JOINT VIRTUAL MACHINE EMBEDDING AND WIRELESS DATA CENTER TOPOLOGY MANAGEMENT

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

BEYZA BÜTÜN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
COMPUTER ENGINEERING

MAY 2022

Approval of the thesis:

JOINT VIRTUAL MACHINE EMBEDDING AND WIRELESS DATA CENTER TOPOLOGY MANAGEMENT

submitted by **BEYZA BÜTÜN** in partial fulfillment of the requirements for the degree of **Master of Science in Computer Engineering Department, Middle East Technical University** by,

Prof. Dr. Halil Kalıpçılar
Dean, Graduate School of **Natural and Applied Sciences** _______________

Prof. Dr. Halit Oğuztüzün
Head of Department, **Computer Engineering** _______________

Prof. Dr. Ertan Onur
Supervisor, **Department Of Computer Engineering, METU** _______________

Examining Committee Members:

Prof. Dr. İbrahim Körpeoğlu
Department Of Computer Engineering, Bilkent University _______________

Prof. Dr. Ertan Onur
Department Of Computer Engineering, METU _______________

Assist. Prof. Dr. Pelin Angın
Department Of Computer Engineering, METU _______________

Date: 10.05.2022

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Surname: Beyza Bütün

Signature :

ABSTRACT

JOINT VIRTUAL MACHINE EMBEDDING AND WIRELESS DATA CENTER TOPOLOGY MANAGEMENT

Bütün, Beyza
M.S., Department of Computer Engineering
Supervisor: Prof. Dr. Ertan Onur

May 2022, 54 pages

With emerging technologies such as the Internet of Things and 5G, generated data grows enormously. Hence, Data Center Networks (DCNs) have an important duty to store and process a significant amount of data, which makes them a critical component of the network. To meet the massive amount of traffic demands, wired DCNs need to deploy large numbers of servers and power-hungry switches, and huge lengths of wires. An enormous increase in the usage of cables causes high cabling complexity and cost while deploying large numbers of servers and power-hungry switches causes high-level power consumption. To address these problems, Wireless DCNs (WDCNs) have emerged. WDCNs utilize the bandwidth efficiently thanks to their adaptation ability to the dynamically changing conditions and decrease high cost, cabling complexity, and power consumption problems with the help of their wireless structure. WDCNs are also a solution to the oversubscription problem. However, establishing reliable communication, satisfying changing traffic demands dynamically, and increasing throughput are challenges of WDCNs that need to be emphasized. In this work, we introduce a WDCN using internationally available 60 GHz bands and aim to handle the dynamic traffic demands in the WDCN and maximize the through-

put. To achieve it, we jointly deploy virtual machines (i.e., services) into physical machines (i.e., servers) considering traffic demand between virtual machines and the capacity of physical machines and establish wireless links between two transceivers deployed on different racks. To maximize throughput in the WDCN, we formulate a mixed-integer programming (MIP) problem. We also propose two heuristics, named Heuristic for Wireless Link and Service Deployment (HWSD) and Improved Heuristic for Wireless Link and Service Deployment (I-HWSD) to decrease the time it takes to solve the MIP problem with an optimization solver. These two heuristics approach the optimal solution by about 20% with a significant improvement in the time performance and adapt to dynamically changing traffic demands. Unlike top-of-rack (ToR)-to-ToR WDCN, deploying virtual machines makes a significant contribution to the total traffic demand we can complete in WDCN.

Keywords: wireless data center network, 60 GHz, wireless communication, mixed-integer programming, optimization

ÖZ

ORTAK SANAL MAKİNE YERLEŞTİRME VE KABLOSUZ VERİ MERKEZİ TOPOLOJİSİ YÖNETİMİ

Bütün, Beyza
Yüksek Lisans, Bilgisayar Mühendisliği Bölümü
Tez Yöneticisi: Prof. Dr. Ertan Onur

Mayıs 2022 , 54 sayfa

Nesnelerin İnterneti, beşinci nesil vb. gibi gelişen teknolojilerle birlikte üretilen veriler muazzam bir şekilde büyüyor. Bu nedenle, günümüzde Veri Merkezi Ağları (VMAlar), önemli miktarda veriyi depolamak ve işlemek için önemli bir göreve sahiptir ve bu da onları ağın kritik bir bileşeni haline getirir. Çok büyük miktarda trafik talebini karşılamak için kablolu VMA'lar çok sayıda sunucu, güç tüketen anahtarlar ve çok uzun kabloların kullanımına ihtiyaç duyarlar. Kabloların kullanımındaki büyük artış, yüksek kablolama karmaşıklığına ve maliyetine neden olurken, çok sayıda sunucu ve güç tüketen anahtarlar, yüksek düzeyde güç tüketimine neden olur. Bu sorunları gidermek için Kablosuz Veri Merkezi Ağları (Kablosuz VMAlar) ortaya çıkmıştır. Kablosuz VMAlar, dinamik olarak değişen koşullara uyum sağlama yetenekleri sayesinde bant genişliğini verimli bir şekilde kullanır ve kablosuz yapıları sayesinde yüksek maliyet, kablolama karmaşıklığı ve güç tüketimi sorunlarını azaltır. Kablosuz VMAlar ayrıca hotspot problemine umut vaat eden bir çözüm sunar. Kablosuz VMAlar umut vaat eden çözümler sunmanın yanı sıra, güvenilir iletişim kurmak, dinamik olarak değişen trafik taleplerini karşılamak ve veri merkezinde karşılanan trafik miktarını artırmak gibi bazı zorlukları da beraberinde getirir. Bu çalışmada, 60 GHz bandını kullanan dinamik topolojiye sahip bir Kablosuz VMA öneriyoruz ve Kablosuz VMA'daki dinamik trafik taleplerini karşılamayı ve verimi en üst düzeye çıkarmayı amaçlıyoruz. Bunu başarmak için, sanal makineler (yani uygulamalar) arasındaki trafik talebini ve fiziksel makinelerin (yani sunucuların) kapasitesini göz önünde bulundurarak eş zamanlı bir şekilde sanal makineleri fiziksel makinelere verimli bir şekilde dağıtıp farklı raflarda konuşlandırılmış iki alıcı-verici arasında kablosuz bağlantılar kuruyoruz. Kablosuz VMA'daki verimi en üst düzeye çıkarmak için bir karma tamsayılı programlama (KTP) problemi formüle ediyoruz. Ayrıca, KTP problemini bir optimizasyon çözücü ile çözmek için gerekli olan süreyi azaltmak için Kablosuz Bağlantı ve Hizmet Dağıtımı (KBHD) ve Geliştirilmiş Kablosuz Bağlantı ve Hizmet Dağıtımı (G-KBHD) adlı iki buluşsal yöntem öneriyoruz. Bu iki buluşsal yöntem, önerilen problemi çözmek için gerekli süreyi önemli ölçüde azaltırken bir optimizasyon çözücü tarafından elde edilen optimal çözüme yaklaşık %20 oranında yaklaşır ve VMA'da dinamik olarak değişen trafik taleplerine uyum sağlar. Daha önce yapılmış ToR-to-ToR Kablosuz VMA çalışmasından farklı olarak, tanıtılan çalışma, sanal makinelerin fiziksel makinelere yerleştirilmesini de önererek, Kablosuz VMAlarda karşılanan trafik talebi oranını artırmaya önemli bir katkı sağlar.

Anahtar Kelimeler: kablosuz veri merkezi ağı, 60 GHz, kablosuz iletişim, karma tamsayılı programlama, optimizasyon

To my family, close friends, and playmates Luna and Duru

ACKNOWLEDGMENTS

I would like to express my gratitude to Prof. Dr. Ertan Onur for his support and guidance throughout my bachelor's and master's. It was an honor and a great opportunity for me to work with him.

I also want to thank my parents, sister, and brother for their unconditional support and love. When I fell into despair, feeling them by my side always encouraged me. My most special thanks go to my dear Eda, who beautifies every moment of my life and gives me hope and power.

Finally, this work is supported by Vodafone Telecommunication Inc., Turkey within the framework of the 5G and Beyond Joint Graduate Support Programme coordinated by Information and Communication Technologies Authority. I would like to thank all the authorities who led this program and supported students in their academic journey.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

LIST OF ABBREVIATIONS

5G	Fifth Generation
BPSK	Binary Phase Shift Keying
DCN	Data Center Network
FSO	Free-space Optical Communication
HWSD	Heuristic for Wireless Link and Service Deployment
I-HWSD	Improved Heuristic for Wireless Link and Service Deployment
LoS	Line-of-Sight
MIP	Mixed-Integer Programming
PM	Physical Machine
SC	Single Channel
TCP	Transmission Control Protocol
ToR	Top-of-Rack
VM	Virtual Machine
WDCN	Wireless Data Center Network
WSD	Wireless Link and Service Deployment

CHAPTER 1

INTRODUCTION

Data Center Networks (DCNs) are an essential component of networks for storing and processing data. They are responsible for the smooth running of information storage, computing, web services, and telecommunications activities. They become a more crucial component of networks as the data produced grows enormously. Today, a massive amount of data are stored and processed. Hence, it is a critical requirement of DCNs continues to operate without any problems. That's why many researchers are working on preventing performance drops or interruptions.

Data is stored on physical machines and processed by virtual machines running on physical machines. Each physical machine is deployed in the metallic racks carrying devices responsible for storage, computing, and networking operations [1]. These devices connect to a switch located at the top of each rack, and in conventional data centers, each switch connects to a higher-level switch [1]. Cables are required for these connections. Racks are generally arranged in rows, and there is an aisle between each row. This structure helps to perform intra-rack, inter-rack, and inter-data center communications. Figure 1.1 shows the architecture of a conventional tree-based DCN. Top-of-rack (ToR) switches are responsible for intra-rack and inter-rack communications, and each of them connects to a higher-level switch. Gateway router plays a critical role in inter-data center communications.

With technological advancements such as artificial intelligence (AI), fifth-generation (5G), the Internet of Things, machine-to-machine communication, and cloud services, the data generated in DCNs grows significantly [2]. This growth increases the importance of DCNs. To handle the massive amount of data generated by applications, the number of devices such as servers and switches deployed in DCNs needs to be in-

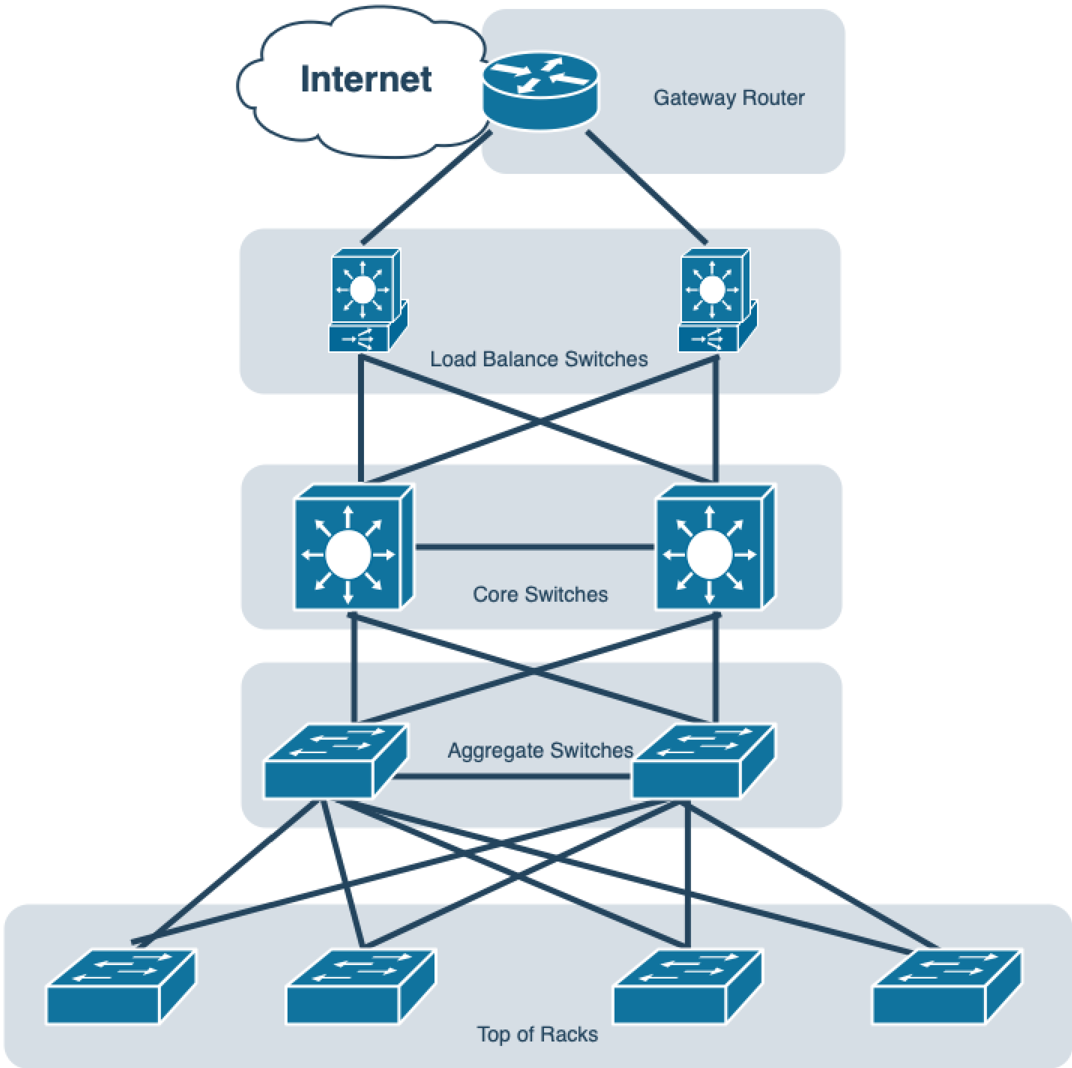

Figure 1.1: Conventional tree-based DCN architecture.

creased. Adding more network resources needs costly re-deployment processes, and it makes DCNs more complex, which increases the cost for further re-deployment operations and creates obstacles to the scalability and reconfigurability of DCNs [3]. Network entities connect via uniform and fixed capacity cables. Hence, increasing the number of network devices requires deploying a large length of wires, which is another obstacle to the scalability of the network and the reason for inefficient space utilization. Besides, deploying large bundles of wires creates difficulties in the cooling process because wires become obstacles to the flow of chilled air, which is the reason behind inefficient cooling [3], [4], [5]. Inefficient cooling exacerbates energy consumption, as with conventional DCNs using large numbers of power-hungry switches.

With the enormous increase in data generated, the problem of oversubscription is becoming more frequent, especially with conventional DCNs such as fat-tree [6] which is the most popular today. In this topology, there are three layers named access, aggregate, and core, and servers arranged in racks are connected through a hierarchy of the switches of these layers [4]. At the core layer, a few switches serve as root nodes [3], [7]. Hence, during transmissions of ultra-dense traffic via fixed capacity cables, it is inevitable to face a bottleneck problem at the core layer, which causes oversubscription in transmissions between servers deployed in different racks [3], [7]. The oversubscription problem is one of the crucial reasons for the decrease in overall throughput [3].

