

Research Article

Layout-Independent Wireless Facility Constructing and Scheduling for Data Center Networks

Xianglin Wei and Qin Sun

Nanjing Telecommunication Technology Research Institute, Nanjing 210007, China

Correspondence should be addressed to Xianglin Wei; wei_xianglin@163.com

Received 27 June 2017; Accepted 27 August 2017; Published 8 October 2017

Academic Editor: Dharma P. Agrawal

Copyright © 2017 Xianglin Wei and Qin Sun. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The control packets in the data center networks (DCNs) have to contest with the data packets although they are usually much shorter in size and much more important in network management. Moreover, the uneven distribution of the packets may create potential traffic hotspots in the DCN which could degrade network performance drastically. To bridge these gaps, a layout-independent constructing algorithm and a scheduling method are put forward towards layout-independent wireless facility in data centers. First of all, a conflict aware spanning tree algorithm is developed to construct the wireless facility network (WFN). Secondly, a scheduling method which contains three steps, route calculation, traffic estimation, and flow scheduling, is presented. In the route calculation step, a route set between each node pair is calculated in advance for later usage. The scheduler estimates the traffic loads on the links on a regular basis in the traffic estimation step. Then, arrived data and control flows are scheduled according to multiple policies based on given route sets and scheduling objectives in the flow scheduling step. Finally, a series of experiments have been conducted on NS3 based on two typical data center layouts. Experimental results in both scenarios have validated our proposal's effectiveness.

1. Introduction

With the rapid development of modern cloud-based network services, more and more data centers are built around the world since the data centers (DCs) are the best infrastructure to support these applications. Cisco's Global Cloud index report forecasts that DC traffic will triple to reach 15.3 zettabytes annually by 2020 [1–3]. In contrast, traditional tree based data center network (DCN) structure faces a lot of challenges to carry out so huge network traffic, such as structure design, network management, and load balancing. Therefore, many novel or revolutionized network structures, such as FatTree [4] and Jellyfish [5], have been presented to replace or enhance the traditional tree based DCNs. However, it is hard to implement these architectures in practice since they have introduced too much overheads like cabling or multiports servers.

To reduce the cabling overhead, there are also a few wireless proposals which augment existing DCNs with multi-gigabit wireless links to provide extra bandwidth for the DCs. These novel DCNs which consist of both wired links as well as

wireless ones are named as hybrid DCNs (HDCNs). Recent researches have revealed that HDCNs could remarkably reduce the number of cables and switches and thus can reduce not only equipment costs but also server installation and reconfiguration costs. In comparison with existing wired structures, HDCNs are more energy efficient. Moreover, wireless links can be deployed in an incremental fashion and can be used to update outdated and tightly coupled DCs.

However, existing HDCNs are proposed for the general purpose without paying special attention to control packets transmission although these packets are usually more important and time-sensitive. Moreover, some of the proposals assume a regular DC layout and their performance relies on this assumption. Actually, almost every data center has its own unique layout based on its scale and deployed applications. How to efficiently construct and schedule a layout-independent wireless facilities in the DCNs has drawn little attention. This is also the focus of this paper.

Here, we want to construct a DC layout-independent wireless facility network (WFN) which can efficiently adapt to different DC layouts. The WFN can carry control flows

and be utilized to alleviate traffic hotspots as well as handle link failure. To be specific, three main design requirements for our WFN are as follows: first of all, the construction process of the WFN should be independent of underlying DC layouts. Secondly, the wireless facilities should carry the control flows in priority since these flows are usually smaller but more important than the data flows in network management. Finally, the scheduling method should not rely on complicated data collection since it may introduce too much measurement cost and cause long scheduling delay.

To meet these challenges, a WFN construction method named conflict aware spanning tree algorithm is presented firstly which can connect the racks through wireless links based on abstracting DC layouts and taking the mutual-conflict among wireless links into consideration. Then, a WFN scheduling algorithm is put forward, which includes three steps, that is, route calculation, traffic estimation, and flow scheduling. In the route calculation step, a route set is computed for each node pair based on the topology of the WFN. Then, the scheduler estimates the traffic loads on the links on a regular basis in the traffic estimation step. Finally, the arrival flows are scheduled based on scheduling policies in the last step. A series of simulation experiments are conducted to evaluate the effectiveness of the proposed methods. Compared with existing solutions, our main contributions are threefold:

- (i) The proposed layout-independent solution can adapt to various DC layouts. As for some particular DC layout, our construction method can transfer its abstract layout to a graph and place the wireless links while considering their mutual interference.
- (ii) The performance of the control flows can be greatly promoted through efficiently utilizing the deployed wireless links. Actually, separating control and data flows can also benefit data flows since they have more available wired bandwidth.
- (iii) The constructed WFN can maintain the connectivity of the DCN when wired link failure happens. This is very important to ensure the availability of the services deployed in the DCs.

The WFN can also provide the managers or maintainers of the DCs with options to upgrade or update their DCs to incorporate newly emerged computing paradigms. Moreover, timely control flow delivery can benefit data center managers a lot since they can rapidly find and handle potential anomalies.

This rest of the paper is organized as follows. Section 2 summarizes related work. Section 3 gives background about 60 GHz wireless links and formulates the problem. Section 4 presents the design of our construction and scheduling methods. A series of simulation experiments are conducted in Section 5. Finally, we conclude our main work in Section 6.

2. Related Work

To enhance traditional DCNs, a few proposals have been proposed based on introducing wireless links or optical links to the data centers.

2.1. Wireless Solutions. Current hybrid electrical/wireless switch architectures using 60 GHz wireless technologies can be mainly divided into two categories:

- (i) Using wireless switching network as a control plane for wired DCNs to promote the availability of the control packets transmission.
- (ii) Augmenting wired DCNs with wireless links to alleviate traffic hotspots.

In the former category, a wireless facilities network named ANGORA for data centers using low-cost, 60 GHz beamforming radios is proposed in [6]. ANGORA provides robust paths decoupled from the wired network and flexibility to adapt to workloads and network dynamics. In ANGORA, a wired data plane is in charge of switching data packets while the wireless facilitates network based on wireless links delivers control traffic only. ANGORA adopts pretuned antenna directions (i.e., fixed topology) to remove the need of link coordination. Moreover, they implemented Kautz graphs using wireless links while taking wireless link interference and network failures into consideration.

In the latter category, there are a few proposals. In [7], Halperin et al. have explored using 60 GHz wireless technology to relieve hotspots in oversubscribed DCNs. By experimenting with prototype equipment, they have shown that the DC environment is well suited to a deployment of 60 GHz links contrary to concerns about interference and link reliability. Using directional antennas, many wireless links can run concurrently at multi-Gbps rates on top of rack (ToR) switches. The wired DC network can be used to sidestep several common wireless problems. By analyzing production traces of DC traffic for four real applications, they have shown that adding a small amount of network capacity in the form of wireless Flyways to the wired DC network can improve performance.

