A Multi-Floor Arrayed Waveguide Grating Based Architecture with Grid Topology for Datacenter Networks

IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received April 24, 2020, accepted May 4, 2020, date of publication May 11, 2020, date of current version June 18, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2993933

A Multi-Floor Arrayed Waveguide Grating Based Architecture With Grid Topology for Datacenter Networks

YUANYUAN HONG[1], DAN ZHANG[2], BING YANG[2], GUANG CHEN[1], YANG WANG[1], QIMAN CHEN[2], XUEZHI HONG[2], AND JIAJIA CHEN[3]

[1]School of Electronics and Information Engineering, Taizhou University, Taizhou 318000, China
[2]ZJU-SCNU Joint Research Center of Photonics, South China Normal University, Guangzhou 510631, China
[3]Chalmers University of Technology, 41296 Gothenburg, Sweden

Corresponding authors: Xuezhi Hong (xuezhi.hong@coer-scnu.org) and Jiajia Chen (jiajiac@chalmers.se)

This work was supported in part by the National Natural Science Foundation of China under Grant 61550110240 and Grant 61671212, in part by the National Natural Science Foundation of Guangdong Province under Grant 2016A030313445, in part by the National Natural Science Foundation of Zhejiang Province under Grant LQ20F010006, Grant LQ18F030002, and Grant LY18F010005, in part by the Swedish Foundation of Strategic Research, in part by the Swedish Research Council, and in part by the Göran Gustafsson Foundation.

ABSTRACT This paper proposes a grid topology based passive optical interconnect (POI) architecture that is composed of multiple floors of arrayed waveguide grating routers (AWGRs) to offer high connectivity and scalability for datacenter networks. In the proposed POI signal only needs to pass one AWGR, and thus can avoid the crosstalk accumulation and cascaded filtering effects, which exist in many existing POI architectures based on cascaded AWGRs. Meanwhile, due to high connectivity, the proposed grid topology based POI also has the potential advantage of high reliability. Simulation results validate the network performance. With a proper node degree, the proposed grid topology can achieve acceptable blocking probability. Besides, steady performance is kept when the number of floors increases, indicating good scalability of the proposed POI.

INDEX TERMS Arrayed waveguide grating router (AWGR), multi-floor, grid topology, network performance.

I. INTRODUCTION

Nowadays datacenters are experiencing an exponential increase in network traffic [1]. As modern datacenters have already consumed 1-2% [2] of global energy, next-generation datacenter network architectures are required to be capable of providing high-bandwidth connectivity and supporting a large scale in an energy-efficient way. Conventional interconnect architectures using commodity electrical switches (such as [3], [4]) are not sustainable with a dramatic traffic increase. Moreover, the topologies employed in these conventional architectures, e.g., fat-tree [3], FiConn [4] and multi-root tree [5], require highly complex cabling, which brings great challenges to the system maintenance and upgrades. Optical fiber communication is considered promising for the datacenter networks because of the ability to offer higher throughput and significantly lower energy consumption compared with electrical interconnects [6]. The wireless communication has attracted the researchers' attention with the ability to offer reconfigurability while reducing the cost by eliminating cabling and switches [7]. However, the wireless solutions of datacenter networks are not efficient for large-scale datacenter networks. Therefore, the optical fiber based interconnects that employ industrially standardized optical fiber communication components become the dominant solutions, particularly for large-scale datacenter networks. In the past years, several optical interconnect architectures have been proposed for high-bandwidth datacenter networks, e.g., Helios [8], Mordia [9], DOS [10] and OSA [11]. To further enhance the energy efficiency of optical datacenter networks, passive optical interconnect (POI) architectures that avoid using active components for interconnections have been proposed.

In the existing POI solutions, there are two major categories, namely the wavelength-routing and broadcast-and-select. The former category is based on wavelength division multiplexing (WDM) devices (e.g., arrayed waveguide

The associate editor coordinating the review of this manuscript and approving it for publication was Yue Zhang.

grating AWG) [12]–[17] and has been found to offer better scalability than the latter one [18]–[20] that typically needs high insertion loss devices, like optical power combiners/splitters. LIONS [12] and H-LIONS [13] use cyclic AWGs and wavelength-tunable transmitters to achieve optical interconnection. The cyclic AWGs with multiple input and output ports are also referred to as AWG routers (AWGRs), performing a key role for wavelength routing. [14] proposes a modular scheme to construct the AWG based interconnection network, which possesses a much lower cabling complexity and improved scalability and reliability. In the cascaded AWG-based (CA)-POI [15] and its extended version [16], two cascaded stages of wavelength routing devices/modules are employed to offer multiple connections. [17] proposes a novel WDM ring network by exploring spatial wavelength reuse of 2×2 AWGRs for datacenters. However, these architectures have limited connectivity. For instance, in [15], [16] as there is only one wavelength routing module placed at the top stage, the number of directly connected links between any two bottom wavelength routing modules is limited to one. Once such a link is occupied or failed, the connection has to pass the other nodes, resulting in the occupation of extra resources and increasing the risk of blocking the upcoming requests.

To further enhance the connectivity, a POI architecture with the heatsink topology capable of supporting multiple concurrent connections in datacenter networks has been proposed [21]. Both the bottom and top stages in the heatsink based POI architecture contain multiple AWGRs, allowing several directly connected links between any pair of bottom AWGRs. However, a connection between a pair of end nodes that are associated with different bottom AWGRs needs to pass several AWGRs, which may significantly degrade the quality of transmission due to the crosstalk accumulation and cascaded filtering effects.

In this regard, in this paper we propose a novel AWGR-based POI architecture that can offer high connectivity and scalability while maintaining transmission quality by avoiding passing too many AWGRs. The proposed approach is referred to as grid topology based architecture, which is composed of multiple floors and can be extended easily. In this paper, we concentrate on network performance and investigate the blocking probability of the proposed grid topology based architecture. Here the blocking probability is defined as the number of the blocked connection requests over the total number of connection requests. The simulation results show that the blocking probability is determined by many factors, including proper routing and wavelength assignment (RWA) algorithms, node degree, the number of floors, and the size of AWGR. A guideline is provided on how to properly configure the grid topology based POI to optimize the blocking probability.

The remainder of the paper is organized as follow. Section II depicts the architecture of the proposed grid topology based POI and discusses the characteristics of the physical architecture and logical topology. Section III presents the network performance of the grid architecture by exploring the impacts of the implemented RWA algorithms and system configurations. A performance comparison with the heatsink topology is also given in this section. Finally, conclusions are drawn in Section IV.

II. GRID TOPOLOGY BASED POI ARCHITECTURE

In this section, we investigate the grid topology based POI from two aspects: the physical architecture and the logical topology. The physical architecture part depicts the connection rules among the connected nodes, while the logical topology part shows the routing possibilities of the grid topology.

A. PHYSICAL ARCHITECTURE

The proposed architecture can be applied in different tiers in datacenters, which means the end nodes connected can represent server, rack or cluster. To illustrate our idea, in this paper we consider the end node represents the server and each AWGR is used to connect servers within a rack. The grid topology consists of several floors, each of which comprises a group of racks. One rack is composed of an AWGR that interconnects different servers. Fig. 1 shows the overall physical architecture of the grid topology based POI. The three floors are sequentially arranged from bottom to top. There are $2N$ racks on each floor. Rack(i, j) denotes the jth ($1 \leqslant j \leqslant 2N$) rack on the ith floor. Each rack has $2N$ servers. The kth ($1 \leqslant k \leqslant 2N$) server in Rack($i$, j) is denoted as Server(i, j, k). The first half of the servers (i.e., $1 \leqslant k \leqslant N$) in every rack on the ith floor are connected to the second half of the servers (i.e., $N < k \leqslant 2N$) in every rack on the $(i+1)$th floor by a specifically defined ''interconnection method'', as shown in Fig. 1. In such a way, more floors (i.e., floors of racks) can be easily added either on top or at the bottom of the architecture, thereby providing good scalability for the datacenter networks.

