A Power Efficient Server-to-Server Wireless Data Center Network Architecture Using 60 GHz Links

Sree Gowrishankar Umamaheswaran
su7970@rit.edu

Follow this and additional works at: https://scholarworks.rit.edu/theses

Recommended Citation
A Power Efficient Server-to-Server Wireless Data Center Network

Architecture Using 60 GHz Links

by

Sree Gowrishankar Umamaheswaran

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of
Master of Science in Computer Engineering

Supervised by

Dr. Amlan Ganguly
Department of Computer Engineering
Kate Gleason College of Engineering
Rochester Institute of Technology
Rochester, NY
November 2017

Approved By:

____________________
Dr. Amlan Ganguly
Primary Advisor – R.I.T. Dept. of Computer Engineering

____________________
Dr. Andres Kwasinski
Secondary Advisor – R.I.T. Dept. of Computer Engineering

____________________
Dr. Mineosk Kwon
Secondary Advisor – R.I.T. Dept. of Computer Science
Acknowledgements

I would like to acknowledge my thesis advisor, Dr. Amlan Ganguly for guiding me throughout my thesis. I would like to extend my gratitude to Dr. Andres kwasinski and Dr. Mineosk kwon for serving on my thesis committee and providing their support.
Abstract

Data Centers have become the digital backbone of the modern society with the advent of cloud computing, social networking, Big data analytics etc. They play a vital role in processing the large amount of information generated. The number of Data Centers and the servers present in them has been on the rise over the last decade which has eventually led to the increase in the power consumption of the Data Center due to the power-hungry interconnect fabric which consists of switches and routers. The Significant portion of the power consumption is due to the power-hungry switching fabric necessary for communication in the datacenter. Additionally, the complex cabling in traditional datacenters pose design and maintenance challenges and increase the energy cost of the cooling infrastructure by obstructing the flow of chilled air. In this work, we address these problems of traditional datacenters by designing a unique new server-to-server wireless datacenter network (DCN) architecture.

The proposed design methodology uses 60GHz unlicensed millimeter-wave bands to establish direct communication links between servers in a DCN without the need for a conventional fabric. This will reduce the power consumption of the DCN significantly and increase the in-dependency of servers from switching fabric.

In our work, the previous traffic models of a data center network are studied and a new traffic model very similar to the actual traffic in data center is modeled and used for simulating the DCN environment. We first demonstrate that such a power-efficient wireless DCN can sustain the traffic requirements encountered and provide
data rates that are comparable to traditional DCNs. Having established the feasibility of a server-to-server wireless DCN in terms of performance, we estimate that its power consumption is lowered by six to eighteen times in comparison to a conventional DCN fabric. We have also compared the efficiency, performance, scalability of our DCN architecture with some of the other practically used architectures like Dcell, Bcube with the same traffic.
Contents

Acknowledgements ii

Abstract iii

1 Introduction 1
   1.1 Thesis Contributions 3
   1.2 Thesis Layout 4

2 Background & Related work 5
   2.1 Motivation 5
      2.1.1 Power Consumption 6
      2.1.2 Cabling Complexity 6
      2.1.3 Traffic Congestion 8
   2.2 Related Work 9
      2.2.1 Data Center Networks 9
         Fat-Tree 10
         Dcell 10
         Bcube 11
      2.2.2 Wireless Data Centers 11
      2.2.3 The 60 GHz band 12
      2.2.4 60 GHz Standards 13
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.4 Simulation Platform</td>
<td>42</td>
</tr>
<tr>
<td>4.5 Performance Analysis</td>
<td>43</td>
</tr>
<tr>
<td>4.5.1 Traffic Flow Completion Duration</td>
<td>44</td>
</tr>
<tr>
<td>4.5.2 Throughput</td>
<td>45</td>
</tr>
<tr>
<td>4.6 Conclusions</td>
<td>46</td>
</tr>
<tr>
<td>4.7 Future Work</td>
<td>47</td>
</tr>
</tbody>
</table>
List of Figures

2.1 Cabling Complexity ................................................. 7
2.2 60 GHz .............................................................. 13
3.1 Transmission Rate in small data center ......................... 20
3.2 Transmission Rate in large data center ......................... 21
3.3 Flow Duration in small data center ............................... 22
3.4 Flow Duration in large data center ............................... 23
3.5 Flow Size in small data center ................................... 24
3.6 Flow Size in large data center ................................... 24
4.1 Server picture ....................................................... 26
4.2 DC Layout ........................................................... 27
4.3 HP Communication .................................................. 28
4.4 VP Communication .................................................. 29
4.5 Algorithm Flowchart ............................................... 32
4.6 Intra Rack ............................................................ 34
4.7 HP-Inter Rack ......................................................... 35
4.8 VP-Inter Rack ........................................................ 36
4.9 Two hop-Inter Rack .................................................. 37
4.10 Power Consumption comparison ................................. 42
4.11 Flow completion duration ........................................ 44
List of Tables

4.1 Power Consumption of different DCN components ....... 39
4.2 Packet latency in different devices ...................... 46
4.3 Comparison between different data center architectures .... 47
List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DCN</td>
<td>Data Center Network</td>
</tr>
<tr>
<td>RF</td>
<td>Radio Frequency</td>
</tr>
<tr>
<td>ToR</td>
<td>Top of Rack</td>
</tr>
<tr>
<td>LoS</td>
<td>Line of Sight</td>
</tr>
<tr>
<td>MAC</td>
<td>Medium Access Control</td>
</tr>
<tr>
<td>UHF</td>
<td>Ultra-High Frequencies</td>
</tr>
<tr>
<td>OFDM</td>
<td>Orthogonal Frequency Division Multiplexing</td>
</tr>
<tr>
<td>SC</td>
<td>Single Channel</td>
</tr>
<tr>
<td>NS-3</td>
<td>Network Simulator 3</td>
</tr>
<tr>
<td>BER</td>
<td>Bit Error Rate</td>
</tr>
<tr>
<td>TCP</td>
<td>Transmission Control Protocol</td>
</tr>
<tr>
<td>MTU</td>
<td>Maximum Transmission Unit</td>
</tr>
</tbody>
</table>
Dedicated to my Loving Parents
Chapter 1

Introduction

Datacenters play a vital role in modern day as they provide processing power and storage required by billions of connected devices. One of the major challenges is, however, the enormous power consumption of datacenters. A census conducted by Natural Resources Defense Council (NRDC) shows that the energy consumption by datacenters reached 91 billion kWh energy in 2013, and will increase to 140 billion kWh by 2020 [1]. As the number of servers within a datacenter grows, the role of the interconnection infrastructure becomes paramount in terms of both energy or power efficiency and performance. Traditionally, datacenter networks (DCNs) are interconnected in tree-based topologies using wired links and multiple hierarchical levels of aggregation. These tree-based networks and other recent alternative technologies exhibit inherent limitations in scalability and oversubscription as they rely on copper and optical cable-based links [2].