Cui et al. have presented a novel DCN architecture, Diamond, which nests the wired DCN with radios equipped on all servers [8]. To harvest the gain allowed by the rich reconfigurable wireless resources, they proposed the low-cost deployment of scalable 3D Ring Reflection Spaces (RRSs) which are interconnected with streamlined wired herringbone to enable large number of concurrent wireless transmissions through high-performance multireflection of radio signals over metal.

By taking advantage of clean LOS channels on top of server racks, Katayama et al. have implemented a robust wireless packet-switching network for wired DCN using mmWave LOS channels [9, 10]. They have presented the architecture considerations for the wireless crossbar switching. Moreover, packet routing and relaying mechanism at each wireless node is presented.

Shin et al. have discussed the feasibility of completely wireless data centers [11]. They have introduced an architecture of the Cayley data center and have presented how this architecture affects the positioning of the 60 GHz transceivers in a wireless data center which in turn defines the network topology. Moreover, they introduced a novel geographical routing protocol for this unique topology and adopt a MAC layer protocol to address the hidden terminal problem. Aktas et al. have put forward WiCOD, which relies on a wireless control plane serving an all-optical data plane [12]. Li and Santini have investigated energy-aware scheduling of coflows and antenna directions in a server-centric DCN which includes both wire and wireless links [13]. Roh et al. have presented a novel approach to flow and virtual machine placement problems in hybrid data center networks [14]. In [14], a threshold-based, wireless link-aware flow placement algorithm with low complexity is put forward. Luo et al. have presented VLCcube, an interrack wireless solution that extends the design of wireless DCN through organizing all racks into a wireless Torus structure via the emerging visible light links [15].

A detailed analysis of applying wireless links to DCs can be referred to [16].

2.2. Optical Solutions. Besides, there are a few proposals which introduce optical links and switches in to the DCNs to construct an optical/electricity switch structure. In [17], Wang et al. have presented a hybrid packet and circuit-switched DCN architecture which augments the traditional hierarchy of packet switches with a rack-to-rack optical circuit-switched network. Based on this architecture, they have built a prototype system called *c-Through* [17]. Farrington et al. have put forward a hybrid electrical/optical switch architecture for modular data centers named Helios, which can deliver significant reductions in the number of switching elements, cabling, cost, and power consumption relative to recently proposed data center network architectures [18]. They have also identified hardware requirements for improving the performance of hybrid electrical-packet-switched/optical-circuit-switched data center networks [19]. Then, a hybrid optical circuit-switched/electrical packet-switched network for datacenters called Mordia is presented [20].

In reference [21], Bazzaz et al. have investigated several challenges arising in incorporating switched optical circuits into the modern data center. Then, an Observe-Analyze-Act framework is presented to meet the requirements for data center optical circuit controllers. In [22], Singla et al. have put forward Proteus, an all-optical architecture targeting unprecedented topology flexibility, lower complexity, and higher energy efficiency. Bowers et al. have provided analysis of the economic benefits of hybrid packet Optical-Circuit-Switched (OCS) networks for DCN and described a practical management and control plane architecture [23]. In [24], Hamedazimi et al. have presented FireFly, an interrack network solution using free-space optics (FSO). Hamza et al. have presented OWCcells, a class of optical wireless cellular data center network architectures in which fixed line of sight (LOS) optical wireless communication (OWC) links are used

to connect racks of servers arranged in regular polygonal topologies [25].

From the above analysis, we can see that existing proposals mainly concentrate on the structure of the HDCN. Little attention has been paid on how to efficiently schedule the wireless facilities in the DCNs.

3. Problem Formulation

In this section, we firstly introduce 60 GHz wireless technology and DC layouts. Then, the WFN construction and scheduling problem is formulated.

3.1. 60 GHz Wireless Technology. License-free 60 GHz radios can achieve multi-gigabit Radio Frequency (RF) links using the allocated sufficient spectrum. Moreover, the very narrow beam associated with 60 GHz radios enables multiple 60 GHz radios to be easily and accurately installed by a nonexpert installer on the same rooftop or mast, even if they are all operating at the same transmit and receive frequencies. However, 60 GHz signals can be absorbed by the Oxygen. This limits the distances that 60 GHz links can cover, but it also offers interference and security advantages when compared to other wireless technologies. In a word, these unique characteristics make 60 GHz wireless technology suitable for short-to-medium distance, high-bandwidth applications.

In the DCNs, 60 GHz radios are usually installed on the top of the racks and are connected to the ToR switches. The number of radios that can be setup on the top of a rack depends on the size of the rack. For example, today's standard rack is 4 ft \times 2 ft and a 60 GHz radio is 1 ft \times 1 ft [26], so at most 8 radios can sit atop each rack. Because 60 GHz links are highly directional, each rack can only communicate with a small, constant number of peers in parallel. This limited "node degree" makes it hard for a rack to reach any other rack with bounded delay [6]. For instance, ANGORA placed 8 radios on each rack while Halperin et al. placed 1 radio on each rack [6, 7].

To efficiently utilize the available space while reducing the interference among the 60 GHz wireless links, three typical 60 GHz wireless links (as shown in Figure 1) can be built in the DC:

- (i) Direct wireless link: two 60 GHz radios deployed at neighbor racks can directly establish a LOS 60 GHz wireless link.
- (ii) Raised wireless link: a raised radio can establish a wireless link with another raised radio with the same height.
- (iii) 3D beamforming link: two radios can build a wireless link through the reflector (or a mirror) deployed on the ceiling.

Here, all these three link types can be adopted. As long as there is no barrier between two racks and their distance is less than some particular distance, there may be a wireless link between them. Then, a few wireless links can be built between the racks. The number of the wireless links can be determined by the given cost constraint.

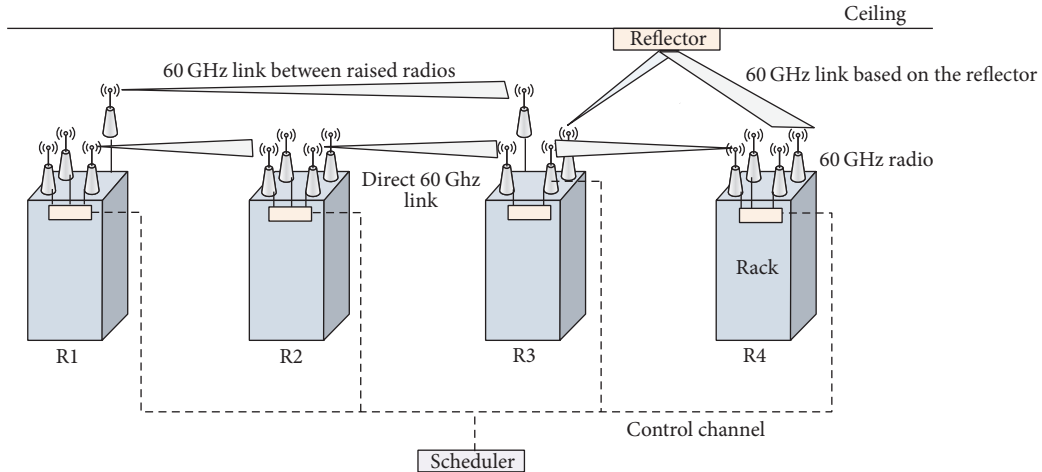


FIGURE 1: 60 GHz interrack wireless link.

