On the Feasibility of Completely Wireless Data Centers

Paper #22, 14 pages

ABSTRACT

Conventional data centers, 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 data centers 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 data centers built with this methodology can potentially attain higher aggregate bandwidth, lower latency, and substantially higher fault tolerance than a conventional wired data center while improving ease of construction and maintenance.

1. Introduction

Performance, reliability, cost of the switching fabric, power consumption, and maintenance are some of the issues that plague conventional wired data centers [2, 16, 17]. Current trends in cloud computing and high-performance data center applications indicate that these issues are likely to be exacerbated in the future [1, 5].

In this paper, we explore a radical change to the construction of data centers 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, 42] 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 data center? 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 data center design—because its network connectivity subgraphs belong to a class of Cayley graphs [7], we call our design a Cayley data center. 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; the connections together form a densely connected mesh. To achieve this, we replace the network interface card (NIC) of a modern 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, aggregation switches, copper and/or 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 data center. We show that it is possible for 60GHz technology to serve as the sole and central means of communication in the demanding data center 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 data centers. Using a set of 60GHz transceivers, we demonstrate that signals in Cayley data centers do not interfere with each other. We also show that, compared to a fat-tree [37, 38] and a conventional data center, our proposal exhibits higher bandwidth, substantially improved latency due to the switching fabric being integrated into server nodes, lower power consumption, and easier maintenance as a result of the plug-and-play simplicity of connecting servers. Cayley data centers exhibit strong fault tolerance due to the routing scheme that can fully explore the mesh: a Cayley data centers can maintain connectivity to over 99% of live nodes until up to 31% of total racks or 55% of total nodes fail.

The remainder of this paper explores the assumptions, feasibility and technical challenges related to our proposal.

Section 2 provides background information regarding the 60GHz wireless technology and Section 3 presents the proposed wireless data center architecture. Section 4 details evaluation of interference using 60GHz transceivers in a Cayley data center design and Section 5 describes the performance evaluation of our proposal. Section 6 summarizes our findings, Section 7 outlines the related work and Section 8 concludes the paper.

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 data center.

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. To date, 60GHz as a technology has been mostly pursued as a wireless replacement for HDMI (high-definition multimedia interface) connections [46]. 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, 51], WiGig 7Gbps standard with beam-forming [52], and ECMA-387/ISO DS13156 6.4Gbps spec [15] based upon Georgia Tech's design [42].

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 [42].

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:

$$\begin{aligned} M &= P_{TX} + G_{TX+RX} \\ &\quad -L_{Fade} - L_{Implementation} \\ &\quad -FSL - NF - SNR, \end{aligned} \tag{1}$$

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 [53]. 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

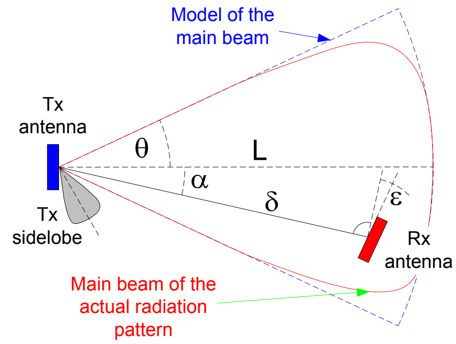


Figure 1: Geometric communication model.

$BER = \frac{1}{2}\mathrm{erfc}(\sqrt{SNR})$ 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 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 $25° \leq \theta \leq 45°$ as a constraint that simplifies antenna design with suppressed side lobes. Based on design parameters from the antenna prototypes developed by the same team, we limit ε and γ to be smaller than θ, 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 data center design and lead to a higher bandwidth and less noisier transceiver design compared to off-the-shelf 60GHz transceivers for HDMI. We validate the assumptions behind these parameters in Section 4 with a conservative 60GHz hardware prototype built by Terabeam/HXI [25]. More research in 60GHz RF design with a focus on Cayley data centers can further improve performance.

[1] Typically, reconnection after beam-steering involves training of communication codebooks involving delays on the order of microseconds. This may be tolerated in home networking scenarios but not in the data center.

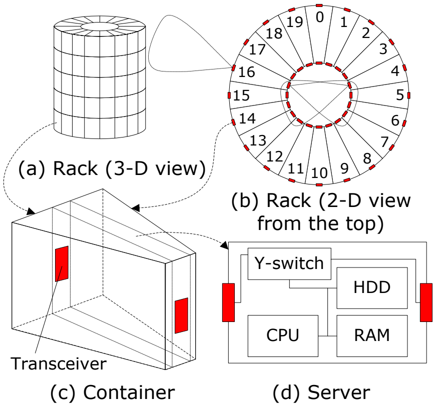


Figure 2: Rack and server design.

3. Cayley Data Center Design

This section introduces Cayley data center architecture, the positioning of the 60GHz transceivers in a wireless data center, 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 problem.

3.1 Component Design

In order to maximize opportunities for resource multiplexing in a wireless data center, 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 data center divided by that of a single antenna beam. We focus on the network topology that would optimize key performance characteristics, namely latency and bandwidth.

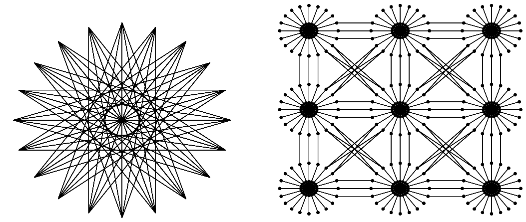
To separate the wireless signals for communications within a rack and among different racks, we propose cylindrical racks (Figure 2.a) which store servers in prism-shaped containers (Figure 2.c). This choice is appealing, because it partitions the data center volume into two regions: intra- and inter-rack 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 sequentially starting from the 12 o'clock position from 0 to 19 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 [7] 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 [47].



