

# A Survey on Wireless Flyways in Data Centers

Bojie Li, *Student Member, IEEE*

**Abstract**—Modern data centers contains potentially hundreds of server racks, at which scale full-bisection bandwidth is unrealistic and over-subscription is common. The 60 GHz wireless technology can deliver multi-Gigabit throughput indoor, and with directional antennas multiple wireless links would not interfere with each other. By creating wireless “flyways” on-demand, hotspots in data center can be eliminated and the performance of a wireless-augmented over-subscribed wired network can match that of a static non-over-subscribed network.

**Index Terms**—Data Center Networking, Wireless Technology, Beamforming

## I. INTRODUCTION

IN this survey we present several papers to utilize the 60 GHz wireless technology in a data center to augment the wired network and alleviate hot-spots.

Traditional data center networks are tree-structured, with an over-subscription ratio of 1:2 to 1:4 [1, 2]. Servers are packed into racks, on top of which there is a *Top-of-Rack* (ToR) switch. ToR switches are then each connected to one *Aggregation Switch*. Each rack can hold up to 40 servers each with 1 Gbps link, and the link from ToR switch to aggregation switch is typically 10 Gbps, leading to a 1:4 over-subscription ratio. *Core switches* provide links to outside data center and each core switch connects to every aggregation switch. Packets in data center network is IP-routed and ECMP routing is employed to provide load balancing and failover between aggregation switch and core switch layers [6].

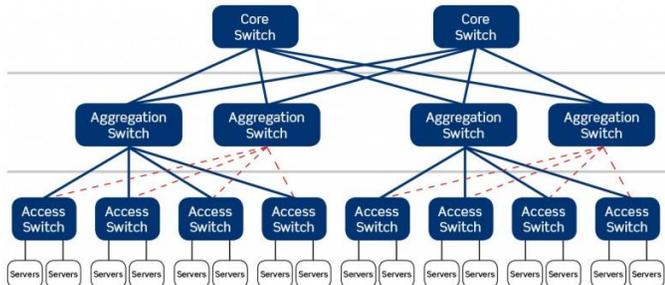


Fig. 1. Traditional data center topology. Dashed lines are for failover and do not carry traffic at normal time.

This survey is submitted on Dec. 31, 2014 as a course paper for Wireless Ad-hoc Networks. This survey does not contain unpublished ideas or confidential information, so it is free to redistribute this work.

Bojie Li (SA14011026) is a graduate student with Univ. of Science and Technology of China, Hefei, Anhui, China 230026. He is now on a joint PhD program with Microsoft Research Asia. His research interests include software defined networking and data center networking (e-mail: boj@mail.ustc.edu.cn).

Over-subscription potentially leads to temporary hot-spots in ToR layer, in which case SLA (Service Level Agreement) on network bandwidth may be violated. Recent researches address this problem by adding more links between ToR switch and aggregation switch layer and employ multipath routing schemes [6, 7]. This approach essentially reduces the over-subscription ratio, but comes with large cabling cost [6].

Recent advances in 60 GHz wireless technology made it available for commodity use, and supports multiple Gbps data rates within several meters [3], a magnitude faster than 802.11n running on 2.4 GHz bands, making it comparable to state-of-art wired technologies. Furthermore, directional antennas can beamform waves and reduce inference, therefore multiple wireless links can be deployed within a data center. The wireless links, called *flyways* in [1], can respond to abrupt bandwidth requirements and create temporary wireless channels among source and destination racks, adding bandwidth to the static wired network and alleviates hot-spots caused by over-subscription.

The rest of this survey is organized as follows. Section II introduces 60 GHz wireless technology. Section III illustrates the hot-spot problem in data center. Section IV introduces the first approach to utilize wireless links in data center. Section V introduces a smarter and more efficient approach. We wrap up with Section VI.

## II. 60 GHz WIRELESS TECHNOLOGY

57-64 GHz sub-band in RF is unlicensed and only suitable for short-range radio communications because signals in this band is severely attenuated due to resonance with oxygen molecules in the air [8]. Furthermore, the 60 GHz band is 100 times wider than its 802.11 counterpart and therefore 20 dB noisier.

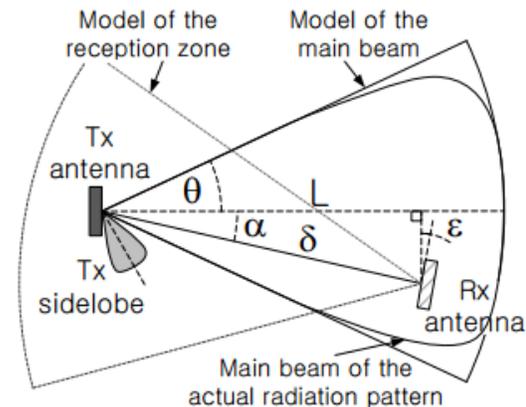


Fig. 2. Geometric communication model of directional antenna [8]

Fortunately, directionality is inversely proportional to the square of the wavelength [8], enabling highly directional antennas to achieve high gain and avoid non-line-of-sight (NLOS) blockage by obstacles. Racks in data center are fixed and highly directional antennas can be deployed on top of each rack so that transceivers and receivers are in line-of-sight.

