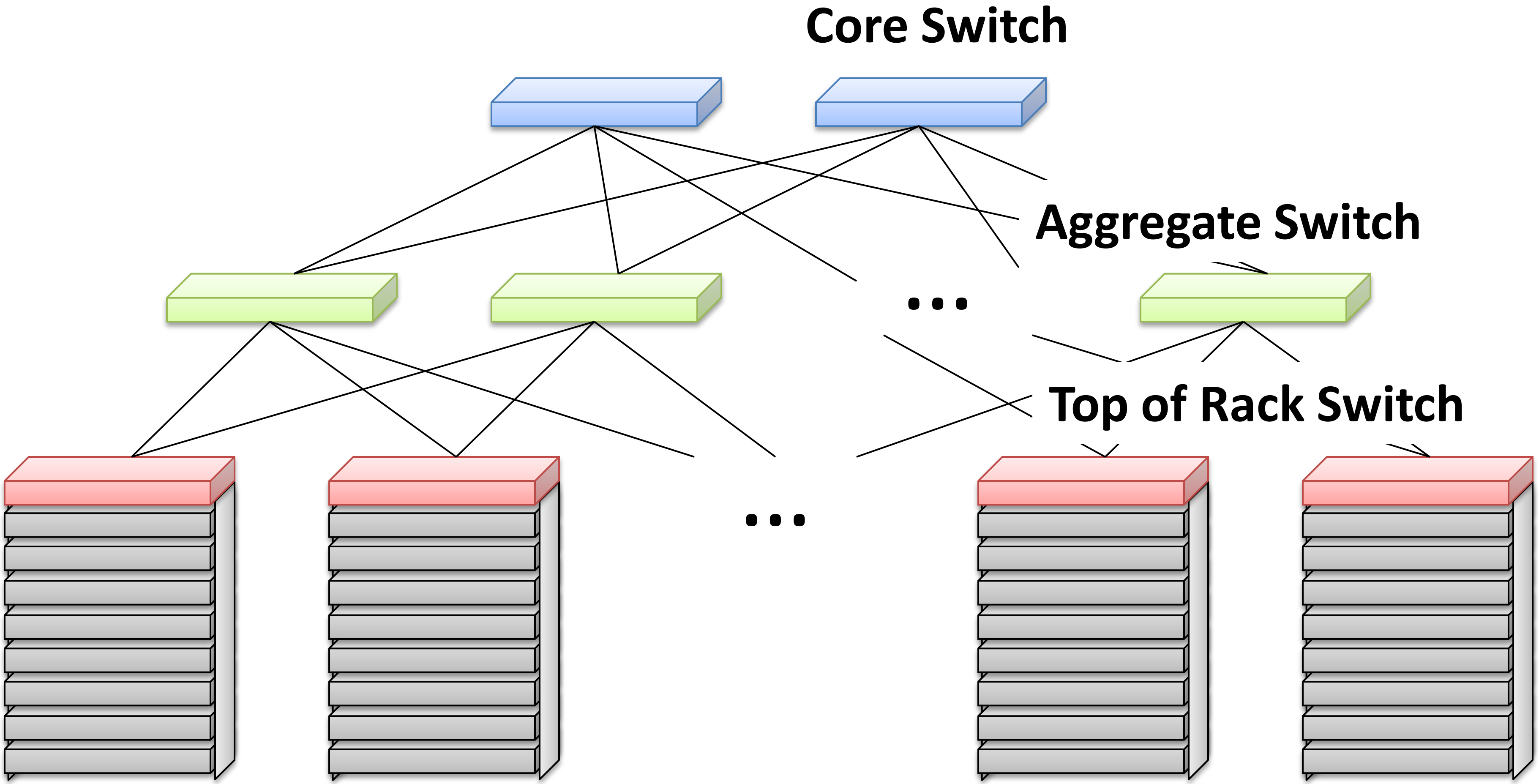
ACM/IEEE ANCS, Oct 29, 2012

On the Feasibility of Completely Wireless Datacenters

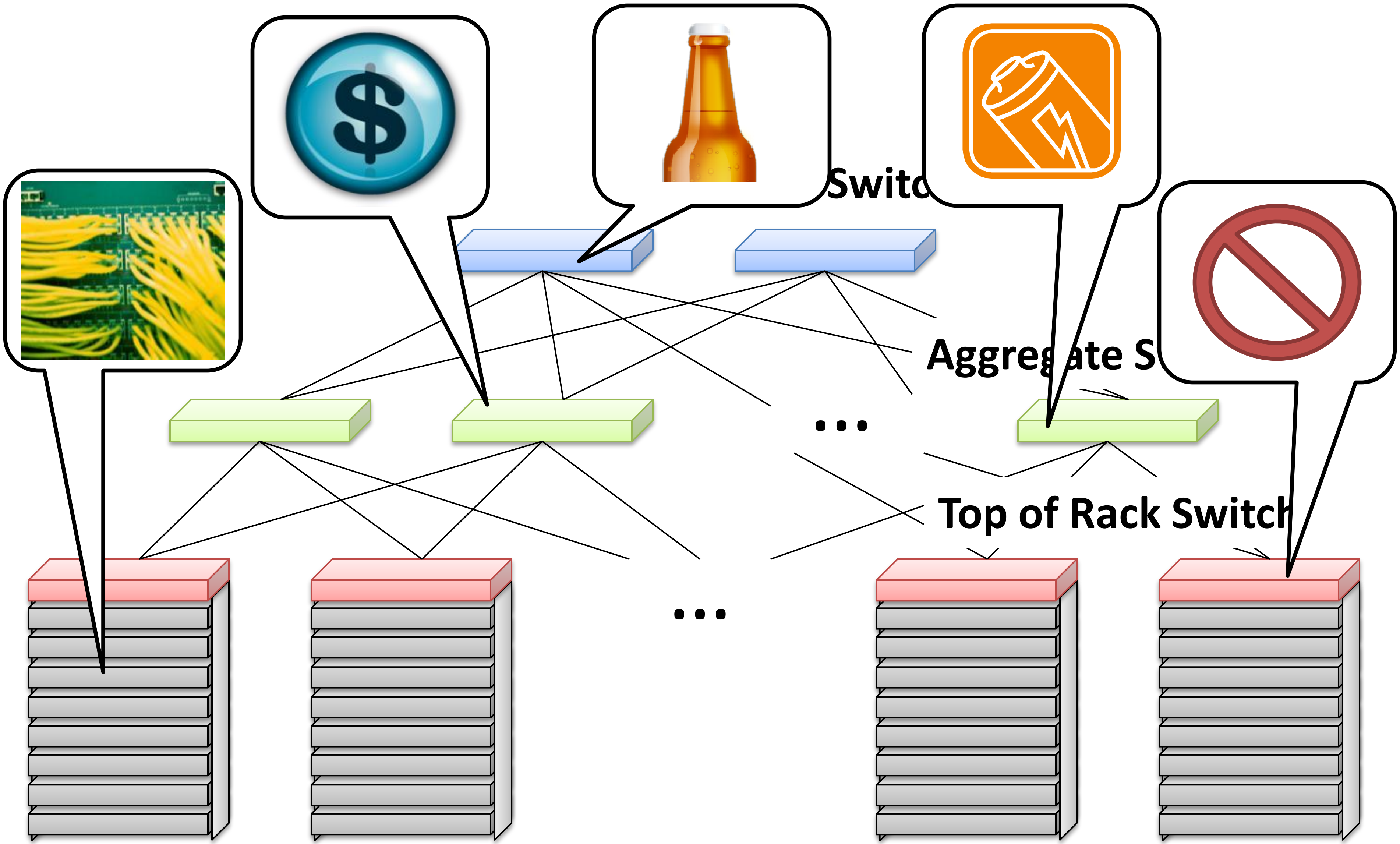
Ji-Yong Shin
Cornell University

In collaboration with
Emin Gün Sirer (Cornell), Hakim Weatherspoon (Cornell) and Darko Kirovski (MSR)

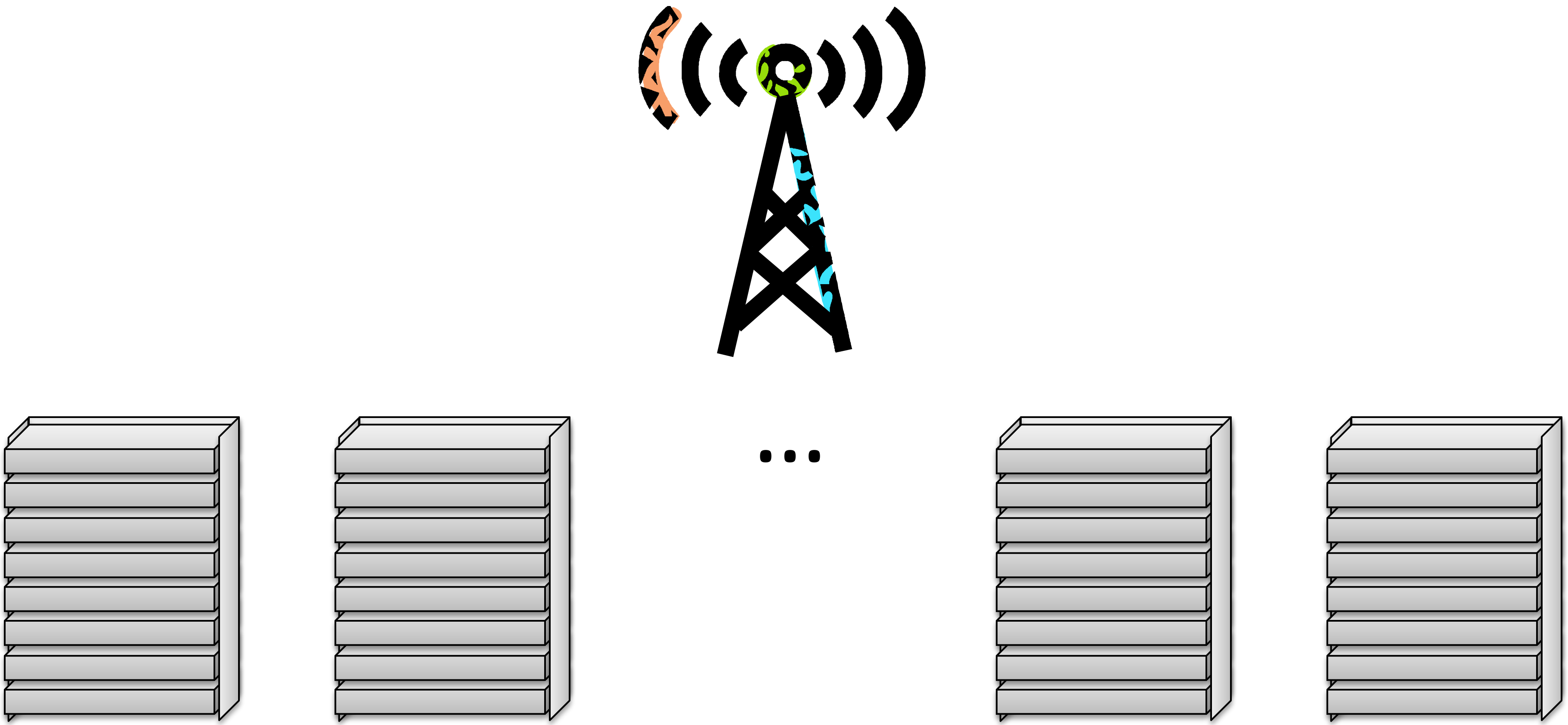
Conventional Datacenter



Conventional Datacenter



Conventional Datacenter



Going Completely Wireless

- Opportunities
  - Low maintenance : no wires
  - Low power: no large switches
  - Low cost: all of the above

  - Fault tolerant: multiple network paths
  - High performance: multiple network paths

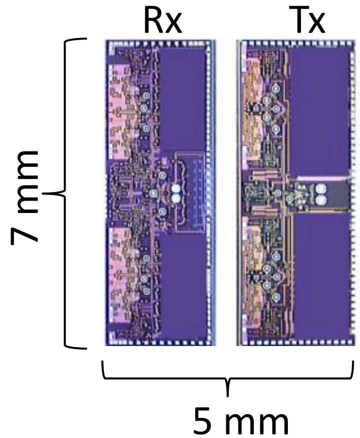
Which wireless technology?

60GHz Wireless Technology

- Short range
  - Attenuated by oxygen molecules
- Directional
  - Narrow beam
- High bandwidth
  - Several to over 10Gbps
- License free
  - Has been available for many years

Why now?

