On the Feasibility of Completely Wireless Datacenters

Ji-Yong Shin, Emin Gün Sirer, and Hakim Weatherspoon
Dept. of Computer Science, Cornell University
Ithaca, NY, USA
{jyshin, egs, hweather}@cs.cornell.edu

Darko Kirovski
Microsoft Research
Redmond, WA, USA
darkok@microsoft.com

ABSTRACT

Conventional datacenters, based on wired networks, entail high wiring costs, suffer from performance bottlenecks, and have low resilience to network failures. In this paper, we investigate a radically new methodology for building wire-free datacenters based on emerging 60GHz RF technology. We propose a novel rack design and a resulting network topology inspired by Cayley graphs that provide a dense interconnect. Our exploration of the resulting design space shows that wireless datacenters built with this methodology can potentially attain higher aggregate bandwidth, lower latency, and substantially higher fault tolerance than a conventional wired datacenter while improving ease of construction and maintenance.

Categories and Subject Descriptors

C.2.1 [**Computer-Communication Networks**]: Network Architecture and Design—*Network topology, Wireless communication*

General Terms

Design, Experimentation, Performance

Keywords

60GHz RF, Wireless data center

1. INTRODUCTION

Performance, reliability, cost of the switching fabric, power consumption, and maintenance are some of the issues that plague conventional wired datacenters [2, 16, 17]. Current trends in cloud computing and high-performance datacenter applications indicate that these issues are likely to be exacerbated in the future [1, 4].

In this paper, we explore a radical change to the construction of datacenters that involves the removal of all but power supply wires. The workhorses of communication in this new design are the newly emerging directional, beamformed 60GHz RF communication channels characterized by high bandwidth (4-15Gbps) and short range ($\leq$ 10 meters). New 60GHz transceivers [40, 43] based on standard 90nm CMOS technology make it possible to realize such channels with low cost and high power efficiency (< 1W). Directional (25° – 60° wide) short-range beams employed by these radios enable a large number of transmitters to simultaneously communicate with multiple receivers in tight confined spaces.

The unique characteristics of 60GHz RF modems pose new challenges and tradeoffs. The most critical questions are those of feasibility and structure: can a large number of transceivers operate without signal interference in a densely populated datacenter? How should the transceivers be placed and how should the racks be oriented to build practical, robust and maintainable networks? How should the network be architected to achieve high aggregate bandwidth, low cost and high fault tolerance? And can such networks compete with conventional wired networks?

To answer these questions, we propose a novel datacenter design—because its network connectivity subgraphs belong to a class of Cayley graphs [6], we call our design a Cayley datacenter. The key insight behind our approach is to arrange servers into a densely connected, low-stretch, failure-resilient topology. Specifically, we arrange servers in cylindrical racks such that inter- and intra-rack communication channels can be established and form a densely connected mesh. To achieve this, we replace the network interface card (NIC) of a server with a Y-switch that connects a server's system bus with two transceivers positioned at opposite ends of the server box. This topology leads to full disappearance of the classic network switching fabric (e.g., no top-of-rack switches, access routers, copper and optical interconnects) and has far-reaching ramifications on performance.

Overall, this paper makes three contributions. First, we present the first constructive proposal for a fully wireless datacenter. We show that it is possible for 60GHz technology to serve as the sole and central means of communication in the demanding datacenter setting. Second, we propose a novel system-level architecture that incorporates a practical and efficient rack-level hardware topology and a corresponding geographic routing protocol. Finally, we examine the performance and system characteristics of Cayley datacenters. Using a set of 60GHz transceivers, we demonstrate that signals in Cayley datacenters do not interfere with each other. We also show that, compared to a fat-tree [37, 38] and a conventional datacenter, our proposal ex-

ANCS'12, October 29–30, 2012, Austin, Texas, USA.

hibits higher bandwidth, substantially improved latency due to the switching fabric being integrated into server nodes, and lower power consumption. Cayley datacenters exhibit strong fault tolerance due to a routing scheme that can fully explore the mesh: Cayley datacenters can maintain connectivity to over 99% of live nodes until up to 55% of total nodes fail.

2. 60GHZ WIRELESS TECHNOLOGY

In this section, we briefly introduce the communication characteristics of the newly emerging 60GHz wireless technology, which is the foundation of our datacenter.

Propagation of RF (radio frequency) signals in the 57-64GHz sub-band is severely attenuated because of the resonance of oxygen molecules, which limits the use of this sub-band to relatively short distances [34]. Consequently 57-64GHz is unlicensed under FCC rules and open to short-range point-to-point applications. Several efforts are aiming to standardize the technology, with most of them tailored to home entertainment: two IEEE initiatives, IEEE 802.15.3c and 802.11.ad [26,55], WiGig 7Gbps standard with beam-forming [56], and ECMA-387/ISO DS13156 6.4Gbps spec [14] based upon Georgia Tech's design [43].

In this paper, we focus on a recent integrated implementation from Georgia Tech whose characteristics are summarized in Table 1:

Category	Characteristic
Technology	Standard 90nm CMOS
Packaging	Single chip Tx/Rx in QFN
Compliance	ECMA TC48
Power	0.2W (at output power of 3dBm)
Range	$\leq$ 10m
Bandwidth	4-15Gbps

Table 1: 60GHz wireless transceiver characteristics [43].

More details about 60GHz transceiver characteristics can be found from a link margin, which models communication between a transmitter (Tx) and a receiver (Rx). The link margin, M, is the difference between the received power at which the receiver stops working and the actual received power, and can be expressed as follows:

$$M = P_{TX} + G_{TX+RX} - L_{Fade} - L_{Implementation} - FSL - NF - SNR, \quad (1)$$

where P_{TX} and G_{TX+RX} represent transmitted power and overall joint transmitter and receiver gain which is dependent upon the geometric alignment of the Tx↔Rx antennae [57]. Free-space loss equals $FSL = 20\log_{10}(4\pi D/\lambda)$, where D is the line-of-sight Tx↔Rx distance and λ wavelength. The noise floor $NF \sim 10\log_{10}(R)$ is dependent upon R, the occupied bandwidth. SNR is the signal to noise ratio in dBs which links a dependency to the bit error rate as $BER = \frac{1}{2}\text{erfc}(\sqrt{SNR})$ for binary phase-shift keying (BPSK) modulation for example. Loss to fading and implementation are constants given a specific system. From Equation 1, one can compute the effects of constraining different communication parameters.

Figure 1 illustrates a planar slice of the geometric communication model we consider in this paper. A transmitter antenna radiates RF signals within a lobe—the surface of the lobe is a level-set whose signal power is equal to one half of the maximum signal power within the lobe. Because the

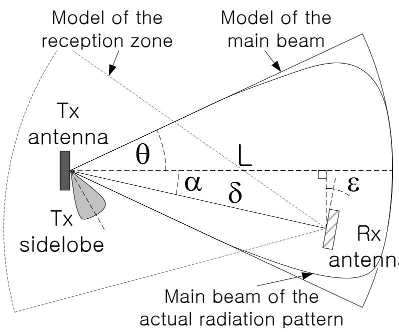


Figure 1: Geometric communication model.

attenuation is very sharp in the 60GHz frequency range, a receiver antenna should be within the bound of a transmitter's beam for communication. The beam is modeled as a cone with an angle θ and length L. Using a spherical coordinate system centered at transmitter's antenna, one can define the position of the receiver antenna with its radius, δ, elevation α, and azimuth β. The plane of the receiver antenna can then be misaligned from the plane of the transmitter antenna by an angle ε along the elevation and γ along the azimuth. We use a modeling tool developed at Georgia Tech to convert $\{\alpha, \beta, \gamma, \varepsilon, \delta, L, \theta\}$ into G_{TX+RX}. Through personal communication with Georgia Tech's design team, we reduced our space of interest to $25^\circ \leq \theta \leq 45^\circ$ as a constraint to suppress side lobes. Based on design parameters from the antenna prototypes developed by the same team, we model a reception zone of the receiver that is in identical shape to the main transmitter beam. We limit ε and γ to be smaller than θ such that the transmitter is located within the reception zone and assume a BER of 10^{-9} at 10Gbps bandwidth within $L < 3$ meters range. We do not utilize beam-steering[1] and assume that the bandwidth can be multiplexed using both time (TDD) and frequency division duplexing (FDD).

The design parameters of the transceiver are optimized for our datacenter design and lead to a higher bandwidth and less noisier transceiver design compared to off-the-shelf 60GHz transceivers for HDMI [49]. More research in 60GHz RF design with a focus on Cayley datacenters can further improve performance.

