied, and a method named complementary HC assignment (CHA) for assigning wavelengths to HCs was proposed to support all-to-all and uniform communication among all pods (each having thousands of servers) using minimum wavelengths [14]. The basic idea of CHA is for different nodes to spatially reuse the same wavelength as the origin of their respective HCs with polynomial time complexity. In addition, compared with electrical counterparts, CHA has shown great promise in terms of reducing the number of needed wires and transceivers (also including a reduction in cost and power consumption).

An integer linear programming (ILP) formulation was proposed in [12]. Based on the ILP scheme, an optimal solution can be found for unidirectional and bidirectional rings under both all-to-all uniform traffic and arbitrary traffic demands; however, this solution is only feasible for small-size networks. Then, a heuristic algorithm, called GAP in this paper, was proposed with lower computational complexity than ILP [13]. However, the wavelength utilization of the heuristic approach is no better than that of ILP and CHA, and the earlier a wavelength is assigned, the heavier the load this wavelength will bear. Furthermore, in our preliminary work, a heuristic scheme named longest path matching and graph coloring (LPMGC) was proposed to address the GWA issue [15]. Specifically, the graph coloring algorithm is applied to implement the wavelength assignment. The results show that LPMGC costs almost the same in wavelengths as CHA. Even though providing better load balancing on the wavelengths than CHA, the time and space complexity of LPMGC is higher.

One thing to note is that the work described above has not taken into consideration the DGWA problem. However, based on part of the work we undertook in [15], this paper focus on addressing the DGWA issue with extension to and modification of the wavelength assignment problem.

3. Problem Formulation

In this section, we will briefly describe the HC construction problem. In this work, we consider an all-optical ring network based on LOBS-HC without wavelength conversion. Because the routing paths based on SPF are fixed for any traffic, the routing issue does not need to be taken into consideration in ring networks. There is a bidirectional link between every two adjacent

nodes. It is assumed that each wavelength can only transmit bursts in one direction (i.e., in the clockwise or counterclockwise direction) at one moment in time.

Fig. 2 illustrates an eight-node ring. In order to make the problem easy to understand, each node is also considered as an end node, i.e., it can transmit or receive traffic as well as forward traffic from other nodes. The numbers of wavelengths in the clockwise and counterclockwise directions are the same. If a flow belonging to a given source-destination node pair (s, d) requires bandwidth x , an HC from the source s to destination d will be established by using one or more wavelengths in order to guarantee the bandwidth that the flow requires.

One wavelength can be shared simultaneously by multiple same-source HCs due to the property of LOBS-HC. In **Fig. 3**, an example is given to describe how HC grouping works in a ring. There are three HCs in the counterclockwise direction: $HC_{1,7}$, $HC_{1,6}$ and $HC_{1,5}$ and their normalized bandwidths are $h_{1,7} = 0.3$, $h_{1,6} = 0.4$ and $h_{1,5} = 0.5$, respectively. If a wavelength can only be assigned to one HC in each link, three wavelengths will be occupied on links (1, 8) and (8, 7) in the counterclockwise direction. However, being from the same-source node, the three HCs can share wavelengths. Because $h_{1,6} + h_{1,5} < 1.0$ and $h_{1,7} + h_{1,6} + h_{1,5} > 1.0$, we can group $HC_{1,6}$ and $HC_{1,5}$ together to share one wavelength, and make $HC_{1,7}$ the only member of a new group using another available wavelength. Finally, only two wavelengths are used on links (1, 8) and (8, 7). To reduce the number of wavelengths being allocated, wavelength sharing can

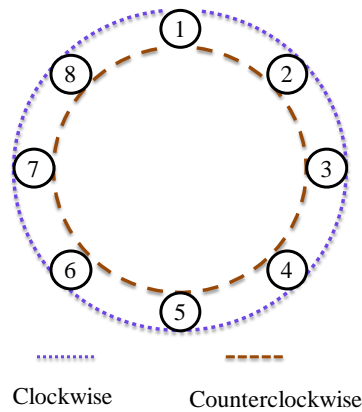


Fig. 2. Bidirectional ring with eight nodes

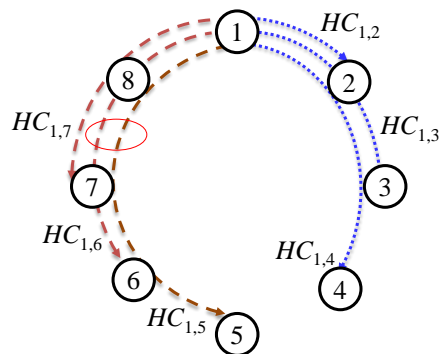


Fig. 3. An example of six HCs from node 1 in which $HC_{1,6}$ and $HC_{1,5}$ are grouped together

also be applied among the HCs, e.g. $HC_{1,2}$, $HC_{1,3}$ and $HC_{1,4}$, in the clockwise direction. In addition, two HCs with the same source, but in different directions, are unavailable in the situation of sharing wavelengths.

LOBS-HC is not wavelength-routed or switched. Instead, the traffic carried by an HC is still optically switched on a burst-by-burst basis and the bursts belong to the traffic flow. Therefore, in this work, we model the dynamic traffic process as the one-by-one arrival and departure process in which only one request arrives or departs at a time. Since the HC departure process requires releasing the relative resources, this work focuses on dealing with the request arrival process, i.e., assigning a befitting wavelength to construct HC.

4. Optimal Path Matching and Symmetric HC Matching

In this section, OPM-SHM is proposed to implement the GWA in dynamic network traffic. OPM solves the grouping issue and tries to share wavelength(s) with other same-source HCs, and then build the HC for any newly arrived request, while SHM will be applied to assign an unused wavelength to flows as long as OPM cannot construct the corresponding HCs successfully.