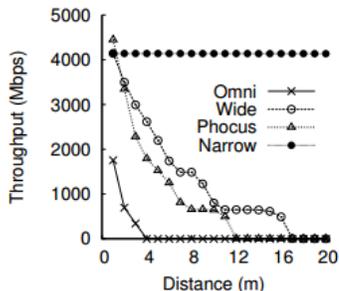


Fig. 3. TCP throughput at various distances with different antennas [2]. This figure underscores the need of directional antennas, where omni-directional antennas can only work within 4m range, wide antennas can go further, and wide antennas provide high throughput consistently within 20m range.

The short wavelength also enable antenna arrays with tens of elements on a single die [3], which has been commercialized by SiBeam’s WirelessHD [10]. The point of antenna array is that the attenuation and phase shift of each antenna can be controlled independently, and the antenna with maximum gain can be chosen to transmit information.

IEEE 802.11 Working Group has been working on 802.11ad standard [9] focusing on short range indoor communications, which is mostly compatible with MAC and PHY of 802.11n, while some changes have been made to support data rates up to 6.76 Gbps.

### III. OVER-SUBSCRIPTION AND HOTSPOTS

Recent researches on data center networking propose to eliminate oversubscription and achieve full bisection. However, some network topologies such as Fat-Tree [11] require so many wires that cabling is a big problem, and other schemes such as BCube [7] and VL2 [12] require fundamental change to the routing hardware and software.

TABLE I  
COMPARISON OF 3 NETWORK TOPOLOGIES [1]

		Traditional Tree	Fat-Tree [11]	VL2 [12]
Oversubscription Ratio		1:2	1:1	1:1
#Links	10G	160	0	640
	1G	3200	10112	3200
#Switches	Agg	1	0	5
	Commodity	0	360	0
	Top-of-rack	160	0	160
Network Cost (approx.)		x	2-3x	4-5x

3200 Servers, 160 port \* 10G Aggregation Switch, 1G Server NIC, 48 port commodity switches for Fat-Tree

Whether or not it is necessary to achieve full bisection? How often do links in core network congest because of over-subscription? [1] measured a 1500-server production data

center and results shown in Figure 4, revealing the following trends:

1. At any time, there are only a few hot racks in the data center.
2. The hot racks appear pseudo-randomly and are changing constantly with time, so this problem is unlikely to be solved by adding a small number of static wires or flow scheduling approaches (e.g. [14]) by identifying elephant flows and predict hotspots.
3. The matrix is sparse, i.e. the hot racks are exchanging data with a small number of racks.

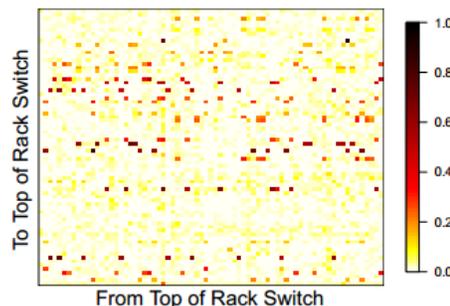


Fig. 4. Traffic heat map among each pair of ToR switch [1]. The dark dots indicate pairs of ToR switches carrying large amount of traffic.

Instead of provisioning for the rarely-used high-traffic case, [1] and its followers utilize wireless channel to add dynamic surplus capacity to respond to traffic hot-spots on demand, and the base wired network only needs to be provisioned for the average case.

The median ToR is exchanging 55% traffic with 10 other ToRs [1], therefore providing an additional link between the hot ToR and its most traffic-demanding destination would not only improve flow completion time between these two ToRs, but also reduce traffic on the base network, thus improving flow completion time between the source ToR and other ToRs.

### IV. IMPLEMENTATION OF 2D FLYWAYS

Stability is a major concern in using wireless links to augment wired network. Variations in wireless deployments typically come from fluctuations of environmental factors, movement of devices and obstacles and interference from other wireless devices. However, data center is a controlled place, and line of sight is always clear when transceivers and receivers are put on top of rack. Stable SNR and TCP throughput with narrow directional antennas has been reported by [2].



Fig. 5. Narrow-beam (left) and wide-beam (right) horn antennas for 60 GHz, used in [2]

Another concern is potential inference among multiple

flyways. [2] suggests that with narrow directional antennas, there can be at least 10 concurrent flyways operating at 6 Gbps, or 30 concurrent flyways at 1 Gbps.

