Free Talk in the Air: A Hierarchical Topology for 60 GHz Wireless Data Center Networks

Chaoli Zhang, Fan Wu, *Member, IEEE*, Xiaofeng Gao, *Member, IEEE*, and Guihai Chen, *Member, IEEE*

***Abstract*— With the development of 60 GHz technology, data centers are going to be wireless. A fundamental challenge in wireless data center networking is how to efficiently use 60 GHz wireless technology to improve the performance. Many existing works have been proposed for this, but most of them do not perform well in connectivity or may not be flexible for different environments in data centers. This paper presents Graphite, a novel network structure that has many desirable features for wireless data centers. The whole architecture can be suitable for several different deployments of data centers. In Graphite, the problem of link blockage can be properly solved. Graphite makes best use of the propagation distance of 60 GHz and allows one server to communicate with as many other servers as possible. Graphite also improves the average node degree, which is more than any other existing wireless topologies on the same condition. Furthermore, Graphite can be suitable for data center with different deployments. We build a small testbed of Graphite to demonstrate its ability to solve the problem of link blockage. Results from theoretical analysis and extensive evaluations show that Graphite is a viable wireless topology for data center networks.**

***Index Terms*— 60 GHz, topology design, wireless data center.**

I. Introduction

MANY large data centers have been built to provide various cloud services such as social networking, data storage, online e-business, etc. The growing number of different clients and a variety of Quality of Service standards require the underlying service platforms to meet special system and communication requirements, including scalability, cost effectiveness, high throughput, and low oversubscription. As a result, a number of recent efforts have investigated techniques for deploying more efficient data center networks (DCNs) [4], [14], [16], [17], [28], [33].

Most state-of-the-art data centers are connected by wired links with specific network topologies [18], [22], [29], [43]. Although such connections can offer a range of benefits like fast transmission speed and multi-path routing mechanisms, they have some inherent drawbacks which cannot be easily solved by the deployment of wired network technology. The first is the difficulty of system scalability. To expand the scale for a wired data center, currently we cannot avoid the complicated recabling procedure, the reconfiguration of the whole network, and the replacement of hardware for compatibility issues, which waste huge amount of time and cost millions of dollars.

The second problem is the low bandwidth utilization efficiency among racks in a data center. The fixed numbers of fibers between switches and servers restrict the upper bound of pairwise available bandwidth, which cannot satisfy the unpredictable bandwidth requirements well (like dealing with elephant flows), or leverage the unbalanced traffic easily (like reducing hotspots in the network). Currently, we can only design complex, yet sometimes unpractical scheduling algorithms to better utilize the bandwidth, and bandwidth expansion is even more complicated.

The third drawback comes from the expensive system construction and maintenance cost. The cabling cost may take up to 3∼8% of the overall infrastructure budget [34], while the significant volume of space occupied by cables may degrade the overall cooling efficiency of the data center and increase the maintenance cost greatly.

To overcome these issues, 60 GHz wireless communication technology has been exploited to replace or supplement the underlying wired topology. This recently popularized wireless technology has enabled high speed data transmissions, which satisfies the transmission requirements of data center networks. The 60 GHz spectrum (57∼64 GHz) was set aside as unlicensed by the FCC in 2001 and the available 7 GHz of spectrum can support multiple wireless links of Gbps speeds.

60 GHz technology was first introduced to data center networks by Ramachandran *et al.* [35] as a solution to reduce the cabling complexity. Later, Kandula *et al.* [25] and Halperin *et al.* [19] implemented this idea and proposed *Flyways* as an incremental overlay network to the original wired network, in which HXI devices with horn antenna are placed on top of existing tree-like topologies in data centers to generate additional wireless links. As shown in Figure 1(a), Flyways can provide fast wireless connections between two racks within a short distance and without tween obstacles, but the distance constraint and obstacle-free requirement restrict the number of wireless links it can generate and the connectivity improvement it can bring to a data center topology.

Recently, Zhou *et al.* [49] proposed *3D Beamforming* to establish indirect line-of-sight communication between

Manuscript received April 22, 2016; revised November 26, 2016 and April 29, 2017; accepted September 6, 2017; approved by IEEE/ACM TRANSACTIONS ON NETWORKING Editor H. Zheng. Date of publication October 17, 2017; date of current version December 15, 2017. This work was supported in part by the State Key Development Program for Basic Research of China (973 project) under Grant 2014CB340303, in part by China NSF under Grant 61672348, Grant 61672353, Grant 61422208, and Grant 61472252, in part by the Shanghai Science and Technology Fund under Grant 15220721300, in part by the Program of International *S&Y* Cooperation under Grant 2016YFE0100300, in part by CCF-Tencent Open Research Fund, and in part by the Scientific Research Foundation for the Returned Overseas Chinese Scholars. *(Corresponding author: Fan Wu.)*

The authors are with the Shanghai Key Laboratory of Scalable Computing and Systems, Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: chaoli_zhang@163.com; fwu@cs.sjtu.edu.cn; gao-xf@cs.sjtu.edu.cn; gchen@cs.sjtu.edu.cn).

Digital Object Identifier 10.1109/TNET.2017.2755670

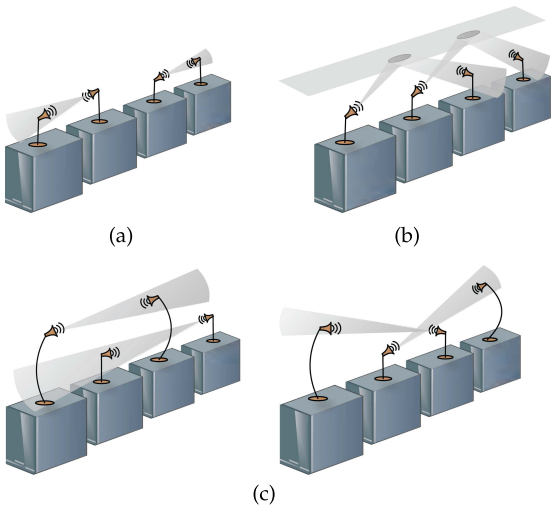


Fig. 1. Beamforming wireless links in (a) Flyways, (b) 3D Beamforming, and (c) Graphite.

two racks by carefully bouncing the 60 GHz signal over the ceiling of a data center, as shown in Figure 1(b). Using such a method, 3D Beamforming can generate more wireless links than Flyways with the same number of devices. Unfortunately, to guarantee the connectivity of data centers, 3D Beamforming has strict requirements on the quality of the ceiling and the height between wireless devices and the ceiling. First, the ceiling has to be as flat as possible, which is not prevalent in data centers. Additionally, the horizontal effective transmission range of wireless devices decrease rapidly when the height between ceiling and wireless device increases, whereas the ceiling is often very high for the cooling purpose. Thus the performance of 3D Beamforming may degrade sharply, and even performs same as Flyways.

Inspired by their works, in this paper we propose a new wireless topology named *Graphite*. As shown in Figure 1(c), Graphite is a multi-layer design derived from traditional mesh topology and has its own new features. Compared with the existing hybrid topologies even using the same number of wireless devices, Graphite has some advantages as follows.

- Compared with Flyways, Graphite efficiently solves the link blockage problem by introducing a liftable and rotatable crank arm. In Graphite each wireless device has the ability to communicate with more peers within its transmission range.
- Compared with 3D Beamforming, Graphite relaxes the strong requirements on the ceilings and the betweenness distance. The performance of Graphite is not affected by the ceiling height and quality, and Graphite increases the average node degree of data center networks, bringing potentially more benefits for routing and load balancing.
- Graphite can alleviate the inflexibility issue in existing data center networks without sacrificing scalability. It generates almost the largest number of wireless connections for a wireless device within its transmission range. Moreover, it is applicable for many different data centers with various environments.

The purpose of our design is to construct a supplementary wireless network to improve the connectivity and scalability of a data center network with the help of 60 GHz wireless technology. Besides the discussions above, we also make the following contributions in this paper.

- First, we provide a detailed system design of Graphite, and illustrate its construction. We implement a small testbed of Graphite, whose results validate the effectiveness of our design. We also introduce a generalization of Graphite such that it can be implemented in many data centers with different environments.
- Second, we precisely depict the mathematical model of Graphite and provide theoretical analysis on the comparison of Graphite with Flyways and 3D Beamforming.
- Finally, we evaluate Graphite by extensive evaluations. Both our theoretical analysis and evaluation results verify that Graphite does perform better than the existing wireless topologies.

The rest of the paper is organized as follows. In Section II, we review related works. In Section III, we present the design of our communication frontend. In Section IV, we describe topology design of Graphite. In Section V, we give our analysis of Graphite. In Section VI, we report numerical results. Finally, Section VII concludes the paper.

II. Related Works

Traditional data centers use commodity-class computers and switches connected by wired links. There are many representative network topologies, briefly divided as switch-centric and server-centric topologies. In switch-centric topologies, switches take charge of majority interconnection and routing works, while most of the servers act as the end users in the network. Tree-like topologies are usually switch-centric, such as Fat-Tree [4], Portland [33], Elastic Tree [22], Aspen Tree [43], Diamond [40], and their variations like VL2 [14] and Monsoon [15]. Other topologies like FBFLY [2] also refer to switch-centric DCNs.

Different from these switch-centric network topologies, in server-centric topologies, servers enable the functions of interconnection and routing while the switches only provide easy crossbar function. Server-centric topologies are often defined recursively with multiple layers, in which a high-level structure consists of certain low-level structures and the structures at the same level are connected with each other by a well-defined way. Typical examples include DCell [17], BCube [16], MDCube [46], FiConn [28], CamCube [3], HCN [18], and SWCube [29]. More details are available in surveys [6], [21].

Although these proposals bring benefits for routing purpose, they significantly increase network construction complexity and scalability difficulties, because of their strict rules on the number of required links or switches and complicated cabling strategies. Optical circuit switching technology can be one of the possible replacement to supply high bandwidth for data centers, like C-Through [44] proposed by Wang *et al.*. There are also many effective works in this area, such as Helios [12], Hedera [5], DOS [47], Proteus [38], REACToR [30], and Quartz [31]. However, such optical circuit switching incurs relatively substantial cost while the wireless data center has the

benefits of reduced cost and deployment complexity. Readers can refer to a survey [24] for details.

Because of the favorable characteristics of 60 GHz signals, several works began to use 60 GHz wireless technology to improve the performance of data center networks. This technology was first proposed to data center networking by Ramachandran *et al.* [35] as a solution to reduce the cabling complexity. They categorized the wireless communication patterns in data centers into three groups, Line-of-Sight (LOS) between racks, indirect Line-of-Sight with reflectors, and multi-hop Non-Line-of-Sight (NLOS). They also proposed some considerable problems and challenges in realizing 60 Ghz wireless technology applied in the data centers.