With the enormous increase in data volume, all these problems have become an obstacle to scalable, reconfigurable, manageable, power-efficient, and cost-efficient DCN. Wireless DCNs (WDCNs) have emerged to address the challenges of wired DCNs. The emergence of WDCNs eliminates the need for switches that tend to consume high power and make errors and the need for deploying large bundles of wires [3], [8]. Hence, it offers a less complex, scalable, reconfigurable and energy, cost, and bandwidth-efficient DCN thanks to its cooperation with wireless communication technologies [3], [8]. Besides point-to-point communication, wireless technologies offer a broadcast communication medium that is suitable to be utilized for different use cases [9]. With the help of established wireless links between different racks, DCNs become more flexible, so WDCNs are efficient to handle dense and unbalanced traffic demands. Even though WDCNs have significant benefits over wired DCNs and provide promising solutions for most of their problems, they have many challenges. Because wireless signals are intolerant to obstructions, it is hard to satisfy reliable communication [3], [10]. It is another challenge to handle dense, unbalanced, and dynamic traffic demands with dynamically established wireless links.

1.1 Contributions

We aim to design a WDCN using an internationally available 60 GHz unlicensed millimeter-wave band and maximize its throughput. To achieve high throughput, we jointly deploy virtual machines into physical machines and establish wireless links between transceivers deployed on the top of racks. While deploying virtual machines into physical machines, we consider traffic demands between virtual machines and the capacity of physical machines in terms of memory and CPU. Deploying virtual machines into physical machines considering average traffic demands between them makes a significant contribution to increasing the throughput we can achieve. We formulate a mixed-integer programming (MIP) problem to maximize throughput in the WDCN. Because the time it takes to solve this problem with an optimization solver is significantly too long, we introduce two heuristics named Heuristic for Wireless Link and Service Deployment (HWSD) and Improved Heuristic for Wireless Link and Service Deployment (I-HWSD). We compare solutions and time performance of the heuristics and the optimization solver. We observe the impact of distance and traffic change in the WDCN on the total traffic met. We also test the adaptability of these heuristics to the dynamic traffic demands and traffic distribution changes. Finally, we compare our findings with a similar study proposing top-of-rack (ToR)-to-ToR WDCN [11] in terms of the ratio of satisfied traffic demands.

1.2 Outline

The rest of the paper is organized as follows:

- In Chapter 2, we present advances and challenges of 60 GHz bands. We also present the research done before on hybrid (wired and wireless) and wireless DCNs. Finally, we highlight the novelty of our study.
- In Chapter 3, we design a WDCN topology considering 60 GHz bands, and we formulate an MIP problem named Wireless Link and Service Deployment (WSD) to find optimal virtual machine deployment and wireless link establishment scenarios.

- We propose two heuristics named Heuristic for Wireless Link and Service Deployment (HWSD) and Improved-HWSD (I-HWSD) to decrease the time it takes to solve the WSD problem with an optimization solver in Chapter 4.
- Analysis of our optimization problem and heuristics are presented in Chapter 5.
- We conclude the paper in Chapter 6.

CHAPTER 2

BACKGROUND AND RELATED WORK

In this section, we talk about the advances and challenges of 60 GHz technology and review the literature on the designs of hybrid and pure wireless data center networks. We then explain the novelty of our study.

2.1 60 GHz Technology

The proposed WDCN uses unlicensed 60 GHz band which is introduced first by Ramachandran et al. [12] for data centers. Today, the 60 GHz band also known as the millimeter-wave band is one of the most known candidates to host wireless applications utilizing multi-gigabits [13]. ECMA387 [14], IEEE802.11.ad [15], and Wireless HD [16] are some of the introduced standardizations of the 60 GHz band. 60 GHz technologies operate in the internationally available interval between 57 GHz to 66 GHz, and these technologies offer a huge bandwidth. It is possible to launch products worldwide thanks to the unlicensed and internationally available bands, and the available spectrum enables the creation of multiple links of Gbps rates. [12]. Table 2.1 shows the available frequency ranges for different regions in the IEEE802.11ad [15], [17].

The bandwidth increases as we move to higher frequencies in the spectrum. Hence, 60 GHz provides higher throughput if we compare it with lower frequencies by supporting a higher data rate, and it is more suitable for indoor environments because it has a short communication range [3]. Short communication range addresses concerns about security issues in data centers and increases spatial reuse of the spectrum [18]. Using 60 GHz bands in indoor environments helps us to decrease the attenuation

Table 2.1: Available frequency ranges in different regions.

Region	Frequency range (GHz)
Europe	57.00-66.00
United States	57.05-64.00
China	59.00-64.00
Australia	59.40-62.90

caused by rain, atmospheric absorption, etc.

Although employing the millimeter-wave band in data centers increases efficiency, it has some notable challenges to deal with. Using millimeter-wave requires coping with high attenuation. Moreover, the high noise power is more apparent because of the usage of larger bandwidths, and the path loss is high due to the usage of high carrier frequencies [19]. The millimeter-wave band is also vulnerable to obstructions such as racks and people moving between racks. Due to its vulnerability to obstructions and highly directional nature, it becomes a need to establish line-of-sight links between transceivers.

2.2 Related Work

For conventional data centers, researchers introduce topologies such as fat-tree [6], DCell [20], BCube [21], VL2 [22], and FiConn [23]. Due to the problems of conventional data centers such as high cabling complexity and cost, usage of power-hungry switches, inefficient cooling, space and bandwidth utilization, and high probability of encountering an oversubscription problem, researchers start researching wireless data center architectures. How wireless data center architectures solve most of these problems is explained in [8], [24] with their drawbacks and challenges.

Higher throughput provided by 60 GHz and its suitability to indoor scenarios motivate researchers to conduct studies on the usability of 60 GHz in data centers. For this purpose, Zaaimia et al. [1] measure 60 GHz channels in a real data center propagation environment to design reliable and robust wireless data center links. Path loss and delay spread are modeled for the use cases such as cross-aisle ToR where line-of-sight (LoS) links satisfy communication between different ToRs located in distinct rack rows separated by an aisle, neighbor ToR where LoS links are established between neighbor racks, Cayley data center [25] which is a completely wireless design, and 3-D Beamforming [26] establishing reflected LoS between ToRs utilizing a ceiling reflector in [27]. Authors of [27] state that the results show great differences between wireless data center environments and regular propagation environments in terms of path loss and delay spread. To handle the problems of fixed cable capacities and huge wires in conventional data centers, Vardhan et al. [18] propose to employ 60 GHz millimeter-wave wireless links in DCNs and state that multi-gigabit data rate and point-to-point wireless links in the millimeter-wave band makes 60 GHz technologies suitable for data centers.

Another candidate technology for wireless data centers is free-space optical (FSO). Hamze et al. [28] compare these two candidate technologies, 60 GHz and FSO, in terms of interference, alignment of the links, and link lengths. Authors of [28] also classify wired and wireless data center networks in terms of communication technologies used in the data center. There are several studies on hybrid wireless DCNs and pure wireless DCNs. In hybrid wireless DCNs, wireless links are used to augment wired DCNs, while in pure DCNs, all the communications are satisfied via wireless links [28]. The authors of [29] review literature about hybrid data centers and completely wireless data centers.

2.2.1 Hybrid Wireless DCNs

To solve problems of conventional DCNs, such as oversubscription and lacking reconfigurability and manageability, researchers come up with hybrid wireless DCNs. Halperin et al. [30] present wireless flyways added to wired DCNs to relieve hotspots. The authors of [30] state that even if there are some concerns about interference and reliability, 60 GHz technologies are suitable for data center environments. They also show that deploying flyways makes network-limited applications with predictable traffic demands speed up by about 45%. We encounter a bottleneck problem in a few ToR switches, which communicate with only a few other switches most of the time

[31]. To eliminate the hotspot problem in these ToRs, the authors of [31] present on-demand flyways. They state that even a few flyways deployed when and where needed improves performance by over 50%. Luo et al. [32] propose to augment wired DCNs with 60 GHz wireless links to alleviate traffic congestion in DCNs by introducing an energy-efficient offline deployment scheme.

Today, data centers need to process and store a massive amount of data. With the growth in data to be processed and stored, utilization and bandwidth needs to grow. Increasing bandwidth with traditional methods is complex and costly in terms of time and resources [33]. To augment bandwidth, authors of [33] introduce a new wireless solution called 3D beamforming that aims to establish in-direct and one-hop 60 GHz line-of-sight links between racks with the help of a data center ceiling reflecting signals. 3D beamforming helps to decrease interference and increase bandwidth. To increase bandwidth and flexibility and decrease cabling complexity, authors of [34] present FireFly that establishes wireless links between transceivers deployed on ToRs with the help of a data center ceiling mirror without removing ToR switches by using free-space optics as a key enabler.

Instead of deploying wireless radios on ToRs, authors of [35] present a novel hybrid DCN architecture named Diamond and a concept named Ring Reflection Space to allow wide deployment of 60 GHz wireless radios at servers and pave the way for an increase in the number of concurrent wireless transmission between servers.

2.2.2 Pure Wireless DCNs

The first constructive proposal for pure wireless DCNs is Cayley data center topology which is densely connected and inspired by Cayley graphs [36], [25]. Authors of [25] state that in Cayley data centers showing strong fault tolerance, servers are arranged in cylindrical racks, and signals don't interfere with each other. Li et al. [37] proposes a flexible pure wireless DCN topology named Spherical Mesh utilizing 60 GHz and the geometric properties of the sphere to alleviate the link blockage problem and decrease network diameter. Authors of [37] state that Spherical Mesh topology reduces maximal routing path by about 20%, and balances the workload of the network. Another proposal that aims to leverage the link blockage problem is a novel pure wireless DCN structure named Graphite [38]. Authors of [38] remark that Graphite outperforms existing topologies like Flyways [30] and 3D Beamforming [33] in terms of connectivity of data center network.

To increase concurrency in DCNs, authors of [39] introduce an algorithm mapping fat-tree data center topology into a hexagonal arrangement for a completely wireless data center networks using 60 GHz. Cao et al. [40] proposes a topology optimization problem for WDCNs with three objectives, which are coverage, propagation intensity, and interference intensity, and with a constraint of connectivity. Cheng et al. [41] evaluate characterization of 300 GHz channel for wireless rack-to-rack data center communications and the effect of obstructions in wireless links.

Establishing ToR-to-Tor wireless links is one of the methods to augment wired DCNs. Mamun et al. [11] evaluate the performance of a ToR-to-ToR WDCN in terms of power consumption and network-level data rate. Authors of [11] remark that the proposed 12-channel 60 GHz WDCN consumes less power than conventional wired DCNs and obtain comparable data rates with these DCNs for typical query-based applications. Authors of [4] introduce a 60 GHz wireless server-to-server DCN (S2S-WiDCN) architecture leveraging line-of-sight links between servers to establish direct communication and propose a novel algorithm to overcome obstructions that may block line-of-sight links. S2S-WiDCN [4] consumes less power by five to seventeen times and obtains comparable flow completion duration and throughput compared with the conventional fat-tree [6] DCN and ToR-to-ToR WDCN [11].

2.3 Novelty

In the scope of pure WDCNs, our proposed study aims to design a WDCN and maximize throughput in the WDCN by jointly deploying virtual machines into physical machines and establishing wireless links between transceivers located on top of the racks. Although we establish ToR-to-ToR communications like in [11], different from other studies, we also deploy virtual machines into physical machines considering traffic and source demand of them to distribute the traffic demands of applications evenly in the network, which helps to increase the total traffic met and makes our study a novel proposal. Our proposed single-channel WDCN outperforms the single-channel ToR-to-ToR WDCN [11] in terms of the ratio of completed demands. Although our proposed single-channel WDCN satisfies demands with more ratio, we do not take interference into account while establishing wireless links like they do.

CHAPTER 3

WIRELESS LINK AND SERVICE DEPLOYMENT (WSD) PROBLEM

Wireless Link and Service Deployment (WSD) problem aims to maximize throughput in WDCN. For this purpose, we first design a WDCN using unlicensed 60 GHz bands and satisfying wireless communication requirements. We then formulate the WSD problem to maximize throughput in the designed WDCN by jointly deploying virtual machines into physical machines and establishing wireless links between transceivers deployed on different racks.

3.1 Wireless Data Center Network Topology

Figure 3.1 shows the layout of the designed WDCN with 15 racks. A transceiver is deployed on the top of each rack and the distance of the transceivers from the ground is equal. In the WDCN, racks are arranged in rows. The location of each transceiver is represented with *x*, *y*, and *z*-coordinates. *z*-axis represents the distance of the transceiver from the ground, *x*-axis represents in a row which rack the transceiver are deployed on, and *y*-axis represents the row containing this rack. The representation of the origin rack in terms of *x, y,* and *z*-coordinates is (0, 0, *h*), where *h* is the height of the deployed transceiver. The values of *x* and *y*-axes increase in the axes' directions. For instance, the location of the transceiver deployed on top of the rack *v* represented with (x_v, y_v, z_v) is (3, 0, *h*) while the location of the transceiver deployed on top of the rack *u* represented with (x_u, y_u, z_u) is (4, 1, *h*). In the designed WDCN, there is a 1-meter aisle between rows. Moreover, we assume that the rack dimensions are 42U in height, where 1U is equal to 44.45mm, 600mm in width, and 1070mm in depth [3], [45]. Hence, the distance between two transceivers deployed on the top of

Table 3.1: List of symbols for topology design.

