On the Feasibility of Completely Wireless Data Centers

Ji-Yong Shin† Emin Gün Sirer† Hakim Weatherspoon† Darko Kirovski‡

Department of Computer Science†
Cornell University
{jyshin, egs, hweather}@cs.cornell.edu

Microsoft Research Redmond‡
darkok@microsoft.com

ABSTRACT

We introduce a novel data center design based on emerging 60 GHz RF technology that uses wires only to deliver power to its server nodes. Fundamental limitation of wireless data centers is that the maximum number of live connections in the network is directly proportional to the full volume occupied by the data center divided by the radiating volume of a single antenna beam. Consequently, we integrate wireless transceivers and switching logic within each server node and collocate them in cylindric racks to establish a semi-regular mesh topology. Our exploration of the resulting design space shows that while attaining comparable bandwidth, our wireless data center exhibits substantially higher fault tolerance, improved latency, lower power consumption, and easier maintenance than a conventional wired data center.

1. INTRODUCTION

Performance, reliability, cost of the switching fabric, power consumption, and difficulty of maintenance are some of the issues that plague the conventional wired data center [17, 18, 2]. 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 workhorse of communication in this new design is a directional, beam-formed 60 GHz RF communication channel characterized by high bandwidth ($4-15$ Gbps) and short range (≤ 10 meters). New 60 GHz modems [39, 37] 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 (up to 10 meters long) beams [39] employed by these modems enable a large number of transmitters to simultaneously communicate with multiple receivers in tight confined spaces.

The unique characteristics of 60 GHz modems pose new challenges and tradeoffs. The most critical questions are those of feasibility and structure: can such networks compete with conventional wired networks? How should the network be architected to achieve high aggregate bandwidth, low cost and high fault tolerance? How should the transceivers be placed and how should the racks be oriented to build practical, robust and maintainable 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 integrate the switching fabric into the server nodes and arrange them into a densely connected, low-stretch, failure-resilient topology. 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. 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. This 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 60 GHz 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 new geographic routing protocol. Finally, we examine the performance and system characteristics of Cayley data centers. Compared to a conventional wired data center, we show that at a comparable total bandwidth, our proposal exhibits 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. Our evaluation shows that the performance in a Cayley data center degrades only under peak traffic load for benchmarks mostly sending packets to servers using large number of network hops. Cayley data centers exhibit strong fault tolerance due to the routing scheme that can fully explore the mesh: it maintains connectivity to all live nodes until up to 14% of total racks fail, or until up to 59% 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 60 GHz wireless technology. Section 3 presents the proposed wireless data center architecture, and Section 4 details the technical evaluation of our proposal. Section 5 outlines the related work and Section 6 summarizes our findings.

2. 60 GHZ WIRELESS TECHNOLOGY

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

Propagation of RF (radio frequency) signals in the $57-64$ GHz sub-band is severely attenuated because of the resonance of oxygen molecules, which limits the use of this sub-band to relatively short distances [33]. Consequently $57-64$ GHz is unlicensed under FCC rules and open to short-range point-to-point applications. To date, 60 GHz as a technology has been mostly pursued as a wireless replacement for HDMI (high-definition multimedia interface) connections [44]. 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 [24, 46], WiGig 7Gbps standard with beam-forming [47], and ECMA-387/ISO DS13156 6.4Gbps spec [14] based upon Georgia Tech's design [39].

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

Characteristic	GEDC
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: 60 **GHz Wireless Transceiver Characteristics. [39].**

A link margin, M, models communication between a transmitter, Tx, and a receiver, Rx. This is the difference between the received power at which the receiver stops working and the actual received power and can be expressed as follows:

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

where P_{TX} and G_{TX+RX} represent transmitted power and overall joint Tx and Rx gain which is dependent upon the geometric alignment of the Tx↔Rx antennae [48]. 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 $BER = \frac{1}{2}\mathrm{erfc}(\sqrt{SNR})$ for binary phase-shift keying (BPSK) modulation for example. Loss to fading and implementation are constants given a specific system. From Equation 1, one could compute the effects of constraining communication parameters. In general, we point to two important trade-offs:

- **Complexity of transceiver vs. constraints imposed on the main and side-lobes**: requirements such as narrow main-lobe and beam-steering complicate transceiver designs and make designs that suppress side-lobes significantly more challenging. In addition, reconnections due to beam-steering usually entail a non-trivial latency cost[1] that may be tolerated in home networking scenarios but not in the data center.
- **Main-lobe volume vs. space multiplexing**: the design of the main-lobe is important because it balances the degree of network connectivity. High degree of connectivity is expected to adversely affect resource multiplexing and conversely poor connectivity increases the average message latency.

Figure 1 illustrates a planar slice of the geometric communication model we consider in this paper. A Tx 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. In 60 GHz, attenuation from oxygen molecules, quickly dissipates power beyond the limits of the lobe—thus, a necessary condition for Tx to communicate with any Rx antenna is for the Rx antenna to be within the bounds of Tx's beam. The beam is modeled as a cone with an angle θ and length L. Using a spherical coordinate system centered at Tx's antenna, one can define the position of the Rx antenna with its radius, δ, elevation α, and azimuth β. The plane of the Rx antenna can then be misaligned from the plane of the Tx antenna by an angle ε along the elevation and γ along the azimuth. We use a modeling tool developed at Georgia Tech to convert $\{\alpha, \beta, \gamma, \varepsilon, \delta, L\}$ into G_{TX+RX}. Through personal communication with Georgia Tech's design team, we reduced our space of interest to $25^\circ \leq \alpha \leq 45^\circ$ as a constraint that simplifies antenna design to efficiently suppress sidelobes. Based on the design parameters from the antenna prototypes developed by the same team, we limit ε and γ to be smaller than α, and assume that a BER of 10^{-9} at lengths shorter than $L < 3$ meters while transmitting at 10Gbps at less than one Watt. We do not assume beam-steering is available. Finally, we assume that the available bandwidth can be multiplexed using both time division (TDD) and frequency division duplexing (FDD).

[1]Typically, reconnection involves training of communication codebooks involving delays on the order of microseconds.