Later, Kandula *et al.* [25] and Halperin *et al.* [19] proposed Flyways to augment data centers networks with embedded wireless links. In their work, The backbone network is a common oversubscribed DCN, while each Top-of-Rack (ToR) commodity switch is attached by one or more 60 GHz wireless devices with electronically steerable directional horn antennas. Several multi-Gbps 60 GHz wireless links can parallel work between pairwise ToR's to supply extra link capacity. Flyways can be viewed as a milestone to implement wireless links as an overlay network on top of the underlying wired networks, which significantly improve the system performance with low cost of wireless devices. Many researchers have paid attention to this area ever since.

From then on, researches about 60 GHz technologies can be divided into two categories: architecture design and performance optimization. The former focuses on the physical characteristics of wireless links, the topological connections of hybrid data center networks (HDCN), and the potential benefits of wireless communications, such as beamforming [42], beamsteering [26], MIMO link design [27], and RF-HYBRID [23]. Researchers even considered the wireless datacenter fabric using free-space optics [20]. The latter mainly discusses the optimization strategies for HDCNs to improve the system performance with the help of wireless connections. For instance, many researches concentrate on the channel allocation of wireless links in data center [7], [9]. Some of them [8], [10] attempt to divert traffics from hotspots by wireless links without interferences to maximize the throughput and utilization of the whole network.

Researchers even considered the feasibility of constructing a completely wireless DCN [36]. They designed a new topology and proposed cylinder racks and sector servers to fully utilize wireless connections. However, their topologies are planar and have high rack density, in which most links are on the same plane with risk of the problem of link blockage. The transmission range of 60 GHz signals also restricts the neighbors each node may connect to, and thus decreases the connectivity greatly.

Recently, an innovative work, 3D Beamforming [48], [49], was proposed to let 60 GHz signals bounce off data center ceilings using beamforming technology. This work establishes indirect line-of-sight connection between any two racks in a data center, solving the link blockage problem efficiently. They also proposed Angora [50], a low-latency facilities network using low-cost, 60 GHz beamforming radios that provide robust paths decoupled from the wired network. However, in modern data centers, ceilings are often very high for cooling purpose, and 3D Beamforming may not be applicable when the height of ceilings is not feasible.

Diamond [11] is another good work in wireless data center network architecture. They proposed the deployment of scalable 3D Ring Reflection Space interconnected with streamlined wired herringbone to enable large number of concurrent wireless transmissions. The novel Diamond nests the wired DCN with radios equipped on all servers. It is a different design from a complete wireless data center network and the preciseness that the multi-reflection of wireless links requires is not easy to achieve in reality.

Consequently, we propose Graphite, which is suitable for most deployments of data center. Our topology not only solves the link blockage problem properly, but also allows each rack in data centers to connect with most of its neighbors directly up to the transmission range of 60 GHz signals.

III. Communication Frontend Design

In this section, we present the design of our liftable 3D Beamforming radio frontend, and show its effectiveness on bypassing obstacles in our testbed.

A. 60 GHz Communication Preliminary

As a license-free wireless communication technology, 60 GHz radios show significant superiorities compared with their companies of 2.4 GHz and 5 GHz radios, and have been adopted by multiple designs of data center network topology (*e.g.*, [25], [35], [49]). Standard 60 GHz protocols, say WirelessHD [45] and WiGig (also known as IEEE 802.11ad) [1], have been released to support up to 28 Gbps and 7 Gbps point-to-point transmissions, respectively. Due to fast attenuation, 60 GHz radios employ beamforming to increase receive/transmit gain by concentrating transmission energy in a specific direction. Off-the-shelf 60 GHz beamforming radios in the form of both horn antennas [32] and antenna arrays [37] have become affordable in recent years. Incorporated with mechanical or electronic mechanisms, they can have fine-grain directional control [19], [41] and less likely to be interfered by other nearby 60 GHz transmissions.

B. Design of Liftable 3D Beamforming Radio Frontend

To cope with link blockage, existing approaches either carefully place 2D beamforming radios on top of the racks [25], or implement an elegant integration of 3D beamforming and a metal reflector [49]. However, while 2D beamforming avoids or reduces link blockage by limiting the rotation flexibility of the radios and relying on bandwidth-consuming multi-hop relay, reflection-based 3D beamforming may considerably shorten span of the communication range in horizontal plane, especially when the reflecting ceiling is far away from the radios.

To overcome the shortages of the aforementioned approaches, we design a hierarchical 3D beamforming

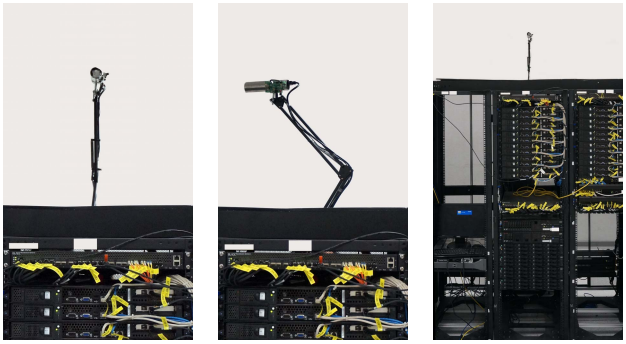


Fig. 2. Graphite's liftable crank arm with 3D beamforming radio frontend. (a) Front view. (b) Side view. (c) Full view.

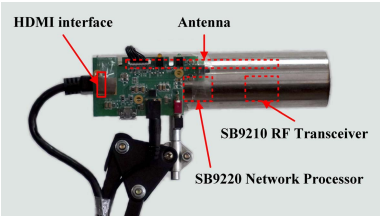


Fig. 3. Graphite's 3D beamforming radio frontend. Parts inside the metal waveguide is indicated using dashed rectangles.

approach, in which we integrate each beamforming horn with a liftable and rotatable crank arm, as shown in Figure 2. With the help of the crank arm, radios can be risen or lowered to appropriate levels and rotated into appropriate directions. Thus, nonblocking radios in the same layer and from different layers can have line-of-sight wireless connections.We adopt the crank arm to avoid link blockage cased by the straight arm, and thus links in lower levels will not be blocked by carefully rotating the crank arm.

Nowadays, several motion control products of FLIR, such as Pan-Tilt Unit-D47 and Pan-Tilt Unit-E46, can achieve max speed of $300°/Sec$ with position resolution as small as $0.003°$ [13]. Thus, although in our prototype, the movement of radios are done manually, automatic rotating and lifting devices for 3D beamforming radio can be easily integrated in the commercial product.

Figure 3 shows our 3D beamforming radio frontend. The radio frontend integrates a SB9220 network processor and a SB9210 RF transceiver providing 4 Gbps over-the-air data transfer rate within 10 meters [39], and is connected with a metal waveguide to direct radio waves in a beam. It implements WirelessHD 1.0 standard on a 7 GHz channel in the 60 GHz extremely high frequency radio band.

According to [49], 3D Beamforming can reach 6+Gbps even at a very low transmit power when the two endpoints are separated by not over $10m$. [49] also shows that the link data rate decreases rapidly when the link distance of 3D (2D) Beamforming is more than $10m$ ($11m$). Thus, we assume the 60 GHz signal propagates $10m$ in our paper. What needs to be pointed out is that this assumption does not cause significant difference to our design.

C. Nonblocking Communication With Liftable 3D Beamforming Radios

To demonstrate the effectiveness of our design, we have built a liftable 3D beamforming radio testbed. The testbed is used to study the feasibility of using liftable 3D beamforming radios to implement nonblocking communications on top of the racks.The testbed consists of a pair of liftable 3D beamforming radios, between which one for transmitting and the other for receiving. We vary the distance between the pair of radios from $1m$ to $9m$ with a step of $0.5m$. By adjusting the liftable heads onto suitable layers and turning the radios into suitable directions, a 4 Gbps communication link between the pair of radios can be established without obvious lag.

Figure 4 shows the deployment of a pair of liftable 3D Beamforming radios onto 2 different layers. By adjusting the liftable heads onto different layers and turning the direction of the radios, a communication link between the pair of radios can be established. Next,We place a metal square plate with the size of $5cm \times 5cm \times 3mm$, which can block 60 GHz wireless signals, to act as an obstacle radio between the pair of radios. As shown in Figure 5, we have tested four representative deployments of the liftable 3D beamforming radios. Our experiment results show that the communication link can be formed in all the four cases. Especially, in case 1 and 3, the communication links are not blocked given the crank arm.

What's more, we also lower the power of the transmitter and measure the signal strength in a $1m \times 1m$ square right in front of the radio in order to infer the interference range. The sniffer (receiver) is an omni-directional antenna.

As we can see in Figure 6, the radio has fine-grained direction control. The beam emitted by the horn antenna propagates in a cone-shape and does not spread out much. Actually, the largest spread-out angle that we calculate from Figure 6(b) is about 2α, that is $2 \times 9.04°$, *i.e.*, the interference area of a radio is a right cone whose aperture is $18.08°$. The base of such a cone with height $h = 10m$ is a circle of radius $0.9m$. Furthermore, we also detect the interference range at the crow-fly distance of $10m$ away from the transmitter in its main transmission direction. The interfered area is roughly a circle with radius $0.9m$, which is also coincident to the result of Flyways got through experiments [49] in the rough.

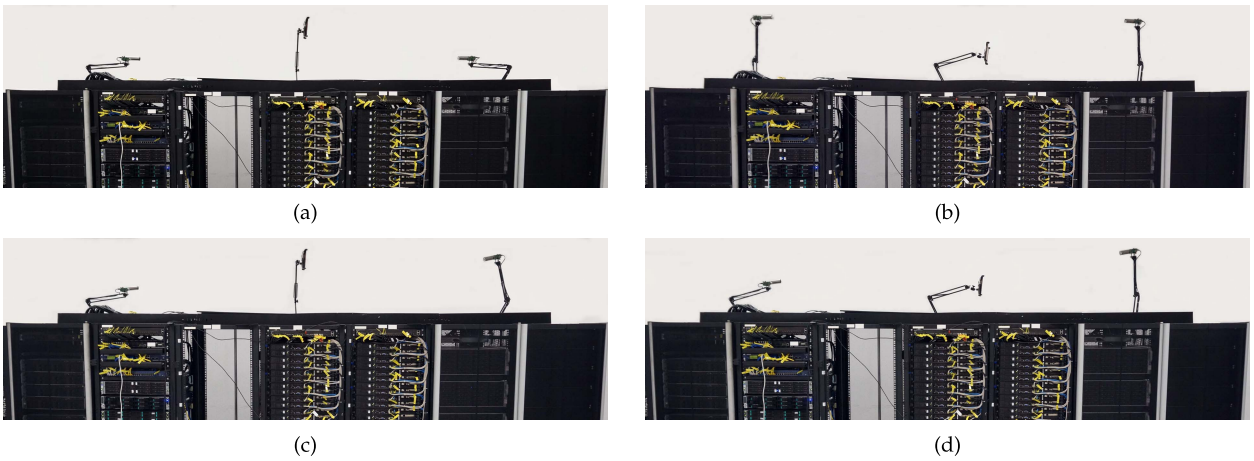
Given the above measurement results, we can make a pair of radios communicate with each other without interfering with the other radios by carefully designing the 3D topology of the 60 GHz wireless network. More specifically, in Graphite, we achieve this goal by setting radios onto different layers and carefully designing the vertical distance between layers which is shown in Section V-E in detail.