The DCN consumes a significant portion of the total power consumption of a datacenter. The proportion of IT power dissipated by the network equipment varies between 10-50% depending upon server utilization [3]. The wired or cabled network technologies require power-hungry switches and create large bundles of cables, causing design overheads and maintenance challenges and obstructions to the
flow of chilled air for cooling [4]. Inefficient cooling only exacerbates the energy efficiency problems that plague the current datacenters. These problems are often worse in small and mid-size datacenters with hundred to a few thousands of individual servers in educational institutions and private enterprises as these are designed and deployed in an ad-hoc manner often leading to structural and functional heterogeneity making regular systematic design impossible. To address these common design issues faced by wired or cabled datacenters wireless datacenter architectures are being investigated as a promising alternative. The capability of the unlicensed 60GHz wireless band to deliver very high communication rates has led to the development and approval of the IEEE 802.11ad wireless local area network (WLAN) standard [5]. Therefore, recently proposed designs leverage newly developed technologies in the unlicensed 60GHz wireless band for wireless DCNs [6] [7].

Advancements in the 60GHz technologies enable the transceivers to consume low power, some even in the milliwatt range [8] [9], and establish multi-gigabit communication channels [10]. Directional horn antennas [11] as well as more recently developed phased arrays of antennas in the 60GHz bands [12] can provide high directional gains and beam steering capability between wireless transceivers. Using such antennas, the 60GHz channels can exhibit spatial re-usability, allowing multiple concurrent links reusing frequency bands to be formed within the same datacenter. The low power consumption combined with the ability to form concurrent multi-gigabit channels makes these transceivers ideal for use in power-efficient wireless DCNs.

In this article, we propose a server-to-server wireless DCN (S2S-WiDCN) architecture, based on the 60GHz wireless links between the individual servers of a small or
mid-size datacenter running typical query-response type of applications like map-reduce and index lookup. Through direct server-to-server wireless links using directional antenna arrays, we propose to eliminate the power-hungry switching fabric of traditional DCNs, resulting in significant power savings in data transfer. The communication between servers in the wireless DCN are achieved along horizontal lines and vertical planes as shown in Fig. 1. We compare the performance as well as power savings of the proposed server-to-server wireless DCN to that of conventional hierarchical DCNs. In this way, we demonstrate that the S2S-WiDCN is able to sustain and provide performance comparable to the conventional counterpart at significantly lower power consumption. To the best of our knowledge, this is the first attempt to measure the network-level characteristics of a server-to-server wireless DCN.

1.1 Thesis Contributions

In this work we propose a new wireless datacenter architecture. The performance of the proposed datacenter architecture is evaluated with that of the traditional wired datacenter architectures like Fat-tree, Dcell, Bcube. Conventional DCNs are arranged in a tree based topology, bringing about significant congestion and oversubscription along wired links as the number of aggregation levels increase. Several solutions have emerged in an attempt to limit the amount of resources used to construct large DCNs, including server centric wired approaches like Dcell, Bcube and wireless ToR-ToR architecture with the use of 60 GHz links.

Many contemporary works studying wireless data centers using 60 GHz links have proposed using these links to supplement existing wired networks in a hybrid style
approach. This proposed work attempts to replace a data center’s wired links entirely in order to maximize the energy savings obtained using wireless technology, while maintaining or improving the bandwidth provided by these links in comparison to a wired network. In addition to the evaluation of an entirely wireless data center, this work will improve upon existing 60 GHz medium access control (MAC) mechanisms to further the wireless data center approach. A primary objective of this proposed work is to demonstrate that a completely wireless data center has the capability to satisfy bandwidth demands of a typical data center while yielding significant energy savings. Furthermore, this work will provide a foundation for future wireless data center investigations using 60 GHz links.

1.2 Thesis Layout

In this research work, first the related work are discussed in detail followed by the explanation on how the traffic is modeled. The proposed wireless architecture is introduced with the routing protocol pertaining to it. Then the power consumption of the proposed is investigated and presented. Finally, the performance is evaluated with respect to throughput, completion duration and in future work’s section other research possibilities are explained in brief.
Chapter 2

Background & Related work

This section will begin with the motivation for this work and will address the central problems afflicting modern data center designs. Mainly, we will discuss about different wired data center architectures like Dcell, Bcube, Fat-tree. All background and related work will then be introduced and discussed. This section will conclude with the significance this work provides.

2.1 Motivation

Research concerning data center design and efficiency has become more prevalent in both industry and academia due to the increasing reliance on data centers to manage complex services. With the advent of cloud computing, data center resources are scaling rapidly and will require new and innovative solutions to meet that demand. Several factors however, limit the scalability and efficiency of modern data centers and have proved challenging to overcome. Some of the most prominent issues needing solution are the increasing power consumption, cabling complexity and traffic handling in data centers. These are discussed in detail to provide a clear picture on why development in data center research arena is required.
2.1.1 Power Consumption

The amount of power consumption by data centers is on the rise ever since the advent of cloud computing. Data centers consume an enormous amount of electricity for their operation. When considered as a whole, they make up a significant portion of overall worldwide electricity usage[13].

Furthermore, authors in [13] indicate that networking equipment already accounts for around 15% of a data center's infrastructure energy budget. This percentage is expected to increase in the future and is projected to grow by up to 50% in the coming years. Due to the fact that servers and cooling equipment represent the majority of a data center's electricity usage, they have received more research and attention [13]. While new technologies make cooling equipment and servers more energy efficient within a data center, the networking infrastructure must receive attention to prevent its energy consumption from growing unchecked. The energy efficiency problems associated with data centers are reinforced by a recent change in focus from the initial capital costs of setting up a new data center, to the cost in electricity of maintaining new data centers. The growth in data center infrastructure has driven down the cost of the initial equipment, however the cost of electricity recurs and adds up over time. The cost of powering a data center can quickly surpass the initial data center's total equipment cost.

2.1.2 Cabling Complexity

A major portion of power consumption in data center is used for cooling the data center facility. Due to the working of IT equipment heat is produced which gets trapped as there is very restricted flow of air. Cooling is mandatory because air
flow in data center is obstructed by complex cabling. Inefficient cooling, resulting from networking and cabling complexities, only exacerbates energy efficiency problems [14]. While structured cabling and raised floor techniques mitigate the cabling complexity challenge, these cables still result in airflow blockage which leads to inefficient cooling [15]. The large bundles of cables and the complexity of cabling can be seen in Figure 2.1. Even with structured cabling approaches, these wires hinder

![Figure 2.1: Cabling Complexity](image)

the ability of chilled air to move around the data center. These cabling overheads result in airflow blockages, additional maintenance costs, and can cost a great deal as a data center scales out. The physical cables within a data center represent a significant portion of the initial data center infrastructure costs. Thus improvements in the reduction of cabling complexity is one of the most challenging arena of research.
2.1.3 Traffic Congestion

This is the most predominant problem till date. As the servers are aggregated up the level for interconnection, traffic congestion becomes unavoidable. IT equipments such as switches and routers thus come in play. Not to mention, these switches and routers are the power hungry equipments in IT infrastructure.

It is estimated that roughly 70% of data center traffic will flow within a data center between servers, placing a large demand on the structure of a data center’s internal network [16]. As the number of servers scales up into the thousands, it becomes impractical and costly to provide full link bandwidth from all servers to every other server. This gives to the rise of importance in data center architecture. There have traditional topologies like Fat-tree, Dcell, Bcube with wired interconnections, switches and routers that either consumes or reason for consumption of large amount of power in data center. Two popular networks known as Fat-Tree and Three-Tier are shown below.

