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

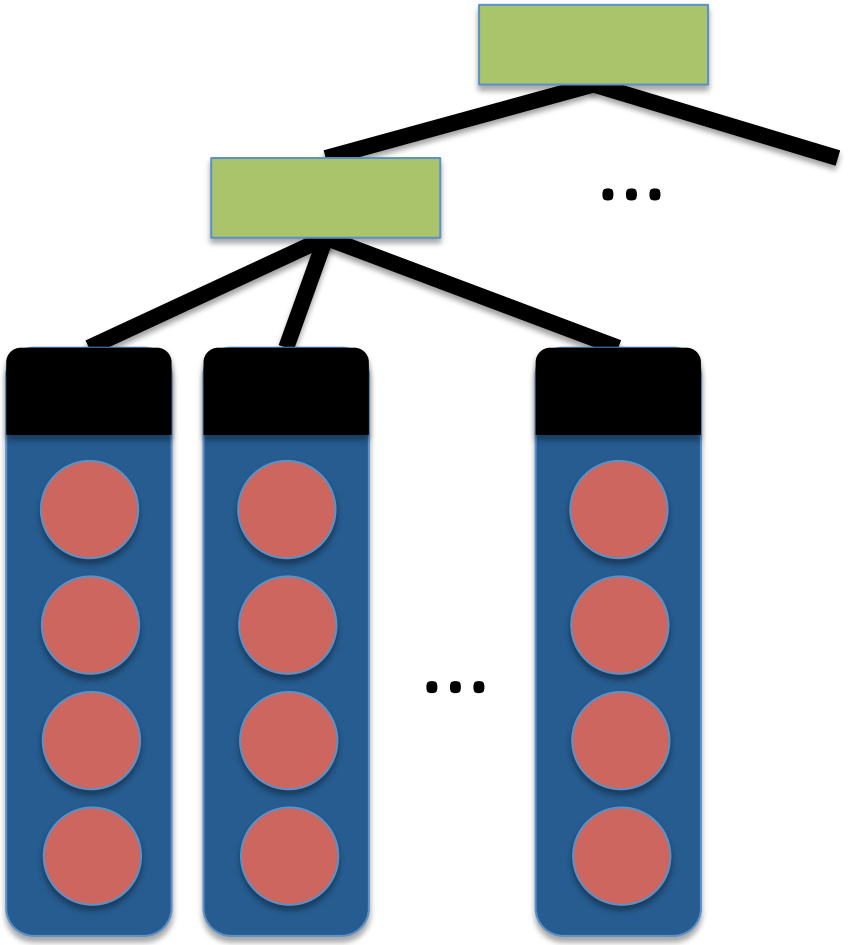
Daniel Halperin

Srikanth Kandula, Jitu Padhye
Victor Bahl, David Wetherall

Microsoft® Research

University of Washington Computer Science & Engineering

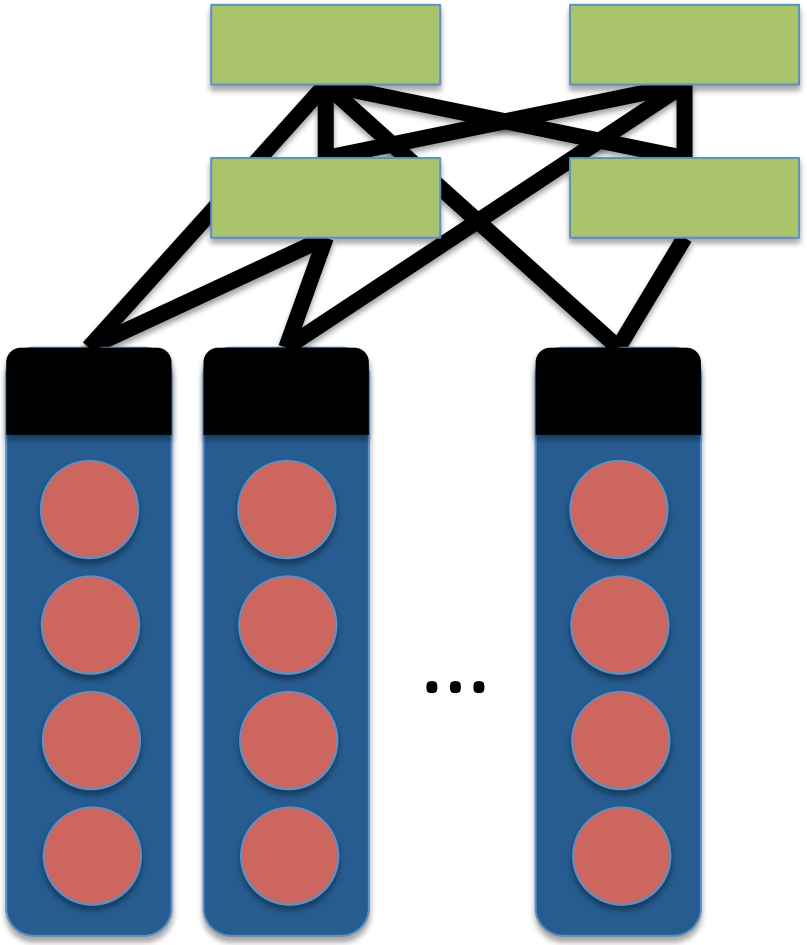
Today's data center networks are oversubscribed in the core



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

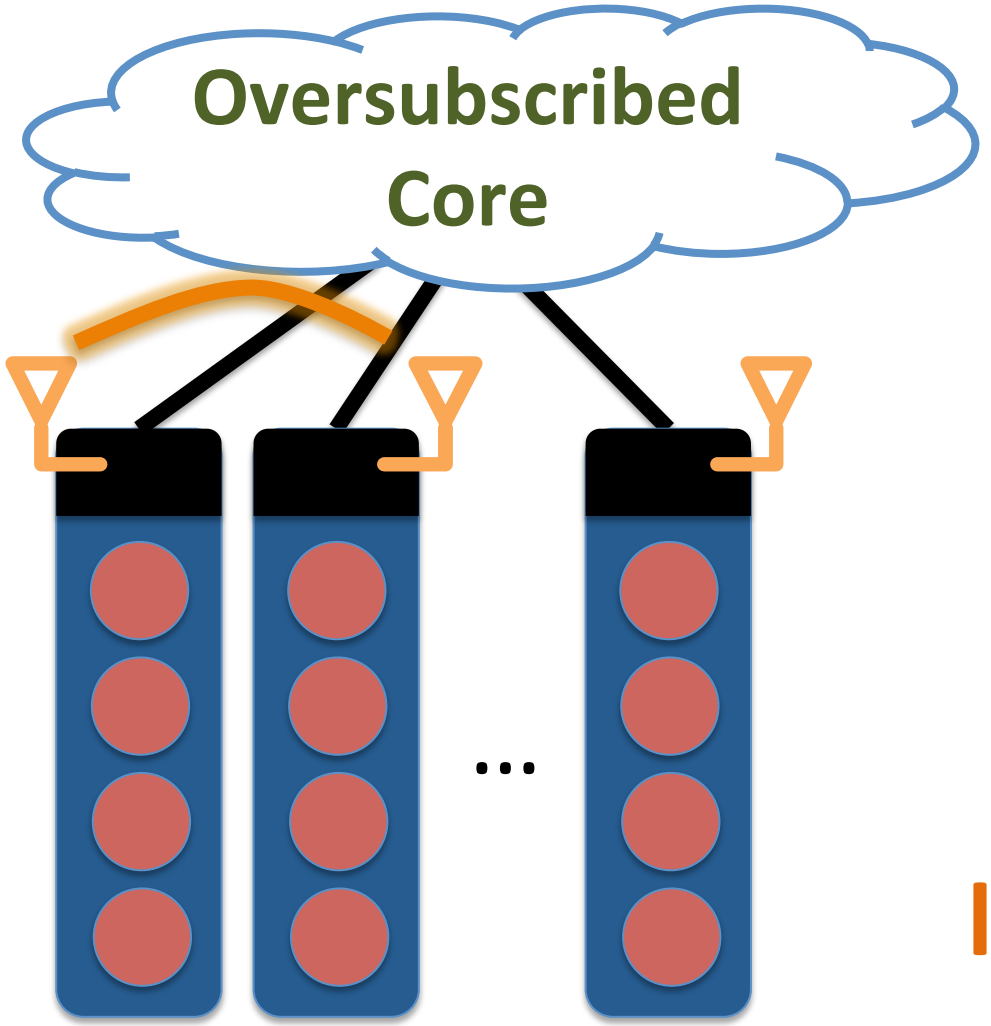
Costly switches

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.