IV. Topology Design

Most of existing designs intelligently place radios on top of the racks in the same horizontal plane (*e.g.*, [25], [49]). However, these approaches are either only effective for some rack pair connections, or significantly shorten the communication range for seeking the help of reflection by ceiling. In contrast to existing designs, we propose a multi-layer topology, namely Graphite, in which the radios are carefully

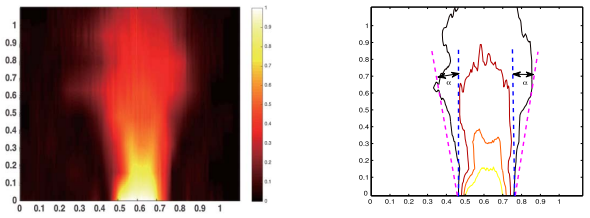
(a) (b) (c) (d)

Fig. 4. Deployment of a pair of liftable 3D Beamforming radios onto 2 different layers. (a) Case 1: Both of the radios are on the lower layer. (b) Case 2: The right radio is higher than the left one. (c) Case 3: The right radio is lower than the left one. (d) Case 4: Both of radios are on the higher layer.



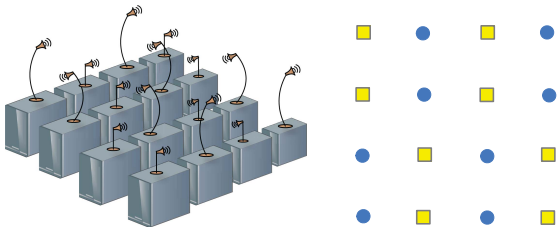
(a) (b) (c) (d)

Fig. 5. Deployment of a pair of liftable 3D beamforming radios onto 2 different layers free of blockage by an obstacle. (a) Case 1: The radios are on a lower layer than the obstacle. (b) Case 2: The radios are on a higher layer than the obstacle. (c) Case 3: The radios are on different layers, and the obstacle is on the same layer as the higher radio. (d) Case 4: The radios are on different layers, and the obstacle is on the same layer as the lower radio.



(a) (b)

Fig. 6. The beam pattern of the radio. (a) Heat map. (b) Contour map.



(a) (b)

Fig. 7. A 4×4 square mesh of radios. Yellow square dots and blue round dots represents the radios residing in the higher layer and the lower layer, respectively. (a) Side view. (b) Top view.

placed onto layers with different heights. We show that such a topology can greatly reduce the link blockage, and thus increase the connectivity among the top-of-rack radios.

The objective of the topology design is to let each radio in the wireless data center network can connect to as many radios as possible within its communication range. We consider that the radios are homogenous, and have the same communication range R.

To clearly present the idea of the topology design, we start from a simple case of mesh unit, in which there is a 4×4 square mesh of radios, as shown in Figure 7. In this mesh unit, the horizontal and vertical distances between two adjacent radios are no more than $R/3$, such that every pair of radios residing in the same row or column in the mesh unit can communicate with each other if the line of sight path between them is not blocked.

Figure 7 also shows our design for the topology of the 4×4 square mesh unit. This topology has two layers, a lower layer and a higher layer. In Figure 7(b), we use yellow square dots and blue round dots to represent the radios residing in the higher layer and the lower layer, respectively.

Such a topology guarantees that any pair of radios within the communication range can talk to each other without

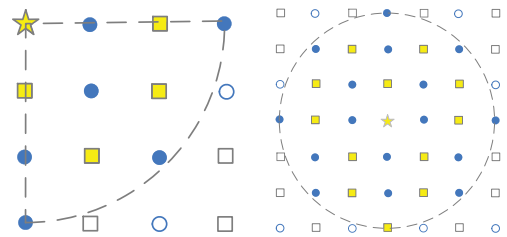


Fig. 8. Coverage of a radio. (a) Coverage in a single mesh unit. (b) Coverage in a mesh tiled by the mesh unit.

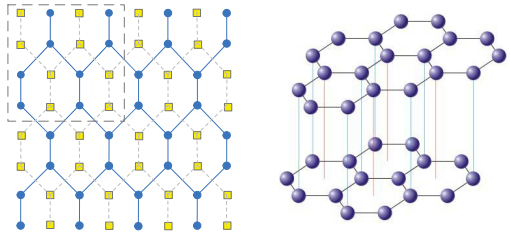


Fig. 9. Topologies of 2-layer Graphite. (a) 2-layer Graphite tiled with a 4×4 mesh unit, which is indicated by a dashed square. (b) Molecular structure of graphite.

blockage. As indicated by Figure 8, the top-left radio can talk to 10 radios in a single mesh unit, and a radio can connect to up to 28 radios within its communication range in a mesh tiled by this mesh unit.

When tiling the mesh with the above mentioned unit, we observe that the 2-layer Graphite looks very similar to the molecular structure of graphite, as shown in Figure 9. Consequently, we name the design of the topology as Graphite. Furthermore, illuminated by the molecular structure of graphite, we extend the 2-layer Graphite to adapt to the general case of l layers. The denser the wireless mesh is, the more layers are needed to facilitate non-blocking communications.

First, we need to determine how many layers are needed to avoid link blockage. We consider that the largest rectangular unit within the range of an $R \times R$ square area is an $m \times n$ mesh of radios.[1] In each row, to avoid having more than 2 radios in the same layer, we need at least $\lceil \frac{n}{2} \rceil$ layers. Similarly, in each column, to guarantee having at most 2 radios in the same layer, we need at least $\lceil \frac{m}{2} \rceil$ layers. Consequently, the minimum number of layers needed is

$$l = \max\left(\left\lceil \frac{m}{2} \right\rceil, \left\lceil \frac{n}{2} \right\rceil\right). \quad (1)$$

We set the layers to have increasing heights from layer 0 to layer $l-1$.

Next, we set the layer for each of the radios in an $x \times y$ mesh of wireless data center network. Illuminated by the molecular structure of graphite, we adjust the layer for every radio in each row, and jump to another layer by every two radios in each column, without loss of generality. Specifically, the layer of the first radio in row i $(0 \leq i < x)$ is set to layer $\lfloor i/2 \rfloor \% l$. Then,

[1]The selection of m (n) should leave enough room for radios to form communication links from different layers in the same column (row). Let δ_{row} (δ_{column}) denote the distance between two adjacent racks in a row (column). Consequently, the selection of m and n should be no larger than $\lfloor R/\delta_{column} \rfloor$ and $\lfloor R/\delta_{row} \rfloor$, respectively.

Algorithm 1 Graphite Topology Construction

Input: An $x \times y$ mesh of wireless data center network, the maximum mesh $m \times n$ unit within the area of an $R \times R$ square.

Output: Number of layers needed l, a profile of layer allocation for the radio in the mesh $h^{x,y}$.

```
l ← max(⌈m/2⌉, ⌈n/2⌉);
for i ← 0 to x − 1 do
    for j ← 0 to y − 1 do
        h_{i,j} ← ((⌊i/2⌋%l) + j)%l;
    end
end
```

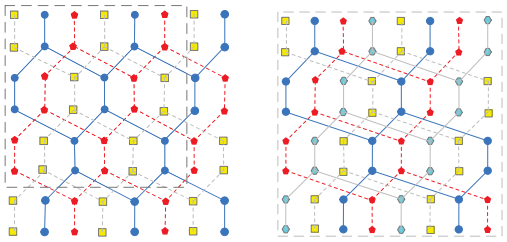


Fig. 10. Topologies of 3-layer and 4-layer Graphite. (a) 3-layer Graphite tiled with a 6×6 mesh unit. (b) 4-layer Graphite's 8×8 mesh unit.

the radio at position (i, j), where $0 \leq i < x$ and $0 \leq j < y$, is set to layer $h_{i,j} = ((\lfloor i/2 \rfloor \% l) + j) \% l$.

Algorithm 1 shows the pseudo-code of the above process for layer determination.

Sample topologies with 3 and 4 layers produced by our algorithm are shown in Figure 10. From the figures, we can see that our algorithm maps the radios to a well-organized graphite structure.

We note that we can use a smaller number of layers than that calculated by Equation (1) to limit the radios within a constrained vertical range, with slightly lowered average node degrees in practice.

What's more, we note that, we can install multiple radios per rack in Graphite similar to what is done in 3D Beamforming and Flyways. Thus, the rack can talk with multiple racks at the same time. The main topology design will not be influenced by this as the radios can rotate to different directions in Graphite, so we omit this discussion not only due to the lack of space, but also to make our design more concise to be understood.

V. Theoretical Analysis

In this section, we will investigate the performance of Graphite and prove theoretically that our design of Graphite is feasible and highly efficient, which makes the best use of the 60 GHz wireless technology.

A. Holistic Approach

We consider an $x \times y$ mesh of ToRs (Top-of-Rack switches embedded with 60 GHz antenna) in the wireless data center network. Number of racks in every row and every column is represented by x and y respectively. The horizontal distance

between two adjacent racks is denoted by d while the vertical interval is w. For the sake of convenience, we denote the reciprocal of d by d_r.

We use graph $G = (V, E)$ to stand for the wireless network in this section. Here $V = \{\text{ToR}_i\}$ consisting of all the ToRs with 60 GHz antenna, and $E = \{(\text{ToR}_i, \text{ToR}_j)\}$ containing pairs of ToRs that can connect directly with one another through wireless links.

In order to demonstrate the higher connectivity of Graphite than other designs, such as Flyways and 3D Beamforming, we introduce two metric Δ and θ to evaluate the performance of various designs. Here Δ represents the average node degree of a wireless network topology, namely

$$\Delta = \frac{1}{|V|} \sum_{v \in V} \deg(v),$$

where $\deg(v)$ is the degree of vertex v. And θ stands for the node coverage percentage for a ToR compared with the number of ToRs in the range of propagate ratio R. More specifically,

$$\theta = \frac{\deg(v)}{|\{v' : \text{dist}(v, v') < R\}|},$$

where v can be any vertex in the center part of network, not considering the boundary effect and $\text{dist}(v, v')$ means the horizonal distance between two nodes v and v'.

From the definition, we know that the larger the parameters Δ and θ are, the more neighbors one node can communicate with averagely, the higher connectivity the wireless network will get and the better application of the wireless technology will be. In the next three subsections, we will analyze and calculate the metric Δ for each design (Flyways, 3D Beamforming, Graphite) respectively, and thereby show the superiority and efficiency of Graphite. First of all, however, we need the following lemma to help us analyze the performance of different designs.

