HORIZON 2020	Deliverable ID: D5.5	Preparation date: 01 May 2019
	Milestone: Final Proposed	
	Title: **Initial NETWORK-layer and management model**	
Terahertz based Ultra High Bandwidth Wireless Access Networks	Editor/Lead beneficiary (name/partner): Sean Ahearne / Dell EMC	
	Internally reviewed by (name/partner): Edward Wasige / UGLA Alan Davy / TSSG	
	Approved by: TBC	

Dissemination level		
PU	Public	X
CO	Confidential, only for members of the consortium (including Commission Services)	

Revisions

Version	Date	Author	Organisation	Details
0.0	13.12.18	Niamh O'Mahony	Dell EMC	Structure proposed
0.1	15.01.19	Niamh O'Mahony	Dell EMC	Revised ToC.
0.2	11.02.2019	Sean Ahearne	Dell EMC	Added Summary, Preliminary Scenarios
0.3	11.03.2019	Sean Ahearne	Dell EMC	Added Multi-hop Scenario. Updated Previous Scenarios.
0.4	09.04.2019	Sean Ahearne	Dell EMC	Added 10Gbps Scenario, Tables, figures, exec summary, finished Chap 1 & 2, refs.
0.5	12.04.2019	Sean Ahearne	Dell EMC	Conclusion, clean-up.
0.6	15.04.2019	Noureddine B.	TSSG	DLL Section
0.7	25.04.2019	Sean Ahearne	Dell EMC	Final Edits
0.8	29.04.2019	Alan Davy	TSSG	Internal Review

The research leading to these results has received funding from the EU H2020 under grant agreement n° 761579.

Table of contents

List of figures

List of tables

Executive summary

The deliverable D5.5 gives an overview of the Network layer testing performed to determine the feasibility of terahertz (THz) wireless links in the datacentre. In this initial deliverable, network testing was performed virtually using network emulation tools and treating the THz links as point-to-point connections. The results show that THz wireless links can be used to assist, replace, or augment existing wired links in a datacentre topology. The output data gathered from this emulation can be passed to the data-link layer for simulation testing. The outline of communication between the data-link and network layers is described here, along with the functionalities provided by the data-link layer.

1 Introduction

1.1 Summary

This deliverable aims to investigate the potential use cases and behaviour of THz wireless links. This investigation will take the form of simulation, emulation, and theoretical modelling of these ultra-high bandwidth wireless links. The motivation for this investigation is to determine the effectiveness of high-speed wireless links in the datacentre, their future potential, and the new features they can provide over traditional optical networks.

The investigation is structured as follows:

1. Determine the potential use cases for THz wireless links in the datacentre.
2. Design simulation and emulation scenarios based on these use cases that show effective use of wireless links.
3. Deconstruct the events and data in each scenario to two of the OSI (Open Systems Interconnection) layers; Network, and Data Link.
4. Collect simulation data and present results from each scenario.

For future features of THz wireless links that cannot yet be simulated, a theoretical evaluation will be performed to determine how to implement such features into a datacentre network. This work will be reported in the later TERAPOD deliverable, D5.6.

1.2 Structure of this document

This document is structured as follows:

- Chapter 2: Overview
  - This chapter will give a summarized explanation of the objectives of this deliverable. It will also provide a background on the state of the art, the emulation tools used, and any assumptions made to complete the emulation/simulation.
- Chapter 3: Network Layer Emulator
  - This chapter goes into detail about the results gathered from initial Network layer testing. These results investigate the use of THz wireless links to assist or augment wired links based on the requirements defined in WP2.
- Chapter 4: Data-Link Layer Model
  - This chapter gives an overview on the possible operation of the DLL (Data-Link Layer) for THz wireless links based on the outputs of chapter 3.

1.3 Relationships with other deliverables

This deliverable is the third deliverable of work package 5. The work presented in this document relates to the following deliverables:

- D2.1.1 – Initial Requirements and Scenario Specifications
  - The scenarios studied in this deliverable serve to demonstrate the ability of TERAPOD technologies to achieve the use-cases outlined in D2.1.1.
- D5.1 – Initial PHY-Layer Model and Simulator
  - Outputs from PHY model and simulator are considered for their feasibility as inputs to the Data-Link Layer.