4.1 Optimal Path Matching (OPM)

OPM enhances the max-matching scheme, which is proposed in our preliminary work in [16], in order to optimize the HC grouping process. To make the matching more comprehensive and balanced, we extend the matching operation between a flow and an HC to that between a flow and an HC group. To facilitate the description, it is assumed that the network is denoted by a graph G consisting of N nodes (i.e., vertexes). An edge in G corresponds to a bidirectional link in the network, e.g., the existence of an edge $e(s, d)$ from node s to node d implies the existence of an edge $e(d, s)$ from node d to node s .

The following is a list of notations used to describe the scheme.

Notations:

N : the total number of nodes in the ring; n_i indicates a single identified node ($i \in [1, N]$).

W : the number of wavelengths on each physical link.

$HC_{s,d}$: a unicast HC or an HC construction request from source node s to destination node d ; $|HC_{s,d}|$ indicates the length of the route of $HC_{s,d}$, or the number of links on the path of $HC_{s,d}$.

$hc_{s,d}$: the normalized traffic demand (or the required bandwidth) of $HC_{s,d}$.

Group: a set of same-source HCs; $Group_{s,k}$ stands for the k th group with source s ; $|Group_{s,k}|$ indicates the number of HCs in $Group_{s,k}$.

We use Fig. 4 to illustrate the basic principle of HC grouping based on OPM. The blue solid line indicates the newly arrived request $HC_{s,d}$, and the black dotted line and the red dashed line indicate two HCs (i.e., $HC_{s,d1}$ and $HC_{s,d2}$) in the same group. It is assumed that the group has enough capacity for $HC_{s,d}$ grouping. In the two cases shown in Fig. 4 (a) and (b), the groups are available for $HC_{s,d}$. However, the group shown in Fig. 4 (c) cannot meet the requirement of $HC_{s,d}$ as $HC_{s,d}$ needs an extra link (d_2, d), but $HC_{s,d1}$ and $HC_{s,d2}$ cannot provide it.

As to how to choose a suitable group, we give the following definitions.

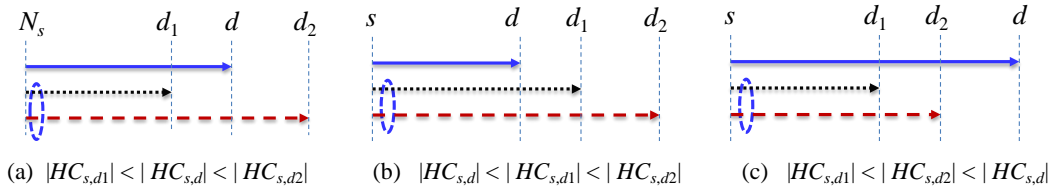


Fig. 4. Relationship between the HC request and the grouped HCs

Definition 1: $\gamma_{hc}(HC_{s,d}, HC_{s,d'})$, the *matching degree* between the routes of the two same-source HCs: $HC_{s,d}$ and $HC_{s,d'}$, which at least have one common physical link on their routes. It is calculated by

$$\gamma_{hc}(HC_{s,d}, HC_{s,d'}) = \begin{cases} \frac{||HC_{s,d'}| - |HC_{s,d}||}{|HC_{s,d'}|}, & \text{if } |HC_{s,d}| - |HC_{s,d'}| < 0 \\ 0, & \text{if } |HC_{s,d}| - |HC_{s,d'}| > 0 \\ 1, & \text{if } |HC_{s,d}| - |HC_{s,d'}| = 0 \end{cases} \quad (1)$$

Definition 2: $\gamma_{group}(HC_{s,d}, Group_{s,k})$, the *matching degree* between $HC_{s,d}$ and the k th group with source s . It can be calculated as in

$$\gamma_{group}(HC_{s,d}, Group_{s,k}) = \frac{\sum \gamma_{hc}(HC_{s,d}, HC_{s,d'})}{|Group_{s,k}|}, \text{ where } HC_{s,d'} \in Group_{s,k}. \quad (2)$$

The calculation of routing is omitted here. We focus on the main process of the OPM algorithm to find the available wavelength to construct an HC for a newly arrived flow. Using $F_{s,d}$ to indicate the new request from s to d , the OPM process to find a proper wavelength for $F_{s,d}$ is presented below.

Optimal Path Matching (OPM)

-
- Step 1:** Find the same-source groups for $F_{s,d}$, and calculate the matching degrees of between $HC_{s,d}$ and the groups.
- Step 2:** Sort the groups in the descending order of matching degrees.
- Step 3:** Check the i th corresponding group $Group_{s,i}$
- If** $Group_{s,i}$ has enough space (here, space is used for capacity) for $F_{s,d}$
 Construct an HC for $F_{s,d}$ sharing the wavelength used by $Group_{s,i}$
 Exit.
 - Else**
 Repeat Step 3 until all the found groups are checked.
-

The wavelength(s) shared within a group should be released from the links when all bursts belonging to the group have passed the reserved links. The group will be deleted if no wavelength is occupied by it.

If no HC can be constructed for the request $F_{s,d}$, some other wavelength will be assigned by using SHM which is described in Section 4.2.

4.2 Symmetric HC Matching (SHM)

In this subsection, the SHM module of OPM-SHM is illustrated. The main job of SHM is to select an appropriate wavelength to build an HC by dynamically maintaining wavelength sequences for all nodes in the network.

In this work, the first fit (FF) method chooses an unused wavelength for a newly arrived flow. Fig. 5 shows the initial state of wavelength sequence and the wavelength sequences are the same for all nodes at the very beginning of the whole process. Each time a new flow arrives, the process moves forward through the sequence from the beginning position to find an unused wavelength for the HC construction and changes intermediate node sequences by obeying following two rules.