Lemma 1: The average node degree Δ of graph $G = (V, E)$ can be derived from:

$$\Delta = \frac{2|E|}{|V|}.$$

It is remarkable that the order of graph G is $|V| = xy$ according to our assumption. Hence, in order to calculate the average node degree Δ with the aid of Lemma 1, we only need to figure out the size of G, namely the number of edges of G.

To compare the performance of three topologies, we define graph G_{flyway}, G_{mirror}, and $G_{graphite}$ to denote the wireless graphs formed by Flyways, 3D Beamforming, and Graphite respectively. Easy to see, these three graphs have the same vertex set V, and the only difference of these three graphs comes from the edge set E.

B. Performance of Flyways

In this subsection, we evaluate the performance of Flyways design. First of all, we calculate the size of graph G_{flyway}.

Theorem 1: In Flyways design, the total number of edges in graph G_{flyway} is

$$|E_{flyway}| = y(x-1) + (y-1)[(2u+1)x - u(u+1)],$$

where $u = \lfloor d_r \sqrt{R^2 - w^2} \rfloor$.

Proof: Due to the limitations and characteristics of wireless technology, wireless connections occur only between adjoin racks or racks within a certain distance from two adjacent rows. Consequently, all edges in E_{flyway} can be classified into two categories:

Case 1: Links in one row.

In this case, the two switches are next to each other. Since there are $x-1$ pairs of adjacent ToRs in every row, the number of such edges adds up to

$$y \cdot (x-1).$$

Case 2: Links between two neighboring rows.

As for a particular ToR in V_{flyway}, it connects to at most $2u+1$ ToRs in one neighboring row considering the limitation of distance, where $u = \lfloor d_r \sqrt{R^2 - w^2} \rfloor$.

However, nodes near the boundary can only connect to a portion of the above maximum number. By subtracting such edges, we could derive the number of edges between two adjacent rows as

$$(2u+1) \cdot x - 2 \sum_{i=1}^{u} i = (2u+1)x - u(u+1).$$

Therefore, the number of edges altogether in this case is

$$(y-1) \cdot [(2u+1)x - u(u+1)].$$

Eventually, we add the above two numbers together and acquire the overall number of edges in graph G_{flyway}, as has been showed in the conclusion of the theorem. □

Then we are able to obtain the average node degree of Flyways design by simply applying Lemma 1.

Theorem 2: In Flyways design, the average node degree is

$$\Delta_{flyway} = 2\left(1 - \frac{1}{x}\right) + 2\left(1 - \frac{1}{y}\right)\left[(2u+1) - \frac{u(u+1)}{x}\right].$$

C. Performance of 3D Beamforming

In this subsection, we will analyze the performance of 3D Beamforming design. Similarly, we use the notation G_{mirror} to represent the wireless graph, derived from Flyways embedded with ceiling reflectors. The height from ToRs to the ceiling is denoted by h.

With similar steps, we firstly calculate the number of edges in graph G_{mirror}. Here $\text{sgn}(\cdot)$ is the sign function, i.e. $\text{sgn}(x)$ equals 1 if x is positive, -1 if x is negative, and 0 if $x = 0$.

Theorem 3: In 3D Beamforming design, the total number of edges of graph G_{mirror}, namely $|E_{mirror}|$, is

$$\sum_{i=0}^{2} \frac{1 + \text{sgn}(i)}{2} (y-i)[(2v_i+1)x - v_i(v_i+1)] - \frac{1}{2}xy,$$

where

$$\begin{aligned} v_0 &= \lfloor d_r R' \rfloor, \\ v_1 &= u = \left\lfloor d_r \sqrt{R^2 - w^2} \right\rfloor, \\ v_2 &= \left\lfloor d_r \sqrt{R'^2 - 4w^2} \right\rfloor, \\ R' &= \sqrt{R^2 - 4h^2}. \end{aligned}$$

Proof: The critical difference between Flyways and 3D Beamforming is that wireless links could be constructed with the help of ceiling reflection in 3D Beamforming design. Therefore, more edges will appear in the wireless network graph. Since the propagation radius is R and the height from the top of rack to the ceiling is h, it is equivalent to say that the actual horizonal propagation radius is $R' = \sqrt{R^2 - 4h^2}$. In other word, the maximum horizonal distance the signal can transmit is R', on the condition that the signal is reflected by the ceiling once.

In a traditional data center network, the height of ceilings is usually 2m to 3m, while the propagation range is about 10m. Hence, the actual propagation radius of 3D Beamforming design is approximately 8m to 9m. In comparison, the distance between two adjacent rows is about 2.4m, and the depth of a rack is around 1.2m [49], so the distance between two adjacent racks in vertical line $w \approx 3.6$m. Therefore, wireless connections could only be constructed between rows that are located with at most one row between them. All edges in graph G_{mirror} are accordingly classified into three categories:

Case 1: Links in one row.

Since ordinary wireless connections occur only between adjoin racks, this sort of links can all be build through ceiling reflections. The actual propagate range is R', so one node could connect to at most $2v_0$ nodes in the same row, where $v_0 = \lfloor d_r R' \rfloor$. Similar to Theorem 1, we can formulate that links in one row is

$$\frac{1}{2} \cdot [2v_0 x - v_0(v_0 + 1)],$$

and the overall number of edges of this kind is

$$y \cdot \frac{1}{2}[2v_0 x - v_0(v_0 + 1)].$$

Case 2: Links between adjacent rows.

In this situation, ToRs can be connected by wireless links directly within the transmission range. As a result, the number of links is exactly the same as what we have acquired in Theorem 1 Case 2, which is

$$(y - 1) \cdot [(2v_1 + 1)x - v_1(v_1 + 1)],$$

where $v_1 = u = \left\lfloor d_r \sqrt{R^2 - w^2} \right\rfloor$.

Case 3: Links between rows with one row in the middle of them.

In this case, wireless links can only happen through ceiling reflection. Similar to the calculation in Theorem 1 Case 2, one ToR could connects to at most $2v_2 + 1$ ToRs in one satisfactory row, where $v_2 = \left\lfloor d_r \sqrt{R'^2 - 4w^2} \right\rfloor$. The number of edges between two such rows is

$$(2v_2 + 1)x - v_2(v_2 + 1).$$

Therefore, the number of edges altogether in this case is

$$(y - 2) \cdot [(2v_2 + 1)x - v_2(v_2 + 1)].$$

In the end, the overall number of edges in graph G_{flyway} can be obtained by summing up the results we get in these three cases. □

In the light of Lemma 1, we can figure out the average node degree of 3D Beamforming as the following theorem.

Theorem 4: In 3D Beamforming design, the average node degree Δ_{mirror} equals

$$\sum_{i=0}^{2} [1 + \mathrm{sgn}(i)] \left(1 - \frac{i}{y}\right) \left[(2v_i + 1) - \frac{v_i(v_i + 1)}{x}\right] - 1.$$

D. Performance of Graphite

In this subsection we try to demonstrate the superiority of our design, Graphite. We use two-layer Graphite as an example and evaluate its connectivity. Easy to see, more layers in Graphite will bring denser connections, which performs even better. $G_{graphite}$ represents the wireless network of two-layer Graphite. We use h_2 to stand for the vertical distance between the first and the second layer in Graphite.

First of all, we figure up the number of edges of graph $G_{graphite}$ by drawing analogies between two-layer Graphite and the previous two designs.

Theorem 5: The size of graph $G_{graphite}$ is

$$\begin{aligned} &|E_{graphite}| \\ &= \frac{1}{4} y(5x - 8) + \frac{1}{2}(y - 1)[(2u + 1)x - u(u + 1)] \\ &\quad + \frac{1}{2} \sum_{i=0}^{2} \frac{1 + \mathrm{sgn}(i)}{2} (y - i)[(2w_i + 1)x - w_i(w_i + 1)], \end{aligned}$$

where

$$\begin{aligned} u &= \left\lfloor d_r \sqrt{R^2 - w^2} \right\rfloor, \\ w_0 &= \lfloor d_r R'' \rfloor, \\ w_1 &= \left\lfloor d_r \sqrt{R''^2 - w^2} \right\rfloor, \\ w_2 &= \left\lfloor d_r \sqrt{R''^2 - 4w^2} \right\rfloor, \\ R'' &= \sqrt{R^2 - h_2^2}. \end{aligned}$$

Proof: Since the wireless links can be constructed between ToRs in the same layer or from different layers, we calculate the number of these two sorts of edges respectively.

Case 1: Links in one layer.

In this case, the ToRs at the same height forms a topology like the molecular structure of graphite. Links can be constructed not only in the same row and between neighboring rows, but also at the principal diagonal in each hexagon. We discuss these three subcases separately.

1) Edges in one row could only take place between adjacent racks, thus the total number of such edges is

$$y \cdot (x - 2).$$

2) As for edges between adjacent rows, the number of this kind of edges is half of that in Flyways because only

half of the other nodes are in the same layer of it. Hence, the overall number of edges in this situation adds up to

$$\frac{1}{2}\cdot(y-1)\cdot[(2u+1)x-u(u+1)],$$

where $u=\lfloor d_r\sqrt{R^2-w^2}\rfloor$.

3) Finally for the remaining edges, there is one edge in each hexagon, and each node is adjacent to exactly one such edge. This yields that the number of such edges is $\frac{1}{2}xy$.

All in all, the total number of edges in Case 1 sums up to

$$y(x-2)+\frac{1}{2}(y-1)[(2u+1)x-u(u+1)]+\frac{1}{2}xy,$$

where $u=\lfloor d_r\sqrt{R^2-w^2}\rfloor$.

Case 2: Links between two layers.

The height difference between two layers is h_2, so the horizonal transmission radius is $R''=\sqrt{R^2-h_2^2}$. Compare with 3D Beamforming design in which the height of the ceiling is set to $h_2/2$, we can easily find out that the degree of each node in this case is half of the corresponding one in 3D Beamforming, because only half of other nodes are in different layer. The only difference between them is that in 3D Beamforming we can only build links with ceiling reflection rather than connection directly. Consequently, the number of edges altogether in this case is

$$\frac{1}{2}\sum_{i=0}^{2}\frac{1+\mathrm{sgn}(i)}{2}(y-i)[(2w_i+1)x-w_i(w_i+1)]-\frac{1}{4}xy,$$

where

$$\begin{aligned}w_0&=\lfloor d_rR''\rfloor,\\ w_1&=\left\lfloor d_r\sqrt{R''^2-w^2}\right\rfloor,\\ w_2&=\left\lfloor d_r\sqrt{R''^2-4w^2}\right\rfloor.\end{aligned}$$

In conclusion, by adding the number of edges in the above two cases, we obtain the size of $G_{graphite}$ and prove the theorem. □

Finally, Lemma 1 yields the following theorem in which the average node degree of two-layer Graphite is formulated.