When links become aggregated, the amount of bandwidth on upstream links is often less than the bandwidth of the downstream links. This difference in bandwidth leads to a concept known as oversubscription. Oversubscription is defined as the ratio of maximal aggregate bandwidth among end hosts in the worst case to the total bisection bandwidth [17]. Oversubscription saves network resources at the expense of possible contention for access to those network resources. As the amount of oversubscription increases, the impact to performance is an undesirable consequence and poses a potential problem if left unaddressed. Thus addressing the traffic congestion with sustain our primary aim of reducing power consumption is a challenging task to solve on. Our work is basically tries to solve this challenge in the most efficient way compared to the existing models.
2.2 Related Work

2.2.1 Data Center Networks

The design of data center architecture plays a vital role in interconnection between servers. Most of the architectures aim to have at least one possible way of traffic flow between any two servers in a data center. The main challenging in architectures is how independent a server is to connect with any other server in the data center. The lesser the intermediate hops or connecting equipments the more independent the network is. Many approaches have been proposed to address data center design issues such as energy consumption, cabling complexity, scalability, and oversubscription. The authors in [2] classify data center networking architectures into four categories, electronic switching technologies, all optical switching, hybrid (optical and electronic switching), and wireless data center technologies. Let us discuss about Fat-Tree, Dcell, Bcube architectures in detail in this section and compare the efficiency with proposed S2S-WiDCN architecture in the next chapter. Each of these research categories seeks to improve the underlying networking architecture to solve various network challenges outlined in [2]. The wireless network approaches in [11] and [18] keep the switching advantages offered by wired networks and employ wireless technologies on top of existing wired approaches. This provides several solutions to unbalanced bursty traffic and oversubscription issues, however these approaches still maintain a large energy consumption due to the underlying wired network. Addressing the energy consumption of wireless DCNs is necessary to allow these networks to compete with emerging alternative network architectures.
Fat-Tree

Hierarchical DCNs are a fat-tree based architecture popularly adopted in the industry. Its three hierarchical layers consist of 160 access, 2 aggregate, and 2 core layer switches similar to the architecture evaluated in [22]. In the architecture, each traditional DCN link from servers to access layer switches has a channel capacity of 1.0Gbps. The links between access layer to aggregate layer are assumed to have a capacity of 1.0Gbps and those between aggregate layer and core layers are considered to be 10.0Gbps.

Dcell

DCell [21][17] is a type of recursive datacenter network topology. The most basic element, which is called DCell0, consists of n servers and one n-port switch. Each server in a DCell0 is connected to the switch in the same DCell0. The number of servers in DCell grows double-exponentially, and the number of levels in DCell is limited by the number of NICs on the servers in it. The next level of DCell is DCell1 which is constructed with n+1 DCell0s where n is the number of servers in DCell0. As a result, the DCell0s are connected to each other, with exactly one link between every pair of DCell0s. Following the same trend, the upper levels of DCells are constructed. For constructing the small-scale DCN network, we constructed a DCell with n=5 and k=2, which can support up to 930 servers. Here, k is number of levels for DCell. So, it is sufficient for supporting the required 800 servers. For the medium scale datacenter, DCell with n=6 and k=2 was constructed which can support up to 1806 servers. We used 1600 servers among these. All links are assumed to have a capacity of 1.0Gbps.
Bcube

BCube [14][16] is also a recursive topology specially designed for shipping container based modular datacenters. The most basic element of a BCube is BCube0. In this, n servers are connected to one n-port switch. The main difference between BCube and DCell is how they scale up. BCube makes use of more switches when constructing higher-level architecture. When constructing a BCube1, n extra switches are used, connecting to exactly one server in each BCube0. Therefore, a BCube1 contains n BCube0s and n extra switches. There are total 2n switches in a BCube1. We constructed the BCube network with n=10 and k=2 for small-scale DCN and n=12 and k=2 for medium scale DCN. Here, k is number of levels for BCube. Just like the DCell, the number of levels in a BCube depends on the number of ports on the servers. The number of servers in BCube grows exponentially with the levels, much slower than DCell. In the DCell and BCube network designs, the servers not only serve as end nodes but also act as relay nodes and take part in traffic forwarding through multiple parallel short paths. All links are assumed to have a capacity of 1.0Gbps.

2.2.2 Wireless Data Centers

Wireless data centers concepts have emerged from advancements in wireless technologies. Wireless frequencies at 60 GHz are capable of providing high data rates over distances of up to 10 meters. Highly directional antennas such as horned antennas or phased array antennas are used to establish wireless links only in one particular direction. These directional antennas enable spatial reusability due to favorable 60 GHz characteristics, however Line-of-Sight (LoS) communication is necessary to facilitate reliable communication [19]. Data center towers, metal frames,
and cooling equipment all represent obstacles that could increase channel losses and prevent links from being established. Several works have addressed these issues using 3D beamforming techniques to reflect 60 GHz signals off of metallic mirrors or through repeaters to achieve reliable communication [20] [18] [22]. The feasibility of establishing wireless 60 GHz links has been studied extensively, however using this technology within a data center environment remains an active area of study.

2.2.3 The 60 GHz band

A part of Millimeter band from 57 GHz to 64 GHz referred to as the 60 GHz band in the United States. This Millimeter band constitutes all frequencies between 30 GHz and 300 GHz. This range of frequencies is known as the millimeter wave band due to the fact that the wavelengths for these frequencies are between roughly 1 mm to 10 mm in length. The 60 GHz portion of the frequency spectrum was made unlicensed by the Federal Communications Commission (FCC) in an attempt to stimulate growth and innovation in order to facilitate the commercialization of 60 GHz technologies [23]. This same portion of the frequency spectrum has been made available worldwide, with small differences in the starting and ending frequencies between various communications regulatory bodies. For example, in Europe the European Telecommunications Standards Institute (ETSI) has made the frequencies from 57 GHz to 66 GHz unlicensed. Due to the global acceptance in making the 60 GHz band open for unlicensed use, this frequency band has received much research and attention. Figure 2.2 depicts the various allocations of the 60 GHz band in different geographic regions.

A common range of bandwidth centers on the spectrum mask at the 60.48 GHz frequency and is universally shared around the world. The 60 GHz band is beneficial
within a data center environment due the large amount of bandwidth available and also because of advantageous channel characteristics. The high carrier frequencies at 60 GHz allows for transceivers to achieve multi-gigabit data rates. Moreover, the limited interference at 60 GHz and potential for spatial reuse makes 60 GHz technologies ideal for use within a wireless data center.

### 2.2.4 60 GHz Standards

The standards for Wireless Communication are emerging day to day to meet the data rate required. Standards such as WirelessHD [24], IEEE 802.15.3c [25], and ECMA 387 [26], are designed for short ranged wireless personal area networking (WPAN) applications. These standards serve to provide short high-speed links for
multimedia scenarios. The WirelessHD standard for example, was designed for high definition video streaming in consumer products. Wireless local area networking (WLAN) standards addressing 60 GHz frequencies include the IEEE 802.11ad standard first published in 2012 and the WiGig standard first published in 2010. These standards address 60 GHz communications over larger distances and provide the natural evolution of Wi-Fi into the 60 GHz spectrum.

IEEE 802.11ay Standard

One of the most recent developments is the emergence of IEEE 802.11ay standard which has a larger bandwidth and data rates compared to the adopted 60GHz band. It is proposed to suffice about 20-40 Gbps for the larger range of upto 500m. The proposed architecture along with 802.11ay standard will prove very efficient or even better than the current wired or optical or both topologies combined. We have discussed in detail about this standard in the future work section.