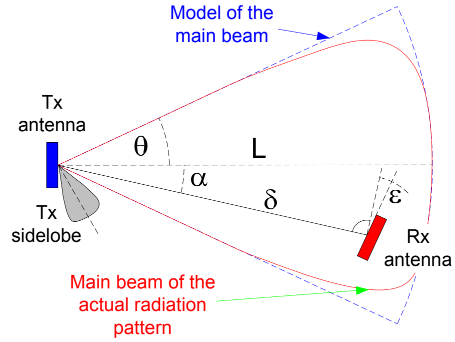


Figure 1: Geometric communication model.

In summary, for a pair of Tx and Rx antennae that conform to the geometric requirements presented in the previous paragraph, we model their connection as a consistent 10Gbps link with a BER of 10^{-9}. We assume that the model is fully symmetric and consistent across the data center, unless discussed otherwise. Technically, this model conforms to the design achievements by the Georgia Tech team.

3. CAYLEY DATA CENTER DESIGN

This section introduces the architecture of the Cayley data center and how it affects the positioning of the 60 GHz transceivers in a wireless data center which in turn defines the network topology. We also introduce a novel 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, there exists a fundamental limitation of wireless data centers: the maximum number of live connections in the network is directly proportional to the full volume occupied by the data center divided by the radiating volume of a single antenna beam. We focus on the

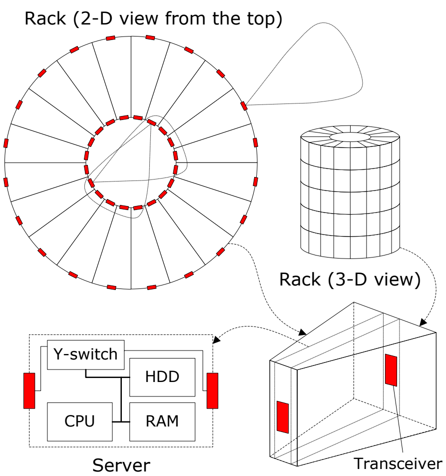


Figure 2: Rack and server design.

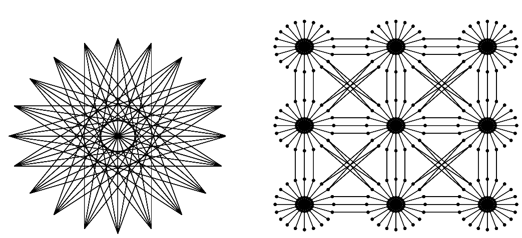
network topology that would optimize key performance characteristics: its latency and total bandwidth.

To separate the wireless signals for communications within a rack and among different racks, we propose cylindrical racks which store servers in prism-shaped containers as illustrated in Figure 2. This choice is appealing, because it partitions the data center volume into two regions: intra-rack and inter-rack free space. This way, 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 the 2-dimensional viewpoint in Figure 2. A rack consists of S stories and each story holds C containers; we constrain $S = 5$ and $C = 20$ for brevity of analysis.

The prism containers have roughly the size to contain commodity half-height blade servers. The transceivers located on opposite sides of the server are connected to a custom built Y-switch. The Y-switch multiplexes an incoming packet to one of the outputs. More details about the Y-switch is discussed at the end of this section.

3.2 Topology

A network architect can position racks and server prisms within a data center, and orient their transceivers to optimize latency and bandwidth. In this paper, we focus on a topology that can be modeled as a mesh of Cayley graphs.



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

Figure 3: Cayley data center topology

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 [45].

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 great flexibility in how a data center can be constructed out of these racks, but we pick the simplest possible 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: small black dots represent the transceivers and the lines indicate the connectivity. A Cayley graph sits in the center of each rack: lines coming out of the Cayley graphs are connections through the Y-switches. 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 other servers in different stories in the same or different racks.

3.3 Routing Protocol

A routing protocol compliant to 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 Cayley data centers use is called diagonal XYZ routing inspired by XY routing [21].

Similar to XY routing, 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 servers, the rest of the servers can identify their own identities by querying the neighbors and propagating the relevant information.

The routing protocol compares the identities of the server holding the packet and the final destination of the packet to find the next destination. Based on rx and ry values, the routing protocol finds an adjacent rack of the server that is closest to the destination. Values s are used to reach the desired story height of the destination that the packet should arrive. Routing protocol uses i values to forward the packet using the shortest path to the destination within the same story. Algorithm 1 describes the details about the routing algorithm and Figure 4 illustrates an example of using this algorithm observed from two different angles.

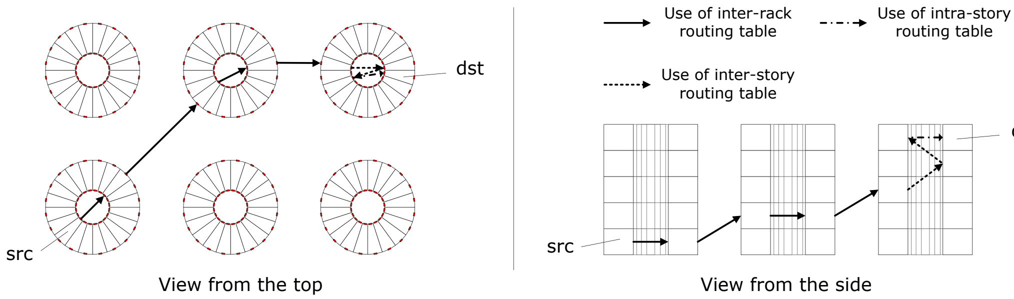


Figure 4: Diagonal XYZ routing. A packet is delivered using the racks with the shortest distance to the destination. Then the story of the packet is adjusted and finally, the packet is delivered to the destination.

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)

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 tale lookup all lead to the same number of hops to the final destination. Thus, the routing protocol pseudorandomly selects one of the choices to evenly distribute the traffic and to allow a TCP flow to follow the same path. We use a pseudorandom hashing of the packet header like the Toepliz Hash function [26] used by Microsoft's RSS specification [36] which ensures the same flow is routed over the same path, and also that traffic is evenly distributed.

The directionality of the radio beam, the presence of multiple transceivers per node and the low latency of the Y-switch makes it possible to deploy an optimization in Cayley data centers that is not possible in traditional wireless communication based on omnidirectional antennas. In particular, cut-through switching [28] which starts routing a packet immediately after receiving and reading the packet header, is used in Cayley data centers along with the proposed routing scheme.