3. CAYLEY DATACENTER DESIGN

This section introduces Cayley datacenter architecture, the positioning of the 60GHz transceivers in a wireless datacenter, and the resulting network topology. We also introduce a geographical routing protocol for this unique topology and adopt a MAC layer protocol to address the hidden terminal and the masked node problem.

3.1 Component Design

In order to maximize opportunities for resource multiplexing in a wireless datacenter, it is important to use open spaces efficiently, because the maximum number of live connections in the network is proportional to the volume of the datacenter divided by that of a single antenna beam. We

[1]Typically, reconnection after beam-steering involves training of communication codebooks involving delays on the order of microseconds, which is not tolerable in datacenters.

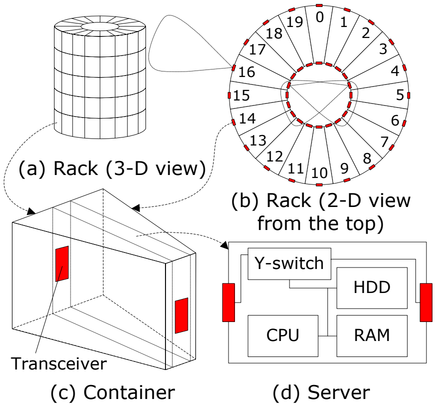


Figure 2: Rack and server design.

focus on the network topology that would optimize key performance characteristics, namely latency and bandwidth.

To separate the wireless signals for communications within a rack and among different racks, we propose cylindrical racks (Figure 2.a) that store servers in prism-shaped containers (Figure 2.c). This choice is appealing, because it partitions the datacenter volume into two regions: intra- and inter-rack free space. A single server can be positioned so that one of its transceivers connects to its rack's inner-space and another to the inter-rack space as the rack illustrated in Figure 2.b. A rack consists of S stories and each story holds C containers; we constrain $S = 5$ and $C = 20$ for brevity of analysis and label servers in the same story from 0 to 19 starting from the 12 o'clock position in a clockwise order.

The prism containers can hold commodity half-height blade servers. A custom built Y-switch connects the transceivers located on opposite sides of the server (Figure 2.d). The Y-switch, whose design is discussed at the end of this section, multiplexes incoming packets to one of the outputs.

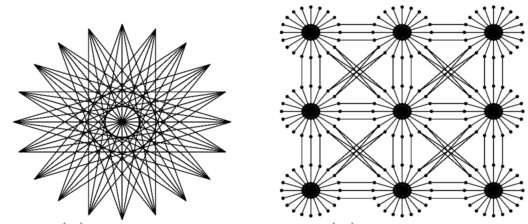
3.2 Topology

The cylindrical racks we propose utilize space and spectrum efficiently and generalize to a topology that can be modeled as a mesh of Cayley graphs.

A Cayley graph [6] is a graph generated from a group of elements G and a generator set $S \subseteq G$. Set S excludes the identity element $e = g \cdot g^{-1}$, where $g \in G$, and $h \in S$ *iff* $h^{-1} \in S$. Each vertex $v \in V$ of a Cayley graph (V, E) corresponds to each element $g \in G$ and edge $(v_1, v_2) \in E$ *iff* $g_1 \cdot g_2^{-1} \in S$. This graph is vertex-transitive, which facilitates the design of a simple distributed routing protocol and is generally densely connected, which adds fault tolerance to the network [50].

When viewed from the top, connections within a story of the rack form a 20-node, degree-k Cayley graph, where k depends on the signal's radiation angle (Figure 3.a). This densely connected graph provides numerous redundant paths from one server to multiple servers in the same rack and ensures strong connectivity.

The transceivers on the exterior of the rack stitch together Cayley subgraphs in different racks. There is a great flexibility in how a datacenter can be constructed out of these racks, but we pick the simplest topology by placing the racks in rows and columns for ease of maintenance. Figure 3.b illustrates an example of the 2-dimensional connectivity of 9 racks in 3 by 3 grids: a Cayley graph sits in the center of each rack and the transceivers on the exterior of the racks



(a) Intra-rack (b) Inter-rack

Figure 3: Cayley datacenter topology when $\theta = 25°$

connect the subgraphs together. Further, since the wireless signal spreads in a cone shape, a transceiver is able to reach servers in different stories, both within and across racks.

3.3 Routing Protocol

A routing protocol for datacenters should enable quick routing decisions, utilize a small amount of memory, and find efficient routes involving few network hops. A geographic routing technique for our topology can fulfill these conditions.

3.3.1 Diagonal XYZ Routing

The uniform structure of Cayley datacenters lends itself to a geographical routing protocol. The routing protocol that we investigate in this paper is called diagonal XYZ routing.

Similar to XY routing [21], diagonal XYZ routing finds an efficient route to the destination at a low computational and storage cost using geographical information. We define the geographical identity g_k of a server k as (rx, ry, s, i), where rx and ry are the x and y coordinates of the rack, s corresponds to the ordinal number for the story, and i is the index of the server within a story. Cayley datacenters use this identity to address the servers.

The geographical identity facilitates finding a path in the Cayley datacenter network. The routing protocol determines the next hop by comparing the destination of a packet to the identity of the sever holding the packet. Based on rx and ry values, the protocol finds an adjacent rack of the server that is closest to the destination. The s value is then used to reach the story height of the destination that the packet should arrive. Finally, the i value is used to forward the packet using the shortest path to the destination server within the same story. Algorithm 1 describes the details about the routing algorithm.

Because the topology has a constant fanout, diagonal XYZ routing requires very little state to be maintained on each host. Every host keeps and consults only three tables to determine the next destination for a packet.

- **Inter-rack routing table**: Maps 8 horizontal directions towards adjacent racks to directly reachable servers on the shortest path to the racks.
- **Inter-story routing table**: Maps 2 vertical directions to directly reachable servers in the same rack of the table owner leading to the desired story.
- **Intra-story routing table**: Maps 20 server index i's to directly reachable servers in the same story in the same rack of the table owner. The servers in the table are on the precomputed shortest path leading to server i.

Inter-rack and inter-story routing tables maintain story s as the secondary index for lookup. Using this index,

Algorithm 1 Diagonal XYZ routing

Require: g_{curr}: geographical identity of the server, where the packet is currently at
  g_{dst}: geographical identity of the packet's final destination
  r_{curr}: rack of g_{curr}
  r_{dst}: rack of g_{dst}
  R_{adj}: set of adjacent racks of r_{curr}
  $T_{InterRack}$: inter-rack routing table of curr
  $T_{InterStory}$: inter-story routing table of curr
  $T_{IntraStory}$: intra-story routing table of curr
Ensure: g_{next}: geographical identity of next destination
  if IsInDifferentRack(g_{curr}, g_{dst}) **then**
    $r_{next} \leftarrow r_{dst}$.GetMinDistanceRack($R_{adj}$)
    $dir \leftarrow r_{curr}$.GetHorizontalDirection(r_{next})
    $G \leftarrow T_{InterRack}$.LookupGeoIDs($dir, g_{dst}.s$)
  else if IsInDifferentStory(g_{curr}, g_{dst}) **then**
    $dir \leftarrow g_{curr}$.GetHorizontalDirection(g_{dst})
    $G \leftarrow T_{InterStory}$.LookupGeoIDs($dir, g_{dst}.s$)
  else if IsDifferentServer(g_{curr}, g_{dst}) **then**
    $G \leftarrow T_{IntraStory}$.LookupGeoIDs($g_{dst}.i$)
  else
    $G \leftarrow g_{dst}$
  end if
  $g_{next} \leftarrow$ RandomSelect(G)

LookupGeoIDs($dir, g_{dst}.s$) returns the identities with the closest s value to $g_{dst}.s$ among the ones leading to dir.

For all three tables, LookupGeoIDs returns multiple values, because a transceiver can communicate with multiple others. The servers found from the table lookup all lead to the same number of hops to the final destination. Thus, the routing protocol pseudo-randomly selects one of the choices to evenly distribute the traffic and to allow a TCP flow to follow the same path. We use a pseudo-random hashing of the packet header like the Toepliz Hash function [28].

The directionality of the radio beam, the presence of multiple transceivers per node and the low latency of the Y-switch makes it possible for Cayley datacenters to deploy cut-through switching [30], which starts routing a packet immediately after receiving and reading the packet header. While this is generally not usable in wireless communication based on omni-directional antennae, unless special methodologies, such as signal cancellation is employed [8,20]— Cayley datacenter servers employ this optimization.

