Mirror Mirror on the Ceiling: Flexible Wireless Links for Data Centers

Xia Zhou, Zengbin Zhang, Yibo Zhu, Yubo Li*, Saipriya Kumar, Amin Vahdat†, Ben Y. Zhao and Haitao Zheng

Department of Computer Science, UC Santa Barbara

**Xi'an Jiaotong University, China*

†*Google and UC San Diego*

Data Centers are Everywhere

- No longer a luxury for tech companies

Google amazon IBM Microsoft YAHOO! CISCO facebook hp rackspace

Retailers

TARGET Walmart Save money. Live better. Coca-Cola THE HOME DEPOT

Governments

NATIONAL SECURITY AGENCY UNITED STATES OF AMERICA

Universities, hospitals

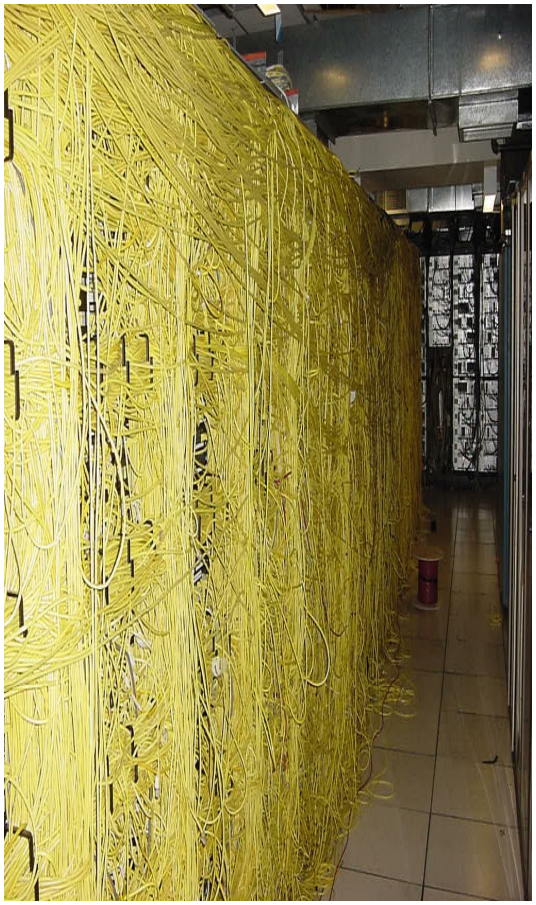
THE UNIVERSITY OF UTAH

SILVER CROSS HOSPITAL The way you *should* be treated.

Children's Hospital of Illinois at OSF Saint Francis Medical Center

Today's Data Centers

- Wiring is complex and costly
  - Planning, deploying, testing 10K+ fibers
  - Takes several weeks or even months
- Difficult to change wiring
  - High labor cost
  - Significant interruptions to operations
- Overprovisioning is difficult
  - Traffic demands unpredictable
  - Limited by hardware costs



Dealing with Traffic Hotspots

- Measurements show sporadic congestion losses caused by traffic hotspots
  - Traffic hotspots are unpredictable, can appear anywhere
  - Can double failure rate for some jobs

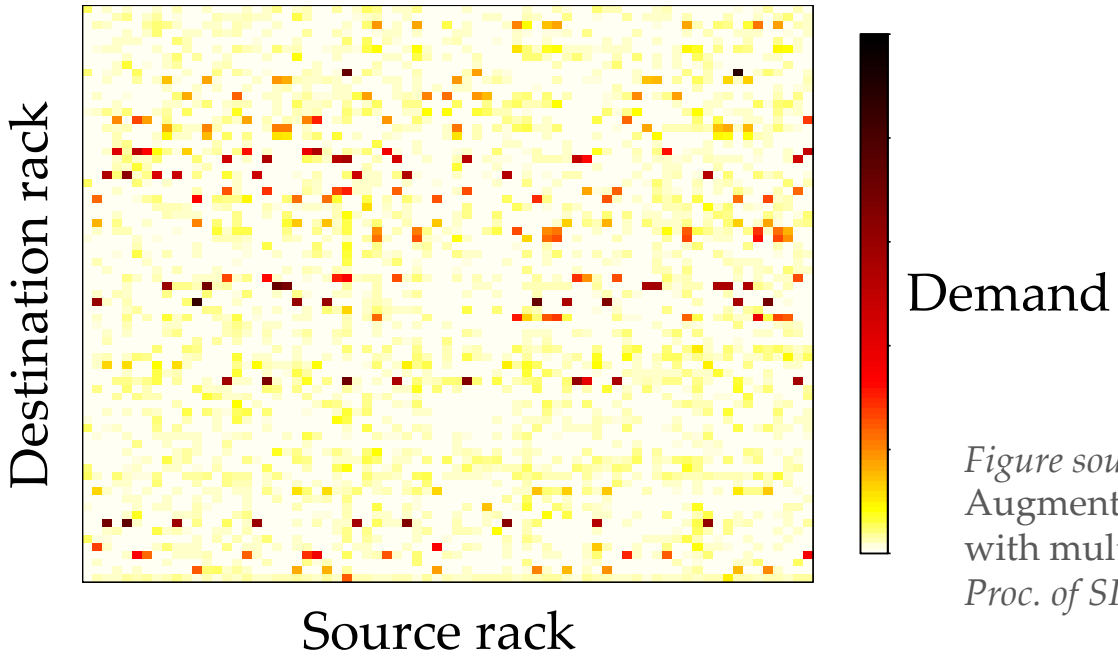


Figure source: Halperin, D., et al. Augmenting data center networks with multi-gigabit wireless links. In *Proc. of SIGCOMM* (2011)

Dealing with Traffic Hotspots

- Measurements show sporadic congestion losses caused by traffic hotspots
  - Traffic hotspots are unpredictable, can appear anywhere
  - Can double failure rate for some jobs

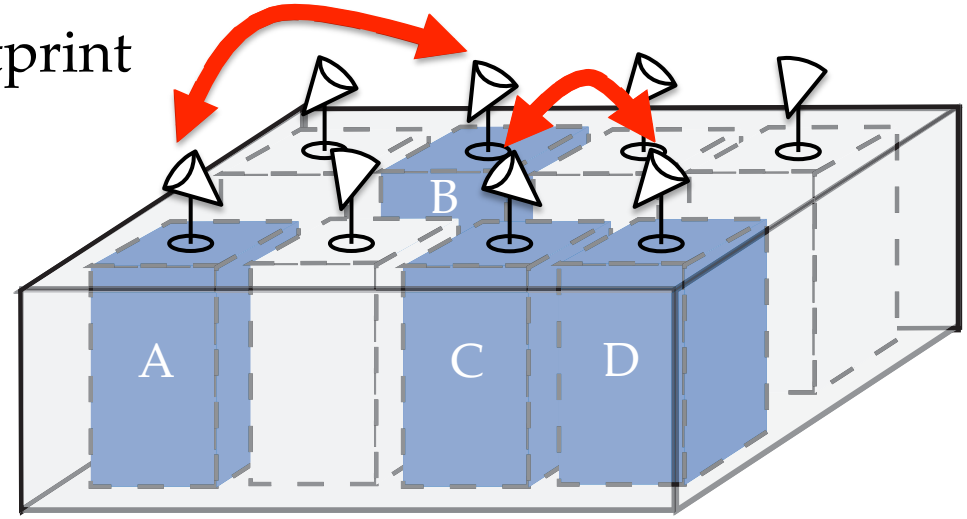
- Hard to add bandwidth using wires
  - ☹ Do not know where to add capacity
  - ☹ Rewiring is complex, high labor cost
  - ☹ Interrupt current operation

Need alternative solutions!