2.2.5 Channel Characteristics

Modeling of 60 GHz frequencies is vital to the development of an accurate wireless data center design. In order to effectively model 60 GHz wireless links, the characteristics of the wireless channel must first be established. Wireless frequencies in lower frequency bands such as the ultra-high frequency (UHF) band (300MHz to 3 GHz) are able to propagate for miles and can penetrate solid obstacles without significant loss. This is due to the relatively large wavelength these frequencies exhibit. 60 GHz frequencies however, are operating at millimeter wavelengths, which limits both the distance they can propagate and the materials they can propagate through. The wavelengths of 60 GHz frequencies are around 5mm, meaning objects
of only a few millimeters will effectively block the propagation of a 60 GHz wave, restricting wireless communication to LoS connections only. For reliable 60 GHz communication to take place, a receiver antenna must receive a signal above a minimum receiver sensitivity level to accurately demodulate the wireless signal. This relationship can be expressed through the following equation.

\[ \text{Received Power} = \text{Transmit Power} + \text{Gains} - \text{Losses} \]

Where transmit power is the power that the transmitter module produces, gains represents the directionality and efficiency of the antenna, and losses represent any deterioration of the signal as it propagates from the transmitter to the receiver. This transmission power is dictated by FCC regulations, limiting the maximum average power level of 60 GHz signals to 40 dBmi [27]. The antenna gain is dependent on the type of antenna used, such as an omni-directional antenna, a horn antenna, or a phased antenna array. The last aspect of the equation is the losses incurred during the transmission. By understanding and mitigating the cause of loss in 60 GHz communications, better data rates can ultimately be achieved. Several factors contribute towards the propagation loss of 60 GHz signals and must be accounted for in order to obtain an accurate representation of the communication channel. 60 GHz signals operating at high data rates are around 55 dB worse than 2.4 GHz wireless links in terms of the signal to noise ratio (SNR) of the channel [11]. This difference can be attributed to two main factors, including free space path loss and the size of the channel. Free space path loss is the main form of loss 60 GHz frequencies face in LoS communication channels. The free space path loss can be expressed in terms of frequency and the distance between any two isotropic antennas shown in Equation 2.1.

\[ L_{\text{FSL}} = (4\pi R/\lambda)^2 \]  
(2.1)
Where $R$ is the distance between the transmitter and receiver antennas and $\lambda$ is the wavelength of the operating frequency. When converting to units of frequency and expressing Equation 1 in terms of dB it becomes:

$$F_{(FSL, dB)} = 92.4 + 20 \log (f) + 20 \log (R) \quad (2.2)$$

This free space path loss becomes significant at high frequencies and accounts for the majority of loss in millimeter wave bands such as the 60 GHz RF band. In addition to the attenuation due to free space, 60 GHz signals also face transmission losses through air due to atmospheric conditions. The 60 GHz frequency coincides with an absorption peak of oxygen, meaning molecules of O2 absorb 60 GHz frequencies. This absorption characteristic of oxygen along with absorption due to water vapor in the air results the attenuation of RF signals. This attenuation results in shorter propagation distances, however it also reduces the interference at other transceivers caused by 60 GHz signals. While path loss accounts for the majority of loss in the transmission of 60 GHz frequencies, the size of 60 GHz channels are roughly 100 times wider than 2.4 GHz channels. This creates a channel that is over 20 dB noisier at high data rates [28]. Additional sources of loss such as multi-path effects are largely mitigated through the use of highly directional antennas [29]. The directionality of the antennas greatly reduces indoor multi-path signal variations.

### 2.2.6 Antenna Characteristics

The antenna technology used in a 60 GHz transmission dictates the radiation pattern the signal produces. Omni-directional antennas propagate in all directions evenly, while directional antennas concentrate a wireless signal in one specific direction.
which is vital to avoid interference. Directional antennas can be created by using a “antenna array” to direct the propagation of the waveform in one direction electronically steered using phased array techniques. These radiation patterns are typically measured in decibels isotropic (dBi), where the forward gain of the directional antenna is compared with a hypothetical isotropic antenna. The antenna array has a very less latency for directionality with comparison to that of the physical directional antennas like horn antenna.

2.2.7 Spatial Reuse

The most important characteristics of our proposed architecture is the spatial reuse of 60 GHz signals. Since these signals are generated using highly directional antennas, multiple transceivers can operate on the same frequency at same time within proximities greater than 24 inches [11]. Spatial reuse isn’t possible with wireless communications in lower frequency ranges for the reason that their signals don’t attenuate in free space as rapidly as 60 GHz signals. The ability to form simultaneous communication channels on the same frequency at extremely high data rates is what makes the 60 GHz band attractive to use within a data center environment. Along with IEEE 802.11ay standard it is very likely to exceed the efficiency of traditional wired data center in terms of both power consumption and performance.

In order to obtain high bandwidth links for use within a data center environment, LoS must be maintained between transceivers [19]. Furthermore, within a data center, when these wireless transceivers and antenna modules are used in a ToR configuration, link blockage becomes a major problem. Dense rack deployments in standard data center layouts presents limitations for the formation of LoS links at the ToR level. Only neighboring racks are able to establish reliable 60 GHz links
Chapter 2. Background & Related work

[30]. This presents a challenge for ToR transceivers that necessitate communications across multiple isles. On the other hand, a medium access challenge exists in determining what non-interfering links can be established in the same channel at the same time. Each of these challenges and considerations are addressed in the proposed approach.

2.2.8 Significance

My work will contribute towards the investigation into wireless data centers utilizing the Server-to-Server Wireless Data center Network(S2S-WiDCN) architecture designed and the unlicensed 60 GHz frequency band. A completely wireless DCN will be explored with the goal of not only achieving efficiency, reduced oversubscription, scalability but also reduced power consumption and increased independence (Less number of intermediate hops, maximum of two hops). My S2S-WiDCN architecture along with IEEE 802.11ay standard will prove to be very efficient and their characteristics, feasibility will be evaluated in future works to be done.
Chapter 3

Traffic Modeling

In order to evaluate the performance of a data center network architecture it is necessary to study and model the traffic in a data center. The nature of traffic in a data center is studied and the characteristics of it are generalized. The main attributes of traffics are Injection time, Source Node, Destination node, Size of Data, Data Rate (often considered as injection rate). The study of traffic in [31] is modeled and is used to stimulate the data center traffic for comparison of proposed S2S-WiDCN architecture with that of practically used architectures like Fat-tree. Some of the characteristics of traffic are discussed in detail in this section.

3.1 General Characteristics

The main attributes of a data center traffic are Injection time, Source Node, Destination node, Size of Data, Data Rate (often considered as injection rate). Injection time specifies traffic’s arrival time at data center. It is specified in seconds. Most of the traffic needs to be surfed at least more than through a couple of server to reach the required information. The source node is the first server of contact and Destination node is the server that contains needed information. In our work we are assigning
the servers with an unique number for identification and is used for specifying the Source and Destination in traffic modeling. The numbering of server is done in a orderly manner to suffice the routing protocol designed in this work. The naming of servers will be more lucid in the routing part the work. Size of data is the amount of data in a single traffic. It can be specified in Bytes, bits. Data rate is the rate at which traffic arrives data center, also known as Injection rate. It is specified in Gigabits per Second or in any standard form of notation like Gbps, Bps.