- D5.3 – Initial Data-Link Layer Model and Simulator
    - Outputs from D5.3 are assessed to be used as inputs to D5.5.
- D6.2.1 – Initial simulation demonstrator
    - A combination of the initial PHY (D5.1), DLL (D5.3), and NET layer (this deliverable) models and simulations will be considered as inputs into D6.3.

1.4 Contributors

The following partners have contributed to this deliverable:

- Sean Ahearne (Dell EMC)
- Niamh O'Mahony (Dell EMC)
- Noureddine Bounjah (TSSG)
- Saim Ghafoor (TSSG)
- Johannes Eckhardt (TUBS)
- Luis Gonzalez Guerrero (UCL)

1.5 Acronyms and abbreviations

OSI – Open **S**ystems **I**nterconnection

PHY – **Phy**sical Layer

DLL – **D**ata-**L**ink **L**ayer

NET – **Net**work Layer

TCP – Transmission **C**ontrol **P**rotocol

UDP – **U**ser **D**atagram **P**rotocol

SDN – Software **D**efined **N**etworking

KPI – **K**ey **P**erformance **I**ndicator

M/Gbps – **M**ega/**G**igabits **p**er **S**econd

SFP – **S**mall **F**orm-factor **P**luggable

LLDP – Link **L**ayer **D**iscovery **P**rotocol

ms – Milli-**s**econds

SNR – **S**ignal to **N**oise **R**atio

UTC-PD – **U**ni-**T**ravelling **C**arrier **P**hoto **D**iode

TOR – **T**op **O**f **R**ack

CRC - Cyclic **R**edundancy **C**heck

BER – Bit **E**rror **R**ate

FER – Frame **E**rror **R**ate

FEC – Forward **E**rror **C**orrection

N/ACK – No / **Ack**nowledgement received

MAC – Medium **A**ccess **C**ontrol

HOL – **H**ead **o**f **L**ine

2 Overview

2.1 Objectives

The objective of this deliverable is to determine the initial possibilities of terahertz (THz) wireless links within the datacentre. To determine these initial possibilities, the THz wireless links will be emulated at the Network Layer. At this initial stage, this document seeks to answer if THz wireless links can be used to replace wired optical links. The means using the wireless links in a point-to-point fashion much like traditional optical links. Wireless links used in this way can prove to be beneficial in a datacentre, as they can provide features such as link redundancy or aggregation with reduced cabling and maintenance efforts or can replace optical connections entirely.

Network layer testing is performed as an emulation. Virtual hosts and network switches are created and linked to each other, with data packets flowing between them in real-time. This method was chosen because network layer tools typically operate in this fashion. A benefit of this method is that the tools used for emulation support Software Defined Networking (SDN). Network devices that support SDN can be controlled in a centralized fashion by an SDN controller[1]. Doing this enables a controller to be topology-aware, meaning it knows the logical location of all network devices it controls, in relation to one another. This is useful as it allows the controller to make decisions based on events within the network, such as responding to link failures and prioritizing traffic types. This feature can be very effective for THz wireless links, as an SDN controller that can be made aware of a topology of wireless links can potentially use them to perform routes and features that are not possible with wired links (i.e. between devices that are not connected by a wire link). A limitation of emulation is that maximum performance is often dictated by the amount of processing power available to the emulator. In this deliverable, scenarios were studied to determine the ability of the THz links to implement point-to-point network features, not to determine the achieved performance statistics. The final scenario, however, does perform some calculations to determine realistic performance.

The information gathered from the network layer tests is passed to the data-link layer. Testing at the DLL and PHY layers is performed via simulation. This means behaviour and information generated at these layers is calculated step-by-step at a fraction of real-time. This method is more time consuming than emulation but can give accurate measurements for Key Performance Indicators (KPIs). Information, such as the packet size and frequency, is used to determine the best methods to use for transmission of the data over a wireless link. The DLL also uses data from the physical layer, such as signal power and noise, to further enhance the determination of transmission parameters. An example of the parameters of each layer and their KPIs is outlined below.