Augmenting via Wireless Links

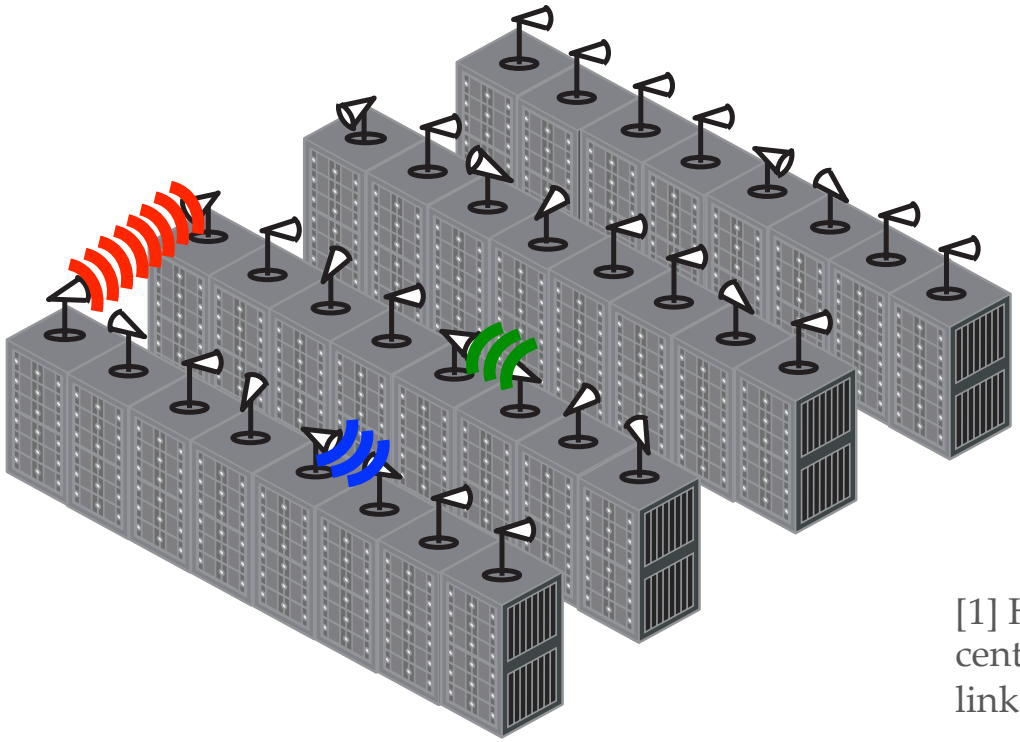
- Key benefit: on-demand links
  - Create links on-the-fly at congestion hotspots
  - Adapt to traffic dynamics

- New wireless technology: 60 GHz beamforming
  - Multi-Gbps data rate
  - Small interference footprint



Existing Work: Connecting Neighboring Racks

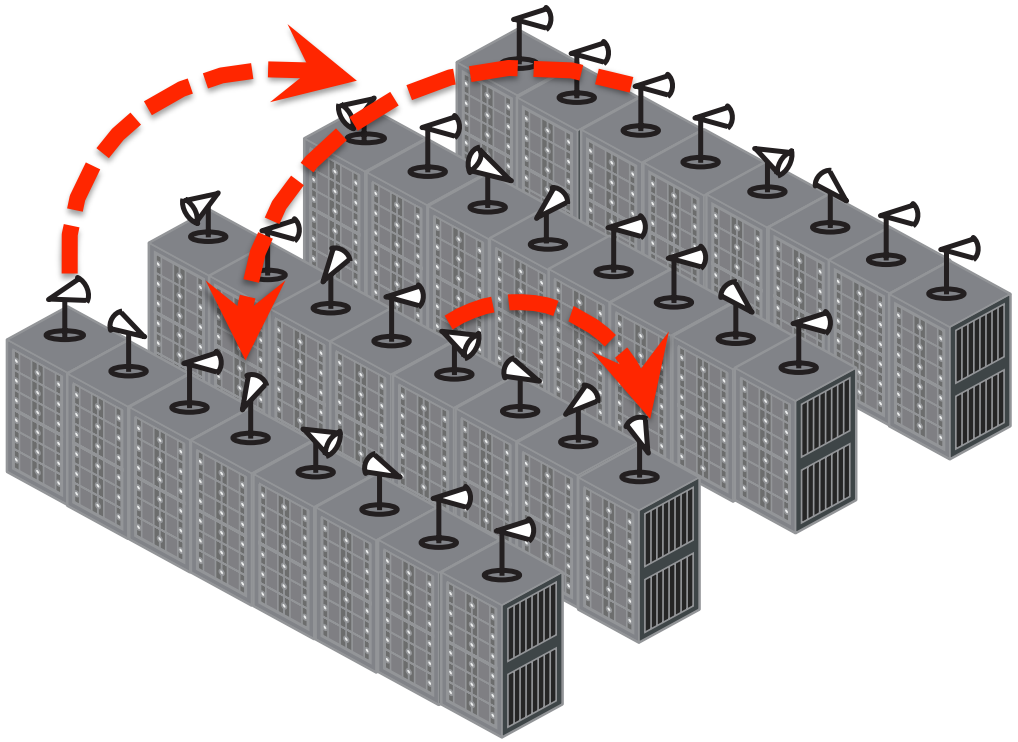
- 60GHz flyways[1] address local traffic hotspots by connecting neighboring racks wirelessly



[1] Halperin, D., et al. Augmenting data center networks with multi-gigabit wireless links. In *Proc. of SIGCOMM* (2011)

Our Goal: Any-to-any Communication

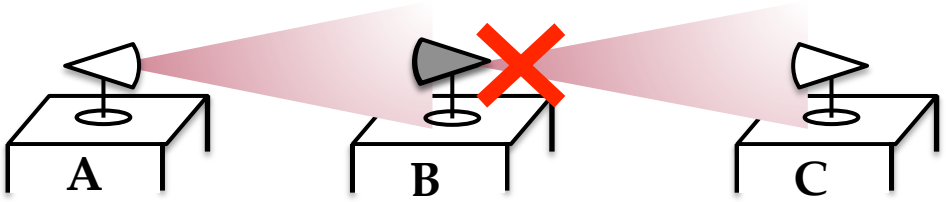
- Traffic hotspots can appear between any rack pair
  → Connect any rack pair wirelessly



Hard to do using existing 60GHz beamforming!

Challenge #1: Link Blockage

- 60GHz transmissions are blocked by small obstacles (anything larger than **2.5mm**!)

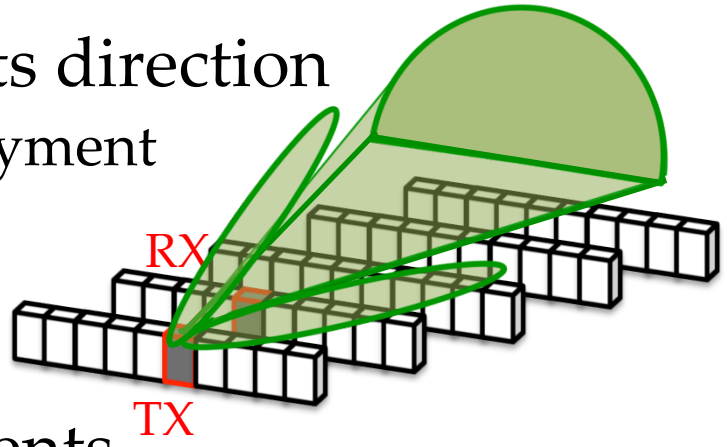


- Confirmed by our testbed measurements
  - Signal strength dropped by 10-30dB
  - Up to 15-90% throughput loss

☹ Must use multi-hop forwarding
  - Antenna rotation delay
  - Reduce throughput by at least half

