Front. Comput. Sci., 2019, 13(5): 1034–1047
https://doi.org/10.1007/s11704-018-7315-6

RESEARCH ARTICLE

# Design and optimization of VLC based small-world data centers

**Yudong QIN[1], Deke GUO (✉)[1], Lailong LUO[1], Geyao CHENG[1], Zeliu DING[2]**

1 Science and Technology on Information System Engineering Laboratory, National University of Defense Technology, Changsha 410073, China

2 School of Electronic Engineering, Naval University of Engineering, Wuhan 430033, China



**Abstract** The visible light communication (VLC) has the potential to provide dense and fast connectivity at low cost. In this paper, we propose a novel VLC enabled *Wireless Small-World Data Center (WSWDC)*. It employs VLC links to achieve a fully wireless data center network (DCN) across racks for the first time. The using of VLC links eliminates hierarchical switches and inter-rack cables, and thus reducing hardware investment, as well as maintenance cost. More precisely, to simplify the configuration and control operations, we propose three DCN design rationales: (1) fully-wireless, all inter-rack links are wireless; (2) easy-deployable, it is not necessary to change the existing infrastructure inside data center; (3) plug-and-play, no extra centralized control operations are required. Previous proposals, however, cannot achieve the three rationales simultaneously. To this end, we first use regular VLC links to interconnect racks as a regular grid DCN and optimize the rack placement to shorten the average path length and the network diameter. To further exploiting the benefits of VLC links, a few random VLC links are carefully introduced to update the wireless grid DCN as a wireless small-world DCN. To avoid the potential interference among VLC links, we deploy VLC transceivers at different heights on the top of each rack. In this way, VLC links would not interfere with others at each height level. Moreover, we design a greedy but efficient routing method for any pair of racks using their identifiers as inputs. Comprehensive evaluation results indicate that our WSWDC exhibits good topological properties and network performance.

Received September 7, 2017; accepted December 8, 2017

E-mail: guodeke@gmail.com

**Keywords** data center network, visible light communication, small-world, wireless links, topology design

## 1 Introduction

Data centers are dominating infrastructures to support various cloud computing applications [1]. Inside each data center, data center network (DCN) interconnects all of the servers and networking devices together to provide massive computing, networking and storage resources [2, 3]. Recently, many novel DCNs have been proposed to improve the network performance inside each data center. The existing wired DCNs suffer from several inherent challenges. Firstly, large-scale DCNs are typically constructed based on a hierarchical topology using multiple layers of switches, e.g., Fat-Tree [4], VL2 [5], Portland [6], etc. In these designs, those core switches become the bottleneck of network performance, and may lead to the one-point-failure [7]. Secondly, in existing designs, transmissions between servers of different racks must pass through network devices and links at upper layers even though they are very close physically. Thirdly, configuring a large-scale data center network is both labor-intensive and time-consuming [8]. Recently, several flat DCNs with homogeneous high-radix switches are proposed to interconnect a large number of racks as complex topologies, such as the generalized hypercube topology and the random regular graph. Although such flat DCNs have tackled those challenging issues at some extent, they suffer non-trivial difficulties of efficient routing, easy configuration and maintenance.

To solve the above problems, researchers offer diverse wireless networks to amend and compensate the wired networks. Advanced wireless communication techniques, e.g., free-space-optical (FSO) [9], 60 GHz radio frequency communication [10] are used to interconnect racks and append a wireless network to the wired network. In addition, there have been some fully inter-rack wireless data centers, such as Firefly [10] and projecToR [11]. All links inside a rack are wired, but all inter-rack links across racks are wireless. Employing these wireless communication technologies to construct wireless DCNs, however, usually impose infrastructure-level alteration and complex manipulation to existing data centers. In this paper, we note that the emerging visible light communication (VLC) technology are competent to construct an inter-rack wireless DCN since the data rate and transmission range of VLC links can satisfy the requirements of DCN.

VLC is being developed rapidly. It transmits data through light and the absence or presence of light represent “0” or “1”, respectively. The VLC links can achieve 10 Gbps data rate without interference in the communication process [12]. Compared with FSO and 60 GHz communication, VLC has the following advantages. First of all, the transceivers of VLC are lower-cost and less-complicated than the other two technologies, since it uses LEDs to generate signals. Second, VLC has higher safety to human. By contrast, 60 GHz will produce electromagnetic radiation and FSO may hurt the eye and skin since it uses lasers. Third, compared with 60 GHz, VLC has better scalability. 60 GHz experiences slower response and lower data reception rate due to the channel contention among the large number of nodes [13]. Thus we believe that VLC is a promising alterative and can perform well in data centers.

In this paper, we envision the following three design rationales to construct high-performance inter-rack wireless data center networks: (1) fully-wireless, all inter-rack links in the network are wireless; (2) easy-deployable, no additional infrastructure-level augment or adjustment within data center is needed; (3) plug-and-play, no extra centralized control are needed to enable an arbitrary connection in the network. The above rationales will bring profound benefits to data centers. Firstly, by constructing a complete wireless data center, networking devices, e.g., hierarchical switches and inter-rack cables, are no longer needed. That is, the hardware investment can be reduced reasonably. Secondly, since it needs no extra center control operations, the configuration and working process of wireless network can be simplified. Thirdly, it does not have to alter the physical infrastructures inside data center, for example, alter ceiling into mirror. Thus it is easy to install and maintain. Fourthly, the elimination of inter-rack cables improves the air flow for more efficient cooling in data center [14].

However, the existing wireless DCNs mainly focus on the flexible reconfiguration of links, and they do not concern the other two rationales. They need to impose infrastructure-level alteration to data centers. For example, to relay the wireless signals, Firefly and 3D BF [15] need to install reflective ceiling, and ProjecToR has to install a micro-mirror array for each launcher. Though eliminating inter-rack cables, they introduce other complex facilities into data centers. Moreover, they impose frequent and complicated control on wireless devices, such as reconfigure links based on traffic prediction.

In this paper, we propose WSWDC, a VLC enabled wireless small-world data center to achieve the three rationales simultaneously. The basic idea is to firstly interconnect existing racks as a regular grid topology [16], one of the most popular topologies, using VLC links. VLC grid is simple and easy to maintain, but suffers from long routing path in the worst cases. To shorten the length of routing path, a few random VLC links are carefully deployed to update the wireless grid DCN as a wireless small-world DCN. More precisely, for any pair of racks, a random link is added with the probability proportional to the $d$th power of their distance, where $d$ is the dimensionality of a lattice network [17]. Thus, the introduce of a few random VLC links can improve the DCN performance considerably, but cause the difficulty for joint routing using both regular and random VLC links. For this reason, we design a greedy but efficient routing method for any pair of racks based on the characteristics of topology. The major contributions of this paper are summarized as follows:

- We design a novel fully wireless data center network WSWDC, which employs VLC links to construct rack-level wireless DCN for the first time. We first interconnect racks as a regular grid topology via VLC links, and then enhance the network performance by appending a few random links according to small-world distribution. WSWDC is fully-wireless, easy-deployable and plug-and-play.
- To fully exploit the benefits of regular as well as random VLC links, a greedy but efficient routing method are proposed based on the characteristics of topology. For any pair of racks, the routing method can generate a desired path between two racks just according to their identifiers.
- Trace-driven experiments are conducted to evaluate the performance of WSWDC under different traffic pat-

terns. The results indicate that our novel DCN can indeed achieve better network performance as expected.