Our approach: Wireless Flyways

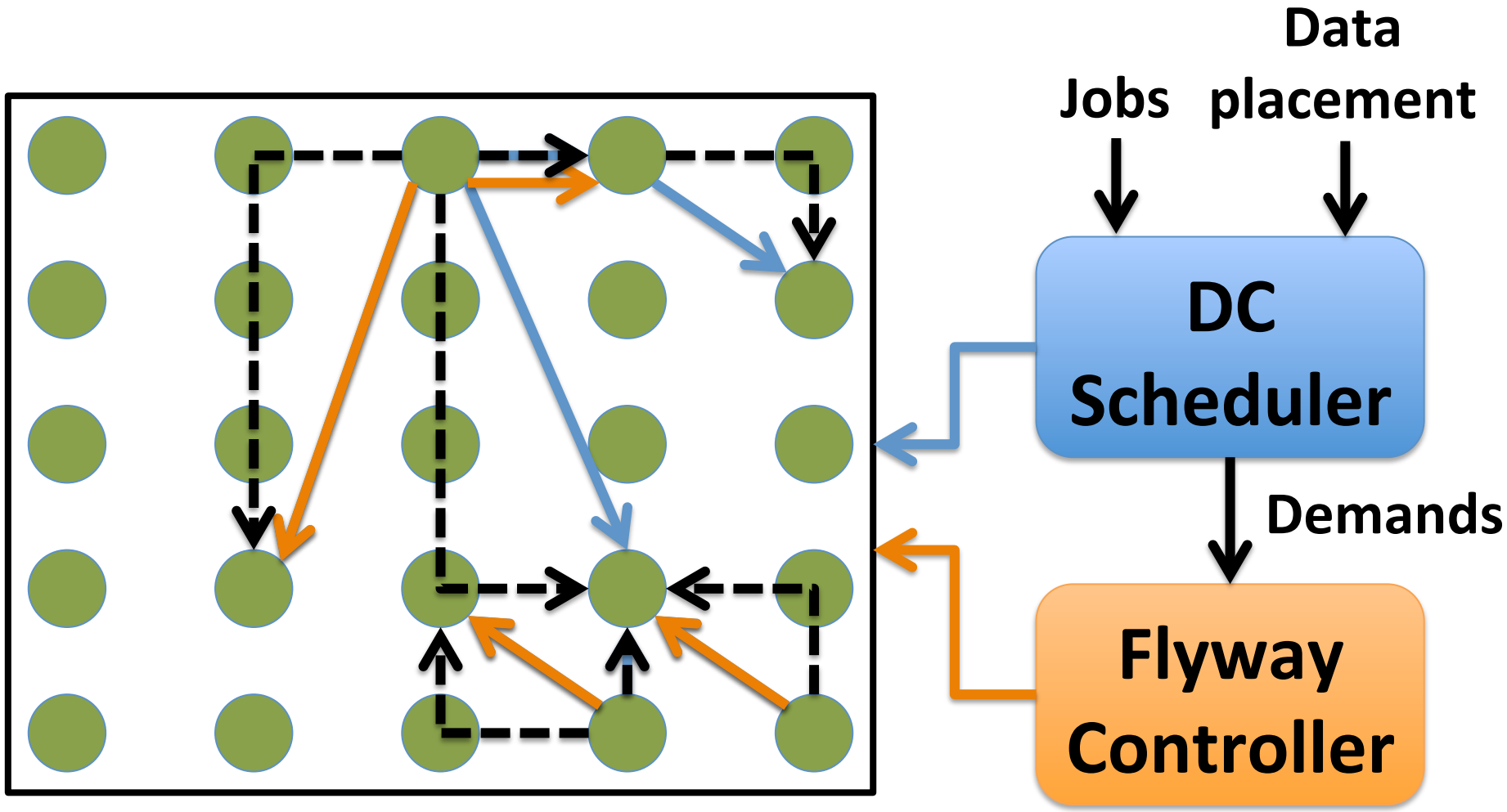


Low cost

Perform well in most cases with job placement

Dynamically inject links where needed

System overview



Outline of the rest of this talk

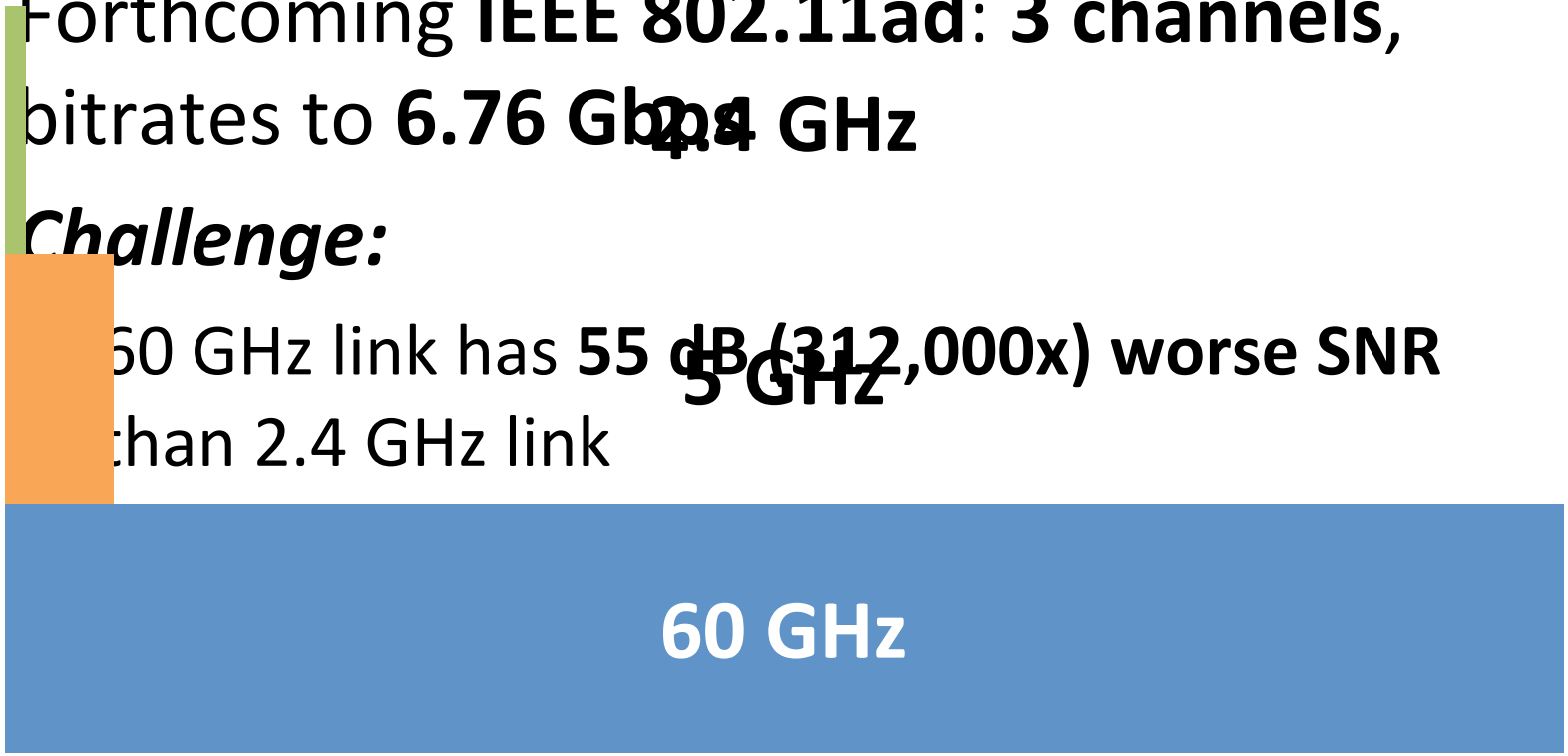
- 60 GHz ***wireless technology***
- Wireless ***flyways system design***
- ***Evaluation*** on ***real data center workloads***

60 GHz WIRELESS

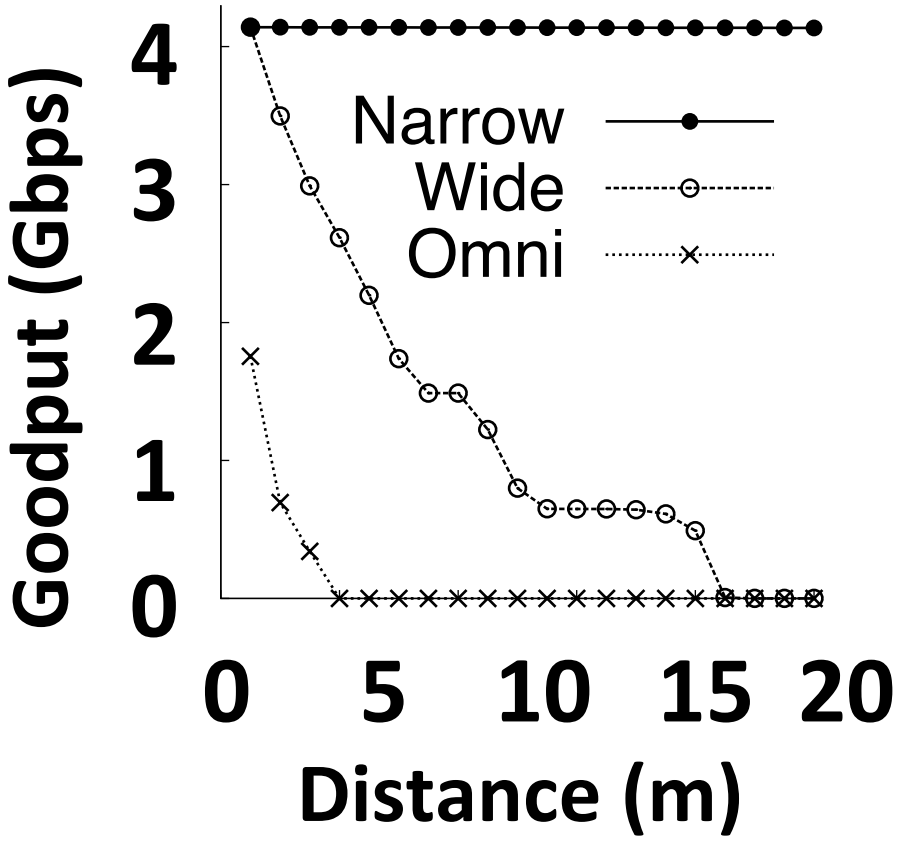
Presented by Daniel Halperin
@SIGCOMM2011

60 GHz primer

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

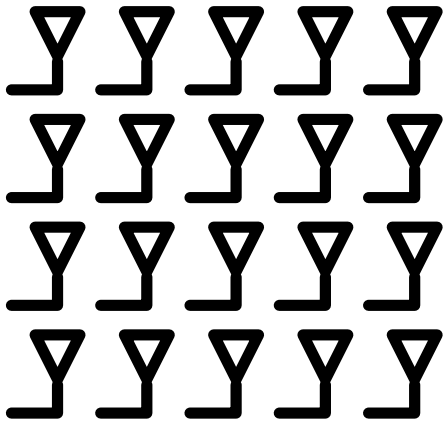


Directionality is crucial

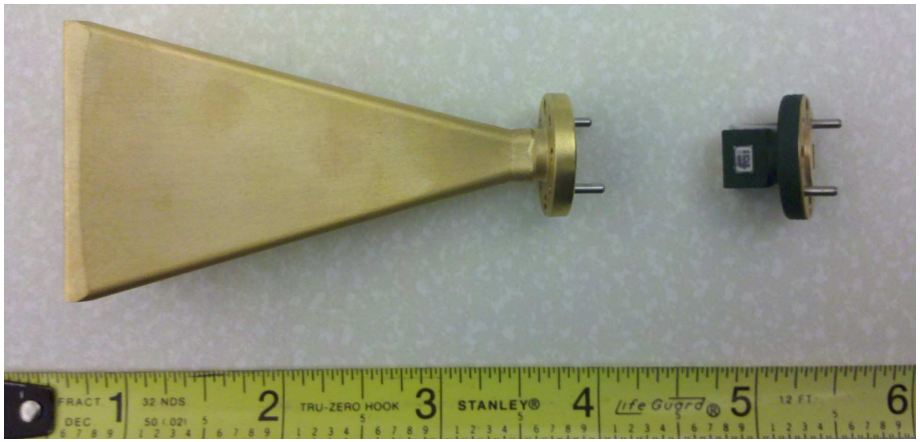


60 GHz directional technology

Phased Array



Physical Antenna



Compact (1 in^2)
Electronic steering (µs)

Fixed position

60 GHz for Flyways

60 GHz links

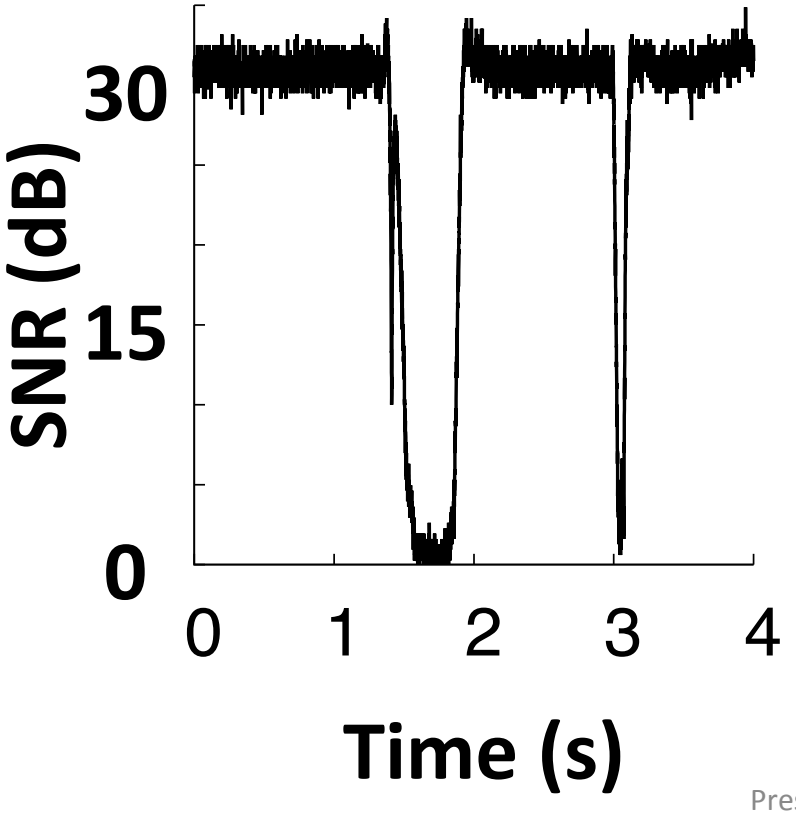
- ***Multi-Gbps***
- ***Directional***
- ***Steerable***