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

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

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 data center 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 connect the subgraphs together. Relatively long lines connecting the transceivers on the exterior of the racks show the wireless inter-rack connections. 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 data centers should enable quick routing decisions, utilize 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 data centers 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 data centers use this identity to address the servers. Once a data center administrator manually configures the identity of several

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)

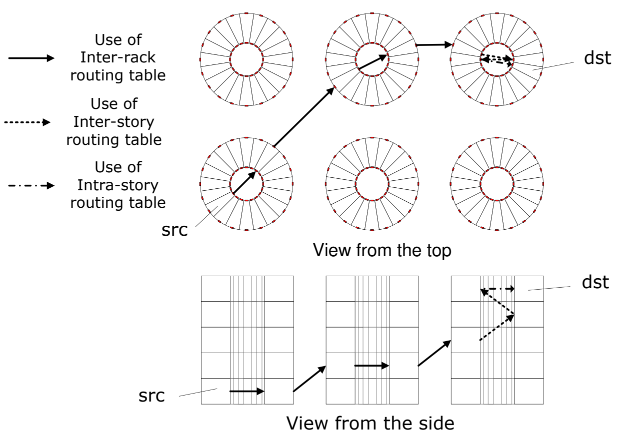


Figure 4: Diagonal XYZ routing example.

servers, the rest of the servers can identify their identities by querying the neighbors and propagating the information.

The geographical identity facilitates finding a path in the Cayley data center 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 and Figure 4 illustrates an example of using this 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, LookupGeoIDs(dir, $g_{dst}.s$) returns geographical 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 data centers 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 [9, 20] – the directional beams and multiple transceivers in Cayley data center servers permit this optimization.

3.3.2 *Adaptive Routing in Case of Failure*

Compared to a conventional data center, a Cayley data center has a distinct failure profile. Conventional data centers are dependent on switches for network connectivity and consequently a switch failure can disconnect many servers. Cayley data centers, 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 [4] (anonymized for blind review) shows that the routing scheme can circumvent randomly failed racks with less than 5us latency overhead in a Cayley data center.

3.4 MAC Layer Arbitration

A transceiver in a Cayley data center 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 transceiver act as hidden terminals for each other. A challenge in a Cayley data center, unusual in common wireless communication, is illustrated in Figure 5. Assume that a transceiver chip 1 is communicating with chip A and chip

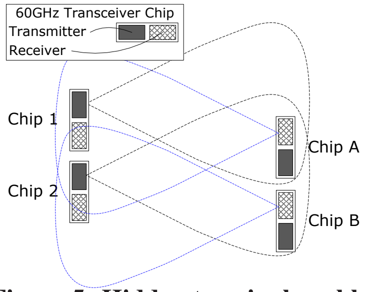


Figure 5: Hidden terminal problem.

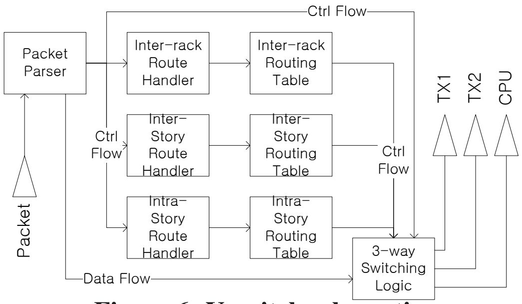


Figure 6: Y-switch schematic.

2 is trying to send signals to chip B. Employing a regular ready-to-send (RTS) /clear-to-send (CTS) based MAC protocol [31] will allow chip B to approve chip 2 to send signals without noticing that chip A is receiving. This can interfere with the communication between chip 1 and chip A.

To mitigate the hidden terminal problem, 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 enables hidden terminal nodes both at the sending and receiving end to know whether other nodes are already using the wireless channel. In the above example, chip A's busy tone will suppress chip 2 from initiating communication. Since DBTMA uses RTS/CTS handshake, erroneous packet reception can be suppressed as well and will have benign effects.

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 data center. High-level schematic of this switch is shown in Figure 6. 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[2]. 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.

[2]Note that the Y-switches could also share the main memory resident on the server to buffer packets if necessary.

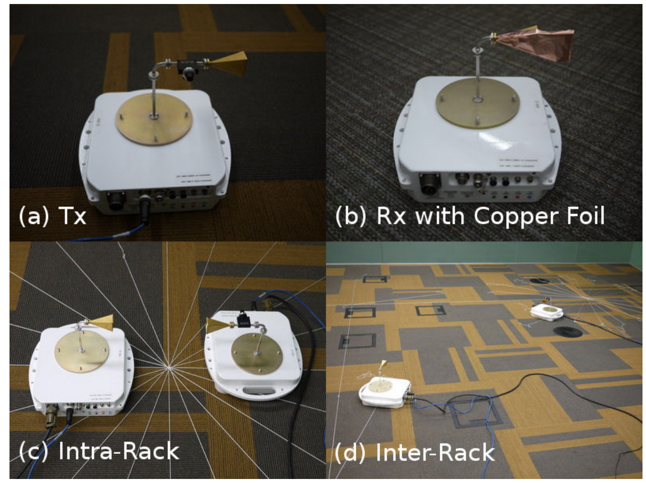


Figure 7: 60GHz Tx, Rx, and measurements on a Cayley data center floor plan

4. Physical Validation

Before evaluating the performance of Cayley data centers, 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 data center.