The reminder of this paper is organized as follows. Section 2 introduces the related works and background. Section 3 reports how to construct a wireless grid topology using VLC links and Section 4 discusses how to deploy a few random links to update the regular grid topology as small-world topology. Section 5 deigns the simple and efficient routing methods for WSWDC. Section 6 evaluates the network performance of WSWDC and Section 7 concludes this paper.

## 2 Related work and background

In this section, related works about wireless DCNs are reviewed. Then we introduce VLC technology and discuss the feasibility of employing the VLC links in data center.

### 2.1 Wireless data center networks

Wired data centers suffer from intrinsic disadvantages, e.g., cabling complexity, dear hardware investment, high maintenance overhead and limited scalability and flexibility. Moreover, cable bundles in wired DCNs may block airflow, reducing cooling efficiency and increasing energy consumption [18].

To solve the above problems, researchers introduce wireless transmission technologies to construct wireless links in data centers. Among them, 60 GHz and FSO are the two most prevalent technologies. Flyway [19] analyzes the feasibility of introducing 60 GHz into data centers for the first time. However, it constructs straight links between rack pairs, which is easy to be blocked. And FSO can also only realize line-of-sight communication. To enable out-of-sight communication, 3D BF alters the ceiling into a mirror, such that the mirror can reflect 60 GHz signals to avoid blocking [9]. Firefly uses the same method to bounce FSO signals [10]. To implement this method, the space above racks need to be complete clearance and the height of ceiling also has strict requirements. However, in most data centers, it is infeasible due to the complex steel structure and air conditioning pipes above racks [20]. Diamond proposes to install the reflector beside racks, and constructs 3D ring reflection spaces for server to server connections. It makes data center enclosed, so that cooling and maintenance can be difficult. By contrast, ProjecToR removes the huge mirror for public use, but assigns micromirror array and mirror assembly for each laser dioxide [11]. However, ProjecToR has really high precision requirement on installation and configuration, and hence is also hard to manage and control.

Indeed, the infrastructure-level alteration and complex control mechanism are the side effects of pursuing flexibility. Most of the existing wireless DCN designs focus on dynamic reconfiguration to alleviate the heavy traffic from the hot spots. Thus they call for precise prediction of the coming traffic patterns, which is certainly time-consuming and costly. To this end, Shin uses 60 GHz technology to construct a static wireless DCN [21]. However, it has poor bisection bandwidth due to severe interference among the 60 GHz links. VLCcube [22] is another static wireless DCN. Its wireless links are fixed and it uses layered arrangement method to avoid blocking without imposing infrastructure-level alterations. Besides, it introduces VLC links into data centers for the first time, and VLC links do not generate interference. However, VLCcube does not eliminate inter-rack cables. It deploys a wired fat-tree and a wireless Torus topology on the same racks together. Inspired by the insight of VLCcube, in this paper, we propose VLC enabled WSWDC and construct a fully inter-rack wireless network with VLC links.

### 2.2 Introduction of visible light communication

Visible light communication (VLC) is a wireless data communication method which uses the visible light between 400 and 800 THz (780-375 nm) as information carrier. It transmits information by generating high-speed optical pulse. VLC turns the light on to represent 1 and off to represent 0 [23]. This on-off keying is simple to generate and decode. In effect, VLC is becoming an alternative choice for next-generation wireless technology by offering low cost, unregulated bandwidth and ubiquitous infrastructures support. This technology is envisioned to be used in a wide range of applications in both indoor and outdoor scenarios. VLC uses light emitting diodes (LEDs), for the dual role of illumination and data transmission. The LED lighting system can achieve lower power consumption and has a longer life-time compared to the fluorescent lamp system. With this leading edge technology, data can be transmitted at high speeds using LED light. Besides, VLC transmits optical signals in free space, we can employ this technology to achieve wireless connection in communication system without using wired transmission media such as optical fiber.

At present, a lot of researchers are working on the development of light-emitting diode (LED) lighting system. The VLC positioning technology and Li-Fi transmission technology have already been commercialized. And it is feasible to introduce VLC links into data centers. The equipment of VLC is cheaper than other wireless communication technology. But in practice, there are few DCNs utilizing VLC links. VLCcube is the first DCN to employ VLC links [22] and it can be seemed as two complete networks constructed on the same set of racks, i.e., the wired Fat-Tree and the wireless Torus. The introduced VLC links can decrease the APL and enhance the network bandwidth. It is a hybrid network structure for data centers, so it does not achieve the fully wireless connection. The usage of light as a source of communication is an innovative and not-yet commercialized technology [24]. And as far as we know, our WSWDC is the first inter-rack wireless DCN based on VLC links.

### 2.3 Feasibility of employing VLC links in DCNs

Here we analyze the feasibility of interconnecting racks using VLC links. Indeed, VLC has many advantages in communication, making it feasible to be introduced into DCNs.

**High data rate** Under the control of microchips, LEDs can complete switching on and off millions of times in a second. Researchers has already achieved the 10 Gbps data rate in 2013. In 2015, this speed has been improved to 50 Gbps under some specific conditions. Hence, it is believed that VLC links are competent to transmit data in a data center. Besides, this kind of flicker can be sensed by photo detectors, but for naked eyes, it is too quick to be caught. So it does no harm to human body.

**No interference** Other wireless communication method, for example, 60 GHz, may generate electromagnetic interference in the communication process. This means the communication signals transmitting towards different directions may interfere with each other. And it also acts as a trouble to some electromagnetic-signal-sensitive devices. By contrast, VLC is free from these problems. There is no radiation and interference among communication signals, which makes it suitable to serve as the intensive links in a data center.

**Sufficient transmission range** At present VLC links are able to transmit data within 2 kilometers, and even high power light emitting diodes are unable to achieve greater distance. But long distance results in limited date rate and vast energy cost. If VLC links are set as that long, the data rate can only reach 500 Mbps [13]. Fortunately, the LED based VLC links can reach 10 Gbps data rate with 10 meters, which are sufficient to connect two close racks in DCNs.

**High security** Visible light cannot penetrate opaque objects. So the information leakage from the severer room need no consideration. But this comes along with a problem. Blockage on transceiver will cause the interrupt of communication. Thus to employ VLC in a data center, light paths should be designed carefully and ensure that they are not blocked by any infrastructure component.

However, VLC has some well-known limitations, such as ambient light influence and line-of-sight requirement. Fortunately, many techniques have been proposed to reduce the background noise, such as Hadamard coding, Manchester coding, turbo coding and so on. Additionally, the ambient light condition inside a data center can be controlled on demand such that the ambient light conditions will not cause serious interference. The line-of-sight requirement can be satisfied by the carefully placement of VLC transceivers to avoid blockage.