The diagonal XYZ routing algorithm can be implemented in layer 3 or 2, depending on the deployment scenario. Since cut-through switching interacts with the TTL decrement, a layer 3 implementation may either choose to avoid this optimization or not decrement the IP TTL while packets are in flight within the data center.

3.3.2 Adaptive Routing in Case of Failure

Compared to a conventional data center, a Cayley data center has a distinct failure profile. While a group of servers within a rack share much of the switching gear (switches and wires) in the former—thus causing dependencies on the failure probability model, in the Cayley data center, these probabilities are both lower (fewer components can fail) and less dependent (servers within a rack share only the same power supply and similar operating temperature) [18]. When they happen, failures of nodes in a Cayley data center can lead to a failure on a network path. In this subsection we show that our proposal is capable of dealing with substantial server failures. Employing a simple timeout scheme in the MAC layer protocol enables easy detection of such failures.

Isolated node failures typically do not lead to disconnection in Cayley data centers, because the servers around the failed server are likely to provide comparable routing functionality as the failed node. Similarly, racks that may be missing individual servers appear like a failed server, so other servers around a "hole" can be used. Therefore, we focus our discussion on two cases with massive, correlated failures: failure of servers in one or more stories of a rack and failure of one or more racks.

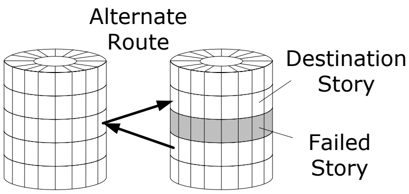


Figure 5: Routing when stories fail.

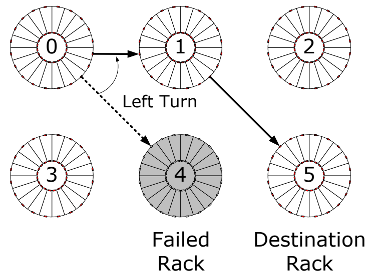


Figure 6: Routing when racks fail.

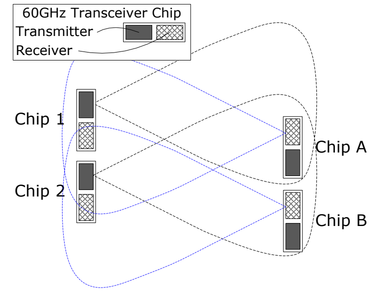


Figure 7: Hidden terminal problem.

Failure of a story can affect paths among stories within the same rack. At the worst case, all entries in inter-story routing table can become useless to deliver a packet to the desired story. In this case, we can use adjacent racks to deliver the packet to different stories as shown in Figure 5.

Rack failures are potentially more catastrophic as they can affect larger number of packets in the inter-rack routing level. We adopt a geographic routing technique based on face routing [25] to deal with these cases. Once the MAC layer detects that a failed rack is blocking the path for a packet, our routing protocol sets up a temporary destination for the packet using the left turn (a.k.a right hand) rule. If the next rack to visit has failed (e.g. the southeast rack in Figure 6), our routing protocol will temporarily route the packet to the the adjacent rack to the left (e.g. the east in the example of Figure 6). Once the packet arrives at its temporary destination, the protocol routes the packet to its original destination.

However, using only the left turn rule can lead to infinite loops. Assume a packet's source is rack 2 and the destination is rack 3 in Figure 6. Routing using only left turns will endlessly route the packet between racks 2 and 5. To prevent this, the failure routing method switches to the right turn (a.k.a left hand) rule, when the following conditions are satisfied: the packet is at the rack on the edge of the grid of racks (e.g. rack 5 in Figure 6) and there is no rack on the very left of the failed rack (e.g. from the viewpoint of rack 5, there is no rack on the very left of rack 4 in Figure 6). In this example, the full path of racks that the packet goes through becomes racks 2 (left turn applies), 5 (right turn applies), 1, and 3, in that order.

Even if we use both left and right turn rules, certain rare failures, such as racks failing in a "⊔" pattern, can lead to live locks. These cases require several racks, which consist of several hundreds of servers, to fail at the same time. However, field data indicates that failures usually involve less than 20 components in data centers with 100K nodes, so such failures are very unlikely [18]. Modification of our failure routing can overcome such failures, but given the rarity of the scenario, we do not discuss them here.

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. 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 7. Assume that a transceiver chip 1 is communicating with chip A and chip 2 is trying to send signals to chip B. Employing a regular ready-to-send (RTS) /clear-to-send (CTS) based MAC protocol [29] 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) [22, 23] channel arbitration/reservation scheme. DBTMA is based on 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 bandwidth for this tone using FDD. We also use a dedicated channel for control messages that does 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 Cayley data center. High-level schematic of this switch is shown in Figure 8. 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 subsection 3.3.1. To analyze the feasibility of the proposed Y-switch design, we implemented the Y-switch design for Xilinx FPGA in Simulink [35] and verified that for an FPGA running at 270MHz, its switching delay is less than 4ns.

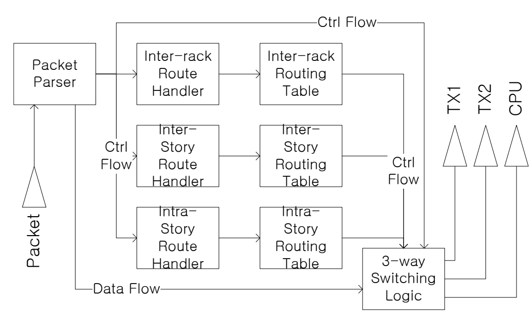


Figure 8: Y-switch schematic.

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

4. FEASIBILITY 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 conventional wired data centers (CDC).

4.1 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 use an analysis tool measuring the number of coexisting flows to explore the design space of a Cayley data center. 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 compared to a CDC?

  By measuring the packet delivery latency using a fine-grain packet level simulation model with different benchmarks, we compare the performance between the two.

- **Failure resilience**: how well can a Cayley data center handle failures?

  Cayley data centers do not have conventional switches, so the characteristics of network failure are different from wired data centers. 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. We measure how many node failures a Cayley data center can tolerate until nodes start to disconnect from the network. We also measure the performance of a Caley data center under various degrees of server and rack failure.

- **Cost**: How cost effective is a Cayley data center compared to a CDC?

  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 the price of network interface cards available in the market. We compare the price between a Cayley data center and a CDC based on the expected price range of 60 GHz transceivers.