We conduct our experiments using Terabeam/HXI 60GHz transmitters [25] (Figure 7.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, 42]. To compensate for the noise stemming from the older horn design, we augment one side of the receiver's horn with a copper foil (Figure 7.b). The devices are statically configured to emit signals in a $\theta = 15^\circ$ 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 data center with a 2 by 3 grid of racks based on Table 2, place transmitter-receiver pairs in their physical locations, and examine whether signal strength is sufficient for communication (Figure 7.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.

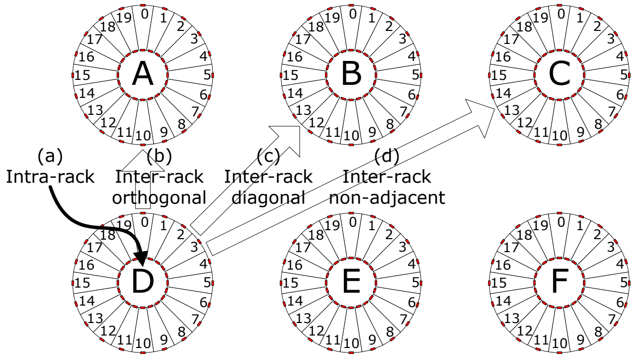


Figure 8: Interference measurement summary

In the following experiments, we primarily examine RSS as a measure of signal quality in relationship to a vendor-defined base[3]. 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 [48], and the typical default noise level we measure in a data-center-like environment was approximately -69dB.

4.1 Received Signal Strength and Facing Directions

The most basic assumption that the Cayley data center design makes of the underlying hardware is that a transmitter and a receiver can communicate when their conical signal zones contains each other. To validate this assumption, we examine the signal strength of a transmitter-receiver pair, placed 1 meter apart, as a function of the facing angle ε. In an ideal scenario with no interference, a receiver would not read any signals when ε exceeds θ.

Figure 9 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 proscribed manner, and more importantly, that there is negligible interference from a transmitter on an unintended receiver whose signal zone does not cover the transmitter.

4.2 Intra-Rack Space

The cylindrical rack structure we propose effectively divides free-space into intra- and inter-rack spaces in order to achieve high free space utilization. These cylindrical racks would not be feasible if there was high interference within the dense intra-rack space (Figure 8.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 10 demonstrates that only the receivers within the 15° main signal lobe of the transmitter (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 data center 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 8.b).

Figure 11 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 signal strength on servers 6-10. Servers on rack A at positions less than 6 or greater than 14 show no signal received. 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 from the inter-rack distance of 1 meter and the signal angle of 15°.

4.4 Diagonal Inter-Rack Space

Cayley data centers 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 8.c).

Figure 12 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 (e.g. 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 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 solve 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 data centers utilize only the wireless links between adjacent racks, it is possible for signals from non-adjacent racks to interfere with each other (Figure 8.d). This experiment examines the attenuation of the signal between non-adjacent racks and quantifies the impact of such interference.

Figure 13 shows the impact of three transmitters on rack D and the non-adjacent rack C. The transmitters are calibrated to communicate with their adjacent racks B and E. The mea-

[3]The raw RSS value we get from the interface is the received signal strength indicator. Due to some missing form factors, which the vendor did not provide, we perform a translation from RSSI to decibels using polynomial regression based on known RSSI to dB mappings. The coefficient of determination R^2 we get from this regression is 0.999993.

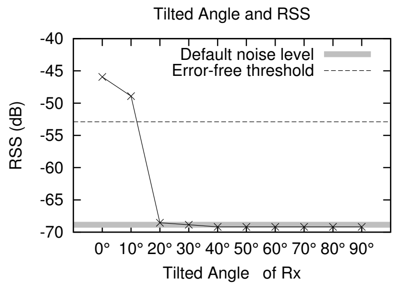


Figure 9: Facing direction of Rx and RSS.

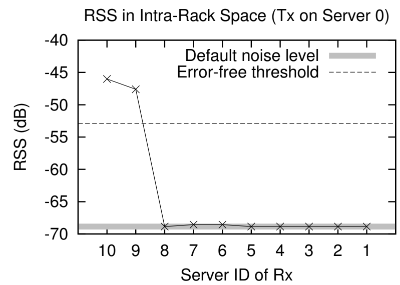


Figure 10: RSS in intra-rack space.

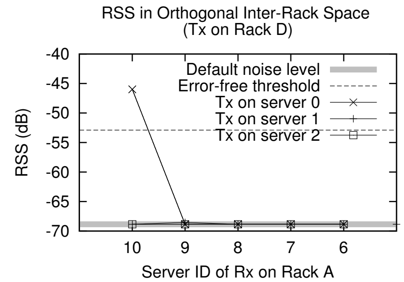


Figure 11: RSS in inter-rack space between racks in orthogonal positions.

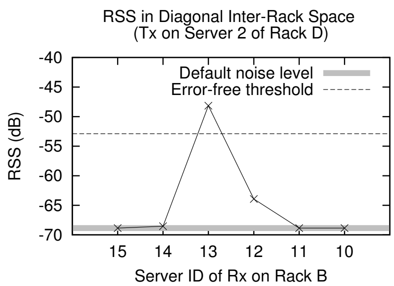


Figure 12: RSS in inter-rack space between racks in diagonal positions.

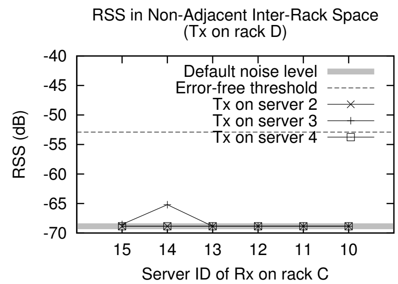


Figure 13: RSS in inter-rack space between non-adjacent racks.