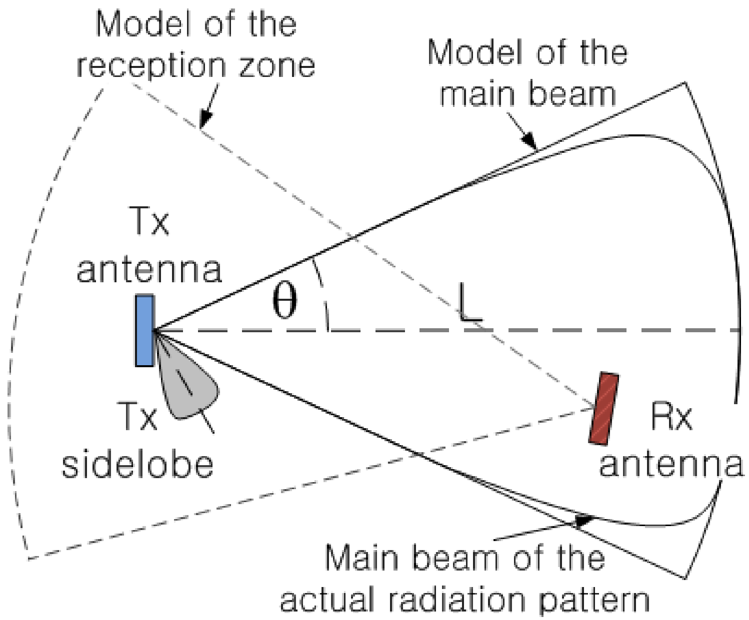
- CMOS Integration
  - Size < dime
  - Manufacturing cost < $1



[Pinel '09]

60 GHz Antenna Model

- One directional
  - *Signal angle between 25° and 45°*
  - *Maximum range < 10 m*
  - No beam steering
- Bandwidth < 15Gbps
  - TDMA (TDD)
  - FDMA (FDD)
- Power at 0.1 – 0.3W



How to integrate to datacenters?

Designing Wireless Datacenters

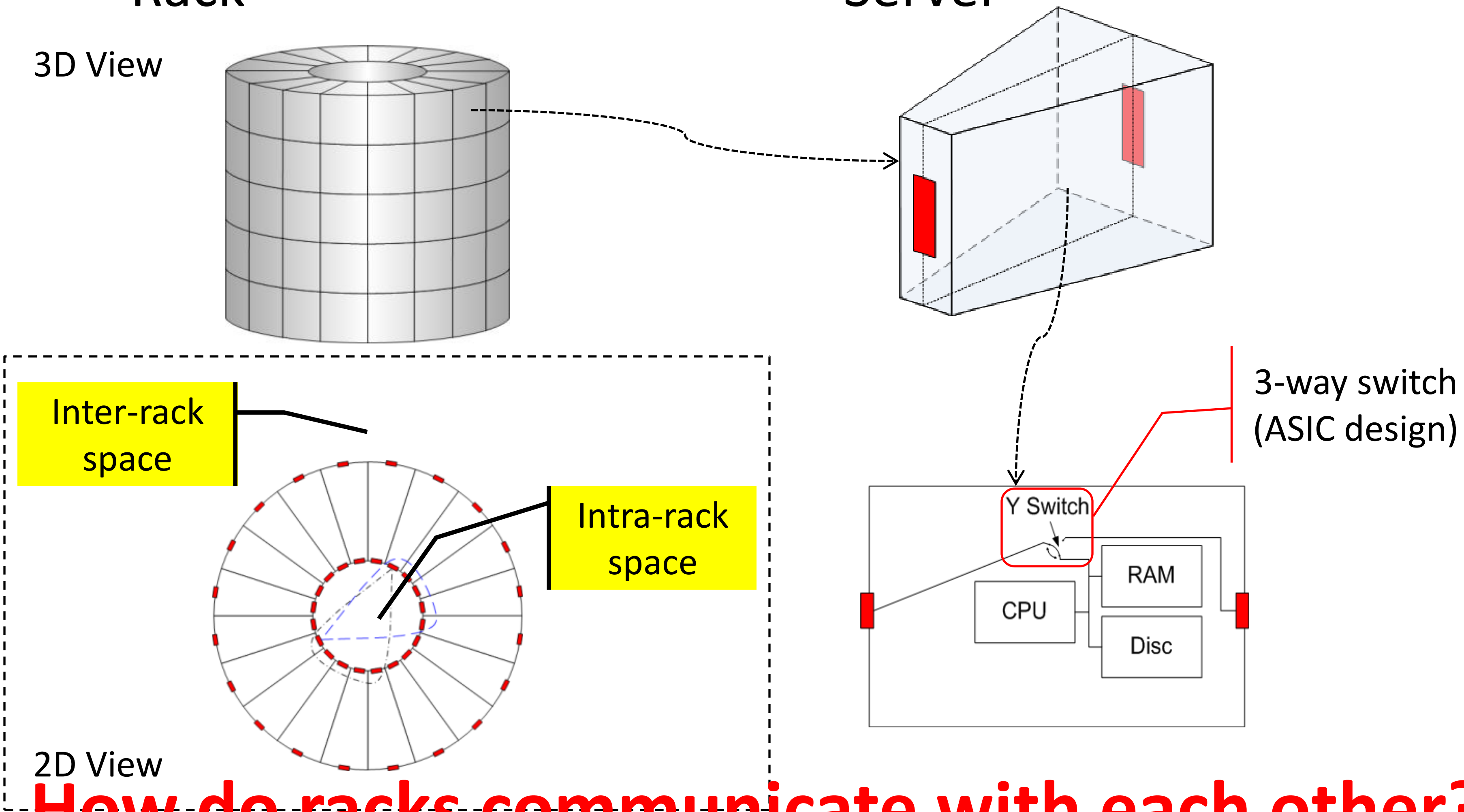
- Challenges
  - How should transceivers and racks be oriented?
  - How should the network be architected?
  - Interference of densely populated transceivers?

Completely Wireless Datacenters

- Motivation
- *Cayley* Wireless Datacenters
  - Transceiver placement and topology
    - Server and rack designs
  - Network architecture
    - MAC protocols and routing
- Evaluation
  - Physical Validation: Interference measurements
  - Performance and power
- Future
- Conclusion