Rule 1: For an arrival flow from s to d , it is assumed that there are several intermediate nodes between s and d on the route. The wavelength assigned to the flow cannot be re-assigned to the remaining arrival flows whose one or more intermediate nodes are the sources of other flows, until it is released.

Based on the initial state shown in Fig. 5, we use Fig. 6 to explain how Rule 1 works. It is assumed that there is a new request from s to d and two nodes n_1 and n_2 are the intermediate nodes. The wavelength λ_k is adopted to construct an HC for the connection request. In this situation, λ_k cannot be assigned to the next arrival flows from source n_1 or n_2 until it is freed. Then, the position of λ_k in the wavelength sequences of nodes n_1 and n_2 will be moved backward. Here in Fig. 6, for simplicity, we just move λ_k to the tail of the wavelength sequence.

Rule 2: For any arriving HC request, SHM must first search the wavelengths of the symmetric group to find one to construct the HC.

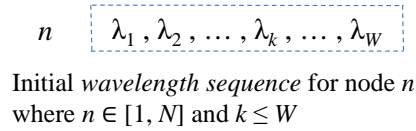


Fig. 5. Wavelength sequence initialization for each node

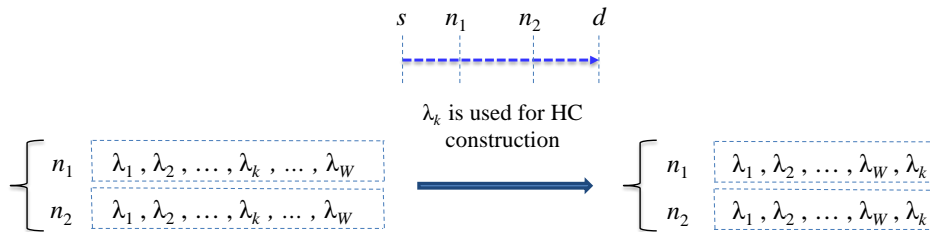


Fig. 6. Maintain the wavelength sequences of intermediate nodes n_1 and n_2 , and simply move λ_k to the end of the sequences because λ_k cannot be re-assigned to any flow from source n_1 or n_2 , otherwise collision will occur.

Two HC groups are symmetric on the basis that their sources are symmetric and their data are transmitted in the same direction. The relationship between two symmetric nodes is shown

below. Assuming a node is presented by its index, s^* , the symmetric node of s , can be calculated according to (3) and (4).

When N is even:

$$\begin{cases} s^* = N, & \text{if } s = N/2 \\ s^* = (s + \frac{N}{2}) \bmod N, & \text{if } s \in [1, \frac{N}{2}) \cup (\frac{N}{2}, N] \end{cases} \quad (3)$$

When N is odd:

$$\begin{cases} s^* = N-1, & \text{if } s = \frac{N}{2} \\ s^* = (s + \frac{N}{2}) \bmod (N+1), & \text{if } s \in [1, \lfloor \frac{N}{2} \rfloor) \cup (\lfloor \frac{N}{2} \rfloor, N] \end{cases} \quad (4)$$

That is to say, if the number of nodes N is odd, there may be a single node without a symmetric node. Then, node n_N (the last node) is ignored, and the group from node n_N should be considered independently. Fig. 7 shows how Rule 2 maintains the wavelength sequence of the symmetric node s^* , or the symmetric node of s . After λ_k is assigned to the new flow, the position of λ_k is simply moved to the head of the sequence of s^* .

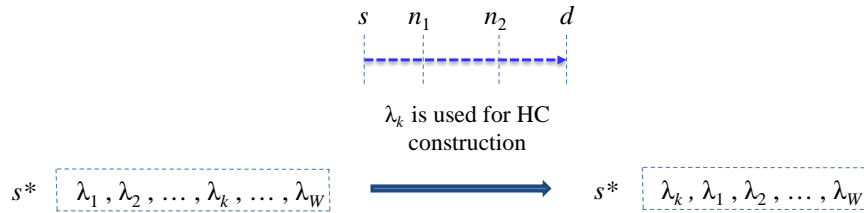


Fig. 7. Update the wavelength sequence of the symmetric node of node s : just move λ_k to the head of the sequence because the wavelengths occupied by the symmetric group are taken into consideration first.

Note that, for a wavelength in the sequence, Rule 1 has higher priority than Rule 2. That is to say, if Rule 1 has been applied to a wavelength, until the wavelength is released, Rule 2 cannot take effect. Furthermore, SHM focuses on the symmetric feature, that is to say, on the distance (or number of hops) from source to destination being no larger than $\lceil N/2 \rceil$ in the clockwise direction. Otherwise, for any traffic in the counterclockwise direction, the upper bound of the distance between source and destination is $\lfloor N/2 \rfloor$.

We divide the main process of SHM into two phases: flow arrival and flow departure.

Symmetric HC Matching (SHM)

Phase 1: When a flow from node s to node d arrives, proceed as follows:

Step 1.1: Search the wavelength sequence of node s to find a proper wavelength (e.g., λ) for the flow.

Construct the corresponding HC, i.e., $HC_{s,d}$.

Step 1.2: Find the intermediate nodes along the path of $HC_{s,d}$.

- Modify the position of λ in the corresponding wavelength sequences by Rule 1.
- Step 1.3: Calculate s^* , the symmetric node of s .
- Modify the position of λ of the corresponding wavelength sequence by Rule 2.
- Phase 2:** With the flow departures, proceed as follows:
- Step 2.1: Remove $HC_{s,d}$ from its group, and release the relative resource.
- Step 2.2: If there are no other HCs in the group, withdraw the group.
-

5. Analysis and Performance Evaluation

This section discusses the time complexity and advantage of OPM-SHM, and presents the simulation work and results demonstrating the network performance in applying OPM-SHM.