Theorem 6: In two-layer Graphite design, the average node degree is

$$\begin{aligned}&\Delta_{graphite}\\ &=\left(\frac{5}{2}-\frac{4}{x}\right)+\left(1-\frac{1}{y}\right)\left[(2u+1)-\frac{u(u+1)}{x}\right]\\ &\quad+\sum_{i=0}^{2}\frac{1+\mathrm{sgn}(i)}{2}\left(1-\frac{i}{y}\right)\left[(2w_i+1)-\frac{w_i(w_i+1)}{x}\right].\end{aligned}$$

Compared with the results from Theorem 2, and Theorem 6, the formula of average node degree indicates that Graphite has much higher connectivity than Flyways. Meanwhile, data centers usually have high ceilings for the purpose of cooling, thus the parameter h is usually large. As a result, Graphite outperforms 3D Beamforming as the average node degree of the latter design declines rapidly when h increases while Graphite solves this problem by setting antennas at distinct levels without utilizing the ceiling (see the comparison between Theorem 4 and Theorem 6). More detailed and specific numerical results will be given in Section VI.

In conclusion, theoretical evaluation demonstrates the considerable connectivity of Graphite, which leads to higher efficiency and better performance than the existing designs, Flyways and 3D Beamforming.

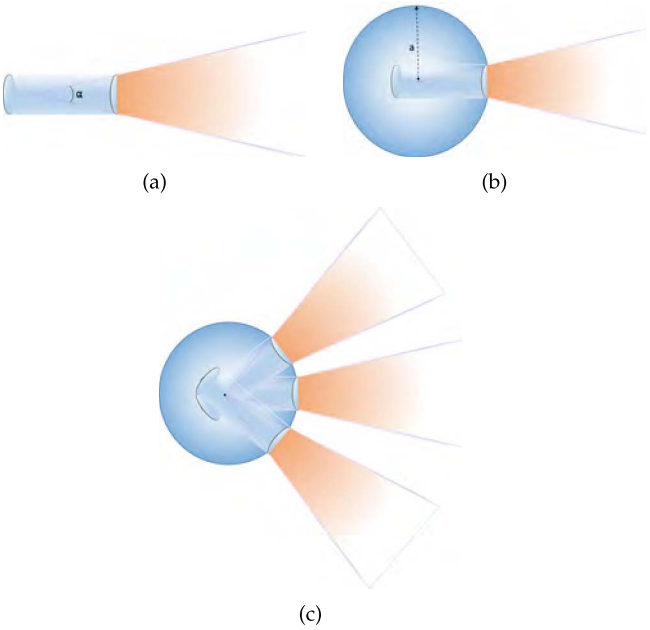


Fig. 11. Radio Model for Calculating Vertical Distance Between Layers. (a) Signal Model of the Radio. (b) Radio Model. (c) Rotating Radio Model.

E. Vertical Distance Between Layers

In our design, we put radios onto different layers to solve the problem of link blockage. However, we cannot simply regard each radio as an infinitely small point. Thus, the vertical distance between two adjacent layers should be large enough to guarantee the existence of line-of-sight communication channel. In fact, they are horns with a definite volume and occupy a certain amount of space. Although the antenna array in Figure 3 is formed by several small antennas, according to what has been discussed in Section III-C and for the sake of concision, it is reasonable to consider that the vertex of the 60 GHz signal cone is one point inside the radio as shown in Figure 11(a). What's more, as the radio can rotate around, we can model that each radio occupies a spherical space with a radius of a and the centre is regarded as the vertex of the 60 GHz signal cone which is shown in Figure 11(c).

According to the Graphite Topology Construction Algorithm 1 and the result of this construction in Figure 9, we can get that if we need to make sure that the signals would not be blocked by the neighboring horns and at the same time the neighboring horns would not be influenced, which could be called the $free\ connection\ condition$, we just need to consider the vertical distance between two layers in the same column. The reason is relatively obvious and can be seen more clearly in Figure 12 that if the radios in the same row and adjacent layers want to communicate with each other, two neighbor horns should be considered to satisfy the free connection condition at most while in the same column no

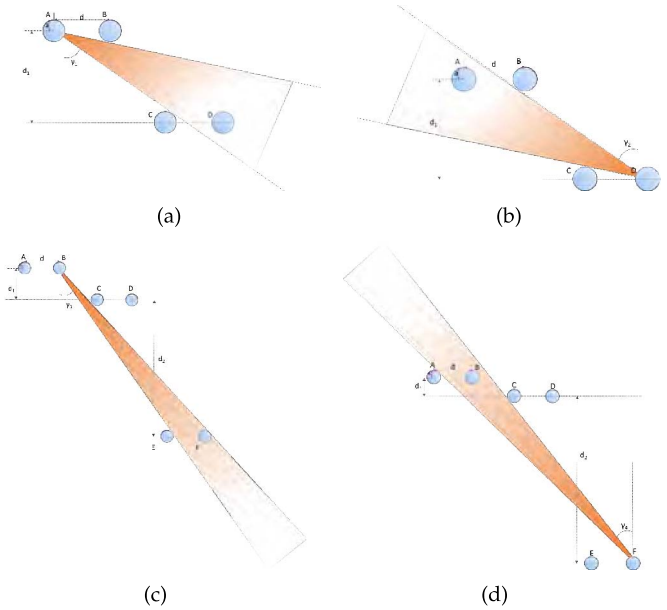


Fig. 12. Free Connection Conditions of 3-Layer Graphite. (a) 1st to 2nd layer. (b) 2nd to 1st layer. (c) 1st to 3rd layer. (d) 3rd to 1st layer.

neighbor horns has the risk of violating the free connection condition. The situation of two layers which are not adjacent can be illustrated in a similar way.

In order to get relatively concise formulas, in the mesh we denote the horizonal distances between two adjacent horns in the same row or column are both d. We set the angle between the margin of the signal cone and the vertical direction as $\gamma \in (0, \frac{\pi}{2})$ which is variable as the radio can rotate around. And the spread-out angle, $i.e.$ the apex angle of the signal cone, is denoted as α. Then we can derive the following results about the vertical distances between layers in 3-Layer Graphite.

Figure 12 shows the constraints of the vertical distance between 1th and 2th layer, d_1, and the vertical distance between 2th and 3th layer, d_2, intuitionisticly. We will show the constraints in a mathematic way. Note that the angle between the margin of the signal cone and the vertical direction is denoted as γ_1, γ_2, γ_3 and γ_4 respectively, for the sake of distinguishing different scenarios in Figure 12.

Firstly, it is obvious that the condition of $1th\ to\ 2th\ layer$ and $2th\ to\ 1th\ layer$ are the same essentially. Then we can formulate the constraints shown in Figure 12(a) and (b), denoted as C_1, below with the denotations given above.

$$\exists \gamma_1 \in (0, \frac{\pi}{2} - \alpha)$$

$$\text{s.t. } d_1 \geq \frac{2d\cos\gamma_1 + a}{\sin\gamma_1} \tag{2}$$

$$\frac{a}{d} \leq \sin(\frac{\pi}{2} - \gamma_1 - \alpha) \tag{3}$$

$$d_1 \leq \frac{3d\cos\gamma_1 + a}{\sin\gamma_1} \tag{4}$$

$$d_1 \geq \frac{3d\cos(\gamma_1 + \alpha) - a}{\sin(\gamma_1 + \alpha)} \tag{5}$$

where

(1) ensures that radio C would not be influenced when radio A connects to D as a transmitter;

(2) ensures that radio B would not be influenced when radio A connects to D as a transmitter;

(3) and (4) ensure that radio D would not be blocked by radio B or C which means that the intersection of the spherical space of radio D and the space of the signal cone is not empty.

All of these above can be got directly using some basic knowledge of space geometry, so we would not show the details.

Similarly, we can formulate the constraints shown in Figure 12(c), denoted as C_3, and Figure 12(d), denoted as C_4.

$$\exists \gamma_3 \in (0, \frac{\pi}{2} - \alpha)$$

$$\text{s.t. } d_1 + d_2 \geq \frac{3d\cos\gamma_3 + a}{\sin\gamma_3} \tag{6}$$

$$d_1 \leq \frac{d\cos(\gamma_3 + \alpha) - a}{\sin(\gamma_3 + \alpha)} \tag{7}$$

$$d_1 + d_2 \leq \frac{4d\cos\gamma_3 + a}{\sin\gamma_3} \tag{8}$$

$$d_1 + d_2 \geq \frac{4d\cos(\gamma_3 + \alpha) - a}{\sin(\gamma_3 + \alpha)} \tag{9}$$

$$\exists \gamma_4 \in (0, \frac{\pi}{2} - \alpha)$$

$$\text{s.t. } d_2 \geq \frac{3d\cos\gamma_4 + a}{\sin\gamma_4} \tag{10}$$

$$d_1 + d_2 \leq \frac{5d\cos(\gamma_4 + \alpha) - a}{\sin(\gamma_4 + \alpha)} \tag{11}$$

$$d_1 + d_2 \leq \frac{4d\cos\gamma_4 + a}{\sin\gamma_4} \tag{12}$$

$$d_1 + d_2 \geq \frac{4d\cos(\gamma_4 + \alpha) - a}{\sin(\gamma_4 + \alpha)} \tag{13}$$

Then we can figure out that in a feasible topology design of 3-Layer Graphite, d_1 and d_2 should satisfy all of the above constraints. As it is extremely difficult to get a mathematical formula about d_1 and d_2 directly, we will give a relatively simple method to solve this problem. In practice, d, a and α are certain constants, so in the following analysis we take $d = 1m$, $a = 0.1m$ and $\alpha = 0.02\pi$ as an example which are reasonable according to our testbed.