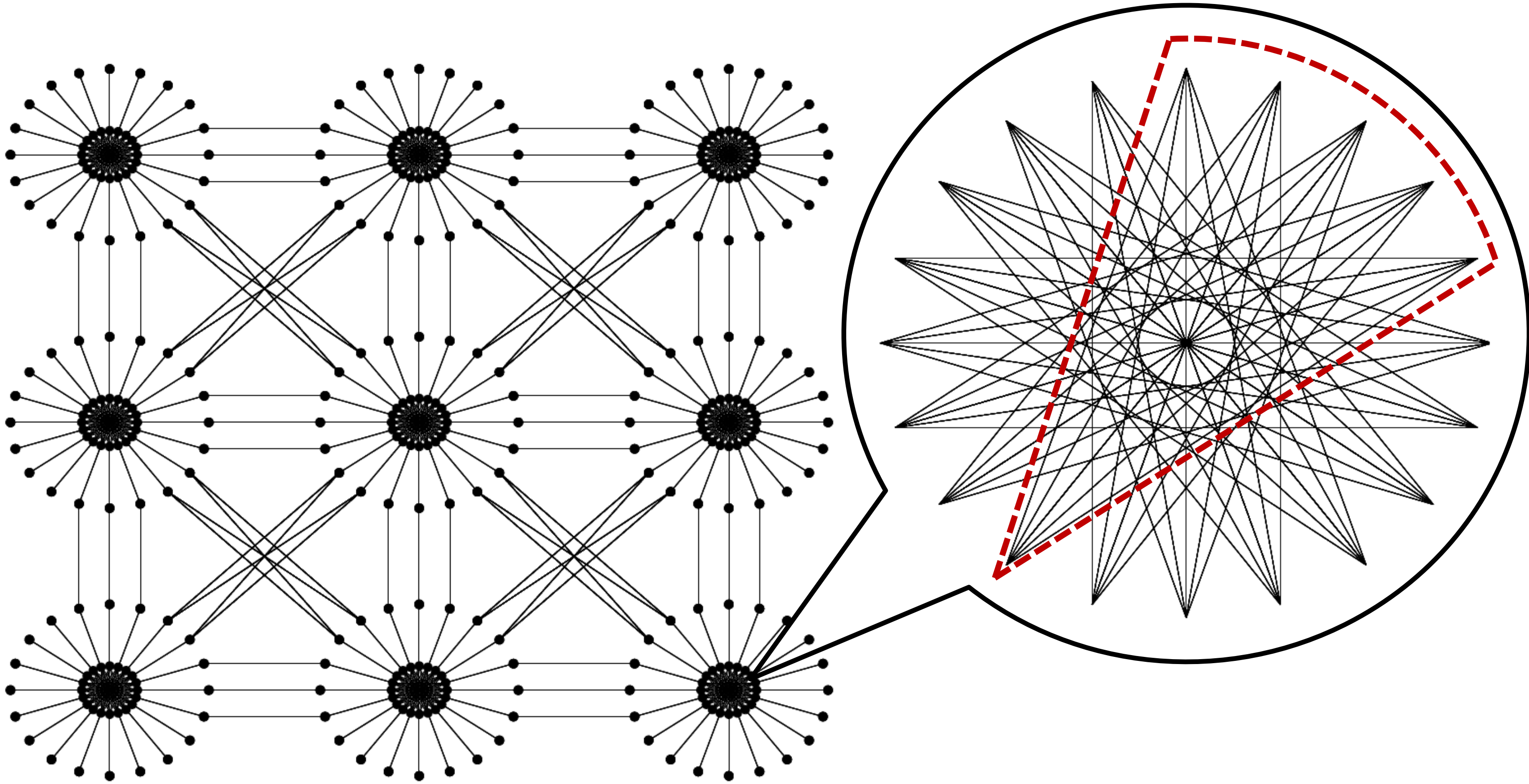
Transceiver Placement: Server and Rack Design

- Rack
- Server



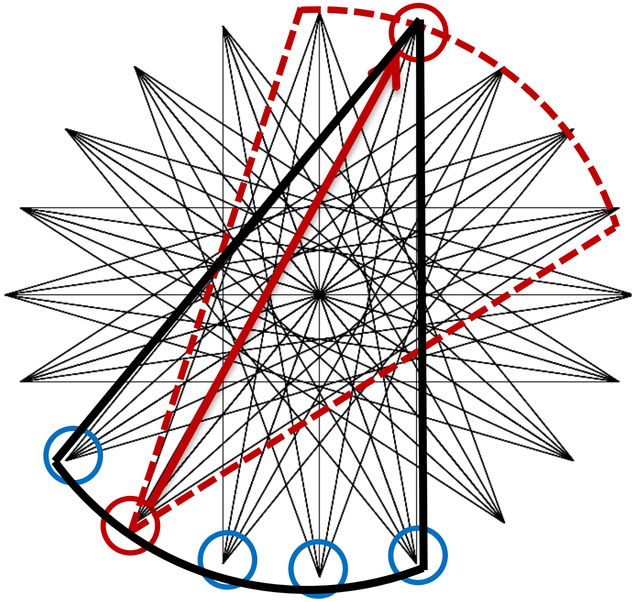
How do racks communicate with each other?

Cayley Network Architecture: Topology



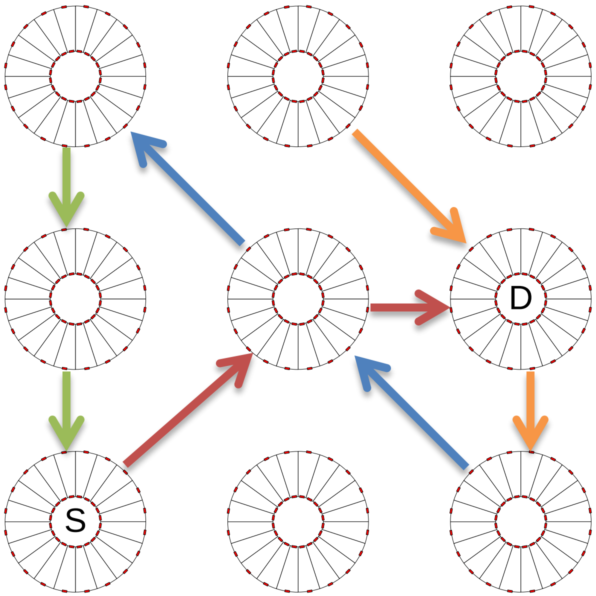
Masked Node Problem and MAC

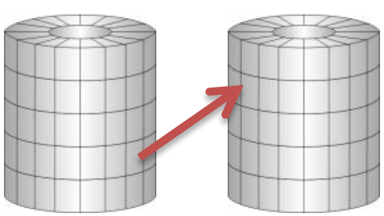
- Most nodes are hidden terminals to others
  - Multiple (>5) directional antennae
    => Masked node problem
  - Collisions can occur
- Dual busy tone multiple access [Hass'02]
  - Out of band tone to preserve channels
  - Use of FDD/TDD channels as the tone



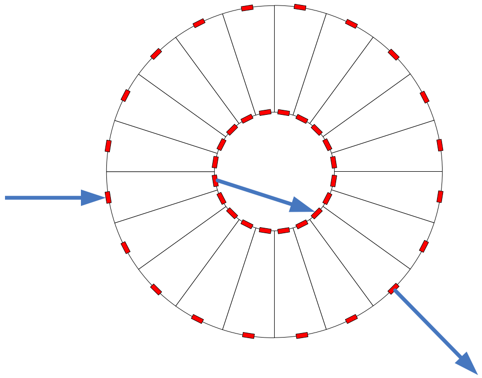
Cayley Network Architecture: Routing

- Geographical Routing
- Inter rack
  – Diagonal XYZ routing

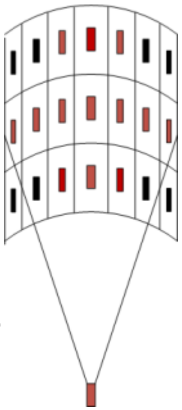




- Turn within rack
  – Shortest path turning



- Within dst rack to dst server
  – Up down to dst story
  – Shortest path to dst serve