Fig. 2 shows a detailed internal structure of Rack(i, j) in Fig. 1. Without loss of generality, Server(i, j, k) in Rack(i, j) is used as an example for a detailed description of the connections in Fig. 2. The server has an optical interface (OI) including two sets of transceivers. The receiver r and transmitter t are responsible for the intra-rack connections (e.g., the red line in Fig. 2), while receiver R and transmitter T are responsible for the inter-rack communications (e.g., the yellow and blue lines in Fig. 2). There are two four-port optical circulators (OCs), denoted as the left OC and right OC, respectively, appended to the OI of the server. For the left OC, the 1st-4th ports are connected to the transmitter t, port k' of the AWGR, the 3rd port of the left OC of Server (i', j', k'), and the receiver R, respectively. For the right OC, the 1st-4th ports are connected to the transmitter T, the 2nd port of the right OC of Server(i', j', k'), port k of the AWGR, and the receiver r, respectively. Here the value of i', j', and k' can be calculated by using (1) and (2). For the first half servers in the rack (i.e., $1 \leqslant k \leqslant N$), an additional fiber Bragg grating (FBG) with a specific reflection wavelength is appended at the 3rd port of the left OC of Server

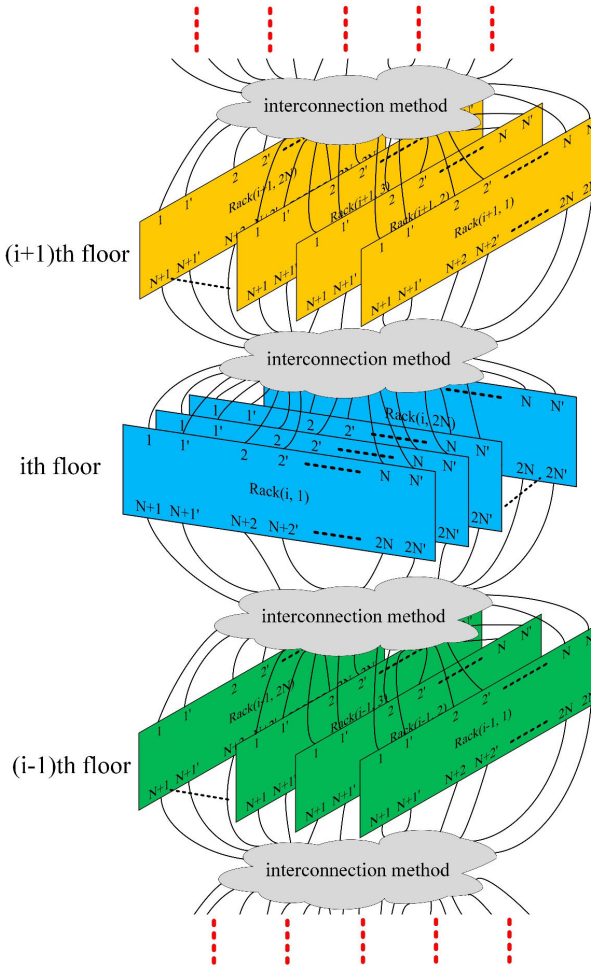

FIGURE 1. grid topology based optical interconnect.

(i, j, k). The "interconnection method" in Fig. 1 represents the connection rules between the servers across the floors, which are expressed by (1) and (2) corresponding to the first half (i.e., $1 \leqslant k \leqslant N$) and the second half servers of each rack (i.e., $N < k \leqslant 2N$), respectively.

$$1 \leq k \leq N \rightarrow \begin{cases} i' = \begin{cases} 1, & i = L \\ i+1, & \text{other} \end{cases} \\ j' = \begin{cases} 2N, & j+k-1 = 2N \\ (j+k-1) \bmod 2N, & \text{other} \end{cases} \\ k' = \begin{cases} 2N, & j = N \text{or} 2N \\ N+j \bmod N, & \text{other,} \end{cases} \end{cases} \tag{1}$$

$$N < k \leq 2N \rightarrow \begin{cases} i' = \begin{cases} L, & i = 1 \\ i-1, & \text{other} \end{cases} \\ j' = \begin{cases} k-N, & -N < a \leq 0 \\ k, & -2N < a \leq -N \\ k, & 0 < a \leq N \end{cases} \\ k' = \begin{cases} a+2N, & -2N < a \leq -N \\ a+N, & -N < a \leq 0 \\ a, & 0 < a \leq N, \end{cases} \end{cases} \tag{2}$$

TABLE 1. Six possible optical paths (I-VI) for inter-rack connections.

Sou. \ Des		$\|i-2\|_{\text{th}}$ FL		$\|i-1\|_{\text{th}}$ FL		i_{th} FL		$\|i+1\|_{\text{th}}$ FL		$\|i+2\|_{\text{th}}$ FL	
		FH	SH	FH	SH	FH	SH	FH	SH	FH	SH
i_{th} FL	FH	-	-	-	-	I	-		V	-	II
	SH	IV	-	VI		-	III	-	-	-	-

Sou.: source server; Des.: destination server; FL: floor; FH: the first half of servers (i.e., $1 \leqslant k \leqslant N$); SH: the second half of servers (i.e., $N < k \leqslant 2N$); the operator $|x|$ stands for modulo L.

where L is the total number of floors of the grid topology, and a equals to $j - k + 1$.

The intra-rack connection can be established by sending signals from the transmitter t via the left OC at the source server to the receiver r at the destination server through its right OC (e.g., the red line in Fig. 2). The wavelength of the intra-rack connection is determined by the wavelength routing table of the AWGR on the same rack. For inter-rack connections in the architecture, there are six different cases, which are summarized in Table 1. The rack on one floor can be connected to the racks on the same floor or on the adjacent two higher/lower floors.

Fig. 3 illustrates the six possible inter-rack optical paths. In the case "$k \leqslant N$", the source Server(i, j, k) belongs to the first half of the racks at the ith floor, and in the case "$k > N$", the source Server(i, j, k) belongs to the second half of the racks at the ith floor.

For cases I-IV (as shown by the yellow and green lines in Fig. 3.), the signal passes through the right OC of the source Server(i, j, k), the right OC of Server(i', j', k'), the AWGR in the Rack(i', j'), the left OC of Server(i', p', q'), a FBG and the left OC of the destination server Server(i, p, q), sequentially. The signal finally arrives at the receiver R of Server (i, p, q). As shown in Fig. 3, for case I, the signal passes through the right OC of the source Server(i, j, k), the right OC of Server$(|i+1|, j1, k1)$, the AWGR in the Rack$(|i+1|, j1)$, the left OC of Server$(i', j2', k2')$, a FBG and the left OC of the destination server Server$(i, j2, k2)$, sequentially. The signal finally arrives at the receiver R of Server $(i, j2, k2)$.

For cases V-VI (as shown by the red lines in Fig. 3), the signal passes through the right OC of the source Server (i, j, k), the right OC of Server(i', j', k'), the AWGR in Rack (i', j') and the left OC of the destination Server(i', j', q) in Rack(i', j'), sequentially. Then the signal is reflected by a FBG and enters the left OC of the destination server again and is detected by the receiver R. For instance, as shown in Fig. 3, for case V, the signal passes through the right OC of the source Server(i, j, k), the right OC of Server$(|i+1|, j1, k1)$, the AWGR in Rack$(|i+1|, j1)$, and the left OC of the destination Server$(|i+1|, j1, k1)$ in Rack$(|i+1|, j1)$ sequentially. Then the signal is reflected by a FBG and enters the left OC of the destination server again and is detected by the receiver R. Note that for a specific server, there only exists three possible optical paths for inter-rack connections, which depend on whether the server belongs to the first or the second half of servers in the rack. For the first (second) half of servers

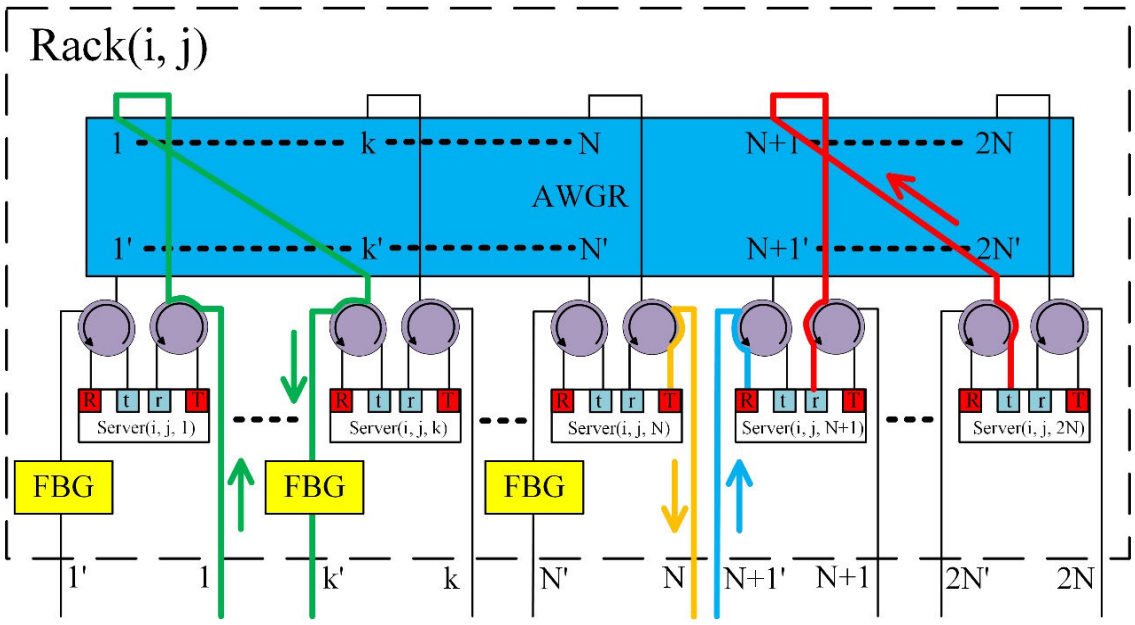

FIGURE 2. The detailed interconnection structure of Rack(*i*, *j*).

in the rack, they can be connected to the first (second) half of servers of other racks in the same floor, servers in the racks that are one floor higher (lower), and the second (first) half of servers in the racks that are two floors higher (lower).