Symbols	Implication	Unit	Default Val.	Ref.
P_t	The transmit power	mW	100	-
P_r	The received signal power	mW	-	-
G_r	Antenna gain of the receiver	dBi	9	[4]
G_t	Antenna gain of the transmitter	dBi	9	[4]
c	The speed of light	m/s	3×10^8	-
f	The carrier frequency	GHz	60.48	[42]
T	The temperature	K	290	-
N_{figure}	Noise figure	dB	6	[43]
N_{floor}	Noise floor	dB	-	-
N_p	Noise power	W	-	-
SNR or $\frac{S}{N}$	Signal-to-noise ratio	-	-	-
$\frac{E_b}{N_0}$	Energy per bit to noise power spectral density ratio	BER	3×10^{-7}	[11]
R_s	The receiver sensitivity	dB	-	-
B	The bandwidth of the channel	GHz	1.632	[44]
R	The bit rate	bps	-	-

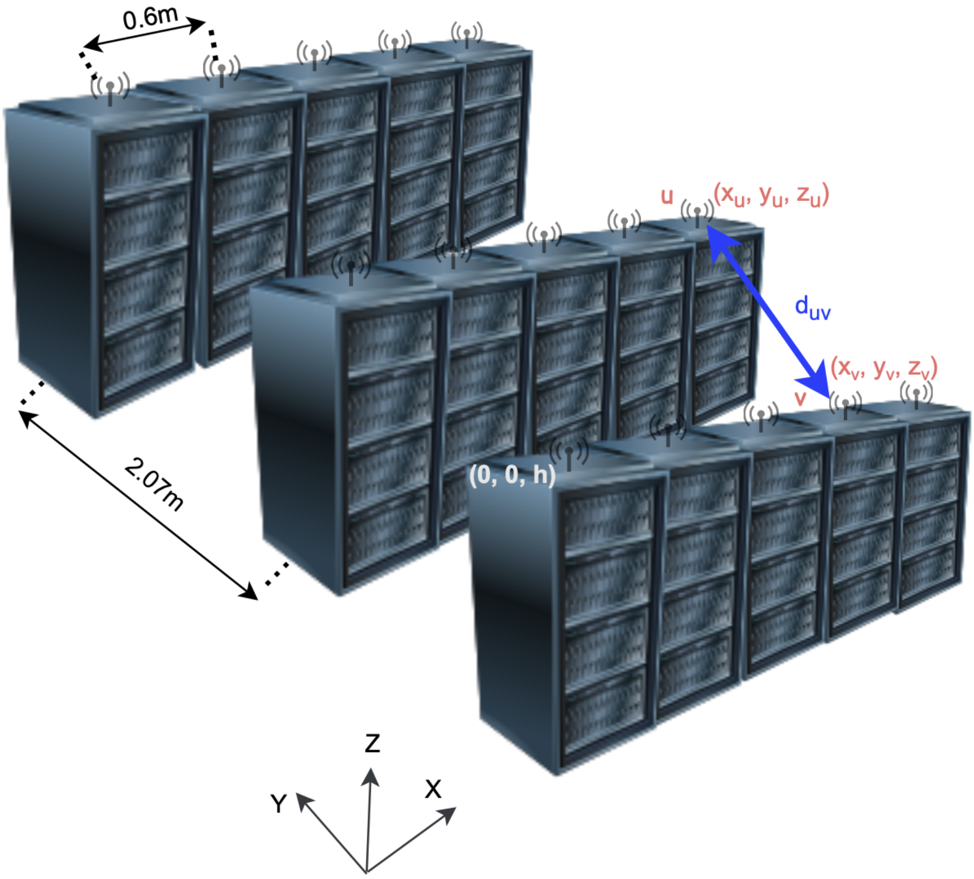

Figure 3.1: Layout of the designed data center with 15 racks.

the side-by-side racks is 0.6 meters and between two transceivers deployed on the top of the back-to-back racks is 2.07 meters.

To find the maximum possible distance between two transceivers deployed on the top of different racks to establish a wireless link is a need to decide how many racks we could deploy in the DC considering 60 GHz channel requirements and to design it. In other words, we need to find the maximum distance that satisfies $P_r \geq R_s$, where P_r represents received signal power and R_s represents receiver sensitivity. We can define receiver sensitivity as the weakest signal that can be identified and processed by the receiver [3] and is calculated as follows:

$$P_r = P_t G_r G_t \left(\frac{c}{4\pi d_{uv} f} \right)^2, \quad (3.1.1)$$

where P_t stands for transmit power taken as 100 milliwatts [3], G_t and G_r are antenna gain for the transceiver and the receiver respectively, and their value is equal to 9 dBi [3], [4]. Moreover, c is the speed of light which is $3 \times 10^8 m/s$, f is the carrier frequency of the channel considered to be 60.48 GHz [3], and d_{uv} is the distance between two transceivers u and v in meters. We can calculate the distance d_{uv} as

$$d_{uv} = \sqrt{(0.6(x_u - x_v))^2 + (2.07(y_u - y_v))^2 + (z_u - z_v)^2} \quad \forall u, v \in V$$

where (x_u, y_u, z_u) and (x_v, y_v, z_v) are the locations of transceivers deployed on top of the racks u and v, respectively, and z_u and z_v, which are the distances of these transceivers from the ground, are equal. Moreover, *V* is the global set of the racks. We can calculate the receiver sensitivity R_s as

$$R_{sdB} = N_{floor_{dB}} + SNR_{dB}, \tag{3.1.2}$$

$$N_{floor_{dB}} = N_{p_{dB}} + N_{figure_{dB}}, \tag{3.1.3}$$

$$N_p = kTB, \tag{3.1.4}$$

$$SNR = \frac{E_b R}{N_0 B} \text{ or } SNR = \frac{P_r}{kTB}, \tag{3.1.5}$$

where the N_{floor} and N_p represents the noise floor and the noise power in decibel, respectively, and N_{figure} in (3.1.3) is the noise figure considered to be 6 dB [3], [43]. Furthermore, in (3.1.4), *k* is the Boltzmann's constant which is equal to $1.38 \times 10^{-23}\ Joules/Kelvin$, T is the room temperature which is assumed to be 290K, and B is the bandwidth of the channel, which is 1.632 GHz [3], [44]. Moreover, in (3.1.5), SNR stands for signal-to-noise ratio, where $\frac{E_b}{N_0}$ is energy per bit to noise power spectral density ratio, and its value is 3×10^{-7} BER which is taken as approximately 11 dB on the linear scale, considering BPSK is used as a modulation scheme [3], [11], and R stands for bit rate in bps. As a wireless communication requirement, the bit rate has to be less than or equal to channel capacity, which is formulated as $R \leq C$, where C is the channel capacity and R is the bit rate. We can calculate channel capacity as

$$C = B \log_2 \left(1 + \frac{S}{N}\right). \tag{3.1.6}$$

Using the formulations from (3.1.1) to (3.1.6), we find the maximum possible distance d_{uv} satisfying $P_r \geq R_s$ between two transceivers u and v. To be able to establish a

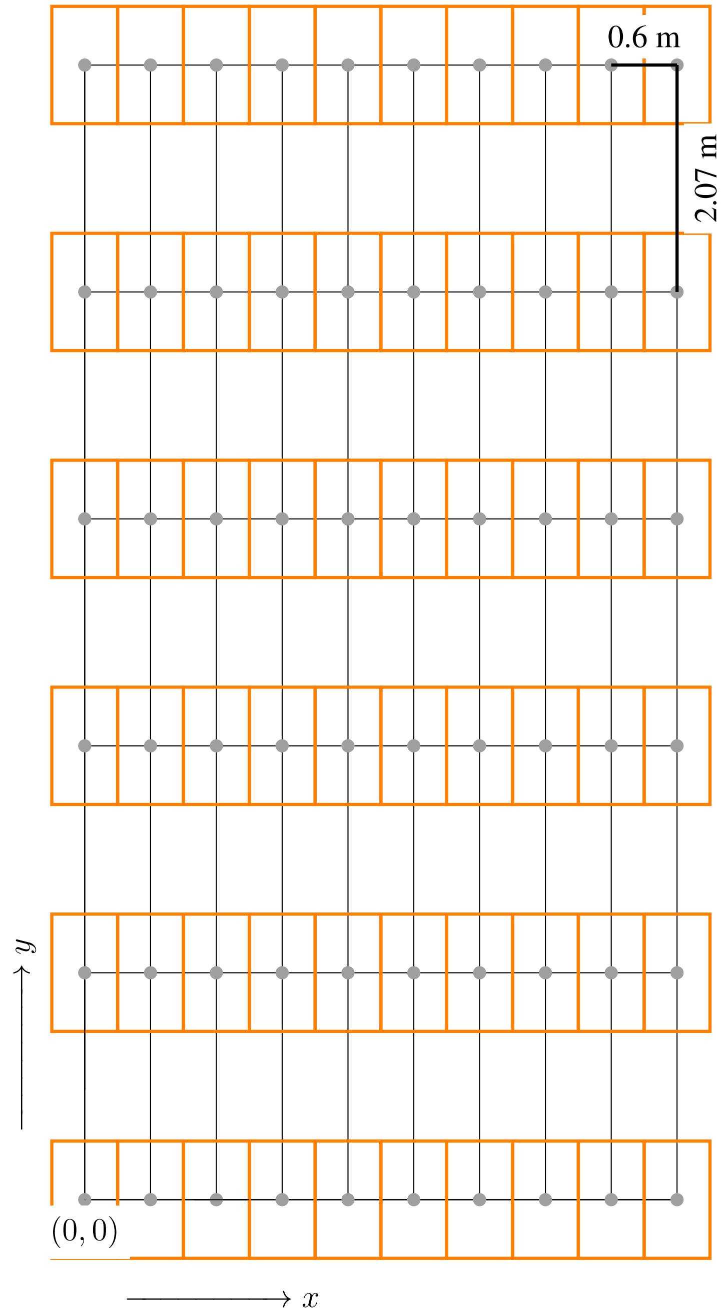

Figure 3.2: Top view of the designed data center

wireless link between any two transceivers deployed on top of the racks u and v in *V*, the distance d_{uv} between these transceivers has to satisfy $d_{uv} \leq 12$, where d_{uv} in meters. In other words, the distance between the farthest transceivers cannot exceed 12 meters. Figure 3.2 shows the grid representing the top view of the deployed data center (DC). The grid is on the *x* and *y*-coordinates. In the grid, orange rectangles represent the top view of racks, while each dot represents the transceiver deployed on top of the related rack. In the light of the found maximum possible distance to establish a wireless link between two transceivers, the designed data center has 6 rows, and each contains 10 racks, as you can see in Figure 3.2.

3.2 Wireless Link and Service Deployment (WSD) Problem Formulation

After designing a WDCN, we formulate the WSD to maximize the throughput in this designed WDCN. WSD problem aims to maximize total satisfied traffic demand in the data center. For this purpose, WSD jointly deploys virtual machines into physical machines and establishes wireless links between transceivers deployed on top of different racks. While deploying virtual machines into physical machines, it considers traffic and source demand between virtual machines and the capacity of physical machines in terms of RAM, cache memory, and CPU. While establishing wireless links between transceivers, it considers the channel capacity between these transceivers. We classify the WSD problem as maximizing the satisfied inter-rack traffic demand only called pure-WSD and maximizing both the intra-rack and inter-rack satisfied traffic demand called hybrid-WSD. We then compare the solutions of pure-WSD and hybrid-WSD problems.

While deploying virtual machines into physical machines, we have to consider that the resource demand of virtual machines cannot exceed the capacity of the physical machines into which they are deployed in terms of RAM, cache memory, and CPU. We introduce three constraints to satisfy it:

$$\sum_{s \in S} \xi_{sp} \alpha_s \leq m_p \qquad \forall p \in P, \tag{3.2.1}$$

$$\sum_{s \in S} \xi_{sp} \beta_s \leq h_p \qquad \forall p \in P, \tag{3.2.2}$$

Table 3.2: List of symbols for problem formulation.

Symbols	Implication	Unit	Default Val
V	The global set of the racks	-	-
P	The global set of the physical machines	-	-
S	The global set of the virtual machines	-	-
m_p	The total memory of the physical machine p	GB	256
h_p	Maximum allowed cache size of the physical machine p	MB	20
c_p	Maximum allowed CPU usage of the physical machine p	-	100%
α_s	The memory usage of the virtual machine s	GB	[4, 64]
β_s	The cache memory usage of the virtual machine s	MB	[1, 12]
γ_s	The CPU usage of the virtual machine s	-	[12.5, 50]%
C_{uv}	The capacity of the channel between the transceivers on top of the racks u and v	bps	-
r_{st}	The average traffic demand from the virtual machine s to the virtual machine t	Gbps	-
d_{uv}	The distance between the transceivers deployed on the racks u and v	meter	-
δ_{pu}	The binary value indicating whether the physical machine p is deployed in the rack u	-	-
ξ_{sp}	The binary value indicating whether the virtual machine s is embedded into the physical machine p	-	-
X_{uv}	The binary value indicating whether there is a wireless link between two transceivers deployed on top of the racks u and v	-	-

$$\sum_{s \in S} \xi_{sp} \gamma_s \leq c_p \qquad \forall p \in P, \tag{3.2.3}$$

where ξ_{sp} is an optimization decision variable indicating whether the virtual machine s is deployed into the physical machine p. If the virtual machine s is embedded into the physical machine p, then ξ_{sp}=1, otherwise ξ_{sp}=0. Figure 3.3 shows the deployment of the virtual machines into the physical machines located in the racks *u* and *v*, and p_1 and r_1 are physical machines while s_1, s_2, and t_1 are virtual machines. The virtual machines s_1 and s_2 are deployed into the physical machine p_1 located in the rack u as we can see in Figure 3.3. Hence, $\xi_{s_1 p_1}$ = 1 and $\xi_{s_2 p_1}$ = 1 while $\xi_{s_1 r_1}$ = 0 and $\xi_{s_2 r_1}$ = 0. Moreover, in the constraints, m_p, h_p, and c_p are the total memory in GB, maximum allowed cache size to be used in MB, and maximum allowed CPU usage in percent of the physical machine p while α_s, β_s, and γ_s are the memory usage in GB, cache memory usage in MB, and CPU usage in percent of the virtual machine s, respectively. These values are known in advance. Furthermore, P and S are the global set of the physical machines and virtual machines, respectively.