We use some mathematical analysis tools to help us get the properties of these constrains as shown in Figure 13. We will show how to solve C_4 as an example. The dashed area in Figure 13(c) is the area satisfying C_4. Set boundary points A_4 and B_4 as $A_4(\gamma_{A_4}, y_{A_4})$ and $B(\gamma_{B_4}, y_{B_4})$ respectively. Then, we can get that constrain C_4 is equivalent to the following constraints

$$\frac{3d\cos\gamma_{B_4} + a}{\sin\gamma_{B_4}} \leq d_2$$

$$d_1 + d_2 \leq \frac{3d\cos\gamma_{A_4} + a}{\sin\gamma_{A_4}}$$

where

$$\frac{3d\cos\gamma_s + a}{\sin\gamma_s} = \frac{5d\cos(\gamma_s + \alpha) - a}{\sin(\gamma_s + \alpha)}, s \in \{A_4, B_4\}, \gamma_{A_4} \leq \gamma_{B_4}$$

Similarly, we can get that C_1 is equivalent to:

$$\frac{2d\cos\gamma_{B_1} + a}{\sin\gamma_{B_1}} \leq d_1$$

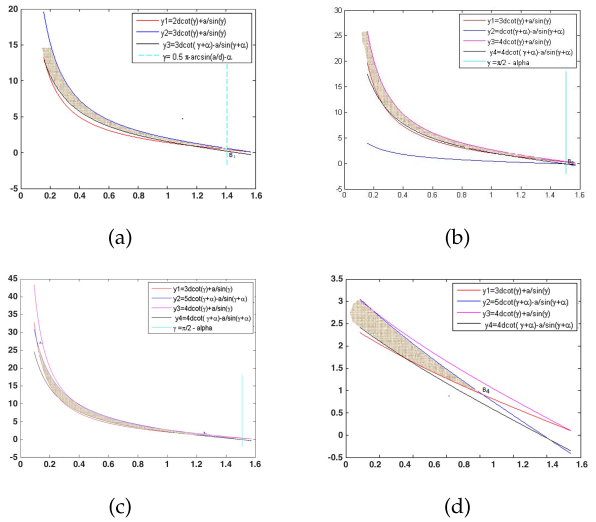


Fig. 13. Mathematical analysis of constrains. (a) Analysis of C_1. (b) Analysis of C_3. (c) Analysis of C_4. (d) Part of (c).

where

$$\gamma_{B_1} = 0.5\pi - \arcsin(a/d) - \alpha$$

C_3 is equivalent to:

$$d_1 + d_2 \geq \frac{3d\cos\gamma_{B_3} + a}{\sin\gamma_{B_3}}$$

where

$$\gamma_{B_3} = \frac{\pi}{2} - \alpha.$$

Finally, we can get that in this feasible topology design of 3-Layer Graphite, d_1 and d_2 should satisfy the following constraints:

$$\begin{aligned}
\frac{2d\cos\gamma_{B_1} + a}{\sin\gamma_{B_1}} &\leq d_1 \\
\frac{3d\cos\gamma_{B_4} + a}{\sin\gamma_{B_4}} &\leq d_2 \\
\frac{3d\cos\gamma_{B_3} + a}{\sin\gamma_{B_3}} \leq d_1 + d_2 &\leq \frac{3d\cos\gamma_{A_4} + a}{\sin\gamma_{A_4}}
\end{aligned}$$

where

$$\begin{aligned}
\gamma_{B_1} &= 0.5\pi - \arcsin(a/d) - \alpha \\
\gamma_{B_3} &= \frac{\pi}{2} - \alpha \\
\frac{3d\cos\gamma_s + a}{\sin\gamma_s} &= \frac{5d\cos(\gamma_s + \alpha) - a}{\sin(\gamma_s + \alpha)}, \\
s &\in \{A_4, B_4\}, \gamma_{A_4} \leq \gamma_{B_4}
\end{aligned}$$

The vertical distances are completely and jointly decided by the size of the radio fronted, the distances between adjacent rows and columns in the mesh and the aperture of the interference-cone, after the number of layers is definitized by the size of the mesh unit in an $R \times R$ square. As the calculation of the vertical distances between the layers is rather complex, we take the vertical distances of 3-Layer Graphite as an example. Some numeric results are shown in Table I, where

TABLE I
VERTICAL DISTANCES BETWEEN ADJACENT LAYERS

d / a	1	2
1	$d_1 = 0.2771, d_2 = 1.2117$	$d_1 = 0.4023, d_2 = 1.63$
2	$d_1 = 0.4303, d_2 = 1.5737$	$d_1 = 0.5542, d_2 = 1.9625$

TABLE II
NODE DEGREE AND COVERAGE RATIO IN 2-LAYER GRAPHITE

2-L	4	5	6	7
4	28, 100%	34, 94.12%	44, 95.45%	54, 92.59%
5		48, 91.67%	62, 93.55%	72, 91.67%
6			80, 88.75%	90, 87.78%
7				112, 86.61%

TABLE III
NODE DEGREE AND COVERAGE RATIO IN 3-LAYER GRAPHITE

3-L	4	5	6	7
4	28, 100%	34, 100%	44, 100%	52, 96.23%
5		48, 100%	62, 100%	70, 97.22%
6			78, 97.5%	86, 95.56%
7				104, 92.86%

- a and d represent the radius of the radio model and the distance of adjacent rows or adjacent columns respectively;
- d_1 and d_2 represent the vertical distance between 1st and 2nd layer, 2nd layer and 3rd layer respectively;
- the omitted specific unit is *meter*.

F. Node Degree and Coverage Ratio

To make the topology's features be understood in a more intuitionistic way, two tables Table II and Table III are given. We put 4×4 to 7×7 racks in an area of 10×10 m^2, and see our topology's node degree and coverage ratio in 2-Layer Graphite and 3-Layer Graphite respectively.

VI. NUMERICAL RESULTS

In this section, we compare the performance of Graphite with state-of-the-art topologies for wireless data center networks, including Flyways [25] and 3D Beamforming [49], using simulations.

In the evaluation, we set the communication range of the radios to be 10 m, and vary the scale of the wireless data center network from 20×20 to 40×40. To evaluate the performance of the topology designs with different densities of the racks, we vary the distance between two adjacent racks in a row/column from 1 m to 4 m with a step of 0.5 m. The distance between top of the racks and the ceiling is set to 4 m.

A. Impact on Node Degree

The degree of a node in a topology of wireless data center network captures the node's opportunities to connect to the

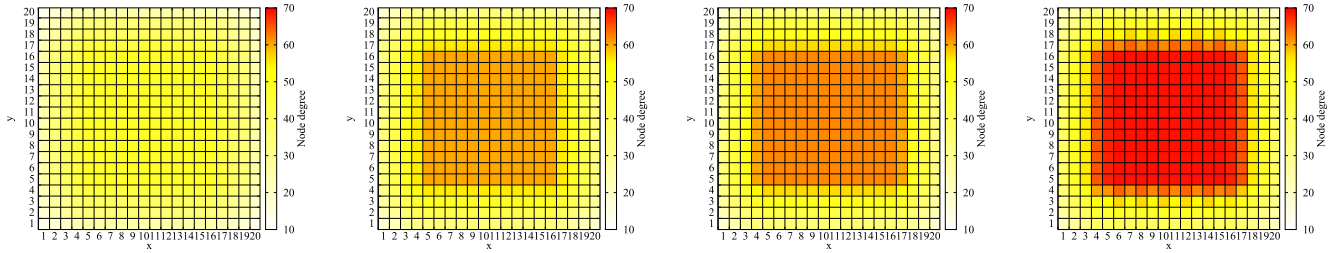


Fig. 14. Heat maps for node degree distribution of a 20×20 mesh achieved by (a) Flyways, (b) 3D Beamforming, (c) 2-Layer Graphite, and (d) 3-Layer Graphite.

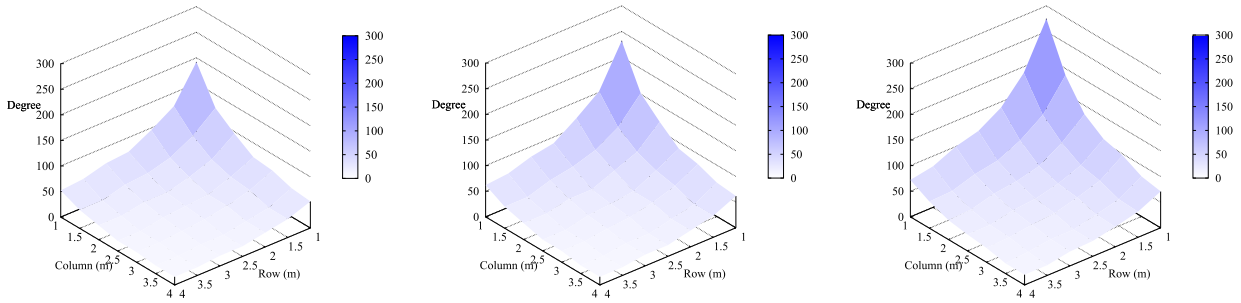


Fig. 15. Average node degrees without boundary effect achieved by (a) Flyways, (b) 3D Beamforming, and (c) Graphite, when the distance between adjacent racks in a row/column varies.

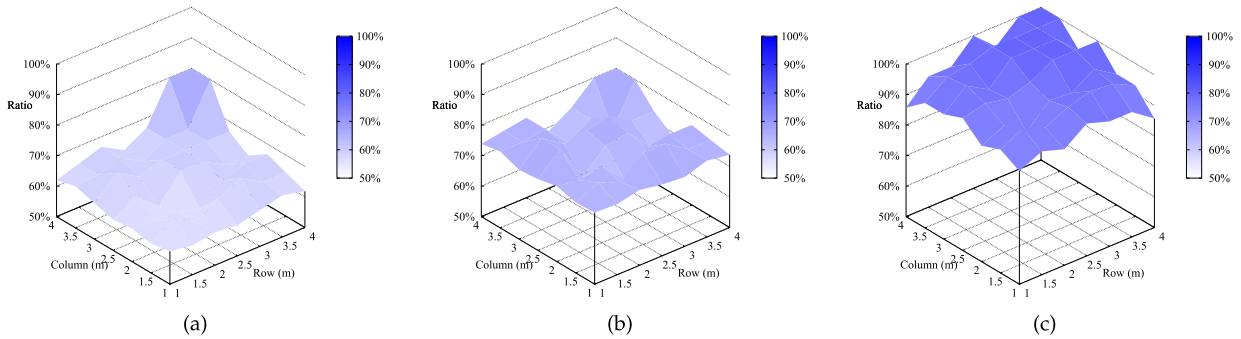


Fig. 16. Coverage ratio without boundary effect achieved by (a) Flyways, (b) 3D Beamforming, and (c) Graphite, when the distance between adjacent racks in a row/column varies.

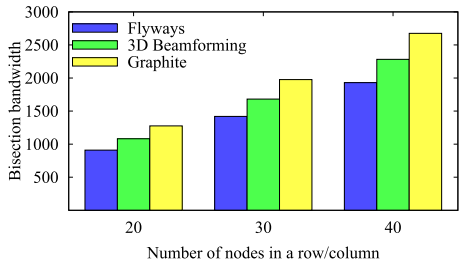


Fig. 17. Bisection bandwidths of Flyways, 3D Beamforming, and Graphite in square mesh topologies.

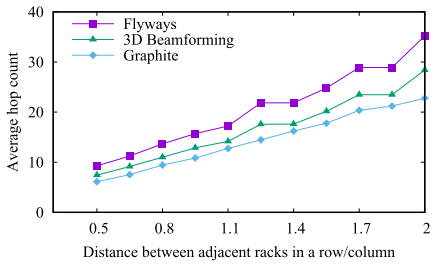


Fig. 18. Average hop count of Flyways, 3D Beamforming, and Graphite in square mesh topologies.

other nodes within its communication range. A higher degree indicates that the node has a higher flexibility to form data flow pathes with the other nodes in the network, and thus may improve the performance of the whole network.