For all cases, the signal passes through the OC four times but the AWGR only once, and then goes through or is reflected by the FBG once. Nevertheless, the wavelength requirement varies for different cases. For cases I-IV, the wavelength that is routed by the AWGR from the ingress port k' to the egress port q' shall not coincide with the reflected wavelength of the FBG. On the contrary, for the inter-rack connections between two adjacent floors that can be represented by cases V-VI, they require the matching of the reflected wavelength of the FBG and the wavelength that is routed by the AWGR from the ingress port k' to the egress port q. The wavelength assignment can be obtained according to the wavelength routing principle listed in Tables 5 and 6 in APPENDIX. Though the constraints on wavelength exist in the proposed architecture, the proposal is still able to provide multiple concurrent connections. The optical paths of any type of interconnects pass one AWG, one FBG and several circulators. With proper design (e.g., athermal packaging, integrated temperature controller, etc.), the off-the-shelf AWG and FBG products can have high wavelength stability against the thermal variation within a large operating temperature range [22], [23]. The reflection passbands of all FBGs are the same as the wavelengths on the diagonal of the wavelength routing table of the AWGR (see Table 5 in APPENDIX). As both the off-the-shelf FBG and AWGR can be made according to the ITU grid with high wavelength accuracy, it is of high feasibility to align their spectrum. With the good matching of the AWGR and FBG passbands [22], [23] and passing the AWGR only once, crosstalk accumulation, which is recognized as a dominating limiting factor in the previous cascaded AWG based schemes (such as that in the heatsink topology [24]), can be avoided without compromising its rich connectivity advantage offered by the multiple concurrent connections.

B. LOGICAL TOPOLOGY

Fig. 4 shows the logical connection of the grid topology on the same floor. The AWGR size is chosen to be 4 as an example. The large dashed circles indicate the racks from Rack(i, 1) to Rack(i, 4), where i ($i = 1, 2, 3 \ldots$) indicates the index of the floor. The small solid circles in the large dashed circles indicate the servers from Server(i, j, 1) to Server(i, j, 4), where j ($j = 1, 2, 3, 4$) indicates the index of the rack. There are three different colors for the links: blue, red, and black. The color of the links represents the wavelength assigned. The connection inside the large dashed circle shows the intra-rack transmission and the connection between the large dashed circles shows the inter-rack transmission on the same floor.

Fig. 5 and Fig. 6 show the logical topology between the different floors of the grid topology. Fig. 5 shows the logical connection of the first half of the ith floor and Fig. 6 shows the logical connection of the second half of the ith floor. The size of the AWGRs of these two figures are 4×4 and all the connections are of the same wavelength. The first half of the ith floor is connected to the second half of the $|i + 1|$th floor (see the blue lines in Fig. 5) and the second half of the $|i+2|$th floor (see the red lines in Fig. 5), respectively. This is consistent with the connections in the physical architecture. Similarly, as we can see in Fig. 6, the second half of the ith

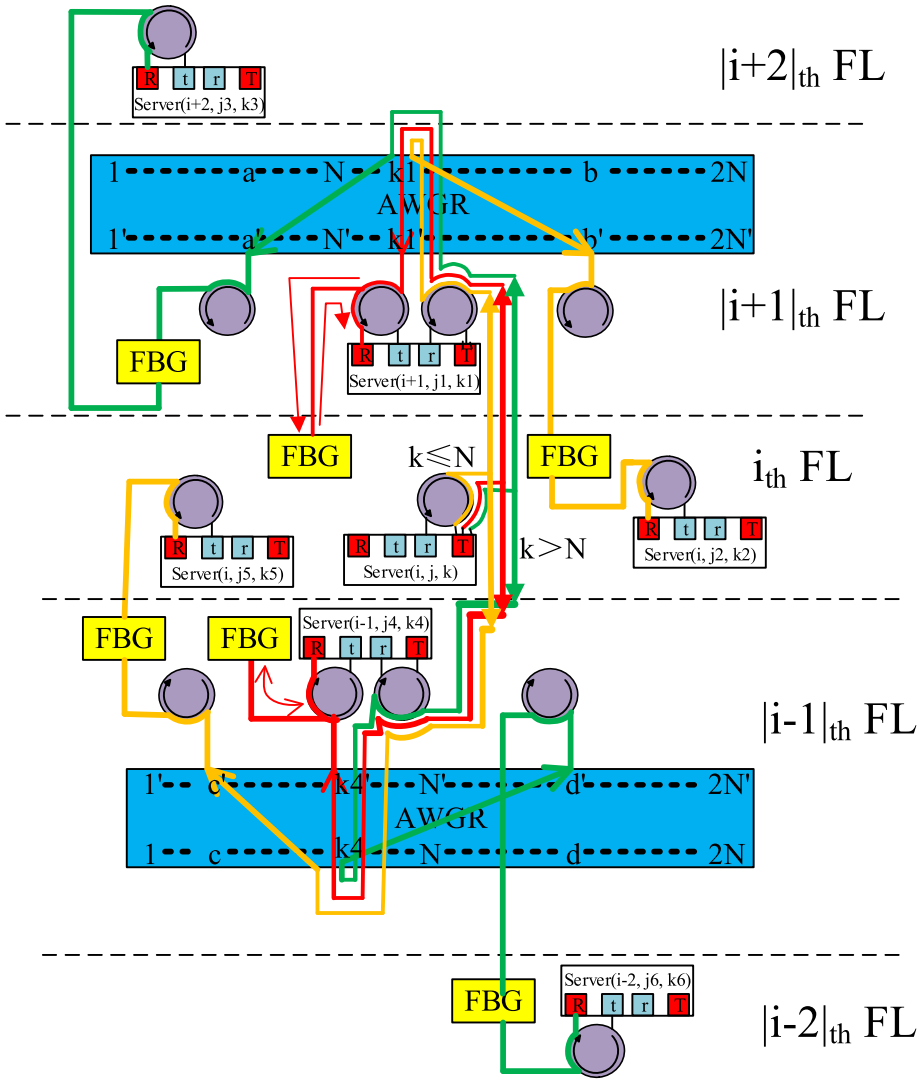

FIGURE 3. Six possible optical paths (I-VI) for inter-rack connections.

floor is connected to the first half of the $|i-1|$th floor (the blue lines in Fig. 6) and the first half of the $|i-2|$th floor (the red lines in Fig. 6), respectively.

For the 4×4 AWGR, the first half of the ith floor has one connection to the second half of the $|i+1|$th floor, and two connections to the second half of $|i+2|$th floor. For a general case with $N \times N$ AWGRs, the first half of the ith floor has one connection to the second half of the $|i+1|$th floor, and $N/2$ connections to the second half of $|i+2|$th floor (as shown in Table 2).

TABLE 2. The number of links for inter-rack connections.

Sou. \ Des.		$\|i-2\|_{th}$ FL		$\|i-1\|_{th}$ FL		i_{th} FL		$\|i+1\|_{th}$ FL		$\|i+2\|_{th}$ FL	
		FH	SH	FH	SH	FH	SH	FH	SH	FH	SH
i_{th} FL	FH	-	-	-	-	$N/2-1$	-		1	-	$N/2$
	SH	$N/2$	-	1		-	$N/2-1$	-	-	-	-

N refers to a half of the number of AWGRs on each floor.