![Figure 3.1: CDF of Flow Transmission rate in Data Center with 800 Servers](image)

### 3.2 Flow Transmission Rate

Flow transmission rate provides us the rate at which the traffic arrives a data center. It is studied that Injection rate follows a Gaussian distribution [31]. In [31] the traffic is studied in a university data center and is modeled with respect to the conclusion from the study. Injection rate plays an important role in deciding the duration required to retrieve the information from server. One of the factors that acts as a
Chapter 3. Traffic Modeling

3.3 Duration

The time required for the traffic from the start at source server till completion is regarded as the duration of that traffic. It is explored that the duration of traffics in a data center follows a Pareto distribution. It is observed that almost about 80% of the traffics duration is less than 11 Seconds each and duration of rest of traffic is higher than that. It can be up to maximum of 200 seconds for 0.1% of traffic. The duration threshold in determining the characteristics of traffic duration is Injection rate. The effect of Injection rate on duration, latency will be discussed on detail in the results of simulation in next chapter. Since the traffic is modeled with respect to [31], the mean of Gaussian injection rate is 4303.69 Bps and variance is 69936.37 Bps$^2$. The value stated above is taken as base case and further investigation on higher rates are explored in this work.
Chapter 3. Traffic Modeling

3.4 Traffic Size

The size of the traffic is the amount of data needed to be transmitted. It can be expressed as a product of Duration and Injection rate. Thus, Size of traffic helps to determine the duration to which the traffic sustains in a data center.

\[
\text{Size of Traffic (Bytes)} = \text{Duration (s)} \times \text{Injection rate (B/s)}
\]

From the expression it can be learned that characteristics of size depends on duration and injection rate. It is dependent of two different distribution and its characteristics is shown in figure 3.5 & 3.6.
3.5 Significance

It is very important to generate a traffic similar to real world. It helps us determine the efficiency and other parameters required for staying in competition with the existing architecture and interconnections. In our study we have designed traffic similar to the characteristics of real university data center and evaluated the efficiency of various traditional architectures and have compared it with our proposed architecture, routing protocol and 60 GHz band.
Figure 3.5: CDF of Flow Size in Data Center with 800 Servers

Figure 3.6: CDF of Flow Size in Data Center with 1600 Servers
Chapter 4

Server-to-Server Wireless Data Center Network (S2S-WiDCN) Design

The architecture of S2S-WiDCN is organized into several parts. First, topology is explained in depth, along with several important design considerations. Next, the 60 GHz antenna technology adopted in this work is examined. Then, the routing path establishment mechanisms are explained and supported. Following route selection is a description of the chosen network protocols used to realize the wireless DCN approach. Finally, this chapter will conclude with a section devoted to the method of power modeling used to evaluate the power consumption of the DCNs.

4.1 Architecture

The proposed S2S-WiDCN architecture mainly depends on the identifying server location and then defining a routing path depending on the server’s location. Thus, numbering of servers will be an important step in the proposed architecture. A patterned numbering will help in locating the source and destination thereby, helping in the process of route selection. This work will then explain both the proposed
wireless architecture and the methodologies utilized in the development of this architecture.

4.1.1 Topology

A completely wireless interconnection is adopted in the proposed S2S-WiDCN architecture, meaning any two servers with LoS have the capability of communicating using 60 GHz wireless transceivers. In the proposed architecture, each server is equipped with a wireless module to establish completely wireless communication. In this work, communication is divided into horizontal and vertical plane commu-

![Antenna placement in a server](image)

**Figure 4.1**: Antenna placement in a server
To establish communication between servers in the vertical plane, an antenna array is placed at the back of the server but the establishment of a connection between any two servers is at the discretion of availability of LoS. Similarly, another antenna array is placed at the top of a server to enable communication in the horizontal plane. Each wireless module contains a transceiver and two antenna array capable of communicating with any other wireless module within the data center. The wireless data center is physically bound by the distance that a 60 GHz wireless signal can reliably transmit to any other module.

Data center racks are laid out in a traditional configuration, with aisles running between rows of data center racks. Space of 1U or 2U is provided between the servers in a rack to support horizontal plane communication.

![Figure 4.2: Layout of the proposed Data Center](image-url)
Horizontal Plane Communication

Antenna array at the top of a server is used to establish communication in the horizontal plane. To avoid interference and obstructions from the rack frames, communication in the horizontal plane are restricted only to a single line between horizontally aligned servers. Horizontal plane communication enables LoS between servers in different row and same column & height. Figure 4.3 shows an example for horizontal communication.

![Figure 4.3: Horizontal Plane Communication](image-url)
Vertical Plane communication

Vertical plane communication helps in connecting servers in the same row. Distance between two aisles helps to spatially reuse the 60 GHz band. This helps to maintain large number of connections simultaneously while increasing the efficiency of wireless communication. The antenna array is arranged in such a manner to enable LoS between all the servers in the same row. This is achieved by mounting the antenna array in an elliptical manner with very less height change.
4.1.2 Antenna Technology

Each server of the datacenter will be equipped with a wireless module consisting of a transceiver and two accompanying antenna arrays [12]. This particular array [12] is fabricated using semiconductor lithography techniques on a single wafer and is hence, extremely compact with a small size. As the radiation pattern suggests, the array provides high directional gain of 9 dBi in the forward and backward directions. Moreover, by adjusting the relative phase of the antenna elements by activating various feed paths, beam-steering can be accomplished over an angle of 60 degrees in the plane of the array. As horizontal communication happens in a single straight line, no steering is required in the antenna arrays on top of the servers. However, as the range of steering angle is 60 degrees for this particular array, 6 antenna arrays are required to cover the entire 360 degree panorama in the vertical plane. Only one out of the 6 arrays will need to be signaled at any given point of time to establish a single link involving that server. Electronic beam-steering for the antenna array has negligible latency compared to mechanically steered horn antennas used in earlier wireless DCNs [20]. Moreover, the antenna array being extremely compact requires very tiny space on top of each server to enable LoS communication in the horizontal direction. The effect of these spaces on the vertical server density in the datacenter racks is discussed and quantified in the results section.

The beam-steering of the transmitting and receiving antennas is achieved by using a separate control interface using 2.4/5 GHz WiFi bands. Although the data rates sustained by the WiFi bands are much lower than the 60GHz bands, it is sufficient for the short control packets. Moreover, the isotropic antennas in the WiFi modules do not require any antenna steering before the control messages can be transmitted.
When a traffic flow between a pair of servers is created, a short control or header packet for the flow will be sent over the WiFi band to enable communicating servers to steer their antennas towards each other when required.

### 4.1.3 Establishment of Wireless Links

A wireless module will be able to connect to any other wireless ToR module within the data center. This however, doesn’t mean that the servers are fully connected. Whenever a wireless module connects to another wireless module, neither of these modules can connect to any other wireless module on the same wireless channel. Furthermore, although multiple wireless links can be formed at the same time, a wireless link has the possibility of creating interference for other transmitting wireless modules. Intelligent link selection is necessary to prevent interference and to optimize the establishment of links.