Flyways must be

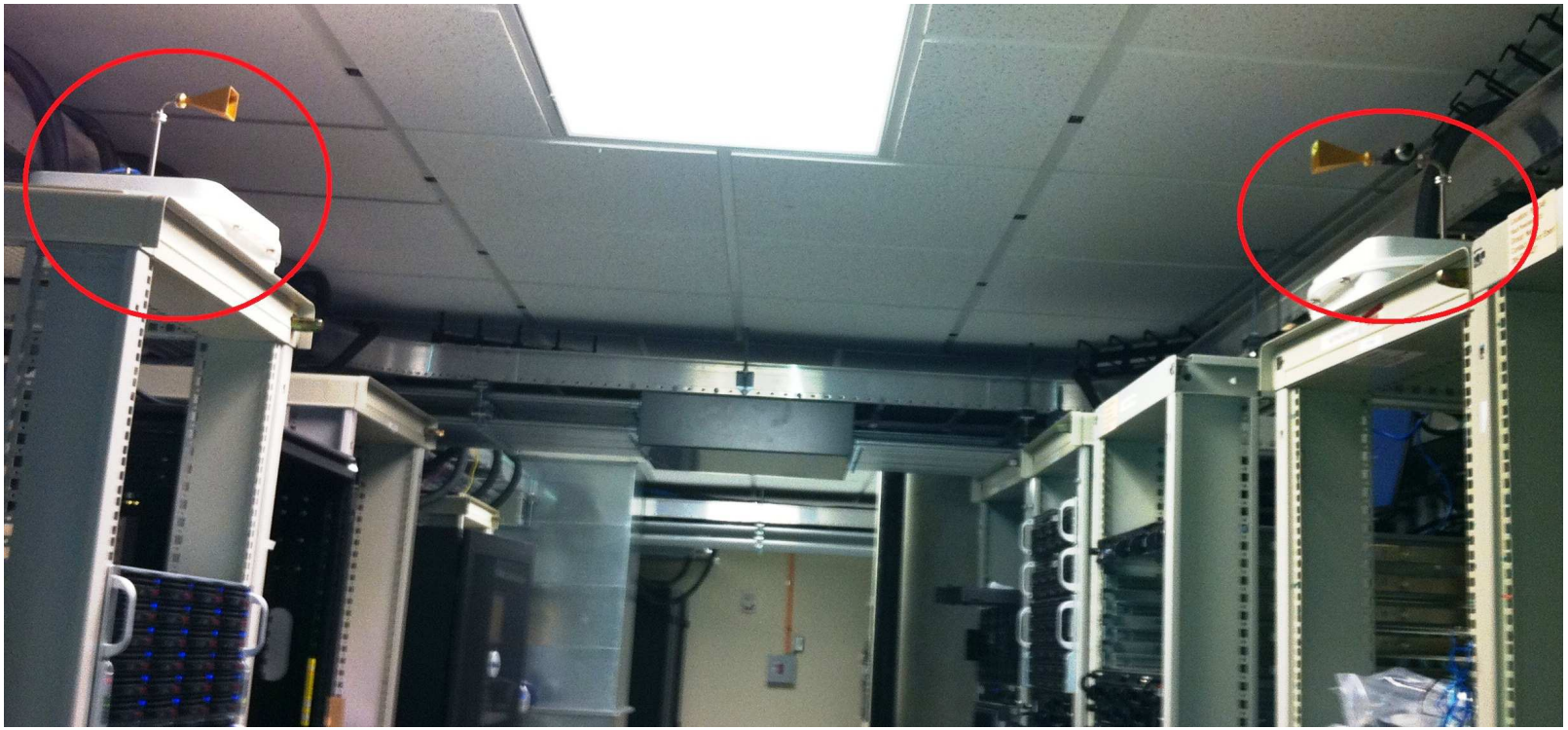
- ***Reliable***
- ***Densely deployed***