5.1 Time Complexity of OPM-SHM

The time complexity in calculating the matching degree is $O(1)$. For the arrival of each new request, the maximum computing times are no more than $N - 1$, because there are at most $N - 1$ same-source HCs at a time. Therefore, the worst time complexity in calculating the matching degree is $O_w(N)$. Then, depending on the calculated matching degrees, we can ascertain the matching degrees between the new flow and the same-source groups. After that, the same-source groups are sorted in descending order according to the average degree in $O(M \lg N)$ time. Finally, the sequence will be searched to find the most proper group for the HC request which costs at most $O(N)$. Overall, the total time complexity of OPM is $O(N + M \lg N)$.

The time complexity of SHM consists of three parts. The first is searching the wavelength sequence which can be done in $O(W)$ time. Applying Rule 1 costs $O(k)$ where k is equal to W being the worst for each intermediate node, and applying Rule 2 requires the same time. Therefore, the total time complexity is $O((2+\delta)*W)$ where δ indicates the number of intermediate nodes, excluding the source node and destination nodes. Overall, combining OPM and SHM, the ultimate time complexity is $O((N+1)*\lg N + (2+\delta)*W)$.

5.2 Analysis of Max Hop Counts by Symmetric HC Matching

It is noted that, the main feature of SHM is considering the symmetric node and group, which is suitable for application in dynamic traffic networks. Therefore, SHM can keep the maximum number of flow routing hops as $\lceil N/2 \rceil$ even in the worst situation, while that of some other methods approaches to $N-1$.

Assuming that there are M flows and M is no larger than $(N-1)^2$, and each flow is identified by $flow_i$ ($i \in [1, M]$) with a route length of route denoted by l_i , depending on these factors, the average hop count ω_{a_hop} can be calculated by:

$$\omega_{a_hop} = \frac{\sum_{i=1}^M l_i}{M} \quad . \quad (5)$$

As introduced in Section 5.1, there are N nodes in the network. SHM divides the ring into two symmetric parts, and this signifies that the maximum number of hops for any flow is no larger than $\lceil N/2 \rceil$. The derivation is given as follows:

$$\frac{\sum_{i=1}^M l_i}{M} \leq \frac{\sum_{i=1}^M \left\lceil \frac{N}{2} \right\rceil}{M}, \text{ and } \frac{\sum_{i=1}^M \left\lceil \frac{N}{2} \right\rceil}{M} = \frac{M * \left\lceil \frac{N}{2} \right\rceil}{M} = \left\lceil \frac{N}{2} \right\rceil$$

s.t.,

$$\omega_{a_hop} \leq \left\lceil \frac{N}{2} \right\rceil. \quad (6)$$

For example, let us compare SHM with another method, e.g. Less-Used that considers the less used links first. In Fig. 8, there are two existing HCs and an arrival flow shown by solid lines and a dotted line, respectively. In Fig. 8 (a), it is easy to calculate that:

$$\omega_{a_hop} = (|HC_{1,2}| + |HC_{1,3}| + |HC_{1,4}|) / 3 = (1 + 2 + 3) / 3 = 2,$$

while the hop count is $(1+2+5) / 3 = 2.7$ in Fig. 8 (b). Although in this case, SHM increases the traffic load of the physical link (1, 2), it decreases the average hop count of flow routing successfully.

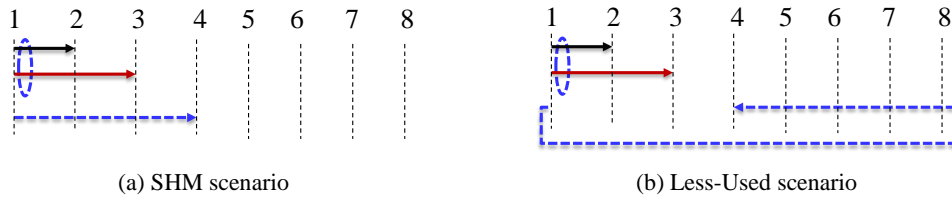


Fig. 8. Comparison of hop count in the SHM and Less-Used methods

5.3 Experiment Results and Performance Evaluation

To verify the aforementioned OPM-SHM solution, we ran a simulation on an eight-node ring network in limited- and sufficient-wavelength scenarios. Each node is the edge and core LOBS-HC node without any optical buffer and corresponds to a pod in the datacenter network. The number of wavelengths per physical link is two or three in the limited-wavelength scenario, and the capacity of each wavelength is 40 Gbps. All the flows are in unicast communication. The source and destination of each flow are chosen randomly and the routing is based on the Dijkstra algorithm. The connection requests of flows dynamically arrive and depart one-by-one in a uniform manner. In addition, the maximum burst size is between 100 and 200 Kbits, and each burst consists of several small-size packets whose average size is 4000 bits with exponential distribution.

In the simulation, the flow arrival rate ρ is calculated as in (7), where W , TTL and ϕ are the number of wavelengths in a physical link, the average survival life of a flow, and the average of the required bandwidth of each flow, respectively. Using O_L as the offered load in the network, generally, W and O_L are the two primary factors affecting ρ .

$$\rho = \frac{W * TTL * O_L}{\varphi} \quad (7)$$

We applied OPM-SHM under the DGWA condition and compared it with some other heuristic schemes, like Random, Least-Used and Most-Used, in terms of the burst loss rate and resource utilization. In Random, the wavelength is selected randomly from the wavelengths having spare bandwidth. The wavelengths occupied by the least and most used links will be assigned with high priorities in Least-Used and Most-Used, respectively.