Challenge #2: Radio Interference

- Beam interferes with racks in its direction
  - Exacerbated by dense rack deployment
  - Signal leakage makes it worse



- Verified via testbed measurements
  - A single link causes 15-20dB drop in signal quality for **15** nearby links

Links interfere with each other

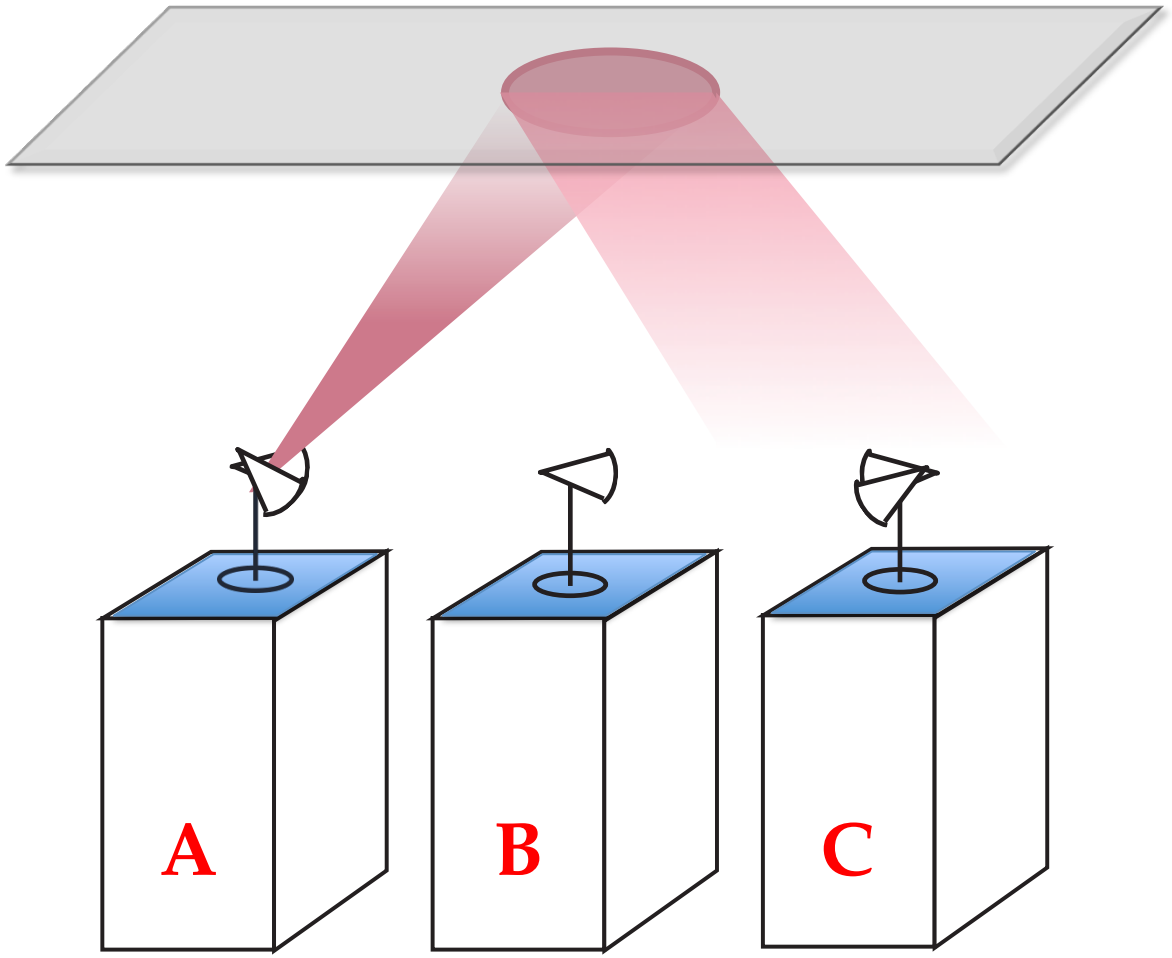
  - Very few links can run concurrently
  - Put a hard limit on aggregate bandwidth

Outline

- Motivation
- **Our solution: 3D beamforming**
- Implications on data centers
- Deployment challenge

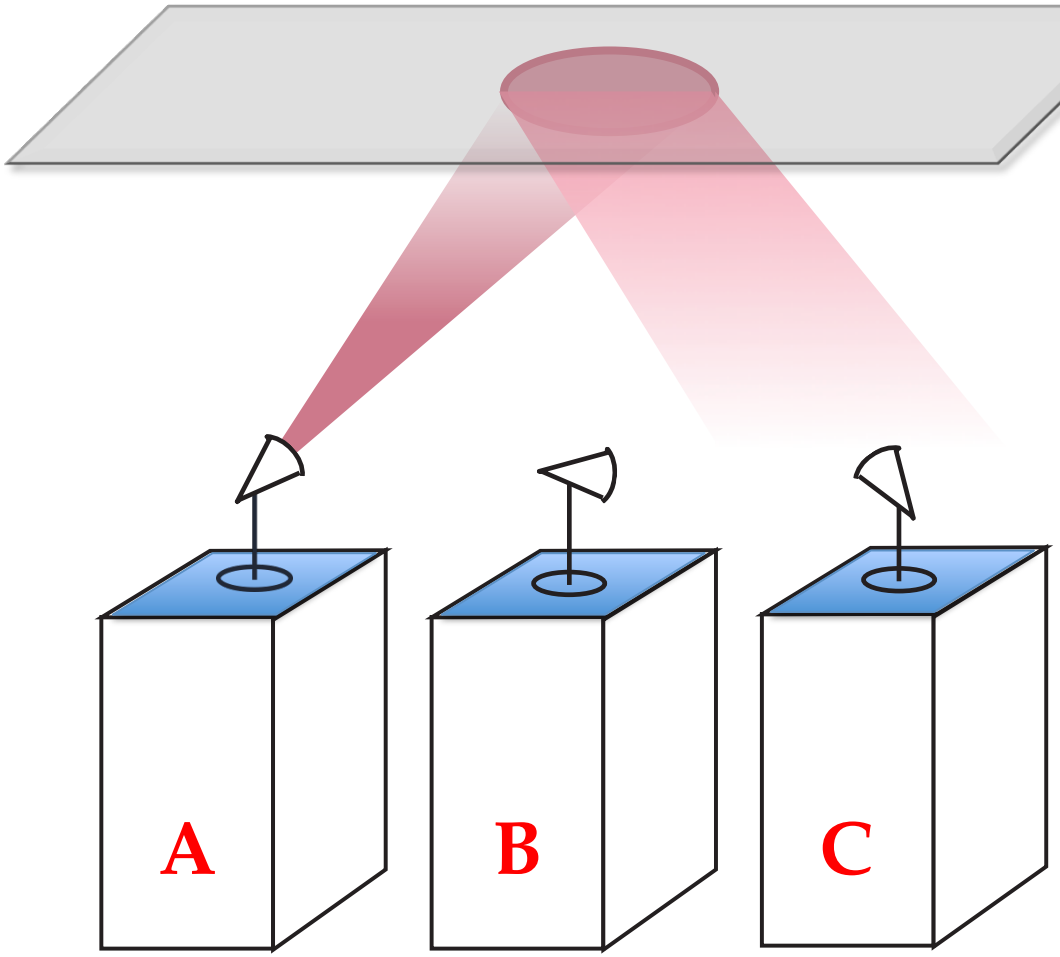
3D Beamforming

Connect racks by reflecting signal off the ceiling!