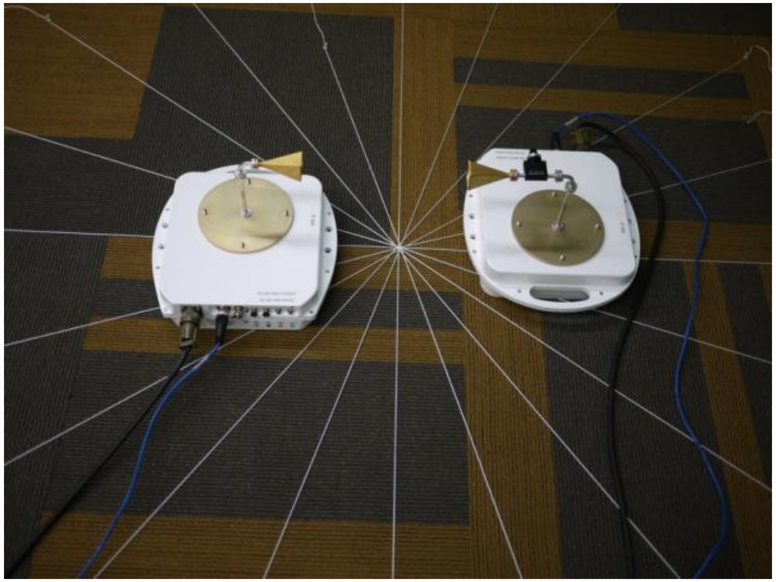


Completely Wireless Datacenters

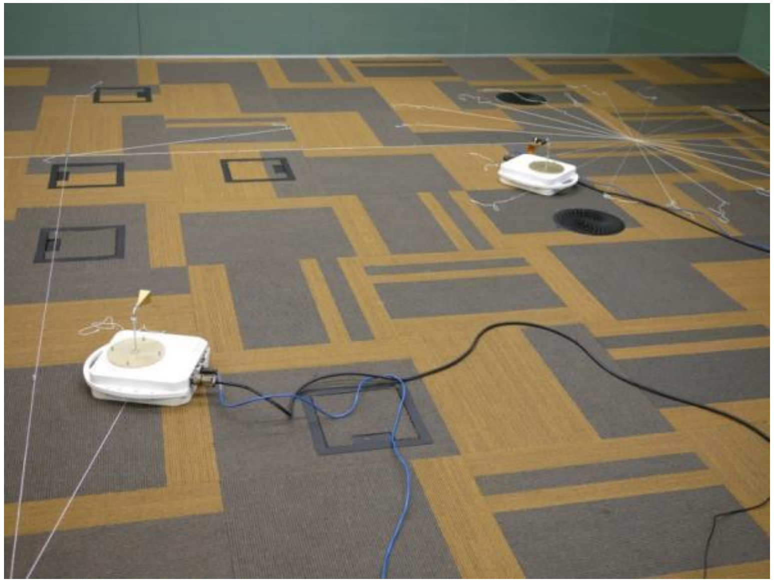
- Motivation
- *Cayley* Wireless Datacenters
  - Transceiver placement and topology
    - Server and rack designs
  - Network architecture
    - MAC protocols and routing
- Evaluation
  - Physical validation: Interference measurements
  - Performance and power
- Future
- Conclusion

Hardware Setup for Physical Validation

- Use of a conservative platform
- Real-size datacenter floor plan setup
- Validation of all possible interferences

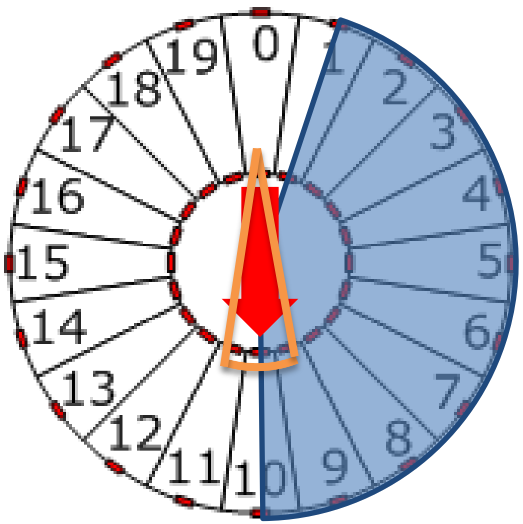


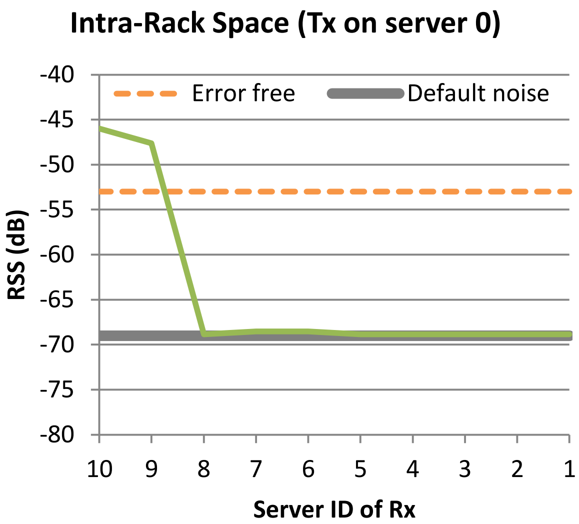
Intra-rack communications



Inter-rack communications

Physical Validation: Interference Evaluation
(Signal angle θ = 15°)