III. NETWORK PERFORMANCE EVALUATION

In this section, the discrete event driven simulator using C++ is employed to investigate the network performance. The discrete event driven simulator is a widely adopted approach to simulate the dynamic system of a communication network and evaluate the network performance [25]. The events (e.g. the arrival and end of communication requests, the assignment and release of the network resources, etc.) are inserted into the event queue following the order of time. Then they are processed in the order of event sequence. The arriving connection request is handled by the RWA algorithm, which decides whether the request is served or blocked. The blocking probability, which equals to the number of blocked requests over the number of total requests, can be calculated after simulation with a sufficient amount of requests. In other words, with the discrete event driven simulator and the RWA

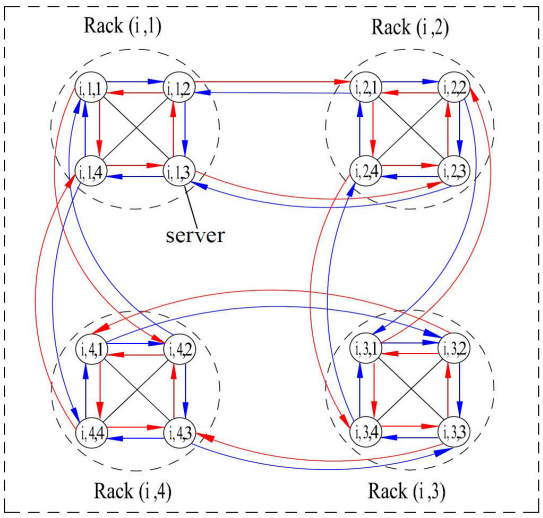

FIGURE 4. The logical topology on the same floor for the proposed grid topology based POI.

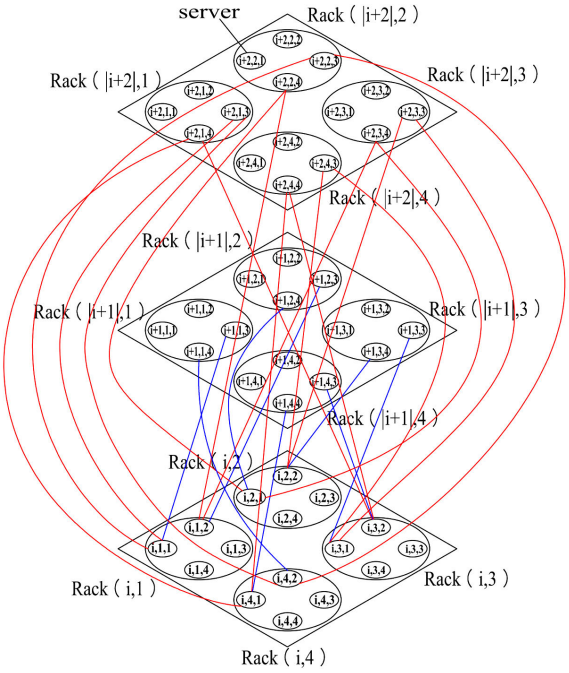

FIGURE 5. Logical topology of cross-floor based optical interconnect (the first half of *i*-th floor).

algorithm, the dynamic process of the communication network system can be simulated and the blocking probability performance can be assessed. In the discrete event driven simulator, the arrival of requests follows the Poisson process and their holding time is exponentially distributed with a mean value normalized to 1 so that the arrival rate per node directly reflects the network load. Every point shown in result figures is obtained by averaging the outcomes of running 4 simulations with different seeds and each seed processing over 200,000 requests. Different traffic distribution patterns have been investigated to emulate different locality levels of traffic distributed in datacenter networks, which represent

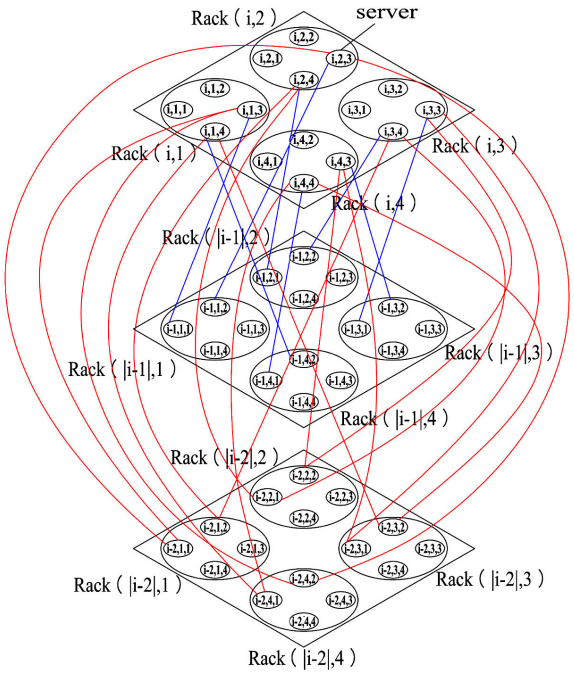

FIGURE 6. Logical topology of cross-floor based optical interconnect (the second half of *i*-th floor).

different types of applications hosted in the datacenters [26]. They are denoted as "uniform" and "50%". For the "uniform" case, the destination of each request follows the uniform distribution, while the "50%" case has a probability of 50% of connection requests staying in the same rack, representing the high traffic locality. In this paper, "$D = a + b$" means that the numbers of transceivers are a and b for intra-rack and inter-rack transmissions, respectively. The number of floors L is set from 2 to 5. The profile used in the simulator including information of network topology, node degree D, number of floors L, traffic distribution pattern, etc. is determined according to the tested scenarios. In this way, the network performance of different topologies in various scenarios is evaluated numerically in terms of blocking probability.

The impacts of the routing and wavelength assignment (RWA) algorithms and system configurations (e.g., the number of floors, the number of transceivers per node, i.e. the node degree D) on network performance in terms of blocking probability are investigated. The performance of the heatsink scheme is also measured for comparison. Based on the numerical results, a guideline is provided on how to properly configure the proposed grid topology based POI to achieve acceptable blocking probability.

A. IMPACT OF THE RWA ALGORITHMS

To understand how the RWA algorithms affect the network performance, two types of RWA algorithms are considered. In the CA-POI architecture, a simple RWA approach is applied, where K-shortest approach is for routing and the first-fit algorithm is for wavelength assignment.

As a benchmark, this approach is referred to as $A1$. The idea of this approach is to find up to Ksp shortest paths for any coming request and choose the first path having available wavelength resource to be assigned. Although this approach performs well for the CA-POI [16], the heatsink and grid topologies have much better connectivity, and hence may not perform efficiently if employing such an algorithm with a fixed routing scheme. Therefore, a wavelength plane approach [27] is implemented, referred to as $A2$. In $A2$, a logical plane that takes into account the available wavelength and transceiver information is set up to find the shortest path. The wavelength plane consisting of the available wavelength information is appended to the original logical topology where the routing and wavelength assignment are executed together. Therefore, the size of the logical topology expands, in which the number of nodes becomes the total number of servers ($2N \times 2N \times L$) multiplied with the number of wavelengths ($2N$). As a result, the complexity of A2 is significantly increased, which can be referred to Table 3. On the other hand, the information of available wavelengths and transceivers is considered in the cost function, i.e. the weight used in the shortest path calculation, which helps to improve the blocking probability. The cost function used in $A2$ is as follow: If a link is not in use, its cost is set to be equal to initial-cost (1 is applied in the simulation); If a link is used in a specific connection request, its cost is set to be a large value, referred to as $COST1$ (500 is applied in the simulation); if a link is not able to be used because all the transceivers at the node are occupied, its cost is set to another large value, referred to as $COST2$ (600 is applied in the simulation to be differentiated from $COST1$). To illustrate the cost function used in A2, the values of cost(x, y), i.e. the cost of link(x, y) directly connecting the node pair(x, y), under three different conditions are listed in (3). The weight of any path equals to the sum of the costs of all links in the path. The path with the minimum cost is obtained by the Dijkstra shortest path algorithm [28]. If the minimum value of total cost (most of the values ranging from 1 to 7) is obviously less than $COST1$/$COST2$, the path is then chosen as the route for the connection request. Otherwise, the connection request is blocked. Once the wavelength is successfully assigned to the request, the costs of the selected links have to be changed from initial-cost to $COST1$. If all of the transceivers at the nodes that are associated with the selected links are occupied, the costs of the links are set to $COST2$. In this approach the information of available transceivers and wavelengths is taken into account to find the shortest path. The wavelength assignment is carried out together with the routing.

$$\cos t(x, y) \rightarrow \begin{cases} 1, & \text{if link}(x, y) \text{ is available,} \\ COST1, & \text{if link}(x, y) \text{ is used,} \\ COST2, & \text{if link}(x, y) \text{ is not able to be used} \\ & \text{because all the transceivers at node} \\ & x \text{ or } y \text{ are occupied.} \end{cases} \tag{3}$$