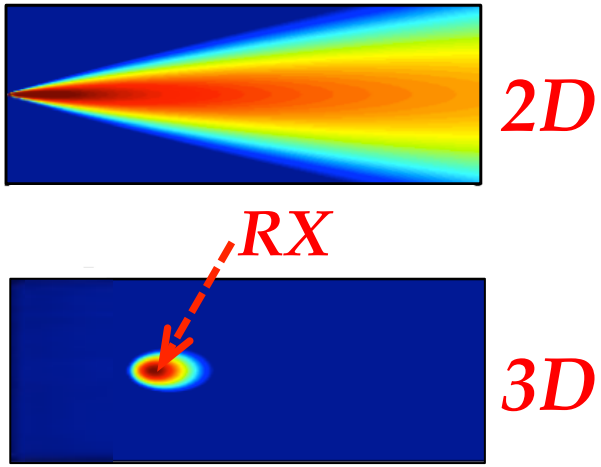
3D Beamforming

Connect racks by reflecting signal off the ceiling!

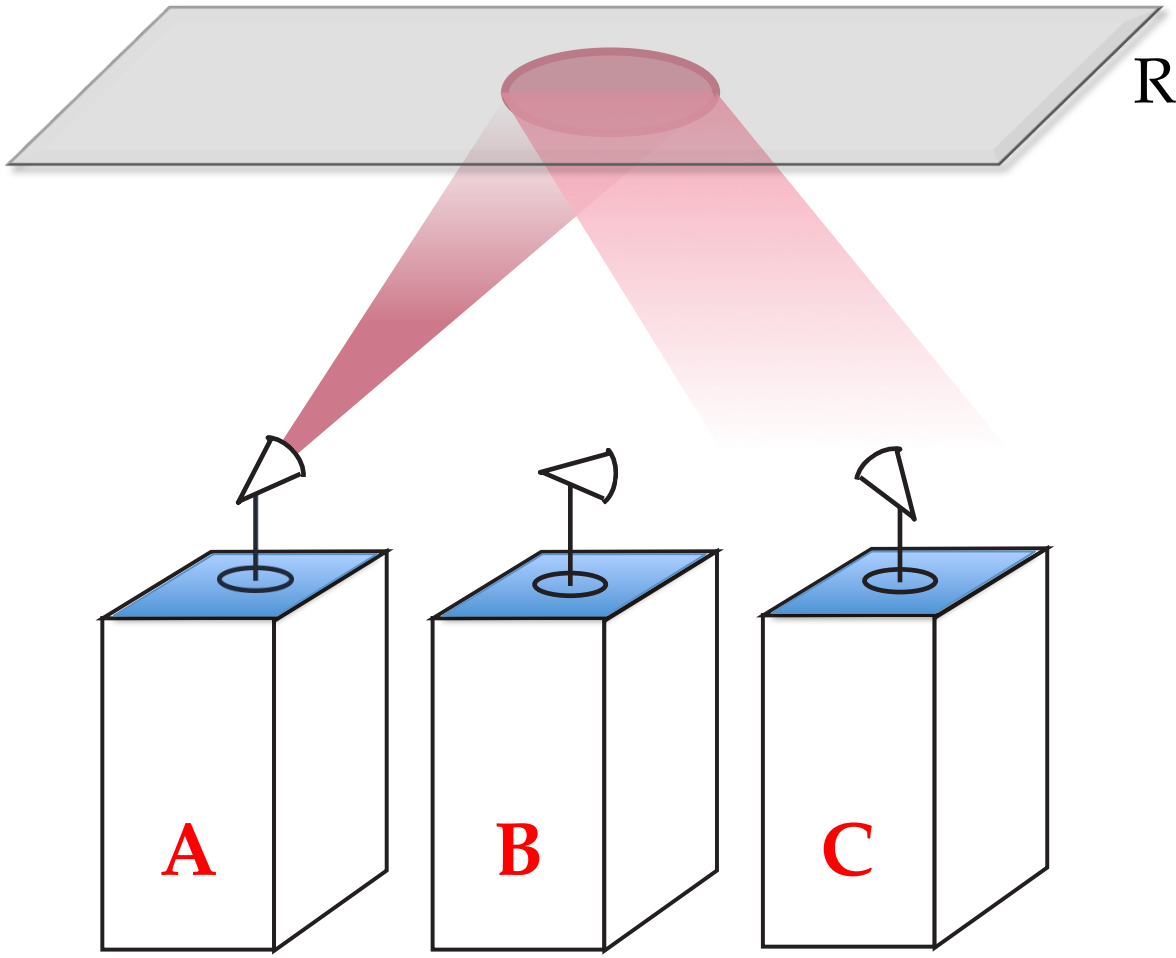


Key Benefits

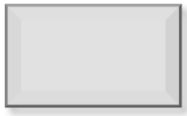
- ✔ No more link blockage
- ✔ Much smaller interference



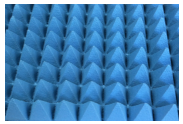
Simple Setup



Reuse existing hardware, low maintenance cost!