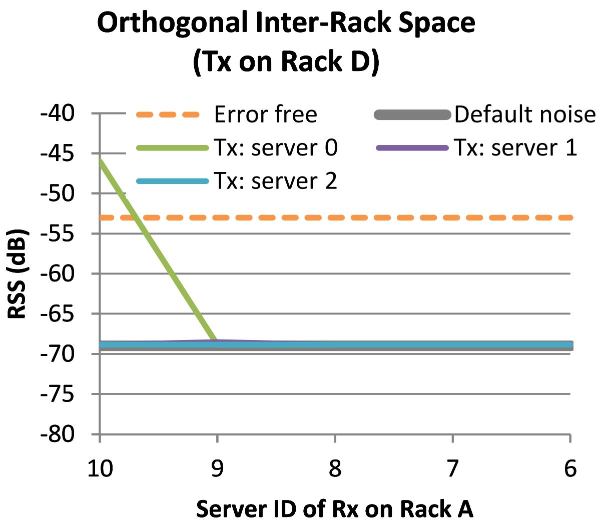


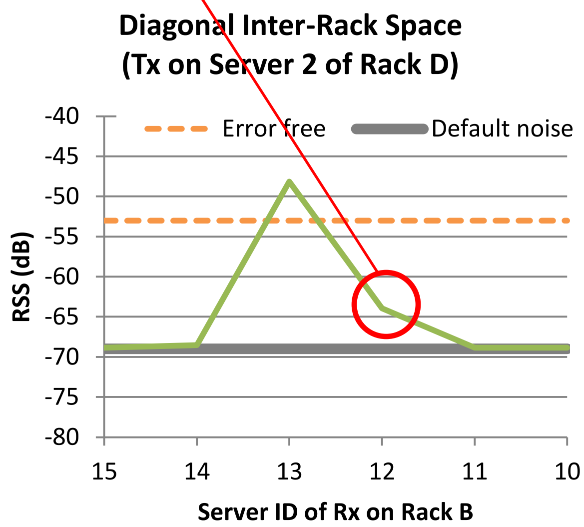


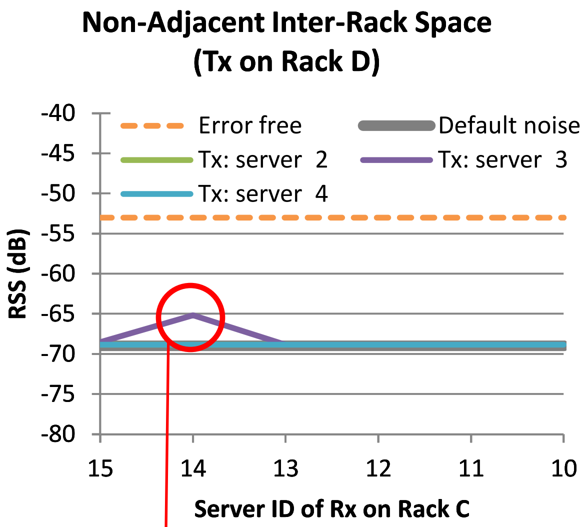
Physical Validation: Interference Evaluation

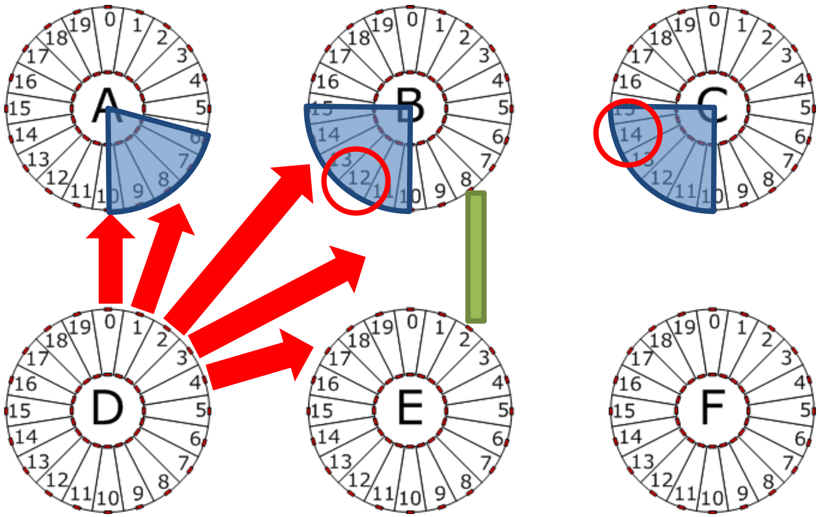
(Signal angle θ = 15°)

Edge of signal: can be eliminated









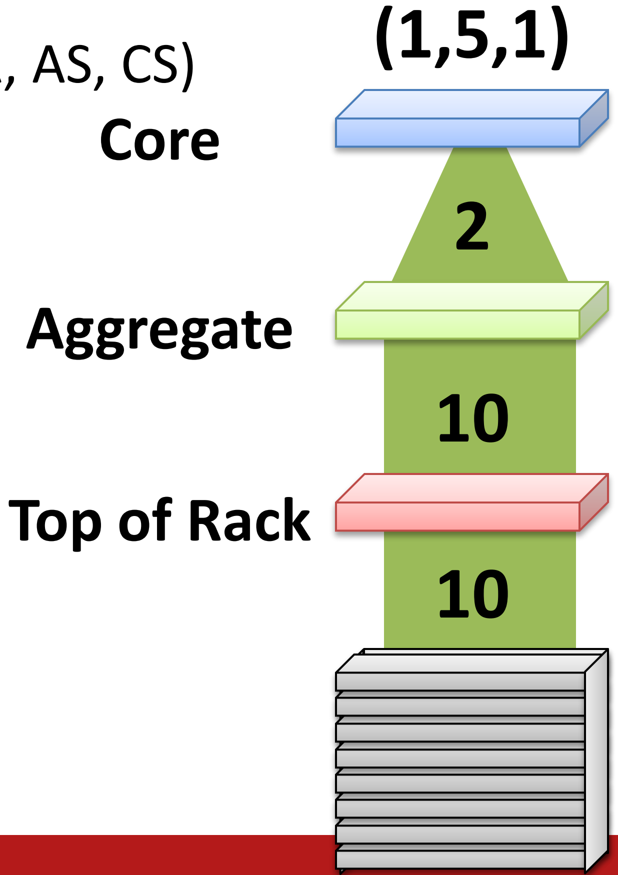
Potential Interference: can be blocked using conductor curtains

Evaluation

- Performance: How well does a Cayley datacenter perform and scale?
  - Bandwidth and latency
- Failure tolerance: How well can a Cayley datacenter handle failures?
  - Server, story, and rack failure
- Power: How much power does a Cayley datacenter consume compared to wired datacenters

Evaluation Setup

- Simulate 10K server datacenter
  - Packet level: routing, MAC protocol, switching delay, bandwidth
- Conventional datacenter (CDC)
  - 3 Layers of oversubscribed switches (ToR, AS, CS)
    - (1, 5, 1), (1, 7, 1) and (2, 5, 1)
    - Latency: 3-6us switching delay
    - Bandwidth: 1Gbps server
- FAT-tree: Equivalent to CDC (1,1,1)
- Cayley wireless datacenter
  - 10Gbps bandwidth
  - 1 Transceiver covers 7 to 8 others
  - Signal spreading angle of 25°
  - Low latency Y-switch (<< 1us)