Directional 60 GHz links are not robust to blockage

Beam Interrupted

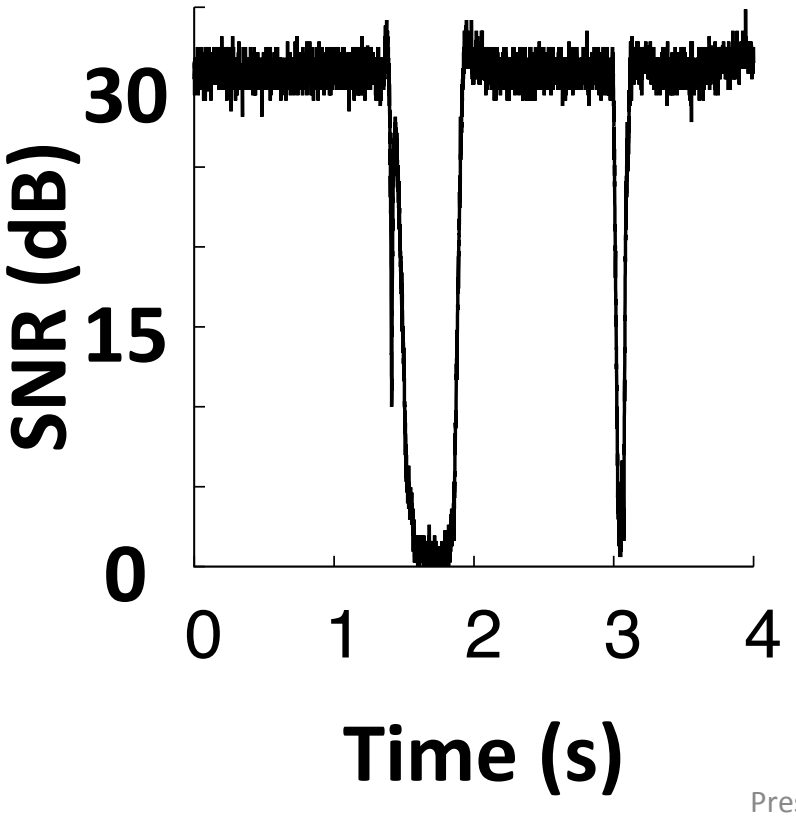


A 60 GHz link in a data center

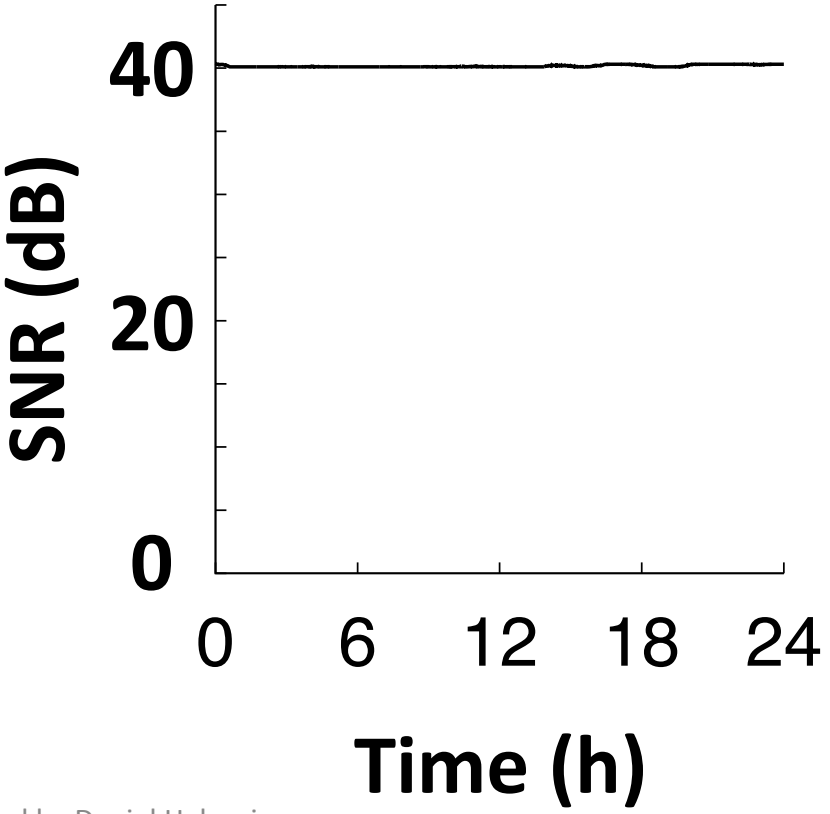


Directional 60 GHz links are stable in a data center

Beam Interrupted



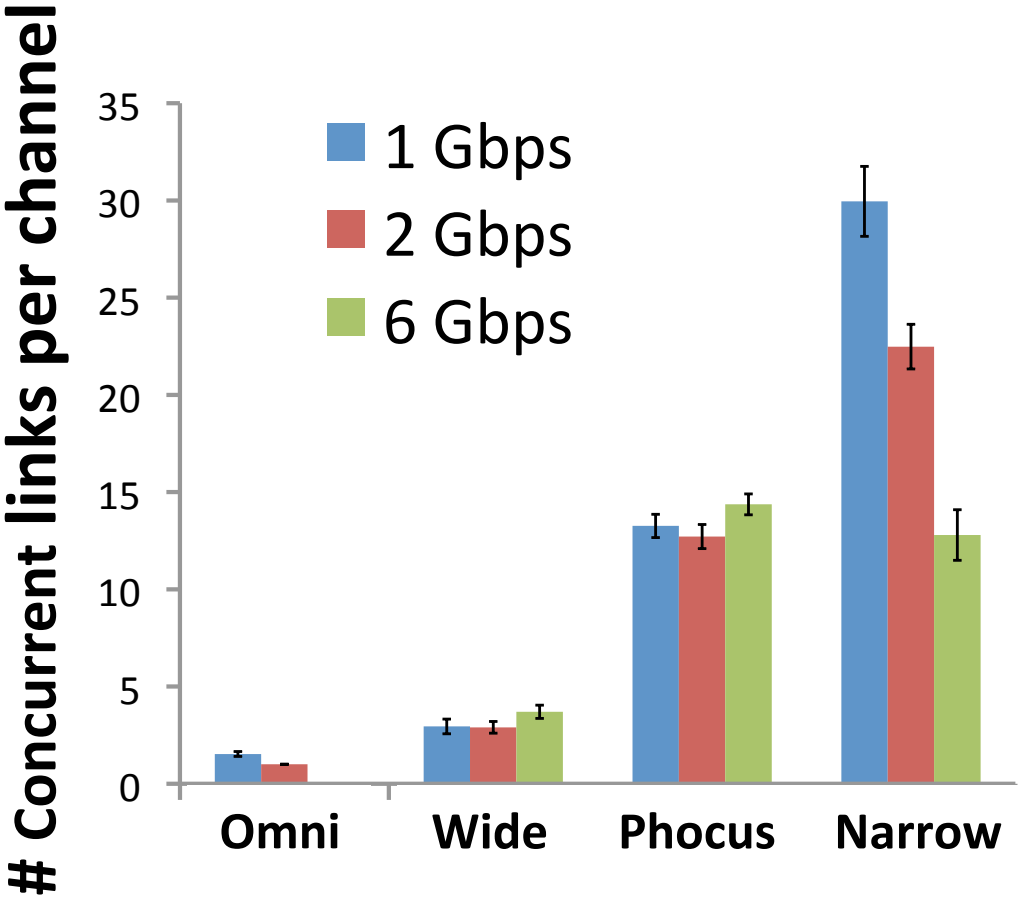
24h in Data Center



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

Flyways can be densely deployed



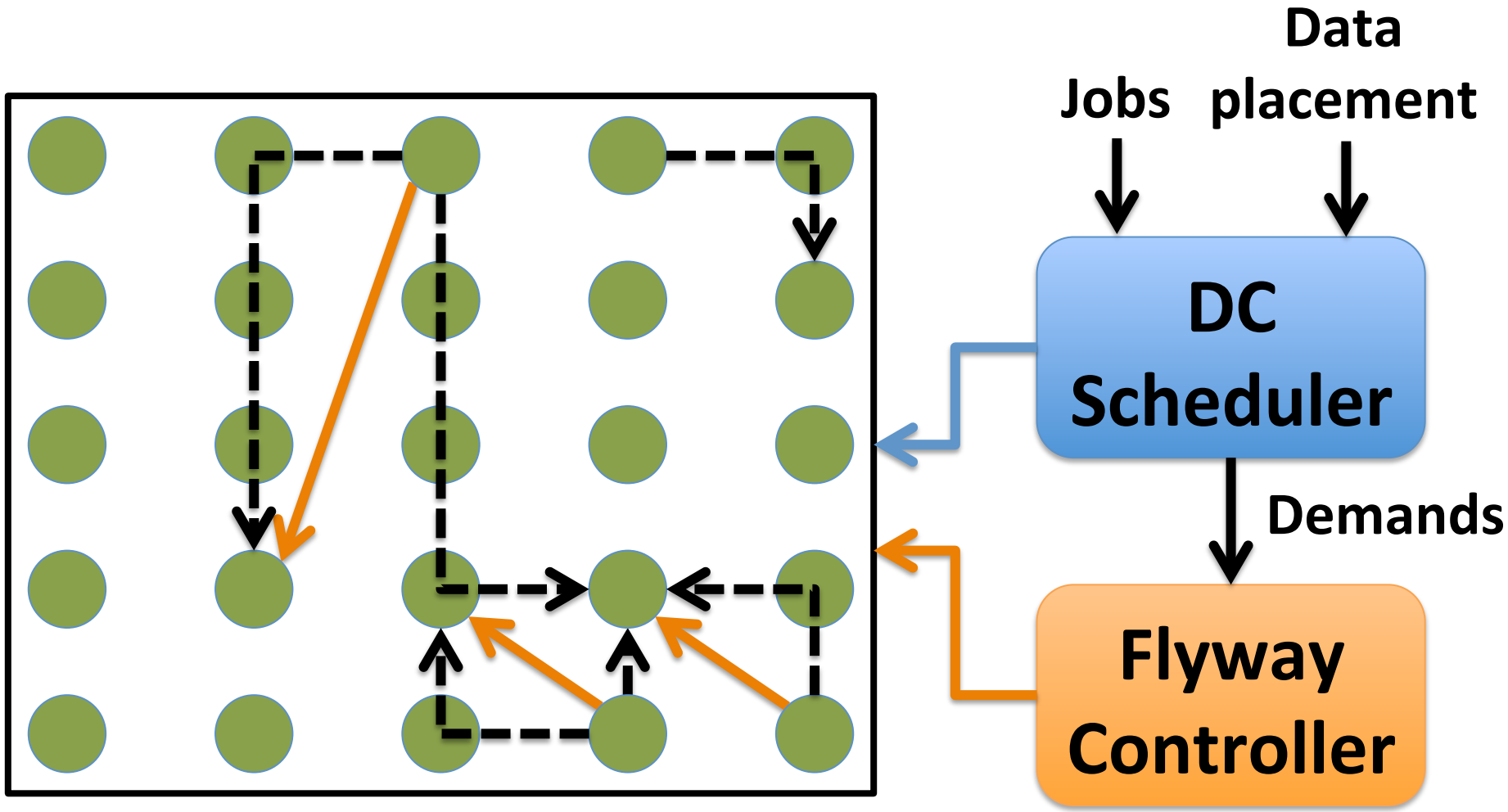
- 160 racks, based on real DC topology
- ***Draw random links*** until no more can be added
- Ensure ***all links meet rate*** threshold
- **12-30 links per channel**, depending on rate

Measurement summary

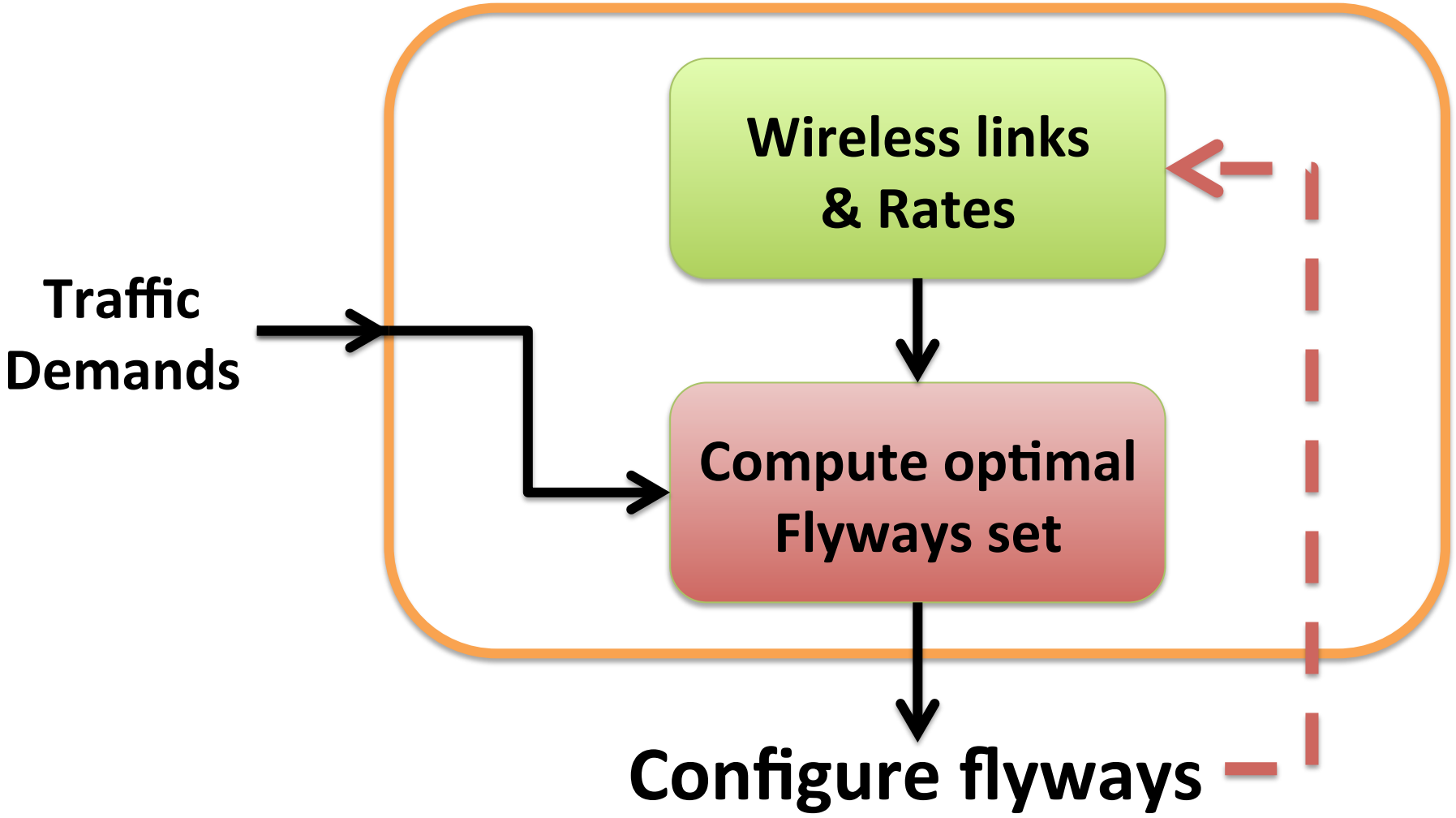
- 60 GHz offers ***multi-Gbps***, ***directional***, ***steerable*** wireless links with IEEE 802.11ad
- Measurements and simulations show
  - Links are ***reliable in data centers***
  - With directionality, links can be ***densely deployed***
- Many additional measurements in paper