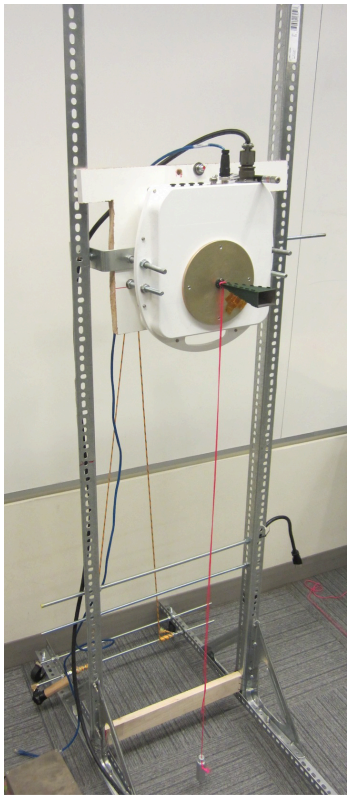
Reflector

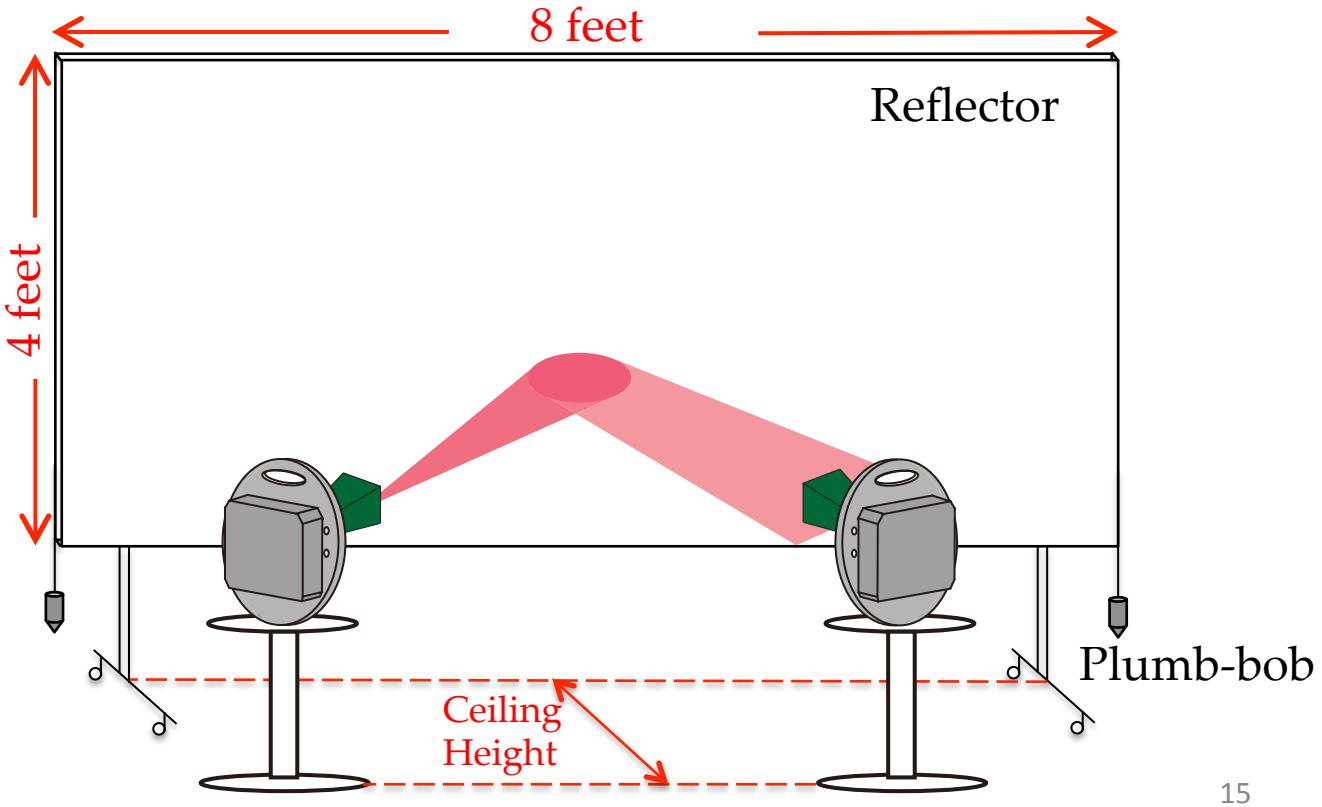


Absorber

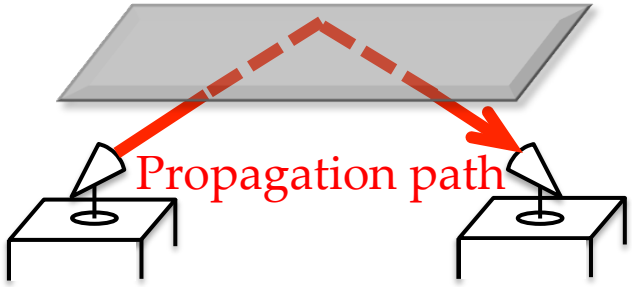
3D Beamforming Testbed

- Off-the-shelf 60GHz radio and horn antenna
  - HXI radio with 0dBm transmission power
  - 10^{o} horn antenna from Flann Microwaves



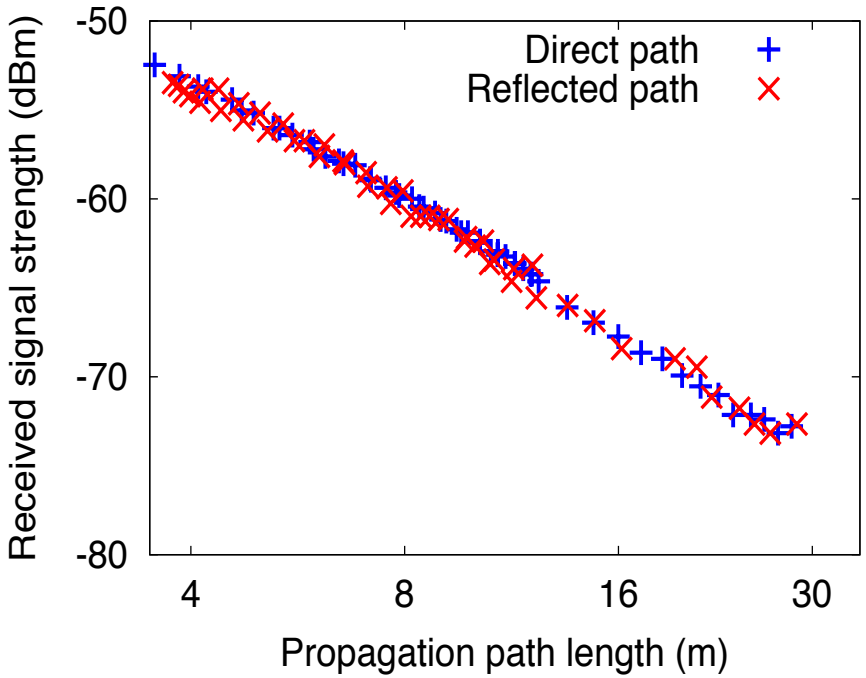


Benchmark #1: Link Connectivity

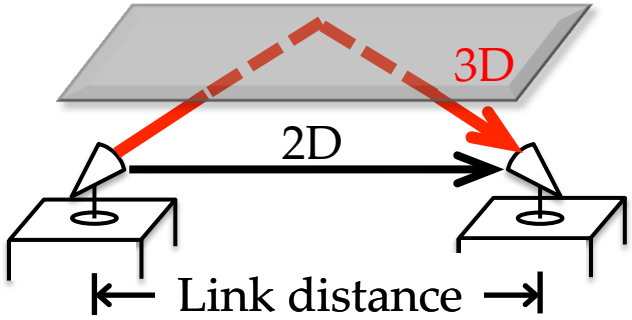


Q1: Does reflection cause any energy loss?

Even cheap metal plate provides perfect reflection!