Layer	Parameter	KPIs
Network	Packet Size, Packet Rate, Route	Latency
Data-Link	Modulation Scheme, Queue Scheme, Media Access Control	Throughput
Physical	Output Power, Gain, Noise	Signal-to-Noise Ratio

Table 1: An example of parameters and KPIs for each layer

2.2 State of the art

In a modern datacentre, connections between hosts and racks are made using an array of network switches. These switches are most commonly arranged in what is known as a fat tree topology[2]. This topology is effective at distributing traffic load across a large-scale datacentre and allows for link aggregation and redundancy. Figure 2.1, below, depicts the typical network architecture found in a fat tree topology. Hosts on single rack connect to an edge switch, which is then connected to an aggregation layer of switches, followed by a core layer[3].

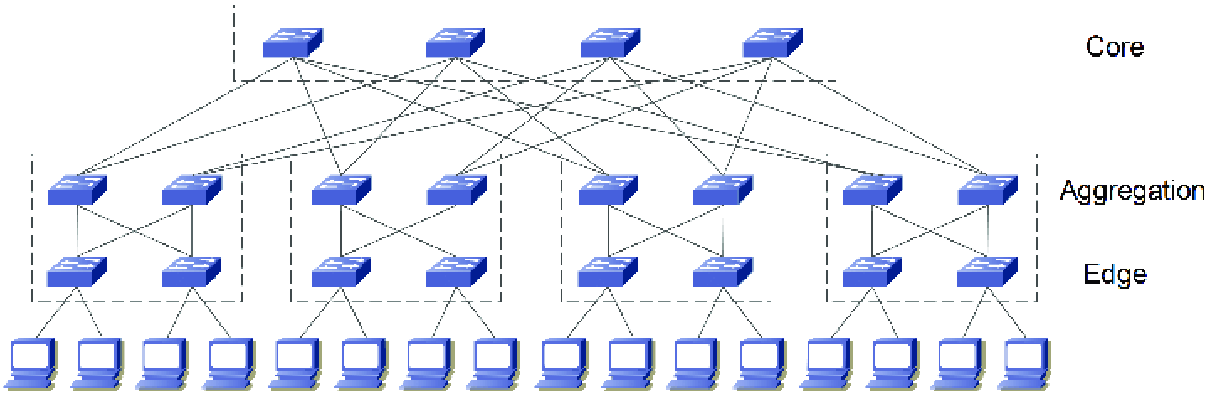


Figure 2.1: A typical fat-tree topology[3]

In terms of bandwidth, hosts connect to an edge switch using a Small Form-factor Pluggable (SFP+) optical transceiver operating at bitrates of approx. 10Gbps[4]. Links between edge and aggregation layer switches may also operate at 10Gbps or may use the Quad-SFP standard with bitrates of 40Gbps[5]. In a similar fashion, links between Core and Aggregation switches often operate at 40Gbps or will use the QSFP28 standard which operates at a bitrate of approx. 100Gbps[6].

For THz wireless links to be feasible in the datacentre, they must be capable of achieving these bitrates. Initial hardware tests of THz wireless links in a benchtop lab environment shows that a single THz wireless channel can achieve a bitrate of 20Gbps[7]. Using wireless channel aggregation, bitrates of 100Gbps are possible[8]. These bitrates show that THz links can compete with existing optical fibre links in a datacentre network and possibly exceed them in the future.

SDN is becoming a critical element for modern datacentre networks, as it allows for centralized and autonomous management of large-scale network fabrics[9]. It also enables features such as Network Function Virtualization (NFV), a feature which is of great importance to 5G and wireless networks, as it enables the virtualization of wireless hardware functions, thereby improving performance and reducing resource usage[10]. Centralizing the control plane into an SDN controller enables the controller to become fully aware of logical topology of both switches and hosts, using the Link Layer Discovery Protocol (LLDP)[11]. Being fully aware of the end-to-end network topology within a datacentre can allow the SDN controller to make more informed routing decisions than traditional switches and react quickly to network events.

2.3 Choice of Tools

The following lists the tools that have been used for the investigations described in this deliverable.

Network Layer

1. Mininet
    a. Network Layer Emulator[12]
2. Wireshark
    a. Packet capture & analytics tool[13]
3. iPerf(3)
    a. Traffic generator[14]
4. OpenDayLight
    a. SDN Controller[15]

Data-Link Layer

1. NS-3
    a. Data-Link Layer Simulator[16]

2.4 Assumptions

As the goal of this initial deliverable is primarily to ascertain the ability of THz wireless links to inter-operate with wired links, the bandwidths used in all scenarios, except the final one, are for reference only and not representative of real-world performance. Low bandwidths of 10-100 Mbps were chosen for several reasons:

- Different bandwidths for the THz and wired links will clearly determine when each is being used.
- Reduced computational resources required to perform emulation.
- Reduced amount of excess data captured during emulation.
- Low bitrates encourage variable packet lengths to be used during transmission, as opposed to the maximum permissible at high bitrates. This is beneficial data to the Data-Link Layer as it is a better representation of the variable types of traffic flows that can be expected in a datacentre.