Besides, the successful practice of VLCcube is a strong evidence of the feasibility. In VLCcube, VLC links are introduced to interconnect racks into wireless Torus, and cables interconnect racks into wired Fat-tree. The introduction of VLC links does not alter infrastructure inside data center, but improves the performance of data center significantly. However, VLCcube is a hybrid DCN. We try to establish a novel DCN which eliminates all inter-rack cables.

## 3 The design of WSWDC

In this section, we illustrate the construction on equipment level, including the placement of racks and construction of VLC links. Specifically, we build a complete VLC enabled inter-rack wireless grid and install transceivers for constructing random links. The method for constructing random links will be discussed in next section.

### 3.1 Overview of WSWDC

We have proposed three design rationales above, i.e., fully-wireless, easy-deployable and plug-and-play. In WSWDC, all racks are interconnected by VLC links, thus it achieves the fully inter-rack wireless connection. Mirrors and other devices used to change the direction of wireless signals are not needed. All VLC links are straight line without inflection points, and they are installed in different heights to avoid being blocked by transceivers. WSWDC is a static architecture. It needs no centralized control operations to reconfigure its topology.

VLC links in WSWDC can be divided into two categories, i.e., regular links and random links. Regular links interconnect racks into grid topology. Grid has been one of the most popular DCN topologies. In a regular grid topology, each node in the network is connected with two neighbors along each dimension. If the network is one-dimensional, the resulting topology is a chain. Then if the network is two-dimensional, the resulting topology is depicted in Fig. 1, and in this paper we mainly focus on the 2D grid topology.

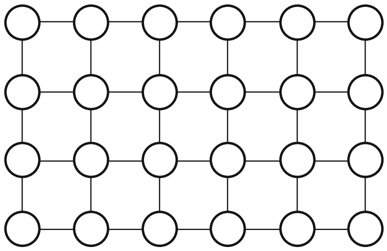

**Fig. 1** An instance of 4×6 2D grid

Random links are introduced to further improve the wireless grid, such that the resulted network diameter can be shorten reasonably. Random links are inspired by small-world phenomenon [25]. Random links are chosen with the probability proportional to the $d$th power of the distance between two racks where $d$ is the dimension numbers of lattice network, thus it can route efficiently [17]. Due to the introduced random VLC links, we finally establish a wireless small-world data center, which is shorten as WSWDC.

### 3.2 Placement of racks

The WSWDC is constructed based on grid, thus racks should be located according to the location of nodes in grid topology. Specifically, there is a one-to-one match between each rack and each node in grid topology. This topology guarantees that all nodes are reachable, and easy to route according to their position. As a data center network, its performance is affected by its average path length (APL).

For an $m \times n$ grid, its APL $d_P$ can be calculated as follows:

$$d_P = \sum_{j=1}^{n}\sum_{i=1}^{m}\frac{1}{m^2n^2}[\frac{j(j-1)m}{2}+\frac{(n-j+1)(n-j)m}{2} + \frac{i(i-1)n}{2}+\frac{(m-i+1)(m-i)n}{2}], = \frac{1}{3}(m+n-\frac{1}{m}-\frac{1}{n}). \quad (1)$$

According to the properties of basic inequalities, it is easy to derive that $d_P$ will achieve minimal value in the case of $m = n$ when $m \times n$ is fixed. Thus given the number of racks, the average path length of an $m \times n$ grid will reach the least if $m$ and $n$ are set as the same value. However, selecting suitable values for $m$ and $n$ can reduce the network diameter, but it is not enough. Note that the average path length of a $n \times n$ grid increases linearly with $n$, leading to poor scalability.

And the grid based DCNs also suffer from other inherent problems. The network diameter is large. Given an $m \times n$ grid, the network diameter is $m + n - 2$. In this case, the two racks locate at the two ends of grid's diagonal line respectively. Compared with Torus of the same size, the network diameter is only $(m + n)/2$. And the average path length of grid is also longer. These problems have a significant impact on performance. So it is necessary to introduce random links to improve its performance.

### 3.3 Regular VLC links

Wireless data centers connect racks without wired links to enable adaptive flexible network topology. Here VLC links are introduced into grid topology to construct regular links. For VLC, transmitting data is achieved by intensity modulation of visible spectrum lighting emitting diodes (LEDs). The presence or absence of light on transceivers represent 1 or 0 respectively. Since we do not intend to introduce any other devices (e.g. mirrors) as many DCNs do, it is unable to change the path of visible light. Thus in WSWDC, each VLC links between racks is straight without bounce. So it need to ensure that all light paths of VLC links will not be blocked by any infrastructure component, for the blockage on transceivers will cause the communication breakdown. Furthermore, it requires that both transceivers on source rack and destination rack can fully face with each other to achieve the best communication effect. If not, the leaked light which is not received by the destination transceiver will lead to the energy loss and even become an interference to other transceivers. So for VLC links, or transceivers, reasonable layout in space is needed.

In WSWDC, the transceivers of regular links and random links are installed hierarchically to avoid mutual blocking. The regular links are only constructed between adjacent racks, while random links often need to be constructed across some racks. Accordingly, transceivers of regular links are installed below the upper surface of racks, and transceivers of random links are installed above racks. For regular links, their transceivers are installed on the side of racks, and each side only install one transceiver. To make data center regular and easy to maintain, transceivers should be placed on the same position of each side. Constructed all regular links, and the racks are connected into wireless grid network. Figure 2(a) illustrates the wireless network of grid. For each transceiver, it only communicates with its partner transceiver on the neighbor rack. The only thing need to be ensured is that the two partner transceivers should be placed on one line, thus the visible light emitted from one transceivers could be received by the other. So there is no need to transform indoor environment or design other complex devices in the installation process.

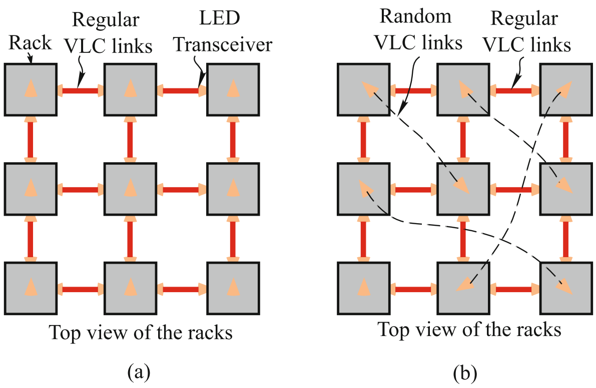


**Fig. 2** The hierarchical layout of transceivers. (a) Wireless grid on top of racks; (b) random links on top of racks

For random links, their transceivers are installed on the top of racks. Any two transceivers can form a random link. Construct all random links randomly, and wireless grid is upgraded into WSWDC, the VLC enabled inter-rack wireless small-world data center. The rule of constructing random links is not discussed in this section, but it still needs to notice that the layout of the random links is more complex than that of regular links, since they can emit towards any direction and the light paths may encounter massive crossover.

## 4 Configuration of random links

In last section, WSWDC has been built by adding random links to wireless grid randomly. However, the complete random method does not make good use of the small-world phenomenon. In this section, the wireless grid DCN is augmented with random links which follow the small-world-inspired distribution. They serve as shortcuts between two nodes and can reduce the network latency. Also, the implement of transceivers is discussed in this section.