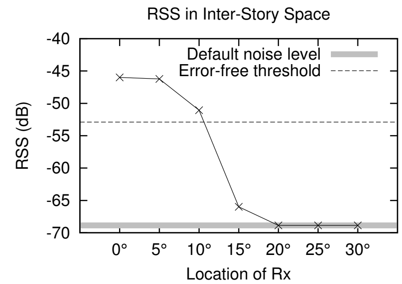


Figure 14: RSS in inter-story space.

surements show that rack C receives no 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 nodes are displaced from 0° to 30° from the orthogonal line.

Figure 14 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 data center 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 that received negligible or no signal in Figures 9 through 14. 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 data center while controlling interference.

5. Performance and Cost Analysis

In this section, we explore the design space and quantify the performance, failure resilience properties, and cost of Cayley data centers in comparison to a fat-tree and a conventional wired data center (CDC).

5.1 Objectives

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

- **Design space**: What are the factors that influence the performance of a Cayley data center?

  We perform an analysis on Cayley data center topologies while varying parameters such as signal's angle and distance between racks. By comparing the measured values and the input, we analyze characteristics of parameters and find suitable configurations.

- **Performance**: How well does a Cayley data center 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 data center handle failures?

  Unlike wired data centers, server failures can affect routing reliability in Cayley data centers 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 data center compared to wired data centers?

  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 data center with a CDC based on the expected price range of 60GHz transceivers.

5.2 Test Environments

Because data centers involve tens of thousands of servers and 60GHz transceivers in Section 2 are not yet massively produced, it is impossible to build a full Cayley data center at the moment. Therefore, we built a fine-grained packet level simulation to evaluate the performance of different data center 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 data centers 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 data centers along with fat-trees and CDCs with 10K server nodes. Racks are positioned in a 10 by 10 grid for Cayley data centers. 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), are used in fat-trees and CDCs in a commonly encountered oversubscribed hierarchical tree [14]. 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 data center are often larger than 10 and can increase to over several hundred [6, 17]. To be conservative, we configure CDCs to have oversubscription rates between 1 and 10, where the rate 1 represents the fat-tree.

Cayley data center parameter	Value
Inner radius	0.25 (meter)
Outer radius	0.89 (meter)
Distance between racks (regular)	1 (meter)
Distance between racks (close)	0.5 (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 data center configurations

Conventional data center 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 data center configurations

The basic configurations for Cayley data centers 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 [10, 11, 13].

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

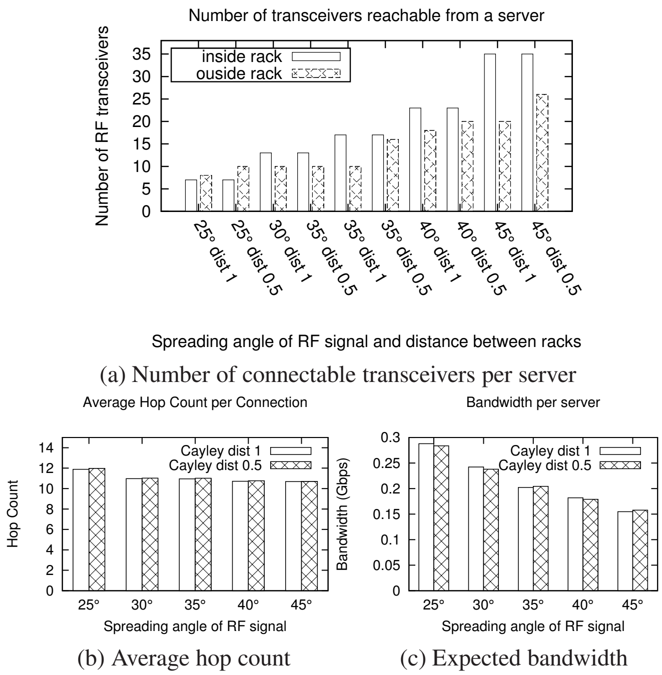
5.4 Exploration of Wireless Data Center Design Space

Before we run any simulations, we explore the design space of a Cayley data center by analyzing the topology to figure out a suitable configuration for the data center.

We examine the maximum number of servers that a server can communicate with when varying the signal's angle and the distance between racks. Figure 15.a shows that the larger the angle and the longer the distance, the larger the number of servers that a server can communicate with. Having larger number of servers to communicate with from a server can imply smaller network hop counts, but it can increase the MAC layer contention.

We measure the hop counts when using the diagonal XYZ routing between random node pairs. Different from our expectation, the average hop counts when varying the parameters do not change much depending on the parameters (Figure 15.b). This is because our geographical routing protocol takes place in the inter-rack level.

Next, we measure the expected bandwidth achievable per server by counting the number of random flows that can coexist (Figure 15.c). When using a 25° signal, the bandwidth is 12% higher than using a 30° signal, and 86% higher than using a 45° signal. The decrease in performance from



(a) Number of connectable transceivers per server

(b) Average hop count

(c) Expected bandwidth

Figure 15: Analysis of Cayley data center topologies.

wider signal angle is mainly due to increased MAC layer contention. Meanwhile, the distance between racks shows small influence on the overall performance.

In summary, the RF signal angle does not significantly affect the hop count, but it can affect the overall performance by influencing the amount of MAC layer contention.

5.5 Performance

In this subsection, we measure the key performance characteristics, the maximum aggregate bandwidth and average and maximum packet delivery latency of Cayley data centers, 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 data center 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 $mod(x + (\text{total \# of servers})/2,\ \text{total \# of servers})$.
- **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 [50], where server nodes compute and exchange information with other nodes during the reduce phase.