In the final scenario in this deliverable, a maximum bandwidth of 10Gbps is chosen for both the THz and wired links as that is the current maximum bandwidth capable with real hardware between existing switch SFP+ optical transceivers and the THz wireless transmitter. The reason for this is that the THz transmitter requires an optical signal input at a wavelength of 1550 nanometres, with current network switch transceivers only supporting that wavelength at a maximum bitrate of 10Gbps. Calculations are completed in the final scenario that verify THz wireless links are capable of 10Gbps at the distances measured in Dell EMC's physical datacentre.

3 Network Layer Emulator

Overview and Scenarios

For the initial network emulation, three simple scenarios will be created to test the feasibility of using THz wireless links. These scenarios are based on the use cases and requirements defined in WP2. These scenarios relate to use cases UC-02A and UC-02B defined in D2.1. They use point-to-point links to perform link redundancy and traffic management for a wired network in order to replace or augment wired connectivity in the datacentre. Scenario 3.2 also relates to UC-01 as it uses real-world data to calculate the feasibility of these links in an existing datacentre.

3.1 Inputs

The inputs used to create these scenarios were the requirements set for THz links in TERAPOD WP2. It was specified in D2.1 that these links must be capable of matching or augmenting the performance or wired optical links. These scenarios show that THz links are can fulfil the specified requirements with redundancy, traffic aggregation/segregation, and load balancing.

3.2 Outputs

Outputs from these scenarios include the performance data gathered above, and raw packet captures of traffic generated during emulation. These packet capture files contain every individual packet created and transmitted/received on a per host basis. Information such as the time stamp, sequence number, MAC/IP address, packet size, protocol and more can be found in these capture files. This information is important to the Data-Link Layer and can be used to determine processes such as the optimal framing and queueing scheme to use.

3.3 Data Formats

As previously mentioned, the format of the data output from these scenarios is in the form of packet capture files (.pcap) created in Wireshark. These files can be processed for important DLL information, and can also be simply converted to plain-text de-limited data (.csv) if necessary to be processed externally.

3.4 Scenarios: Point to Point links

Objective: Reduce network congestion and/or downtime using THz wireless links.