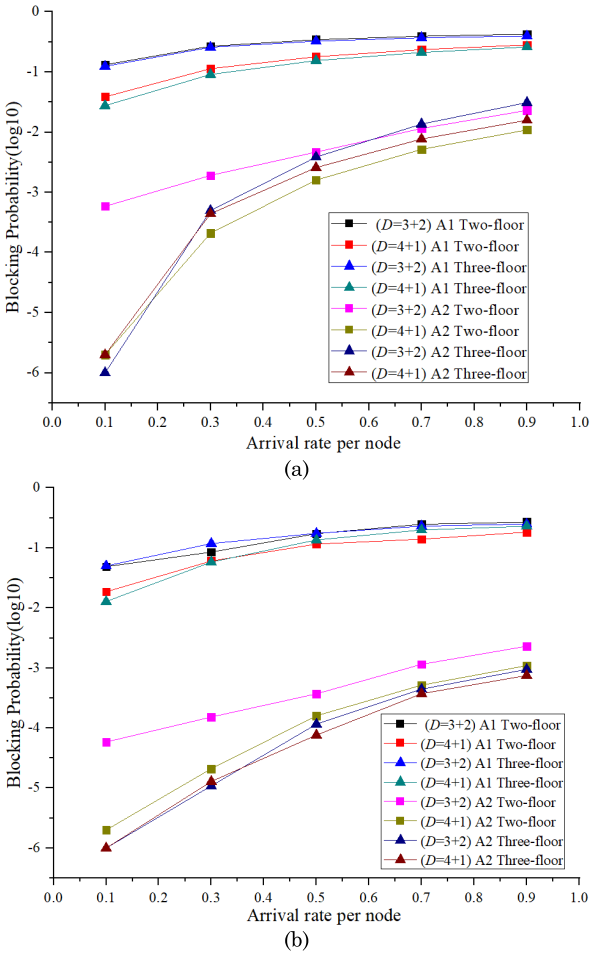

FIGURE 7. Blocking probability when different algorithms are employed in grid topology with $D = 5$ and $N = 16$ in the traffic pattern of (a) "uniform" and (b) "50%".

Fig. 7 shows the blocking probability of the grid topology in two RWA algorithms. The results show the blocking probability as a function of arrival rate per node (i.e., the number of arrival requests per time unit divided by the number of total nodes that are able to be connected to the POI). For the results shown in Fig. 7, the number of shortest paths for $A1$ is set to 5, which is sufficiently large to get statured performance. The size of AWGR is 16 and the number of transceivers (i.e., node degree D, the maximum number of links established for each node) is 5. The traffic pattern is "uniform" in (a) and "50%" in (b). "Two-floor" and "Three-floor" indicate the number of total floors of the grid topology is two and three, respectively.

From Fig. 7 we can see that $A2$ has obviously lower blocking probability than $A1$ in most cases. It clearly shows that a proper algorithm needs to be chosen in order to leverage the increased connectivity in the grid topology. Therefore, in the later part of this section, $A2$ is selected for RWA.

Table 3 shows the difference between A1 and A2. In A2, the routing and wavelength assignment algorithms are carried out together, while they are separate in A1. This leads to better utilization of the high connectivity of the grid topology and

TABLE 3. The comparison of A1 and A2.

Complexity	Relationship of routing and wavelength assignment		Blocking probability	
A1	Routing	Wavelength assignment	Separate	High
	$O(4N^2\times4N^2\times L^2)$	$O(Ksp\times2N)$		
A2	Routing and wavelength assignment		Joint	Low
	$O(4N^2\times4N^2\times L^2\times4N^2)$			

L is the number of floors and $2N$ equals the number of AWGRs on each floor in the grid topology. Ksp indicates the number of shortest paths calculated by the Dijkstra routing algorithm.

TABLE 4. Comparison of heatsink and grid topologies.

	Heatsink	Grid
The number of AWGRs	$K+2N$ $(K<2N)$	$2N\times L$
The number of floors	1	$L(\geqslant 2)$
The number of servers	$2N\times(2N-K)$	$2N\times2N\times L$
Total fiber links	$2N\times(2N-K)\times(2N-1)$	$2N\times2N\times L\times(4N-1)$
Intra-rack fiber links	$2N\times(2N-K)\times(2N-K-1)$	$2N\times2N\times L\times(2N-1)$
Inter-rack fiber links	$2N\times(2N-K)\times K$	$2N\times2N\times L\times2N$
Maximum number of AWGRs that need to be passed for any direct optical connections	3	1

L refers to the number of floors in the grid topology. K refers to the number of top AWGRs in the heatsink topology. For both grid and heatsink topologies, we consider the size of AWGRs is $2N\times2N$.

better blocking probability performance when employing A2. However, this also results in higher complexity in A2 than in A1. The complexity determines the calculation time, that depends on the key parameters of the RWA algorithm (e.g., Ksp) and POI architecture (e.g., L, N). The complexity of the Dijkstra routing algorithm is $O(V^2)$ where V indicates the number of nodes in the logical plane. With A1 the number of nodes equals the number of servers in the topology. With A2 the logical plane is expanded and the number of nodes in the logical topology equals the number of servers multiplied by the number of wavelengths. The complexity shown in Table 3 is for the grid topology. The complexity of the heatsink topology can be deducted similarly.

B. IMPACT OF THE NUMBER OF FLOORS

The grid topology can achieve cross-floor transmission, and therefore the number of total floors is also a key parameter in the network performance. In this section, we choose four cases: "Two-floor", "Three-floor", "Four-floor" and "Five-floor". Fig. 8 shows the blocking probability of the grid topology as a function of arrival rate per node for the system with the size of AWGR of 16. In Fig. 8, the requests are randomly distributed among different floors. The traffic case is "uniform" and the number of transceivers is 5, where $a = 2$ and $b = 3$.

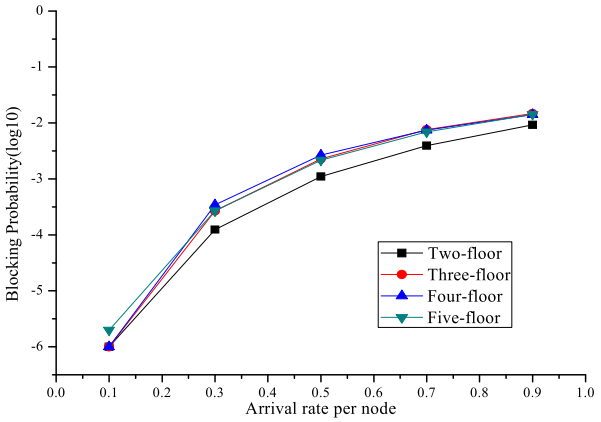

FIGURE 8. Blocking probability of grid topology when the value of N is 16 and D is 5 ($a = 2$ and $b = 3$).

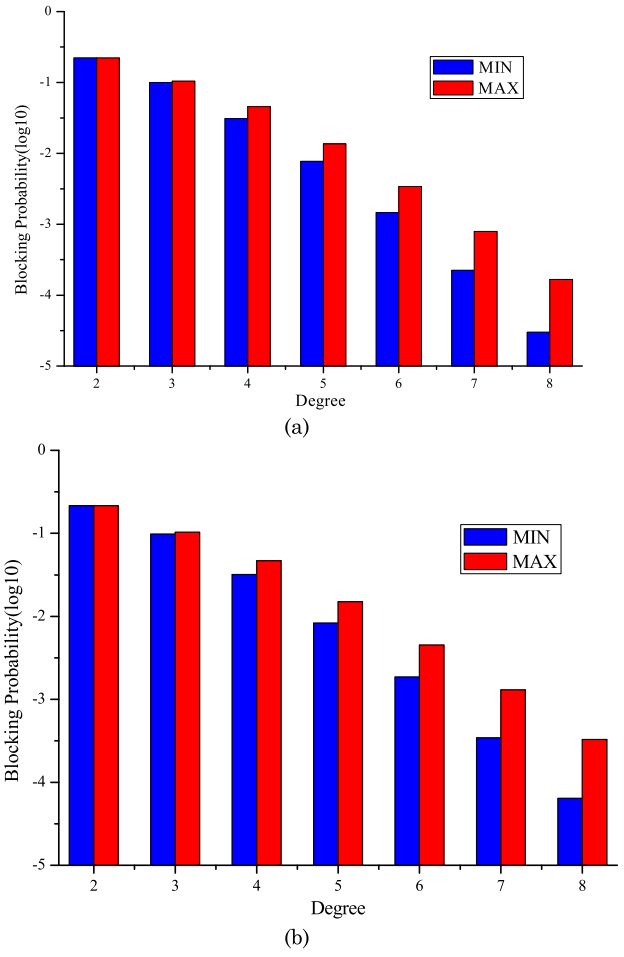

FIGURE 9. Blocking probability of grid topology versus degree with parameters: arrival rate per node = 0.7; the value of total floor is 3 under traffic case of "uniform" when the size of AWGR is: (a) 16 and (b) 32.