The link selection can be thought of as a type of scheduling problem. Given a list of application demands over a period of time, wireless links must be established and removed to satisfy bandwidth demands. These application demands can be generated through trace driven approaches or predictive mechanisms. Using this list of known demands, wireless links are generated until an optimal state is achieved. An optimal state can defined as a state in which the maximum amount of data is transferred globally across the entire data center. Given the large number of possible combinations of links that can be established and the large set of application demands, this scheduling problem becomes NP-Complete in computational complexity. A simple greedy approach is taken to connect the wireless modules with the largest demands on a first come first serve scheme. The greedy algorithm used is depicted in Figure 4.5.
The algorithm establishes as many non-interfering links as possible on a first-come-first-serve basis until all demands are exhausted or the number of available channels runs out. The list of unaddressed demands is ordered by the start time of each demand. While this greedy approach may not reach an optimal state, it serves to achieve a good enough state with minimal computational effort.
Chapter 4. Server-to-Server Wireless Data Center Network (S2S-WiDCN) Design

4.2 Routing Protocol

In this section, we will discuss about the routing protocol designed to implement wireless communication in the proposed S2S-WiDCN architecture in the previous section. The topology helps in spatial re-usability of the frequency and the 60 GHz band helps in attaining the data rates comparable to the existing wired or hybrid architecture. Firstly, we will discuss about the allocating unique ID for servers and then will explore the steps involved in routing.

4.2.1 Server ID

An unique numbering pattern is followed throughout the data center to support the proposed routing protocol. This step plays a vital role in sorting out the path to be traversed in the data center. In this section, we will try to explain the server ID allocating process with an example considering data center having about thousand racks.

The rows are assumed to be in parallel with the X-axis of the 3D Cartesian coordinate system. The servers are staked in racks along the Z-axis with spaces between them to support HP Communication. The servers are allotted with a three digit number with corresponding to the example considered. The hundred’s digit corresponds to the Z-axis. The bottom servers in all racks are assumed to be at zero level and increases by one moving upwards. Thus in our example, the servers on the top of racks are at level 9. The ten’s place implies the row to which the server belongs and the unit’s digit tells the exact position in the row specified by Ten’s digit.


Chapter 4. Server-to-Server Wireless Data Center Network (S2S-WiDCN) Design

4.2.2 Inter and Intra Rack Communication

In our work, classification of communication between servers will help in understanding the proposed routing protocol. Communication in a data center can be classified into Inter Rack and Intra Rack communication based on the servers location. In a Data Center it is studied that almost eighty percent of the traffic is Intra Rack. If servers communicating are present in the same Rack then such a communication is known as Intra Rack communication. Intra Rack communication is depicted in the Figure 4.6.

Communication between servers in different racks at the data center is known as Inter Rack communication. Inter Rack Communication can be further divided into Horizontal Plane and Vertical Plane Communication as discussed in section 4.1.1.
Chapter 4. Server-to-Server Wireless Data Center Network (S2S-WiDCN) Design

All Intra Rack Communications fall under Vertical Plane Communication as servers participating are in the same rack. Let us briefly discuss Inter Rack Communication.

**Horizontal Plane Inter Rack Communication**

Communication between servers in same Z-axis & X-axis level and different Y-axis falls under this category. The antenna array at the top of server is used for communication. Figure 4.7 shows an example for this type of communication.

---

**FIGURE 4.7: Inter Rack Communication along Horizontal Plane**
Chapter 4. Server-to-Server Wireless Data Center Network (S2S-WiDCN) Design

**Vertical Plane Inter Rack Communication**

Servers in the same row communicate using the Vertical Plane, such communications are classified as Vertical Plane Inter Rack Communication. An example for Vertical Plane Inter Rack Communication is shown in Figure 4.8.

![Inter Rack Communication along Vertical Plane](image)

**Figure 4.8: Inter Rack Communication along Vertical Plane**

Apart from the cases discussed above, certain communication require both Vertical plane and Horizontal plane communication. These communications require an intermediate server for routing the traffic to the destination server. The maximum number of hops required to communicate between any two servers in the proposed
model is two. Figure 4.9 depicts an example of communication requiring two hops.

The routing of traffics requiring two hops can be implemented either by executing Horizontal plane communication followed by Vertical plane communication or by first implementing Vertical plane communication followed by Horizontal plane communication. In our model, we have adopted the horizontal plane first method. Thus communication between any two servers in a data center is made possible with the proposed routing protocol. The worst case of the proposed routing protocol in terms of number of the hops required for completion is limited to two.
Chapter 4. Server-to-Server Wireless Data Center Network (S2S-WiDCN) Design

A control packet with instructions for intermediate and destination servers to perform beam-steering in the correct directions is sent over a separate IEEE 802.11 2.4/5 GHz ISM band channel. Each server is equipped with a separate IEEE 802.11 transceiver. As the radiation pattern of the 60GHz antenna array has main lobes in both forward and backward directions with zero phase difference, beam-steering is not required for the horizontal linear communication (see Fig. 4.3). For communications in the vertical planes, the server first sends a control packet to the receiver while simultaneously activating one of six antennas in the array and steering its beam towards the receiver. Upon receipt of this control message, the receiver’s wireless module chooses the antenna array in the correct sector out of the set of six and steers that beam towards the sender though appropriate beamforming (setting the antennas weights).

4.3 Power Consumption

Estimation of power consumed in a data center is a complex task. The power consumption depends on several internal factors such as utilization of computing power, the cooling mechanism, and datacenter networks. It is also affected by external parameters like the geographical location, weather, temperature, and humidity. In particular, the network contributes significantly to power consumption in a datacenter [3]. Our focus is solely on networking, and we only analyze the power consumption involved in networking. In that regard, we assume that the power consumption other than networking is identical in all the cases. We estimate power consumption for wired DCNs using commercially available data from Cisco network switches [32][33][34]. Specifically, we use Cisco 7702 for the core-level switches, Cisco 9508 at the aggregation level, and Cisco 9372 for access-level


Chapter 4. Server-to-Server Wireless Data Center Network (S2S-WiDCN) Design

Table 4.1: Power Consumption of different DCN components

<table>
<thead>
<tr>
<th>Device</th>
<th>Model</th>
<th>Adapted in</th>
<th>Power Consumption (Watts)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Typical</td>
<td>Maximum</td>
</tr>
<tr>
<td>Access Layer Switch</td>
<td>Cisco 9372 Fat-Tree, Optical, ToR WiDCN</td>
<td>210</td>
<td>650</td>
</tr>
<tr>
<td></td>
<td>Cisco 3560CX-8PD DCell</td>
<td>24.20</td>
<td>24.40</td>
</tr>
<tr>
<td></td>
<td>Cisco 3560CX-12PD BCube</td>
<td>28.90</td>
<td>26.50</td>
</tr>
<tr>
<td>Aggregate Layer Switch</td>
<td>Cisco 9508 Fat-Tree</td>
<td>2527</td>
<td>3324</td>
</tr>
<tr>
<td>Core Layer Switch</td>
<td>Cisco 7702 All</td>
<td>837</td>
<td>1305</td>
</tr>
<tr>
<td>Network Interface Card</td>
<td>Silicom PE2G2135 Fat-Tree, Optical, ToR WiDCN</td>
<td>2.64</td>
<td>3.36</td>
</tr>
<tr>
<td></td>
<td>Silicom PE2G4135 DCell, BCube</td>
<td>3.60</td>
<td>5.04</td>
</tr>
<tr>
<td>Antenna and Wireless</td>
<td>Transceiver and Antenna Power ToR WiDCN, S2S-WiDCN</td>
<td>0.50</td>
<td>1</td>
</tr>
<tr>
<td>WiFi Adapter</td>
<td>D-Link DWA-171 S2S-WiDCN</td>
<td>0.22</td>
<td>0.24</td>
</tr>
</tbody>
</table>

switches. We also use the data from Silicom PE2G2135 for the power consumption of the network interface cards (NIC) [35]. More formally, the power consumption of each core and aggregate switches are as follows:

\[
P_{Core} = P_{I/O} + P_{FanTray} + P_{sv}
\]

\[
P_{Agg} = P_{I/O} + P_{FanTray} + P_{sv} + P_{Fabric} + P_{SysCtrl}
\]