4.2 Test Environments

Because data centers involve tens of thousands of servers and because 60 GHz transceivers are not yet massively produced, it is impossible to build a Cayley data center model 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 at a nanosecond scale, packet hops, number of packet collisions, number of packet drops from buffer overflow or timeout and so on. It can correctly build 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 CDC given the number of ports and oversubscription rate of switches in each hierarchy.

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

As a supplementary method to quickly evaluate and explore the design space of Cayley data centers, we also built a flow counter tool that measures the number of network flows that can coexist in a data center at any moment in time. This tool selects random pairs of nodes in a given topology and tries to connect the nodes by reserving the path. If any part of a new path is blocked by another connection, the new connection is aborted. The total number of connections achieved from this tool can be interpreted as bandwidth per server by dividing the combined link capacity with the total number of servers. This tool does not consider the MAC layer protocol and switching delays; instead, we use it to more exhaustively predict the trend of simulation results.

4.3 Base Configurations

Throughout this section, we evaluate both Cayley data centers and CDCs with 10K server nodes. Racks are positioned in a 10 by 10 grid for Cayley data centers and three levels of switches, top of rack switches (TOR), aggregation switches (AS), and core switches (CS), are used in CDC 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 data center are often larger than 10 and can increase to over several hundred [18, 6]. To be conservative, we configure CDCs to have oversubcription rates between 3 and 10.

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, 9, 12].

We focus exclusively on traffic within the data center, which account for more than 80% of the traffic even in client-facing web

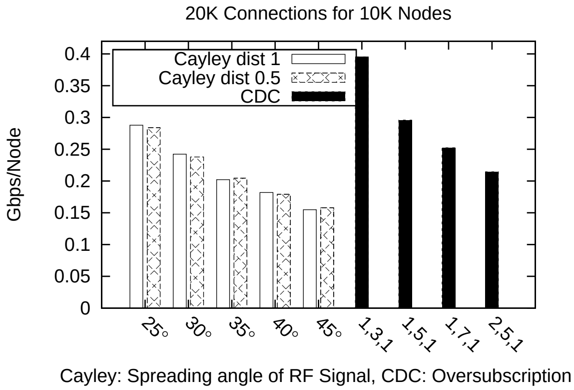


Figure 9: Expected bandwidth per server in data centers.

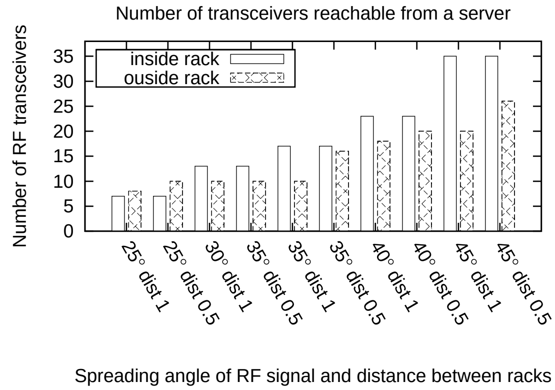


Figure 10: Number of connectable transceivers per server.

clouds [18]. Traffic in and out of the Cayley data center can be accomodated without hot spots through transceivers on the walls and ceiling as well as wired injection points.

4.4 Exploration of Wireless Data Center Design Space

In this subsection, we explore the design space of Cayley data centers by varying design parameters. Using the flow counter tool, we attempt to establish two times as many random connections as the number of servers in the data center to quantify the relationship between input parameter and achievable bandwidth per server. We run 20 tests on each configuration of the Cayley data center and the CDC and average the results. The average of the standard deviation is 0.007. The angles of the wireless signal ranged from 25° to 45° and the distance between the racks are configured to be either 0.5 meters or 1 meter. 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.

The results in Figure 9 show that Cayley data centers with signal angle 25° to 35° perform as well as to CDC (1,5,1), (1,7,1), and (2,5,1). Since real data centers have higher oversubcription rates [6], this result indicates that wireless data centers can perform comparable to existing data centers.

As the signal's angle becomes larger in Cayley data centers, the bandwidth per server decreases. This is because the number of the transceivers contending to reserve a wireless channel increases as the angle increases, adding MAC layer contention.

The distance between racks also affect the number of connectable transceivers from a transceiver in the inter-rack space. The maximum number of RF chips directly reachable from a server through the inside space of a rack and the outside space of the rack is shown in Figure 10 for different signal angles and distances between the racks. Larger numbers indicate that there is more MAC layer contention. The distance between racks does not greatly affect the number of connectable transceivers. Our rack design ensures that varying the distance will not significantly change the bandwidth (Figure 9). Because the intra-rack space is shared by the packets arriving from all 8 directions, there is relatively more contention to reserve the wireless channel inside the rack than the outside of the rack. Thus, this implies that contention inside the rack has greater influence on the overall performance.

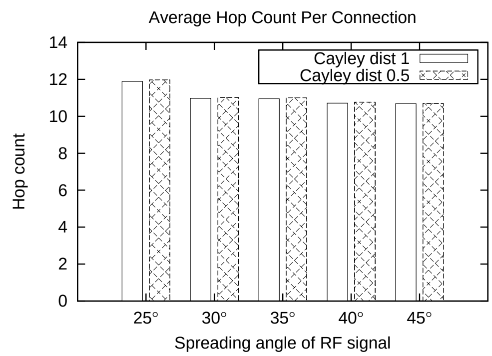


Figure 11: Average Network hop count for Cayley data center.

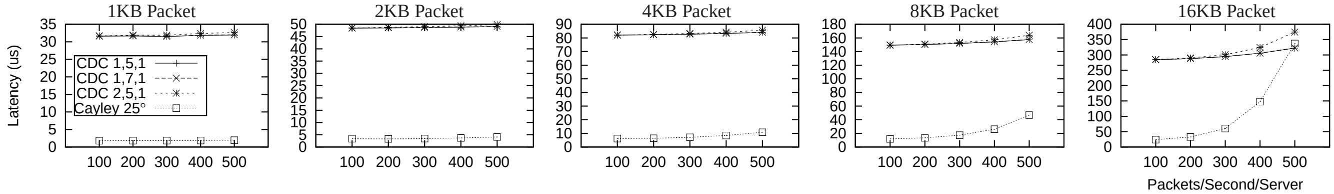
We next examine the average hop count per connection from source to destination. Figure 11 shows that the number of hops in a path does not vary much depending on the signal's angle or the distance between the racks. This is because our geographical routing protocol takes place in the inter-rack level.