From Fig. 8, we can see the blocking probability almost remains the same with the increasing of the number of floors. It is clear that the trend of the network performance is similar regardless of the number of floors, indicating good scalability of the grid topology. Therefore, the number of floors in the grid topology can be added up to request.

TABLE 5. Wavelength routing table of AWGR.

Ingress \ Egress	1'	2'	…	$(k-1)'$	k'	$(k+1)'$	…	$(2N-1)'$	$2N'$
1	λ_1	λ_2	…	λ_{k-1}	λ_k	λ_{k+1}	…	λ_{2N-1}	λ_{2N}
2	λ_{2N}	λ_1	…	λ_{k-2}	λ_{k-1}	λ_k	…	λ_{2N-2}	λ_{2N-1}
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
$k-1$	λ_{2N-k+3}	λ_{2N-k+4}	…	λ_1	λ_2	λ_3	…	λ_{2N-k+1}	λ_{2N-k+2}
k	λ_{2N-k+2}	λ_{2N-k+3}	…	λ_{2N}	λ_1	λ_2	…	λ_{2N-k}	λ_{2N-k+1}
$k+1$	λ_{2N-k+1}	λ_{2N-k+2}	…	λ_{2N-1}	λ_{2N}	λ_1	…	λ_{2N-k-1}	λ_{2N-k}
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
$2N-1$	λ_3	λ_4	…		λ_{k+2}		…	λ_1	λ_2
$2N$	λ_2	λ_3	…		λ_{k+1}		…	λ_{2N}	λ_1

TABLE 6. The components passed in the six optical paths as well as the ingress and egress ports of AWGR.

Comp. \ Case	I	II	III	IV	V	VI
1st	right OC Server(i,j,k)	right OC Server(i,j,k)	right OC Server(i,j,k)	right OC Server(i,j,k)	right OC Server(i,j,k)	right OC Server(i,j,k)
2nd	right OC Server $(i+1,j_1,k_1)$	right OC Server $(i+1,j_1,k_1)$	right OC Server$(i-1,j_4,k_4)$	right OC Server$(i-1,j_4,k_4)$	right OC Server$(i+1,j_1,k_1)$	right OC Server$(i-1,j_4,k_4)$
3rd	AWGR Rack$(i+1,j_1)$	AWGR Rack$(i+1,j_1)$	AWGR Rack$(i-1,j_4)$	AWGR Rack$(i-1,j_4)$	AWGR Rack$(i+1,j_1)$	AWGR Rack$(i-1,j_4)$
4th	left OC Server$(i+1,j_1,b')$	left OC Server $(i+1,j_1,a')$	left OC Server$(i-1,j_4,c')$	left OC Server$(i-1,j_4,d')$	left OC Server$(i+1,j_1,k_1)$	left OC Server$(i-1,j_4,k_4)$
5th	FBG Server(i,j_2,k_2)	FBG Server $(i+1,j_1,a')$	FBG Server$(i-1,j_4,c')$	FBG Server$(i-2,j_6,k_6)$	FBG Server(i,j,k_1)	FBG Server$(i-1,j_4,k_4)$
6th	left OC Server(i,j_2,k_2)	left OC Server$(i+2,j_3,k_3)$	left OC Server(i,j_5,k_5)	left OC Server$(i-2,j_6,k_6)$	left OC Server$(i+1,j_1,k_1)$	left OC Server$(i-1,j_4,k_4)$
Note (cf. Eqs. (1)-(2))	$(i+1,j_1,k_1)$= (i',j',k') $(i+1,j_1,b')$= (i',j_2',k_2')	$(i+1,j_1,k_1)$= (i',j',k') $(i+1,j_1,a')$= $((i+2)',j_3',k_3')$	$(i-1,j_4,k_4)$= (i',j',k') $(i-1,j_4,c')$= (i',j_5',k_5')	$(i-1,j_4,k_4)$= (i',j',k') $(i-1,j_4,d')$= $((i-2)',j_6',k_6')$	$(i+1,j_1,k_1)$= (i',j',k')	$(i-1,j_4,k_4)$= (i',j',k')
AWGR ports: Ingress	k_1	k_1	k_4	k_4	k_1	k_4
AWGR ports: Egress	b'	a'	c'	d'	k_1'	k_4'

C. IMPACT OF NODE DEGREE

Because the transceivers are allocated differently for the intra-rack and inter-rack connections, the blocking probability for the same number of transceivers is not the same. For every node, here we choose two blocking probability values: MIN and MAX, representing the minimum and maximum values of blocking probability that can be achieved when assigning the different number of transceivers for intra-rack and inter-rack connections, respectively. From Fig. 9 one can see that the blocking probability significantly depends on the node degree. The network performance can be improved by increasing the node degree. The size of AWGR is chosen to be 16 or 32 considering the size of commercially available AWGR products [23]. It is clear that the trend of the network performance is similar regardless of the size of AWGRs. The results can be explained as when the degree becomes larger there are more transceivers per node and more available paths can be found to improve the network performance. Therefore, the blocking probability can be decreased drastically with a larger node degree.

D. COMPARISON OF THE HEATSINK AND GRID

We compare the blocking probability as a function of arrival rate per node of the heatsink topology and grid topology. The number of transceivers is 5 and the traffic distribution pattern is "uniform". The size of the AWGR is 16 and 32 in Fig. 10(a) and Fig. 10(b), respectively. Besides, the number of top AWGRs in the heatsink topology K is 1, 2 and 3. Moreover, all the connection requests of the grid topology stay on the same floor. From Fig. 10, we can see the blocking probability of the grid topology is smaller than that in the heatsink topology regardless of the value of K. Furthermore, the simulation results are similar regardless of the size of AWGR. As mentioned in Section I, there are two layers of the AWGRs (i.e., the top and bottom AWGRs) in the heatsink topology. For the nodes associated with the same bottom AWGR, they can communicate with each other with a one-hop direct connection via this AWGR. However, the interconnection between a pair of nodes associated with two different bottom AWGRs has to pass several AWGRs in the heatsink topology. K denotes the number of the top AWGRs

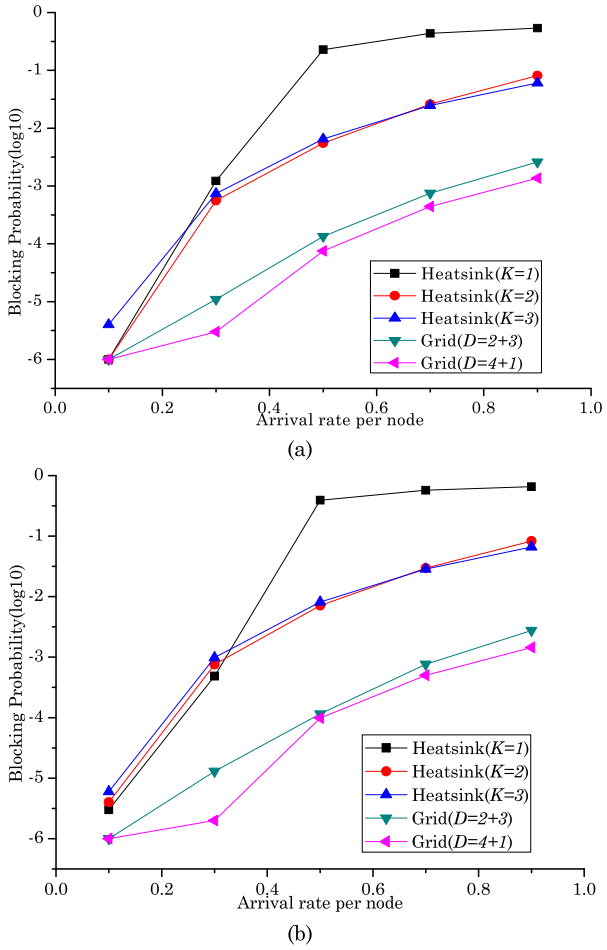

FIGURE 10. Blocking probability of heatsink topology and grid topology when the size of AWGR is: (a) 16 and (b) 32.

in the heatsink topology, which allows up to K paths for the connection of a pair of the nodes associated with different bottom AWGRs. It can be seen from Fig. 10 when increasing K to 2, the blocking performance can be significantly improved. However, such an improvement becomes minor when continuing to increase K.