3.3.2 Adaptive Routing in Case of Failure

Compared to a conventional datacenter, a Cayley datacenter has a distinct failure profile. Conventional datacenters are dependent on switches for network connectivity and consequently a switch failure can disconnect many servers. Cayley datacenters, on the other hand, can compensate for the failure of nodes and racks by utilizing some of the many alternative paths in their rich topology. We employ an adaptive routing scheme such as a variant of face routing [27] with the diagonal XYZ routing. Due to space constraints, we do not detail our adaptive routing scheme, but our previous work [47] shows that the routing scheme can circumvent randomly failed racks with less than 5us latency overhead.

3.4 MAC Layer Arbitration

A transceiver in a Cayley datacenter can communicate with approximately 7 to over 30 transceivers depending on its configuration. As a result, communication needs to be coordinated. However, due to the directionality of the signal, all transceivers that can communicate with the same

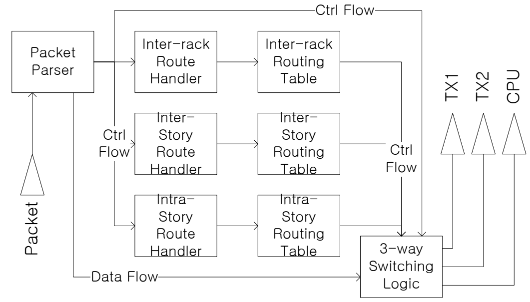


Figure 4: Y-switch schematic.

transceiver act as hidden terminals for each other. Such multiple hidden terminals can lead to a masked node problem [46] that causes collisions if a regular ready-to-send (RTS)/clear-to-send (CTS) based MAC protocol [31] is used.

Therefore, we adopt a dual busy tone multiple access (DBTMA) [23, 24] channel arbitration/reservation scheme. DBTMA is based on an RTS/CTS protocol, but it employs an additional out of band tone to indicate whether the transceivers are transmitting or receiving data. This tone resolves the masked node problem by enabling nodes both at the sending and receiving end to know whether other nodes are already using the wireless channel.

We use a fraction of the dedicated frequency channel for this tone and control messages using FDD so that they do not interfere with the data channel.

3.5 Y-Switch Implementation

The Y-switch is a simple customized piece of hardware that plays an important role in a Cayley datacenter. High-level schematic of this switch is shown in Figure 4. When the Y-switch receives a packet, it parses the packet header and forwards the packet to the local machine or one of the transceivers. The decisions are made by searching through one of the three routing tables described in Section 3.3.1. To analyze the feasibility of the proposed Y-switch design, we implemented the Y-switch design for Xilinx FPGA in Simulink [39] and verified that, for an FPGA running at 270MHz, its switching delay is less than 4ns.

4. PHYSICAL VALIDATION

Before evaluating the performance of Cayley datacenters, we validate the assumptions behind the Cayley design with physical 60GHz hardware. Specifically, we quantify communication characteristics and investigate the possibility of interference problems that may interfere with realizing the Cayley datacenter.

We conduct our experiments using Terabeam/HXI 60GHz transceivers [25] (Figure 5.a). While the Terabeam/HXI transceivers are older and therefore not identical to the Georgia Tech's transceiver described in Section 2, they provide a good baseline for characterizing 60GHz RF signals. This is a conservative platform, previously used in [22], over which modern hardware would provide further improvements. For instance, the Terabeam antennae are large and emit relatively broad side lobes and the signal-guiding horns catch some unwanted signals. In contrast, recently proposed CMOS-based designs can be smaller than a dime, effectively suppress side lobes, and do not use signal-guiding horns at all [36, 43]. To compensate for the noise stemming from the

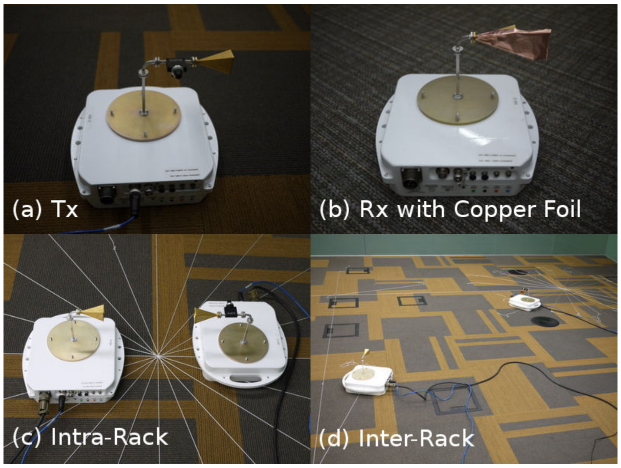


Figure 5: 60GHz Tx, Rx, and measurements on a Cayley datacenter floor plan

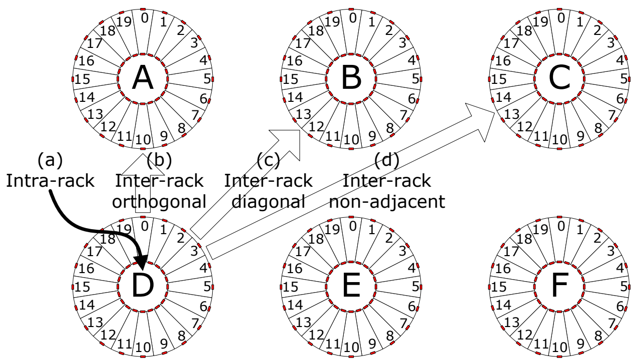


Figure 6: Interference measurement summary

older horn design, we augment one side of the receiver's horn with a copper foil (Figure 5.b). The devices are statically configured to emit signals in a $\theta = 15°$ arc, which is narrower than the Georgia Tech's transceiver.

We validate our model with physical hardware by first measuring how the received signal strength (RSS) varies as a function of the angle between the transmitter and receiver. We then build a real-size floor plan of a Cayley datacenter with a 2 by 3 grid of racks based on Table 2, place transmitter-receiver pairs in their physical locations, and examine whether the signal strength is sufficient for communication (Figure 5.c and d). Finally, we quantify the amount of interference for all possible receiver and transmitter pairs in intra-rack space, in inter-rack space both between adjacent and non-adjacent racks, and in different rack stories. Due to the symmetric circular structure of racks on a regular grid, evaluating a subset of transceiver pairs on the 2 by 3 grid is sufficient to cover all cases.

In the following experiments, we primarily examine RSS as a measure of signal quality in relationship to a vendor-defined base. We configure the transmission power of the Terabeam transmitter for all experiments such that a receiver directly facing the transmitter receives signal at -46dB. This is a conservative level, as the minimum error-free RSS for this hardware is -53dB in a noisy environment [52], and the typical default noise level we measured in a datacenter-like environment was approximately -69dB.

4.1 Received Signal Strength and Facing Directions

The most basic assumption that the Cayley datacenter design makes of the underlying hardware is that a transmitter and a receiver pair can communicate when they are within each other's signal zone. To validate this assumption, we examine the signal strength of a transmitter-receiver pair, placed one meter apart, as a function of the facing angle ε (i.e. $\alpha, \beta = 0°$ and $\delta = 1$ meter in Figure 1). In an ideal scenario with no interference, a receiver would not read any signals when ε exceeds θ.

Figure 7 shows that the received signal strength is significantly above the error-free threshold when $\varepsilon \leq \theta = 15°$ and is negligible when $\varepsilon > 15°$. This confirms that the pair can communicate when oriented in the prescribed manner, and more importantly, that there is negligible interference from a transmitter on an unintended receiver whose reception zone does not cover the transmitter.

4.2 Intra-Rack Space

The cylindrical rack structure we propose divides free space into intra- and inter-rack spaces in order to achieve high free space utilization. Such cylindrical racks would not be feasible if there was high interference within the dense intra-rack space (Figure 6.a). To evaluate if this is the case, we measure the interference within a rack by measuring the signal strength at all receivers during a transmission.

Figure 8 demonstrates that only the receivers within the $15°$ main signal lobe of the transmitter (receivers at positions 9 and 10 for transmitter 0) receive a signal at a reliable level. The rest of the servers do not receive any signal interference. In part, this is not surprising given the previous experiment. But it confirms that any potential side lobes and other leaked signals from the transmitter do not affect the adjacent receivers.