In summary, the RF signal angle does not significantly affect the hop count in a Cayley data center. Figures 11 and 10 together show that the MAC layer contention, which is dependent on the signal's angle, greatly affects the wireless communication performance.

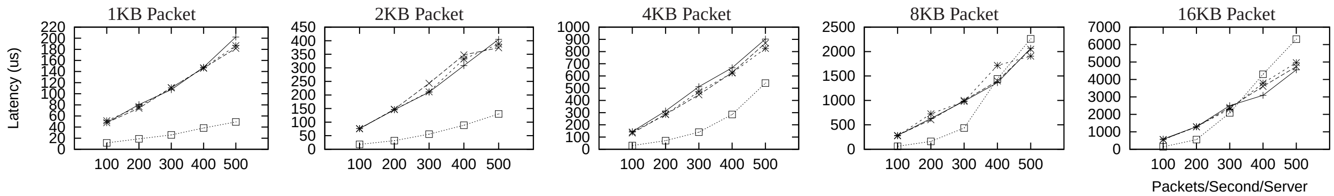
4.5 Performance: Packet Delivery Latency

We measure the average and maximum packet delivery latency of Cayley data centers and CDCs using a detailed packet level simulator. The evaluation involves three synthetic benchmarks varying the packet injection rates and packet sizes:

- **Staggered (Intra-Pod)**: 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.



(a) Average Packet Delivery Latency



(b) Maximum Packet Delivery Latency

Figure 12: Simulation Results for Staggered Traffic.

- **Stride**: source node with a global ID x sends packets to the destination node with ID $x+$ (total # of servers)$/2$.

We use a spreading 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 bandwidth analysis of previous section and 1 meter distance between the racks is more ergonomic than 0.5 meters for people to walk through data centers.

The results for each benchmark, staggered, uniform random, and stride, are shown in Figures 12, 13, and 14, respectively. The first row of each figure shows the average packet delivery latencies for packet size from 1KB to 16KB, and the second row shows the maximum latencies. Packet injection rate per server is tested from 100 to 500 for all packet sizes.

The distances traveled by packets become larger in the order of staggered, uniform random, and stride due to each benchmark's inherent characteristics. Staggered is the most favorable traffic and stride is the least favorable traffic to both Cayley data centers and CDCs from the performance point of view.

Under the staggered workload, packets in a CDC do not travel above AS, so different oversubscription rates in AS do not create a large difference in the performance (Figure 12). Only CDC (2,5,1), where TOR is oversubscribed, performs worse than other CDCs when traffic load becomes large. The average packet hop count in CDCs is 3.89 and that in Cayley data centers is 5.03. The average packet delivery latency of the Cayley data center is an order of magnitude smaller than that of CDC for most cases, due to high bandwidth wireless link and small delay in Y-switch. It performs up to 17 times better than CDC. However, because of MAC layer contention, increase of traffic load quickly degrades the performance of the Cayley data center as shown in the plot for average latency using 16KB packets. The performance decreases slightly when compared to the CDC (1,5,1) and (1,7,1). 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 [31]. Therefore, Cayley data centers are expected to perform significantly better on average than CDCs, except under peak load for applications with similar traffic patterns to staggered.

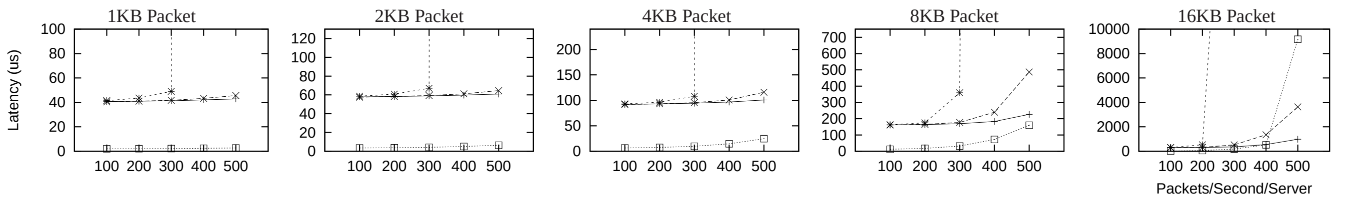
The maximum latency in a Cayley data center is also smaller than a CDC for the staggered benchmark. Although the latency quickly increases and exceeds that of CDC, Cayley data centers perform better overall and the maximum latency is at most 27.2% larger than CDC only when the traffic load reaches its peak.

The average packet hop count for uniform random traffic in CDCs and Cayley data centers are 5.9 and 11.5 respectively. Still, Cayley data centers perform better for most cases: it demonstrates maximum 20 times better performance than CDC (Figure 13). CDC suffers from relatively large switching delay and its 1 GigE link becomes the bottleneck. The performance gap between CDCs and Cayley data centers narrows quicker than with staggered traffic. This is because the traffic in Cayley data centers go through greater number of network hops. Similarly, when the traffic load hits the peak, the performance in the Cayley data center decreases compared to the CDC (1,5,1) and (1,7,1): the maximum latency in the Cayley data center is 9.3 and 2.5 times greater, respectively. CDC (2,5,1) performs the worst for all cases. This is because packets start to drop due to buffer overflow in oversubscribed switches.

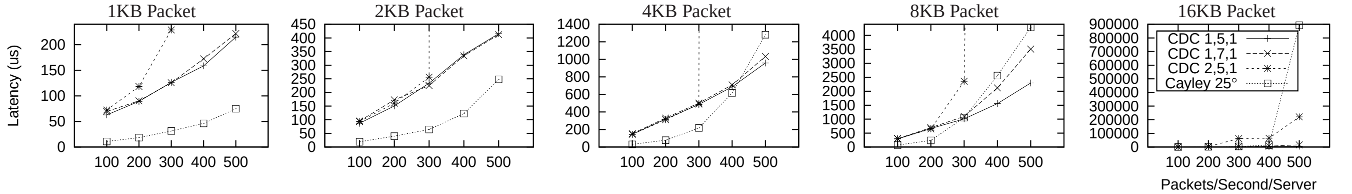
The maximum packet delivery latency shows the potential challenge in a Cayley data center. Although the average latencies are better than CDCs, Cayley data centers show relatively worse maximum latency as traffic load increases. However, except for under the peak traffic load, the maximum latency of Cayley data centers is less than 3.04 times as large as the latency of CDC (1,5,1), and is smaller than CDC for most cases.