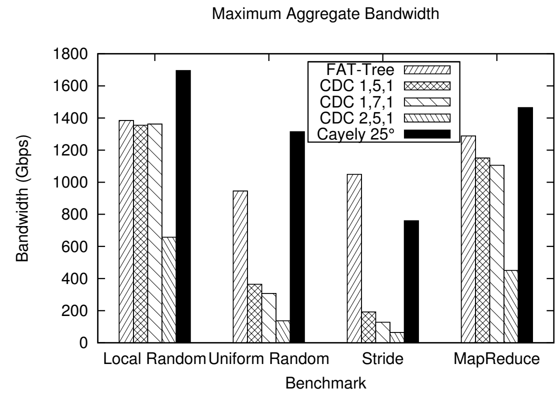


Figure 16: Maximum aggregate bandwidth.

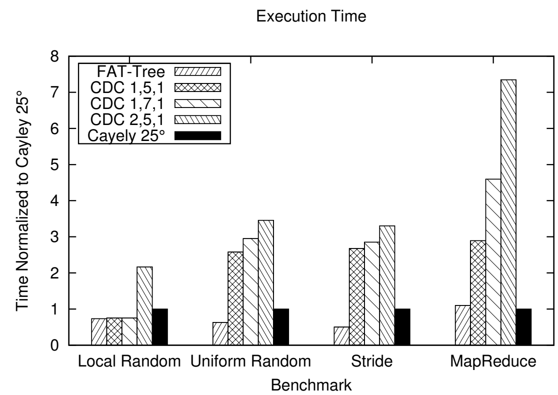


Figure 17: Execution time

We use a signal angle of 25° and distance of 1 meter between racks for Cayley data centers. We choose this configuration because 25° showed the best result in the expected bandwidth analysis of previous subsection and 1 meter distance between the racks is more ergonomic than 0.5 meters for people to walk through data centers.

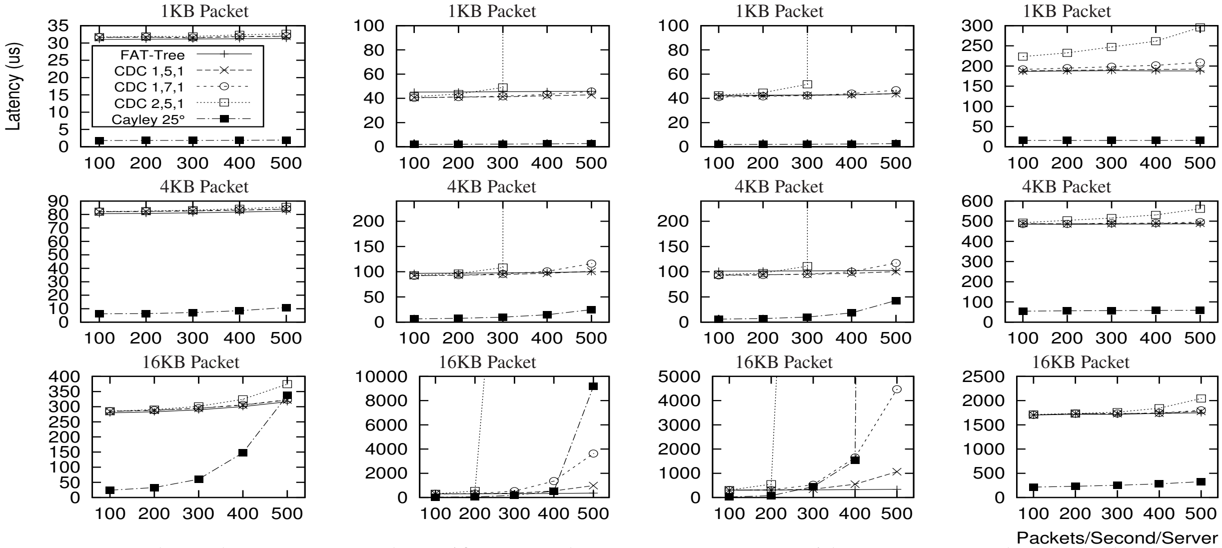
We use different oversubscription rates in each level of switch in the CDC and use three numbers to indicate them: each number represents the rate in TOR, AS, and CS in order. For example, (2,5,1) means the oversubscription rate of TOR is 2, that of AS is 5, and that of CS is 1 and a fat-tree is equivalent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

5.5.1 Bandwidth

We measure the maximum aggregate bandwidth while every node pair is sending a burst of 500 packets[4]. The results are summarized in Figure 16.