### 4.1 Construction of random links

VLC links in small-world data centers are classified into two categories, i.e. random links and regular links. To construct wireless grid structure, the transceivers of regular links are deployed on the side of racks and the room on top of racks is for the deployment of random links. Figure 2(b) gives an example of two random links. The small-world data center topology admits diverse instantiations. For example, the number of random links and degree of each racks can vary depending on the configuration. In this paper, for the purpose of realistic deployment and low-cost, we add the degree of each rack by one. That is to say, only one additional transceiver is introduced on the top of each rack, and the total number of random links is $(m \times n)/2$.

To construct random links is to group the $m \times n$ transceivers into $(m \times n)/2$ pairs. Since any two transceivers can form a random link, it is feasible to group transceivers at fully random, but data centers can perform better by introducing some rules. In this paper, Kleinberg's small-world model [17] is adopted which can promote the efficiency of routing. It requires that random links are chosen with the probability proportional to the $d$th power of the distance (i.e., their shortest path length) between two racks, where $d$ is the number of dimensions of the lattice network. In grid, data can only be transmitted along the regular links, which are either north-south direction or east-west direction. Set the length of all regular links as 1, and the distance of two racks can be calculated as Manhattan distance, i.e., the length between these two racks in the north-south direction plus the length in the east-west direction. Manhattan distance can be expressed in mathematical formula as follows,

$$d_M(x, y) = |x_i - y_i| + |x_j - y_j|, \tag{2}$$

where $x$ and $y$ represent racks, $i$ is row and $j$ is column. Since the grid is a 2-dimension lattice network, thus in realistic deployment, each rack pair is assigned a value which is proportional to the square of their Manhattan distance. The value is normalized to interval (0,1), thus it can represent the probability of their being chosen to interconnect with a random links. Based on these values, all these transceivers can be finally grouped into pairs and form random links.

### 4.2 Spatial layout of random links

Transceivers need to be adjusted to different heights. On one hand, there probably exist some VLC links blocked by transceivers if all transceivers are set to the same height. So these VLC links need to be raised over transceivers which block them. On the other hand, a number of light paths intersect with others. Since the LED light is not as strictly parallel as lasers, and the VLC transceivers have great sensitivity [13], transceivers may receive interference signals from light path which is close to them, and they cannot distinguish these interference signals with their own signals. So we still need to make a hierarchical layout for these random links on top of racks.

To this end, we transform it as the vertex coloring problem. Typically, the vertex coloring problem assigns "colors" to the vertices of a graph so that no edge connects two identically colored vertices. Let $l$ to represent the minimal number of colors to reach this goal. To implement random links based on vertex coloring problem, $interference\ graph\ G$ [20] needs to be constructed first to describe the intersection relations among all random links. Each vertex in $G$ corresponds to a random link, and if two random links intersect, there is a link between them in $G$. Consequently, the layout of random links problem can be reasonably tackled based on the theory of vertex coloring, which has been studied for many years.

The solution to vertex coloring problem derives the minimal $l$. According to the coloring result, random links can be divided into $l$ groups. Random links in the same group are fixed to the same height, thus there are $l$ height planes above racks. Note that an extreme case cannot be solved well where a random link is blocked by transceivers and its transceivers also block other links. Under this case, the coloring result cannot fully reflect the intersection relations.

This method is easy, and has no limitation to the position of transceivers. We can simply install them all on the top center of racks. Thus racks are standardized and easy to reconfigure. However, if the location of transceivers is flexible, we can avoid some intersections by adjusting transceivers' position. In this case, the number of height planes is no more than $l$. But the effect of adjusting fully depends on operator's intelligence. So it is a challenging and interesting work.

### 4.3 Grid with limited range of random links

The method to construct WSWDC based on wireless grid has been discussed above. It can guarantee that random links possess the characteristics of small-world networks. Note that the probability of constructing a random link between two racks is proportional to their Manhattan distance, and that is to say, rack pairs are more likely to be connected by random links if they are far apart. These long random links can reduce the path length significantly, and shorten the network diameter. However, the range of LED based VLC links is limited, and it can keep the data rate at 10 Gbps just within 10 meters. Thus those long random links beyond this range suffer from drastic data rate degradation. By construct, lasers can realize fast long-distance communication (in the order of kilometers) with high data rate, due to its outstanding directionality. Therefore, we can introduce laser transceivers to support these long random links, and use LED transceivers to construct short random links.

There are some cases when random links need to be reconfigured, such as scale expansion of data centers. Note that the introduction of laser transceivers will cause reconfiguration more complexity. It need to reinstall these laser transceivers and even increase its amount when reconfiguration since the layout and amount of long random links may change. Besides, laser transceivers cost more than LED transceivers. So we try to control all random links within their range.

Here restrictions are applied to the length of random links in a simple way and the resulted topology is a variant of small-world networks, in which random links have limited range. To judge whether a random link is beyond range or not, the Euclidean distance between rack pairs is calculated first. If the distance is larger than 10 meters, we assign 0 to their possibility value, instead of being proportional to the square of their Manhattan distance. This can guarantee that they will not be chosen. If the Euclidean distance between adjacent racks is $d_a$, then the two racks connected by random link, such as $x$ and $y$, have to satisfy the following inequality:

$$\frac{10}{d_a} \geqslant \sqrt{|x_i - y_i|^2 + |x_j - y_j|^2}. \tag{3}$$

With this method, we can construct the range-limited WSWDC. Compared with laser-enhanced WSWDC, it has larger APL since its random links are shorter. But for range-limited WSWDC, all racks have the same alteration and that makes convenience for configuration, expansion and management.

## 5 Routing method

In this section, we assign identifier to each rack based on topology of WSWDC to route efficiently. There are two routing strategy in WSWDC. One is using general shortest path algorithms to derive shortest path. The other is a greedy method. It uses identifiers to generate paths without much computation.

### 5.1 Identifier of racks in WSWDC

In a small-world data center, if we simply using MAC addresses or IP addresses to determine a rack, the small-world topology can still provide some benefits. But the using of identifier can bring more efficiency, especially for routing and maintenance.

For each rack, an identifier is assigned to speed up the routing process [26]. Since WSWDC is built based on grid, a two dimensional identifier is sufficient to distinguish every rack in the network. The identifier of a rack located in $i$th row and

$j$th column in WSWDC is assigned as $r_{ij}$. To assign the identifiers reasonably, one of the four racks located on the vertexes of grid should be selected as a landmark. The chosen landmark is denoted as $r_{11}$, which represents that this rack belongs to the first row and the first column. The rest of racks can determine their identifiers autonomously through regular links.

In effect, the identifiers can help a rack get the relative position of any destination rack and generate a path along regular links towards the destination rack easily. This routing method is simple and the time complexity is $O(1)$. But the path generated by this routing method does not contain random links which can serve as shortcuts. Here we have two routing strategies to utilize random links to decrease the number of hops between two racks, 1) searching out the shortest paths with traditional algorithms like Dijkstra and Floyd, or 2) derive a path through a light-weight greedy algorithm. The details are specified in the subsections below.