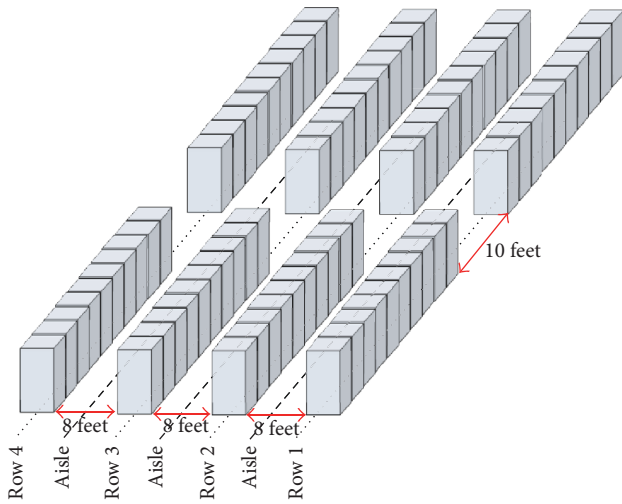


FIGURE 2: Partial top view of data center of a large search provider. Each row has ten 24×48 inch racks. The aisles are 10 and 8 feet wide, as shown.

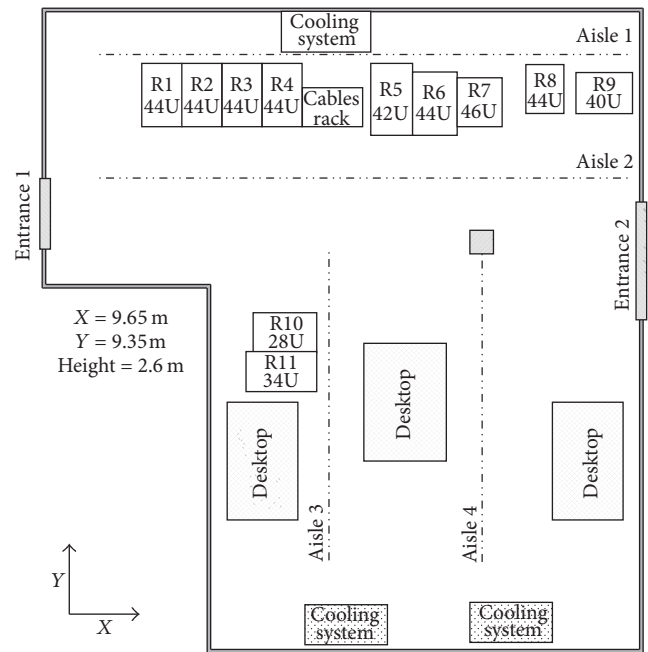


FIGURE 3: The layout of a campus data center.

3.2. *DC Layout.* A DC’s layout can significantly influence its cabling, cooling, and management. Moreover, it will severely impact the deployment of the wireless radios and the availability of the wireless links. Figure 2 shows a medium-size search provider’s regular DC layout [7]. Moreover, Figure 3 illustrates the layout of a campus DC [27]. In this campus DC, the racks have different sizes. Moreover, a pillar exists in the DC which may block the 60 GHz links between rack 9 and rack 11. For the WFN and its scheduling algorithm, we need to adapt to different layouts.

3.3. *Problem Formulation.* Given some particular DC layout and a number of 60 GHz wireless radios, our objective is constructing and scheduling a WFN for the given DC. Our main design considerations include the following:

(i) For the WFN, it can set up different types of wireless links according to the layout of the given DC. This

means that the possible wireless links can be established according to these three link setup methods introduced in Section 3.1. This will increase the number of possible links.

- (ii) For the WFN, it should be adapted to different layouts. This means that it has to derive a list of possible wireless links and their interference relationship based on a given DC’s layout before constructing the WFN.
- (iii) To enable timely network management, the control traffic in the DC should be transmitted with higher priority than data flows.
- (iv) To ensure the service continuity, we should utilize the idle bandwidth of the WFN to carry these isolated

racks' traffic. Moreover, the idle wireless bandwidth can also be utilized to alleviate traffic hotspots.

To be specific, we should firstly construct a WFN based on a given DC's layout. Then, based on its topology, a set of feasible routes for each pair of nodes should be calculated. Note that these sets can be obtained in advance since the topology is relatively static in DCN. After obtaining these sets, we can schedule the given set of flows $\{f_1, f_2, \dots, f_n\}$.

4. WFN Construction and Scheduling

In this section, we firstly present the WFN construction algorithm and then discuss its scheduling method.

4.1. WFN Construction. The WFN construction process includes three steps:

- (1) DC layout abstraction: obtain possible wireless links based on a given DC's layout and the transmission characteristics of the adopted 60 GHz wireless devices.
- (2) Conflict links determination: determine the interference relation among possible wireless links.
- (3) Wireless links arrangement: arrange the wireless links to construct a facility network while reducing their mutual-conflict.

4.1.1. Feasible Wireless Links Determination. The racks in the DC can be modeled as r_1, r_2, \dots, r_N , where N is the number of racks. Their locations are $(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)$, respectively. Note that each rack's location in the DC is usually static and can be known in advance. A few radios can be installed on top of each rack and their positions on the rack are static with adjustable height, which is selected randomly from a height set initially. Then, if two radios are within each other's transmission range, there will be a feasible wireless link between them. The Euclid distance D_{ij} of two radios installed on top of the i th and the j th rack is calculated based on i 's and j 's locations.

For two radios installed on top of rack i and j , respectively, there is a possible link between them if and only if

- (i) rack constraint: these two radios are not installed on the same rack; that is, $i \neq j$;
- (ii) distance constraint: $D_{ij} \leq d$;
- (iii) obstruction free: there are no obstacles between these two racks that can block wireless radio transmission.

Assume there are n_i radios on the i th rack, we have in total $n = \sum_{i=1}^N n_i$ radios and C_n^2 possible unidirectional links. Actually, the total number of feasible wireless links is limited due to the above three constraints.

4.1.2. Conflict Links Determination. Besides the constraints discussed in above section, the mutual-conflict among wireless links should also be considered. For each link in the feasible wireless links, we have to find its corresponding

conflict link set firstly. In other words, a link cannot coexist or cotransmit with the links in its conflict link set. Here, three types of conflict links exist for each link.