4.3 Orthogonal Inter-Rack Space

Eliminating all wires from a datacenter requires the use of wireless communication between racks. Such communication requires that the signals from nodes on a given rack can successfully traverse the free space between racks. We first examine the simple case of communication between racks placed at $90°$ to each other (Figure 6.b).

Figure 9 shows that a transmitter-receiver pair can communicate between racks only when their signal zones are correctly aligned. For clarity, the graph omits symmetrically equivalent servers and plots the received signal strength of servers 6 to 10 on rack A. Other servers on rack A at positions less than 6 or greater than 14 show no signal from servers 0 to 2 on rack D. The graph shows that server 0 on rack D can transmit effectively to server 10 on rack A without any interference to any other servers, as expected.

4.4 Diagonal Inter-Rack Space

Cayley datacenters take advantage of diagonal links between racks in order to provide link diversity and increase bandwidth. We next validate whether the transceivers in our cylindrical racks can effectively utilize such diagonal paths (Figure 6.c).

Figure 10 shows the received signal strength between diagonally oriented racks, and demonstrates that the intended transmitter-receiver pairs can communicate successfully. Once again, the figure omits the symmetrical cases

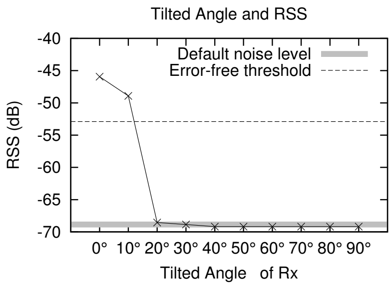


Figure 7: Facing direction of Rx and RSS.

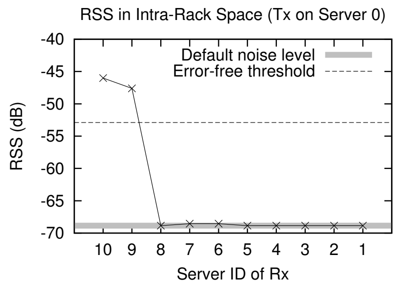


Figure 8: RSS in intra-rack space.

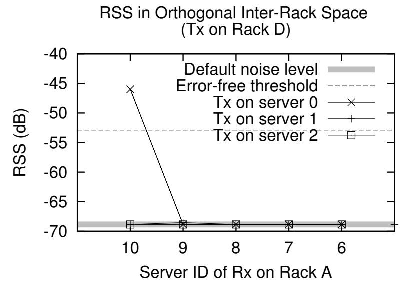


Figure 9: RSS in inter-rack space between racks in orthogonal positions.

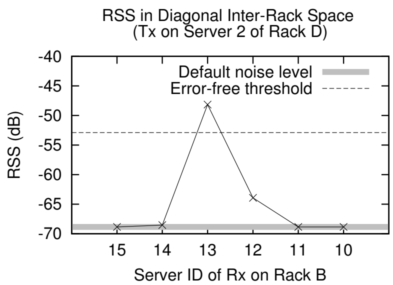


Figure 10: RSS in inter-rack space between racks in diagonal positions.

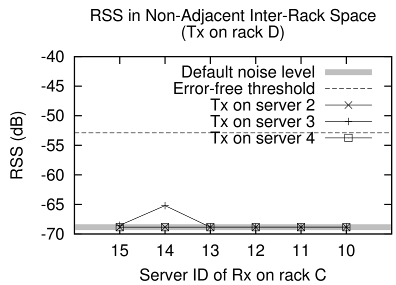


Figure 11: RSS in inter-rack space between non-adjacent racks.

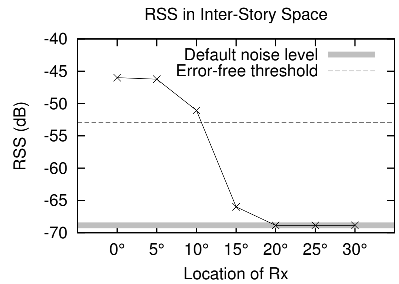


Figure 12: RSS in inter-story space.

(e.g. transmitter on server 3 of rack D), and no signal from far away servers (e.g. 0, 1, 4, 5 of rack D) reaches rack B at all. The signal strength in this experiment is as high as the orthogonal case despite the increased distance due to transmit power adjustment. The case of receiver on server 12 represents an edge case in our model: the signal strength is slightly above the background level because the node is located right at the boundary of the transmission cone. This signal level, while not sufficient to enable reliable communication, can potentially pose an interference problem. To avoid this problem, one can slightly increase the transmitter's signal's angle so that it sends a stronger signal. Alternatively, one can narrow the transmitter's signal angle to eliminate the signal spillover.

4.5 Non-Adjacent Racks

While Cayley datacenters utilize only the wireless links between adjacent racks, it is possible for signals from non-adjacent racks to interfere with each other (Figure 6.d). This experiment examines the attenuation of the signal between non-adjacent racks and quantifies the impact of such interference.

Figure 11 shows the impact of three transmitters on rack D and the receivers on non-adjacent rack C. The transmitters are calibrated to communicate with their adjacent racks B and E. The measurements show that receivers on rack C receive no signal or weak signal not strong enough for communication, but when multiple non-adjacent transmitters send the weak signal (i.e. transmitter on server 3 and receiver on server 14), the noise rate could potentially become too great. For this reason, we propose placing non-reflective curtains, made of conductors such as aluminum or copper foil, that block the unwanted signal. Such curtains can be placed in the empty triangles in Figure 3.b without impeding access.

4.6 Inter-Story Space

Finally, we examine the feasibility of routing packets along the z-axis, between the different stories on racks. To do so, we orient a transmitter-receiver pair exactly as they would be oriented when mounted on prism-shaped servers placed on different stories of a rack, and examine signal strength as the receivers are displaced from 0° to 30° following the z-axis.

Figure 12 shows that the signal is the strongest at the center of the main lobe and drops quickly towards the edge of the signal zone. When the receiver reaches the borderline (15°) of the signal, it only picks up a very weak signal. Once the receiver moves beyond the 15° point, it receives no signal. Overall, the signal strength drops very sharply towards the edge of the signal, and except for the 15° borderline case, transceivers on different stories can reliably communicate.

4.7 Summary

In summary, we have evaluated transceiver pairs in a Cayley datacenter and demonstrated that the signal between pairs that should communicate is strong and reliable, with little interference to unintended receivers. Calibrating the antenna or using conductor curtains can address the few borderline cases when the signal is weaker than expected or where there is potential interference. Although not described in detail, we also tested for potential constructive interference. We verified with two transmitters that even when multiple nodes transmit simultaneously, the signals do not interfere with the unintended receivers, namely the receivers in positions that received negligible or no signal in Figures 7 through 12. Overall, these physical experiments demonstrate that extant 60GHz transceivers achieve the sharp attenuation and well-formed beam that can enable the directed communication topology of a Cayley datacenter, while controlling interference.

5. PERFORMANCE AND COST ANALYSIS

In this section, we quantify the performance, failure resilience, and cost of Cayley datacenters in comparison to a fat-tree and a conventional wired datacenter (CDC).

5.1 Objectives

We seek to answer the following questions about the feasibility of wireless datacenters:

- **Performance**: How well does a Cayley datacenter perform and scale?

  By measuring the maximum aggregate bandwidth and packet delivery latency using a fine-grain packet level simulation model with different benchmarks, we compare the performance with fat-trees and CDCs.

- **Failure resilience**: How well can a Cayley datacenter handle failures?

  Unlike wired datacenters, server failures can affect routing reliability in Cayley datacenters because each server functions as a router. Thus, we measure the number of node pairs that can connect to each other under an increasing number of server failures.

- **Cost**: How cost effective is a Cayley datacenter compared to wired datacenters?

  The wireless transceivers and Y-switches are not yet available in the market. We estimate and parameterize costs based on the technologies that wireless transceivers use and compare the price of a Cayley datacenter with a CDC based on the expected price range of 60GHz transceivers.

5.2 Test Environments

Because datacenters involve tens of thousands of servers and 60GHz transceivers are not yet massively produced, it is impossible to build a full Cayley datacenter at the moment. Therefore, we built a fine-grained packet level simulation to evaluate the performance of different datacenter designs.

We model, simulate, and evaluate the MAC layer protocol including busy tones, routing protocol, and relevant delays in the switches and communication links both for Cayley datacenters and CDCs. From the simulation, we can measure packet delivery latency, packet hops, number of packet collisions, number of packet drops from buffer overflow or timeout and so on. The simulator can construct the 3-dimensional wireless topology depending on the parameters such as the transceiver configurations, the distance between racks, and the size of servers. We also model, simulate, and evaluate the hierarchical topology of a fat-tree and a CDC given the number of ports and oversubscription rate of switches in each hierarchy.