### 5.2 Routing with shortest path algorithms

Obviously, transmitting packets with the shortest paths consumes least time consumption. In fact, given a pair of source rack and destination rack, general algorithms, e.g., Dijkstra or Floyd algorithms can be employed to deriving the shortest path. In this strategy, each rack needs to precompute the shortest path to all destination racks and maintain a routing table. Computing the shortest paths in advance can reduce congestion and latency, thus increase the bandwidth.

There are a few disadvantages of computing the routing table. First, every rack needs to maintain an routing table. The size of this table is proportional to the size of data center, and each rack needs to store such a table locally. Thus such routing method is neither space-efficient nor scalable. Second, the algorithm to find the shortest path usually has high complexity. For example, the Floyd algorithm has the computing complexity of $O(n^3)$, and the Dijkstra algorithm has the computing complexity of $O(n^2)$. For these reasons, we prefer a greedy yet light-weight alternative to decrease the resulted computation cost.

### 5.3 Routing with greedy algorithms

To achieve both low touting complexity and acceptable path length, we propose a greedy routing method. With this method, the next hop can be determined only based on the identifiers of the current rack and the destination rack. Specifically, assuming that a packet's current location is $r_{ij}$, and the destination rack is $r_{mn}$. There are five racks at most connecting with $r_{ij}$ through VLC links. It needs to compute the Manhattan distance between these racks and $r_{mn}$ respectively. And the neighboring rack with minimal Manhattan distance to $r_{mn}$ will be selected as the next hop for the packet. The regular links guarantee that it can always find a rack with smaller Manhattan distance to $r_{mn}$ than $r_{ij}$. And the random links may shorten the path length significantly in one step.

Compared with the routing based on shortest paths, this method performs better in terms of computing complexity and time complexity. Besides, each rack only needs to store the information of racks directly connected with it. However, greedy routing method may generate longer paths rather than shortest paths, since the shortest paths are searched out based on the global topology. Thus to derive a shorter path with the proposed greedy routing algorithm, we allow $r_{ij}$ to search its neighbors up to $k$ hops away to make greater use of random links. We call this method as greedy_$k$ routing, and describe as follow:

Among all racks at most $k$ hops away from $r_{ij}$, find the rack with the minimal distance which is computed as the Manhattan distance between itself and $r_{mn}$ plus the hops from itself to $r_{ij}$. Then choose the shortest path from $r_{ij}$ to this selected rack.

To implement the greedy_$k$ routing algorithm, each rack has to maintain a routing table to record its neighbors within $k$ hops. Each entry in this table represents one rack, containing its identifier, the shortest path from it to the current rack and the length of this path. Algorithm 1 depicts this strategy. Our experiments show that $k = 3$ can achieve the best balance between the reduced path length with the increased storage.

**Algorithm 1** Greedy_$k$ routing algorithm

**Require:** source rack $r_S$, destination rack $r_D$ and $k$.
**Ensure:** the greedy routing path $P$

**while** $r_S \neq r_D$ **do**
  **for** $i = 1$ to $k$ **do**
    $S_i \leftarrow$ racks $i$ hops away from $r_S$;
    **for** $j = 1$ to $|S_i|$ **do**
      $r_j \leftarrow$ the $j$th rack in $S_i$;
      $d_M \leftarrow$ Manhattan distance between $r_j$ and $r_D$;
      $r_j.d \leftarrow d_M + i$;
    **end for**
  **end for**
  $r_{next} \leftarrow find(r, min\{r.d\})$;
  $P_n \leftarrow$ the shortest path from $r_S$ to $r_{next}$ in routing table;
  $P \leftarrow P \cup P_n$;
  $r_S \leftarrow r_{next}$;
**end while**
**return** $P$;

## 6 Performance evaluation

In this section, we evaluate the performance of our WSWDC. We first compare WSWDC with other wireless DCNs qualitatively. Then we introduce the settings and methodologies to quantitatively compare WSWDC with grid in terms of both topological properties and network performance.

### 6.1 Qualitative comparison

We first compare WSWDC with three wireless data centers qualitatively, i.e., Firefly, 3D BF and VLCcube. Firstly, the three topologies employ the Laser, 60 GHz and VLC to establish the wireless links respectively. Our WSWDC also employs the VLC links, and it achieves the fully wireless connection in inter-rack level while VLCcube still needs wired links to connect racks. Secondly, Firefly and 3D BF are flexible, and the links can be reconfigured according to the underlying traffic demand. By contrast, VLCcube and WSWDC are static architectures, and they maintain stable topologies over racks. In this way, the non-trivial computing cost caused by reconfiguration can be significantly avoided, and the routing complexity can be simplified. Thirdly, both Firefly and 3D BF have to impose infrastructure-level alteration on the existing production data centers, e.g., redecoration of the ceilings and introducing complex devices. However, VLCcube and WSWDC do not need these alterations.

Fourthly, VLCcube and WSWDC are plug-and-play. Once the fabrics are established, no extra control or coordination mechanisms are needed any more. On the contrary, Firefly and 3D BF design their wireless links based on complicated mechanical or electrical control operations. Moreover, Firefly and 3D BF need precise traffic prediction which is also a time-consuming task in dynamic network systems.

Therefore, WSWDC achieves the inter-rack wireless and stable connection with VLC links. It meets the three design rationales proposed in this paper.

### 6.2 Setting and methodology of evaluation

In this section, NS3, a professional network simulator, is employed to compare our WSWDC with grid in terms of throughput and latency. For both WSWDC and grid, placing the racks into a square leads to shorter network diameter than a rectangle. Thus, we generate these two structures with the size of $n \times n$. The bandwidth of each VLC link is set as 10 Gbps. For both networks, according to setting in [27], the link delay is set as 1 ms. With the above basic setting, we compare their topological characteristics, the complexity of routing algorithms, and moreover, the network performance. Besides, in WSWDC, we set the distance between two adjacent racks as 1.5 meters. And we also compare the WSWDC only using VLC links with WSWDC enhanced by lasers.

In our evaluations, we consider two traffic patterns which are introduced in [16, 28]:

**Uniform random traffic**: In this case, source racks select random destination racks with uniform random probability. Thus packets are distributed evenly in the whole network so that no persistent bottlenecks are generated [16]. This pattern is like the shuffle phase of a MapReduce task with uniform distributions [28].

**Local random traffic**: This traffic pattern is generated to evaluate the performance of network under unbalanced workload. Here we generate random but nonuniform flows in network. Specifically, we let the first 12.5 percent of the network (the lowest rows) bear the 25 percent of traffic, and the remaining 75 percent traffic is spread uniformly among the whole network [16].

### 6.3 Topological properties

In this part, we qualify several topological metrics, including the bandwidth, path length and time complexity of different routing methods. We compare these parameters among different topologies and under diverse network scale. The path length is generated by several routing patterns.