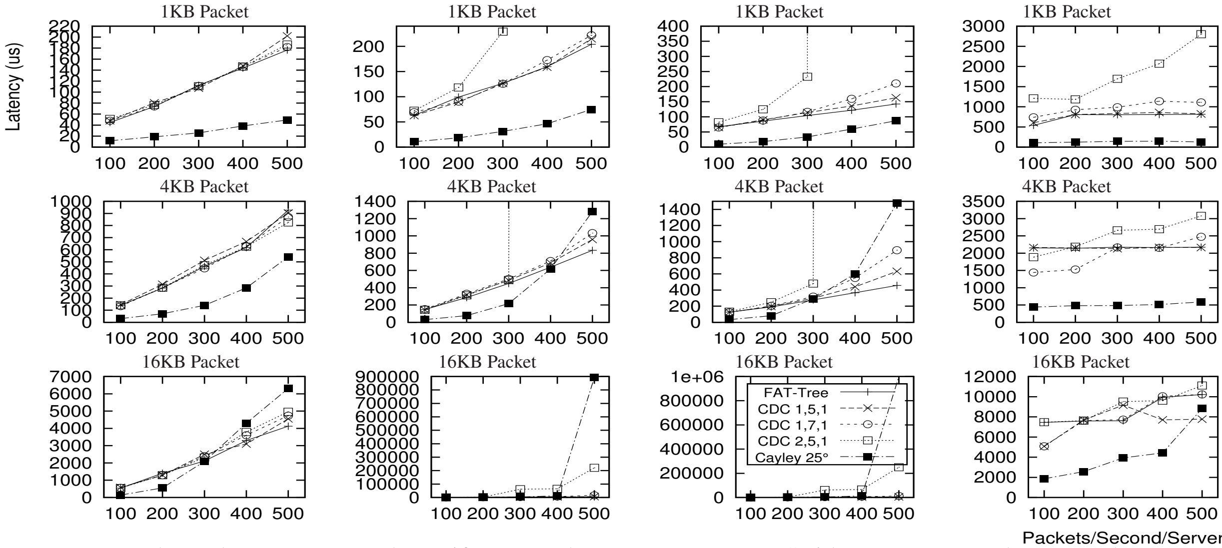
For all cases, the Cayley data center shows higher maximum aggregate bandwidth than any CDC. A Cayley data center takes advantage of high bandwidth, oversubscription-free wireless channels. The figure clearly shows the disadvantage of having oversubscribed switches in CDCs: when the majority of packets travel outside of a rack or above a

[4]We configure the MapReduce benchmark to generate equivalent amount of packets.



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

Figure 18: Average packet delivery latency.



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

Figure 19: Maximum packet delivery latency.

AS, as in uniform random and stride, the bandwidth falls below 50% of Cayley data center's bandwidth.

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

Although Cayley data centers generally show higher maximum aggregate bandwidth for most cases, they have long bandwidth tails. Comparing the execution time, Cayley data centers generally have shorter execution time than CDCs but slightly longer execution time than fat-trees (Figure 17). The issue is that the MAC layer contention allows only one transceiver to send at a time among 7 to 8 others that share the overlapping signal space. Still, for the realistic MapReduce benchmark, the Cayley data center performs the best.

5.5.2 Packet Delivery Latency

We measure packet delivery latencies by varying the packet injection rate and packet size. Figure 18 and 19 show the average and maximum latencies, respectively. The columns separate the type of benchmarks and the rows divide the packet sizes that we use for the experiments. 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 data centers 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 data centers 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 data center switches have relatively larger switching delay than the custom designed

Y-switch and Cayley data centers use wider communication channels. For local random and MapReduce benchmarks that generate packets with relatively small network hops (Figure 18.a and d), Cayley data centers outperform fat-trees and CDCs for almost all cases.

For all other benchmarks, CDC (2,5,1) performs noticeably worse than all others specially 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 starts to drop due to buffer overflow (Figures 18.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 in this plot 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 data center [32].

Cayley data centers generally maintain lower latency than fat-tree and CDCs. The only case when Cayley data centers' latency quickly degrades is near the peak load. When running uniform random and stride benchmarks under the peak load, Cayley data centers deliver packets slower than fat-tree, CDC (1,5,1), and CDC (1,7,1) (the last row of Figures 18.b and c). The numbers of average network hop count for a Cayley data center are 11.5 and 12.4 whereas those of the tree-based data centers 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 data centers 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 data center (Figure 19). Although the average latencies are better than CDCs, Cayley data centers 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 Cayley data centers 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 data centers 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, Cayley data centers perform better than or comparable to fat-trees and CDCs. As the average number of hops per packet increases, the performance of Cayley data centers quickly decreases. This shows that Cayley data centers may not also be as scalable as CDC, which has stable wired links with smaller number of network hops. Cayley data centers may not be suitable to handle applications requiring large number of network hops per packet, but this type of applications also penalize the CDC performance as we observed for CDC (2,5,1). In reality, data center applications are usually not designed to generate traffic resembling uniform random or stride benchmarks. In particular, MapReduce rather resembles the local random benchmark which does not saturate oversubscribed (aggregate) switches [6,32] and the experimental results also demonstrate that Cayley data centers perform the best for MapReduce. Consequently, Cayley data centers may be able to speed up a great portion of data center applications. Even for larger scale data centers, engineering the application's traffic pattern as in [3] will enable applications to run in Cayley data centers more efficiently than in fat-trees and CDCs.

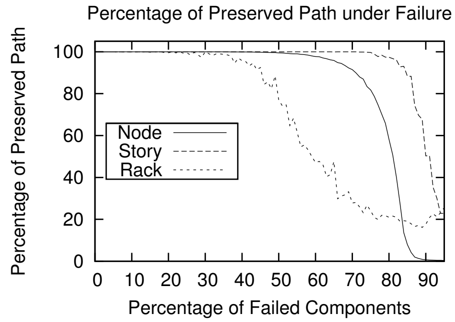


Figure 20: Percentage of preserved path under failure.

5.6 Failure Resilience

We evaluate how tolerant Cayley data centers are to failures by investigating the impact of server failures on connections between live nodes (Figure 20). 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 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. Assuming failure rates of servers are the same in wireless data centers, fat-tree based data centers and CDCs, a Cayley data center can be more resilient to network failures. This is mainly because wireless data centers do not have conventional switches which can be critical points of failure and the failures catastrophic enough to partition a Cayley data center is very rare [17].