5.3 Base Configurations

Throughout this section, we evaluate Cayley datacenters along with fat-trees and CDCs with 10K server nodes. Racks are positioned in a 10 by 10 grid for Cayley datacenters. We use the smallest configurable signal angle of 25° to maximize the number of concurrent wireless links in the Cayley datacenter and distance of one meter between racks for ergonomic reasons [47].

Cayley datacenter parameter	Value
Inner radius	0.25 (meter)
Outer radius	0.89 (meter)
Distance between racks	1 (meter)
Height of each story	0.2 (meter)
# of servers per story	20
# of stories per rack	5
# of servers per rack	100
Bandwidth per wireless data link	10 Gbps
Bandwidth per wireless control link	2.5 Gbps
Switching delay in Y-switch	4 ns

Table 2: Cayley datacenter configurations

Conventional datacenter parameter	Value
# of servers per rack	40
# of 1 GigE ports per TOR	40
# of 10 GigE port per TOR	2 to 4
# of 10 GigE port per AS	24
# of 10 GigE port per CS sub-unit	32
Buffer per port	16MB
Switching delay in TOR	6 μs
Switching delay in AS	3.2 μs
Switching delay in CS	5 μs

Table 3: Conventional datacenter configurations

For CDCs and fat-trees, we simulate a conservative topology consisting of three levels of switches, top of rack switches (TOR), aggregation switches (AS), and core switches (CS) in a commonly encountered oversubscribed hierarchical tree [13]. Oversubscription rate x indicates that among the total bandwidth, the rate of the bandwidth connecting the lower hierarchy to that connecting the upper hierarchy is $x : 1$. The oversubscription rates in a real datacenter are often larger than 10 and can increase to over several hundred [5,17]. To be conservative, we configure CDCs to have oversubscription rates between 1 and 10, where the rate 1 represents the fat-tree.

The basic configurations for Cayley datacenters and CDCs are described in Tables 2 and 3 respectively. The number of switches used for CDC varies depending on the oversubscription rate in each switch. The configuration and delays for the switches are based on the data sheets of Cisco products [9,10,12].

We focus exclusively on traffic within the datacenter, which account for more than 80% of the traffic even in client-facing web clouds [17]. Traffic in and out of the Cayley datacenter can be accommodated without hot spots through transceivers on the walls and ceiling as well as wired injection points.

5.4 Performance

In this subsection, we measure the key performance characteristics, maximum aggregate bandwidth and average and maximum packet delivery latency of Cayley datacenters, fat-trees and CDCs, using a detailed packet level simulator. The evaluation involves four benchmarks varying the packet injection rates and packet sizes:

- **Local Random**: A source node sends packets to a random destination node within the same pod. The pod of a CDC is set to be the servers and switches connected under the same AS. The pod of a Cayley

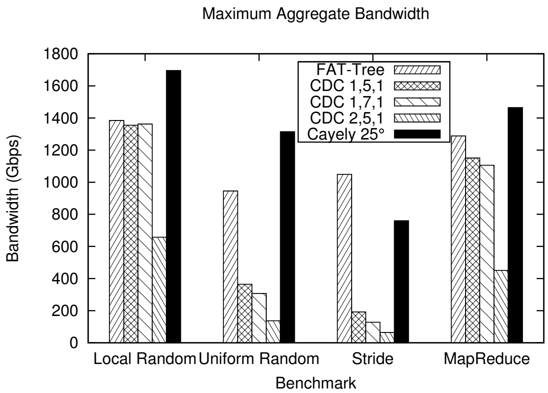


Figure 13: Maximum aggregate bandwidth.

datacenter is set to be the servers in a 3 by 3 grid of racks.

- **Uniform random**: Source and destination nodes for a packet are randomly selected among all nodes with uniform probability.
- **Stride**: Source node with a global ID x sends packets to the destination node with ID $mod(x +$ (total # of servers)/2, total # of servers).
- **MapReduce**: (1) A source node sends messages to the nodes in the same row of its rack. (2) The nodes that receive the messages send messages to the nodes in the same columns of their racks. (3) All the nodes that receive the messages exchange data with the servers in the same pod and outside the pod with 50% probability each. This benchmark resembles the MapReduce application used in Octant [54].

We use different oversubscription rates in each level of switch in the CDC and use three numbers to indicate them: each number represents the rate in TOR, AS, and CS in order. For example, (2,5,1) means the oversubscription rate of TOR is 2, that of AS is 5, and that of CS is 1 and a fat-tree is equivalent to (1,1,1).

5.4.1 Bandwidth

We measure the maximum aggregate bandwidth while every node pair is sending a burst of 500 packets. The results are summarized in Figure 13.

For all cases, the Cayley datacenter shows higher maximum aggregate bandwidth than any CDC. A Cayley datacenter takes advantage of high bandwidth, oversubscription-free wireless channels. The figure clearly shows the disadvantage of having oversubscribed switches in CDCs: when the majority of packets travel outside of a rack or above a AS, as in uniform random and stride, the bandwidth falls below 50% of Cayley datacenter's bandwidth.

Fat-trees perform noticeably better than all CDCs except for local random, where no packet travels above AS's. However, Cayley datacenters outperform fat-trees for all cases except the stride benchmark. Packets from the stride benchmark travel through the largest amount of hop counts, thus it penalizes the performance of the Cayley datacenter.

5.4.2 Packet Delivery Latency

We measure packet delivery latencies by varying the packet injection rate and packet size. Figure 14 and 15 show the average and maximum latencies, respectively. The columns separate the type of benchmarks and the rows divide the packet sizes that we use. Packets per server per second injection rates ranged from 100 to 500.

Local random is the most favorable and stride is the least favorable traffic for all datacenters from a latency point of view: packets travel a longer distance in order of local random, MapReduce, uniform random, and stride.

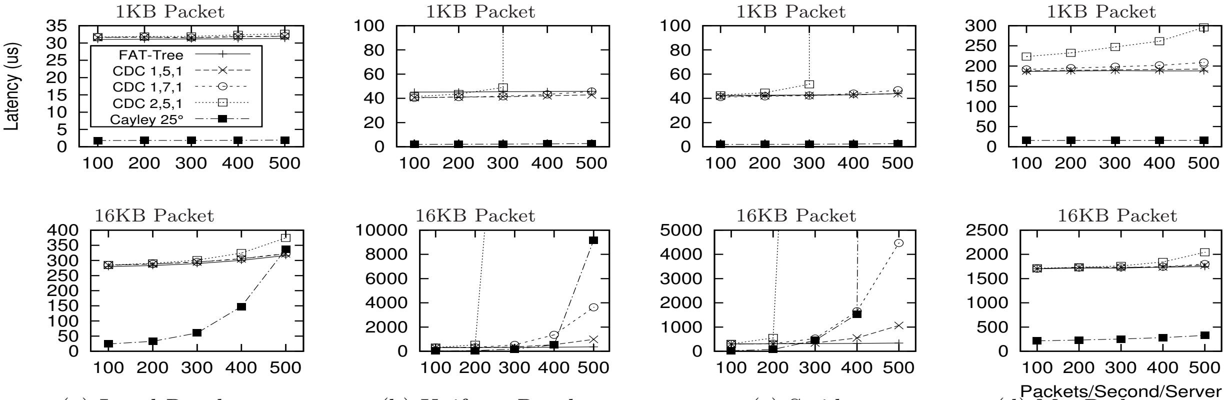
Overall, the average packet delivery latencies of Cayley datacenters are an order of magnitude smaller (17 to 23 times) than those of fat-trees and all CDCs when the traffic load is small. This is because datacenter switches have relatively larger switching delay than the custom designed Y-switch and Cayley datacenters use wider communication channels. For local random and MapReduce benchmarks that generate packets with relatively small network hops (Figure 14.a and d), Cayley datacenters outperform fat-trees and CDCs for almost all cases.

For all other benchmarks, CDC (2,5,1) performs noticeably worse than all others, especially when traffic load is large, because the TOR is oversubscribed. The latency of CDC (2,5,1) skyrockets once uniform random and stride traffic overloads the oversubscribed switches and packets start to drop due to buffer overflow (Figures 14.b and c). Besides CDC (2,5,1), fat-tree and other CDCs maintain relatively stable average latencies except for during the peak load. The amount of traffic increases up to 8MBps per server: 8MBps per server is approximately the same amount of traffic generated per server as the peak traffic measured in an existing datacenter [32].