Since the completion time of the last flow within a task is a major metric of flow scheduling, the main point of flyways is to eliminate the stragglers. For a fixed number of flyways, neither too many flyways at one ToR or too many ToRs with one flyway is a good choice. Placement of flyways has been formulized as an optimization problem in [1].

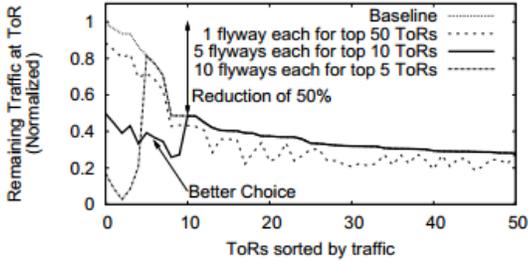


Fig. 6. How placement of flyways influence performance [1].

In the pioneering works of wireless flyways [1, 2], transceivers and receivers are fixed and they point to their corresponding ones in neighboring rack. Distant transceivers and receivers cannot point to each other directly because 60 GHz signals are extremely sensitive to obstacles on the line-of-sight, and transceivers and receivers of other racks are obstacles in the grid.

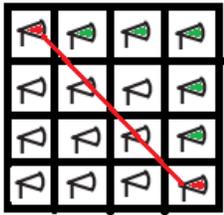


Fig. 7. Line-of-sight is blocked by other transceivers and receivers (red line), so a hop-by-hop forwarding scheme must be employed (green).

For communication among distant racks, the central controller plans a route from the source rack to the destination rack, and instructs nodes on the route to forward the packets. In this multi-hop relay scheme, each intermediate node can only serve for one route, which restricts concurrency and scalability of wireless flyways.

## V. INFERENCE AND 3D BEAMFORMING

Another serious problem of 2D Flyways is link inference that limits concurrent transmissions in dense data centers [3].

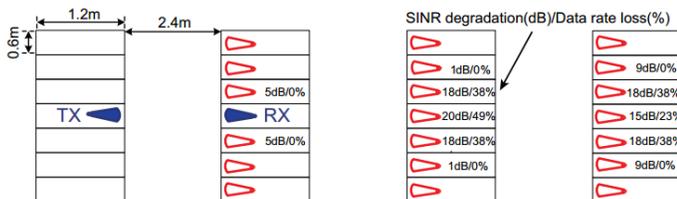


Fig. 8. Inference of transmitter (blue) to victim receivers (red) behind the designated receiver (blue) [3].

Alternative data center physical layouts have been proposed

[13] to reduce inference among neighboring racks. However, this approach requires rack containers to be redesigned, which cannot be incrementally deployed in existing data centers.

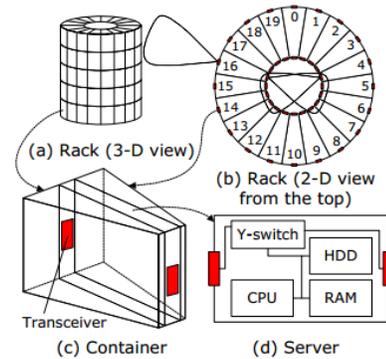


Fig. 9. Cylindrical rack design to utilize both space and RF spectrum in [13]

A more practical design [3] utilized 3D space and puts a mirror on the ceiling of data center to reflect a focused beam from the transceiver directly to the receiver. This approach does not need a multi-hop relay, and eliminates inference of transceiver to the receivers in neighboring racks.

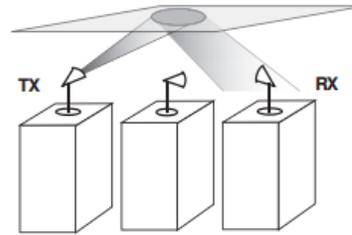


Fig. 10. 3D beamforming with mirror on the ceiling [3]

Since 3D beamforming constrains beam in a horn-shaped space reflected by the mirror, the intersection area of such beam and the top-of-rack plane where receivers reside is much smaller than the 2D beamforming approach.

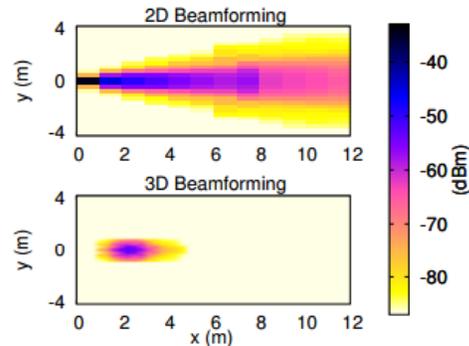


Fig. 11. Inference map of 2D Beamforming and 3D Beamforming on top-of-rack plane [3]

Furthermore, since 3D beamforming can create wireless link between any two pairs of racks, multiple radios can be deployed in each rack to enable one rack to simultaneously transfer packets to multiple racks using different radios. Results of [3] show that number of concurrent links supported by 3 channels can grow almost linearly with the number of radios per rack, enabling the majority of rack pairs to be connected wirelessly.