WIRELESS FLYWAYS SYSTEM DESIGN

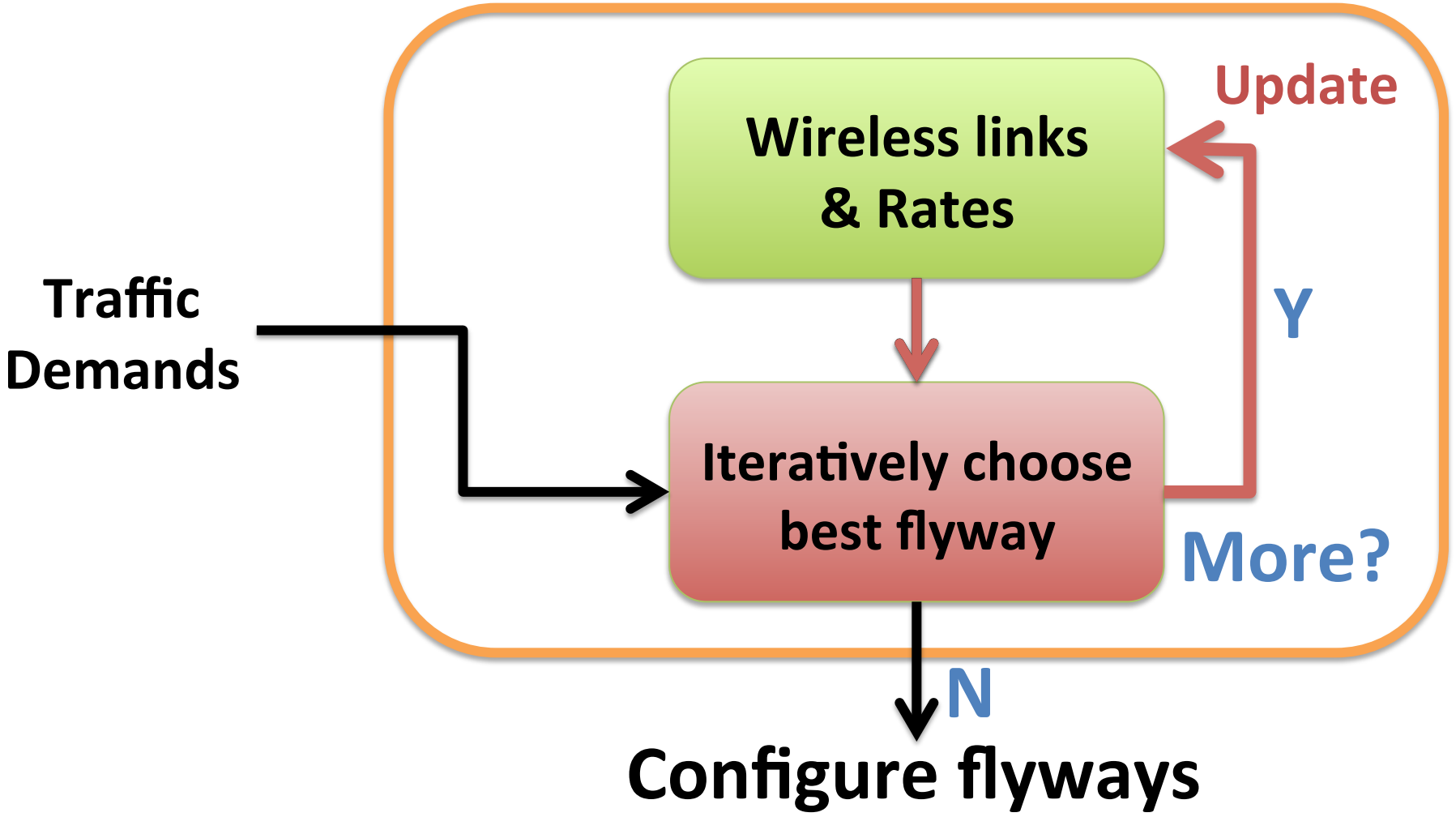
System overview



Flyway controller architecture



Flyway controller architecture



Flyway controller architecture

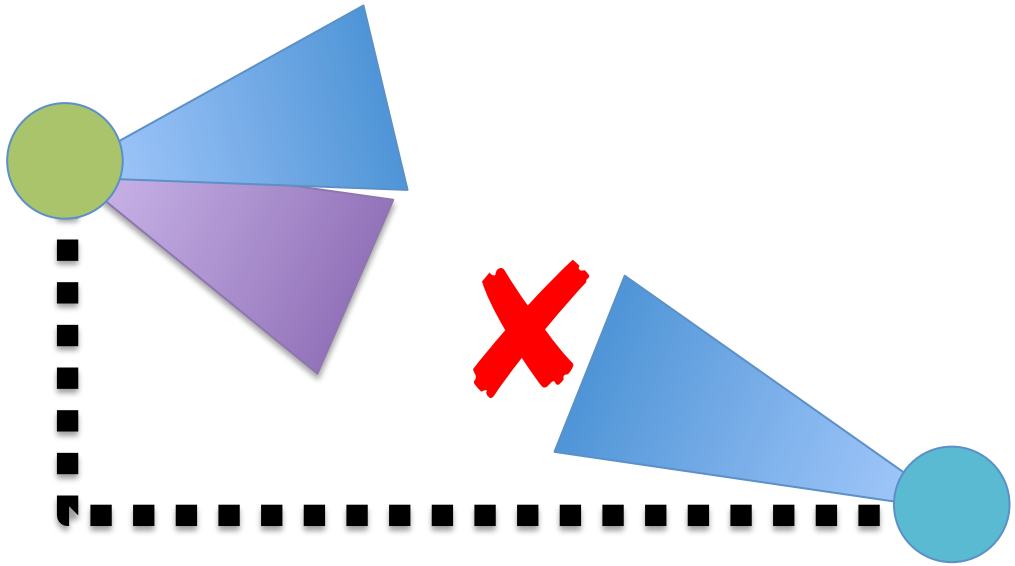
Wireless links & Rates

***nks*,**
***s*,**
erence

Iteratively choose best flyway

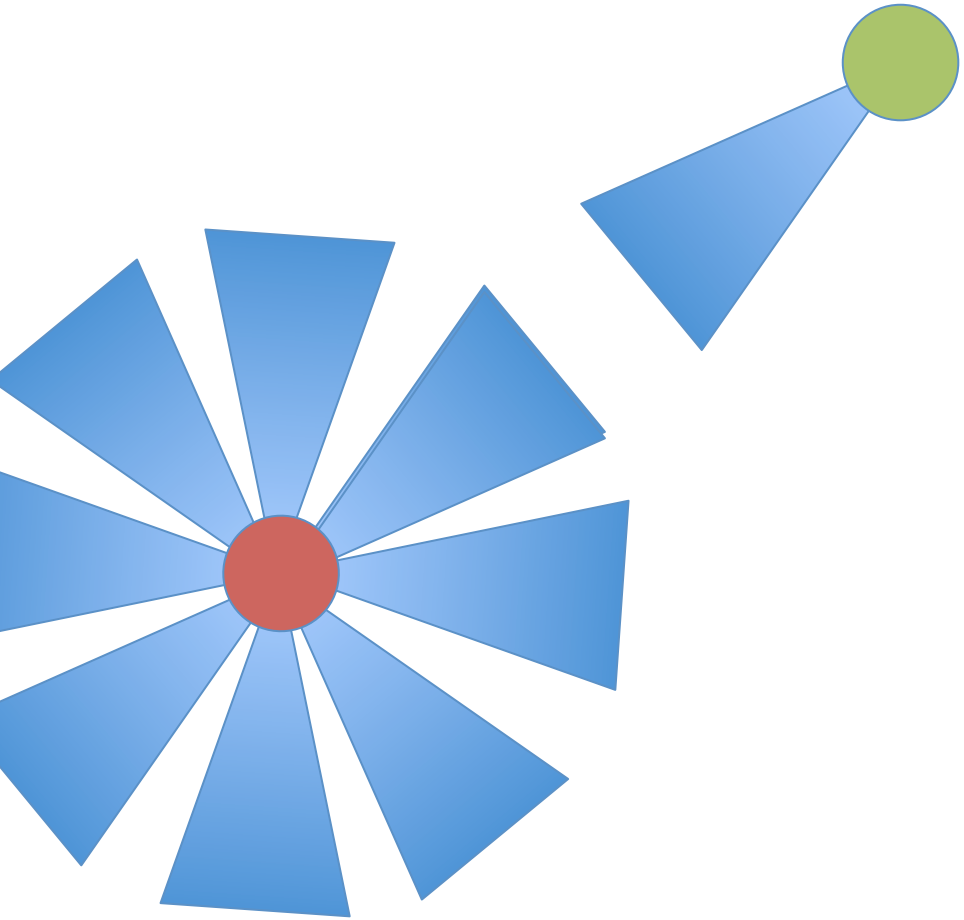
H **s** that will
ance

Coordinating devices



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

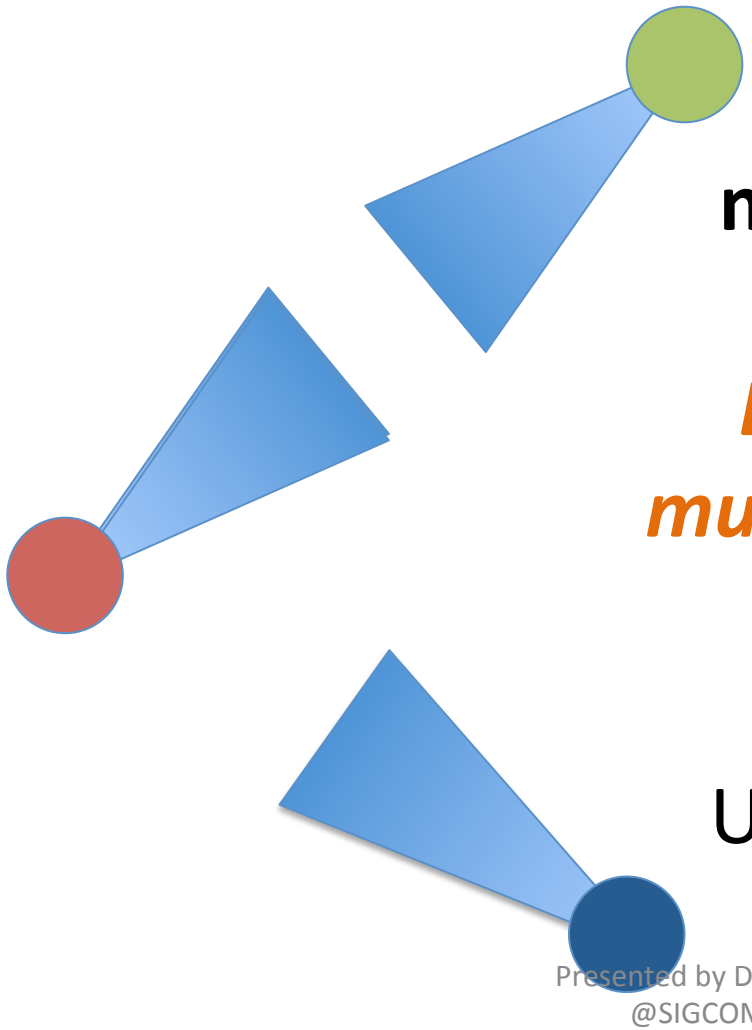
Orienting antennas



Traditional algorithms *search*, e.g. **sector sweep**

Data center topology is ***known*** *and* ***stable***

Predicting bitrate

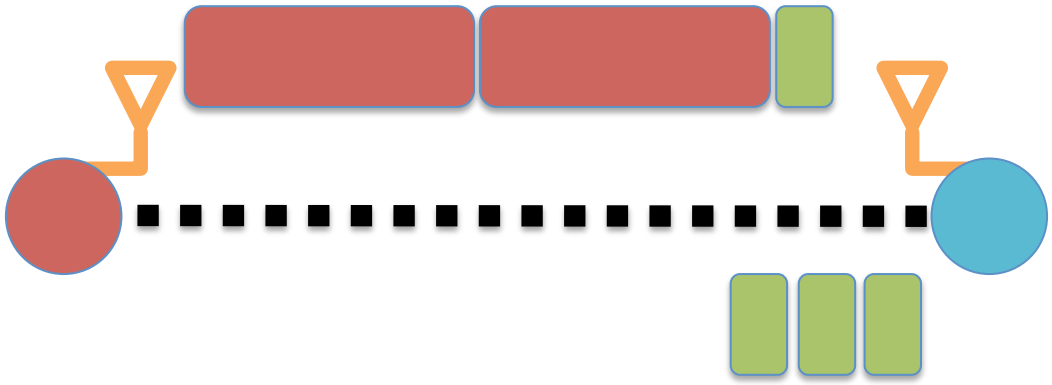


This is **hard in multi-path** environments

Directionality alleviates multi-path: SNR lookup table
[DIRC, SIGCOMM'09]

Use **SINR for interference**

High-efficiency MAC



Offload small reverse TCP packets to wired network:

+25% wireless goodput

Flyway controller architecture

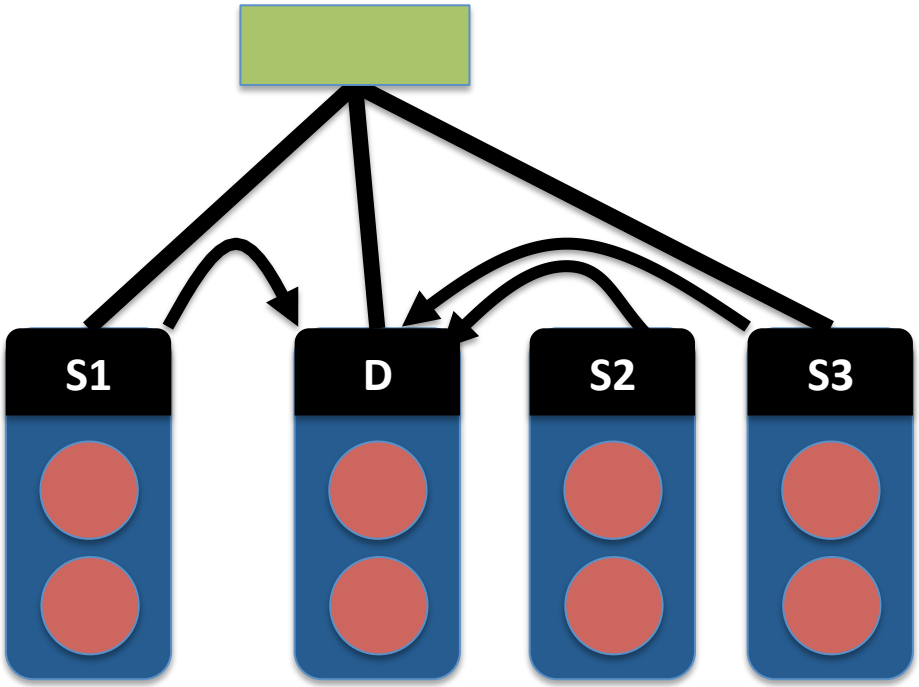
Wireless links & Rates

How to ***setup links***, ***predict bitrates***, and ***manage interference***

Iteratively choose best flyway

How to ***select flyways*** that will ***improve performance***

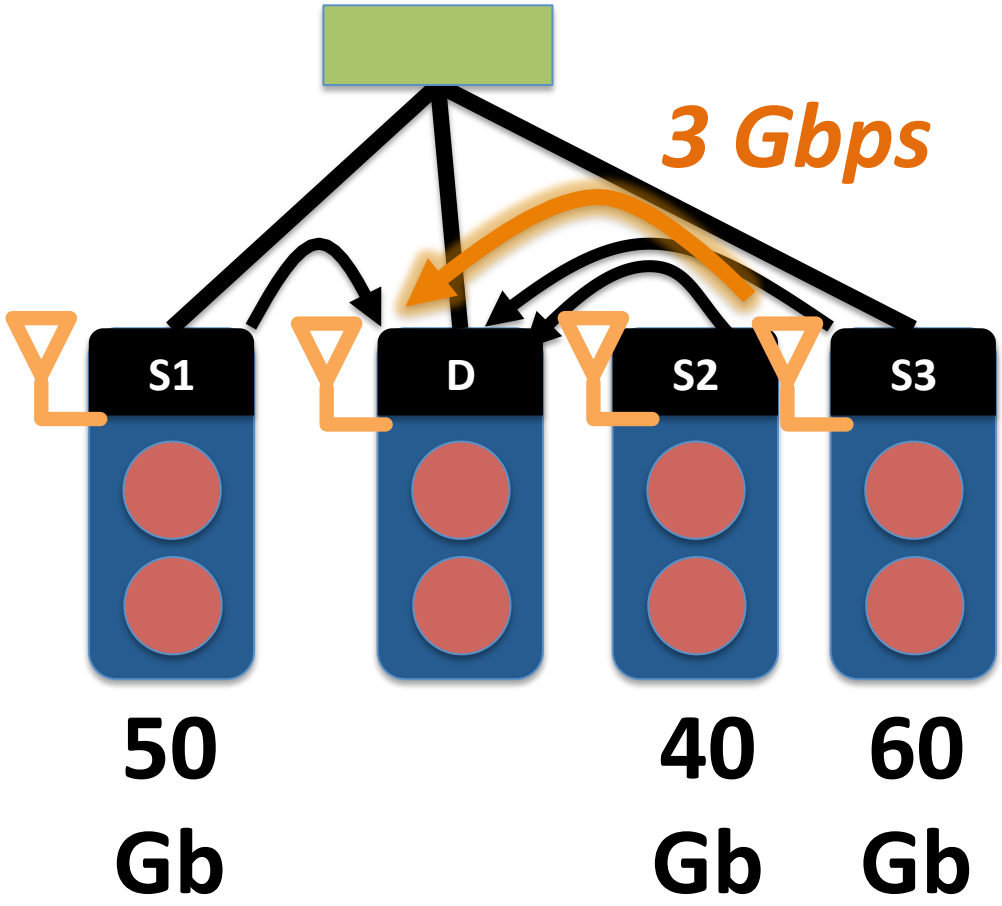
Selecting flyways: Simple example



Base 10 Gbps network:

- **15 seconds**

“Straggler”: Flyway at largest hotspot



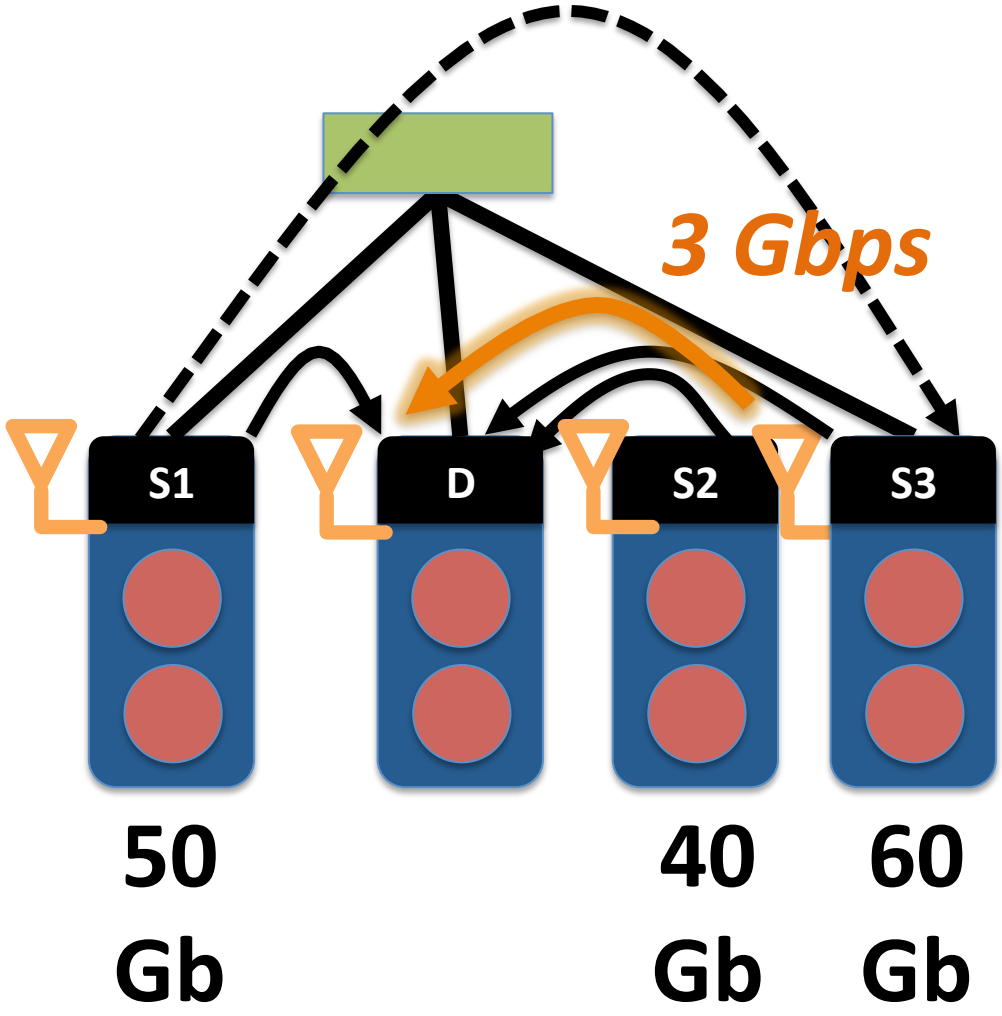
Base 10 Gbps network:

- 15 seconds

Straggler:

- **12.2 seconds**

"Transit": Forward traffic on flyway



Base 10 Gbps network:

- 15 seconds

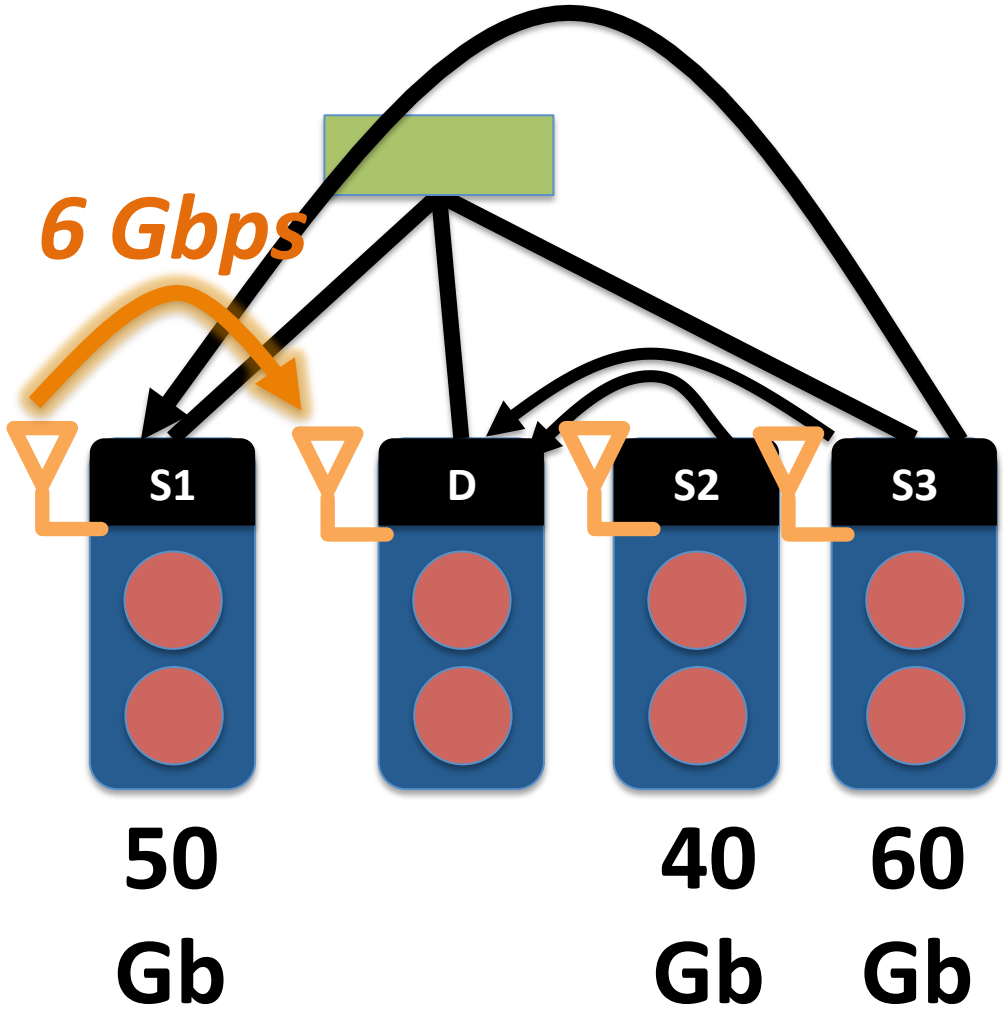
Straggler:

- 12.2 seconds

Transit:

- **11.7 seconds**

"Greedy": Choose faster flyways



Base 10 Gbps network:

- 15 seconds

Straggler:

- 12.2 seconds

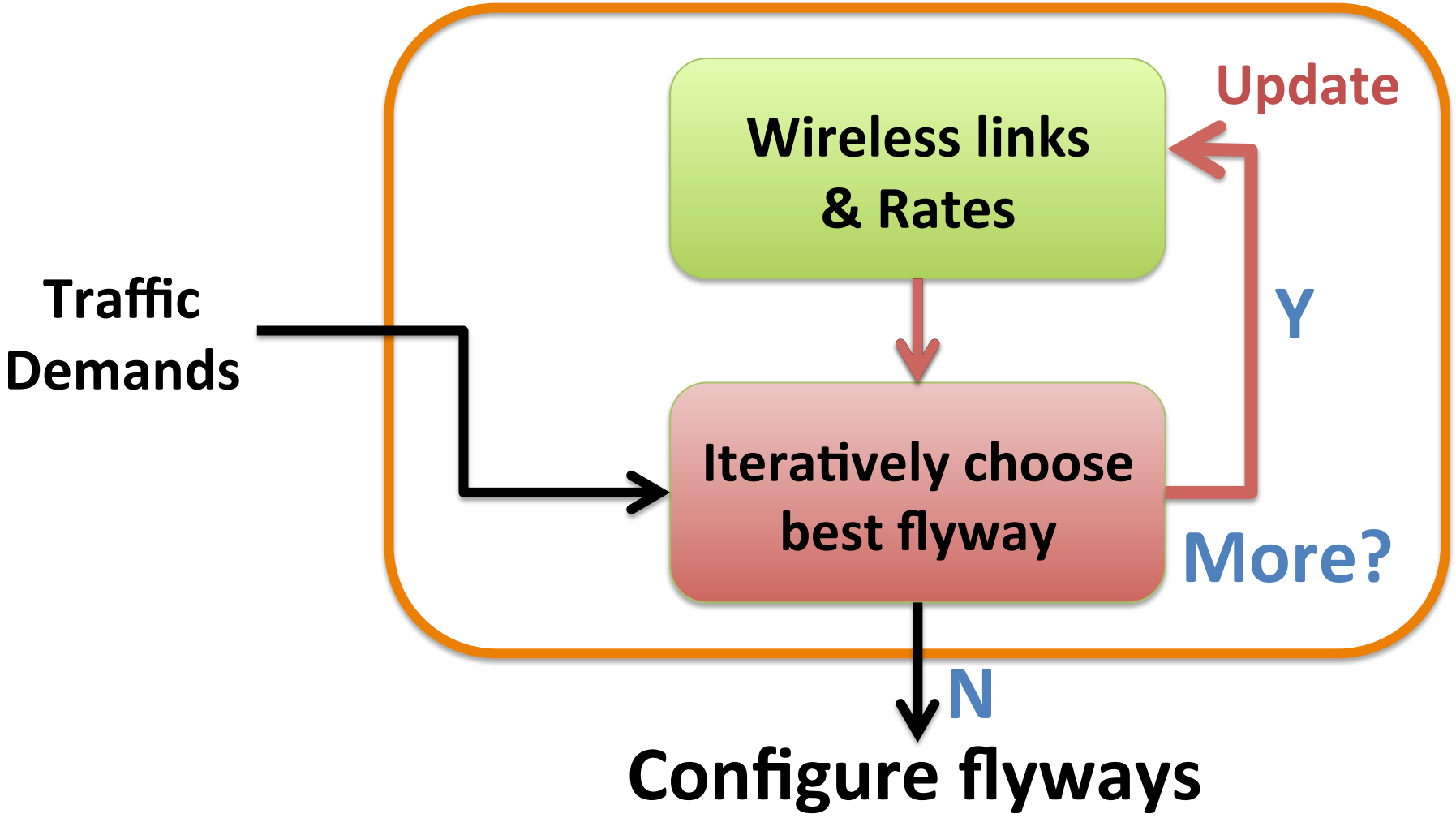
Transit:

- 11.7 seconds

Greedy:

- **9.4 seconds**

Flyway controller architecture

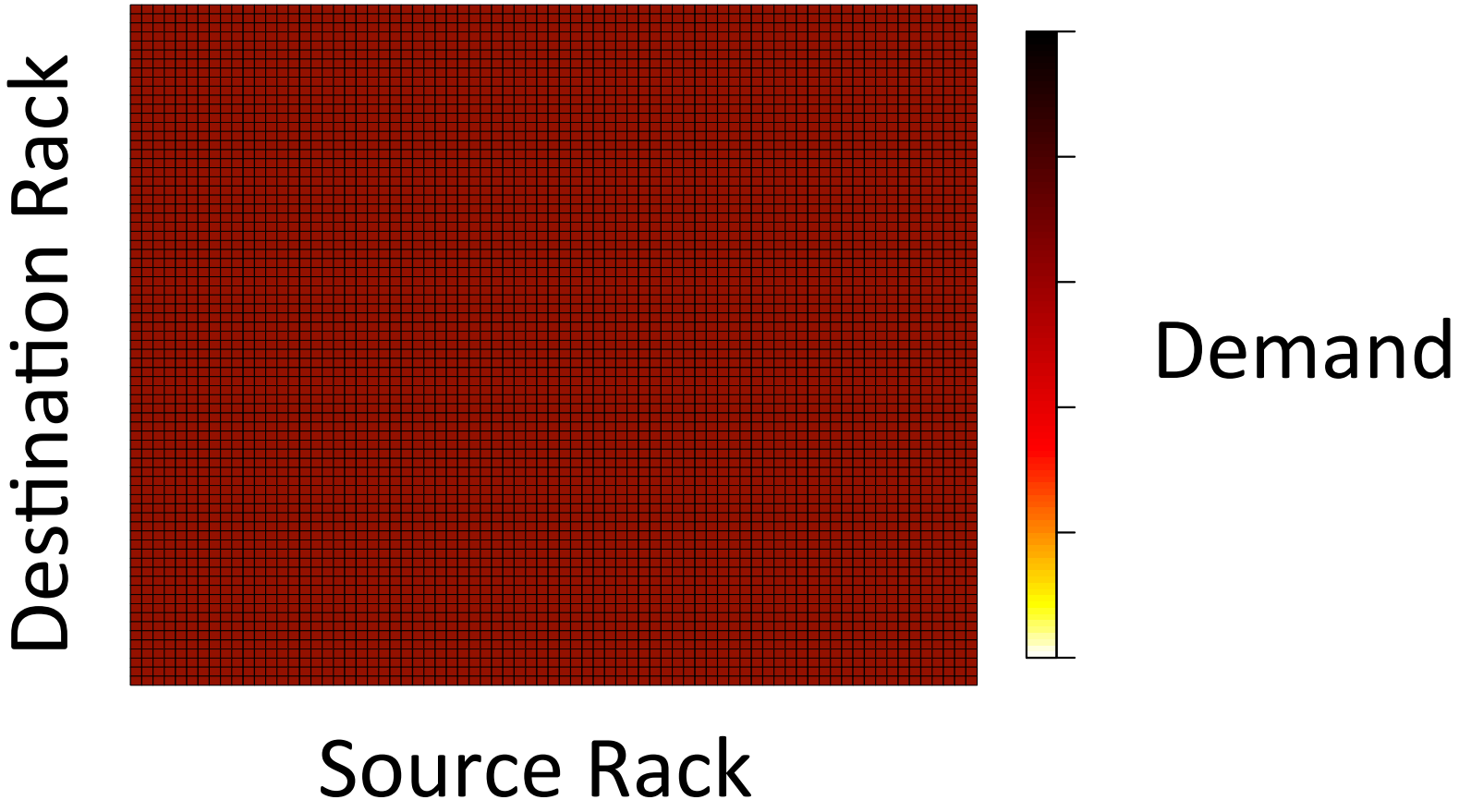


EVALUATION

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

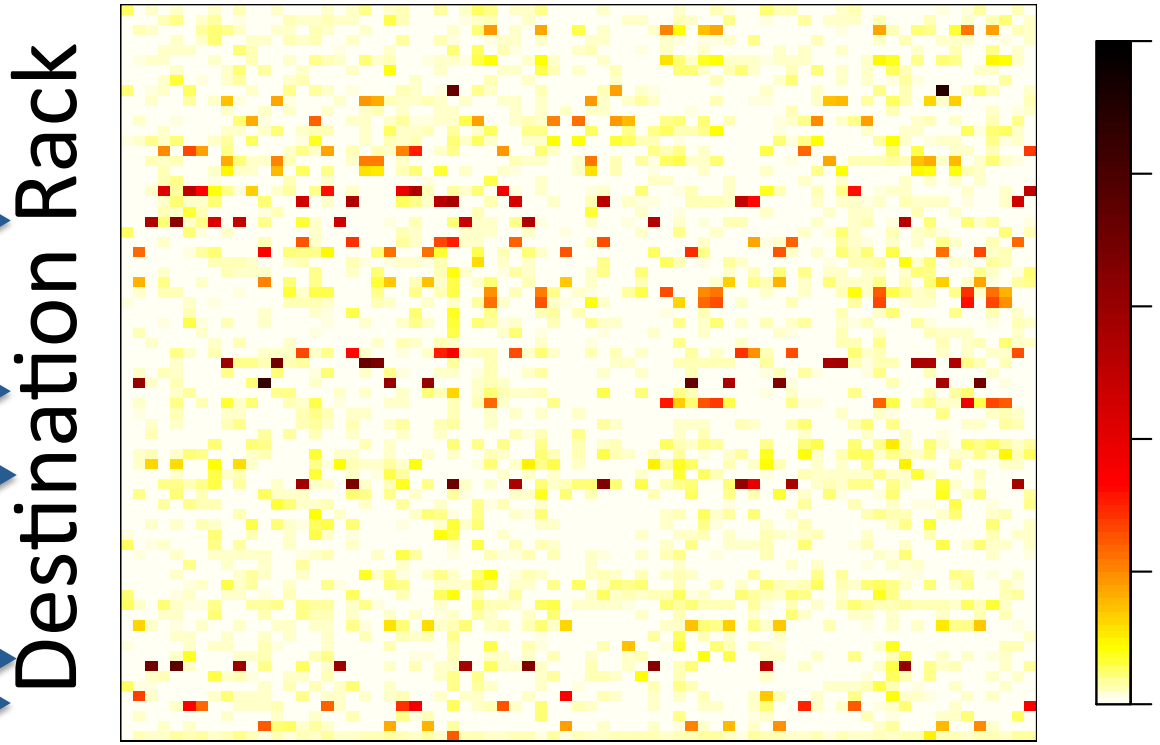
Hypothetical demand matrix needs full-bisection



Real traces have localized hotspots

Very few hotspots!