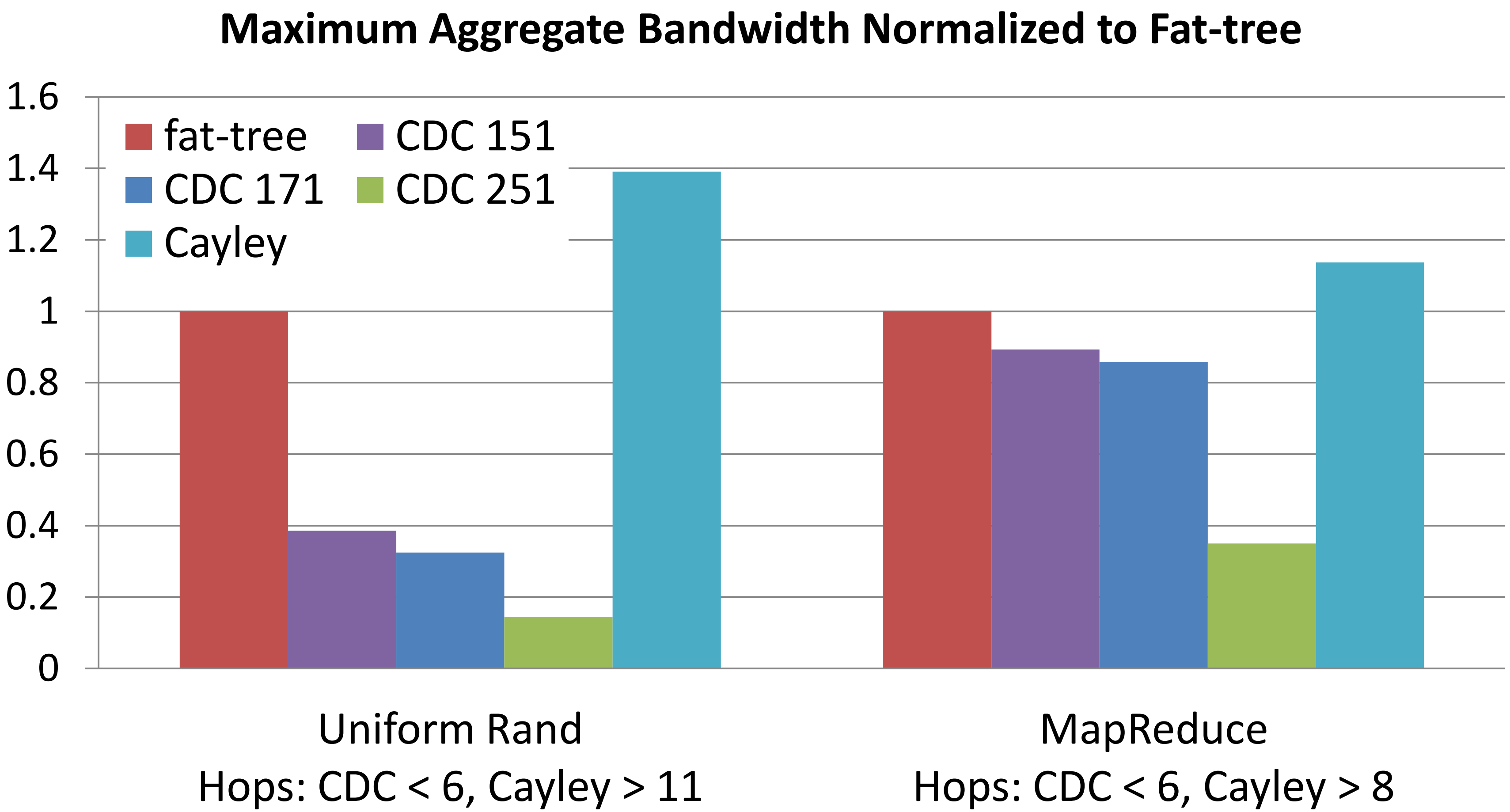


Evaluation Setup

- Uniform random
  - Src and dst randomly selected in entire datacenter
- MapReduce
  - Src sends msg to servers in same row of rack
  - Receiver sends msg to servers in same column of rack
  - Receivers send msg to servers inside same pod with 50% probability

Bandwidth

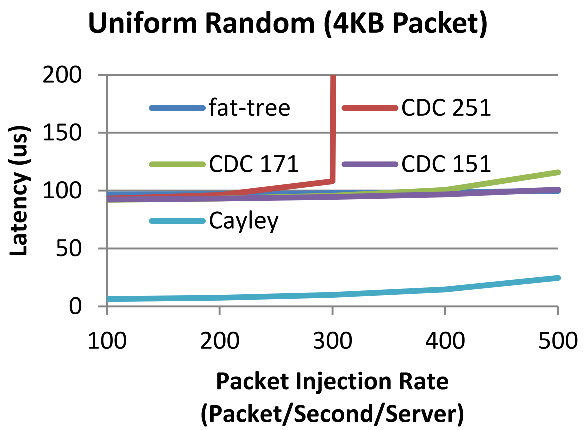
- Burst of 500 x 1KB packets per server sent