Table 4 shows the differences between the heatsink and grid topologies. Here $2N \times 2N$ AWGRs are used in two topologies and the number of top AWGRs is K in the heatsink topology. From Table 4 we can see that the grid topology achieves more connections per AWGR compared with the heatsink topology, especially for inter-rack links. It means the grid topology has higher connectivity, leading to better blocking probability performance as shown in Fig. 10. Moreover, the grid topology can be extended vertically with L floors, showing good scalability. On the other hand, the grid topology has a higher complexity.

IV. CONCLUSION

We have proposed one passive optical interconnect architecture, named as the grid topology. We introduce the physical structure and logical topology of the proposed architecture. Compared with the heatsink topology, the grid topology can increase connectivity, scalability and reliability, while reducing the number of AWGRs that has to pass through for each connection. The simulation results have demonstrated that the network performance of the grid topology is better than the heatsink topology and the grid topology is more appropriate for the optical interconnection in datacenters. Furthermore, the numerical results show that for the grid topology the network performance almost remains the same when the number of floors increases and can be improved by increasing the node degree.

Moreover, it should be noted that although the required numbers of FBGs and AWGRs of the grid topology increase when the number of servers to be connected grows, the maximum numbers of hops and required wavelengths are not necessarily high. The rich connectivity and high energy and cost efficiency have also made the POI approaches like the grid topology become attractive solutions for datacenter networks [19], [29]. On the other hand, the proposed scheme employs circuit switching, which is more suitable for the high bandwidth application in datacenter networking [30]–[32]. The nodes in the grid topology cannot only represent servers in the edge tier, but also the racks/clusters in the aggregation and core tiers. When the grid topology is used for the inter-rack/cluster interconnections, the advantage of circuit switching of the grid topology makes it more suitable for the datacenter networks. On the other hand, the control signals are sent on the packet basis in the control plane [18] which is separated from the circuit switching in the data plane. Software-defined networking (SDN) as a centralized controller recently is considered as a promising protocol that can be employed for the proposed scheme. We believe the grid topology could be a step forward to show the possibility to have circuit switching to support scalable data transmission but at the same time brings a challenge on control signaling that calls for research efforts.

The advantages of network connectivity and reliability make the grid topology a promising approach for datacenter networks. The advantages are provided by the multiple concurrent connections and passing only one AWGR for any direct optical connections. To be more specific, the topologies with multiple AWGRs that need to be passed for direct connections [21] suffer from the degradation of transmission quality, resulting in the restriction of some connections due to unacceptable transmission quality. The grid topology relieves the restriction and makes more connections available. On the other hand, the grid topology offers rich concurrent connections, especially for the inter-rack communications, with the proper design of architecture. The superiority of the number of connections per AWGR, especially the inter-rack connections, of the grid topology over the heatsink topology is shown in Table 4. This leads to better blocking probability performance of grid topology as shown in Fig. 10. When compared to the POI architectures [15], [16] where the number of directly connected links is limited to one, the advantage of grid topology is expected to be more significant. Moreover,

the rich connections not only offer better connectivity but also provide backup solutions once the original links are occupied or failed. This leads to higher reliability and the reduction of failures of the arriving requests. The further analysis of reliability and resilience of the proposed grid topology based datacenter networks will be included in our future research. The problem of reliability analysis can be rather complex, especially when the datacenters become the essential infrastructure to support applications like 5G, Internet of Things (IoT), etc. For future work, we will further analyze the reliability of the proposed datacenter networks in complex dynamic scenarios [33]–[36]. It should be noted that the paper focuses on the architecture design and blocking probability analysis of the grid topology based datacenter networks. Multi-objective optimization [37], [38] will be an important research direction for future work to improve the blocking probability and reliability at the same time, especially when the networks are facing the significant challenge of continuously increasing traffic while satisfying the requirement of low energy consumption.

APPENDIX
AWGR ROUTING TABLE AND COMPONENTS IN THE SIX OPTICAL PATHS

The wavelength routing table of $N \times N$ AWGR is shown in Table 5, and the components together with ingress and egress ports of the AWGR in the six optical paths are shown in 6.

ACKNOWLEDGMENT

(Yuanyuan Hong and Dan Zhang are co-first authors.)

REFERENCES

[1] M. N. Rahman and A. Esmailpour, "A hybrid electrical and optical networking topology of data center for big data network," presented at the ASEE Zone I Conf., Bridgeport, CT, USA, Apr. 3–5, 2014.

[2] New Scientist. *Internet Responsible for 2 Percent of Global Energy Usage*. Accessed: Dec. 7, 2019. [Online]. Available: https://cacm.acm.org/news/139386-internet-responsible-for-2-percent-of-global-energy-sage/fulltext

[3] G. Guelbenzu, N. Calabretta, and O. Raz, "Hybrid fat-tree: Extending fat-tree to exploit optical switch transparency with WDM," *Opt. Fiber Technol.*, vol. 44, pp. 89–101, Aug. 2018.

[4] D. Li, C. Guo, H. Wu, K. Tan, Y. Zhang, S. Lu, and J. Wu, "Scalable and cost-effective interconnection of data-center servers using dual server ports," *IEEE/ACM Trans. Netw.*, vol. 19, no. 1, pp. 102–114, Feb. 2011.

[5] M. F. Bari, R. Boutaba, R. P. Esteves, L. Z. Granville, M. Podlesny, M. Rabbani, Q. Zhang, and M. F. Zhani, "Data center network virtualization: A survey," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 2, pp. 909–928, 2nd Quart., 2012.

[6] M. Xu, C. Liu, and S. Subramaniam, "PODCA: A passive optical data center network architecture," *J. Opt. Commun. Netw.*, vol. 10, no. 4, pp. 409–420, Apr. 2018.

[7] A. Celik, B. Shihada, and M.-S. Alouini, "Optical wireless data center networks: Potentials, limitations, and prospects," *Proc. SPIE*, vol. 10945, Feb. 2019, Art. no. 109450.

[8] 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," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 40, no. 4, pp. 339–350, Aug. 2010.

[9] N. Farrington, A. Forencich, P.-C. Sun, S. Fainman, J. Ford, A. Vahdat, G. Porter, and G. Papen, "A 10 μs hybrid optical-circuit/electrical-packet network for datacenters," presented at the Opt. Fiber Commun. Conf., Anaheim, CA, USA, Mar. 19–21, 2013.

[10] X. Ye, Y. Yin, S. J. B. Yoo, P. Mejia, R. Proietti, and V. Akella, "DOS: A scalable optical switch for datacenters," in *Proc. 6th ACM/IEEE Symp. Archit. Netw. Commun. Syst. (ANCS)*, San Diego, CA, USA, 2010, p. 24.

[11] 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 Trans. Netw.*, vol. 22, no. 2, pp. 498–511, Apr. 2014.

[12] Y. Yin, R. Proietti, X. Ye, C. J. Nitta, V. Akella, and S. J. B. Yoo, "LIONS: An AWGR-based low-latency optical switch for high-performance computing and data centers," *IEEE J. Sel. Topics Quantum Electron.*, vol. 19, no. 2, Mar. 2013, Art. no. 3600409.

[13] R. Proietti, Z. Cao, C. J. Nitta, Y. Li, and S. J. B. Yoo, "A scalable, low-latency, high-throughput, optical interconnect architecture based on arrayed waveguide grating routers," *J. Lightw. Technol.*, vol. 33, no. 4, pp. 911–920, Feb. 15, 2015.

[14] T. Ye, T. T. Lee, M. Ge, and W. Hu, "Modular AWG-based interconnection for large-scale data center networks," *IEEE Trans. Cloud Comput.*, vol. 6, no. 3, pp. 785–799, Jul. 2018.

[15] X. Hong, Y. Gong, Y. Yang, and J. Chen, "AWG based passive optical interconnects for datacenters," in *Proc. Adv. Photon. (IPR, NOMA, Sensors, Netw., SPPCom, SOF)*, Vancouver, BC, Canada, Jul. 2016, Paper NeTu1C.3.

[16] X. Hong, Y. Yang, Y. Gong, and J. Chen, "Passive optical interconnects based on cascading wavelength routing devices for datacenters: A cross-layer perspective," *J. Opt. Commun. Netw.*, vol. 9, no. 4, pp. C45–C53, Apr. 2017.

