60 GHz Flyways:
Adding multi-Gbps wireless links to data centers

Daniel Halperin
Srikanth Kandula, Jitu Padhye
Victor Bahl, David Wetherall
Today’s data center networks are oversubscribed in the core

Perform well in average case with job placement

Bottlenecks in core can be workload “hotspots”
Eliminating oversubscription is expensive

No core hotspots
No job placement
Costly switches
Complex wiring
Our goal: Flyways

To enable a network with an oversubscribed core to act like a non-oversubscribed network by dynamically injecting high-bandwidth links.

Presented by Daniel Halperin
@SIGCOMM2011
Our approach: Wireless Flyways

Low cost
Perform well in most cases with job placement
Dynamically inject links where needed

Presented by Daniel Halperin @SIGCOMM2011
Outline of the rest of this talk

• 60 GHz wireless technology

• Wireless flyways system design

• Evaluation on real data center workloads
60 GHz WIRELESS
60 GHz primer

• **7 GHz** of unlicensed spectrum @60 GHz
• Forthcoming IEEE 802.11ad: 3 channels, bitrates to **6.76 Gbps** at 2.4 GHz
• **Challenge:** a 60 GHz link has **55 dB** (312,000x) worse SNR than a 2.4 GHz link

Presented by Daniel Halperin @SIGCOMM2011
Directionality is crucial

![Graph showing goodput (Gbps) vs distance (m) for different antenna types: Narrow, Wide, Omni.]

Presented by Daniel Halperin
@SIGCOMM2011
60 GHz directional technology

Phased Array

Compact (1 in²)
Electronic steering (µs)

Physical Antenna

Fixed position

Figure 1: Vendor A device, paired with a horn antenna.

Figure 2: Narrow-beam (left) and wide-beam (right) horn antennas for 60 GHz. Note the small size.

These devices interface with removable antennas using the standard 60 GHz WR-15 waveguide. We use two physical directional antennas. A wide-beam horn antenna that was marketed as a 10 dBi/60° gain antenna, and a narrow-beam horn antenna, marketed as a 23 dBi/15° gain antenna. Figure 2 shows how small these antennas are. We measured their radiation patterns in a large, free-space environment, and show the results in Figure 3. Also somewhat expected, the actual gain values differ slightly from manufacturer claims. Thus, we will refer to these two antennas as wide-beam (WB) and narrow-beam (NB), respectively.

3.2 Signal propagation

To study 60 GHz propagation, we conduct experiments using Vendor A devices and NB antennas. We evaluate a few environments in our building: the atrium, which resembles a free-space environment with no walls closer than 40m from either end of the link, and a 1.5 m wide interior hallway, where multiple paths exist. We focused on line-of-sight environments as they come closest to the space on top of racks in our data centers. We set up one sender and one receiver and varied the distance between the two, measuring the signal strength at the receiver at each step.

We present the results in Figure 4. We see that signal strength degrades rapidly with distance. The path exponent is 2, reflecting near-perfect Friis free-space propagation. Prior studies show that line-of-sight links in multi-path environments (waves in the 900–2400 MHz frequency with omni-directional antennas) have path exponents between 1.6–1.8. Thus, we believe that our directional antennas effectively mitigate the impact of multi-path. In fact, even at distances of 25 m, the signal variation (likely due to multi-path) is no more than 3 dB in the atrium and 5 dB in the hallway. This conclusion is supported by prior 60 GHz measurements that showed that directionality at just one side of the link greatly reduced indoor multi-path effects.

These results show that the Friis model is appropriate for indoor line-of-sight 60 GHz links when the endpoints use narrow directional antennas.

3.3 Link stability

The adjective "flaky" is often associated with performance of wireless links, and is a potential concern for using wireless links in the DC. However, the variability notable in WLAN/Wi-Fi deployments comes from device mobility, environmental movement (people, doors opening and closing), temperature changes, and interference. The data center offers a stable, temperature-controlled environment, with infrequent movement of equipment, people, or doors. With devices mounted on top of racks and using directional antennas, the impact of these movements is even less. There is also no external interference in the 60 GHz band. Thus, we expect individual links to be extremely stable.

To verify link stability, we set up a 60 GHz link in our data center using Vendor A devices with NB antennas. We deployed the devices atop two racks, facing each other across...
60 GHz for Flyways

60 GHz links

- Multi-Gbps
- Directional
- Steerable

Flyways must be

- Reliable
- Densely deployed

Presented by Daniel Halperin
@SIGCOMM2011
Directional 60 GHz links are not robust to blockage

Beam Interrupted

![SNR vs Time Graph]

Presented by Daniel Halperin @SIGCOMM2011
A 60 GHz link in a data center
Directional 60 GHz links are stable in a data center

Beam Interrupted  
24h in Data Center

![Graphs showing SNR (dB) over time for Beam Interrupted and 24h in Data Center scenarios.](image)
Measurement-based 802.11ad simulator

• Simulator to *evaluate many concurrent flyways*
  – Channel model from indoor/DC RF measurements
  – Measured 60 GHz antenna patterns
  – Also compared to 8-element 2.4 GHz “Phocus” array

• *Implementation in ns-3*
  – 802.11ad physical layer and protocol
  – TCP and UDP packet simulations
  – Dozens of concurrent multi-Gigabit links