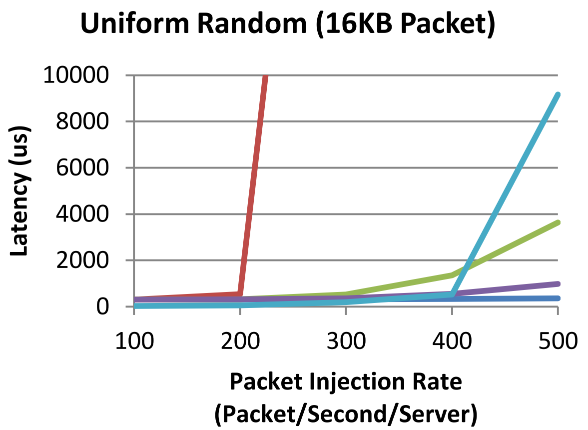


Cayley datacenters have the most bandwidth

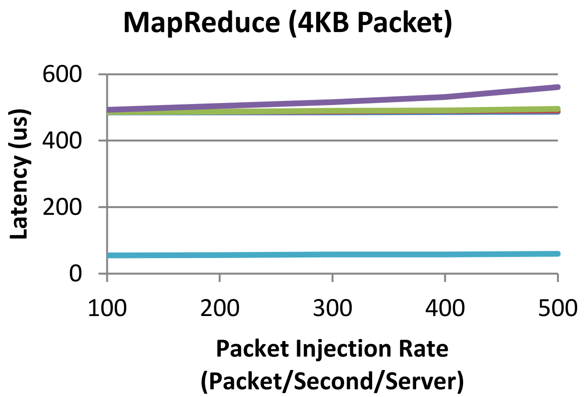
Latency

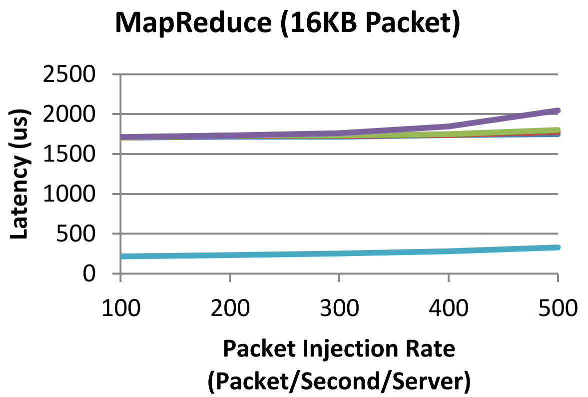
- Uniform random benchmark





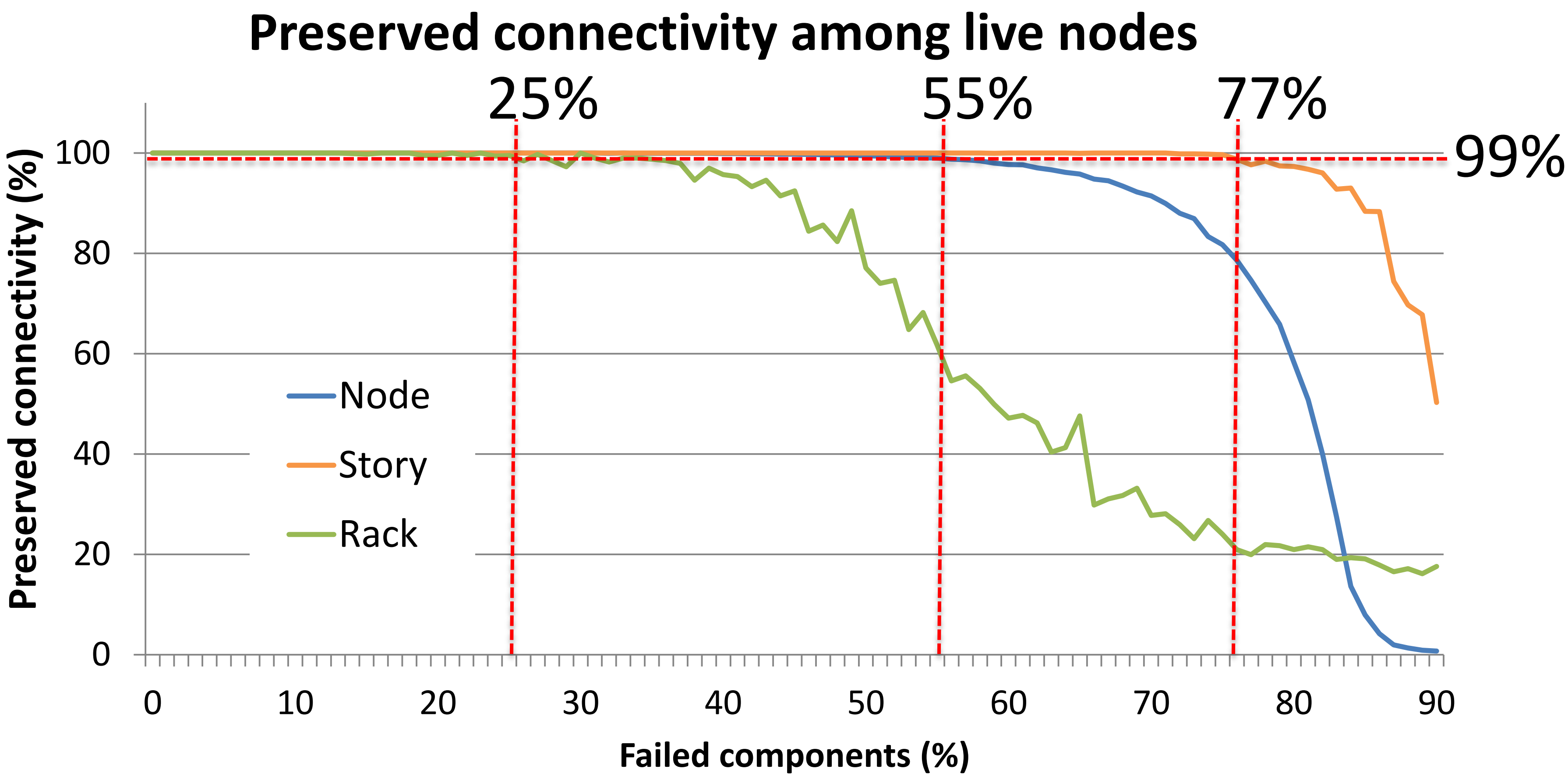
- MapReduce benchmark





Cayley datacenters typically performs the best

Fault Tolerance



Cayley datacenters are extremely fault tolerant

Power Consumption to Connect 10K Servers

- Conventional datacenter (CDC) *

Switch Type	Typical Power
Top of rack switch (ToR)	176W
Aggregation switch (AS)	350W
Core switch (CS)	611W

  – Depending on the oversubscription rate **58KW to 72KW**

- Cayley datacenter
  – Transceivers consume < 0.3W
  – Maximum power consumption: **6KW**

- **Less than 1/10 of CDC power consumption**

* Cost and spec of Cisco 4000, 5000, 7000 series switches

Discussion and Future Work

- Only scratched the surface
  - How far can wireless datacenters go with no wires?
- Need larger experiment/testbed
  - Interference and performance of densely connected datacenter?

- Scaling to large datacenters (>100K servers)?
- Scaling to higher bandwidth (> 10Gbps)?

Conclusion

- Completely wireless datacenters *can be* feasible
- Cayley wireless datacenters exhibit
  - Low maintenance
  - High performance
  - Fault tolerant
  - Low power
  - Low cost

References

- S. Pinel, P. Sen, S. Sarkar, B. Perumana, D. Dawn, D. Yeh, F. Barale, M. Leung, E. Juntunen, P. Vadivelu, K. Chuang, P. Melet, G. Iyer, and J. Laskar. 60GHz single-chip CMOS digital radios and phased array solutions for gaming and connectivity. IEEE Journal on Selected Areas in Communications, 27(8), 2009.
- Z.J. Hass and J. Deng. Dual busy tone multiple access (DBTMA)-a multiple access control scheme for ad hoc networks. IEEE Transactions on Communications, 50(6), 2002.
- PEPPM. Cisco Current Price List. http://www.peppm.org/Products/cisco/price.pdf, 2012.