Resource utilization is evaluated by using three metrics: wavelength utilization rate, light-link utilization rate, and the standard deviation of light-link utilization. Note that CHA and GAP mentioned in Section 2 cannot be compared here, since they are only available under the SGWA condition. If an HC fails to be set up for a flow, the flow will be dropped at the source, and the blocked bursts are discarded due to there being no optical buffer in the LOBS-HC core nodes.

In addition, the wavelength utilization rate is achieved by using (8), where w_i indicates the number of wavelengths being used on the i th physical link and $|E(G)|$ is the number of physical links.

$$v_\lambda = \frac{\sum_{i=1}^{|E(G)|} \left(\frac{w_i}{W}\right)}{|E(G)|} \quad (8)$$

Besides, light-link corresponds to a physical link with a certain wavelength, and the light-link utilization rate $v_{light-link}$ can be calculated by (9), where $n_{light-link_used}$ is the number of light-links used.

$$v_{light-link} = \frac{n_{light-link_used}}{2|E(G)W|} \quad (9)$$

The mean-variance of the light-links used is calculated by (10), in which $E|b|$ is the average number of transmitted packets of each light-link used, and b_i stands for the number of packets in the i th link.

$$\sigma = \sqrt{\frac{1}{2|E(G)W|} \sum_{i=1}^{2|E(G)W|} (b_i - E|b|)^2} \quad (10)$$

The results shown in **Fig. 9** and **Fig. 10** are both achieved in the limited-wavelength scenario. This demonstrates that OPM-SHM can achieve a lower burst loss rate and higher wavelength utilization than the other three methods when the offered load is less than 1. As shown in (7), the flow arrival rate becomes larger as the offered load or the number of wavelengths in a physical link increases, which means more flows are injected into the network from the source nodes and more burst congestions will take place. In the OPM

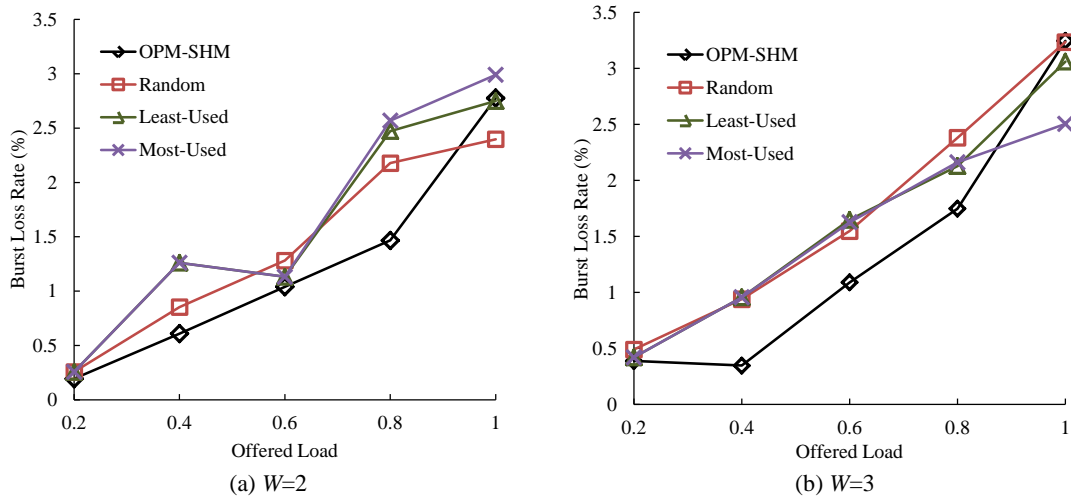


Fig. 9. Burst loss rate in an 8-node ring with limited wavelengths

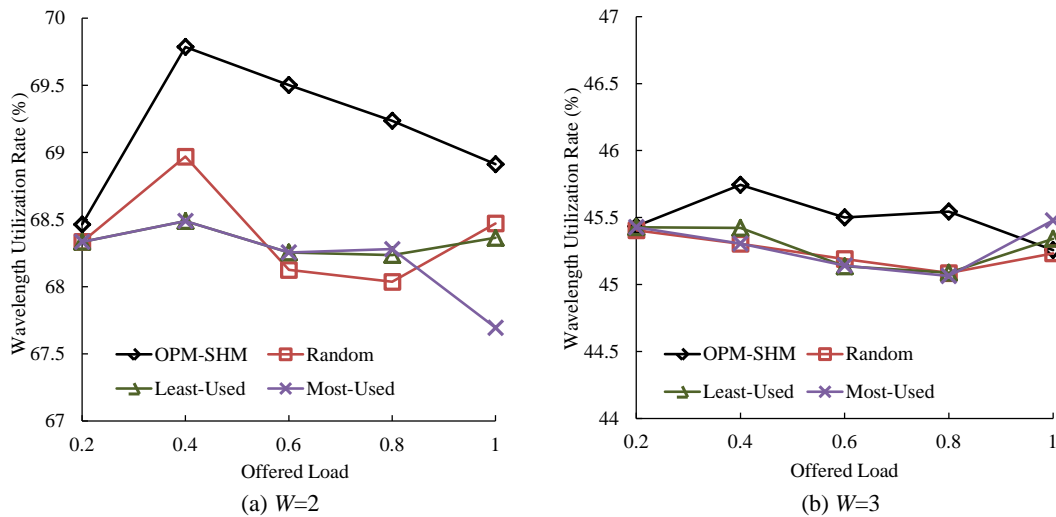


Fig. 10. Wavelength utilization rate in an 8-node ring with limited wavelengths

module, abiding by (1) and (2), we try to make the same-source HCs with less hop distance group together and assign the wavelength bandwidth more compactly. Therefore, compared with the other three methods, the heuristic method deals with the huge amount of traffic more flexibly.