5.7 Cost Comparison

It is complicated to compare two technologies when one is commercially mature and the other is yet commercialized. We can easily measure the cost of a fat-tree and a CDC, but the cost of a Cayley data center is not accurately measurable. However, we parameterize the costs of Cayley data centers and compare the cost for different values of 60GHz transceiver cost.

Hardware cost: We compare the cost of the wireless and the wired data centers 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 [41]—and the Y-switch. In our system, we replace the NIC with the

Component	Price ($)	Min Unit	Min unit price ($)
TOR	8000	1	4,800
AS	9,000	1	10,000
CS			358,500
CS subunit	42,000	1	42,000
CS chassis	12,000	1	12,000
CS power supply	3500	3	10,500

Table 4: CDC switches [43]

Config	#TOR	#AS	#CS sub-unit	#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 5: CDC networking equipment cost for 10K nodes

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 data centers stem from the wireless transceivers and the switches. The prices of TOR, AS, and CS, and the cost required for CDC and fat-tree to connect 10K servers using data center switches are summarized in Tables 4 and Table 5. 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 as in [38] and the cost can vary much depending on the switch configuration. For this reason, we mainly focus on the comparison between CDCs and Cayley data centers.

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 data centers can be expressed as a function,

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

where $cost_t$ is the price for a transceiver and N_{server} is the number of servers in a data center. From this function, we can find out that as long as $cost_t$ is less than US$90, Cayley data centers can connect 10K servers with lower price than a CDC. Similarly, if $cost_t$ becomes US$10, the cost of transceivers in Cayley data centers 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[5] 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 [42]. 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 [10–12]. In total, wired switches typically consumes 58 kilowatts to 72 kilowatts depending on the oversubscription rate for data center with 10K servers. Thus, a Cayley data center 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 coming from the absence of wires and substantially increased tolerance to failure can be a strong point for wireless data centers. In summary, we argue that 60GHz could revolutionize data center construction and maintenance.

[5]Fully integrated 60GHz wireless HDTV transceiver boxes are currently US$130 at an online store. (http://www.amazon.com/RocketfishTM-4-Port-Wireless-Kit-Rf-whd200/dp/B004HSU07E)

6. Putting It All Together

The summary of our findings throughout the evaluation of Cayley data centers are as follows. The merits of Cayley, or wireless, data centers over fat-trees and CDCs are:

- **Ease of maintenance through inherent fault tolerance**: Densely connected wireless data centers have significantly greater resilience to failures than wired data centers, 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 data centers can perform better than or comparable to fat-trees and CDCs. Cayley data centers 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 data center 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 data centers 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 wireless channel followed by MAC layer contention greatly influence the overall performance: the lower the contention the greater the performance.
- **Hop count**: The performance depends on the number of network hops, because each hop entails MAC layer arbitration. However, the signal's angle does not greatly affect the overall hop count in Cayley data centers.
- **Scalability**: Due to multi hop nature of the topology, the scalability is not as good as CDC. Yet, this can be overcome by tuning the applications.

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

7. Related Work

Ramachandran et al. [44] outlined the benefits and challenges for removing wires and introducing 60GHz communication within a data center and Vardhan et al. [49] 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 data centers. 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 data centers.

Flyways [22] and [35] are wireless networks based on 60GHz or 802.11n organized on top of wired data center 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 data centers.

Zhang et al. [54] proposed using 3D beamformation and ceiling reflection of 60GHz signals in data centers using networks like Flyways to reduce interference. Cayley data centers use cone-shape 3D beams, but use a novel cylindrical rack design to isolate signals and avoid interference.

A scalable data center network architecture by Al-Fares et al. [2] and Portland [38] employ commodity switches in lieu of expensive high-performance switches in data centers and provide a scalable oversubscription-free network architecture. They achieve high performance at low cost, but at the cost of larger 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 data center design.

The MAC layer protocol that we used [23, 24] is not developed specifically for Cayley data centers; as a result, there may be inefficiencies that arise. Alternatively, there are other MAC layer protocols developed specifically for 60GHz technology and directional antennae [8, 33, 45], but they require global arbitrators or multiple directional antennae collectively pointing to all directions. These are not suitable for data centers. Designing a specialized MAC layer protocol for wireless data centers is an open problem.

While our design adopted XY routing for Cayley data centers, other variations of routing protocols for interconnection network, such as [19, 21, 29], can be adapted to our design.

8. Conclusion

In this paper, we proposed a radically novel methodology for building data centers 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 data centers: while maintaining higher bandwidth, Cayley data centers substantially outperform conventional data centers 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 data centers open up many avenues for future work. One could focus on each aspect of systems research related to data centers 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 of 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 data center? 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 data centers.

9. 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), Apr 2010.

[2] M. Al-Fares, A. Loukissas, and A. Vahdat. A scalable, commodity data center network architecture. *ACM SIGCOMM Conference on Data Communication*, Aug 2008.

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

[4] Anonymized. Anonymized for blind review.

[5] R. Buyya, C. S. Yeo, and S. Venugopal. Market-oriented cloud computing: vision, hype, and reality for delivering IT services as computing utilities. *IEEE International Conference on High Performance Computing and Communications*, Sep 2008.

[6] T. Benson, A. Akella, and D. A. Maltz. Network traffic characteristics of data centers in the wild. *Conference on Internet Measurement*, Nov 2010.

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

[8] X. Chen, J. Lu, and Z. Zhou. An enhanced high-rate WPAN MAC for mesh networks with dynamic bandwidth management. *Conference on Global Telecommunications*, Nov 2005.