Presented by Daniel Halperin
@SIGCOMM2011
Flyways can be densely deployed

- 160 racks, based on real DC topology
- **Draw random links** until no more can be added
- Ensure **all links meet rate** threshold
- **12-30 links per channel**, depending on rate

Presented by Daniel Halperin
@SIGCOMM2011
Measurement summary

• 60 GHz offers *multi-Gbps, directional, steerable* wireless links with IEEE 802.11ad
• Measurements and simulations show
  – Links are *reliable in data centers*
  – With directionality, links can be *densely deployed*
• Many additional measurements in paper

Presented by Daniel Halperin
@SIGCOMM2011
WIRELESS FLYWAYS
SYSTEM DESIGN

Presented by Daniel Halperin
@SIGCOMM2011
System overview

Jobs placement

DC Scheduler

Data

Demands

Flyway Controller

Presented by Daniel Halperin
@SIGCOMM2011
Flyway controller architecture

Traffic Demands → Wireless links & Rates

→ Compute optimal Flyways set

→ Configure flyways

Presented by Daniel Halperin
@SIGCOMM2011
Flyway controller architecture

Traffic Demands

Wireless links & Rates

Iteratively choose best flyway

Configure flyways

Presented by Daniel Halperin @SIGCOMM2011
Flyway controller architecture

- Wireless links & Rates
- Iteratively choose best flyway
Coordinating devices

Leverage the *wired backbone* to *sidestep issues of coordination*
Orienting antennas

Traditional algorithms search, e.g. sector sweep

Data center topology is known and stable

Presented by Daniel Halperin @SIGCOMM2011
Predicting bitrate

This is hard in multi-path environments

*Directionality alleviates multi-path: SNR lookup table*

[DIRC, SIGCOMM’09]

Use SINR for interference

Presented by Daniel Halperin
@SIGCOMM2011
High-efficiency MAC

Offload small reverse TCP packets to wired network:
+25% wireless goodput
Flyway controller architecture

How to setup links, predict bitrates, and manage interference

Wireless links & Rates

Iteratively choose best flyway

How to select flyways that will improve performance

Presented by Daniel Halperin
@SIGCOMM2011
Selecting flyways: Simple example

Base 10 Gbps network:

• 15 seconds

Presented by Daniel Halperin
@SIGCOMM2011
“Straggler”: Flyway at largest hotspot

Base 10 Gbps network:
• 15 seconds

Straggler:
• 12.2 seconds

Presented by Daniel Halperin
@SIGCOMM2011
“Transit”: Forward traffic on flyway

Base 10 Gbps network:
- 15 seconds

Straggler:
- 12.2 seconds

Transit:
- 11.7 seconds

Presented by Daniel Halperin
@SIGCOMM2011
“Greedy”: Choose faster flyways

Base 10 Gbps network:
• 15 seconds
Straggler:
• 12.2 seconds
Transit:
• 11.7 seconds
Greedy:
• 9.4 seconds

Presented by Daniel Halperin
@SIGCOMM2011
Flyway controller architecture

Traffic Demands

Wireless links & Rates

Iteratively choose best flyway

More?

Update

N

Configure flyways

Presented by Daniel Halperin
@SIGCOMM2011
EVALUATION

Presented by Daniel Halperin
@SIGCOMM2011
Evaluation using real DC workloads

• We studied *four live data centers*
  – Mix of applications (Cosmos, IndexSrv, 2xHPC)
  – Pre-production and production servers

• *76 hours of traces, 114 TB of traffic*
  – Measured application demand
Hypothetical demand matrix needs full-bisection

Figure 12(a) depicts a heat map of the demands between pairs of the racks. The color palette is on a logarithmic scale, i.e., black indicates zero demand, and white indicates a very high demand. The majority of the ToR pairs are yellow, i.e., less than half as loaded as the most loaded link. A few trends are apparent. First, only a few ToR pairs are hot, i.e., send or receive a large volume of traffic (darker dots). The bulk of the ToR pairs are yellow, i.e., less than half as loaded as the most loaded link.
Real traces have localized hotspots

Very few hotspots!

Affect only a few racks

Source Rack

Destination Rack

Presented by Daniel Halperin
@SIGCOMM2011
Evaluation setup

• Evaluated 60 GHz flyways improvements on *real demand matrices* in an ns-3 *topology based on real DC layout*

• **Metric:** Completion time of Demands (CTD)
  – Relative to non-oversubscribed network
  – CTD of 1 ➞ *same as non-oversubscribed*
  – CTD of 2 ➞ *same as 1:2 oversubscribed*
1 flyway device / node

CDF over Demand Matrices

Smallest possible CTD for 50% of traces

CTD < 1.5 for >80% of traces

Presented by Daniel Halperin
@SIGCOMM2011
Incremental benefit of strategies

CDF over Demand Matrices

Presented by Daniel Halperin
@SIGCOMM2011
1-3 devices / node

CDF over Demand Matrices

Presented by Daniel Halperin
@SIGCOMM2011
Conclusions

• 60 GHz flyways can substantially improve performance in oversubscribed DC

• Traffic redirection crucial for practical benefit in real workloads

• Novel techniques leverage wired backbone to dramatically simplify and speed hybrid system

Read more: http://r.halper.in/paper/flyways_sigcomm11

Presented by Daniel Halperin
@SIGCOMM2011