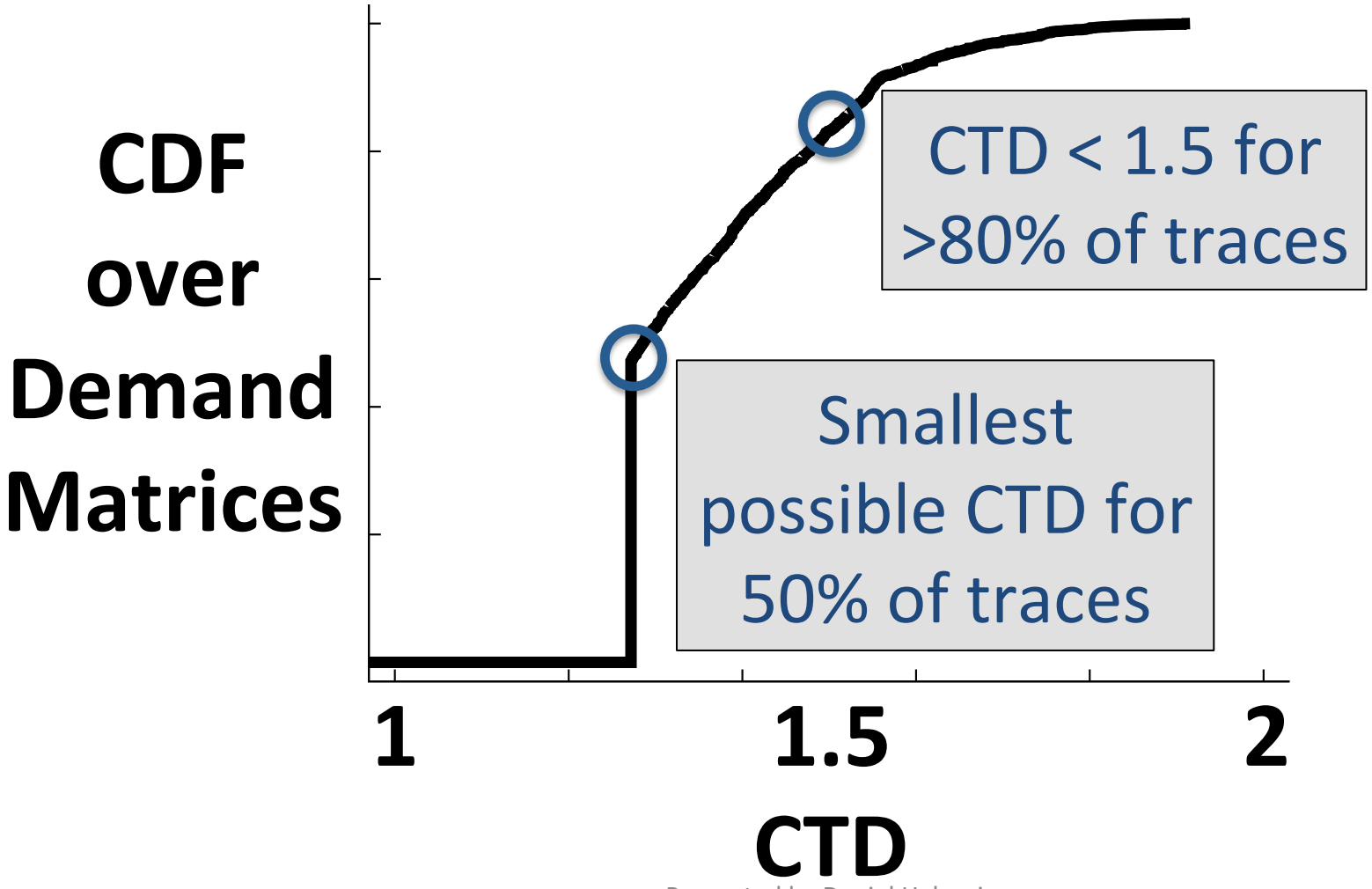
Affect only a few racks



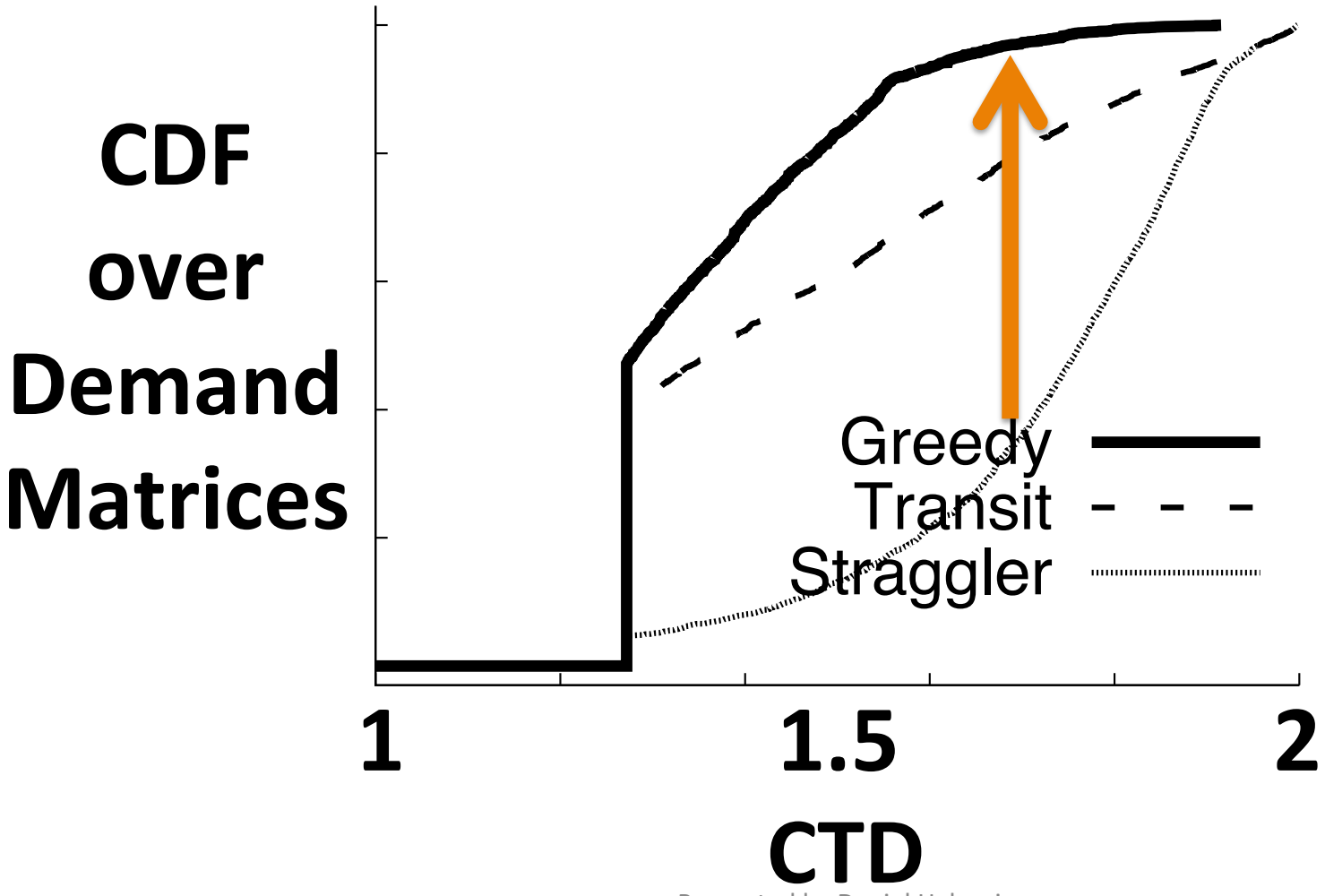
Evaluation setup

- 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***

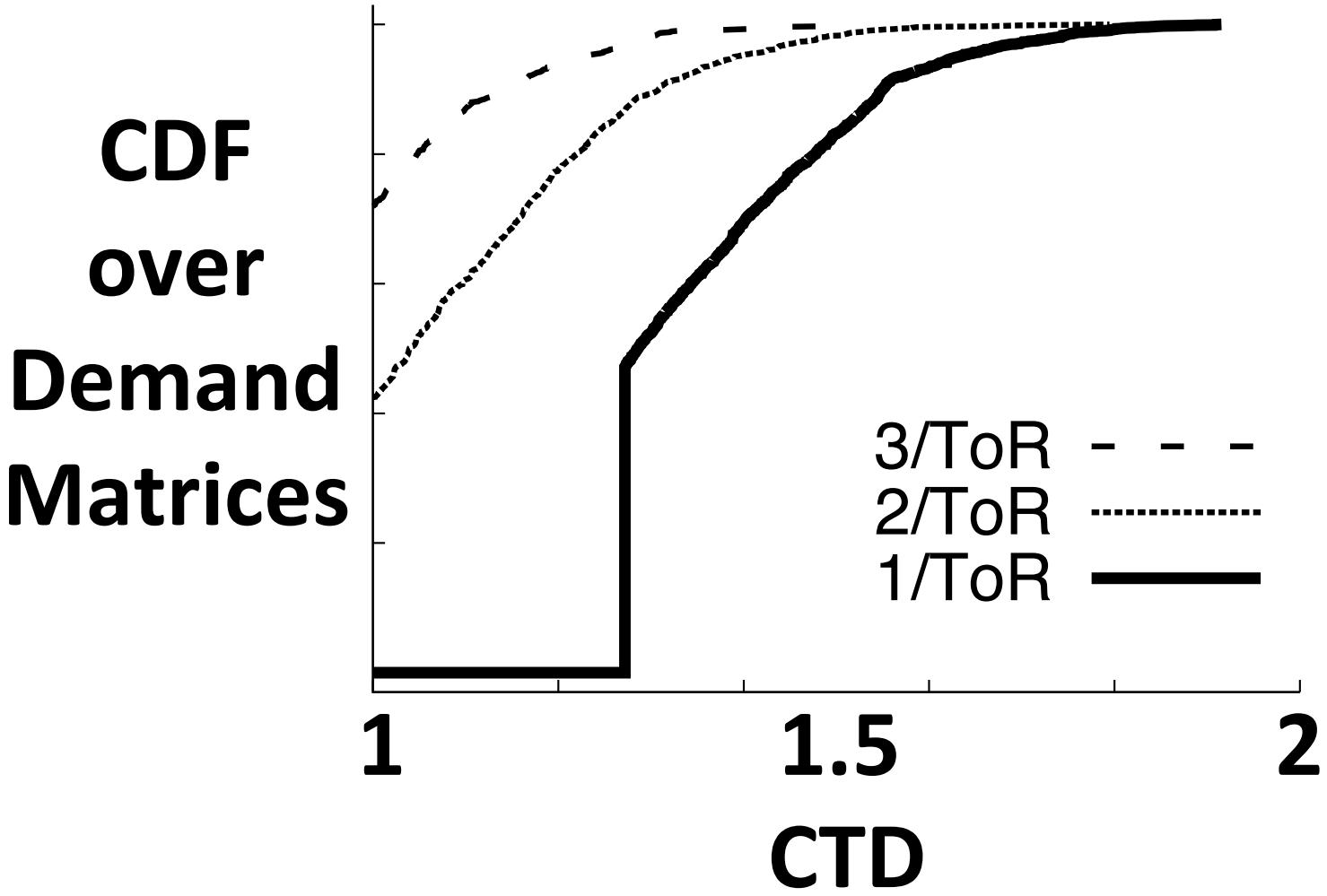
1 flyway device / node



Incremental benefit of strategies



1-3 devices / node



Conclusions

- **60 GHz flyways** can substantially ***improve performance in oversubscribed DC***
- ***Traffic indirection 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