- (i) Shared wireless radio: two links will conflict with each other if they share the same wireless radio.
- (ii) Link blockage: link's radios may to some extent block another link since 60 GHz transmissions can be easily blocked by small obstacles.
- (iii) Radio interference: a link's beam may interfere with racks and the radios (links) installed on the racks in its direction.

To eliminate link blockage and reduce radio interference, Zhou et al. have presented 3D beamforming, which places reflectors on the ceiling and electromagnetic absorbers around the radios [26]. This method can greatly reduce the number of conflict links for each link. Besides, we can also raise the height of the radios to avoid the interference among links. But we still have to cope with the scenarios where for some racks it is hard to place the reflectors on the ceiling or raise the radios.

For simplicity of the problem abstraction, we assume that a radio r 's coordinate is (x_r, y_r, z_r) , where x_r and y_r are r 's location on the rack while z_r is its height which can be selected from a few fixed values. Two radios with different height values will not interfere with each other. At the same height, two links may conflict with each other when a link's radio is located between another link's two radios or a radio's arrived beam is within some particular angle of another radio. The conflict angle can be measured or estimated [26, 27] and can be calculated by the locations of the radios. In summary, a link l 's conflict link set contains the following links:

- (i) Links which share at least one radio with l .
- (ii) Links built on those radios which locate between l 's two radios, that is, blocking links.
- (iii) Links built on radios within the conflict angle of l 's two radios, that is, interference links.

According to this principal, we can determine each link's conflict link set. Then, we can eliminate some conflict wireless links through adjusting its radios' positions. For instance, changing its radio height can move it from one "congested" conflict plane to another one and leaving its position on the rack intact. After adjustment, those nonconflict links will be removed from a link's conflict link set.

4.1.3. Wireless Links Arrangement. After determining each link's conflict link set, we then need to arrange the wireless links based on specific objectives. Our basic motivation is that the WFN we want to construct can connect each rack with low-cost. Therefore, the WFN is at least a tree which connects all the racks by wireless links. However, we cannot directly apply the minimum spanning tree (MST) algorithm developed in [28] since there exists conflict among links.

To achieve this goal, we devise a conflict aware spanning tree algorithm to arrange the wireless links in the DC as shown in Algorithm 1. The wireless network can be modeled

```

Input: feasible wireless links set F, racks set R
Output: A: The spanning tree; B: The racks in A
(1) A =  $\emptyset$ , B =  $\emptyset$   $\triangle$  B is the racks covered by A;
(2) Tabu list T =  $\emptyset$ ;
(3) while A does not form a spanning tree do
(4)   if no links available for selection then
(5)     break;
(6)   end
(7)   Randomly select a link  $l(u, v)$  from F, where
       $r(u) \in \mathbf{B}$  and  $r(v) \in \mathbf{R} - \mathbf{B}$   $\triangle$   $r(u)$  and  $r(v)$ 
      are the racks which setup wireless radio  $u$  and  $v$ 
      respectively, i.e.  $l(u, v)$  is a wireless link crossing B
      and  $\mathbf{R} - \mathbf{B}$ ;
(8)   F = F -  $\{l(u, v)\}$   $\triangle$  Remove  $l(u, v)$  from F;
(9)   Determine  $l(u, v)$ 's conflict link set L based on the
      principals listed in Section 4.1.2;
(10)  if  $(\mathbf{R} - \mathbf{B}) \neq \emptyset$  and there is no crossing link between
       $\mathbf{B} \cup \{u, v\}$  and at least one of the connected
      components of the graph
       $G((\mathbf{R} - \mathbf{B}) \cup \{u, v\}, \mathbf{F} - \mathbf{L})$  then
(11)    T = T  $\cup \{l(u, v)\}$ ;
(12)    continue;
(13)  else
(14)    A = A  $\cup \{l(u, v)\}$ ;
(15)    B = B  $\cup \{r(v)\}$ ;
(16)    F = F - L;
(17)  end
(18) end
(19) if  $\mathbf{R} - \mathbf{B} \neq \emptyset$   $\triangle$  There exist isolated areas in the
      network then
(20)   Try to establish new wireless radios or adjusting their
      positions on the isolated racks to increase the amount
      of feasible links;
(21)   Deploy multiple reflectors for these racks' radios to
      increase the amount of feasible links;
(22)   return to Step (3);
(23) end
(24) return A and B

```

ALGORITHM 1: Conflict aware spanning tree algorithm.

as an undirected graph $\mathbf{G} = (\mathbf{V}, \mathbf{E})$, where \mathbf{V} is the set of racks and \mathbf{E} is the set of feasible links between pairs of racks. Assume that set **A** keeps a subset of the MST for each iteration.

Steps (1)-(2) initialize the spanning tree and the tabu list. Whenever there are available wireless links and there are isolated racks, Algorithm 1 chooses a wireless link from the link set in steps (7)–(16). Note that in order to ensure the connectivity of the WFN, we may need to adjust some radios' positions or install new reflectors to increase the number of feasible wireless links in steps (20) and (21).

Note that our basic spanning tree tries to connect all the racks. Therefore, at least $|\mathbf{R}| - 1$ wireless links will be added to the spanning tree, which will incur $O(|\mathbf{R}|)$ time complexity. To facilitate algorithm design, we need to store all possible C_n^2 wireless links at first. Therefore, Algorithm 1's space complexity is about $O(n^2)$.

Besides, we can also add more wireless links to the MST and provide multiple paths between each pair of radios. This will greatly enhance the service capacity of the WFN while introducing more cost.

4.2. WFN Scheduling. As mentioned before, there are a few requirements for WFN scheduling, like hotspot alleviating and failure tolerance. Here, WFN's topology is relatively stable after construction and multiple paths may exist between each two wireless nodes. Therefore, we can calculate the routes between each node pair in advance. Then, the WFN scheduler can estimate the load on each link based on ongoing flows on the routes. Finally, the newly arrived flows will be assigned to the routes based on the traffic estimation results. To be specific, the scheduling process is divided into three steps, that is, routes calculation, traffic estimation, and flow scheduling.

```

Input: Topology graph  $G(V, E)$ ,  $i, j$ , route length limit  $k$ 
Output: a set of routes for  $(i, j)$ 
(1) Tabu list  $T = \emptyset$ , Route set  $S = \emptyset$ , Node set  $N = \{i\}$ ;
(2) for each node  $b$  in set  $N$  do
(3)   if the length of the path from  $i$  to  $b$  is larger than  $k$ 
      then
(4)     Break;
(5)   end
(6)   for each neighbor  $a$  of node  $b$  do
(7)     if  $a$  is equal to  $j$  then
(8)       Add the path from  $i$  to  $a$  in the tree to  $S$ ;
(9)       Add the nodes in the path to  $T$  except  $i$ ;
(10)    else
(11)      if  $a$  is already included in  $T$  then
(12)        Cut  $a$  and the subtree rooted at  $a$  from
          the tree;
(13)      end
(14)    end
(15)    Add  $a$  to  $N$  if  $a \notin N$ ;
(16)  end
(17)  Remove  $b$  from  $N$ ;
(18) end
(19) Return  $S$ 

```

ALGORITHM 2: Route calculation.