Figure 14 shows the heat maps for node degree distribution of a 20×20 mesh achieved by Flyways, 3D Beamforming, 2-Layer Graphite, and 3-Layer Graphite. In this evaluation, the distance between two adjacent racks in a row/column is set to 2 meters. Then, the largest rectangular unit (allowing communication between any pair of nodes within the same row/column) within the range of a $10\text{m} \times 10\text{m}$ square area is a 5×5 mesh of radios. According to Algorithm 1, the number of layers needed to guarantee non-blocking communication in a row/column is 3. Here, to demonstrate the adaptability of

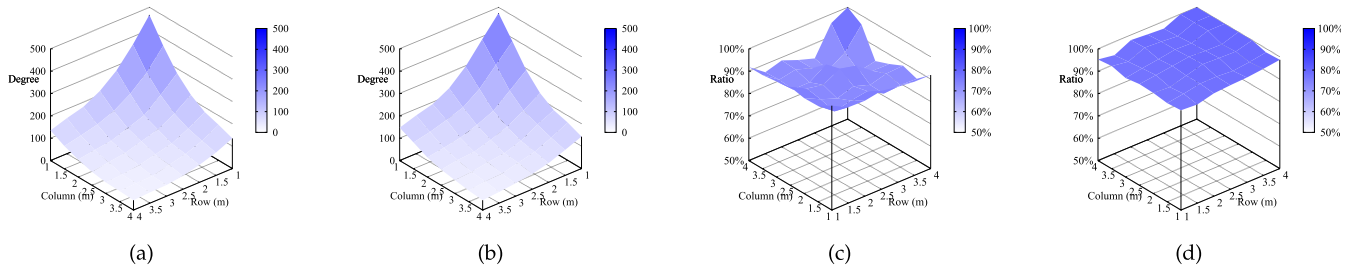


Fig. 19. Comparison results when the communication range is $15m$. (a) Average node degree achieved by 3D Beamforming. (b) Average node degree achieved by Graphite. (c) Coverage ratio achieved by 3D Beamforming. (d) Coverage ratio achieved by Graphite.

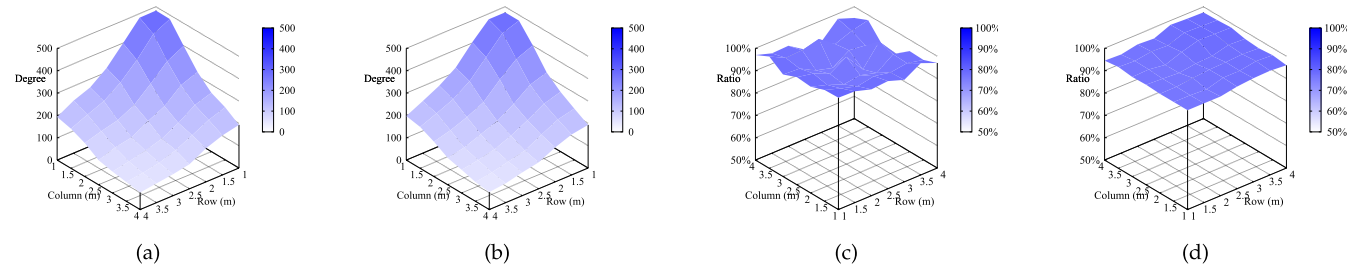


Fig. 20. Comparison results when the communication range is $20m$. (a) Average node degree achieved by 3D Beamforming. (b) Average node degree achieved by Graphite. (c) Coverage ratio achieved by 3D Beamforming. (d) Coverage ratio achieved by Graphite.

Graphite with insufficient number of layers, we also show the results for a topology of 2-Layer Graphite. Although node degree increases from the border to the heart for all the four evaluated topologies, Graphite achieves much higher node degree when approaching the central part of the mesh, which indicates that Graphite can provide more connection opportunities for the nodes, especially the inner nodes, in the wireless data center network.

Noting that boundary nodes have smaller degrees than interior nodes due to the boundary effect, we further evaluate average node degrees of interior nodes for Flyways, 3D Beamforming, and Graphite. Figure 15 shows the evaluation results on average node degree without boundary effect achieved by Flyways, 3D Beamforming, and Graphite, when the distance between adjacent racks in a row/column varies. We can observe that the average node degree without boundary effect of Graphite is always higher than those of Flyways and 3D Beamforming by 42.5% and 15%, respectively.

B. Impact on Coverage Ratio

Given the high directionality of 60 GHz radio signals, two nodes within the communication range may not be able to talk to each other due to the blockage of the intermediate obstacles. So we define coverage ratio to capture the percentage of nodes that a node can directly connect to within its communication range. Coverage ratio reflects how well a topology can avoid link blockage by its radios allocated in a wireless data center network.

Figure 16 demonstrates the coverage ratio without boundary effect achieved by Flyways, 3D Beamforming, and Graphite for varying distance between adjacent racks in a row/column. Flyways can only achieve a coverage ratio of up to 80.0% in relatively sparse network topologies, while most of its coverage ratios are no more than 68.6%. 3D Beamforming performs better than Flyways, especially in relatively dense network topologies, but its coverage ratio is still limited to 80.0% due to need of reflection with the ceiling. Overall, Graphite archives the best coverage ratio in all the cases, and can even reach 100% coverage in relatively sparse network topology.

C. Impact on Bisection Bandwidth

We define the bisection bandwidth of a topology for wireless data center network as the number of possible wireless links between two equally split parts of the network.

In this evaluation, we consider square meshes of wireless data center networks with 20×20, 30×30, and 40×40 racks. The distance between two adjacent racks in a row/column is set to 2 meters.

Figure 17 shows bisection bandwidths achieved by Flyways, 3D Beamforming, and Graphite in square mesh topologies. The results demonstrate that Graphite always outperform the other two approaches in terms of bisection bandwidth.

D. Impact on Average Hop Count

In computer networking, store and forward and other latencies are incurred trough each hop, so the flow completion time is mainly influenced by the hop count. We simulate a simple scenario in a data center consisting 10000 (100×100) servers to give the full picture in terms of how the design performs in practice and how the gain in node degree translates into shorter job completion time.

Considering that wireless data center network is exploited to replace or supplement the underlying wired topology to deal with the problems caused by the extremely large data center requirement nowadays, we mainly focus on the hop counts of source-destination pairs with long distances. In such a 100×100 wireless network data center, we choose the servers

in the first row and the first column as sources, and the servers in the last row and the last column as destinations, respectively. That is, the hop count of every source-destination pair in such two groups will be calculated to get the average hop count of them. The propagation range of the 60 GHz signal is set as $10m$ and the ceiling height is $4m$. We vary the distance between adjacent racks in a row/column from $0.5m$ to $2m$. The results are shown in Figure 18.

We can get that Graphite can over perform 3DB and Flyways as high as 25% and 55% respectively. Generally, in most of the cases, the ratio are 20% and 45% respectively.

E. Results for Other Communication Ranges of the Radios

Although we have analysed the performance of Graphite and compared Graphite with the other two state-of-art designs, *i.e.*, Flyways and 3D Beamforming, in theory in Section V where the communication range of the radios is set as R for generality and in simulation in Section VI-A, VI-B, VI-C where the communication range of the radios is set as $10m$, we show the numerical results for more different communication ranges in this section to make this issue clearer.

As 3D Beamforming always outperforms Flyways and the heat map shows little information, to make the comparison more concise, we mainly give the results of the average node degree without boundary effect and the coverage ratio without boundary effect achieved by 3D Beamforming and Graphite when the communication ranges are set as $15m$ and $20m$ respectively. Note that, all the simulation settings keep the same expect the communication range.

The results are shown in Figure 19 and Figure 20 below. We can get similar conclusions to those when the communication range is $10m$. That is, Graphite still outperforms 3D Beamforming when the communication range varies.

VII. Conclusion

In this paper, we have implemented a liftable 60 GHz 3D beamforming wireless radio frontend and proposed a new topology of wireless data center network, namely *Graphite*, to replace or supplement the current underlying wired topology in data centers, such that the connectivity, throughput, and transmission quality of the data center network can be potentially improved. By carefully adjusting our radio frontend to layers with different heights and into suitable directions, Graphite can achieve direct line-of-sight communication between a good number of rack pairs in a data center with different deployments. Our testbed measurements have confirmed that links in Graphite will not be obstructed or interfered when transmitters and receivers are carefully placed at different layers. We have also provided theoretical analyses and simulations to evaluate the performance of Graphite. Both of the results validate the superior connectivity feature of Graphite, compared with existing topologies like Flyways and 3D Beamforming.

Acknowledgement

The authors thank Y. Li and Z. Chen for helping to design of liftable 3D beamforming radio frontend, and X. Zhang for some useful data about the radio signal. The opinions, findings, conclusions, and recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the funding agencies or the government.