#### 6.3.1 Network bandwidth

Bandwidth is the data transferred at one time [29]. Fig. 3(a) compares the bandwidth between WSWDC and grid. Definitely, bandwidth in both networks increases rapidly with the scale of networks. WSWDC offers higher network bandwidth than grid, due to those introduced random VLC links. The bandwidth of WSWDC is approximately 1.26 times on average higher than that of grid, and there exist a slightly decrease in this ratio with the growth of $n$. This ratio varies from 1.28 to 1.26 when $n$ increases from 10 to 50. This is because the growth rate of regular links is higher than that of random links with the increase of $n$.

#### 6.3.2 Average path length

The length of a routing path denotes the number of hops from the source to destination along the routing path [30]. Using general shortest path algorithms can find the shortest path. We determine the shortest path length from each node to all others using Dijkstra algorithm, and calculate the average length of all these paths. Figure 3(b) demonstrates the shortest path length under different network scale, and makes comparisons between WSWDC and the traditional grid topologies. The legend "WSWDC_R" represents the WSWDC whose links have range limitation, and "WSWDC_L" represents the WSWDC enhanced by laser which conquers the limited range of VLC links.

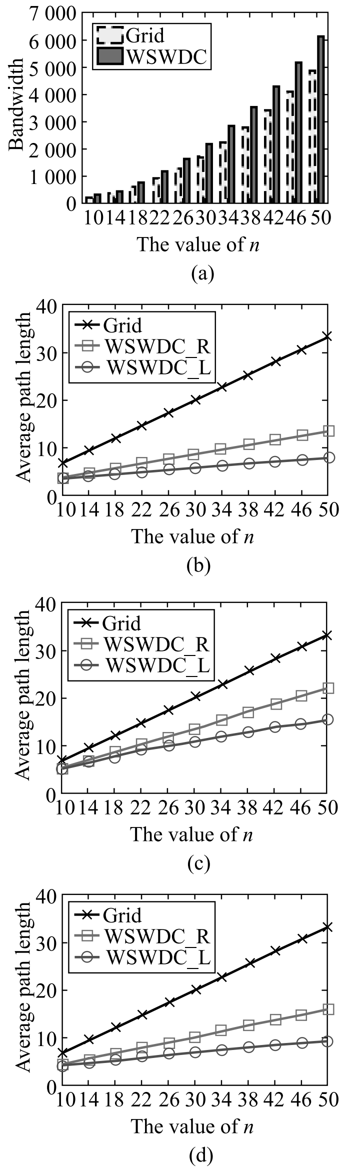


**Fig. 3** The impact of the network size on the performance of three methods under data center networks. (a) Total network bandwidth; (b) dijkstra path length; (c) greedy_1 path length; (d) greedy_3 path length

We can see that the average path length grows with the network size linearly. Under all network scale, the shortest path lengths in WSWDC are much shorter than that of grid, and is about 68% less on average. Among the WSWDC, employing laser achieved shorter path length, and the advantage is amplified with the increase of network size.

Figures 3(c) and 3(d) demonstrate the path length achieved by greedy algorithm. Greedy algorithm is less cost than Dijkstra algorithm, and we use it to generate path between all nodes on each topology. The pattern in which each node is only aware of its direct neighbors is named as greedy_1. Here we also compare the performance with greedy_3 algorithm. In greedy_3 pattern, each node using greedy routing is aware of its three hop neighbors. When make comparisons between WSWDC and grid, we can get the similar conclusion to what we analyzed for shortest path algorithm. Moreover, the greedy_3 routing performs far better than the greedy_1 routing, and it makes a 36% decrease on path length on average.

Since each WSWDC node only uses a small amount of information for greedy routing, the average path length is longer than the Dijkstra path length which is determined using global information. The greedy_1 path length is 43% longer, while the greedy_3 path length is just 11% longer than Dijkstra path length on average.

### 6.3.3 Routing complexity

We further measure the consumed time caused by calculating routing paths in different topologies. As depicted in Fig. 4, the time-consumption of these different routing algorithms shows similar increasing trends. But we can also see that they are not on the same order of magnitude. Greedy_1 routing is the fastest method. It is about 10 times faster than greedy_3 routing. The time consumption of Dijkstra is extremely high, and it is about two orders of magnitude higher than greedy_3 routing. Consider that the gap of path length between them is not big, the optimal routing using Dijkstra algorithm may not be widely accepted.

On the other hand, the greedy routing algorithm running on WSWDC_R is more time-consuming than that on WSWDC_L when $n$ is relatively high. Because this algorithm needs more hops to deliver a packet on average due to its shorter random links. But for Dijkstra algorithm, the two results are similar, since it has approximately the same computation load on these two topologies.

## 6.4 Network performance

In this section, we evaluate the network performance of WSWDC and grid in terms of network throughput and latency. The latency here is depicted as the average finish time of all flows. Under each of the two traffic patterns, we vary the network scale by adjusting the value of $n$, and observe the

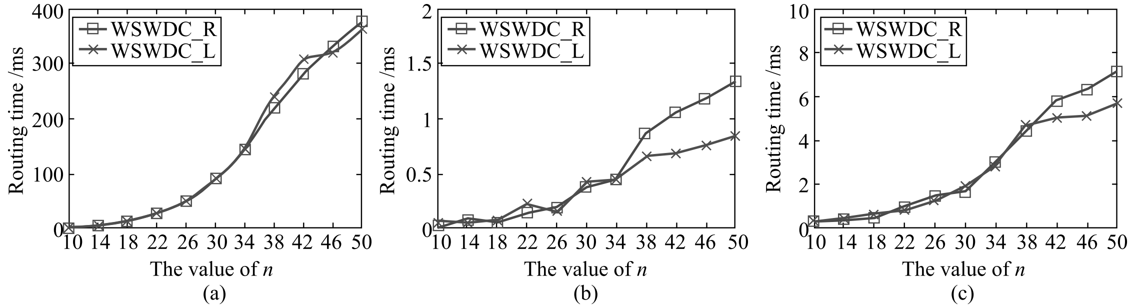


**Fig. 4** Routing complexity. (a) Dijkstra; (b) greedy_1; (c) greedy_3

changing trends of the network throughput and latency. To reveal the impact of flow size on the network performance, we use two kinds of flows, and the average size of each flow are 4 Mb and 12 Mb respectively. In each test, the network throughput is depicted as the ratio of the real throughput between WSWDC and grid in each kind of parameter setting.

### 6.4.1 Network performance under uniform random traffic

In the setting of uniform random traffic pattern, the source and destination server of each flows are all selected randomly. We inject $n^2$ batched random flows into each network topology to evaluate the network performance.

We first fix the maximum flow size as 4 Mb, and adjust the network scale by increasing $n$ from 10 to 30. For all networks the throughput increases significantly. That is because the networks are capable to accommodate more flows with the incensement of $n$. Figure 5 depicts the ratio of the real throughput of WSWDC and grid. We can see that the throughput of WSWDC_R and WSWDC_L are 1.45 and 1.70 times greater on average than that of grid respectively when the flow size is fixed to 4 Mb. Increase the flow size to 12 Mb, the real throughput of the three networks grow dramatically, while the ratio decrease to 1.19 and 1.25 respectively. The two types of WSWDCs outperform grid for all cases due to their shorter path length.