Where \(P_{I/O}, P_{FanTray}, P_{sv}, P_{Fabric}, P_{SysCtrl}\) represent the power consumption of the input/output card, fanout ports, supervisor controller, cables and system controller respectively. Then the total power is calculated as follows:

\[
P_{Total} = N_{Core}P_{Core} + N_{Agg}P_{Agg} + N_{Acc}P_{Acc} + N_SP_{NIC}
\]
Where $N_{\text{Core}}$, $N_{\text{Agg}}$, $N_{\text{Acc}}$, $N_S$ are the number of core, aggregation, access switches, and the total number of servers respectively.

Unlike the fat-tree wired DCN, both DCell and BCube use low-power commodity switches instead of power-hungry aggregate- and access-layer switches. For this, we consider Cisco 3560CX-8PC-S [34] and Cisco 3560CX-12PD-S [34] for small and medium-scale datacenters, respectively. For the NIC, we assume Silicom PE2G4I35 with multiple connecting ports essential for these architectures. We also consider two core switches that act as gateways to the external network. We again use $P_{\text{Core}}$ equation to estimate their power consumption but without aggregate-layer switches.

The fixed power dissipation of the infrastructure dominates the power consumption of network devices. The power consumption is not affected by data rates much, and therefore this dependency is largely ignored in the estimation.

In S2S-WiDCN, however, no core, aggregate or access layer switches are needed, but only antennas, transceivers and NICs are required for wireless communication. The power consumption of the wireless 60GHz transceiver is measured based upon the assessment of emerging 60GHz transceivers such as [8]. The NICs of S2S-WiDCN are equipped with two transceivers for horizontal and vertical communication. In the traditional DCN, external connections are established via the two core switches. To provide equivalent connectivity in S2S-WiDCN, we employ two servers to work as gateways, and their power consumption is modeled as that of Cisco 7702. We use the D-link DWA-171 WiFi adapter for the control channel. Now the power consumption of communication per server in S2S-WiDCN is calculated as:

$$P_{\text{Wireless}} = P_{60 \text{GHz Tran}} + P_{\text{Antenna}} + P_{\text{WifiControl}}$$

where $P_{60\text{GHz Tran}}$ is the power consumption of the 60GHz transceiver, $P_{\text{WifiControl}}$ is
Chapter 4. Server-to-Server Wireless Data Center Network (S2S-WiDCN) Design

the power consumption of the WiFi adapter for the control channel, and $P_{\text{Antenna}}$ is
the power transmitted through the antennas. We conservatively adopt $P_{60\text{Ghz Tran}} + P_{\text{Antenna}}$
to be 500mW for an average case, and 1W the for maximum case from [8]. We con-
sider $P_{\text{Wifi Control}}$ to be 220mW for the datasheet of D-link DWA-171. Finally, the total
power consumption in S2S-WiDCN becomes:

$$P_{\text{Total WiDCN}} = N_{\text{Core}}P_{\text{Core}} + N_{s}P_{\text{Wireless}}$$

4.3.1 Analysis of Power Consumption

The primary advantage of S2S-WiDCN is lowered power consumption. To study
this more deeply, we show the total power consumption estimated for the typical
and maximum cases for all the DCNs in Figure 4.10. In the typical scenario, the
average power consumption of every device is used, while the maximum power
consumption is considered in the maximum cases. For small-sized and mid-sized
DCNs, the result shows at least 6-fold and 18-fold reduction in both maximum and
typical power consumption of S2S-WiDCN compared to the traditional DCN topol-
ogy respectively. The complete elimination of power-hungry aggregate and access-
layer switches contribute to this drastic reduction primarily. In comparison to both
DCell and BCube, roughly 6-fold power reduction is observed for S2S-WiDCN for
both small and medium-sized datacenters, DCell and BCube consume less power
compared to the fat-tree DCN due to the low power commodity switches that these
DCNs use. This power saving, however, comes at the cost of huge cabling com-
plexity. Unlike DCell and BCube, S2S-WiDCN can save power consumption sig-
nificantly, specifically at least 6-fold and 18-fold reduction. A media converter is
needed for every link in the optical network to convert electrical signals to optical
ones, and these converters use non-negligible power. Moreover, the signals are also converted in switches and routers ultimately increasing the total power consumption of the switches and routers. Since access-level switches are needed per rack in the ToR-WiDCN, its reduction in power consumption not as significant as the optical network. These estimates indicate the potential reduction in power consumption is up to 18-fold if wireless interconnections are used in DCNs. Next, we evaluate the performance of S2S-WiDCN and its comparison with alternate DCNs.

4.4 Simulation Platform

We used the NS-3 network simulator [36] that supports testing the characteristics of wireless propagation as well as network-level communications. It is critical to simulate the characteristics of both wireless propagation and network-level communication accurately to obtain realistic performance results. We used a modified version of NS-3 extended with features of wireless datacenters including the 60GHz band.
and the IEEE 802.11ad standard as discussed in [11]. This extension incorporates interference modeling, bit error rates, and directional antenna modeling. The accuracy of these parameters is verified with physical layer measurements from their prototype 60GHz hardware [11]. Additionally, we introduced criteria for wireless link selection to enable many concurrent links, and modify the IEEE 802.11ad physical layer channel to allow multiple OFDM channels to operate at gigabit speeds. The latencies for the exchange of header packets via the WiFi channels and electronic beam-steering are negligible as they are orders of magnitude smaller than the average flow transmission duration. The overall latency includes the transmission delay, propagation delay, queuing delay, and the delay introduced by beam steering and control packet exchange. We consider two datacenter sizes in the experiment. The first one is a small-sized datacenter with a total of 800 servers representing those in an educational institution. The second one is a mid-sized datacenter, and has 1600 servers representing those in private enterprises. In both cases, the servers are arranged in a 20 by 8 array of racks. There are 10 racks arranged in a single row and two columns of 8 rows, totaling 160 racks. Each rack occupies an area of 0.6m by 0.9m and is 2m high containing 5 servers. Adjacent rows are separated by 1m and the width of the central aisle is 2m. This separation between rows coupled with the directional radiation pattern of the antennas ensures that the vertical planes of wireless communication are spatially isolated meaning that communications in one plane do not interfere with those in the next

4.5 Performance Analysis

In this section, the simulation results of S2S-WiDCN along with a comparative analysis against other DCNs in terms of flow completion duration and throughput.
4.5.1 Traffic Flow Completion Duration