Furthermore, in the SHM module, the symmetric HCs, which have different source nodes, can share the same wavelength, which is different from the original LOBS-HC. It takes the length of routing and load balancing as one of its measurements and can also reduce the burst collisions to a certain extent. On the contrary, the other three methods only focus on the cases using wavelength allocation. This results in great congestion on some certain wavelengths or links. The above results indicate that the OPM-SHM scheme can provide high wavelength utilization while maintaining a low burst loss rate with limited wavelengths.

In order to evaluate the network performance by applying OPM-SHM when the bandwidth resource is infinite, we changed the number of wavelengths per physical link W to 100 which is enough for the traffic demand. The simulation results are depicted in Fig. 11 and Fig. 12.

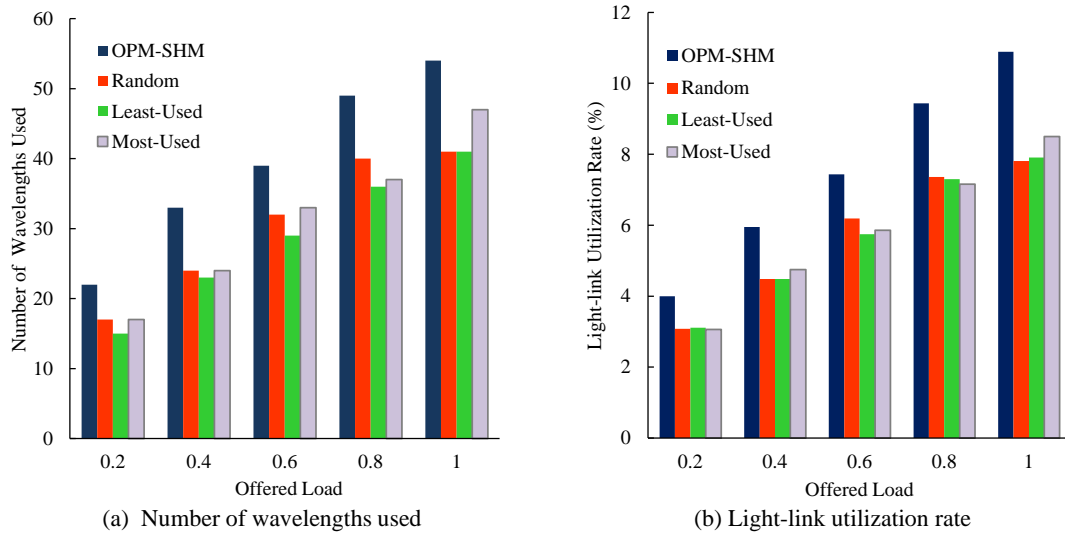


Fig. 11. Number of wavelengths used & light-link utilization rate in an 8-node ring with enough wavelengths

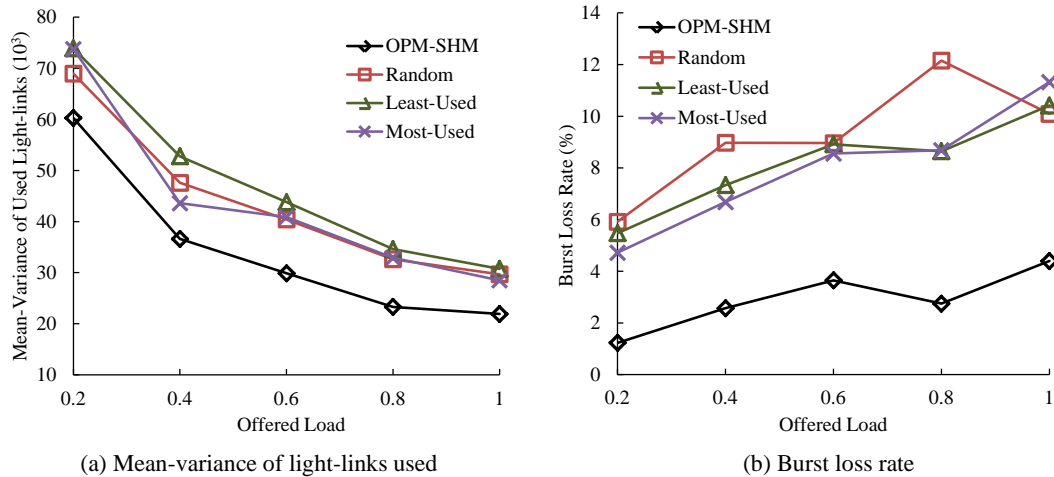


Fig. 12. Mean-variance of light-links used & burst loss rate in an 8-node ring with enough wavelengths

As presented in **Fig. 11(a)** and **(b)**, OPM-SHM makes the most use of the wavelengths and light-links provided by the network. Therefore, more HCs are constructed in OPM-SHM than in the other three methods. Moreover, as plotted in **Fig. 12(a)**, OPM-SHM can keep a lower mean-variance σ while the light-link utilization increases. This indicates that the traffic flows are evenly distributed in the light-links. Therefore, besides the limited-wavelength scenario having the same advantages as OPM-SHM mentioned above, OPM-SHM can achieve a lower burst loss rate than the other three methods, as demonstrated by the result given in **Fig. 12(b)**. To be more specific, we can see that, first, OPM-SHM can achieve a low burst loss rate with limited resources, and second, OPM-SHM will cost more in terms of wavelength when there are sufficient wavelengths. Nevertheless, OPM-SHM uses the extra wavelengths on reducing the congestion probability and tries to distribute the traffic equally to the wavelengths in the network. It turns out that this strategy can decrease the burst loss rate effectively.