[17] I.-F. Chao and C.-H. Lee, "AWG-based WDM ring networks: high-performance and low-cost system designs," *Comput. Netw.*, vol. 145, pp. 64–75, Nov. 2018.

[18] M. Fiorani, S. Aleksic, M. Casoni, L. Wosinska, and J. Chen, "Energy-efficient elastic optical interconnect architecture for data centers," *IEEE Commun. Lett.*, vol. 18, no. 9, pp. 1531–1534, Sep. 2014.

[19] Y. Gong, X. Hong, Y. Lu, S. He, and J. Chen, "Passive optical interconnects at top of the rack: Offering high energy efficiency for datacenters," *Opt. Express*, vol. 23, no. 6, pp. 7957–7970, Mar. 2015.

[20] P. E. Green, L. A. Coldren, K. M. Johnson, J. G. Lewis, C. M. Miller, J. F. Morrison, R. Olshansky, R. Ramaswami, and E. H. Smithand, "All-optical packet-switched metropolitan-area network proposal," *J. Lightw. Technol.*, vol. 11, nos. 5–6, pp. 754–763, May/Jun. 1993.

[21] B. Yang, Y. Gong, X. Hong, and J. Chen, "Heatsink topology based optical interconnect for datacenters," presented at the IEEE/OSA/SPIE ACP, Guangzhou, China, Nov. 10–13, 2017.

[22] MPB Communications Inc. *50/100 GHz DWDM Gratings*. Accessed: Dec. 7, 2019. [Online]. Available: https://mpbcommunications.com/en/site/products/fiber_Bragg_gratings/dwdm/index.html

[23] Enablence Technologies Inc. *NxN AWG Multiplexers and Demultiplexers Router Module (APRTE)*. Accessed: Dec. 7, 2019. [Online]. Available: http://www.enablence.com/technologies/products/awg/nxn-awg-multiplexers-and-demultiplexers-router-module-aprte/

[24] D. Zhang, B. Yang, Q. Chen, Y. Yang, and J. Chen, "Heatsink topology offering high connectivity for datacenters," presented at the SPIE ICOCN, Zhuhai, China, 2018, pp. 16–19.

[25] H. T. Mouftah and R. P. Sturgeon, "Distributed discrete event simulation for communication networks," *IEEE J. Sel. Areas Commun.*, vol. 8, no. 9, pp. 1723–1734, Dec. 1990.

[26] Y. Yang and J. Chen, "Network performance analysis of an AWG-based passive optical interconnect for datacenters," in *Proc. Asia Commun. Photon. Conf.*, Wuhan, China, 2016, Paper ATh3E.1.

[27] P. Wiatr, J. Chen, P. Monti, and L. Wosinska, "Energy efficiency versus reliability performance in optical backbone networks," *IEEE/OSA J. Opt. Commun. Netw.*, vol. 7, no. 3, pp. A482–A491, Mar. 2015.

[28] E. W. Dijkstra, "A note on two problems in connexion with graphs," *Numerische Math.*, vol. 1, no. 1, pp. 269–271, Dec. 1959.

[29] J. Chen, Y. Gong, M. Fiorani, and S. Aleksic, "Optical interconnects at the top of the rack for energy-efficient data centers," *IEEE Commun. Mag.*, vol. 53, no. 8, pp. 140–148, Aug. 2015.

[30] K.-I. Sato, "Realization and application of large-scale fast optical circuit switch for data center networking," *J. Lightw. Technol.*, vol. 36, no. 7, pp. 1411–1419, Apr. 1, 2018.

[31] H. Nagai, Y. Mori, H. Hasegawa, and K. Sato, "Design and verification of large-scale optical circuit switch using ULCF AWGs for datacenter application," *J. Opt. Commun. Netw.*, vol. 10, no. 7, pp. B82–B89, Jul. 2018.
[32] K. Ishii, T. Inoue, and S. Namiki, "Toward exa-scale optical circuit switch interconnect networks for future datacenter/HPC," *Proc. SPIE*, vol. 10131, Jan. 2017, Art. no. 1013105.
[33] S. Hong, J. Zhu, L. A. Braunstein, T. Zhao, and Q. You, "Cascading failure and recovery of spatially interdependent networks," *J. Stat. Mech., Theory Exp.*, vol. 2017, no. 10, 2017, Art. no. 103208.
[34] S. Hong, B. Wang, X. Ma, J. Wang, and T. Zhao, "Failure cascade in interdependent network with traffic loads," *J. Phys. A, Math. Theor.*, vol. 48, no. 48, Dec. 2015, Art. no. 485101.
[35] S. Hong, H. Yang, T. Zhao, and X. Ma, "Epidemic spreading model of complex dynamical network with the heterogeneity of nodes," *Int. J. Syst. Sci.*, vol. 47, no. 11, pp. 2745–2752, Aug. 2016.
[36] S. Hong, C. Lv, T. Zhao, B. Wang, J. Wang, and J. Zhu, "Cascading failure analysis and restoration strategy in an interdependent network," *J. Phys. A, Math. Theor.*, vol. 49, no. 19, May 2016, Art. no. 195101.
[37] Z. Fei, B. Li, S. Yang, C. Xing, H. Chen, and L. Hanzo, "A survey of multi-objective optimization in wireless sensor networks: Metrics, algorithms, and open problems," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 1, pp. 550–586, 1st Quart., 2017.
[38] Y. Chang, X. Yuan, B. Li, D. Niyato, and N. Al-Dhahir, "Machine-learning-based parallel genetic algorithms for multi-objective optimization in ultra-reliable low-latency WSNs," *IEEE Access*, vol. 7, pp. 4913–4926, 2019.

YUANYUAN HONG was born in Guizhou, China, in 1991. She received the B.S. degree in information engineering and the Ph.D. degree in optical engineering from Zhejiang University, in 2012 and 2018, respectively. She is currently a Lecturer with Taizhou University. Her research interests include optical datacenter networks, routing, and spectrum assignment algorithms.

DAN ZHANG was born in Henan, China, in 1994. She received the B.S. degree from the Zhong Yuan University of Technology, in 2016, and the M.S. degree from South China Normal University, in 2019. Her main research interest includes datacenter optical networks.

BING YANG was born in Shandong, China, in 1994. He received the B.S. degree from the China University of Petroleum (Huadong), in 2014, and the M.S. degree from South China Normal University, in 2019. His main research interest includes datacenter optical networks.

GUANG CHEN was born in Zhejiang, China, in 1987. He received the B.S. degree in automation engineering and the Ph.D. degree in control theory and engineering from the Zhejiang University of Technology, Hangzhou, China, in 2009 and 2012, respectively. He currently serves as a Lecturer at Taizhou University, Taizhou, China. His research interests include traffic modeling and optimization of complex networks, data mining, and deep learning algorithm.

YANG WANG received the B.S. degree in electronic engineering from the Zhejiang University of Technology, China, in 2002, and the M.S. degree in signal and information processing from Shanghai University, China, in 2005. He is currently pursuing the Ph.D. degree with Hohai University, China. He is also a Lecturer with the Department of Electric Engineering, Taizhou University, China. His research interests focus on computer-aided diagnosis and biomedical image processing.

QIMAN CHEN was born in Jiangxi, China, in 1995. He received the B.S. degree from the EAST China University of Technology, Nanchang, China, in 2017. He is currently pursuing the M.S. degree with the South China Academy of Advanced Optoelectonics, South China Normal University. His main research interests include datacenter optical networks and service function chain.

XUEZHI HONG received the B.E. degree from Wuhan University, Wuhan, China, in 2006, and the Ph.D. degree from Zhejiang University, Hangzhou, China, in 2012. He held a postdoctoral position at the KTH Royal Institute of Technology, Stockholm, Sweden, in 2016 and 2017. He is currently working as an Associate professor with South China Normal University, Guangzhou, China, working on visible light communication and positioning, optical fiber communication, and datacenter networks.

JIAJIA CHEN received the bachelor's degree from Zhejiang University, China, in 2004, and the Ph.D. degree from the KTH Royal Institute of Technology, Sweden, in 2009. She is currently a Professor with the Chalmers University of Technology, Sweden. She and her group have made advances in optical network architecture design along with the supporting transmission techniques and resource allocation strategies, which greatly improve capacity, reliability, energy efficiency, and cost efficiency in access, core, and data center networks. She has been sharing her knowledge with the technical community with many peer-reviewed journal articles, conference papers, invited talks (more than 200 publications to-date), and ten patent applications. She has organizing several IEEE international conferences in the field and serving on technical program committees, including OFC (Subcommittee Chair 2019), ECOC, Globecom (Symposium Co-Chair 2020), and ICC.