4.2.1. Routes Calculation. There are a few algorithms which can calculate the shortest routes among nodes, like Dijkstra or Ford algorithms [28]. However, as there are a few paths among each node pair in WFN, we need to combine the breadth-first searching algorithm and tabu search [29] to determine a few noninteraction routes for each node pair.

We can treat the WFN as a graph $G(V, E)$, where V and E are the sets of links and edges, respectively. To decide the routes between any two nodes i and j , then, we start from node i and apply the breadth-first search until node j is found in the tree or the depth of the searching tree reaches k , which is the maximum length of a route. A tabu list is kept during the searching process to cut those tree branches encountered before or overlapping with existing routes. The algorithm of this searching process for a pair of nodes i and j is shown in Algorithm 2.

Algorithm 2's time complexity is roughly $O(N^3 \times \beta)$, where β is the average number of neighbors for each wireless node. β is a small constant and it is strictly less than the number of wireless radios placed on top of each rack. The space complexity of Algorithm 2 is $O(N^2)$ since we have to store the routes between each node pair.

4.2.2. Traffic Estimation. After determining routes for each node pair, each flow will be assigned to one or more than one routes. Then, the load on each link on the route can be calculated based on the assignment results. For instance, if a flow f_i is assigned to a route r_i , each link on the route will increase by the assigned bandwidth of f_i . Based on the start and end times of each flow, the scheduler can estimate all the links' traffic load.

4.2.3. Flow Scheduling. According to various scheduling objectives, we can implement different scheduling policies. To be specific, the following four methods can be adopted here:

- (i) Wired-only: only the wired networks are utilized for carrying the data and control flows. In other words, the WFN is not adopted.
- (ii) Wireless-for-control: the wireless links are only utilized for carrying control flows while the data flows are carried on the wired network.
- (iii) Partial-for-data-with-priority: the wireless links are utilized for carrying control flows and it can carry a few offloaded data flows if there is idle bandwidth available. Here, if a data flow is partially offloaded to the WFN, we move 10% of its total packets to the WFN and the remaining 90% packets will still be transmitted through the wired network. Moreover, the control flow's priority is higher than the offloaded data flows. The offloaded data flows will be transferred back to the wired network if the control flows need to use its occupied wireless capacity.
- (iv) Failure-tolerant: the wireless links are used to carry control flows and they carry the flows for those ToRs whose links to the aggregation switches are failure.

5. Simulation and Results

To evaluate the performance of the WFN construction and its scheduling methods proposed in this paper, a series of simulation experiments are conducted based on extending

the Flyways simulator developed on NS3 [7, 30]. To fulfill our simulation requirements, we have extended the simulator by WFN construction, routes calculation, and flow scheduling.

5.1. Experimental Settings

Parameter Settings. We adopted two typical DC layouts shown in Figures 2 and 3 to conduct the experiments. The FatTree topology is adopted in the wired network [31] for both layouts. The bandwidths of the aggregation wired and wireless links are set to be 10 Gbps and 3 Gbps, respectively. 10 groups of flows are generated in the experiments and each flow group is composed of 3 data flows and 2 control flows. Moreover, these 5 flows have the same destination host and their source hosts are randomly selected from the hosts in the DC. Each data flow's size is set to be a number randomly selected between 80 and 100 Mbps. In contrast, the size of the control flow is relatively small and is randomly selected between 8 and 10 Mbps. In total, we have 30 data flows and 20 control flows. The time intervals between the flows are set to be $t + \Delta t$, where t is 0.1s and Δt is a random number selected between 0.02s and 0.12s. Note that we do not specify the applications that generate these flows and just aim to simulate the traffic characteristics in the DCs. The adopted data flows can incorporate typical MapReduce-like DC flows while the control flows here may be the control flows generated by network management, the control flows in Software Defined Networking (SDN), and so on. To evaluate the fault tolerant scenario, two wire links are randomly removed from the network. For the simplicity of the description, the layouts shown in Figures 2 and 3 will be referred to be Layout-A and Layout-B, respectively, in the following analysis.

Evaluation Metrics. To evaluate different scheduling policies' performance, we evaluate the following two metrics of each flow:

- (i) Completion time of the flow (CTF): CTF is defined as $t_2 - t_1$, where t_2 is the time when all the packets in the flow are received at the flow's destination and t_1 is time when the first packet of the flow is generated.
- (ii) Average one-way delay: it is defined as the average value of a flow's all packets' one-way delays. A packet's one-way delay is defined as its receiving time at the destination minus its sending time at the source.

5.2. Evaluation Results for Layout-A

5.2.1. DCN Topology. Note that different number of wireless radios can be installed on each rack. Here, we assume that no more than 8 radios can be installed on each rack and the wireless links are randomly established among the racks. Based on this setting, the wireless links among the racks shown in Figure 2 are shown in Figure 4. In Figure 4, each red rectangle represents a rack, the blue dots on the racks are the wireless radios, and the dotted line between two radios represents a wireless link. These wireless links can provide at least one path between every two racks.

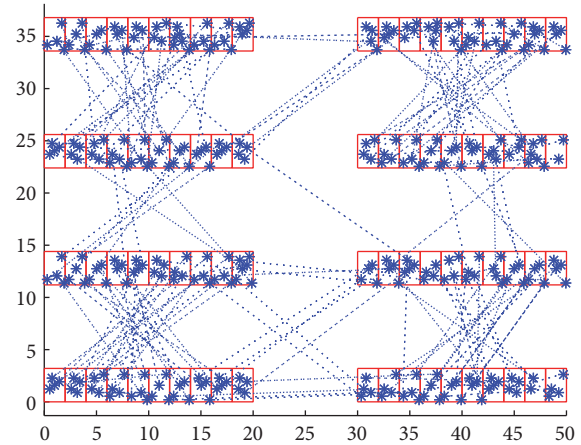


FIGURE 4: The deployed wireless links for the DC layout in Figure 2.

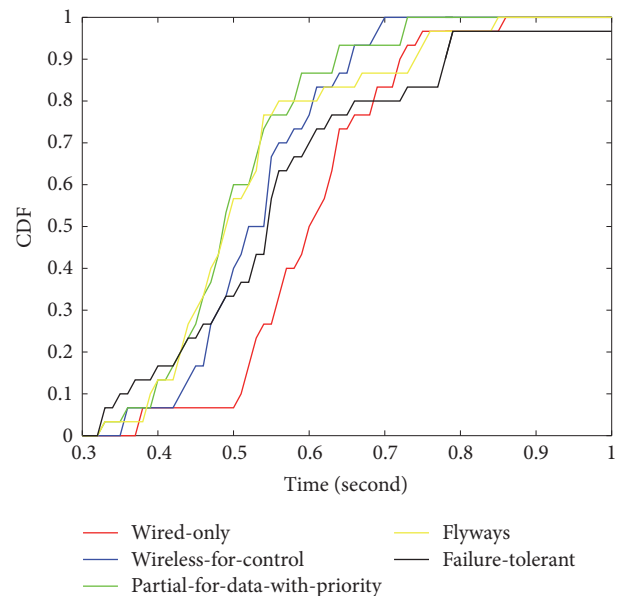


FIGURE 5: The CTF of all the data flows for Layout-A.