Besides, a virtual machine cannot be deployed into more than one physical machine:

$$\sum_{p \in P} \xi_{sp} = 1 \qquad \forall s \in S. \tag{3.2.4}$$

To contribute to an increase in throughput, we need to embed virtual machines into physical machines in a way that distributes the traffic demand of virtual machines evenly in the WDCN as much as possible. For this purpose, we introduce a constraint satisfying that if there is a wireless link between transceivers deployed on top of the racks u and v, the total traffic demand of virtual machines embedded into the physical machines located in the rack u to the virtual machines embedded into the physical machines located in the rack v cannot exceed the capacity of the wireless link established from u to v, and vice versa. For instance, as we can see in Figure 3.3, the total traffic demand from the rack u to v cannot exceed the channel capacity C_{uv}. The related constraint is

$$\sum_{s,t \in S} \sum_{p,r \in P} \delta_{pu} \xi_{sp} X_{uv} \delta_{rv} \xi_{tr} r_{st} \leq C_{uv} \qquad \forall u, v \in V, \tag{3.2.5}$$

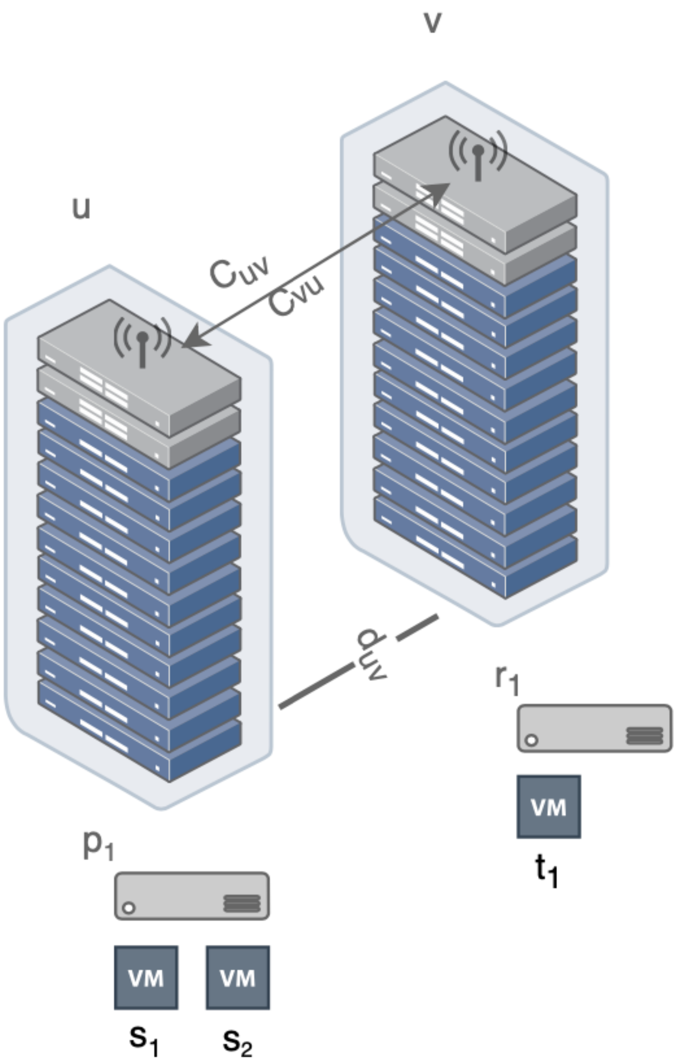

Figure 3.3: Virtual machine deployment in the WDCN.

where X_{uv} is an optimization parameter indicating whether there is a wireless link between transceivers deployed on top of the racks u and v. If there is a wireless link between them, then $X_{uv} = 1$, otherwise $X_{uv} = 0$. Figure 3.4 shows a WDCN model with 10 racks and the established wireless links between transceivers deployed on top of the rack pairs u_2-u_3 and u_5-v_5. Hence, $X_{u_2u_3} = 1$, $X_{u_3u_2} = 1$, $X_{u_5v_5} = 1$, and $X_{v_5u_5} = 1$. Moreover, δ_{pu} indicates whether the physical machine p is located in the rack u. If it is, then $\delta_{pu} = 1$, otherwise $\delta_{pu} = 0$. The value of δ_{pu} is known in advance. Furthermore, r_{st} is the average traffic demand in bps from the virtual machine s to t, and C_{uv} is the channel capacity of the wireless link established from the transceiver

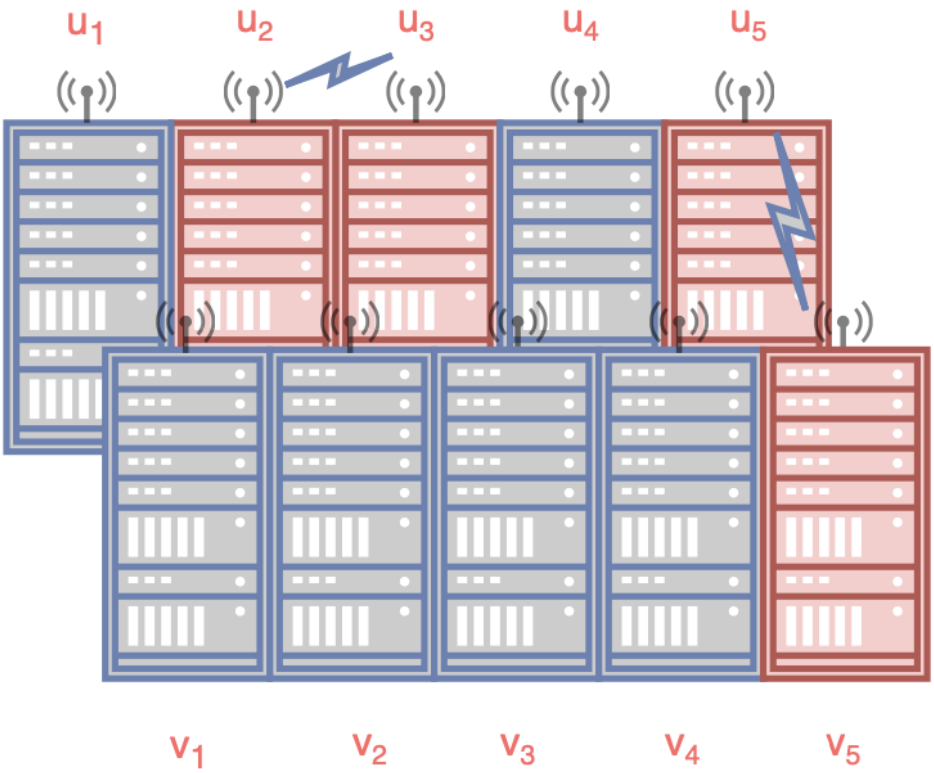

Figure 3.4: Wireless links establishment in the WDCN.

deployed on top of the rack u to the transceiver deployed on top of the rack v.

If there is a wireless link between transceivers, there will be two-way communication, and a transceiver can only transmit to one transceiver at the same time. We introduce two constraints to meet these two requirements:

$$\sum_{v \in V} X_{uv} \leq 1 \qquad \forall u \in V, \qquad (3.2.6)$$

$$X_{uv} - X_{vu} = 0 \qquad \forall u, v \in V. \qquad (3.2.7)$$

The received power has to be greater than or equal to the receiver sensitivity to process a signal. Hence, $P_r \geq R_s$ has to be satisfied to establish a wireless link between two transceivers. This equality and the equations of (3.1.2) to (3.1.5) help us to generate a relation between the received power P_r and the channel capacity C_{uv} of the wireless link established between two transceivers deployed on top of the racks u and v:

$$N_{p_{dB}} = -204_{dB/Hz} + 10\log_{10}(B),$$

$$SNR_{dB} = 11_{dB} + 10log_{10}(R) - 10log_{10}(B),$$

$$R_{sdB} = -204_{dB} + 6_{dB} + 11_{dB} + 10log_{10}(R),$$

$$P_{rdB} \geq -187_{dB} + 10\log_{10}(C_{uv}), \tag{3.2.8}$$

where the bit rate R of the wireless link between two transceivers deployed on top of any two racks is assumed to be equal to the channel capacity between these transceivers considering the requirement $R \leq C$.

The constraints of (3.2.1) to (3.2.8) are valid for both pure and hybrid-WSD problems. In the hybrid-WSD problem, we consider intra and inter-rack communication, therefore, we introduce a constraint setting the value of X_{uv} as 1 when u and v are equal. It helps us to maximize both inter and intra-rack traffic demand satisfied in the hybrid-WSD problem. The constraint is

$$X_{uv} = 1 \qquad \forall u, v \in V \text{ such that } u = v. \tag{3.2.9}$$

With the constraints of (3.2.1) to (3.2.9), we aim to embed the virtual machines into the physical machines considering the source and traffic demand of the virtual machines and the capacity of the physical machines and establish wireless links between transceivers considering the traffic demand of virtual machines and the channel capacity.

The objective is to maximize the total traffic demand satisfied in the WDCN with virtual machine deployment and wireless link establishment decisions:

$$\max_{X_{uv},\xi_{pu},\xi_{rv}} \sum_{s,t\in S}\sum_{p,r\in P}\sum_{u,v\in V} \delta_{pu}\xi_{sp}X_{uv}\delta_{rv}\xi_{tr}r_{st}. \tag{3.2.10}$$

For completeness, the mathematical formulation of hybrid and pure-WSD problems is as follows:

$$\max_{X_{uv},\xi_{pu},\xi_{rv}} \sum_{s,t\in S}\sum_{p,r\in P}\sum_{u,v\in V} \delta_{pu}\xi_{sp}X_{uv}\delta_{rv}\xi_{tr}r_{st}. \tag{3.2.11}$$

s.t.

$$\sum_{s\in S}\xi_{sp}\alpha_s \leq m_p \qquad \forall p \in P,$$

$$\sum_{s\in S}\xi_{sp}\beta_s \leq h_p \qquad \forall p \in P,$$

$$\sum_{s\in S}\xi_{sp}\gamma_s \leq c_p \qquad \forall p \in P,$$

$$\sum_{p\in P}\xi_{sp} = 1 \qquad \forall s \in S,$$

$$\sum_{s,t\in S}\sum_{p,r\in P}\delta_{pu}\xi_{sp}X_{uv}\delta_{rv}\xi_{tr}r_{st} \leq C_{uv} \qquad \forall u, v \in V,$$

$$\sum_{v\in V}X_{uv} \leq 1 \qquad \forall u \in V,$$

$$X_{uv} - X_{vu} = 0 \qquad \forall u, v \in V,$$

$$P_{rdB} \geq -187_{dB} + 10\log_{10}(C_{uv}),$$

If the problem is hybrid-WSD problem, then we have also:

$X_{uv} = 1 \qquad \forall u, v \in V$ such that $u = v$.

Except for the specified constraint, the formulation is same for the hybrid and pure-WSD problems.

CHAPTER 4

METHODOLOGY

Due to the high complexity of the problem, it takes a long time to solve the WSD problem with an optimization solver. Hence, we introduce two heuristics named Heuristic for Wireless Link and Service Deployment (HWSD) and Improved Heuristic for Wireless Link and Service Deployment (I-HWSD). HWSD aims to decrease the time it takes to solve the WSD problem and approximate the optimal solution found by a MIP solver, while the goal of I-HWSD is to increase the total traffic demand obtained by HWSD in a reasonable time.

4.1 Heuristic for Wireless Link and Service Deployment (HWSD)

Heuristic for Wireless Link and Service Deployment (HWSD) aims to obtain approximately optimal throughput in the designed WDCN by proposing a virtual machine embedding and wireless link establishment scenario. If the WSD problem is pure-WSD, the goal is to maximize satisfied traffic demand between racks via wireless links, while if the WSD problem is hybrid-WSD, to maximize both inter and intra-rack traffic demand satisfied is the aim. For this purpose, it establishes wireless links between transceivers deployed on top of the closest racks and embeds virtual machines into the physical machines located in these racks considering the channel capacity between transceivers, the source, and traffic demand of the virtual machines, and the capacity of the physical machines.

As described in Algorithm 1, HWSD takes the optimization model containing the decision variables as an input. We first get the decision variables X_{uv} and ξ_{sp} and set the value of the heuristic solution as 0 for a fresh start at the first three lines. Then, we

Algorithm 1 Greedy Heuristic - HWSD.

Input: $model$

Output: $model, h$

X_{uv} $\leftarrow$: get_wireless_links($model$)
ξ_{sp} $\leftarrow$: get_vm_deployment($model$)
$h \leftarrow$: 0
$w \leftarrow$: get_WSD_type()
$r_{current} \leftarrow$: 1
$r_{next} \leftarrow$: find_closest_rack($r_{current}$)
X_{uv} $\leftarrow$: set_curr(1), where $u, v = r_{current}, r_{next}$
$v_{current} \leftarrow$: find_max_demanded_service()
decrease_pm_capacities($P1$, $v_{current}$)
while ($v_{current}$ **is** unvisited) **do**
    $v_{next} \leftarrow$: find_next_service($r_{current}, r_{next}, w$)
    pm_found $\leftarrow$: 0
    if (Link capacity between $r_{current}$ and r_{next} is enough to deploy v_{next}) **then**
        if (There is at least one PM p in the r_{next} whose capacity is enough to deploy v_{next}) **then**
            pm_found $\leftarrow$: 1
            ξ_{sp} $\leftarrow$: set_curr(1) , where $s = v_{next}$
            update(h)
            decrease_link_capacities($r_{current}, r_{next}, v_{next}$)
            decrease_pm_capacities(p, v_{next})
            $v_{current}$ $\leftarrow$: v_{next}
            exchange_values($r_{current}$ r_{next})
        end if
    end if
    if pm_found **is** 0 **then**
        $r_{current} \leftarrow$: choose_rack()
        $r_{next} \leftarrow$: find_closest_rack($r_{current}$)
        $v_{current} \leftarrow$: find_max_demanded_service()
        $p_{current} \leftarrow: {}'P' + str((r_{current} - 1) * pm\_per\_rack + 1)$
        X_{uv} $\leftarrow$: set_curr(1), where $u, v = r_{current}, r_{next}$
        ξ_{sp} $\leftarrow$: set_curr(1) , where $s, p = v_{current}, p_{current}$
        decrease_pm_capacities($p_{current}$, $v_{current}$)
    end if
end while
return $model, h$

get the type of the WSD problem if it is hybrid-WSD or pure-WSD as shown in the fourth line. After choosing the $rack1$ as the current rack and the closest rack to the $rack1$ as the next rack, we establish a wireless link between them and set the related decision variables, as seen between the lines 4 and 8. We choose the service having the most traffic demand to/from other services as the current service and decrease the capacity of the physical machine in which the current service will be embedded with the help of the defined functions, as shown between the lines 7 and 10.

After embedding the current service, the next service which will be embedded is chosen according to the WSD problem type. If the WSD problem is pure-WSD, we choose the service having the most total traffic demand from/to the services already deployed in the current rack. If the WSD problem is hybrid-WSD, we choose the service having the most total traffic demand from/to the services already deployed in the next and current rack. The related part in Algorithm 1 is the line 11, and we define a function for this purpose. If there is a physical machine having enough capacity to serve the chosen service's source demand and if the channel capacity between the chosen racks is enough to satisfy traffic demand of this service, we embed it to the physical machine chosen, as seen between the lines 12 and 23. Whenever we embed a virtual machine to a physical machine, we set the value of the decision variable related with virtual machine deployment and update the heuristic solution, physical machine capacities, and link capacities, as shown between the lines 15 and 20. For further deployments, we need to set the current service as the next service and exchange values of the current and next rack, as seen in the lines 20 and 21.

If the channel capacity between chosen racks is not enough to deploy the chosen service, or there is no physical machine with enough capacity for this service, we choose a new rack as the current rack and the closest rack to the current rack as the next rack. After this process, similar actions are taken with the beginning of the algorithm, as shown between the lines 24 and 32. All of these processes continue until all the services are visited.

Algorithm 2 Improved HWSD - I-HWSD.