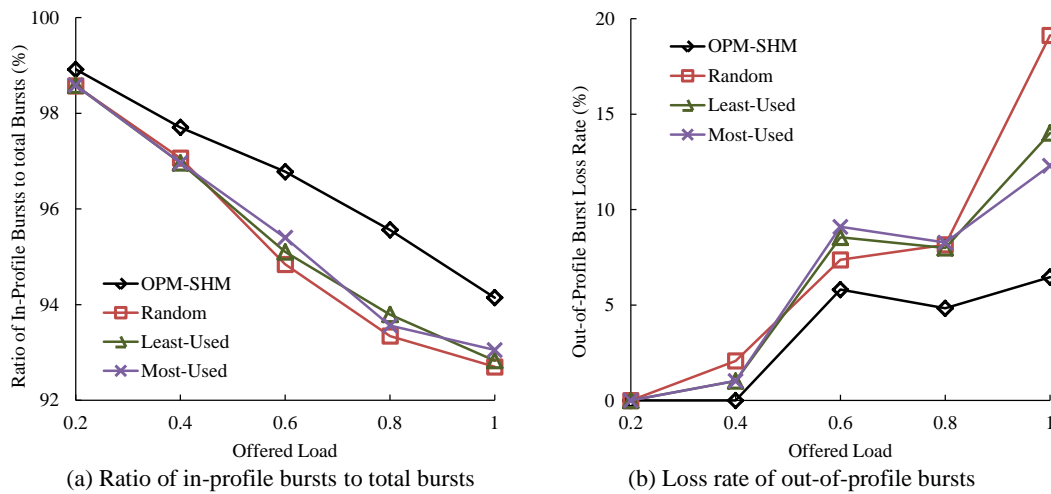


Fig. 13. Ratio of in-profile bursts to total bursts and loss rate of out-of-profile bursts ($W=3$)

The bursts with the loss rate presented in Fig. 9 and Fig. 12(b) include in-profile bursts and out-of-profile bursts. The percentage rates of in-profile bursts decrease as the offered load increases in all four schemes, and the numbers of in-profile bursts in the OPM-SHM cases are the highest in all the scenarios. For instance, as shown in Fig. 13(a), the ratio of in-profile bursts to total bursts is larger than 94% in the OPM-SHM cases when $W=3$. In addition, OPM-SHM can also provide better transmission for out-of-profile bursts.

6. Conclusion

How to dynamically construct a home circuit is one of the crucial points of the DGWA in LOBS-HC. In this paper, we have proposed an OPM-SHM scheme consisting of two parts, OPM and SHM to overcome this problem for ring datacenter networks. As a result, the performance of the network applying OPM-SHM has been evaluated by simulation under both limited- and sufficient-wavelength situations. OPM tries to build the HC by sharing wavelength(s) with the same-source HCs. The wavelength, which is occupied by a group and has the maximum matching degree, will be chosen to construct an HC for a dynamically arriving flow. As the remedy module following OPM, SHM takes into consideration the wavelengths being allocated to the symmetric group of the HC construction request of the flow, and it maintains the wavelength sequences of certain intermediate nodes in real time. In this way, SHM can guarantee the hop count ceiling of HCs to be $\lceil N/2 \rceil$.

Overall, the results have shown that the proposed OPM-SHM can reduce burst collisions, and allocate wavelength bandwidth effectively. Furthermore, OPM-SHM is able to balance the traffic and obtain a low burst loss rate in the datacenter network using the LOBS-HC ring.

References

- [1] M. Chen, S. Mao and Y. Liu, "Big data: a survey," *ACM/Springer Mobile Networks and Applications*, vol. 19, no. 2, pp. 171-209, 2014. [Article \(CrossRef Link\)](#)
- [2] M. Chen, H. Jin, Y. Wen and V. C. M. Leung, "Enabling technologies for future data center networking: a primer," *IEEE Network*, vol. 27, no. 4, pp. 8-15, 2013. [Article \(CrossRef Link\)](#)