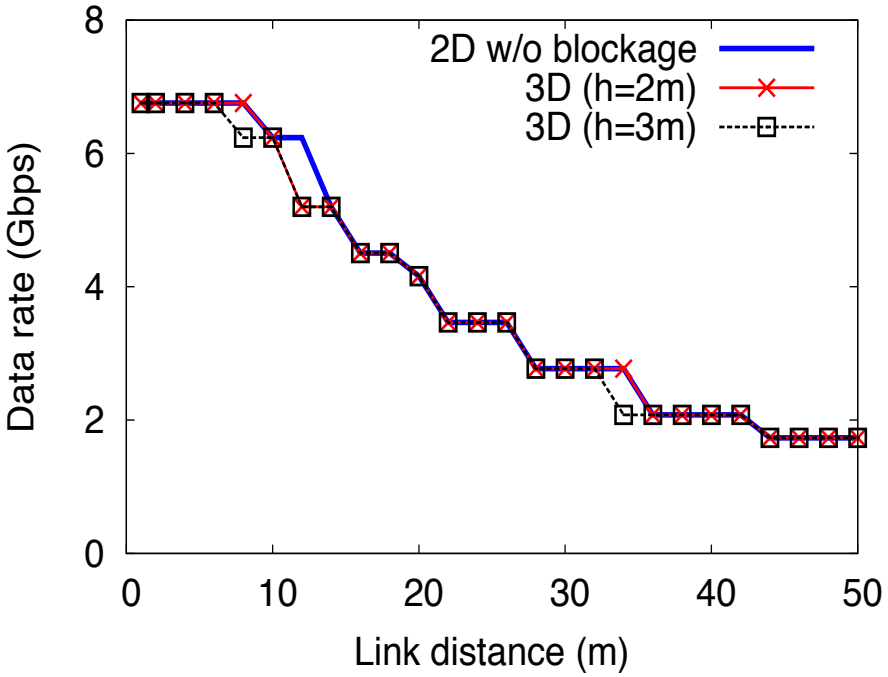


Benchmark #1: Link Connectivity



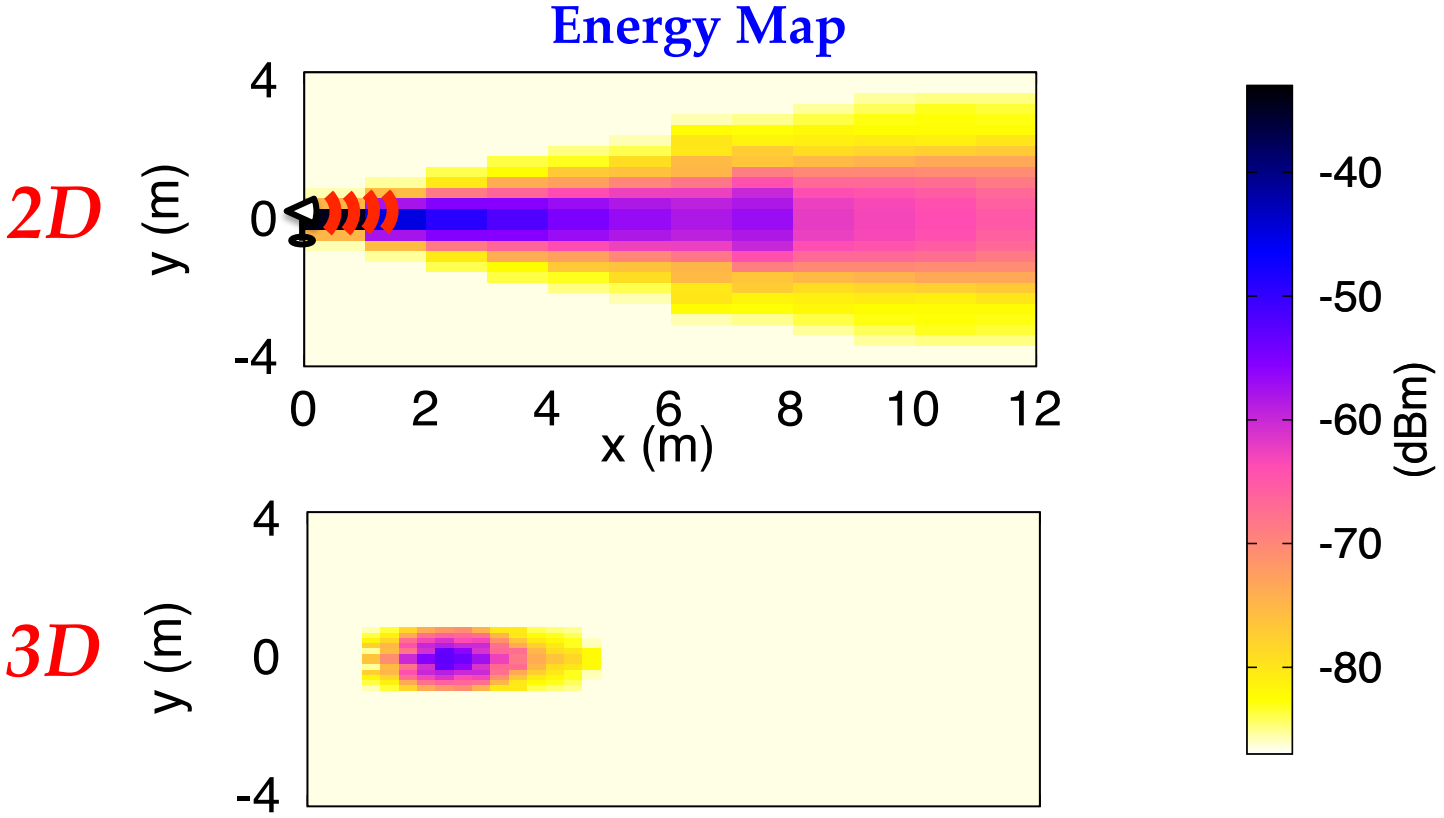
Q2: How does longer propagation path impact data rate?

Negligible data rate loss



Benchmark #2: Interference Footprint

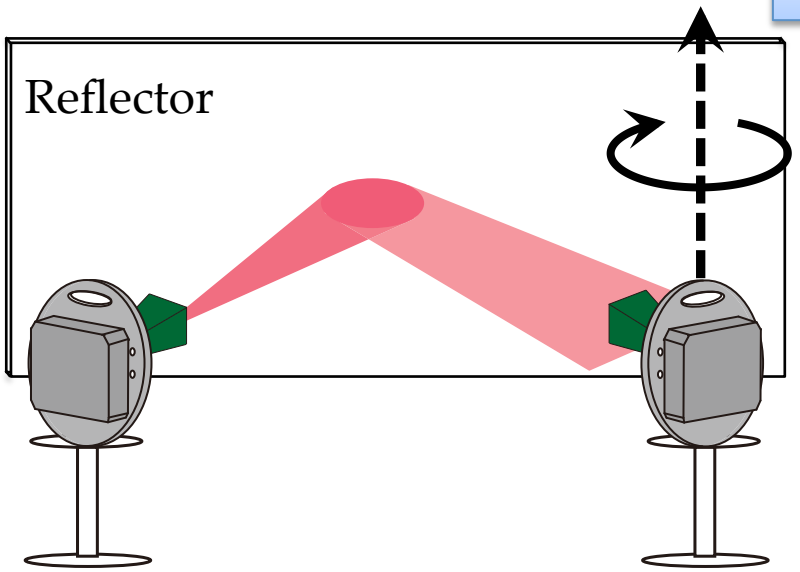
- A transmitter (0,0) communicates with a receiver (2,0)
- Measure the received energy at multiple locations

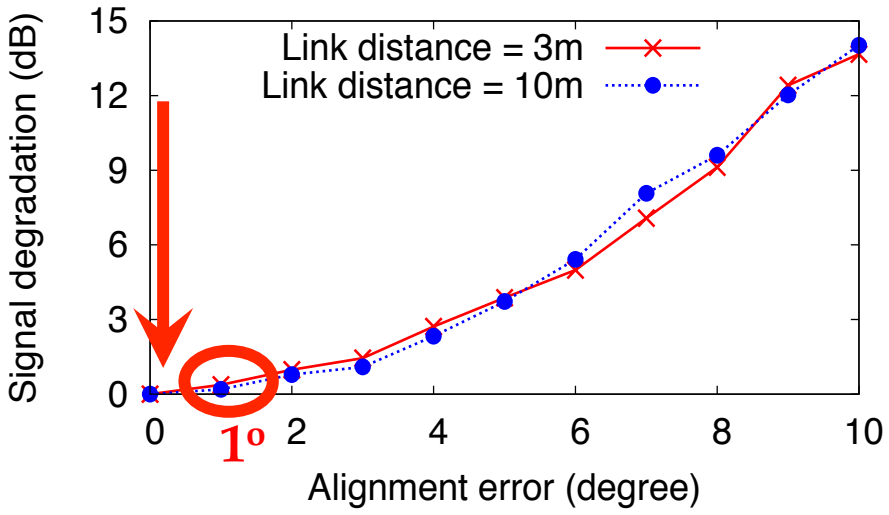


Benchmark #3: Robustness to Alignment Errors

- How does alignment accuracy impact signal strength?
- Fine grain experiment
  - Measure received signal when antennas perfectly tuned
  - Measure signal strength while introducing artificial alignment errors at 1° increments