3.4.1 Scenario 1: Reducing Downtime

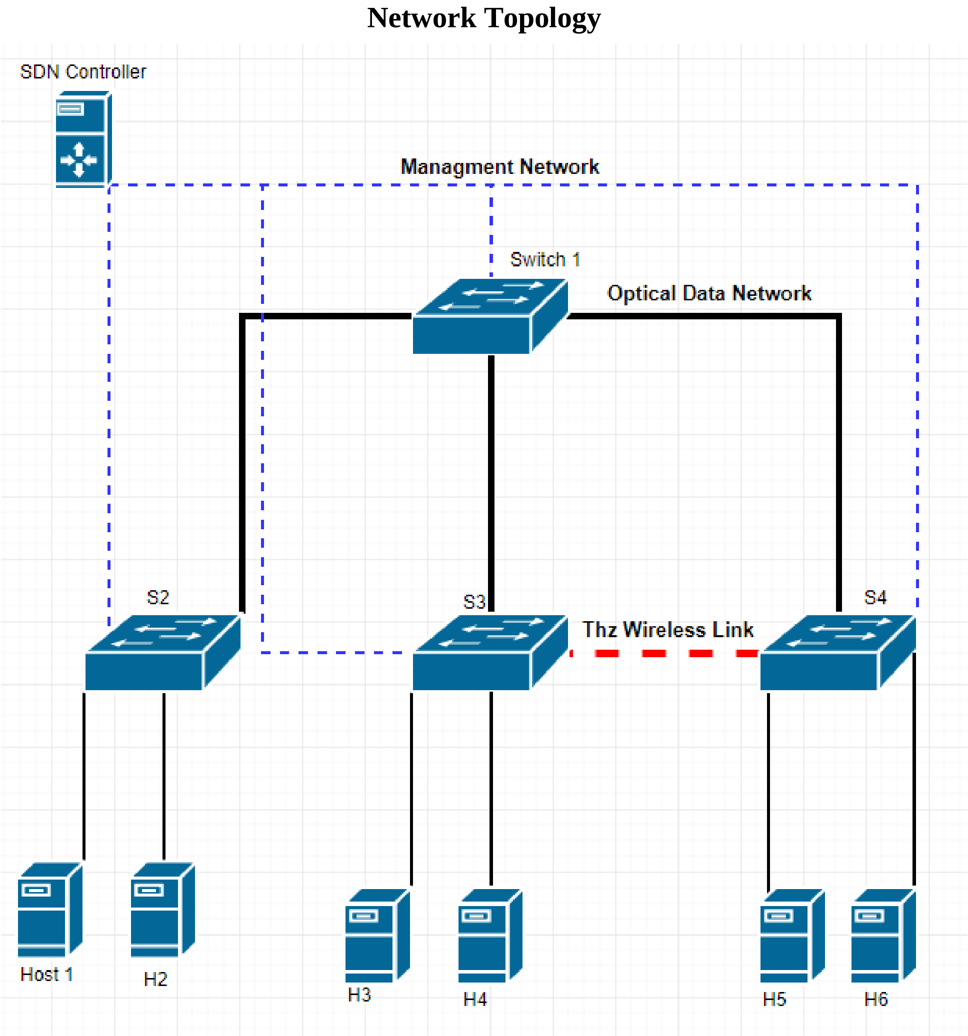


Figure 3.1: Network topology for Scenario 1 showing wired and wireless links

Events

1. A simple TCP stream of data will flow between H3 and H5.
2. This high bandwidth stream will represent a Virtual Machine migration between hosts.
3. During this stream, the optical link between Switch 1 and Switch 4 will fail.
4. This scenario requires the SDN controller to reconfigure the network topology to use the THz wireless link instead of the failed optical link.
5. It must do this in a timely fashion such that the TCP connection does not time-out or otherwise fail.
6. Ideally the SDN controller will also change back to using the optical link as the primary data path, should it be restored.

Link Characteristics

1. **All optical links**
   - 100Mbps Bandwidth
2. **THz Wireless Link**
   - 10Mbps Bandwidth

Timeline

1. A TCP stream will flow for 10.3 seconds from H5 (10.0.0.3) to H3 (10.0.0.2) on the optical network only to verify link characteristics.
2. In this capture, it occurs from packet No. 13 at time 45.74 seconds, to packet no. 113996 at time 56.09 seconds. Recorded bandwidth was 85.0 Mbps
3. A TCP stream will flow for 12.1 seconds from H5 to H3 on the THz wireless link only to verify link characteristics.
4. In this capture, it occurs from packet No. 114001 at time 75.38 seconds, to packet no. 129222 at time 87.56 seconds. Recorded bandwidth was 9.51 Mbps
5. A TCP stream will flow for 12.5 seconds from H5 to H3, beginning on the optical network. The link will then fail, with the TCP connection migrating to the THz link.
6. In this capture, the TCP stream begins at packet No. 129229 at time 117.67 seconds. The link failure event occurs at time 118.66 seconds (packet no. 138909), with connectivity restored at time 122.69 seconds. The stream finishes at packet no. 148450 at time 130.23 seconds. Recorded bandwidth was 11.16 Mbps
7. A TCP stream will flow for 10.1 seconds from H5 to H3, beginning on the THz link. The optical link will then be restored, with the TCP connection migrating to the optical link.
8. In this capture, the TCP stream begins at packet No. 148456 at time 158.99 seconds. The link restoration event occurs at time 162.35 seconds (packet no. 149808). The stream finishes at packet no. 2251584 at time 169.16 seconds. Recorded bandwidth was 55.6 Mbps