The average hop count in CDCs and Cayley data centers are 6 and 12.4 for the stride workload. When traffic load is small, Cayley data centers perform maximum 23 times better than CDCs (Figure 14). Cayley data centers show better average performance when using 1, 2, and 4KB packets and shows slightly better performance than CDC (1,7,1) when using 8KB packets. As the traffic reaches the peak when using 16KB packets, the latency in the Cayley data center skyrockets to 179ms. Competing for a data channel at each hop with relatively large packets significantly degrades the performance of Cayley data centers. The maximum latency shows similar trend, but the Cayley data center's latency reaches CDCs latency more quickly as the load increases.

In summary, except for handling the peak traffic for uniform random and stride benchmark, Cayley data centers perform better

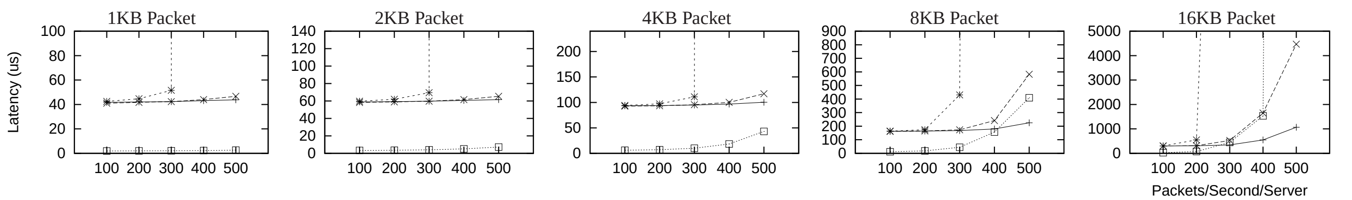


(a) Average Packet Delivery Latency

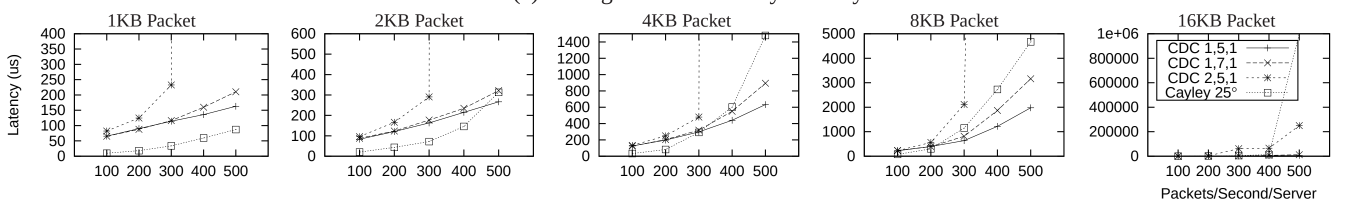


(b) Maximum Packet Delivery Latency

Figure 13: Simulation Results for Uniform Random Traffic.



(a) Average Packet Delivery Latency



(b) Maximum Packet Delivery Latency

Figure 14: Simulation Results for Stride Traffic.

than or comparable to CDCs. As the average number of hops per packet increases, the performance of Cayley data centers quickly decreases. This shows that Cayley data center may not 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, applications, such as MapReduce, resemble staggered benchmark which do not saturate oversubscribed (aggregate) switches [31, 6]. 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 CDCs.

4.6 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 15). 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 all cases is 6.5%.

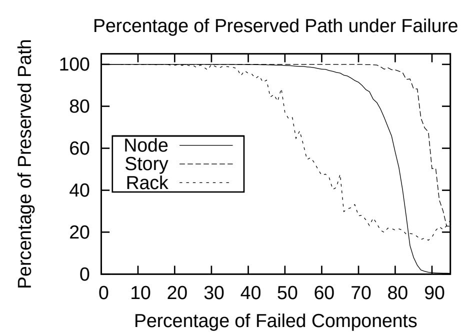


Figure 15: Percentage of preserved path under failure.

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 50% 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 and CDCs, a Cayley data center can be more resilient to network component failures. This is mainly because wireless data centers do not have conventional switches which can be critical points of failure.

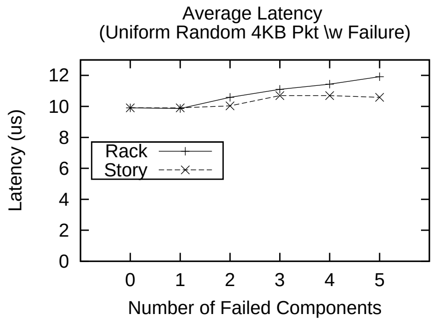


Figure 16: Average latency in case of failure.

4.7 Packet Delivery Latency under Failure

While the previous subsection showed the connectivity of a Cayley data center, we explore the packet delivery latency in the presence of failed components in this subsection. Figure 16 shows the performance variation as the number of randomly selected racks and stories with failure increases from 0 to 5. We exclude cases where failures were uncorrelated. In our 10K node setting, failing one rack means failing 100 servers, or 1 percent of total servers. Typically only less than 20 devices fail at the same time and they are mostly fixed within a day in data centers with 100K servers [18]. With commodity servers, the rate of simultaneously failing nodes increases, but it is still within 5 percent range [15].

The average number of hops increases by a maximum of 0.04 for all cases compared to the base case without failure. The hop count shows that alternate routes, involving almost the same number of hops as the original route, can be found for most cases using our routing method. The performance degrades by maximum 20.3 percent and 7.9 percent compared to the base case as the number of failures for racks and stories increases, respectively. The degradation of performance mainly originates from detecting the failed nodes using timeout. However, once the failure is detected, packets can be directly forwarded to alternate routes. The latencies involving the failures are still an order of magnitude smaller than the latencies of CDC without failure.

4.8 Cost Comparison

It is complicated to compare two technologies when one is commercially mature and the other exists only as a concept. We can easily measure the cost of 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 between two data centers for different values of 60 GHz transceiver cost.