Input: $model, h, U, V, S$

Output: h

$X_{uv} \leftarrow:$ get_wireless_links($model$), $\xi_{sp} \leftarrow:$ get_vm_deployment($model$)
$g_{total} \leftarrow: 0.1$, $\epsilon \leftarrow: 0$, $w \leftarrow:$ get_WSD_type()
while ($g_{total} > \epsilon$) **do**
$g_{total} \leftarrow: 0$
for v in S **do**
$t_{satisfied} \leftarrow: 0$
if v **not in** U **then**
$r_{current} \leftarrow:$ get_rack_info(v) , $r_{neighbour} \leftarrow:$ get_neighbour_rack(v)
if w **is** $hybrid\_WSD$ **then**
$t_{satisfied} \leftarrow:$ get_total_demand($r_{neighbour}, v$) + get_total_demand($r_{current}, v$)
else
$t_{satisfied} \leftarrow:$ get_total_demand($r_{neighbour}, v$)
end if
end if
$re\_deployment \leftarrow: 0$, $d \leftarrow: null$
$t_{demand} \leftarrow: 0$, $t_{contr} \leftarrow: t_{satisfied} - t_{demand}$
for r **in** V **do**
$r_{deploy} \leftarrow:$ get_connected_rack(r)
if w **is** $hybrid\_WSD$ **then**
$t_{demand} \leftarrow:$ get_total_demand(r, v) + get_total_demand(r_{deploy}, v)
else
$t_{demand} \leftarrow:$ get_total_demand(r, v)
end if
if $t_{demand} > t_{satisfied}$ **then**
if (There is at least one PM p in the r_{deploy} whose capacity is enough to re-deploy v) **then**
$d \leftarrow:$ get_deployment_info(p, r_{deploy}, v)
$t_{satisfied} \leftarrow: t_{demand}$
$re\_deployment \leftarrow: 1$
end if
end if
end for
if $re\_deployment$ **then**
deploy_service(d), update(h), update(X_{uv}), update(ξ_{sp})
$g_{total} \leftarrow: g_{total} + (t_{satisfied} - t_{contr})$
end if
end for
end while
return h, X_{uv}, ξ_{sp}

4.2 Improved Heuristic for Wireless Link and Service Deployment (I-HWSD)

Improved Heuristic for Wireless Link and Service Deployment (I-HWSD) aims to improve the solution of the WSD problem found by HWSD in a reasonable time. For this purpose, it checks every service if it needs to be re-deployed or not according to the increase in the heuristic solution. As shown in Algorithm 2, I-HWSD takes the model, heuristic solution, undeployed services', and the list of racks and services as input from HWSD. We first get the decision variables X_{uv} and ξ_{sp}, set the value of the total gain and epsilon for a fresh start, and get the type of the WSD problem, as seen between the lines 1 and 3. Total gain is the total traffic demand gain after the re-deployment process of all services, while epsilon is the limiting factor of the re-deployment process. As long as total gain is greater than epsilon, re-deployment process continues, as shown between the lines 3 and 37.

After choosing the service to be re-deployed in the line 5, we check if it is deployed into a physical machine in HWSD. If it is, we get the information of the rack that it is deployed and the rack that connects to the rack that it is deployed. Then, according to the WSD problem type, we get the contribution of this service to the total traffic met. The related lines are between 7 and 14. If the service is not deployed in HWSD, then it does not have a contribution to the total traffic demand satisfied. After this process, the aim is to find out if a service contributes more to the total traffic demand met when it is embedded into another physical machine, as seen between the lines 17 and 31. If it does, we find the physical machine where the service contributes the most to the total satisfied traffic demand when embedded, with the help of the line 24. After checking for every rack, we deploy the service into its new physical machine and update the heuristic solution and decision variables, as shown between the lines 32 and 35. All the processes continue as long as the total gain is greater than epsilon. Finally, I-HWSD returns the improved heuristic solution and decision variables.

In Figure 4.1, the flowchart demonstrates all these processes of I-HWSD. The algorithm checks if there is a virtual machine that is not visited, and if it is, it chooses one of them. Then, it checks whether there is a physical machine that maximizes the contribution to the throughput when the selected virtual machine is deployed. If it finds a physical machine, the deployment process occurs, and the total gain is calculated. If

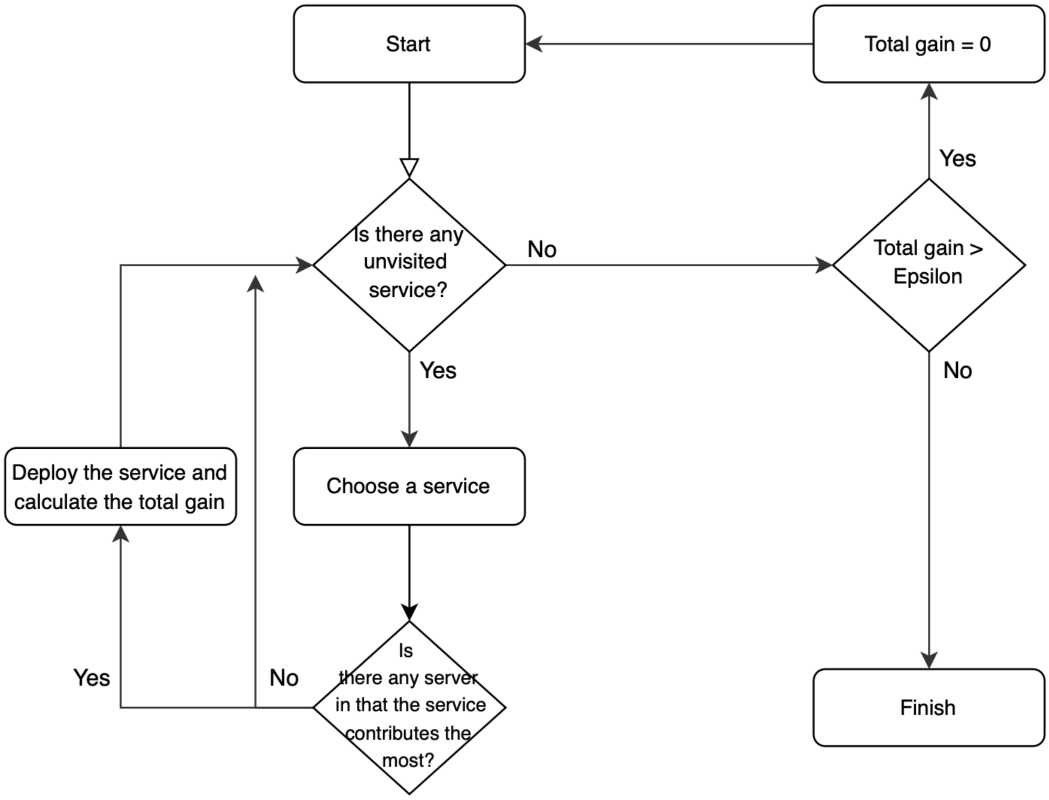

Figure 4.1: The flowchart of I-HWSD.

there is no virtual machine to deploy, and the total gain is less than or equal to epsilon, the process ends. If there is no virtual machine to deploy, and the total gain is greater than epsilon, the process starts again.

CHAPTER 5

RESULTS AND DISCUSSION

5.1 Experiment Setup

In this section, we analyze the numerical solutions of the WSD problem, which is an optimization problem solved by an optimization solver and heuristics named HWSD and I-HWSD. We use Gurobi Optimizer [46] as an optimization solver with the version of 9.1.1 to solve the WSD problem. To implement our proposed heuristics named HWSD and I-HWSD, we use Python programming language with the version of 3.8.10 [47]. The source code repository of the heuristics is [48]. We work on a server with 64 GB RAM and 10-core Intel Xeon Silver-4114 2.20 GHz. We repeat each experiment 30 times and obtain the results by taking the average of the findings in each experiment.

For the experiments:

- Table 5.1 shows 13 WDCN designs to use in experiments. These data centers are small-sized and generally used for educational purposes [49]. Each data center contains different numbers of racks, virtual machines, and physical machines. Racks are arranged in columns, rows, and groups. We assume that the distance is 1 meter between two groups of columns and rows. Each rack contains five physical machines [4].
- We assume that RAM and cache memory of physical machines are 256 GB and 20 MB, respectively [3].
- As shown in Table 5.2, we use Gaussian and Poisson distributions with different parameters for traffic generation between virtual machines. In the experiments,

Table 5.1: WDCN designs for the experiments.

# of Racks	Column x Row x Group	# of Servers	# of Services
2	2x1x1	10	25
4	2x2x1	20	50
6	3x2x1	30	75
8	4x2x1	40	100
10	5x2x1	50	125
12	4x3x1	60	150
20	5x4x1	100	250
30	5x6x1	150	375
40	5x8x1	200	500
60	5x6x2	300	750
80	5x8x2	400	1000
100	10x5x2	500	1250
120	10x6x2	600	1500

we use different derivations of Gaussian traffic distributions with the mean of 4303.0 bps and the standard deviation of 264.45 bps which is considered as query-response based traffic for small-sized DCNs [50], and with the mean 100 $Mbps$ and the standard deviation 114.153 $Mbps$ which is considered as multimedia traffic for small-sized DCNs [4]. We also utilize Uniform distribution generating transmission rates between the range of 10 $Mbps$ and 1 $Gbps$. We use different traffic distributions to observe the adaptability of our proposed WDCN and heuristics to different conditions. Source and destination virtual machines between which we generate traffic demand are chosen uniformly.

- We uniformly generate the source demand of virtual machines, such as RAM, cache memory, and CPU, from the given value interval in Table 3.2.

- We assume that if two physical machines are deployed on the same rack, the channel capacity between these physical machines is high, so we do not need to take the channel capacity between physical machines deployed on the same rack into account while solving the WSD problem.

Table 5.2: Traffic distributions used in the experiments.

Type	Mean (bps)	Std. (bps)	Rate (bps)	Range	Reference
Gauss	2151.5	264.45	-	-	The third model is taken from [50]. Others are derived from this model.
	4303.0	132.225	-	-	
	4303.0	264.45	-	-	
	4303.0	528.9	-	-	
	8606.0	264.45	-	-	
	172012.0	264.45	-	-	
Gauss	100×10^6	114.153×10^6	-	-	The first model are taken from [4]. Others are derived from this model.
	400×10^6		-	-	
	800×10^6		-	-	
	1000×10^6		-	-	
	2000×10^6		-	-	
Poisson	-	-	2151.5	-	The models are derived from the models in [50] and [4].
	-	-	4303.0	-	
	-	-	8606.0	-	
	-	-	100×10^6	-	
	-	-	400×10^6	-	
	-	-	800×10^6	-	
Uniform	-	-	-	[10 $Mbps$, 1 $Gbps$]	Taken from [11].

In the experiments:

- We introduce two heuristics to shorten the time to solve the WSD problem by Gurobi MIP Solver. We compare time performance and solutions found by the Gurobi MIP solver and the heuristics in Subsection 5.2.1.
- We analyze the ratio of the wireless traffic demand met to the total traffic demand met after solving both the pure-WSD and hybrid-WSD problems with the heuristics in Subsection 5.2.1.
- We evaluate the effect of some parameters such as distance between racks and

traffic distribution on the found solution in Subsection 5.2.2.

- We examine the reaction of our proposed solution to dynamic traffic changes such as adding and removing traffic flow and changing traffic distribution type in Subsection 5.2.3.
- We compare the findings of our proposed solution with the results of ToR-to-ToR WDCN [11] in terms of the ratio of satisfied traffic demand in Subsection 5.2.4.

5.2 Results

5.2.1 Solution to the WSD Problem

In Chapter 3, we introduce the pure and hybrid-WSD problems. Because these optimization problems are too complex, solving them by the Gurobi MIP Solver takes too long. Hence, we propose two heuristics in Chapter 4. HWSD aims to solve the formulated problems in a much shorter time, while the goal of I-HWSD is to improve the total traffic demand met found by HWSD in a reasonable time.

In the experiments, we solve these two problems with heuristics. Figure 5.1 demonstrates the total traffic demand met after running HWSD and I-HWSD on the pure and hybrid-WSD problems in *bps*, respectively. We run heuristics on the designed WDCNs shown in Table 5.1 and calculate the total traffic met by considering both the intra and inter-rack traffic demands. We generate the traffic demands between virtual machines normally with the mean 100 $Mbps$ and the standard deviation 114.153 $Mbps$ [4]. Because we do not take intra-rack communication into account in the formulation of the pure-WSD problem and try to maximize only the total inter-rack traffic demand satisfied via wireless links, the result is less than the found by solving the hybrid-WSD problem. In the formulation of the hybrid-WSD problem, we try to maximize both the total intra and inter-rack traffic demand met so we obtain more throughput in all WDCN models. Figure 5.1 also shows how much I-HWSD improves the solution found by HWSD. The improvement changes between 1% and 9% in the solution of the pure-WSD problem while 1% and 10% in the solution of the

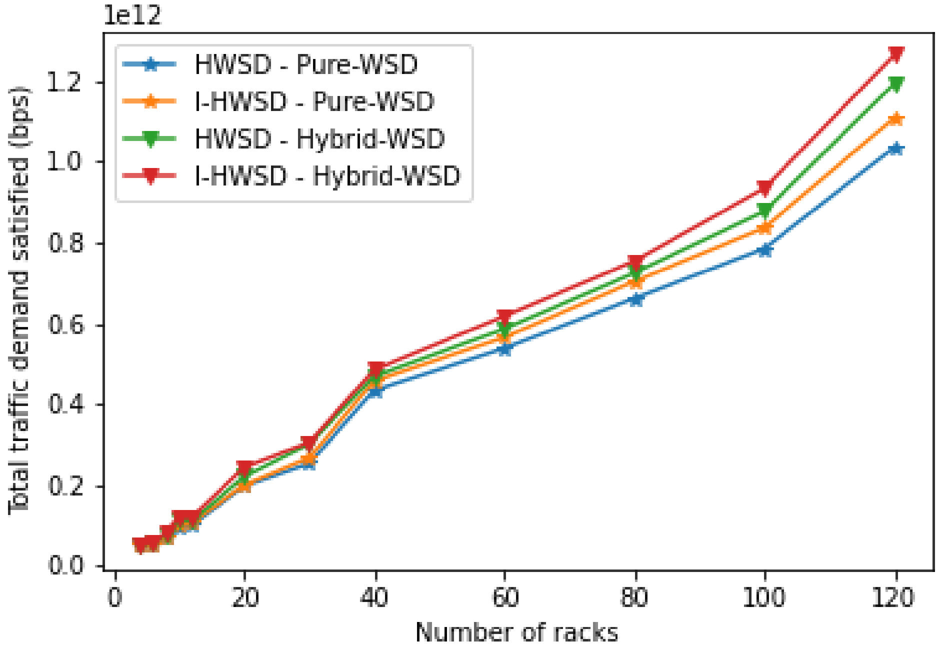

Figure 5.1: The total traffic demand met after heuristics.

hybrid-WSD problem.

Figure 5.2 shows the throughput after solving the pure-WSD problem in the designed WDCNs with 2 to 30 racks by the heuristics HWSD and I-HWSD and the Gurobi MIP Solver. These two heuristics approach the optimal solution by about 20% with a significant improvement in the time performance. Table 5.3 demonstrates the time performance of HWSD, I-HWSD, and the Gurobi MIP solver in finding a solution to the pure-WSD problem. HWSD does not take more than 9 seconds to find an optimal virtual machine deployment and wireless link establishment scenario for all WDCN designs. I-HWSD improves the solution found by HWSD in a reasonable time, but it takes much more time when the data center gets bigger. Remarkably, proposed heuristics take much less time compared to the Gurobi MIP Solver, but the time performance of the heuristics gets poor when the data center grows.

