Cutting the Cord: A Robust Wireless Facilities Network for Data Centers

Yibo Zhu, Xia Zhou§, Zengbin Zhang, Lin Zhou, Amin Vahdat†, Ben Y. Zhao and Haitao Zheng

U.C. Santa Barbara, §Dartmouth College, †U.C. San Diego & Google

yibo@cs.ucsb.edu
Data Center Networks (DCN)

- DCN: key infrastructures for mobile and big data applications

- Large and dynamic $\rightarrow$ management complexity
  - Highly dynamic data traffic
  - Shared by changing customers
  - Frequent failure, maintenance and upgrades
Beyond Data Plane

• Various control messages
  – Flow scheduling
  – Monitoring environment & power
  – Virtual machine imaging and configuration
  – Failure recovery
  – Bootstrap upgraded devices

• Must deliver timely and reliable
  – Not interfered by congested data traffic
  – Even when data plane not working

Upgrade ~100 servers per day on average
A Facilities Network

Proposed DCN architecture
## Requirements of Facilities Network

<table>
<thead>
<tr>
<th>Performance</th>
<th>Fault isolation</th>
<th>Robustness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low bandwidth</td>
<td>Not fate-sharing</td>
<td>Always connected</td>
</tr>
<tr>
<td>1Gbps enough</td>
<td>Ideally physically separated</td>
<td></td>
</tr>
<tr>
<td>Bounded delay⁠¹</td>
<td></td>
<td></td>
</tr>
<tr>
<td>One packet message &lt;10ms</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1MB Large message &lt;500ms</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

¹*Devoflow, SIGCOMM’11*

Must remain working even racks taken off
Option: Wired Facilities Network

- Connect all devices using cables
  - In-band: share w/ data plane
  - Out-of-band

- Advantage: large capacity

- Challenges
  - Out-of-band: high cost, wiring headache
  - Poor fault isolation/robustness
    - Zero fault isolation for in-band
    - Even out-of-band interrupted by cable tray maintenance
Option: Wireless Facilities Network

• Add radios to racks
  – WiFi (1.3Gbps), 60GHz (6.76Gbps)
  – Enough bandwidth

• Advantages
  – Cost: low (no additional switches/cables)
  – Fault isolation: physically isolated from data plane
  – Robustness: automatically reform links

• Challenge: delay from wireless interference
Choice of Wireless Technology

WiFi

Widely available
Well-understood
• Omni-directional
• Contend for channel
Large interference footprint
• Poor in dense DC
• Unpredictable delay

60GHz 3D Beamforming
Recently available
Less-understood
• Highly directional
• Need coordination
Small interference footprint
• Good for dense DC
Outline

• Motivation

• System design
  – Angora: a 60GHz facilities network
  – Wireless overlay design
  – Minimizing link interference
  – Fault recovery

• Evaluation

• Conclusion
Angora: a 60GHz Overlay

- Highly directional signal + limited radios per rack \(\rightarrow\) limited connections per rack
- Antenna alignment \(\rightarrow\) extra delay

- Angora: fixed topology overlay
  - Multi-hop \(\rightarrow\) any-to-any connectivity
  - Fixed topology \(\rightarrow\) no link coordination \(\rightarrow\) no extra controllers, minimize delay
Structured Overlay Graph

- Key goal: minimize delay (hop count)
- The constraint: constant number of radios per rack → constant degree graph

- We choose **Kautz** graph
  - Smallest diameter given node degree and the number of nodes.
- Hop count: **Kautz < Random**\(^1\) << **Fat-tree**
  - Wired networks prefer Fat-tree due to low wiring complexity

\(^1\) *Jellyfish*, NSDI’12
Kautz Graph

• Simple digit-shift routing

• Graph diameter = length of IDs = $\sim \log_k(N)$
  – $N$: # of nodes, $k$: node degree (4)

• Challenge: Kautz only supports specific $N$
  – We design an algorithm to handle arbitrary $N$
Node Naming and Interference

• Nodes naming affects interference
  – 60GHz interference: function of angular separation

• Goal: maximize angular separation between links

• Designed an optimal naming scheme
  – Achieved 14° angular separation in practice
Failure Recovery Algorithms

• Link failure $\rightarrow$ remove a graph edge
  – May happen when radio fails, or signal blocked
  – Leverage Kautz structure to re-route the traffic

• Rack failure $\rightarrow$ remove a graph node
  – Similar deterministic algorithm
Failure Recovery Results

- Structural fault recovery $\rightarrow$ good robustness
- Deterministic algorithms $\rightarrow$ no extra coordinator

0.1% paths fail when 20% of links fail

100% connectivity until >50% collocated racks fail (leverage far-away siblings)
Outline

• Motivation

• System design

• Evaluation
  – Testbed
  – Simulation

• Conclusion
Testbed Validation

• Two testbeds
  – HXI: horn antennas
  – Wilocity: 2x8 arrays, affordable for multi-hop

• Single link performance
  – Measured per-second TCP throughput over one month
  – Average ~900Mbps (capped by 1Gbps NIC)
  – Standard variation <1% average throughput → as stable as a wired link
Testbed Validation (Multi-hop)

- Without interference
  - Throughput not affected
  - Latency scales with hops

<table>
<thead>
<tr>
<th>Path Length</th>
<th>TCP Thpt¹</th>
<th>10KB Latency</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 hops</td>
<td>662Mbps</td>
<td>2.5ms</td>
</tr>
<tr>
<td>3 hops</td>
<td>654Mbps</td>
<td>3.1ms</td>
</tr>
<tr>
<td>4 hops</td>
<td>665Mbps</td>
<td>3.5ms</td>
</tr>
</tbody>
</table>

Multi-hop paths have low interference $\rightarrow$ high throughput and predictable latency.

- Different channels $\rightarrow$ no throughput loss
  - 802.11ad defines 3 channels $\rightarrow$ low self-interference

- Cross-path interference mitigated by node naming

¹Throughput lower than single link due to software port forwarding overhead
Large-scale Simulation

• We implement Angora in NS-3
  – Antenna: horns and arrays
  – 3D beamforming signal reflection
  – 802.11ad PHY/MAC
  – Kautz overlay routing
  – Medium size (320~480 racks) DCN layouts

• Micro-benchmark: path hop count, concurrency, fault-tolerance

• End-to-end performance: single flow, Poisson flows, synchronized flows
End-to-end Performance

- Worst case: synchronized flows

- Tail delay satisfies facilities network requirements
- Structural (Kautz) >> random at tails
Conclusion

• Motivation: build an orthogonal facilities network as a core tool for managing DCN.

• We propose Angora, a Kautz overlay built on 60GHz 3D beamforming links.

• Addressed challenges
  – Wireless interference
  – Robustness to failures
  – Incomplete Kautz graph
Thank you!

Questions?