There are several challenges to the 3D beamforming approach:

*Rotation Accuracy.* Rotator accuracy can achieve  $0.006^\circ \sim 0.09^\circ$  [3], with 30m distance there is only 3mm  $\sim$  47mm offset, which is sufficiently precise.

*Rotation Delay.* Rotation delay can be 0.01  $\sim$  1 second [3], which is longer than the duration of some short flows. This is an overhead of 3D beamforming over its 2D counterpart. The scheduler needs to be aware of such overhead.

*Reflection Energy Degradation.* Flat metal plates offer perfect reflection without noticeable energy or changes to path loss characteristics [3].

*Reflection and Scattering Near Antenna.* [3] placed electromagnetic absorbers near transmitting and receiving antennas.

*Accumulative Inference.* Although inference of each individual transmitter is negligible, due to the density of data center racks, the accumulation of inferences from all transmitting neighbors must be taken into account. The scheduler in [3] applies Kelleher's universal horn pattern to model the radiation of antennas.

*Reflector Placement.* Above racks there are typically cables and cooling pipes [15] which are not only obstacles for wireless beams, but also a construction challenge for putting a large, flat mirror on the ceiling.

*Reflector Flatness.* A small curvature of the mirror would create observable multipath signals which would be inference on the receiver side. [3] reported visible increase in RSS and even connection failure when the mirror ceiling is curved.

While 3D beamforming face several physical challenges, it still have significant benefits over 2D beamforming. 3D beamforming can not only be applied to data centers, but also be used as an enhancement to indoor wireless networks [16].

## VI. CONCLUSION

Data center networks are typically designed with some ratio of over-subscription. In order to cover unpredictable surpass traffic in today's data centers, a most researched approach is to add additional wires and switches and deploy more complicated network topology to reduce over-subscription ratio, even achieving full bisection, which incurs high cabling, device and maintenance cost. However, hot-spots generating surpass traffic is measured to be sparse, therefore could be covered by a few dynamic "flyways" among racks.

With directional antennas, 60 GHz band can deliver multi-Gbit bandwidth, enabling wireless flyways to be constructed on the top of racks. Following works have been published to address the inference problem in the pioneering 2D beamforming design by utilizing spatial reflection, making wireless flyways more efficient.

Despite challenges in deployment, wireless flyways is a promising approach to tackle the temporary hot-spot problem in over-subscribed modern data centers.

## REFERENCES

- [1] S. Kandula, J. Padhye, and P. Bahl. Flyways to de-congest data center networks. In HotNets, Nov. 2009.
- [2] D. Halperin, et al. Augmenting data center networks with multi-gigabit wireless links. In Proc. SIGCOMM 2011.
- [3] X. Tie, et al. On 60 GHz Wireless Link Performance in Indoor Environments. In Proc. PAM 2012.
- [4] X. Zhou, et al. Mirror Mirror on the Ceiling: Flexible Wireless Links for Data Centers. In Proc. SIGCOMM 2012.
- [5] H. Huang, et al. The Architecture and Traffic Management of Wireless Collaborated Hybrid Data Center Network. In Proc. SIGCOMM 2013.
- [6] M. Al-Fares et al. A scalable, commodity data center network architecture. In SIGCOMM, 2008.
- [7] C. Guo et al. BCube: High performance, server-centric network architecture for data centers. In SIGCOMM, 2009.
- [8] J. Shin, et al. On the Feasibility of Completely Wireless Datacenters. In Proc. ANCS 2012.
- [9] WG802.11 - Wireless LAN Working Group. <http://standards.ieee.org/develop/project/802.11ad.html>
- [10] SiBeam. <http://sibeam.com/whitepapers/>
- [11] M. Al-Fares, A. Loukissas, and A. Vahdat. A Scalable Commodity Data Center Network Architecture. In SIGCOMM, 2008.
- [12] A. Greenberg, J. Hamilton, N. Jain, S. Kandula, C. Kim, P. Iahiri, D. Maltz, P. Patel, and S. Sengupta. VL2: A scalable and flexible data center network. In SIGCOMM, 2009.
- [13] W. Zhang, et al. 3D beamforming for wireless data centers. In HotNets, 2011.
- [14] M. Al-Fares et al. Hedera: Dynamic flow scheduling for data center networks. In NSDI, 2009
- [15] Sun modular datacenter s20/d20 overview. <http://download.oracle.com/docs/cd/E19115-01/mod.dc.s20/index.html>
- [16] Z. Shi, et al. Three-dimensional spatial multiplexing for directional millimeter-wave communications in multi-cubicle office environments, IEEE GLOBECOM 2013.