Cayley datacenters generally maintain lower latency than fat-trees and CDCs. The only case when the Cayley datacenters' latency is worse, is near the peak load. When running uniform random and stride benchmarks under the peak load, Cayley datacenters deliver packets slower than fat-tree, CDC (1,5,1), and CDC (1,7,1) (the last row of Figures 14.b and c). The numbers of average network hops for a Cayley datacenter are 11.5 and 12.4 whereas those of the tree-based datacenters are 5.9 and 6 for uniform random and stride benchmarks. Competing for a data channel at each hop with relatively large packets significantly degrades the performance of Cayley datacenters compared to fat-trees and CDC (1,5,1) and (1,7,1).

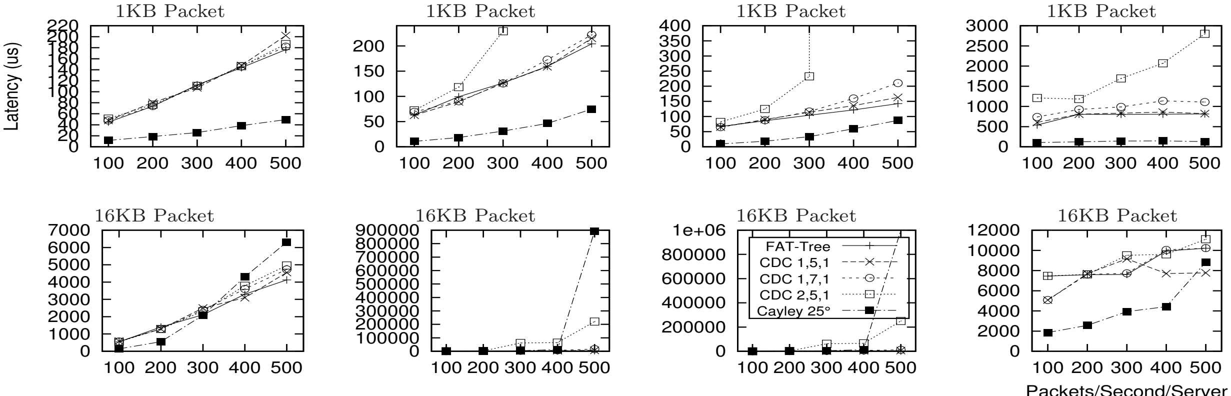
The maximum packet delivery latency shows the potential challenge in a Cayley datacenter (Figure 15). Although the average latencies are better than CDCs, Cayley datacenters show a relatively steep increase in maximum latency as traffic load increases. Therefore, the gap between average and maximum latency for packet delivery becomes larger depending on the amount of traffic. However, except for under the peak traffic load, the maximum latency of the Cayley datacenter is less than 3.04 times as large as the latency of a fat-tree, and is smaller than CDCs for most cases. Therefore, Cayley datacenters are expected to show significantly better latency on average than fat-tree and CDCs, except under peak load for applications similar to stride.

In summary, except for handling the peak traffic for uniform random and stride benchmark, the Cayley datacenter performance is better than or comparable to fat-tree and CDC. As the average number of hops per packet increases, the performance of Cayley datacenters quickly decreases. This shows that Cayley datacenters may not also be as scalable as CDC, which has stable wired links with



(a) Local Random (b) Uniform Random (c) Stride (d) MapReduce

Figure 14: Average packet delivery latency.



(a) Local Random (b) Uniform Random (c) Stride (d) MapReduce

Figure 15: Maximum packet delivery latency.

smaller number of network hops. Cayley datacenters may not be suitable to handle applications requiring large number of network hops per packet, but this type of applications also penalizes the CDC performance as we observed for CDC (2,5,1). In reality, datacenter applications such as MapReduce usually resembles the local random benchmark, which does not saturate oversubscribed (aggregate) switches [5,32]. Further, the experimental results demonstrate that Cayley datacenters perform the best for MapReduce. Consequently, Cayley datacenters may be able to speed up a great portion of datacenter applications. Even for larger scale datacenters, engineering the application's traffic pattern as in [3] will enable applications to run in Cayley datacenters more efficiently than in fat-trees and CDCs.

5.5 Failure Resilience

We evaluate how tolerant Cayley datacenters are to failures by investigating the impact of server failures on connections between live nodes (Figure 16). We select the failing nodes randomly in units of individual node, story, and rack. We run 20 tests for each configuration and average the results. The average of standard deviation for the 20 run is less than 6.5%.

Server nodes start to disconnect when 20%, 59%, and 14% of the nodes, stories, and racks fail, respectively. However, over 99% of the network connections are preserved until more than 55% of individual nodes or stories fail. Over 90% of the connections are preserved until 45% of racks fail.

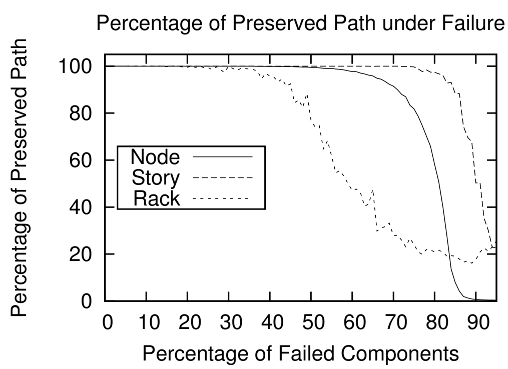


Figure 16: Percentage of preserved path under failure.

Assuming failure rates of servers are the same in wireless datacenters as fat-tree based datacenters and CDCs, then a Cayley datacenter can be more resilient to network failures. This is mainly because wireless datacenters do not have conventional switches which can be critical points of failure and the failures catastrophic enough to partition a Cayley datacenter is very rare [17].

5.6 Cost Comparison

It is complicated to compare two technologies when one is commercially mature and the other is yet to be commer-

Config	#TOR	#AS	#CS	#CS chassis	Cost ($)
2,5,1	250	26	8	1	1,818,500
1,7,1	250	48	12	2	2,229,000
1,5,1	250	52	16	2	2,437,000
fat-tree	250	88	96	10	6,337,000

Table 4: CDC networking equipment cost for 10K nodes

cialized. We can easily measure the cost of a fat-tree and a CDC, but the cost of a Cayley datacenter is not accurately measurable. However, we parameterize the costs of Cayley datacenters and compare the cost for different values of 60GHz transceiver cost.

Hardware cost: We compare the cost of the wireless and the wired datacenters based on the network configurations that we used so far. The price comparison can start from the NIC—typically priced at several tens of dollars [42]—and the Y-switch. In our system, we replace the NIC with the proposed simple Y-switch and at least two transceivers. Y-switches consist of simple core logic, host interface, such as a PCI express bus, and interface controllers. Thus, we expect the price of a Y-switch to be comparable to a NIC.

The price differences between wireless and wired datacenters stem from the wireless transceivers and the switches. The cost required for CDC and fat-tree to connect 10K servers based on the price of TOR, AS, and CS [44] are summarized in Table 4. The total price ranges from US$1.8M to US$2.4M for CDCs and US$6.3M for a fat-tree. Since the cost of a fat-tree can be very high, it should be able to use commodity switches [38] and the cost can vary much depending on the switch configuration. Thus, we mainly focus on the comparison between CDCs and Cayley datacenters.

60GHz transceivers are expected to be inexpensive, due to their level of integration, usage of mature silicon technologies (90nm CMOS), and low power consumption which implies low-cost packaging. We cannot exactly predict the market price, but the total cost of network infrastructure excluding the Y-switch in Cayley datacenters can be expressed as a function,

$$Cost_{Cayley}(cost_t, N_{server}) = 2 \times cost_t \times N_{server}, \qquad (2)$$

where $cost_t$ is the price for a transceiver and N_{server} is the number of servers in a datacenter. From this function, we can find out that as long as $cost_t$ is less than US$90, Cayley datacenters can connect 10K servers with lower price than a CDC. Similarly, if $cost_t$ becomes US$10, the cost of transceivers in Cayley datacenters can be 1/9 of CDC switches. Considering the rapidly dropping price of silicon chips [18] we expect the transceiver's price to quickly drop to less than US$90 even if it starts with a high cost. This comparison excludes the wire price for CDC, so there is an additional margin, where $cost_t$ can grow higher to achieve lower cost than CDC.