Figure 5.3 shows how much of the satisfied total traffic demand found after running HWSD and I-HWSD is satisfied via wireless links established between transceivers deployed on top of the racks in percent. Because solving the pure-WSD problem requires considering only inter-rack traffic, the ratio of the wireless traffic demand

Table 5.3: Solution time comparison of HWSD, I-HWSD, and the Gurobi MIP Solver in seconds.

# of Racks	HWSD	I-HWSD	Gurobi MIP Solver
2	0.001	0.001	110
4	0.003	0.001	115672
6	0.006	0.014	354289
8	0.009	0.025	
10	0.014	0.092	
12	0.022	0.149	
20	0.065	0.671	
30	0.196	2.406	More than 2 days.
40	0.421	5.558	
60	1.29	21.348	
80	3.027	43.0	
100	5.507	88.92	
120	8.752	374.068	

met to total traffic demand met is much more than the one in the solution of the hybrid-WSD problem. This ratio is not consistent in small data centers with less than 20 racks. For WDCNs with more than 30 racks, the ratio is usually steady, about 97% for the solution of the pure-WSD and about 68% for the solution of the hybrid-WSD problem. Hence, we can conclude that the hybrid-WSD problem ensures to embed virtual machines into physical machines in a way that more evenly distributes the traffic demands in the WDCN in terms of intra and inter-rack demands.

5.2.2 Effects on Results

While deploying virtual machines into physical machines and establishing wireless links between transceivers, the capacity of both physical machines and the channel between two transceivers plays an important role. The distance between two transceivers is one of the parameters having an impact on the channel capacity between these two transceivers. Figure 5.4 shows the effect of the distance between

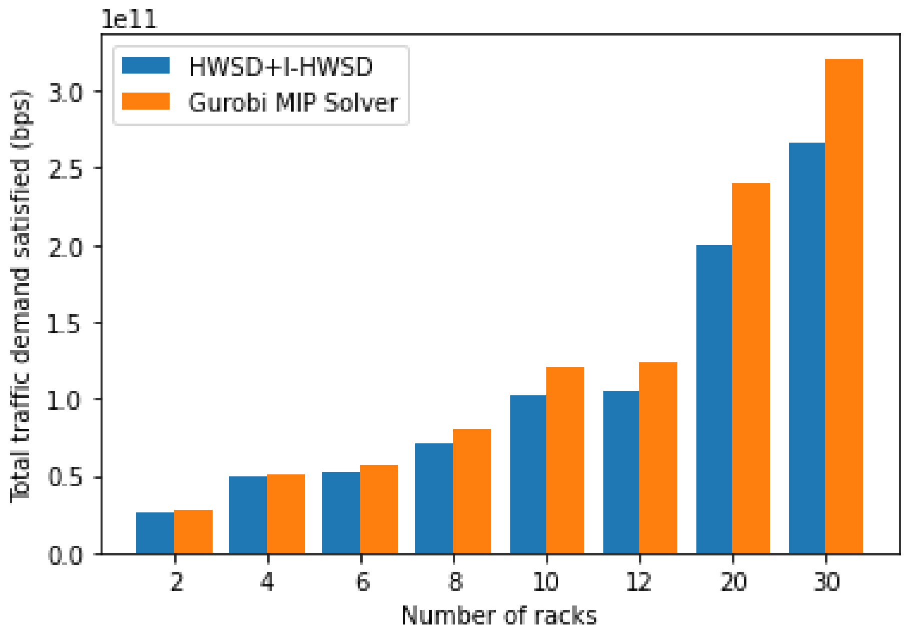

Figure 5.2: The comparison of total traffic demand satisfied obtained by the proposed heuristics and the Gurobi MIP Solver.

racks on the total traffic demand satisfied in the designed WDCN with 80 racks. HWSD and I-HWSD run on the hybrid-WSD problem with normally distributed traf-

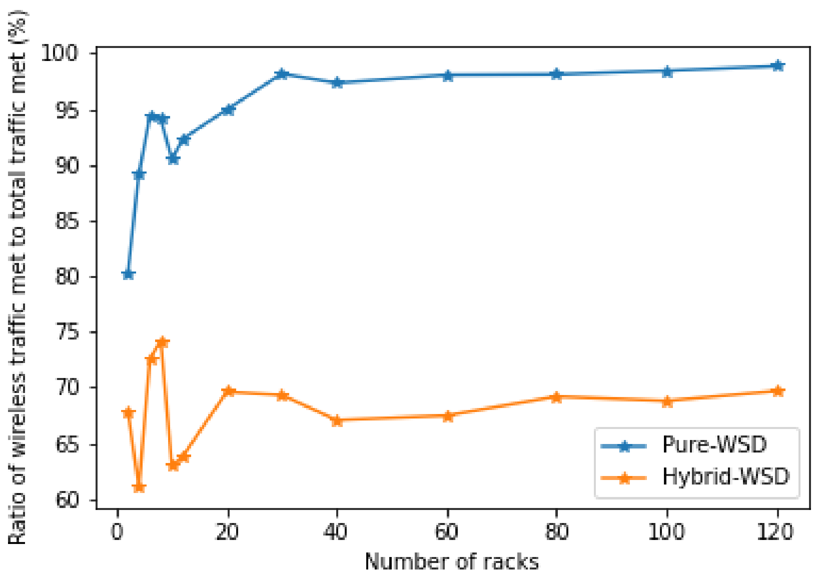

Figure 5.3: Ratio of the wireless traffic met to the total traffic met after heuristics.

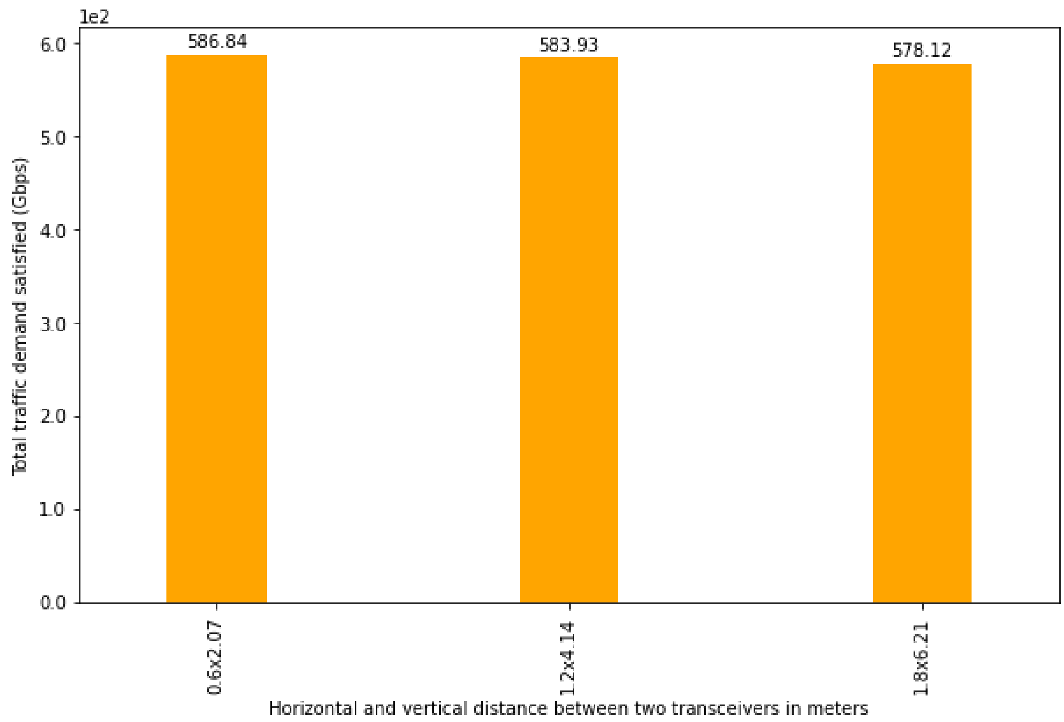

Figure 5.4: The impact of the distance between the racks on the throughput achieved by HWSD and I-HWSD.

fic demands with the mean of 2 $Gbps$ and standard deviation of 114.153 $Mbps$. In the designed WDCNs, the distance between two transceivers deployed on the top of the side-by-side racks is 0.6 meters and between two transceivers deployed on the top of the back-to-back racks is 2.07 meters. We analyze the impact of doubling and tripling the distance between two transceivers deployed on the racks in the same row and in the same column, respectively. This increase causes to satisfy fewer traffic demands because the capacity between transceivers decreases with the increase in distance. The decrease is not significant, so we can conclude that the physical machine capacity is the main limiting factor of deploying virtual machines.

We also evaluate the impact of the traffic increase in the WDCN on the throughput obtained by HWSD and I-HWSD run on the hybrid-WSD problem with normally distributed traffic demands of the mean 100 $Mbps$ and standard deviation 114.153 $Mbps$, as shown in Figure 5.5. The 60% and 90% decrease and the 60% increase in the traffic demand cause approximately the same amount of decrease and increase in the throughput obtained by HWSD and I-HWSD. The 90% increase in the traffic

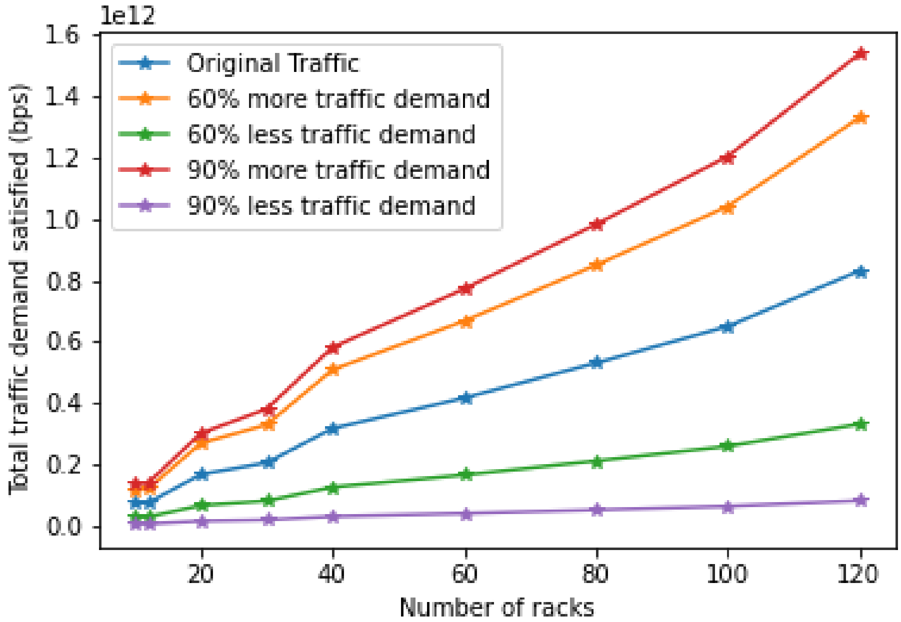

Figure 5.5: The impact of the traffic demand change of the services on the throughput achieved by HWSD and I-HWSD.

demand causes an increase in the throughput, but the increase is much less than 90%. The reason is that the channel capacity between transceivers does not allow to meet all demand increases, and it affects the virtual machine deployment decisions.

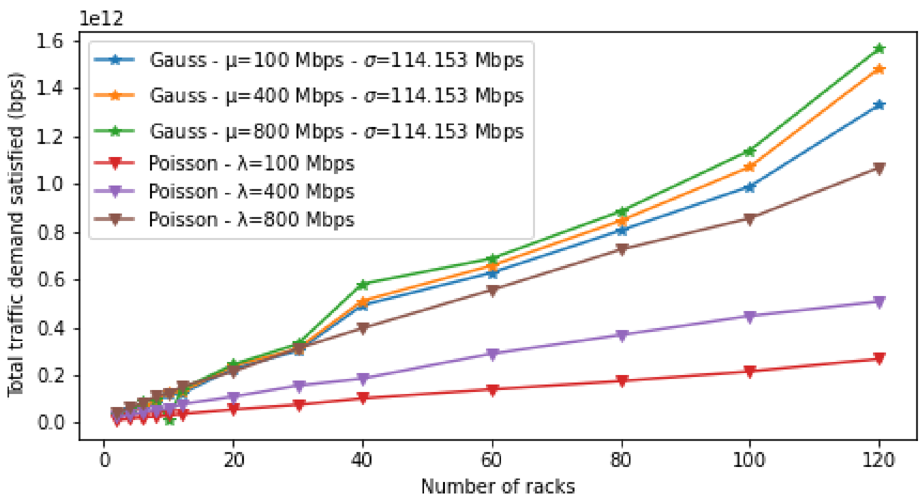

Figure 5.6: The total traffic demand met after type of traffic distribution changes.

Figure 5.6 demonstrates the throughput obtained by HWSD and I-HWSD on the hybrid-WSD with the traffic demands distributed with Gaussian and Poisson distributions between virtual machines. When the mean increases with the constant standard deviation, the obtained throughput increases in Gaussian distribution by less degree than Poisson distribution. When the traffic is generated by Gaussian and Poisson distributions with the same mean and rate, respectively, the throughput obtained with the normally distributed traffic is always higher. The results show us our WSD problem and heuristics to solve it are applicable for different traffic distributions.

5.2.3 Dynamic Traffic Demands

To analyze the behavior of our designed WDCN and hybrid-WSD problem and heuristics to the dynamically changing traffic demands between virtual machines, we expose them to a dynamic environment. This dynamic environment contains the traffic flow addition and deletion and the change of the traffic distribution in the WDCN. The flow is the traffic rate between a source service and a destination service in bps. In the flow addition and deletion experiment, we add and delete elephant flows and or-

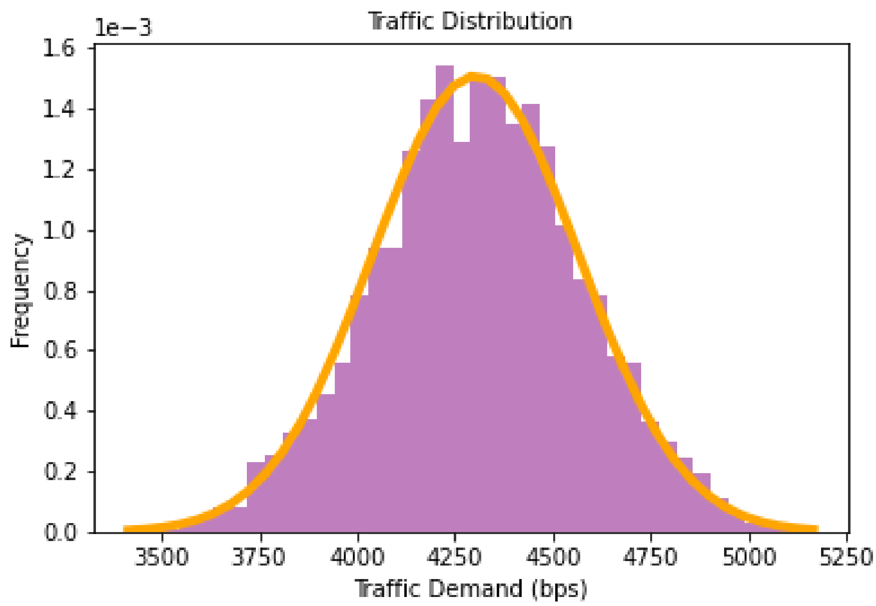

Figure 5.7: Gaussian distribution with the mean 4303 bps and the standard deviation 264.45 bps for traffic generation.

dinary flows. Elephant flows are extremely large TCP flows produced by applications running in data centers and increase the chance of network congestion [51], [52]. The ordinary flow is a bit large flow in terms of flow sizes generated in the WDCN.