5.2.2. WFN Scheduling

CTF. To have a comprehensive overview of the network flows' CTF, the Cumulative Distribution Function (CDF) of the CTF of the data flows and control flows are shown in Figures 5 and 6, respectively. Moreover, Figure 7 shows the average CTF of the data and control flows.

From Figures 5 and 7, we can see that, for the data flows, Partial-for-data-with-priority performs better than other four policies since it can efficiently utilize the WFN through dedicated load distribution between wire and wireless networks. To be specific, it uses wireless links while ensuring that offloaded traffic will not exceed the capacity of the WFN. The performance of wired-only policy is the worst among the five methods since it cannot use wireless links. Flyways achieves the second best since it can support indirect traffic offload which can greatly shorten data flows' CTF. With

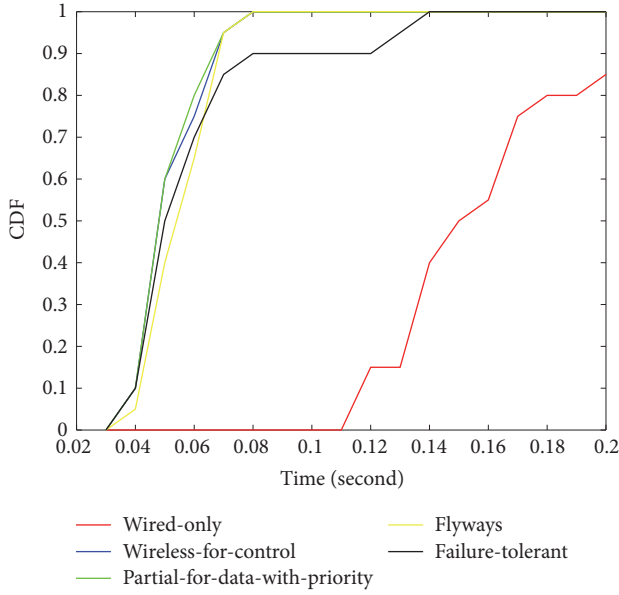


FIGURE 6: The CTF of all the control flows for Layout-A.

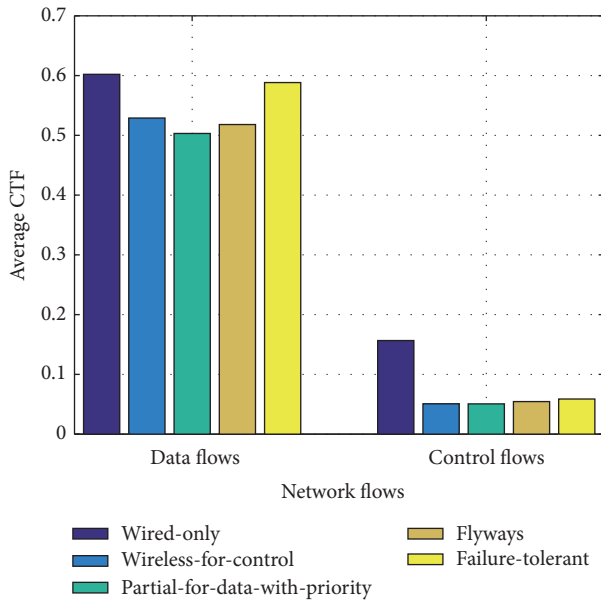


FIGURE 7: Average CTF of the data and control flows for Layout-A.

failure-tolerant policy, the data flows experience long CTF due to congestion due to link failure.

Combining Figures 5 and 7, we can draw the following observation conclusions: (1) wired-only scenario has the worst performance among the five considered scenarios. This is due to the fact that the latter four scenarios can provide more bandwidths for the control flows. (2) Wireless-for-control and Partial-for-data-with-priority achieve the best performance for control flows since they can ensure that control flows have the highest priority in wireless link scheduling. Compared with Flyways, they can reduce the average CTF by

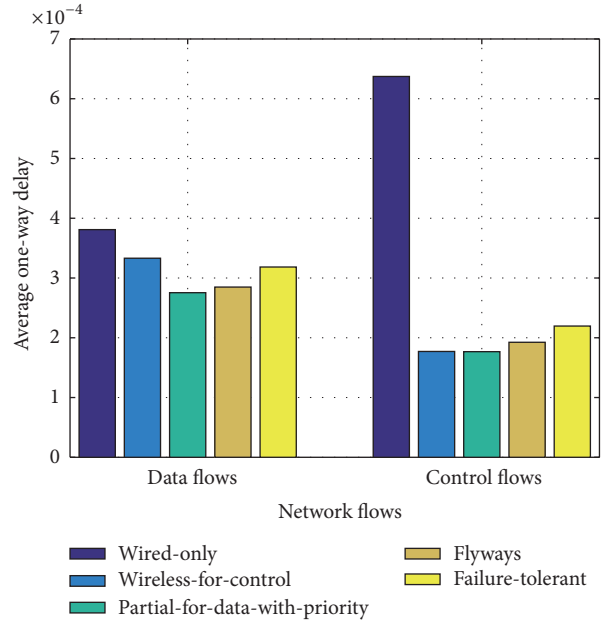


FIGURE 8: Average one-way delay of the data and control flows for Layout-A.

about 7%. (3) In the Failure-tolerant scenario, the WFN can be utilized to carry data flows for the isolated racks although its performance is poor compared with normal scenarios.

Average One-Way Delay. The average one-way delays of the data and control flows are shown in Figure 8. From Figure 8, we can see that separating the control flows can dramatically reduce the average one-way delay of the control flows. For the control flows, this is due to the fact that the wireless links can provide them with high bandwidths. The offload of the control flows may reduce the burden on the wire links and thus promote the performance of the data flows. This can greatly benefit efficiently network management since the control packets can be delivered to their destination much faster than before. Compared with Flyways, Wireless-for-control and Partial-for-data-with-priority could reduce the average one-way delay by roughly 8% since Flyways does not pay special attention to control flows' performance.

5.3. Evaluation Results for Layout-B. The constructed WFN for the layout illustrated in Figure 3 is shown in Figure 9. In Figure 9, each red rectangle represents a rack, the blue dots on the racks are the wireless radios, and the dotted line between two radios represents a wireless link. We can see that our WFN construction method can also work in this irregular DC layout.