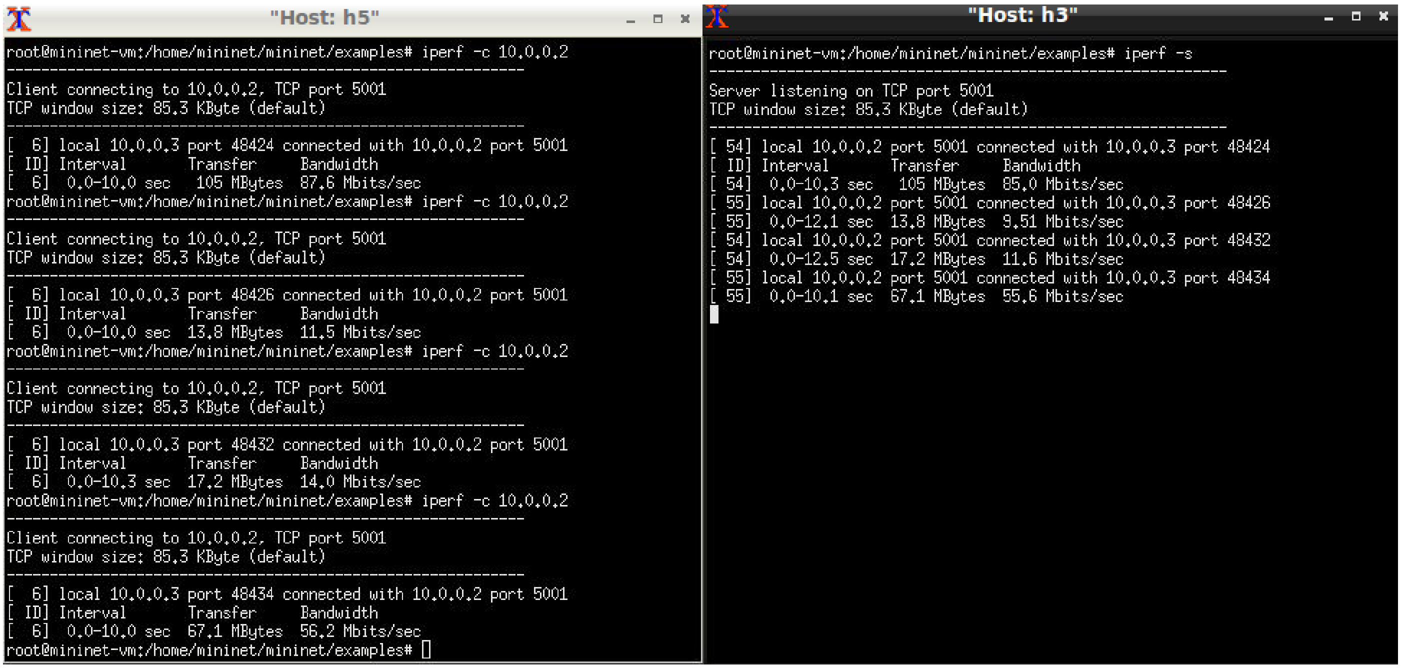


Figure 3.2: Network performance from H3 to H5, showing wired link bandwidth, wireless, and failover bandwidth. Note TCP connection does not time out.

3.4.2 Scenario 2.1: Reducing network congestion: Traffic segregation

1. Topology:
   - Same as 1

Event

1. A simple TCP stream of data will flow between H5 and H3
2. A simple UDP stream of data will also flow between H5 and H3
3. The simple TCP stream will represent a VM migration.
4. The simple UDP stream will represent a low-priority video stream.
5. The SDN controller will reconfigure the network such that it enables the TCP stream to use all bandwidth available on the optical link, with UDP traffic forwarded using the THz link.

Link Characteristics

1. All optical links:
   - Same as 1
2. THz Wireless Link
   - Same as 1
3. TCP Traffic:
   - TCP traffic will use Jumbo Frames to determine the effect of this on the data link layer.

Timeline

1. The SDN controller will be configured to forward TCP traffic from H5 to H3 over the optical network, and UDP traffic from H5 to H3 over the THz wireless link.
2. A TCP stream will flow for 11.4 seconds from H5 (10.0.0.5) to H3 (10.0.0.3)
3. In this capture, it occurs from packet 19 at time 28.17 seconds, to packet no. 13453 at time 39.57 seconds. Recorded bandwidth was 78.4 Mbps (Recorded by H3)
4. A UDP stream will flow for 11.2 seconds from H5 to H3. The stream will be set at a bandwidth of 100 Mbps, meaning packet loss should occur over the 10Mbps THz link.