The flow completion durations for all the DCNs are shown in Figure 4.11. The average completion duration of S2S-WiDCN is lower than that of all the wired networks for both 800 and 1,600 servers. S2S-WiDCN and the other DCNs considered in this work were able to successfully complete routing of all the flows that were exchanged between the servers over the duration of simulations. Fewer hops are involved in data delivery in S2S-WiDCN, it results in lower time of flight and switching overheads. Table 4.2 shows the packet latencies of representative network devices for various DCN architectures that are considered in this paper. As can be seen in table each NIC has a packet delay of around 30s. Due to higher number of hops through the NICs in BCube and DCell compared to S2S-WiDCN signified by higher hop-counts, the flow control duration in them are higher. DCell shows the highest average flow completion durations because it requires a large number of hops ranging from 2 to 10. BCube performs slightly better than DCell because of lower number of hops involved in data delivery. This advantage mainly comes from more number of alternate paths created by more number of switches which reduces the average
number of hops in BCube. In the traditional fat-tree wired network, two servers even in a single rack need to go through their access layer ToR switch to communicate, traversing at least 2 hops. Even intra-rack flows incur the access layer switching delay as shown in Table 4.2 in the fat-tree DCN. Inter-rack flows incur higher delay through the other switching layers as well shown in Table 4.2. In S2S-WiDCN, however, two intra-rack servers can communicate directly in a single hop incurring only the packet delay from the NICs as shown in Table 4.2, transceiver delay and time of flight. The transceiver delay in 60GHz wireless transceivers is 0.1 ns per bit [8]. Therefore, the packet latency of the wireless links is 150 ns. The time of flight is in the order of nanoseconds and is accounted for, by NS-3 based on the distance between the communicating servers. Therefore, the direct server-to-server wireless communication in S2S-WiDCN through the NICs and the wireless transceivers has lower flow completion duration compared to the hierarchical Fat-Tree DCN where the flows have to traverse through slower upper-layer switches in addition to the NICs. The beam-steering latency is 266 microseconds and is considered while computing the flow completion duration of S2S-WiDCN in Figure 4.11. The majority of datacenter communication is known to be intra-rack [31], and this makes the wireless S2S-WiDCN reduce delays for most of the packets, which are intra-rack due to single-hop wireless links along the vertical planes.

4.5.2 Throughput

Throughput is also measured for these different DCN architectures under the traffic conditions described. We are interested in whether S2S-WiDCN imposes any adverse impact on throughput. The results show that S2S-WiDCN provides comparable throughput to the wired fat-tree DCN for both the DCN sizes. Also, all the
Chapter 4. Server-to-Server Wireless Data Center Network (S2S-WiDCN) Design

<table>
<thead>
<tr>
<th>Device</th>
<th>Model</th>
<th>Adapted in</th>
<th>Average Packet Delay</th>
</tr>
</thead>
<tbody>
<tr>
<td>Access Layer Switch</td>
<td>Cisco 9372</td>
<td>Fat-Tree, Optical, ToR WiDCN</td>
<td>1.5 μs</td>
</tr>
<tr>
<td></td>
<td>Cisco 3560CX-8PD</td>
<td>DCell</td>
<td>4.9 μs</td>
</tr>
<tr>
<td></td>
<td>Cisco 3560CX-12PD</td>
<td>BCube</td>
<td>4.9 μs</td>
</tr>
<tr>
<td>Aggregated Layer Switch</td>
<td>Cisco 9508</td>
<td>Fat-Tree</td>
<td>3.6 μs</td>
</tr>
<tr>
<td>Core Layer Switch</td>
<td>Cisco 7702</td>
<td>Fat-Tree</td>
<td>5 μs</td>
</tr>
<tr>
<td>Network Interface Card</td>
<td>Silicom PE2G2135</td>
<td>Fat-Tree, Optical, ToR WiDCN</td>
<td>~30 μs</td>
</tr>
<tr>
<td></td>
<td>Silicom PE2G4135</td>
<td>DCell, BCube</td>
<td>~30 μs</td>
</tr>
<tr>
<td>Antenna and Wireless</td>
<td>Transceiver</td>
<td>ToR WiDCN, S2S-WiDCN</td>
<td>~150 μs</td>
</tr>
</tbody>
</table>

DCNs achieve throughput that closely matches the flow transmission rates.

The main reason to have a lower number of hops in S2S-WiDCN is to achieve fewer intermediate nodes with the possibility of lesser network congestion. All these results and analysis indicate that S2S-WiDCN outperforms the fat-tree DCN as well as other server-centric networks including DCell and BCube.

4.6 Conclusions

The challenges in current Data Center Network’s are high design and maintenance cost, huge power consumption, high cabling complexity which are hard to keep accurate per-cable information and inefficient cooling. Structured cabling bundle incur significant initial effort and cost to setup and still may cause airflow blockage. All these challenges can be overcome by using the proposed completely wireless
Chapter 4. Server-to-Server Wireless Data Center Network (S2S-WiDCN) Design

4.3 Comparison between different data center architectures

<table>
<thead>
<tr>
<th></th>
<th>No. of Servers</th>
<th>Wired Fat-Tree</th>
<th>DCell</th>
<th>BCube</th>
<th>ToR-WiDCN</th>
<th>S2S-WiDCN</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Average Flow Completion Duration (Seconds)</strong></td>
<td>800</td>
<td>1.128</td>
<td>1.2108</td>
<td>1.1810</td>
<td>1.1283</td>
<td>1.0847</td>
</tr>
<tr>
<td></td>
<td>1600</td>
<td>1.134</td>
<td>1.2514</td>
<td>1.1914</td>
<td>1.136</td>
<td>1.063</td>
</tr>
<tr>
<td><strong>Average Throughput (KBps)</strong></td>
<td>800</td>
<td>0.971</td>
<td>0.961</td>
<td>0.970</td>
<td>0.959</td>
<td>0.999</td>
</tr>
<tr>
<td></td>
<td>1600</td>
<td>0.964</td>
<td>0.949</td>
<td>0.964</td>
<td>0.951</td>
<td>0.989</td>
</tr>
<tr>
<td><strong>Minimum No. of Hops</strong></td>
<td>800</td>
<td>6</td>
<td>10</td>
<td>8</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>1600</td>
<td>6</td>
<td>10</td>
<td>8</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td><strong>Average No. of Hops</strong></td>
<td>800</td>
<td>2.54</td>
<td>3.71</td>
<td>3.25</td>
<td>2.203</td>
<td>1.172</td>
</tr>
<tr>
<td></td>
<td>1600</td>
<td>2.60</td>
<td>3.77</td>
<td>3.29</td>
<td>2.214</td>
<td>1.182</td>
</tr>
</tbody>
</table>

server-to-server DCN architecture. We observe that the S2S-WiDCN improves the flow completion duration and throughput compared to conventional fat-tree based DCNs for typical query/response based applications, while reducing the power consumption by six to eighteen times.

4.7 Future Work

The emergence of IEEE802.11ay provides high possibilities of extended transmission, resulting in increase in data center size. The proposed wireless server-to-server architecture along with incorporation of other interconnect medium like optical can contribute to increase in efficiency of data center for known type of traffic. Such Hybrid configurations are potential arenas to investigate to enhance the proposed architecture. It will also help to reduce the sole dependency on wireless bands. Caching more frequently used data to servers that are under-utilized may help in
reducing the power consumption from software prospective. As the dependency on social network increases, the need for improvements in data center becomes vital and this research work will serve as a starting point for completely wireless communication in an indoor communication environment.
Bibliography