CTF. The CTF of all the network flows is shown in Figure 10. From Figure 10, we can draw the following observation conclusions: firstly, adding extra wireless links can greatly promote the performance of the network flows, especially those control flows since they are given high priority for transmission. Secondly, the average CTF of the data flows

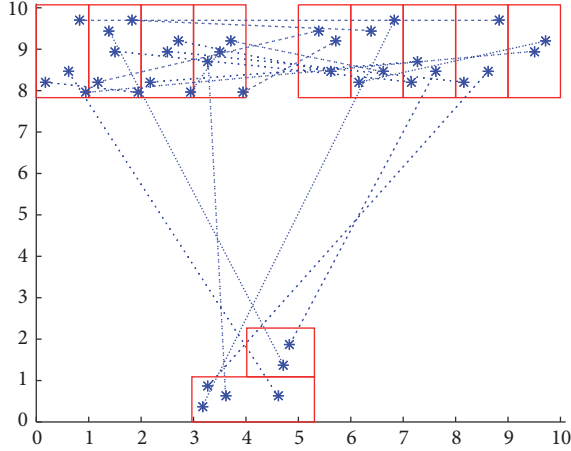


FIGURE 9: The deployed wireless links for the DC layout in Figure 3.

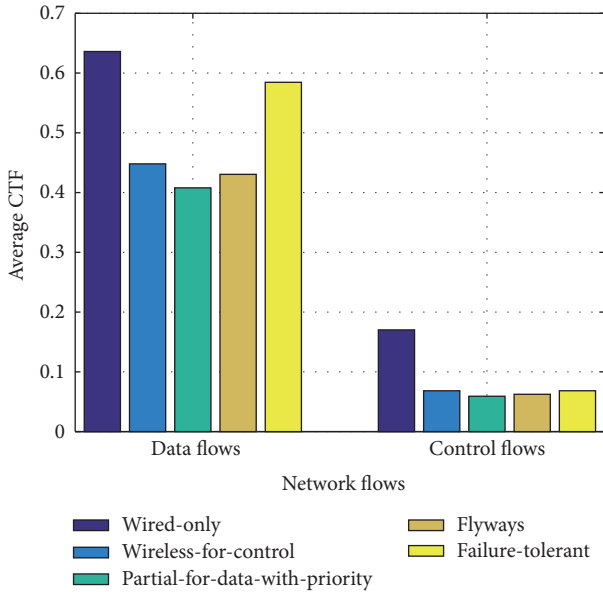


FIGURE 10: Average CTF of all the network flows for Layout-B.

is much larger than that of control flows since data flows are usually much larger than data flows. Thirdly, link failure can influence the whole network's performance other than those servers that are directly connected to failed links since link failure can potentially influence the route selection of the whole network and cause performance churn to the data center. Failure-tolerant scenario has the second worst performance since it has to transmit the isolated racks' traffic through wireless links.

One-Way Delay. All the network flows' average one-way delays are shown in Figure 11. From Figure 11, we can draw similar conclusion with Figure 8, that is, introducing extra wireless links can greatly reduce the one-way delay of the control flows and thus promote the performance of the whole

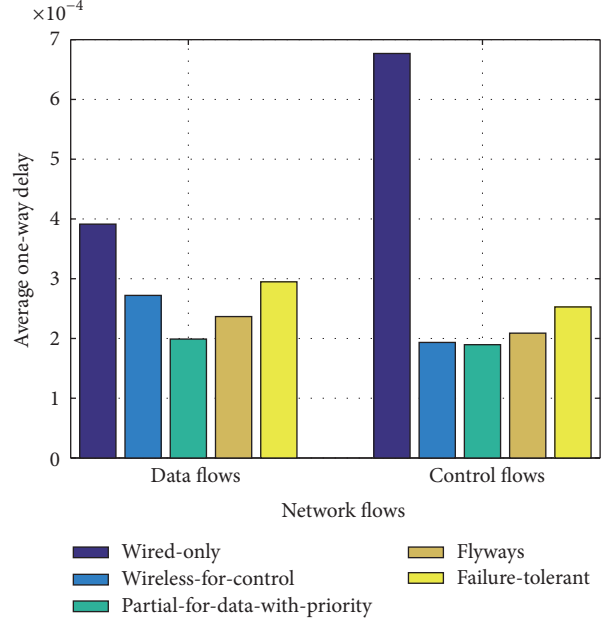


FIGURE 11: Average one-way delay of all the network flows for Layout-B.

network. Among the available five policies, Partial-for-data-with-priority always performs better than other policies.

6. Conclusion

Recent evidences have shown that it is feasible to introduce wireless links into the data center networks (DCNs) to relieve the hotspot wired links and to promote the performance of the applications. Therefore, a few recent proposals have been put forward to carry the control flows or offload data flows in the DC through utilizing wireless links.

In this paper, we have presented a wireless links or facilities construction and scheduling method to efficiently utilize the wireless resources in the DC. A conflict aware spanning tree algorithm is developed to construct the wireless facility network (WFN) which takes the mutual-conflict of the wireless links into consideration. The scheduling method mainly contains three steps, that is, routes calculation, traffic estimation, and flow scheduling. Routes calculation is in charge of computing the routes between each node pair in advance since the topology of the DCN is stable. Then, the traffic on each link is estimated by a scheduler based on the historical assignment results. In the last step, the newly arrived flows are assigned to the routes obtained in the first step based on four different kinds of scheduling policies.

We have implemented our scheduling method on NS3 through extending an existing simulator. Experimental results have validated the effectiveness of the proposed construction and scheduling algorithms.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was supported in part by the National Natural Science Foundation of China under Grant no. 61402521 and no. 61201216 and Jiangsu Province Natural Science Foundation of China under Grant nos. BK20140068, BK20140070, and BK20150201.