We compare the cost of wireless and conventional 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 [38]—and the Y-switch. In our system we replace this card with the proposed simple Y-switch and at least two transceivers. Y-switches consist of simple core logic, host interface, such as 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 to connect 10K servers using switches are summarized in Tables 4 and Table 5. The total

Component	Price ($)	Min Unit	Min unit price ($)
TOR	8000	1	8,000
AS	9,000	1	9,000
CS			342,500
CS subunit	40,000	1	40,000
CS chassis	12,000	1	12,000
CS power supply	3500	3	10,500

Table 4: CDC switches [40]

Config	#TOR	#AS	#CS sub-unit	#CS chassis	Cost ($)
2,5,1	250	26	8	1	2,576,500
1,7,1	250	48	12	2	2,957,000
1,5,1	250	52	16	2	3,153,000

Table 5: CDC networking equipment cost for 10K nodes

price ranges from US$2.5M to US$3.1M.

On the other hand, 60 GHz 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 $Cost_{Cayley}(cost_t, N_{server})$, where $cost_t$ is the price for a transceiver and N_{server} is the number of servers in a data center. Servers in Cayley data centers have a pair of transceivers, so the function,

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

From this function, we can easily find out that as long as $cost_t$ is less than US$125, Cayley data centers can connect 10K servers with lower price than CDC. Similarly, if $cost_t$ becomes US$10, the cost of transceivers in Cayley data centers can be less than 1/12 of CDC switches. Considering the rapidly dropping price of silicon chips—CPU price drops by 40% within 9 month [19]— we expect the transceiver's price to quickly drop to less than US$100 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.

Considering the power consumption for the connections, the power consumption of a 60 GHz transceiver is less than 0.3 watts [39]. 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, 9, 11]. In total, the switches typically consumes 58 kilowatts to 72 kilowatts depending on the oversubscription rate for data center with 10K servers. This is 10 to 12 times as large as the maximum power consumption in the transceivers.

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 60 GHz could revolutionize the simplicity of integrating and maintaining data centers. Ultimately, we expect that there exists a high-bar on the performance-price curve, where CDC becomes advantageous.

4.9 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 CDCs are:

- **Performance**: Cayley data centers can perform better than or comparable to CDCs depending on the applications: traffic pattern with small number of packet hops with moderate traffic load provide the best condition for Cayley data centers to perform better than CDCs. Also Cayley data centers can maintain good performance under failure.
- **Failure resilience characteristic**: densely connected wireless data centers inherently have greater resilience than CDC. This is because wireless data centers do not have switches, which can be cause for correlated loss of connectivity. Further, densely connected networks enable communication even under large number of node failures. Cayley data centers can handle up to 59% of node failure before a path between two live nodes become disconnected.
- **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. Power consumption and expected maintenance cost is much lower than CDC.

The characteristics and limitations of Cayley data centers are:

- **MAC layer contention**: sharing of wireless channel followed by MAC layer contention greatly influence the overall 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 our 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 altogether summarize the challenges and open problems for designing a wireless data center and present strong benefits and feasibility of adopting the wireless technology into data centers.

5. RELATED WORK

Ramachandran et al. [42] outlined the benefits and challenges for removing wires and introducing 60 GHz communication within a data center. We share many their insight in our work and also conclude that 60 GHz wireless networks can improve conventional data centers. Further, we address some of the problems identified by the authors. In this paper, we propose a novel rack-level architecture, use realistic parameters for 60 GHz transceivers, and provide an extensive evaluation of the performance of the proposed wireless data centers.

Flyways is a wireless network based on 60 GHz or 802.11n organized on top of wired data center racks [30]. It provides a supplementary network for relieving congested wired links. In contrast, wireless links are the main communication channels in Cayley data centers and there are no wired connections.

A scalable commodity data center network architectures by Al-Fares et al. and Portland [2, 34] employ commodity switches to replace expensive switches in data centers and to provide a scalable and oversubscription-free network architecture. It achieves high performance with small 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 [22, 23] 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 60 GHz technology and directional antennas [8, 41, 43, 32], but they require global arbitrators or multiple directional antennas collectively pointing to all directions. These are not suitable for data centers. Designing a full MAC layer protocol for wireless data centers is an open problem.

While our design adopted XY routing for Cayley data centers, other variations of oblivious routing protocols, such as [16, 4, 21], and adaptive routing protocols, such as [27, 20], can be adapted to our design.

6. CONCLUSION

In this paper, we propose a radically novel methodology for building data centers which displaces the existing massive switching fabric including all switches, routers, and wired connections, with wireless transceivers integrated within server nodes. A routing protocol directs packets within a Y-switch inside each server; each Y-switch connects all server's transceivers as well as its system bus. We identify that the resulting dense connectivity subgraph belongs to a class of Cayley graphs and outline its well-understood properties.

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 having total bandwidth on par with wired data centers, Cayley data centers substantially outperform conventional data centers 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.

Directions for future work are numerous. 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 congestions 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 subband shifts higher in frequency? Would some degree of wired connectivity among servers internal to a single rack benefit performance? As the mm-wave technology matures, we expect novel wireless networking architectures to be realized in data centers and many of the issues mentioned here to be resolved.

7. ACKNOWLEDGMENT

While pursuing this work, we received a lot of help from a variety of people. First of all, we thank the Georgia Tech team for providing us with the specifications and the communication model of the 60 GHz wireless transceivers. We appreciate Han Wang for helping us implement the Y-switch for FPGAs. We would also like to thank Microsoft's staffs including Dave Harper and Dave Maltz who are involved in data center management and research, Deniz Altinbuken, and Tudor Marian for sharing their thoughts about our work.

8. 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. In *Communications of the ACM*, 53(4):50-58, April 2010.

[2] M. Al-Fares, A. Loukissas, and A. Vahdat. A scalable, commodity data center network architecture. In *ACM SIGCOMM Conference on Data Communication*, pages 63-74, Seattle, WA, USA, August 2008.

[3] H. Abu-Libdeh, P. Costa, A. Rowstron, G. O'Shea, and A. Donnelly. Symbiotic routing in future data centers. In *ACM SIGCOMM Conference on Data Communication*, pages 51-62, New Delhi, India, August 2010.

[4] H.G. Badr and S. Podar. An optimal shortest-path routing policy for network computers with regular mesh-connected topologies. In *IEEE Transactions on Computers*, 38(10):1362-1371, October 1989.