For flow addition and deletion experiments, we assume that the traffic distributed normally with the mean 4303 bps and the standard deviation 264.45 bps in the designed WDCN with 4 to 40 racks, and we separately add and delete ten elephant and ordinary flows. Figure 5.7 shows the Gaussian traffic distribution in the WDCN with 20 racks, while Figure 5.8 shows the rate of 10 elephant flows added in $Gbps$. We can observe how big the rate of the elephant flows compared to the generated data rates.

We demonstrate the performance values such as success, total gain, and the dynamicity of virtual machine deployment and wireless link establishment processes obtained after adding and deleting elephant and ordinary flows in Table 5.4. Success represents the ratio of the number of flow added and met to the total number of flow added, while total gain represents the ratio of the satisfied traffic demand gain after a flow is added and met to the rate of the added flow. The highest value of success is equal to 1. After a flow is added or deleted, we first optimize the links to find an optimal wireless link establishment scenario, and then we call I-HWSD to find an optimal deployment for each virtual machine. Hence, virtual machine deployment dynamicity represents the rate of service deployment in I-HWSD, while wireless link dynamicity represents the rate of wireless link re-establishment in percent. We calculate virtual machine deployment dynamicity for each WDCN design by dividing the number of virtual machines re-deployed in each I-HWSD round by the total number of virtual machines in the WDCN. We also calculate wireless link dynamicity by dividing the number of wireless links re-established by the total number of wireless links. These values are the average of the values obtained after the experiments in each WDCN design.

Table 5.4 demonstrates that the rate of meeting the added elephant flow is higher than meeting the added ordinary flow because satisfying the elephant flow tends to increase the optimal solution much more. When an elephant flow is satisfied, because of the re-deployment process of virtual machines, the total gain is not equal to the size of the elephant flow. The situation is different when we add an ordinary flow. After the re-

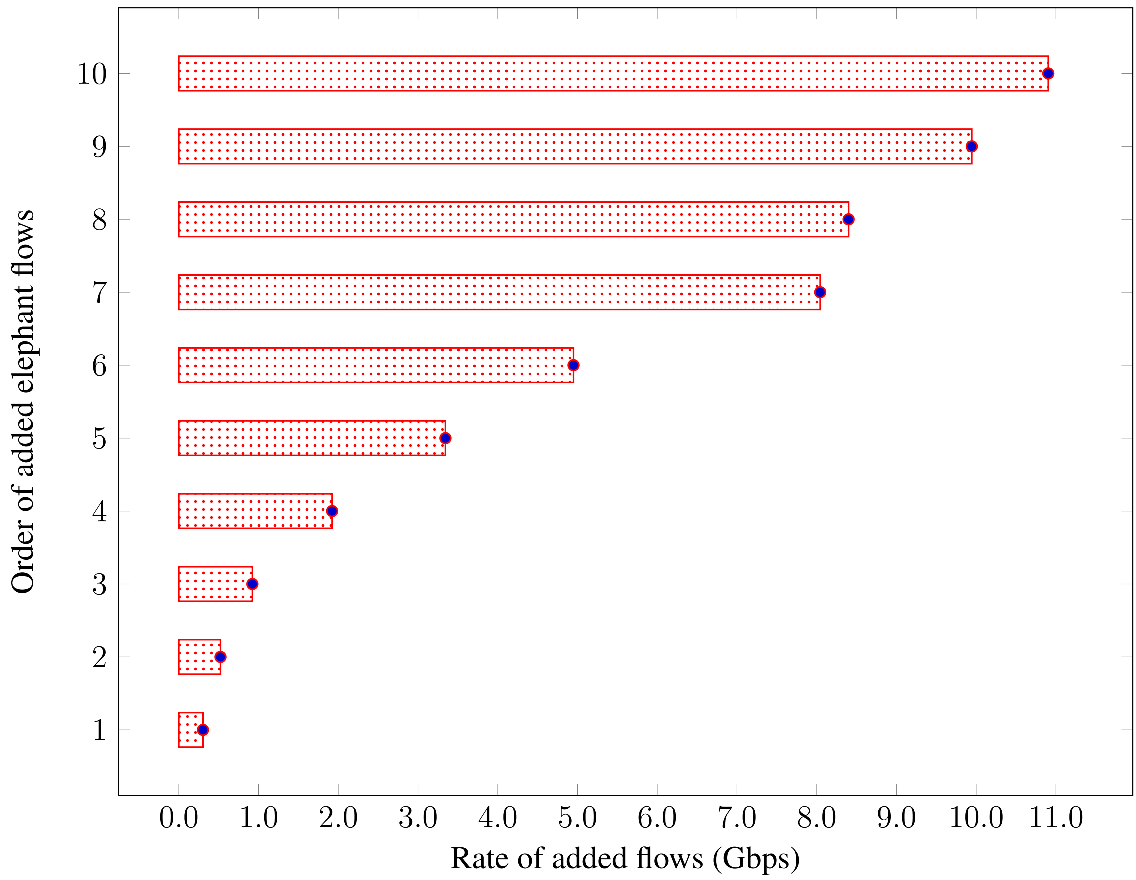

Figure 5.8: The rate of elephant flows added in order.

Table 5.4: Results of flow addition and deletion experiments.

Flow Type	Elephant	Elephant	Ordinary	Ordinary
Addition	✓	✗	✓	✗
Deletion	✗	✓	✗	✓
Success	1.0	-	0.87	-
Total Gain	0.98	-	1.02	-
VM Dynamicity (%)	5.52	0.74	4.32	0.02
Wireless Link Dynamicity (%)	45.36	5.5	26.17	0.0

deployment process, some flows other than the added one are also satisfied which is why the total gain rate is greater than 1. Moreover, wireless links are more susceptible to flow additions than the deployment of virtual machines. Because re-deployment of virtual machines is costly, the results indicate a decrease in the deployment costs. Flow deletions require much less virtual machine re-deployment and wireless link re-establishment processes which contribute to the dynamicity of the WDCN.

We also evaluate the reaction of our heuristics to the changes in traffic distribution. In this experiment, we continuously change the traffic distribution in the WDCN with 4 to 20 racks. Then, we re-establish wireless links and re-deploy virtual machines for adapting to the traffic changes in the network. Figure 5.9 demonstrates the ratio of the satisfied traffic to the total traffic in the network for each WDCN design in percent with different traffic generation distributions. The distributions in Figure 5.9 are shown in Table 5.2 in the fifth, first, sixth, twelfth, and tenth rows, respectively. When the distribution changes, the hybrid-WSD problem adapts to these continuous changes and satisfies a similar ratio of the total traffic. It also demonstrates the decrease in the ratio when the WDCN grows.

Figure 5.10 shows the average dynamicity of virtual machine deployment and wireless link establishment processes in percent to adapt to the continuously changing

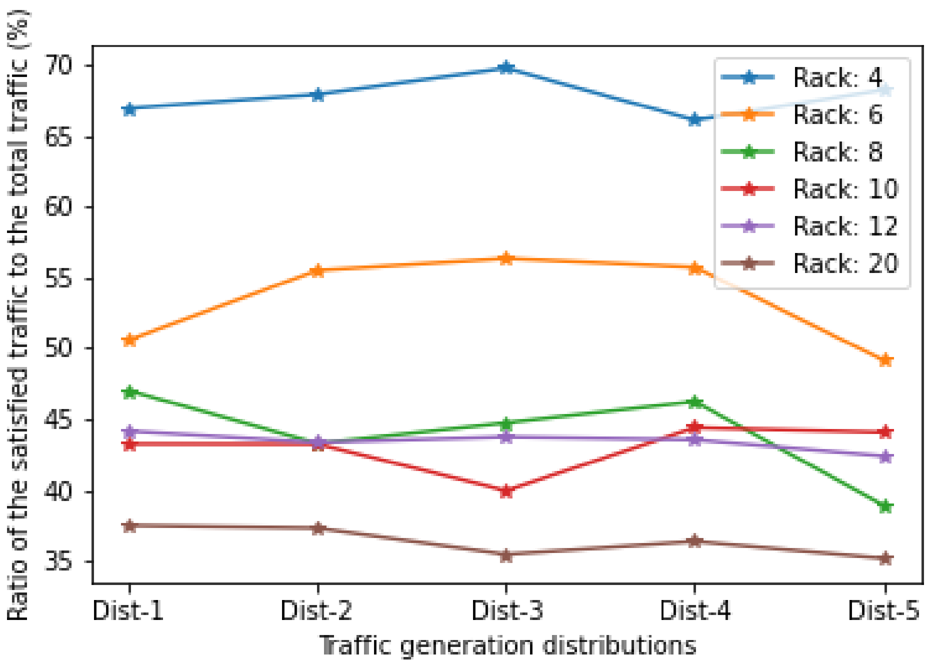

Figure 5.9: The ratio of the satisfied traffic demand to the total traffic demand with continuously changing traffic distributions.

traffic distributions. We calculate the dynamicity of virtual machine deployment for each WDCN design with 4 to 20 racks by dividing the number of virtual machines re-deployed in each I-HWSD round to adapt to ten continuously changing traffic distributions by the total number of virtual machines in the WDCN. Wireless link dynamicity is calculated by dividing the number of wireless links re-established to keep pace with ten different traffic distributions by the total number of wireless links. These values are the average of values obtained after each traffic distribution change. The sensitivity of wireless links to the dynamic traffic changes in the WDCN causes the wireless link dynamicity to be greater than the virtual machine deployment dynamicity. Adapting to the changing traffic demands caused by changes in traffic generation distribution with a minimum amount of virtual machine re-deployment decreases the re-deployment costs. The cost increases with the increase in virtual machines re-deployment when the WDCN grows.

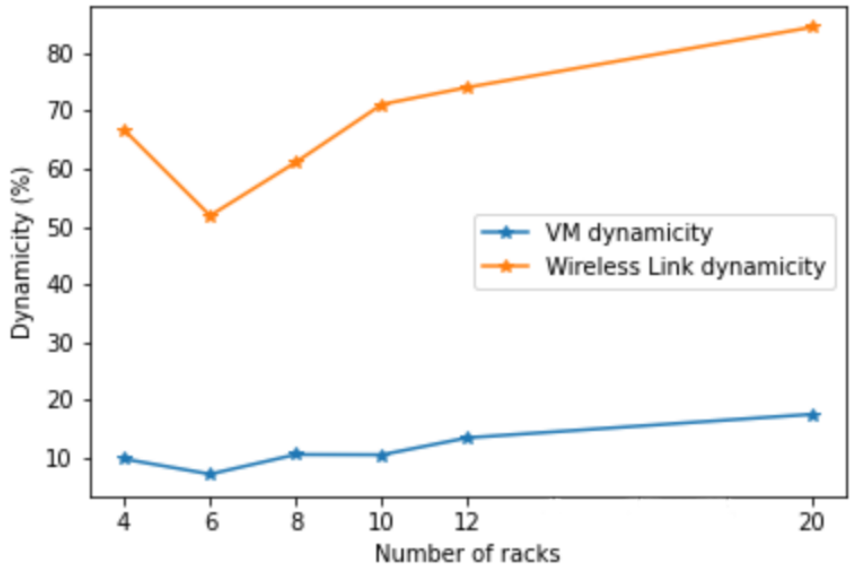

Figure 5.10: The average dynamicity of virtual machine deployment and wireless link establishment processes after continuously changing traffic distributions.

5.3 Demand Completion Success

Mamun et al. [11] propose a WDCN where established ToR-to-ToR wireless links satisfy the traffic demands between different racks, and evaluate the performance of ToR-to-ToR WDCN in terms of the number of completed demands. Authors of [11] work on a simulation environment and run the experiments to obtain performance values with 1000 flows generated over 100 seconds. The sizes of the generated flows are randomly distributed between 1 MB and 1 GB, and data rates are uniformly selected between 10 $Mbps$ and 1 $Gbps$ [53]. We only work on the snapshot of the network different from ToR-to-ToR WDCN so we find the total flow number of when the ToR-to-ToR WDCN oversubscribes, and we work with these flows.

We compare the ratio of completed demands achieved by top-of-rack (ToR)-to-ToR WDCN and our proposed WDCN that are evaluated with a single 60-GHz channel. In this experiment, we work on a WDCN with 160 racks. Figure 5.11 demonstrates the ratio of completed demands of both the single-channel ToR-to-ToR WDCN pro-

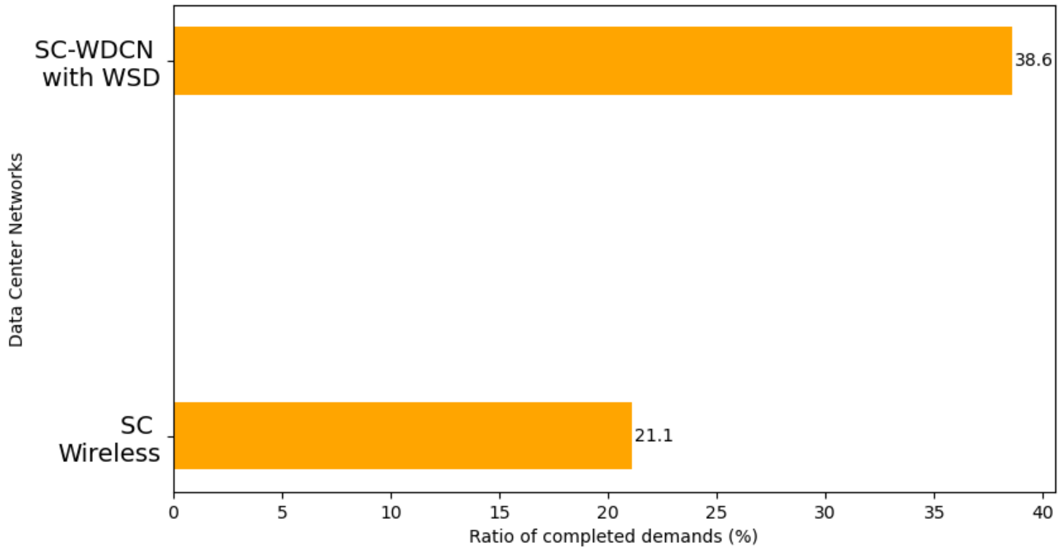

Figure 5.11: The ratio of completed demands.

posed by [11] and our designed single-channel WDCN obtained by the hybrid-WSD problem. Our proposed single-channel WDCN outperforms the single-channel ToR-to-ToR WDCN by completing traffic demands with about a 39% ratio, while the ratio for single-channel ToR-to-ToR WDCN is about 21%. The reason is that, unlike ToR-to-ToR WDCN, deploying virtual machines makes a significant contribution to the traffic demands we can complete.