References

- [1] Cisco, "Cisco global cloud index: Forecast and methodology 2015-2020 (white paper)," 2016.
- [2] P. Li, J. Li, Z. Huang et al., "Multi-key privacy-preserving deep learning in cloud computing," *Future Generation Computer Systems*, vol. 74, pp. 76–85, 2017.
- [3] W. Xia, P. Zhao, Y. Wen, and H. Xie, "A Survey on Data Center Networking (DCN): Infrastructure and Operations," *IEEE Communications Surveys and Tutorials*, vol. 19, no. 1, pp. 640–656, 2017.
- [4] M. Al-Fares, A. Loukissas, and A. Vahdat, "A scalable, commodity data center network architecture," in *Proceedings of the ACM SIGCOMM Conference on Data Communication (SIGCOMM '08)*, pp. 63–74, Seattle, Wash, USA, August 2008.
- [5] A. Singla, C. Y. Hong, L. Popa, and P. B. Godfrey, "Jellyfish: Networking data centers randomly," in *Proceedings of the 9th USENIX Symposium on Networked Systems Design and Implementation*, 2012.
- [6] Y. Zhu, X. Zhou, Z. Zhang et al., "Cutting the cord: A robust wireless facilities network for data centers," in *Proceedings of the 20th ACM Annual International Conference on Mobile Computing and Networking, MobiCom 2014*, pp. 581–592, usa, September 2014.
- [7] D. Halperin, S. Kandula, J. Padhye, P. Bahl, and D. Wetherall, "Augmenting data center networks with multi-gigabit wireless links," in *Proceedings of the ACM SIGCOMM 2011 Conference, SIGCOMM'11*, pp. 38–49, Toronto, Canada, August 2011.
- [8] Y. Cui, S. Xiao, X. Wang et al., "Diamond: nesting the data center network with wireless rings in 3d space," in *Proceedings of the in Usenix Conference on Networked Systems Design and Implementation*, pp. 657–669, 2016.
- [9] Y. Katayama, K. Takano, Y. Kohda, N. Ohba, and D. Nakano, "Wireless data center networking with steered-beam mmWave links," in *Proceedings of the 2011 IEEE Wireless Communications and Networking Conference, WCNC 2011*, pp. 2179–2184, mex, March 2011.
- [10] Y. Katayama, T. Yamane, Y. Kohda, K. Takano, D. Nakano, and N. Ohba, "MIMO link design strategy for wireless data center applications," in *Proceedings of the 2012 IEEE Wireless Communications and Networking Conference, WCNC 2012*, pp. 3302–3306, fra, April 2012.
- [11] J.-Y. Shin, E. G. Sirer, H. Weatherspoon, and D. Kirovski, "On the feasibility of completely wireless datacenters," in *Proceedings of the 8th ACM/IEEE Symposium on Architectures for Networking and Communications Systems (ANCS '12)*, pp. 3–14, Austin, Tex, USA, October 2012.
- [12] T. Aktas, C.-H. Wang, and T. Javidi, "WiCOD: Wireless control plane serving an all-optical data center," in *Proceedings of the 2015 13th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks, WiOpt 2015*, pp. 299–306, ind, May 2015.
- [13] T. Li and S. Santini, "Energy-aware coflow and antenna scheduling for hybrid server-centric data center networks," in *Proceedings of the ICC 2017 - 2017 IEEE International Conference on Communications*, pp. 1–7, Paris, France, May 2017.
- [14] H. Roh, C. Jung, K. Kim, S. Pack, and W. Lee, "Joint flow and virtual machine placement in hybrid cloud data centers," *Journal of Network and Computer Applications*, vol. 85, pp. 4–13, 2017.
- [15] L. Luo, D. Guo, J. Wu, T. Qu, T. Chen, and X. Luo, "VLCcube: A VLC Enabled Hybrid Network Structure for Data Centers," *IEEE Transactions on Parallel and Distributed Systems*, vol. 28, no. 7, pp. 2088–2102, 2017.
- [16] A. S. Hamza, J. S. Deogun, and D. R. Alexander, "Wireless communication in data centers: A survey," *IEEE Communications Surveys and Tutorials*, vol. 18, no. 3, pp. 1572–1595, 2016.
- [17] G. Wang, D. G. Andersen, M. Kaminsky et al., "C-Through: Part-time optics in data centers," in *Proceedings of the 7th International Conference on Autonomic Computing, SIGCOMM 2010*, pp. 327–338, ind, September 2010.
- [18] N. Farrington, G. Porter, S. Radhakrishnan et al., "Helios: A hybrid electrical/optical switch architecture for modular data centers," in *Proceedings of the 7th International Conference on Autonomic Computing, SIGCOMM 2010*, pp. 339–350, ind, September 2010.
- [19] N. Farrington, Y. Fainman, H. Liu, G. Papen, and A. Vahdat, "Hardware requirements for optical circuit switched data center networks," in *Proceedings of the 2011 Optical Fiber Communication Conference and Exposition and the National Fiber Optic Engineers Conference, OFC/NFOEC 2011*, 3, 1 pages, March 2011.
- [20] N. Farrington, A. Forencich, G. Porter et al., "A multiport microsecond optical circuit switch for data center networking," *IEEE Photonics Technology Letters*, vol. 25, no. 16, pp. 1589–1592, 2013.
- [21] H. H. Bazzaz, M. Tewari, G. Wang et al., "Switching the optical divide: Fundamental challenges for hybrid electrical/optical datacenter networks," in *Proceedings of the 2nd ACM Symposium on Cloud Computing, SOCC 2011*, 30, Cascais, Portugal, October 2011.
- [22] A. Singla, A. Singh, K. Ramachandran, L. Xu, and Y. Zhang, "Proteus: A topology malleable data center network," in *Proceedings of the 9th ACM SIGCOMM Workshop on Hot Topics in Networks, HotNets-9*, Monterey, Calif, USA, October 2010.
- [23] J. Bowers, A. Raza, D. Tardent, and J. Miglani, "Advantages and control of hybrid packet optical-circuit-switched data center networks," in *Proceedings of the Photonics in Switching, PS 2014*, San Diego, Calif, USA, July 2014.
- [24] N. Hamedazimi, Z. Qazi, H. Gupta et al., "FireFly: A reconfigurable wireless data center fabric using free-space optics," in *Proceedings of the 2014 ACM Conference on Special Interest Group on Data Communication, SIGCOMM 2014*, pp. 319–330, Chicago, Ill, USA, August 2014.
- [25] A. S. Hamza, S. Yadav, S. Ketan, J. S. Deogun, and D. R. Alexander, "OWCell: Optical wireless cellular data center network architecture," in *Proceedings of the ICC 2017 - 2017 IEEE International Conference on Communications*, pp. 1–6, Paris, France, May 2017.
- [26] X. Zhou, Z. Zhang, Y. Zhu et al., "Mirror mirror on the ceiling: flexible wireless links for data centers," *ACM SIGCOMM Computer Communication Review*, vol. 42, no. 4, pp. 443–454, 2012.
- [27] M. Z. Zaaimia, R. Touhami, L. Talbi, M. Nedil, and M. C. E. Yagoub, "60-GHz Statistical Channel Characterization for Wireless Data Centers," *IEEE Antennas and Wireless Propagation Letters*, vol. 15, pp. 976–979, 2016.

- [28] T. H. Cormen, C. E. Leiserson, R. Rivest, and C. Stein, *Introduction to Algorithms*, The MIT Press, 2009.
- [29] S. Edelkamp, S. Schroedl, and S. Koenig, *Heuristic Search: Theory and Applications*, Morgan Kaufmann Publishers Inc, 2010.
- [30] NS3, <https://www.nsnam.org/>.
- [31] A. Greenberg, P. Lahiri, D. A. Maltz, P. Patel, and S. Sengupta, "Towards a next generation data center architecture: Scalability and commoditization," in *Proceedings of the SIGCOMM 2008 Conference and the Co-located Workshops - PRESTO'08: ACM Workshop on Programmable Routers for Extensible Services of Tomorrow*, pp. 57–62, Seattle, WA, USA, August 2008.