- [3] K. Chen, A. Singla, A. Singh, K. Ramachandran, L. Xu, Y. Zhang, X. Wen and Y. Chen, "OSA: an optical switching architecture for data center networks with unprecedented flexibility," *IEEE/ACM Transactions on Networking*, vol. 22, no.2, 2014. [Article \(CrossRef Link\)](#)
- [4] A. Drakos, T. G. Orphanoudakis and A. Stavdas. "Performance benchmarking of core optical networking paradigms," *Optics Express*, vol. 20, no. 16, pp. 17421-17439, 2012. [Article \(CrossRef Link\)](#)
- [5] M. A. González-Ortega, C. Qiao, A. Gonzalez, X. Liu and J. Lopez-Ardao, "LOBS-H: an enhanced OBS with wavelength sharable home circuits," in *Proc. of IEEE Int. Conf. on Communications (ICC2010)*, pp. 1-5, 2010. [Article \(CrossRef Link\)](#)
- [6] P. Pavon-Marino and F. Neri, "On the myths of optical burst switching," *IEEE Transactions on Communicaions*, vol. 59, no. 9, pp. 2574-2584, 2011. [Article \(CrossRef Link\)](#)
- [7] C. Qiao and M. Yoo, "Optical burst switching (OBS) – a new paradigm for an optical Internet," *Journal of High Speed Networks*, vol. 8, no. 1, pp. 69-84, 1999. [Article \(CrossRef Link\)](#)
- [8] N. Farrington, A. Forencich, P.-C. Sun, J. E. Ford, Y. Fainman, G. C. Papen and A. Vahdat, "A multiport microsecond optical circuit switch for data center networking," *IEEE Photonics Technology Letters*, vol. 25, no. 16, pp. 1589-1592, 2013. [Article \(CrossRef Link\)](#)
- [9] Y. P. Aneja, "Routing in wavelength routed optical networks," in *Proc. of 2001 IEEE Workshop on High Performance Switching and Routing*, pp.155-158, 2001. [Article \(CrossRef Link\)](#)
- [10] P. Saengudomlert, E. Modiano and R. G. Gallager, "On-line routing and wavelength assignment for dynamic traffic in WDM ring and torus networks," *IEEE/ACM Transactions on Networking*, vol.14, no.2, pp. 330-340, 2006. [Article \(CrossRef Link\)](#)
- [11] H. Zhang, J. P. Jue and B. Mukherjee, "A review of routing and wavelength assignment approaches for wavelength-routed optical WDM networks," *Optical Networks Magazine*, pp. 47-60, 2000. [Article \(CrossRef Link\)](#)
- [12] J. Li, X. Gao, W. Tang and C. Qiao, "Optimal home circuit grouping in LOBS-HC rings," in *Proc. of IEEE Photonics Conference (IPC2012)*, pp. 858-859, 2012. [Article \(CrossRef Link\)](#)
- [13] J. Li, H. Zou, W. Tang, X. Gao and C. Qiao, "Home circuit grouping in LOBS-HC ring networks: ILP and heuristic approaches," in *Proc. of 2013 IEEE/CIC Int. Conf. on Communications in China (ICCC2013)*, pp. 171-176, 2013. [Article \(CrossRef Link\)](#)
- [14] L. Peng, C. H. Youn, W. Tang and C. Qiao, "A novel approach to optical switching for intradatecenter networking," *Journal of Lightwave Technology*, vol. 30, no. 2, pp. 252-266, 2012. [Article \(CrossRef Link\)](#)
- [15] X. Yang, B. Yi, W. Tang and J. Li, "Heuristic scheme for home circuit grouping and wavelength assignment in LOBS-HC networks," *Journal of Optical Communications*, vol. 35, no. 3, pp. 211-219, 2014. [Article \(CrossRef Link\)](#)
- [16] W. Tang, X. Yang, B. Yi and R. Zhu, "Home circuit sharing for dynamic wavelength assignment in LOBS-based datacenter networks," *IEICE Transactions on Information and Systems*, pp. 2660-2662, 2014. [Article \(CrossRef Link\)](#)
- [17] N. Farrington, G. Porter, S. Radhakrishnan, H. H. Bazzaz, V. Subramanya, Y. Fainman, G. Papen and A. Vahdat, "Helios, a hybrid electrical optical switch architecture for modular data centers," in *Proc. of ACM Annual Conf. of the Special Interest Group on Data Communication (SIGCOMM 2010)*, pp. 339-350, 2010. [Article \(CrossRef Link\)](#)
- [18] H. Bazzaz, M. Tewari, G. Wang, G. Porter, T. S. Eugene Ng, D. G. Andersen, M. Kaminsky, M. A. Kozuch and A. Vahdat, "Switching the optical divide: fundamental challenges for hybrid electrical/optical datacenter networks," in *Proc. of the 2nd ACM Symposium on Cloud Computing (SOCC2011)*, Article no. 30, 2011. [Article \(CrossRef Link\)](#)
- [19] N. Farrington, A. Forencich, G. Porter, P.-C. Sun, J. E. Ford, Y. Fainman, G. C. Papen and A. Vahdat, "A multiport microsecond optical circuit switching for data center networking," *IEEE Photonics Technology Letters*, vol. 25, no. 16, pp. 1589-1592, 2013. [Article \(CrossRef Link\)](#)
- [20] G. Wang, D. G. Andersen, M. Kaminsky, M. Kozuch, T. S. Eugene Ng, K. Papagiannaki and M.

Ryan, “c-Through: part-time optics in data centers,” in *Proc. of ACM Annual Conf. of the Special Interest Group on Data Communication (SIGCOMM2010)*, 2010. [Article \(CrossRef Link\)](#)



network security, etc.

Wan Tang received the B.S. and M.S. degrees in Computer Application Technology from South Central University for Nationalities, Wuhan, China, in 1995 and 2001, respectively, and received the Ph.D. degree in Communication and Information Systems from Wuhan University, China in 2009. She is currently an associate professor in the College of Computer Science of South-Central University for Nationalities. Also, from 2001 to 2002, she worked as a visiting scholar at the Department of Computer Engineering, Chonbuk National University, South Korea. Furthermore, from 2012 to 2013, she worked as a visiting scholar at the Department of Computer Science and Engineering, SUNY at Buffalo, USA. Her research interests include protocols for optical/wireless networks, software defined networking,



Bo Yi is a M.S. student at the College of Computer Science, South-Central University for Nationalities. He received the B.S. in the College of Computer Science from South-Central University for Nationalities in 2012. His research interests lie in optical networking, software-defined networking, data center networking, etc.



Ximin Yang received the B.S. and M.S. degrees in Computer Application Technology from South-Central University for Nationalities, China in 1994 and 2003, respectively, and received the Ph.D. degree in Computer Architecture from Huazhong University of Science and Technology, China, in 2010. He is currently an associate professor in the College of Computer Science of SCUN. His research interests include network storage system and advance networks.



Jingcong Li is an associate professor in the College of Information Engineering at Minzu University of China. He received the B.S. in Physics from Inner Mongolia University in 1991, and M.S. in Measuring and Testing Technologies and Instruments from Nanjing University of Science and Technology in 1996, respectively. He obtained the Ph.D. in communication and information systems from Peking University, China in 2003. Li was a visiting scholar in the Department of Computer Science and Engineering of SUNY at Buffalo, USA from 2011 to 2012. His research interests include design, modelling and optimization for optical networks.