CHAPTER 6

CONCLUSIONS AND FUTURE WORK

In this work, the aim is to maximize the throughput in WDCNs with dynamic traffic demands. For this purpose, we design a WDCN utilizing internationally available 60 GHz bands, which are useful for indoor environments. We work with 2 to 120 racks in this designed WDCN. We also introduce two MIP optimization problems named pure-WSD and hybrid-WSD. The goal of these problems is to maximize the throughput in the designed WDCN with different rack numbers, by jointly establishing wireless links between transceivers deployed on top of the racks and embedding virtual machines into physical machines. The pure-WSD aims to maximize the inter-rack traffic demand met, while the goal of the hybrid-WSD problem is to maximize the total intra and inter-rack traffic demand satisfied. Hence, the hybrid-WSD problem distributes the traffic demand in the WDCN more evenly.

We introduce two heuristics named Heuristic for Wireless Link and Service Deployment (HWSD) and Improved Heuristic for Wireless Link and Service Deployment (I-HWSD) because solving these problems with the Gurobi MIP solver, which is an optimization solver, takes a too long time. These two heuristics shorten the solution time taken by the Gurobi MIP solver, and I-HWSD improves the solution found by HWSD. HWSD and I-HWSD approximate the optimal solution found by the Gurobi MIP solver by about 20%, and I-HWSD improves the throughput obtained by HWSD on both pure-WSD and hybrid-WSD problems by between 1% and 10%. We also analyze the limiting factors, such as the distance between racks and the capacity of physical machines, on the obtained throughput in the WDCN.

We evaluate the reactions of our proposed heuristics to the dynamically changing traffic demands such as flow addition and deletion between virtual machines and changing traffic generation distributions. With dynamic virtual machine embedding and wireless link establishment, the proposed heuristics meet the added flow and meet other virtual machines' traffic demands when a flow is removed. When the distribution type of traffic demands changes, they adapt to these changes by obtaining similar success in the traffic demand satisfied over the total traffic demand. We can conclude that our proposed WSD problems and heuristics comply with the dynamically changing traffic demands.

As we compare with ToR-to-ToR WDCN, our single-channel WDCN performs better than the single-channel ToR-to-ToR WDCN in terms of the ratio of completed demands. The reason why our proposed WDCN outperforms the single-channel ToR-to-ToR WDCN can be the fact that different from other studies, we deploy virtual machines into physical machines considering traffic demands between them.

As future work, we aim to work with a simulation environment to simulate the designed WDCN with dynamically changing traffic demands and to show wireless link establishment and virtual machine deployment processes.

REFERENCES

[1] M. Z. Zaaimia, R. Touhami, V. A. Fono, L. Talbi, and M. Nedil, “60 ghz wireless data center channel measurements: Initial results,” in *2014 IEEE International Conference on Ultra-WideBand (ICUWB)*, pp. 57–61, 2014.

[2] S. Rommel, T. R. Raddo, U. Johannsen, C. Okonkwo, and I. T. Monroy, “Beyond 5G - wireless data center connectivity,” in *Broadband Access Communication Technologies XIII* (B. B. Dingel, K. Tsukamoto, and S. Mikroulis, eds.), vol. 10945, pp. 125 – 133, International Society for Optics and Photonics, SPIE, 2019.

[3] B. Butun and E. Onur, “Joint virtual machine embedding and wireless data center topology management,” in *2021 17th International Conference on Network and Service Management (CNSM)*, pp. 63–69, 2021.

[4] S. A. Mamun, S. G. Umamaheswaran, A. Ganguly, M. Kwon, and A. Kwasinski, “Performance evaluation of a power-efficient and robust 60 ghz wireless server-to-server datacenter network,” *IEEE Transactions on Green Communications and Networking*, vol. 2, no. 4, pp. 1174–1185, 2018.

[5] M. Pedram, “Energy-efficient datacenters,” *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 31, no. 10, pp. 1465–1484, 2012.

[6] M. Al-Fares, A. Loukissas, and A. Vahdat, “A scalable, commodity data center network architecture,” *SIGCOMM Comput. Commun. Rev.*, vol. 38, p. 63–74, aug 2008.

[7] Y. Cui, H. Wang, X. Cheng, and B. Chen, “Wireless data center networking,” *IEEE Wireless Communications*, vol. 18, no. 6, pp. 46–53, 2011.

[8] A. Celik, B. Shihada, and M.-S. Alouini, “Wireless data center networks: Advances, challenges, and opportunities,” 2018.

[9] K. Kant, "Data center evolution: A tutorial on state of the art, issues, and challenges," *Computer Networks*, vol. 53, no. 17, pp. 2939–2965, 2009.

[10] S. Mollahasani and E. Onur, "Evaluation of terahertz channel in data centers," in *NOMS 2016 - 2016 IEEE/IFIP Network Operations and Management Symposium*, pp. 727–730, 2016.

[11] S. A. Mamun, S. G. Umamaheswaran, S. S. Chandrasekaran, M. S. Shamim, A. Ganguly, and M. Kwon, "An energy-efficient, wireless top-of-rack to top-of-rack datacenter network using 60ghz links," in *2017 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*, pp. 458–465, 2017.

[12] K. Ramachandran, R. Kokku, R. Mahindra, and S. Rangarajan, "60 ghz datacenter networking: Wireless worry less?," *NEC Research Paper*, vol. 1, 01 2008.

[13] A. Al-Saman, M. Cheffena, O. Elijah, Y. A. Al-Gumaei, S. K. Abdul Rahim, and T. Al-Hadhrami, "Survey of millimeter-wave propagation measurements and models in indoor environments," *Electronics*, vol. 10, no. 14, 2021.

[14] "Standard ecma-387 2nd edition: High rate 60 ghz phy, mac and pals,," 2010. `https://www.ecma-international.org/wp-content/uploads/ECMA-387_2nd_edition_december_2010.pdf`.

[15] "Ieee standard for information technology–telecommunications and information exchange between systems–local and metropolitan area networks–specific requirements-part 11: Wireless lan medium access control (mac) and physical layer (phy) specifications amendment 3: Enhancements for very high throughput in the 60 ghz band," *IEEE Std 802.11ad-2012 (Amendment to IEEE Std 802.11-2012, as amended by IEEE Std 802.11ae-2012 and IEEE Std 802.11aa-2012)*, pp. 1–628, 2012.

[16] wirelesshd consortium, "Wireless high-definition (wirelesshd)."

[17] E. Perahia, C. Cordeiro, M. Park, and L. L. Yang, "Ieee 802.11ad: Defining the

next generation multi-gbps wi-fi," in *2010 7th IEEE Consumer Communications and Networking Conference*, pp. 1–5, 2010.

[18] H. Vardhan, N. Thomas, S.-R. Ryu, B. Banerjee, and R. Prakash, "Wireless data center with millimeter wave network," in *2010 IEEE Global Telecommunications Conference GLOBECOM 2010*, pp. 1–6, 2010.

[19] F. Al-Ogaili and R. M. Shubair, "Millimeter-wave mobile communications for 5g: Challenges and opportunities," in *2016 IEEE International Symposium on Antennas and Propagation (APSURSI)*, pp. 1003–1004, 2016.

[20] C. Guo, H. Wu, K. Tan, L. Shi, Y. Zhang, and S. Lu, "Dcell: A scalable and fault-tolerant network structure for data centers," *SIGCOMM Comput. Commun. Rev.*, vol. 38, p. 75–86, aug 2008.

[21] C. Guo, G. Lu, D. Li, H. Wu, X. Zhang, Y. Shi, C. Tian, Y. Zhang, and S. Lu, "Bcube: A high performance, server-centric network architecture for modular data centers," *SIGCOMM Comput. Commun. Rev.*, vol. 39, p. 63–74, aug 2009.

[22] A. Greenberg, J. R. Hamilton, N. Jain, S. Kandula, C. Kim, P. Lahiri, D. A. Maltz, P. Patel, and S. Sengupta, "Vl2: A scalable and flexible data center network," *SIGCOMM Comput. Commun. Rev.*, vol. 39, p. 51–62, aug 2009.

[23] D. Li, C. Guo, H. Wu, K. Tan, and S. Lu, "Ficonn: Using backup port for server interconnection in data centers," *IEEE INFOCOM 2009*, pp. 2276–2285, 2009.

[24] E. Baccour, S. Foufou, R. Hamila, and M. Hamdi, "A survey of wireless data center networks," in *2015 49th Annual Conference on Information Sciences and Systems (CISS)*, pp. 1–6, 2015.

[25] J.-Y. Shin, E. G. Sirer, H. Weatherspoon, and D. Kirovski, "On the feasibility of completely wireless datacenters," in *2012 ACM/IEEE Symposium on Architectures for Networking and Communications Systems (ANCS)*, pp. 3–14, 2012.

[26] X. Zhou, Z. Zhang, Y. Zhu, Y. Li, S. Kumar, A. Vahdat, B. Y. Zhao, and H. Zheng, "Mirror mirror on the ceiling: Flexible wireless links for data centers," SIGCOMM '12, (New York, NY, USA), p. 443–454, Association for Computing Machinery, 2012.

[27] M. Z. Zaaimia, R. Touhami, L. Talbi, M. Nedil, and M. C. E. Yagoub, “60-ghz statistical channel characterization for wireless data centers,” *IEEE Antennas and Wireless Propagation Letters*, vol. 15, pp. 976–979, 2016.

[28] A. S. Hamza, J. S. Deogun, and D. R. Alexander, “Wireless communication in data centers: A survey,” *IEEE Communications Surveys Tutorials*, vol. 18, no. 3, pp. 1572–1595, 2016.

[29] C. Terzi and I. Korpeoglu, “60 ghz wireless data center networks: A survey,” *Computer Networks*, vol. 185, p. 107730, 2021.

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

[31] S. Kandula, J. Padhye, and P. Bahl, “Flyways to de-congest data center networks,” in *HotNets*, 2009.

[32] M. Luo, H. Sun, M. Sheng, Y. Shi, and J. Li, “Energy-efficient deployment of fat-tree hybrid data center networks,” in *2018 10th International Conference on Wireless Communications and Signal Processing (WCSP)*, pp. 1–6, 2018.

[33] X. Zhou, Z. Zhang, Y. Zhu, Y. Li, S. Kumar, A. Vahdat, B. Y. Zhao, and H. Zheng, “Mirror mirror on the ceiling: Flexible wireless links for data centers,” in *Proceedings of the ACM SIGCOMM 2012 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communication*, SIGCOMM ’12, (New York, NY, USA), p. 443–454, Association for Computing Machinery, 2012.

[34] N. Hamedazimi, Z. Qazi, H. Gupta, V. Sekar, S. R. Das, J. P. Longtin, H. Shah, and A. Tanwer, “Firefly: A reconfigurable wireless data center fabric using free-space optics,” in *Proceedings of the 2014 ACM Conference on SIGCOMM*, SIGCOMM ’14, (New York, NY, USA), p. 319–330, Association for Computing Machinery, 2014.

[35] Y. Cui, S. Xiao, X. Wang, Z. Yang, S. Yan, C. Zhu, X.-Y. Li, and N. Ge, “Diamond: Nesting the data center network with wireless rings in 3-d space,” *IEEE/ACM Trans. Netw.*, vol. 26, p. 145–160, feb 2018.

[36] A. Cayley, "On the theory of groups," *American Journal of Mathematics*, vol. 11, pp. 139–157, 1889.

[37] Y. Li, F. Wu, X. Gao, and G. Chen, "Sphericalmesh: A novel and flexible network topology for 60ghz-based wireless data centers," in *2014 IEEE/CIC International Conference on Communications in China (ICCC)*, pp. 796–800, 2014.

[38] C. Zhang, F. Wu, X. Gao, and G. Chen, "Free talk in the air: A hierarchical topology for 60 ghz wireless data center networks," *IEEE/ACM Transactions on Networking*, vol. 25, no. 6, pp. 3723–3737, 2017.

[39] H. Vardhan and R. Prakash, "Concurrency in polygonally arranged wireless data centers with all line-of-sight links," in *2014 International Conference on Computing, Networking and Communications (ICNC)*, pp. 716–720, 2014.

[40] B. Cao, J. Zhao, P. Yang, Y. Gu, K. Muhammad, J. J. P. C. Rodrigues, and V. H. C. de Albuquerque, "Multiobjective 3-d topology optimization of next-generation wireless data center network," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 5, pp. 3597–3605, 2020.

[41] C.-L. Cheng and A. Zajić, "Characterization of 300 ghz wireless channels for rack-to-rack communications in data centers," in *2018 IEEE 29th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, pp. 194–198, 2018.

[42] B. Csuka and Z. Kollar, "R-dft-based parameter estimation for wigig," *Periodica Polytechnica, Electrical Engineering*, vol. 61, pp. 224–230, 01 2017.

[43] S. K. Yong and C.-C. Chong, "An overview of multigigabit wireless through millimeter wave technology: Potentials and technical challenges," *EURASIP J. Wirel. Commun. Netw.*, vol. 2007, p. 50, Jan. 2007.

[44] R. Daniels and R. Heath, "60 ghz wireless communications: Emerging requirements and design recommendations," *Vehicular Technology Magazine, IEEE*, vol. 2, pp. 41 – 50, 10 2007.

[45] N. Hu and W. Zhou, "How to choose an it rack." `https://www.se.com/ww/en/download/document/SPD_VAVR-9G4MYQ_EN/`.

[46] L. Gurobi Optimization, "Gurobi optimizer reference manual," 2021. `http://www.gurobi.com`.

[47] G. Van Rossum and F. L. Drake, *Python 3 Reference Manual*. Scotts Valley, CA: CreateSpace, 2009.

[48] B. Butun, "hwsd." `https://github.com/beyzabutun/hwsd`.

[49] T. Benson, A. Akella, and D. A. Maltz, "Network traffic characteristics of data centers in the wild," in *Proceedings of the 10th ACM SIGCOMM Conference on Internet Measurement*, IMC '10, (New York, NY, USA), p. 267–280, Association for Computing Machinery, 2010.

[50] S. S. Chandrasekaran, "Understanding traffic characteristics in a server to server data center network," 2017.

[51] A. H. Alawadi, M. Zaher, and S. Molnár, "Methods for predicting behavior of elephant flows in data center networks," *CoRR*, vol. abs/1911.11673, 2019.

[52] J. A. Rashid, "Sorted-gff: An efficient large flows placing mechanism in software defined network datacenter," *Karbala International Journal of Modern Science*, vol. 4, no. 3, pp. 313–331, 2018.

[53] S. Kandula, S. Sengupta, A. Greenberg, P. Patel, and R. Chaiken, "The nature of data center traffic: Measurements amp; analysis," in *Proceedings of the 9th ACM SIGCOMM Conference on Internet Measurement*, IMC '09, (New York, NY, USA), p. 202–208, Association for Computing Machinery, 2009.