Figure 6 shows the average packet delivery latency. The measured packet delivery latency reflects the trends of the paths lengths. The latency of WSWDC_L is a bit smaller than that of WSWDC_R, while the latency of grid is the largest. When the flow size is 4 Mb, the latencies grow steadily, and gaps among different networks expand with the increase of $n$. The average latency of grid is about 8% and 7% larger than that of WSWDC_L and WSWDC_R respectively. When the flow size grows to 12 Mb, the latency also increases. Besides, the latency and throughput in WSWDC_R is just slightly less than that in WSWDC_L, which means the WSWDC can achieve satisfying effect without introducing lasers under this traffic pattern.

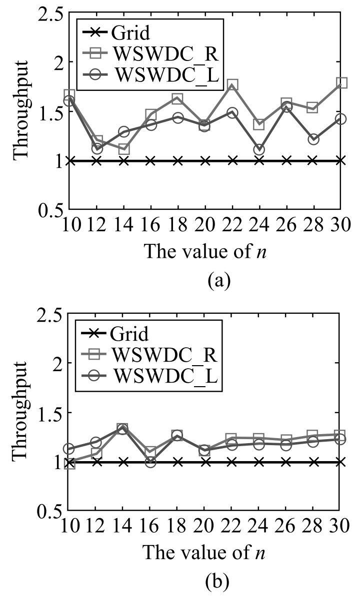


**Fig. 5** Throughput under uniform random traffic. (a) Using 4Mb flows; (b) using 12Mb flows

### 6.4.2 Network performance under local random traffic

In this traffic pattern, workload is unbalanced. We generate random but nonuniform flows in network. Specifically, we let the first 12.5 percent of the network (the lowest rows) bear the 25 percent of traffic, and the remaining 75 percent traffic is spread uniformly among the whole network [16].

The lack of uniformity in this traffic pattern generates an unbalanced utilization of network resource. In this experiment, we first set the flow size as 4 Mb, and measure the throughput and latency under diverse network scales by ranging $n$ from 10 to 30. Figure 7 and Fig. 8 report the evaluation results. Figure 7 depicts the ratio of the real throughput of WSWDC and grid. The throughput here has similar trends to that under uniform random traffic pattern. WSWDC can achieve higher throughput due to the introducing of random links. We can also find that the WSWDC_L which enhanced by lasers can achieve higher throughput than WSWDC_R when the flow size is 4 Mb. When the flow size is 12 Mb, the two kinds of WSWDC achieve almost the same throughput.

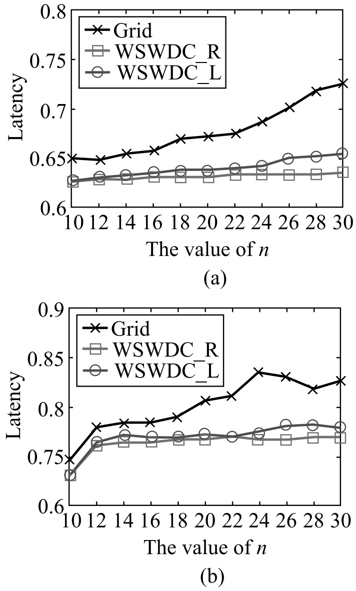


**Fig. 6** Latency under uniform random traffic. (a) Using 4Mb flows; (b) Using 12Mb flows

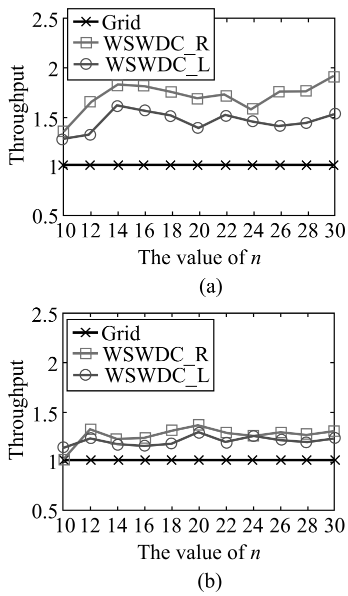


**Fig. 7** Throughput under local random traffic. (a) Using 4Mb flows; (b) using 12Mb flows

Figure 8 shows the average packet delivery latency. In this traffic pattern, the unbalanced workloads affect the overall performance of networks. So it causes higher latency than uniform random traffic. For WSWDC_R and WSWDC_L, the latency grows smoothly and slowly just like the trends under uniform random traffic. But for grid, the latency grows more sharply and fluctuates more widely, since its path is longer in length and less in quantity.

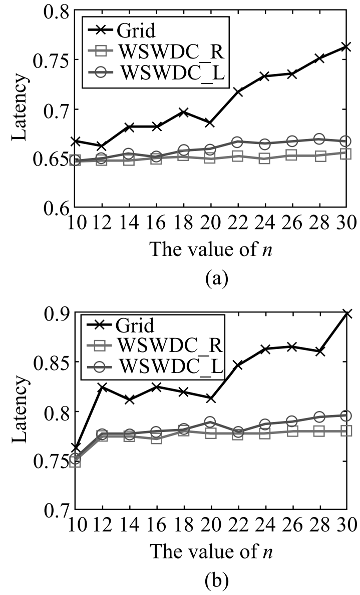


**Fig. 8** Latency under local random traffic. (a) Using 4Mb flows; (b) Using 12Mb flows

In summary, WSWDC achieves better network performance than grid under both uniform and local random traffic. It achieves the larger throughput and less latency.

## 7 Conclusion

In this paper, we propose VLC enabled WSWDC, an easy-deployable and high-performance inter-rack wireless DCN structure. It satisfies the three rationales we proposed. WSWDC employs VLC technology to achieve the fully inter-rack wireless connections for the first time. It is constructed based on grid. Inspired by small-world phenomenon, random links are introduced to optimize topology parameters, such as average path length and network diameter. The placement of racks is also considered to optimize these parameters. Thus it can achieve better network performance, such as throughput and latency. Besides, greedy algorithm based on its topology are proposed for efficiency routing. Using identifiers assigned to racks, it is easy to generate a routing path for packets. The experiments show that under both uniform random and local random traffic, the WSWDC can achieve larger throughput and less packet delivery latency than Grid.

**Acknowledgements** This work was partially supported by the National Natural Science Foundation for Outstanding Excellent young scholars of China (61422214), the National Natural Science Foundation of China (Grant No. 61772544), National Basic Research Program (973 program)

(2014CB347800), the Hunan Provincial Natural Science Fund for Distinguished Young Scholars (2016JJ1002), and the Guangxi Cooperative Innovation Center of cloud computing and Big Data (YD16507 and YD17X11).