Today's rotators: 0.006°- 0.09° accuracy





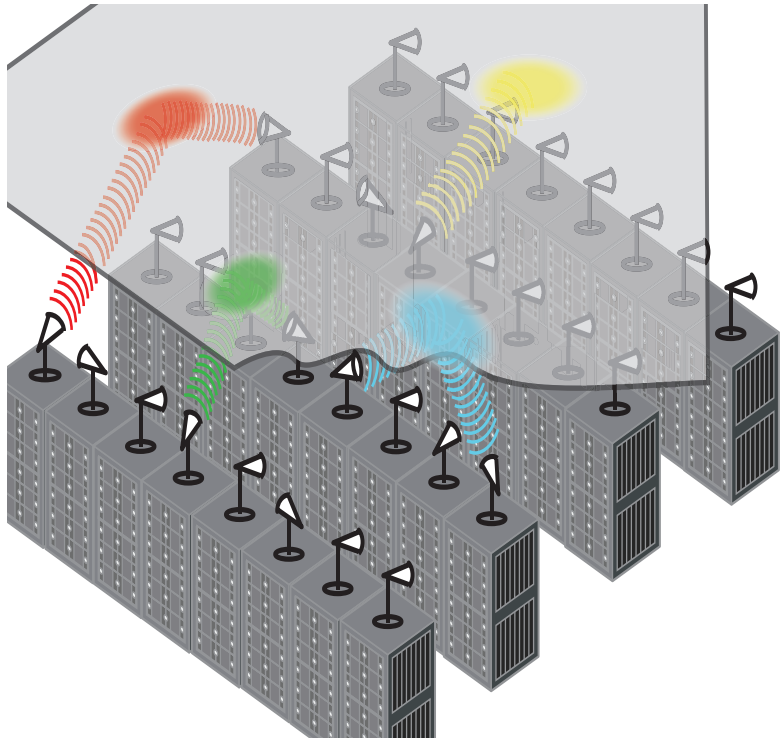
Benefits of 3D Beamforming

- Reflection overcomes link blockage

Connect any rack pair w/ indirect LOS

- Bouncing the beam minimizes interference footprint

Many links can run concurrently



Outline

- Motivation
- Our solution: 3D beamforming
- **Implications on data centers**
- Deployment challenge

Link Concurrency in Data Centers

- Example data center scenario
  - Medium-sized data center: 250 racks in a 42m x 15m room
  - One 60GHz radio per rack
  - 125 randomly chosen bidirectional links w/ 5+Gbps data rate

Results

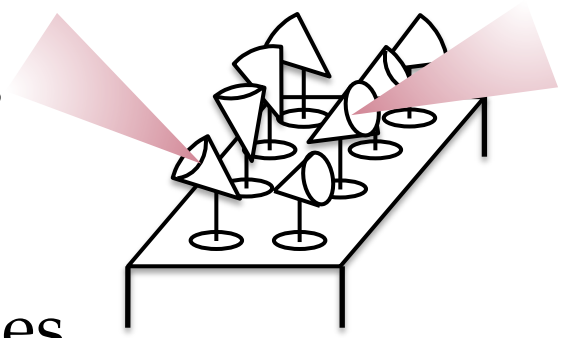
Connect any two racks via a **single hop;**

70% of links run **concurrently** w/ 5+Gbps rate!

Create a highly flexible network with data rates "close" to wired networks

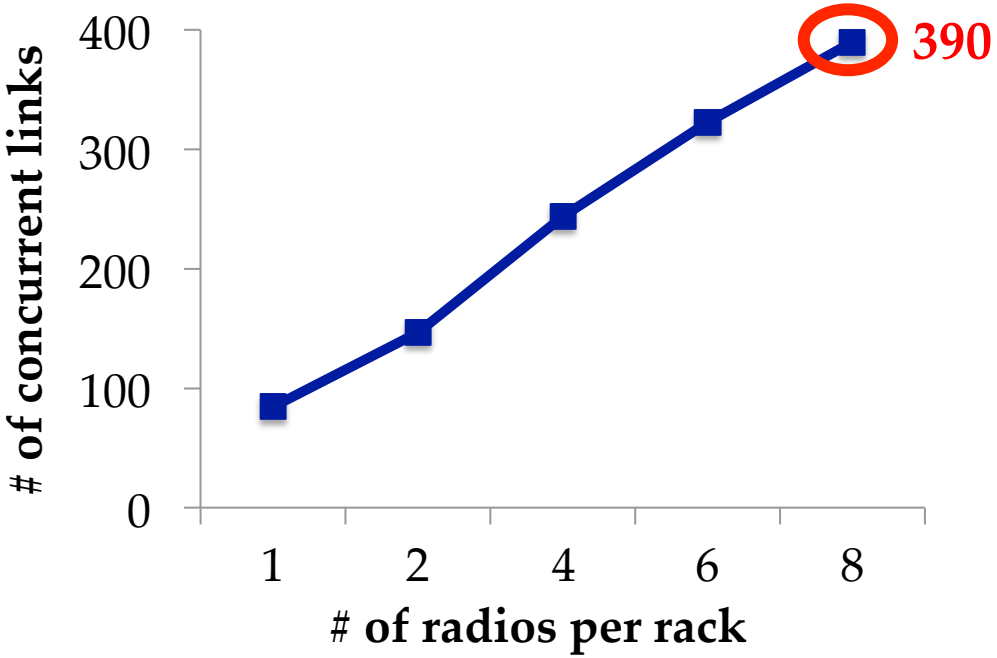
Multiple Radios per Rack

- Each rack can talk to multiple racks concurrently



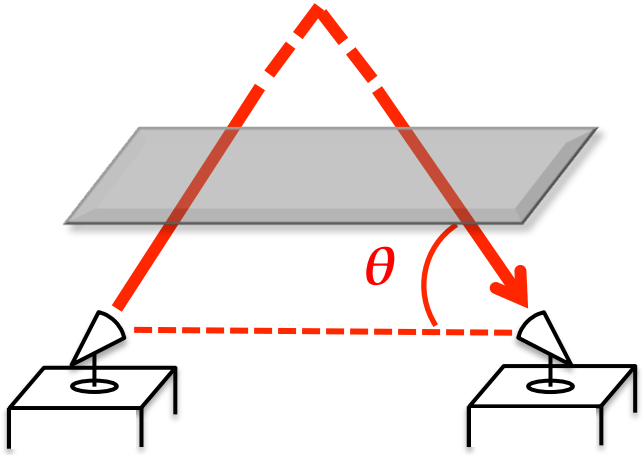
- Number of concurrent links increases linearly w/ the number of radios per rack!

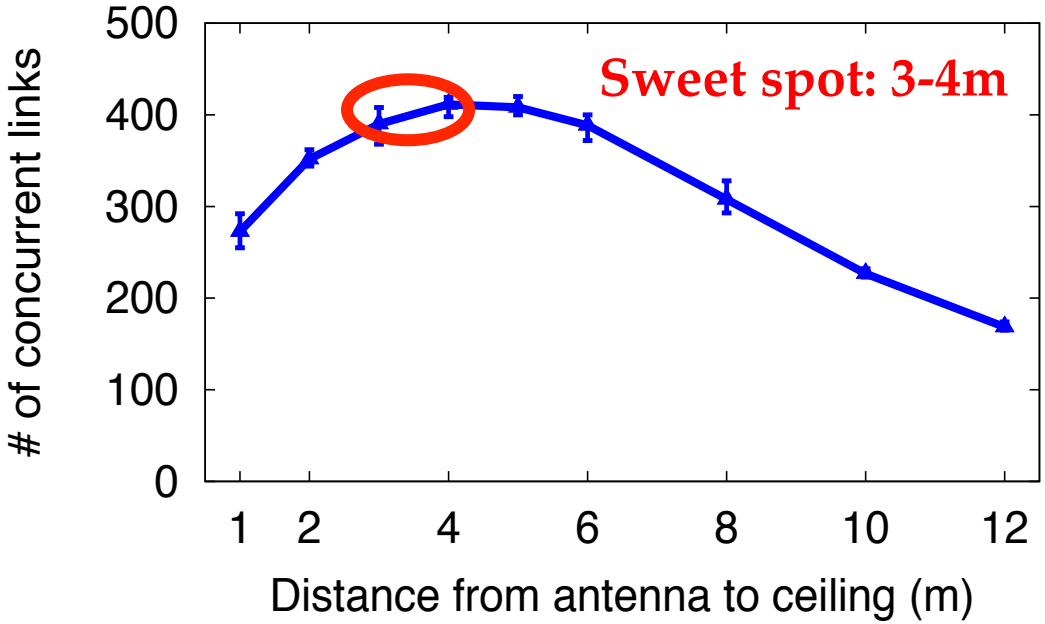
250 racks
5+Gbps links



Impact of Ceiling Height

- How does ceiling height impact performance?
  - Higher ceiling increases signal arrival angle → smaller interference region
  - Also has longer propagation path → signal degradation