Power consumption: The maximum power consumption of a 60GHz transceiver is less than 0.3 watts [43]. If all 20K transceivers on 10K servers are operating at their peak power, the collective power consumption becomes 6 kilowatts. TOR, AS, and a subunit of CS typically consume 176 watts, 350 watts, and 611 watts, respectively [9–11]. In total, wired switches typically consumes 58 kilowatts to 72 kilowatts depending on the oversubscription rate for datacenter with 10K servers. Thus, a Cayley datacenter can consume less than 1/12 to 1/10 of power to switch packets compared to a CDC.

Besides the lower price and power, lower maintenance costs stemming from the absence of wires and substantially increased tolerance to failure can be a strong point for wireless datacenters. In summary, we argue that 60GHz could revolutionize datacenter construction and maintenance.

6. PUTTING IT ALL TOGETHER

The summary of our findings throughout the evaluation of Cayley datacenters are as follows. The merits of completely wireless Cayley datacenters over fat-trees and conventional datacenters are:

- **Ease of maintenance through inherent fault tolerance**: Densely connected wireless datacenters have significantly greater resilience to failures than wired datacenters, in part because they do not have switches which can cause correlated loss of connectivity and in part because the wireless links provide great path diversity. Additionally, installing new or replacing failed components can be easier than in a CDC, since only rewiring power cables is necessary.
- **Performance**: Cayley datacenters can perform better than or comparable to fat-trees and CDCs. Cayley datacenters achieve the highest maximum aggregate bandwidth for most benchmarks and deliver packets at a significantly lower latency, especially for MapReduce-like benchmarks and when traffic load is moderate.
- **Cost**: The price of networking components in a Cayley datacenter is expected to be less than those in CDC depending on the market price of wireless transceivers for comparable performance. Power consumption and expected maintenance costs are significantly lower than CDC.

Characteristics and limitations of Cayley datacenters are:

- **Interference**: Orientation of transceivers on the cylindrical racks and characteristics of 60GHz signals limit the interference and enable reliable communication.
- **MAC layer contention**: Sharing of the wireless channel followed by MAC layer contention greatly influence the overall performance: the lower the contention, the greater the performance.
- **Hop count**: Performance depends on the number of network hops, because each hop entails MAC layer arbitration.
- **Scalability**: Due to the multi hop nature of the topology, scalability is not as good as CDC. Yet, this limitation can be overcome by tuning applications to exhibit spatial locality when possible.

These points summarize the challenges, open problems, opportunities, benefits, and feasibility for designing a wireless datacenter.

7. RELATED WORK

Ramachandran et al. [45] outlined the benefits and challenges for removing wires and introducing 60GHz communication within a datacenter and Vardhan et al. [53] explored the potentials of 60GHz antennae emulating an existing tree-based topology. We share many of their insights and also conclude that 60GHz wireless networks can improve conventional datacenters. Further, we address some of the problems identified by the authors. We propose a novel rack-level architecture, use real 60GHz transceivers and realistic parameters, and provide an extensive evaluation of the performance of the proposed wireless datacenters.

Although we focused on Georgia Tech's transmitter design [43], other research groups are also developing CMOS-based 60GHz transceivers [15, 51]. While the technology was developed initially for home entertainment and mobile devices, other groups are looking at deploying it more broadly [41]. Our work on building completely wireless datacenters extends this line of research and tests the limits of 60GHz technology.

Flyways [22] and [35] are wireless networks based on 60GHz or 802.11n organized on top of wired datacenter racks. They provide supplementary networks for relieving congested wired links or for replacing some of the wired switches. In contrast, wireless links are the main communication channels in Cayley datacenters.

Zhang et al. [58] proposed using 3D beamformation and ceiling reflection of 60GHz signals in datacenters using networks like Flyways to reduce interference. Cayley datacenters use cone-shape 3D beams, but use a novel cylindrical rack design to isolate signals and avoid interference.

A scalable datacenter network architecture by Al-Fares et al. [2] and Portland [38] employ commodity switches in lieu of expensive high-performance switches in datacenters and provide a scalable oversubscription-free network architecture. They achieve high performance at a lower cost, but significantly increase the number of wires.

CamCube consists of a 3-dimensional wired torus network and APIs to support application specific routing [3]. Although the motivation and goal of our paper is different from those of CamCube, combining their approach of application specific routing is expected to enhance the performance of our Cayley datacenter design.

The MAC layer protocol that we used [23, 24] is not developed specifically for Cayley datacenters; as a result, there may be inefficiencies that arise. Alternatively, there are other MAC layer protocols developed specifically for 60GHz technology and directional antennae [7, 33, 48], but they require global arbitrators or multiple directional antennae collectively pointing to all directions. These are not suitable for datacenters. Designing a specialized MAC layer protocol for wireless datacenters is an open problem.

While our design adopted XY routing for Cayley datacenters, other variations of routing protocols for interconnecting networks, such as [19, 21, 29], can be adapted to our design.

8. CONCLUSION

In this paper, we proposed a radically novel methodology for building datacenters which displaces the existing massive wired switching fabric, with wireless transceivers integrated within server nodes.

For brevity and simplicity of presentation, we explore the design space under the assumption that certain parameters such as topology and antenna performance are constant. Even in this reduced search space, we identify the strong potential of Cayley datacenters: while maintaining higher bandwidth, Cayley datacenters substantially outperform conventional datacenters and fat-trees with respect to latency, reliability, power consumption, and ease of maintenance. Issues that need further improvements are extreme scalability and performance under peak traffic regimes.

Cayley datacenters open up many avenues for future work. One could focus on each aspect of systems research related to datacenters and their applications and try to understand the ramifications of the new architecture. We feel that we have hardly scratched the surface of this new paradigm and that numerous improvements are attainable. Some interesting design considerations involve understanding the cost structure of individual nodes and how it scales with applications: is it beneficial to parallelize the system into a substantially larger number of low-power low-cost less-powerful processors and support hardware? What data replications models yield best reliability vs. traffic overhead balance? Could an additional global wireless network help with local congestion and MAC-layer issues such as the hidden terminal problem? What topology of nodes resolves the max-min degree of connectivity across the network? How should software components be placed within the unique topology offered by a Cayley datacenter? How does performance scale as the communication sub-band shifts higher in frequency? Would some degree of wired connectivity among servers internal to a single rack benefit performance? As the 60GHz technology matures, we expect many of the issues mentioned here to be resolved and novel wireless networking architectures to be realized in datacenters.

9. ACKNOWLEDGMENTS

We'd like to thank the Georgia Tech team for providing us with the specifications and the communication model of the 60GHz wireless transceivers; Han Wang for helping implement the Y-switch for FPGAs; Srikanth Kandula, Jitendra Padhye, Victor Bahl, Dave Harper, and Dave Maltz for providing us with the 60GHz antennae for physical validations; Daniel Halperin for his help in experimenting with the 60GHz antennae; Bobby Kleinberg for helping name the project; and Deniz Altinbuken and Tudor Marian for insightful feedback on an earlier version of this paper. This work was supported in part by National Science Foundation grants No. 0424422, 1040689, 1053757, 1111698, and SA4897-10808PG.

10. REFERENCES

[1] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, and M. Zaharia. A view of cloud computing. *Communications of the ACM*, 53(4), 2010.

[2] M. Al-Fares, A. Loukissas, and A. Vahdat. A scalable, commodity data center network architecture. *SIGCOMM*, 2008.

[3] H. Abu-Libdeh, P. Costa, A. Rowstron, G. O'Shea, and A. Donnelly. Symbiotic routing in future data centers. *SIGCOMM*, 2010.

[4] R. Buyya, C. S. Yeo, and S. Venugopal. Market-oriented cloud computing: vision, hype, and reality for delivering IT services as computing utilities. *HPCC*, 2008.

[5] T. Benson, A. Akella, and D. A. Maltz. Network traffic characteristics of data centers in the wild. *IMC*, 2010.

[6] A. Cayley. On the theory of groups. *American Journal of Mathematics*, 11(2), 1889.

[7] X. Chen, J. Lu, and Z. Zhou. An enhanced high-rate WPAN MAC for mesh networks with dynamic bandwidth management. *GLOBECOM*, 2005.

[8] J. I. Choi, M. Jain, K. Srinivasan, P. Levis, and S. Katti. Achieving single channel, full duplex wireless communication. *MOBICOM*, 2010.