[9] J. I. Choi, M. Jain, K. Srinivasan, P. Levis, and S. Katti. Achieving single channel, full duplex wireless communication. *International Conference on Mobile Computing and Networking*, Sep 2010.
[10] 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.
[11] Cisco. Cisco Catalyst 4948 switch. http://www.cisco.com/en/US/prod/collateral/switches/ps5718/ ps6021/product_data_sheet0900aecd8017a72e.pdf, 2010.
[12] Cisco. Cisco Nexus 7000 series environment. http://www.cisco.com/en/US/prod/collateral/switches/ps9441/ ps9402/ps9512/Data_Sheet_C78-437759.html, 2010.
[13] Cisco. Cisco Nexus 7000 F-series modules. http://www.cisco.com/en/US/prod/collateralswitches/ps9441/ ps9402/at_a_glance_c25-612979.pdf, 2010.
[14] 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, Mar 2010.
[15] Ecma International. Standard ECMA-387: high rate 60GHz PHY, MAC and HDMI PAL. http://www.ecma-international.org/publications/standards/Ecma-387.htm, Dec 2008.
[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), Dec 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. *ACM SIGCOMM Conference on Data Communication*, Aug 2009.
[18] C. Gianpaolo, D. M. Xavier, S. Regine, V. W. L. N., and W. Linda. Inventory-driven costs. *Harvard Business Review*, 83(3), Oct 2005.
[19] P. Gratz, B. Grot, and S.W. Keckler. Regional congestion awareness for load balance in networks-on-chip. *International Symposium on High Performance Computer Architecture*, Feb 2008.
[20] S. Gollakota, S. D. Perli, and D. Katabi. Interference alignment and cancellation. *ACM SIGCOMM Conference on Data Communication*, Aug 2009.
[21] C. J. Glass, L. M. Ni, and L. M. Ni. The turn model for adaptive routing. *International Symposium on Computer Architecture*, May 1992.
[22] D. Halperin, S. Kandula, J. Padhye, P. Bahl, and D. Wetherall. Augmenting data center networks with multi-gigabit wireless links. *ACM SIGCOMM Conference on Data Communication*, Aug 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), Jun 2002.
[24] Z. Huang, C.-C. Shen, C. Srisathapornphat, and C. Jaikaeo. A busy-tone based directional MAC protocol for ad hoc networks. *International Conference for Military Communications*, 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*, Aug 1999.
[28] H. Krawczyk. LFSR-based Hashing and Authentication. *International Cryptology Conference on Advances in Cryptology*, Aug 1994.
[29] J. Kim, D. Park, T. Theocharides, N. Vijaykrishnan, and C. R. Das. A low latency router supporting adaptivity for on-chip interconnects. *Design Automation Conference*, Jul 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*, Sep 1990.
[32] S. Kandula, S. Sengupta, A. Greenberg, P. Patel, and R. Chaiken. The nature of data center traffic: measurements & analysis. *Conference on Internet Measurement*, Nov 2009.
[33] T. Korakis, G. Jakllari, and L. Tassiulas. A MAC protocol for full exploitation of directional antennas in ad-hoc wireless networks. *International Symposium on Mobile Ad Hoc Networking & Computing*, Jun 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), Jan 2007.
[35] Y. Katayama, K. Takano, N. Ohba, and D. Nakano. Wireless data center networking with steered-beam mmWave links. *IEEE Wireless Communications and Networking Conference*, Mar 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), Aug 2005.
[37] C. E. Leiserson. Fat-trees: universal networks for hardware-efficient supercomputing. *IEEE Transaction on Computers*, 34(10), Oct 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. *ACM SIGCOMM Conference on Data Communication*, Aug 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), Jan 2007.
[41] 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, Jan 2012.
[42] 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), Oct 2009.
[43] PEPPM. Cisco > Current Price List. http://www.peppm.org/Products/cisco/price.pdf, Jan 2012.
[44] K. Ramachandran, R. Kokku, R. Mahindra, and S. Rangarajan. 60 GHz data-center networking: wireless => worry less? *NEC Technical Report*, Jul 2008.
[45] S. Singh, F. Ziliotto, U. Madhow, E.M. Belding, and M.J.W. Rodwell. Millimeter wave WPAN: cross-layer modeling and multi-hop architecture. *IEEE International Conference on Computer Communications*, May 2007.
[46] SiBeam White Paper. Designing for high definition video with multi-gigabit wireless technologies. http://www.sibeam.com/whtpapers/ Designing_for_HD_11_05.pdf, Nov 2005.
[47] K.W. Tang and R. Kamoua. Cayley pseudo-random (CPR) protocol: a novel MAC protocol for dense wireless sensor networks. *IEEE Wireless Communications and Networking Conference*, Mar 2007.
[48] Terabeam. Terabeam Gigalink field installation and service manual. 040-1203-0000 Rev B, Dec 2003.
[49] H. Vardhan, N. Thomas, S.-R. Ryu, B. Banerjee, and R. Prakash. Wireless data center with millimeter wave network. *Conference on Global Telecommunications*, Dec 2010.
[50] B. Wong and E. G. Sirer. Approximate Matching for Peer-to-PeerOverlays with Cubit. *Computing and Information Science Technical Report, Cornell University*, Dec 2008.
[51] WG802.11 - Wireless LAN Working Group. http://standards.ieee.org/develop/ project/802.11ad.html.
[52] Wireless Gigabit Alliance. http://wirelessgigabitalliance.org, 2010.
[53] 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), Jan 2007.
[54] W. Zhang, X. Zhou, L. Yang, Z. Zhang, B. Y. Zhao, and H. Zheng. 3D Beamforming for Wireless Data Centers. *Workshop on Hot Topics in Networks*, Oct 2011.