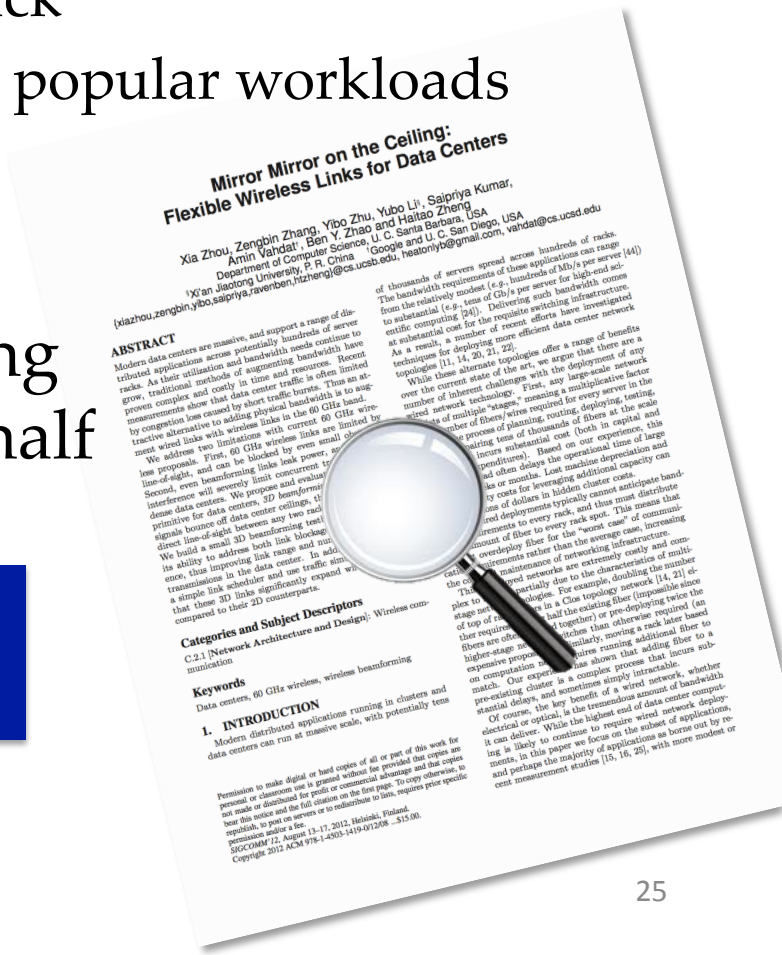




Addressing Traffic Hotspots

- Large-scale data center simulations
  - 250 racks (5K servers), 8 radios/rack
  - Synthetic hotspot traffic based on popular workloads
  - Create 60GHz links for hotspots

- Result: Adding 3D beamforming links cuts completion time by half


Mirror Mirror on the Ceiling:
Flexible Wireless Links for Data Centers

Xia Zhou, Zengbin Zhang, Yibo Zhu, Yubo Li[†], Saipriya Kumar,
Amin Vahdat[‡], Ben Y. Zhao and Haitao Zheng
Department of Computer Science, U. C. Santa Barbara, USA
[†]Xi'an Jiaotong University, P. R. China [‡]Google and U. C. San Diego, USA
{xiazhou,zengbin,yibo,saipriya,ravenben,htzheng}@cs.ucsb.edu, heatonlyb@gmail.com, vahdat@cs.ucsd.edu

ABSTRACT

Modern data centers are massive, and support a range of distributed applications across potentially hundreds of server racks. As their utilization and bandwidth needs continue to grow, traditional methods of augmenting bandwidth have proven complex and costly in time and resources. Recent measurements show that data center traffic is often limited by congestion loss caused by short traffic bursts. Thus an attractive alternative to adding physical bandwidth is to augment wired links with wireless links in the 60 GHz band.

We address two limitations with current 60 GHz wireless proposals. First, 60 GHz wireless links are limited by line-of-sight, and can be blocked by even small [illegible]. Second, even beamforming links leak power, [illegible] interferences will severely limit concurrent [illegible] dense data centers. We propose and evaluate [illegible] primitive for data centers, 3D beamforming [illegible] signals bounce off data center ceilings, [illegible] direct line-of-sight between any two racks. [illegible] We build a small 3D beamforming testbed [illegible] its ability to address both link blockage [illegible] ence, thus improving link range and number [illegible] transmissions in the data center. In addition, [illegible] a simple link scheduler and use traffic simulations [illegible] that these 3D links significantly expand wireless [illegible] compared to their 2D counterparts.

Categories and Subject Descriptors

C.2.1 [Network Architecture and Design]: Wireless communication

Keywords

Data centers, 60 GHz wireless, wireless beamforming

1. INTRODUCTION

Modern distributed applications running in clusters and data centers can run at massive scale, with potentially tens of thousands of servers spread across hundreds of racks. The bandwidth requirements of these applications can range from the relatively modest (e.g., hundreds of Mb/s per server [44]) to substantial (e.g., tens of Gb/s per server [24]). Delivering such bandwidth comes at substantial cost for the requisite switching infrastructure. As a result, a number of recent efforts have investigated techniques for deploying more efficient data center network topologies [11, 14, 20, 21, 22].

While these alternate topologies offer a range of benefits over the current state of the art, we argue that there are a number of inherent challenges with the deployment of any wired network technology. First, any large-scale network consists of multiple "stages," meaning a multiplicative factor number of fibers/wires required for every server in the process of planning, routing, deploying, testing, pairing tens of thousands of fibers at the scale incurs substantial cost (both in capital and penditures). Based on our experience, this ad often delays the operational time of large s or months. Lost machine depreciation and y costs for leveraging additional capacity can ns of dollars in hidden cluster costs. red deployments typically cannot anticipate bandrements to every rack, and thus must distribute mount of fiber to every rack, spot. This means that overdeploy fiber for the "worst case" of communiments rather than the average case, increasing maintenance of networking infrastructure. ed networks are extremely costly and comsentially due to the characteristics of multiologies. For example, doubling the number stage network requires in a Clos topology network [14, 21] elplex to half the existing fiber (impossible since of top of together) or pre-deploying twice the ther requires ches than otherwise required (an fibers are often Similarly, moving a rack later based higher-stage tures running additional fiber to a expensive proposed shown that incurs substantial delays, and sometimes simply intractable. pre-existing cluster is a complex process that incurs sub-

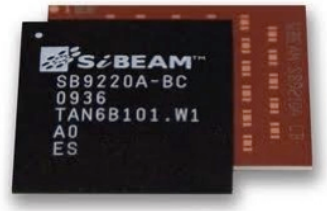
Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
SIGCOMM'12, August 13-17, 2012, Helsinki, Finland.
Copyright 2012 ACM 978-1-4503-1419-0/12/08 ...$15.00.

Deploying 3D Beamforming

- Need clearance between ceiling and top of rack
  - Raised floor to hide wires under racks
  - Cover wires by aluminum-plated ducts
  - Reuse wall or existing metal surface

Deploying 3D Beamforming

- Cost of 60GHz radios
  - Affordable thanks to the low-cost silicon implementation
  - A pair costs ~ $130 (25m+ LOS range)
  - Antenna arrays becoming the cheaper option



Transmitter

Receiver

Mirror Mirror on the Ceiling: Flexible Wireless Links for Data Centers

www.cs.ucsb.edu/~xiazhou/ (on the job market)