[9] Cisco. Cisco Nexus 5000 series architecture: the building blocks of the unified fabric. http://www.cisco.com/en/US/prod/collateral /switches/ps9441/ps9670/white_paper_c11-462176.pdf, 2009.

[10] Cisco. Cisco Catalyst 4948 switch. http://www.cisco.com/en/US/prod/collateral/switches/ps5718/ ps6021/product_data_sheet0900aecd8017a72e.pdf, 2010.

[11] Cisco. Cisco Nexus 7000 series environment. http://www.cisco.com/en/US/prod/collateral/switches/ps9441/ ps9402/ps9512/Data_Sheet_C78-437759.html, 2010.

[12] Cisco. Cisco Nexus 7000 F-series modules. http://www.cisco.com/en/US/prod/collateralswitches/ps9441/ ps9402/at_a_glance_c25-612979.pdf, 2010.

[13] Cisco. Cisco data center infrastructure 2.5 design guide. http://www.cisco.com/en/US/docs/solutions/Enterprise/ Data_Center/DC_Infra2_5/DCI_SRND_2_5_book.html, 2010.

[14] Ecma International. Standard ECMA-387: high rate 60GHz PHY, MAC and HDMI PAL. http://www.ecma-international.org/publications/standards/Ecma-387.htm, 2008.

[15] B. Floyd, S. Reynolds, U. Pfeiffer, T. Beukema, J. Grzyb, and C. Haymes. A silicon 60GHz receiver and transmitter chipset for broadband communications. *ISSCC*, 2006.

[16] A. Greenberg, J. Hamilton, D. A. Maltz, and P. Patel. The cost of a cloud: research problems in data center networks. *SIGCOMM Compututer Communication Review*, 39(1), 2008.

[17] A. Greenberg, J. R. Hamilton, N. Jain, S. Kandula, C. Kim, P. Lahiri, D. A. Maltz, P. Patel, and S. Sengupta. VL2: a scalable and flexible data center network. *SIGCOMM*, 2009.

[18] C. Gianpaolo, D. M. Xavier, S. Regine, V. W. L. N., and W. Linda. Inventory-driven costs. *Harvard Business Review*, 83(3), 2005.

[19] P. Gratz, B. Grot, and S.W. Keckler. Regional congestion awareness for load balance in networks-on-chip. *HPCA*, 2008.

[20] S. Gollakota, S. D. Perli, and D. Katabi. Interference alignment and cancellation. *SIGCOMM*, 2009.

[21] C. J. Glass, L. M. Ni, and L. M. Ni. The turn model for adaptive routing. *ISCA*, 1992.

[22] D. Halperin, S. Kandula, J. Padhye, P. Bahl, and D. Wetherall. Augmenting data center networks with multi-gigabit wireless links. *SIGCOMM*, 2011.

[23] 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.

[24] Z. Huang, C.-C. Shen, C. Srisathapornphat, and C. Jaikaeo. A busy-tone based directional MAC protocol for ad hoc networks. *MILCOM*, volume 2, 2002.

[25] HXI. http://www.hxi.com, 2012.

[26] IEEE 802.15 Working Group for WPAN. http://www.ieee802.org/15/.

[27] E. Kranakis, H. Singh, and J. Urrutia. Compass routing on geometric networks. *Canadian Conference on Computational Geometry*, 1999.

[28] H. Krawczyk. LFSR-based Hashing and Authentication. *CRYPTO*, 1994.

[29] J. Kim, D. Park, T. Theocharides, N. Vijaykrishnan, and C. R. Das. A low latency router supporting adaptivity for on-chip interconnects. *DAC*, 2005.

[30] P. Kermani and L. Kleinrock. Virtual cut-through: a new computer communication switching technique. *Computer Networks*, 3, 1979.

[31] P. Karn. MACA - a new channel access method for packet radio. *ARRL Computer Networking Conference*, 1990.

[32] S. Kandula, S. Sengupta, A. Greenberg, P. Patel, and R. Chaiken. The nature of data center traffic: measurements & analysis. *IMC*, 2009.

[33] T. Korakis, G. Jakllari, and L. Tassiulas. A MAC protocol for full exploitation of directional antennas in ad-hoc wireless networks. *MobiHoc*, 2003.

[34] V. Kvicera and M. Grabner. Rain attenuation at 58 GHz: prediction versus long-term trial results. *EURASIP Journal on Wireless Communications and Networking*, 2007(1), 2007.

[35] Y. Katayama, K. Takano, N. Ohba, and D. Nakano. Wireless data center networking with steered-beam mmWave links. *WCNC*, 2011.

[36] M. M. Khodier and C. G. Christodoulou. Linear array geometry synthesis with minimum sidelobe level and null control using particle swarm optimization. *IEEE Transactions on Antennas and Propagation*, 53(8), 2005.

[37] C. E. Leiserson. Fat-trees: universal networks for hardware-efficient supercomputing. *IEEE Transaction on Computers*, 34(10), 1985.

[38] R. N. Mysore, A. Pamboris, N. Farrington, N. Huang, P. Miri, S. Radhakrishnan, V. Subramanya, and A. Vahdat. PortLand: a scalable fault-tolerant layer 2 data center network fabric. *SIGCOMM*, 2009.

[39] Mathworks. Simulink—simulation and model-based design. http://www.mathworks.com/products/simulink/.

[40] J. Nsenga, W. V. Thillo, F. Horlin, A. Bourdoux, and R. Lauwereins. Comparison of OQPSK and CPM for communications at 60 GHz with a nonideal front end. *EURASIP Journal on Wireless Communications and Networking*, 2007(1), 2007.

[41] A. M. Niknejad. Siliconization of 60GHz. *IEEE Microwave Magazine*, 11(1), 2010.

[42] newegg.com. Intel PWLA8391GT 10/100/1000 Mbps PCI PRO/1000 GT desktop adapter 1 x RJ45. http://www.newegg.com/Product/Product.aspx? Item=N82E16833106121, 2012.

[43] 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.

[44] PEPPM. Cisco Current Price List. http://www.peppm.org/Products/cisco/price.pdf, 2012.

[45] K. Ramachandran, R. Kokku, R. Mahindra, and S. Rangarajan. 60 GHz data-center networking: wireless => worry less? *NEC Technical Report*, 2008.

[46] S. Ray, J.B. Carruthers, and D. Starobinski. Evaluation of the masked node problem in ad hoc wireless LANs. *IEEE Transactions on Mobile Computing*, 4(5), 2005.

[47] J.-Y. Shin, E. G. Sirer, H. Weatherspoon, and D. Kirovski. On the feasibility of completely wireless data centers. *Cornell CIS Tech Report*, 2011.

[48] S. Singh, F. Ziliotto, U. Madhow, E.M. Belding, and M.J.W. Rodwell. Millimeter wave WPAN: cross-layer modeling and multi-hop architecture. *INFOCOM*, 2007.

[49] SiBeam White Paper. Designing for high definition video with multi-gigabit wireless technologies. http://www.sibeam.com/whtpapers/ Designing_for_HD_11_05.pdf, 2005.

[50] K.W. Tang and R. Kamoua. Cayley pseudo-random (CPR) protocol: a novel MAC protocol for dense wireless sensor networks. *WCNC*, 2007.

[51] M. Tanomura, Y. Hamada, S. Kishimoto, M. Ito, N. Orihashi, K. Maruhashi, and H. Shimawaki. TX and RX front-ends for 60GHz band in 90nm standard bulk CMOS. *ISSCC*, 2008.

[52] Terabeam. Terabeam Gigalink field installation and service manual. 040-1203-0000 Rev B, 2003.

[53] H. Vardhan, N. Thomas, S.-R. Ryu, B. Banerjee, and R. Prakash. Wireless data center with millimeter wave network. *GLOBECOM*, 2010.

[54] B. Wong, I. Stoyanov, and E. G. Sirer. Octant: a comprehensive framework for the geolocalization of internet hosts. *NSDI*, 2007.

[55] WG802.11 - Wireless LAN Working Group. http://standards.ieee.org/develop/ project/802.11ad.html.

[56] Wireless Gigabit Alliance. http://wirelessgigabitalliance.org, 2010.

[57] S. K. Yong and C.-C. Chong. An overview of multigigabit wireless through millimeter wave technology: potentials and technical challenges. *EURASIP Journal on Wireless Communications and Networking*, 2007(1), 2007.

[58] W. Zhang, X. Zhou, L. Yang, Z. Zhang, B. Y. Zhao, and H. Zheng. 3D Beamforming for Wireless Data Centers. *HotNets*, 2011.