References

[1] *IEEE Standard 802.11ad-2012*, Dec. 2012, pp. 1–628.
[2] D. Abts, M. Marty, P. Wells, P. Klausler, and H. Liu, "Energy proportional datacenter networks," *ACM SIGARCH Comput. Archit. News*, vol. 38, no. 3, pp. 338–347, 2010.
[3] H. Abu-Libdeh, P. Costa, A. Rowstron, G. O'Shea, and A. Donnelly, "Symbiotic routing in future data centers," *ACM SIGCOMM*, vol. 40, no. 4, pp. 51–62, 2010.
[4] M. Al-Fares, A. Loukissas, and A. Vahdat, "A scalable, commodity data center network architecture," *ACM SIGCOMM*, vol. 38, no. 4, pp. 63–74, 2008.
[5] M. Al-Fares, S. Radhakrishnan, B. Raghavan, N. Huang, and A. Vahdat, "Hedera: Dynamic flow scheduling for data center networks," in *Proc. USENIX NSDI*, 2010, p. 19.
[6] K. Bilal *et al.*, "A comparative study of data center network architectures," in *Proc. ECMS*, 2012, pp. 526–532.
[7] Y. Cui, H. Wang, and X. Cheng, "Channel allocation in wireless data center networks," in *Proc. IEEE INFOCOM*, Apr. 2011, pp. 1395–1403.
[8] Y. Cui, H. Wang, X. Cheng, and B. Chen, "Wireless data center networking," *IEEE Wireless Commun.*, vol. 18, no. 6, pp. 46–53, Jun. 2011.
[9] Y. Cui, H. Wang, X. Cheng, D. Li, and A. Ylä-Jääski, "Dynamic scheduling for wireless data center networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 24, no. 12, pp. 2365–2374, Dec. 2013.
[10] Y. Cui, S. Xiao, C. Liao, I. Stojmenovic, and M. Li, "Data centers as software defined networks: Traffic redundancy elimination with wireless cards at routers," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 12, pp. 2265–2672, Dec. 2013.
[11] Y. Cui *et al.*, "Diamond: Nesting the data center network with wireless rings in 3D space," in *Proc. 13th USENIX Symp. Netw. Syst. Design Implement. (NSDI)*, 2016, pp. 657–669.
[12] N. Farrington *et al.*, "Helios: A hybrid electrical/optical switch architecture for modular data centers," *ACM SIGCOMM*, vol. 40, no. 4, pp. 339–350, 2011.
[13] *FLIR*. Accessed: 2017. [Online]. Available: http://www.flir.com/mcs/products/
[14] A. Greenberg *et al.*, "VL2: A scalable and flexible data center network," in *ACM SIGCOMM*, vol. 39, no. 4, pp. 51–62, 2009.
[15] A. Greenberg, P. Lahiri, D. Maltz, P. Patel, and S. Sengupta, "Towards a next generation data center architecture: Scalability and commoditization," in *Proc. ACM PRESTO*, 2008, pp. 57–62.
[16] C. Guo *et al.*, "BCube: A high performance, server-centric network architecture for modular data centers," *ACM SIGCOMM*, vol. 39, no. 4, pp. 63–74, 2009.
[17] C. Guo *et al.*, "DCell: A scalable and fault-tolerant network structure for data centers," in *Proc. ACM SIGCOMM*, 2008, pp. 75–86.
[18] D. Guo *et al.*, "Expandable and cost-effective network structures for data centers using dual-port servers," *IEEE Trans. Comput.*, vol. 62, no. 7, pp. 1303–1317, Jul. 2013.
[19] D. Halperin, S. Kandula, J. Padhye, P. Bahl, and D. Wetherall, "Augmenting data center networks with multi-gigabit wireless links," *ACM SIGCOMM*, vol. 41, no. 4, pp. 38–49, 2011.
[20] N. Hamedazimi *et al.*, "FireFly: A reconfigurable wireless datacenter fabric using free-space optics," *ACM SIGCOMM*, vol. 44, no. 4, pp. 319–330, 2014.
[21] A. Hammadi and L. Mhamdi, "A survey on architectures and energy efficiency in data center networks," *Comput. Commun.*, vol. 40, pp. 1–21, Mar. 2014.
[22] B. Heller *et al.*, "ElasticTree: Saving energy in data center networks," in *Proc. USENIX NSDI*, 2010, pp. 249–264.
[23] H. Huang *et al.*, "The architecture and traffic management of wireless collaborated hybrid data center network," *ACM SIGCOMM*, vol. 43, no. 4, pp. 511–512, 2013.
[24] C. Kachris and I. Tomkos, "A survey on optical interconnects for data centers," *IEEE Commun. Surveys Tuts.*, vol. 14, no. 4, pp. 1021–1036, 4th Quart., 2012.
[25] S. Kandula, J. Padhye, and P. Bahl, "Flyways to de-congest data center networks," in *Proc. ACM HotNets*, 2009, pp. 1–6.
[26] Y. Katayama, K. Takano, Y. Kohda, N. Ohba, and D. Nakano, "Wireless data center networking with steered-beam mmwave links," in *Proc. IEEE WCNC*, Mar. 2011, pp. 2179–2184.

[27] Y. Katayama *et al.*, "MIMO link design strategy for wireless data center applications," in *Proc. IEEE WCNC*, Apr. 2012, pp. 3302–3306.
[28] D. Li *et al.*, "FiConn: Using backup port for server interconnection in data centers," in *Proc. IEEE INFOCOM*, Apr 2009, pp. 2276–2285.
[29] D. Li and J. Wu, "On the design and analysis of data center network architectures for interconnecting dual-port servers," in *Proc. IEEE INFOCOM*, Apr. 2014, pp. 1851–1859.
[30] H. Liu *et al.*, "Circuit switching under the radar with REACToR," in *Proc. USENIX NSDI*, 2014, pp. 1–15.
[31] Y. Liu, X. Gao, B. Wong, and S. Keshav, "Quartz: A new design element for low-latency dcns," *ACM SIGCOMM*, vol. 44, no. 4, pp. 283–294, 2014.
[32] *Hxi LCC*. Accessed: 2017. [Online]. Available: http://www.hxi.com
[33] R. Mysore *et al.*, "PortLand: A scalable fault-tolerant layer 2 data center network fabric," *ACM SIGCOMM*, vol. 39, no. 4, pp. 39–50, 2009.
[34] L. Popa, S. Ratnasamy, G. Iannaccone, A. Krishnamurthy, and I. Stoica, "A cost comparison of datacenter network architectures," in *Proc. ACM CoNEXT*, 2010, p. 16.
[35] K. Ramachandran, R. Kokku, R. Mahindra, and S. Rangarajan, "60 GHz data-center networking: Wireless⇒worryless?" NEC Lab. Amer. Inc., Princeton, NJ, USA, Tech. Rep., 2008.
[36] J. Shin, E. Sirer, H. Weatherspoon, and D. Kirovski, "On the feasibility of completely wireless datacenters," *IEEE/ACM Trans. Netw.*, vol. 21, no. 5, pp. 1666–1679, Oct. 2012.
[37] *SiBeam*. Accessed: 2017. [Online]. Available: http://www.sibeam.com/
[38] A. Singla, A. Singh, K. Ramachandran, L. Xu, and Y. Zhang, "Proteus: A topology malleable data center network," in *Proc. ACM SIGCOMM Workshop Hot Topics Netw.*, 2010, p. 8.
[39] *S. Store*. Accessed: 2017. [Online]. Available: http://www.semiconductorstore.com/
[40] Y. Sun, J. Chen, Q. Liu, and W. Fang, "Diamond: An improved fat-tree architecture for large-scale data centers," *J. Commun.*, vol. 9, no. 1, pp. 91–98, 2014.
[41] A. Valdes-Garcia *et al.*, "Single-element and phased-array transceiver chipsets for 60-GHz Gb/s communication," *IEEE Commun. Mag.*, vol. 49, no. 4, pp. 120–131, Apr. 2011.
[42] H. Vardhan, N. Thomas, S. Ryu, B. Banerjee, and R. Prakash, "Wireless data center with millimeter wave network," in *Proc. IEEE GLOBECOM*, Dec. 2010, pp. 1–6.
[43] M. Walraed-Sullivan, A. Vahdat, and K. Marzullo, "Aspen trees: Balancing data center fault tolerance, scalability and cost," in *Proc. ACM CoNEXT*, 2013, pp. 85–96.
[44] G. Wang *et al.*, "c-Through: Part-time optics in data centers," *ACM SIGCOMM*, vol. 40, no. 4, pp. 327–338, 2010.
[45] *WirelessHD*. Accessed: 2017. http://www.wirelesshd.org/
[46] H. Wu, G. Lu, D. Li, C. Guo, and Y. Zhang, "MDCube: A high performance network structure for modular data center interconnection," in *Proc. ACM CoNEXT*, 2009, pp. 25–36.
[47] X. Ye *et al.*, "DOS: A scalable optical switch for datacenters," in *Proc. ACM/IEEE ANCS*, Oct. 2010, p. 24.
[48] W. Zhang, X. Zhou, L. Yang, Z. Zhang, B. Y. Zhao, and H. Zheng, "3D beamforming for wireless data centers," in *Proc. ACM HotNets*, 2011, p. 4.
[49] X. Zhou *et al.*, "Mirror mirror on the ceiling: Flexible wireless links for data centers," *ACM SIGCOMM*, vol. 42, no. 4, pp. 443–454, 2012.
[50] Y. Zhu *et al.*, "Cutting the cord: A robust wireless facilities network for data centers," in *Proc. ACM MobiCom*, 2014, pp. 581–592.

Chaoli Zhang received the B.Sc. degree in information security and bachelor of laws from Nankai University, China, in 2015. She is currently pursuing the degree with the Department of Computer Science and Engineering, Shanghai Jiao Tong University, China. Her research interests encompass wireless networks, pricing model, and crowdsourcing.

Fan Wu received the B.S. degree in computer science from Nanjing University in 2004, and the Ph.D. degree in computer science and engineering from The State University of New York at Buffalo in 2009. He has visited the University of Illinois at Urbana–Champaign as a Post-Doctoral Research Associate. He is currently a Professor with the Department of Computer Science and Engineering, Shanghai Jiao Tong University. He has authored over 150 peer-reviewed papers in technical journals and conference proceedings. His research interests include wireless networking and mobile computing, algorithmic game theory and its applications, and privacy preservation. He was a recipient of the First Class Prize for the Natural Science Award of the China Ministry of Education, the NSFC Excellent Young Scholars Program, the ACM China Rising Star Award, the CCF-Tencent Rhinoceros Bird Outstanding Award, the CCF-Intel Young Faculty Researcher Program Award, and the Pujiang Scholar. He has served as the Chair for CCF YOCSEF Shanghai, on the Editorial Board of *Computer Communications*, and as a member of technical program committees of over 60 academic conferences.

Xiaofeng Gao received the B.S. degree in information and computational science from Nankai University, China, in 2004, the M.S. degree in operations research and control theory from Tsinghua University, China, in 2006, and the Ph.D. degree in computer science from The University of Texas at Dallas, USA, in 2010. She is currently an Associate Professor with the Department of Computer Science and Engineering, Shanghai Jiao Tong University, China. She has authored over 120 peer-reviewed papers and 7 book chapters in the related area, including well-archived international journals, such as the IEEE/ACM TRANSACTIONS ON NETWORKING, the IEEE TRANSACTIONS ON COMPUTERS, the IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS, the IEEE TRANSACTIONS ON MOBILE COMPUTING, the IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, and also in well-known conference proceedings, such as INFOCOM, SIGKDD, and ICDCS. Her research interests include wireless communications, data engineering, and combinatorial optimizations. She has served on the Editorial Board of *Discrete Mathematics, Algorithms and Applications,* and as the PCs and peer-reviewers for a number of international conferences and journals.

Guihai Chen received the B.S. degree from Nanjing University in 1984, the M.E. degree from Southeast University in 1987, and the Ph.D. degree from The University of Hong Kong in 1997. He is currently a Distinguished Professor with Shanghai Jiaotong University, China. He had been invited as a Visiting Professor by many universities, including the Kyushu Institute of Technology, Japan, in 1998, The University of Queensland, Australia, in 2000, and Wayne State University, USA, from 2001 to 2003. He has authored over 200 peer-reviewed papers, and over 120 of them are in well-archived international journals, such as the IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS, the *Journal of Parallel and Distributed Computing, Wireless Networks*, *The Computer Journal*, the *International Journal of Foundations of Computer Science*, and *Performance Evaluation*, and also in well-known conference proceedings, such as HPCA, MOBIHOC, INFOCOM, ICNP, ICPP, IPDPS, and ICDCS. He has a wide range of research interests with focus on sensor networks, peer-to-peer computing, high-performance computer architecture, and combinatorics.