## References

1. Luo L, Guo D, Li W, Zhang T, Xie J, Zhou X. Compound graph based hybrid data center topologies. Frontiers of Computer Science, 2015, 9(6): 860–874
2. Guo D, Xie J, Zhou X, Zhu X, Wei W, Luo X. Exploiting efficient and scalable shuffle transfers in future data center networks. IEEE Transactions on Parallel and Distributed Systems, 2015, 26(4): 997–1009
3. Zhang J, Chen J, Luo J, Song A. Efficient location-aware data placement for data-intensive applications in geo-distributed scientific data centers. Tsinghua Science and Technology, 2016, 21(5): 471–481
4. Al-Fares M, Loukissas A, Vahdat A. A scalable, commodity data center network architecture. ACM Sigcomm Computer Communication Review, 2008, 28(4): 63–74
5. Greenberg A, Hamilton J R, Jain N, Kandula S, Kim C, Lahiri P, Maltz D A, Patel P, Sengupta S. VL2: a scalable and flexible data center network. ACM Sigcomm Computer Communication Review, 2011, 39(4): 51–62
6. Mysore R N, Pamboris A, Farrington N, Huang N, Miri P, Radhakrishnan S, Subramanya V, Vahdat A. PortLand: a scalable fault-tolerant layer 2 data center network fabric. ACM Sigcomm Computer Communication Review, 2009, 39(4): 39–50
7. Zhao Y, Wu J, Liu C. On peer-assisted data dissemination in data center networks: analysis and implementation. Tsinghua Science and Technology, 2014, 19(1): 51–64
8. Cui Y, Wang H, Cheng X, Chen B. Wireless data center networking. IEEE Wireless Communications, 2011, 18(6): 46–53
9. Zhou X, Zhang Z, Zhu Y, Li Y, Kumar S, Vahdat A, Zhao B Y, Zheng H. Mirror mirror on the ceiling: flexible wireless links for data centers. ACM Sigcomm Computer Communication Review, 2012, 42(4): 443–454
10. Hamedazimi N, Qazi Z, Gupta H, Sekar V, Das S R, Longtin J P, Shah H, Tanwer A. Firefly: a reconfigurable wireless data center fabric using free-space optics. ACM Sigcomm Computer Communication Review, 2014, 44(4): 319–330
11. Ghobadi M, Mahajan R, Phanishayee A, Devanur N, Kulkarni J, Ranade G, Blanche P A, Rastegarfar H, Glick M, Kilper D. Projector: agile reconfigurable data center interconnect. In: Proceedings of ACM Sigcomm Conference. 2016, 216–229
12. Hussein A T, Elmirghani J M. Mobile multi-gigabit visible light communication system in realistic indoor environment. Journal of Lightwave Technology, 2015, 33(15): 3293–3307
13. Vijay S, Geetha K. A survey on visible light communication appliances used in inter-vehicular and indoor communication. International Journal of Applied Engineering Research, 2016, 11(7): 4893–4897
14. Katayama Y, Takano K, Kohda Y, Ohba N, Nakano D. Wireless data center networking with steered-beam mmwave links. In: Proceedings of Wireless Communications and Networking Conference. 2011, 2179–2184
15. Zhang W, Zhou X, Yang L, Zhang Z, Zhao B Y, Zheng H. 3D beamforming for wireless data centers. In: Proceedings of ACM Workshop on Hot Topics in Networks. 2011, 4
16. Camara J M, Moreto M, Vallejo E, Beivide R, Miguel-Alonso J, Martinez C, Navaridas J. Twisted torus topologies for enhanced interconnection networks. IEEE Transactions on Parallel and Distributed Systems, 2010, 21(12): 1765–1778
17. Kleinberg J. The small-world phenomenon: an algorithmic perspective. In: Proceedings of ACM Symposium on Theory of Computing. 2000, 163–170
18. Ranachandran K, Kokku R, Mahindra R, Rangarajan S. 60GHz datacenter networking: wireless=> worryless. NEC Technical Report, 2008
19. Halperin D, Kandula S, Padhye J, Bahl P, Wetherall D. Augmenting data center networks with multi-gigabit wireless links. In: Proceedings of ACM Sigcomm Conference. 2011, 38–49
20. Cui Y, Xiao S, Wang X, Yang Z, Zhu C, Li X, Yang L, Ge N. Diamond: nesting the data center network with wireless rings in 3D space. In: Proceedings of Usenix Conference on Networked Systems Design and Implementation. 2016, 657–669
21. Shin J Y, Sirer E G, Weatherspoon H, Kirovski D. On the feasibility of completely wirelesss datacenters. IEEE/ACM Transactions on Networking, 2013, 21(5): 1666–1679
22. Luo L, Guo D, Wu J, Rajbhandari S, Chen T, Luo X. VLCcube: a VLC enabled hybrid network structure for data centers. IEEE Transactions on Parallel and Distributed Systems, 2017, 28(7): 2088–2102
23. Louvros S, Fuschelberger D. VLC technology for indoor LTE planning. In: Sklavos N, Hübner M, Goehringer D, Kitsos P, eds. System-Level Design Methodologies for Telecommunication. Springer, Cham, 2014, 21–41
24. Sagotra R, Aggarwal R. Visible light communication. International Journal of Computer Trends and Technology, 2013, 4(4): 906–910
25. Kleinfeld J S. The small world problem. Society, 2002, 39(2): 61–66
26. Guo D, Li C, Wu J, Zhou X. Dcube: a family of network structures for containerized data centers using dual-port servers. Computer Communications, 2014, 53: 13–25
27. Mittal R, Dukkipati N, Blem E, Wassel H, Ghobadi M, Vahdat A, Wang Y, Wetherall D, Zats D. TIMELY: RTT-based congestion control for the datacenter. ACM Sigcomm Computer Communication Review, 2015, 45(4): 537–550
28. Shin J Y, Wong B, Sirer E G. Small-world datacenters. In: Proceedings of ACM Symposium on Cloud Computing. 2011, 1–13
29. Ji S, Li B. Wide area analytics for geographically distributed datacenters. Tsinghua Science and Technology, 2016, 21(2): 125–135
30. Guo D, Wu J, Liu Y, Jin H, Chen H, Chen T. Quasi-kautz digraphs for peer-to-peer networks. IEEE Transactions on Parallel and Distributed Systems, 2011, 22(6): 1042–1055

Yudong Qin received the BS degree in management science and engineering from National University of Defense Technology, China in 2016. He is currently working towards the MS degree in College of Systems Engineering, National University of Defense Technology, China. His research interests include data centers and software-defined networks.

Deke Guo received the BS degree in industry engineering from the Beijing University of Aeronautics and Astronautics, China in 2001, and the PhD degree in management science and engineering from the National University of Defense Technology, China in 2008. He is currently a professor with the College of Systems Engineering, National University of Defense Technology. His research interests include distributed systems, software-defined networking, data center networking, wireless and mobile systems, and interconnection networks. He is a member of the ACM.

Lailong Luo received the BS and MS degrees from the School of Information System and Management, National University of Defense Technology, China in 2013 and 2015, respectively, where he is currently pursuing the PhD degree with the College of Systems Engineering. His current research interests include data centers and software defined networks.

Geyao Cheng received the BS degree in management science and engineering from National University of Defense Technology, China in 2017. She is currently working towards the MS degree in College of Systems Engineering, National University of Defense Technology, China. Her research interests include data centers and software-defined networks.

Zeliu Ding received the MS and PhD degrees from the National University of Defense Technology, China in 2007 and 2012, respectively. He is currently a lecturer with the School of Electronic Engineering, Naval University of Engineering, China. His research interests include complex information system, cloud computing and data center networks.