[5] R. Buyya, C. S. Yeo, and S. Venugopal. Market-oriented cloud computing: vision, hype, and reality for delivering IT services as computing utilities. In *IEEE International Conference on High Performance Computing and Communications*, pages 5-13, Dalian, China, September 2008.

[6] T. Benson, A. Akella, and D. A. Maltz. Network traffic characteristics of data centers in the wild. In *Conference on Internet Measurement*, pages 267-280, Melbourne, Australia, November 2010.

[7] A. Cayley. On the theory of groups. In *American Journal of Mathematics*, 11(2):139-157, 1889.

[8] X. Chen, J. Lu, and Z. Zhou. An enhanced high-rate WPAN MAC for mesh networks with dynamic bandwidth management. In *Conference on Global Telecommunications*, pages 3408-3412, St. Louis, MO, USA, November 2005.

[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, March 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, December 2008.

[15] D. Ford, F. Labelle, F. I. Popovici, M. Stokely, V.-A. Truong, L. Barroso, C. Grimes, and S. Quinlan. Availability in globally distributed storage systems. In *USENIX Conference on Operating Systems Design and Implementation*, pages 1-7, Vancouver, BC, Canada, October 2010.

[16] W.-c. Feng and K. G. Shin. Impact of selection functions on routing algorithm performance in multicomputer networks. In *International Conference on Supercomputing*, pages 132-139, Vienna, Austria, July 1997.

[17] A. Greenberg, J. Hamilton, D. A. Maltz, and P. Patel. The cost of a cloud: research problems in data center networks. In *SIGCOMM Compututer Communication Review*, 39(1):68-73, December 2008.

[18] 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. In *ACM SIGCOMM Conference on Data Communication*, pages 51-62, Barcelona, Spain, August 2009.

[19] C. Gianpaolo, D. M. Xavier, S. Regine, V. W. L. N., and W. Linda. Inventory-driven costs. In *Harvard Business Review*, 83(3):135-141, October 2005.

[20] P. Gratz, B. Grot, and S.W. Keckler. Regional congestion awareness for load balance in networks-on-chip. In *International Symposium on High Performance Computer Architecture*, pages 203-214, Salt Lake City, UT, USA, February 2008.

[21] C. J. Glass, L. M. Ni, and L. M. Ni. The turn model for adaptive routing. In *International Symposium on Computer Architecture*, pages 278-287, Gold Coast, Queensland, Australia, May 1992.

[22] Z.J. Hass and J. Deng. Dual busy tone multiple access (DBTMA)-a multiple access control scheme for ad hoc networks. In *IEEE Transactions on Communications*, 50(6):975-985, June 2002.

[23] Z. Huang, C.-C. Shen, C. Srisathapornphat, and C. Jaikaeo. A busy-tone based directional MAC protocol for ad hoc networks. In *International Conference for Military Communications*, volume 2, pages 1233-1238, 2002.

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

[25] E. Kranakis, H. Singh, and J. Urrutia. Compass routing on geometric networks. In *Canadian Conference on Computational Geometry*, pages 51-54, Vancouver, British Columbia, Canada, August 1999.

[26] H. Krawczyk. LFSR-based Hashing and Authentication. In *International Cryptology Conference on Advances in Cryptology*, pages 129-139, London, UK, August 1994.

[27] J. Kim, D. Park, T. Theocharides, N. Vijaykrishnan, and C. R. Das. A low latency router supporting adaptivity for on-chip interconnects. In *Design Automation Conference*, pages 559-564, Anaheim, California, USA, July 2005.

[28] P. Kermani and L. Kleinrock. Virtual cut-through: a new computer communication switching technique. In *Computer Networks*, 3:267-286, 1979.

[29] P. Karn. MACA - a new channel access method for packet radio. In *ARRL Computer Networking Conference*, pages 134-140, London, Ontario Canada, September 1990.

[30] S. Kandula, J. Padhye, and P. Bahl. Flyways to de-congest data center networks. In *Workshop on Hot Topics in Networks*, New York City, NY, USA, October 2009.

[31] S. Kandula, S. Sengupta, A. Greenberg, P. Patel, and R. Chaiken. The nature of data center traffic: measurements & analysis. In *Conference on Internet Measurement*, pages 202-208, Chicago, IL, USA, November 2009.

[32] T. Korakis, G. Jakllari, and L. Tassiulas. A MAC protocol for full exploitation of directional antennas in ad-hoc wireless networks. In *International Symposium on Mobile Ad Hoc Networking & Computing*, pages 98-107, Annapolis, Maryland, USA, June 2003.

[33] V. Kvicera and M. Grabner. Rain attenuation at 58 GHz: prediction versus long-term trial results. In *EURASIP Journal on Wireless Communications and Networking*, 2007(1):46-46, January 2007.

[34] 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. In *ACM SIGCOMM Conference on Data Communication*, pages 39-50, Barcelona, Spain, August 2009.

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

[36] Microsoft. Scalable networking: eliminating the receive processing bottleneck—introducing RSS. WinHEC, April 2004.

[37] 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. In *EURASIP Journal on Wireless Communications and Networking*, 2007(1):51-51, January 2007.

[38] 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, January 2011.

[39] 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. In *IEEE Journal on Selected Areas in Communications*, 27(8):1347-1357, October 2009.

[40] PEPPM. Cisco > Current Price List. http://www.peppm.org/Products/cisco/price.pdf, January 2011.

[41] J. Qiao, L.X. Cai, and X. Shen. Multi-hop concurrent transmission in millimeter wave WPANs with directional antenna. In *IEEE International Conference on Communications*, pages 1-5, Cape Town, South Africa, May 2010.

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

[43] S. Singh, F. Ziliotto, U. Madhow, E.M. Belding, and M.J.W. Rodwell. Millimeter wave WPAN: cross-layer modeling and multi-hop architecture. In *IEEE International Conference on Computer Communications*, pages 2336-2340, Anchorage, AK, USA, May 2007.

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

[45] K.W. Tang and R. Kamoua. Cayley pseudo-random (CPR) protocol: a novel MAC protocol for dense wireless sensor networks. In *IEEE Wireless Communications and Networking Conference*, pages 361-366, Kowloon, Hong Kong, March 2007.

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

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

[48] S. K. Yong and C.-C. Chong. An overview of multigigabit wireless through millimeter wave technology: potentials and technical challenges. In *EURASIP Journal on Wireless Communications and Networking*, 2007(1):50-50, January 2007.