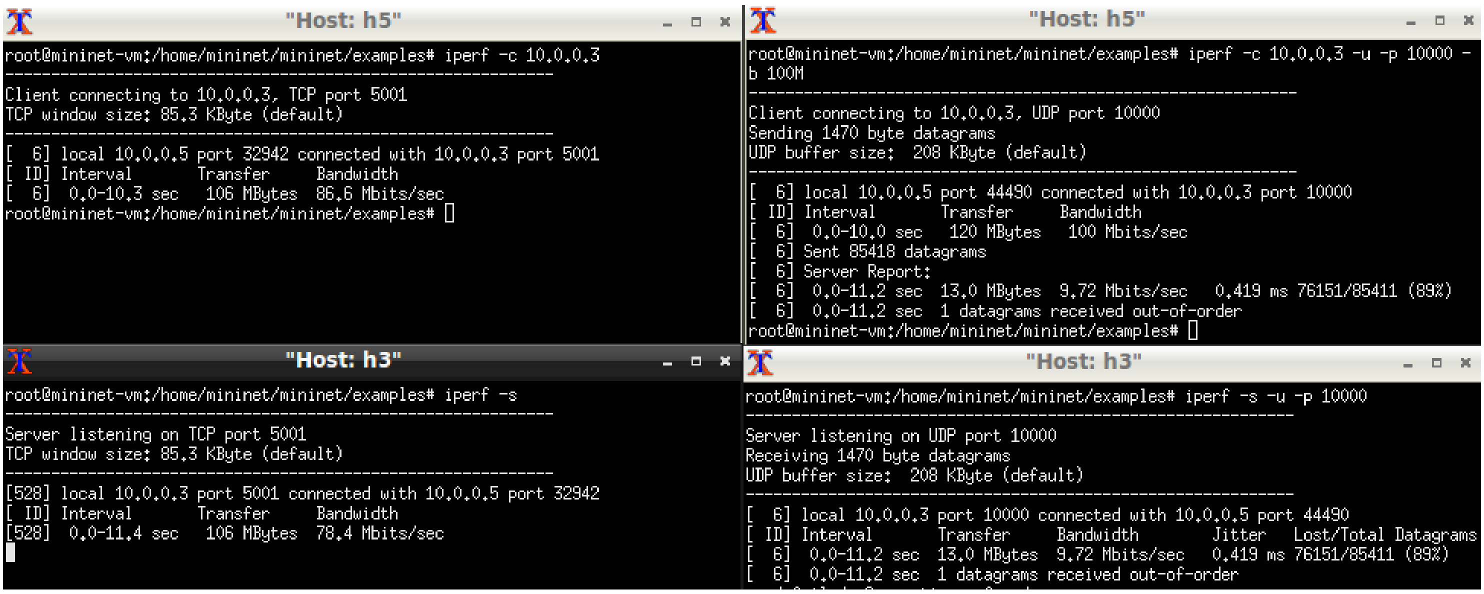


Figure 3.3: TCP and UDP connections between H5 and H3 respectively

5. In this capture, it occurs from packet No. 13457 at time 44.81 seconds, to packet no. 22733 at time 56.02 seconds. The result was a bandwidth of 9.72 Mbps, with a packet loss ratio of 89% (Recorded by H3)

3.4.3 Scenario 2.2: Reducing network congestion: Link aggregation

1. Topology:
    - Same as 2.1

Event

- Same sequence as A.2.1, except both traffic streams will occur at the same time to the same host to determine if the net bandwidth between hosts has increased.

Link Characteristics

1. All optical links:
    - Same as 2.1
2. THz wireless link
    - 50 Mbps
3. TCP traffic:
    - Same as 2.1

Timeline

1. The SDN controller will be configured to forward TCP traffic from H5 to H3 over the optical network, and UDP traffic from H5 to H3 over the THz Wireless link.
2. A TCP stream will flow for 11.4 seconds from H5 (10.0.0.5) to H3 (10.0.0.3)
3. Simultaneously, a UDP stream will flow for 10.2 seconds from H5 to H3. The stream will be set at a bandwidth of 100 Mbps, meaning packet loss is expected to occur over the 50Mbps THz Link.
4. In this capture, the TCP stream begins at packet No. 4 at time 4.45 seconds, the UDP stream begins at packet No. 165 at time 4.62 seconds. The TCP stream completes at packet No. 45550 at time 15.89 seconds, the UDP stream completes at packet No. 44546 at time 14.87 seconds.

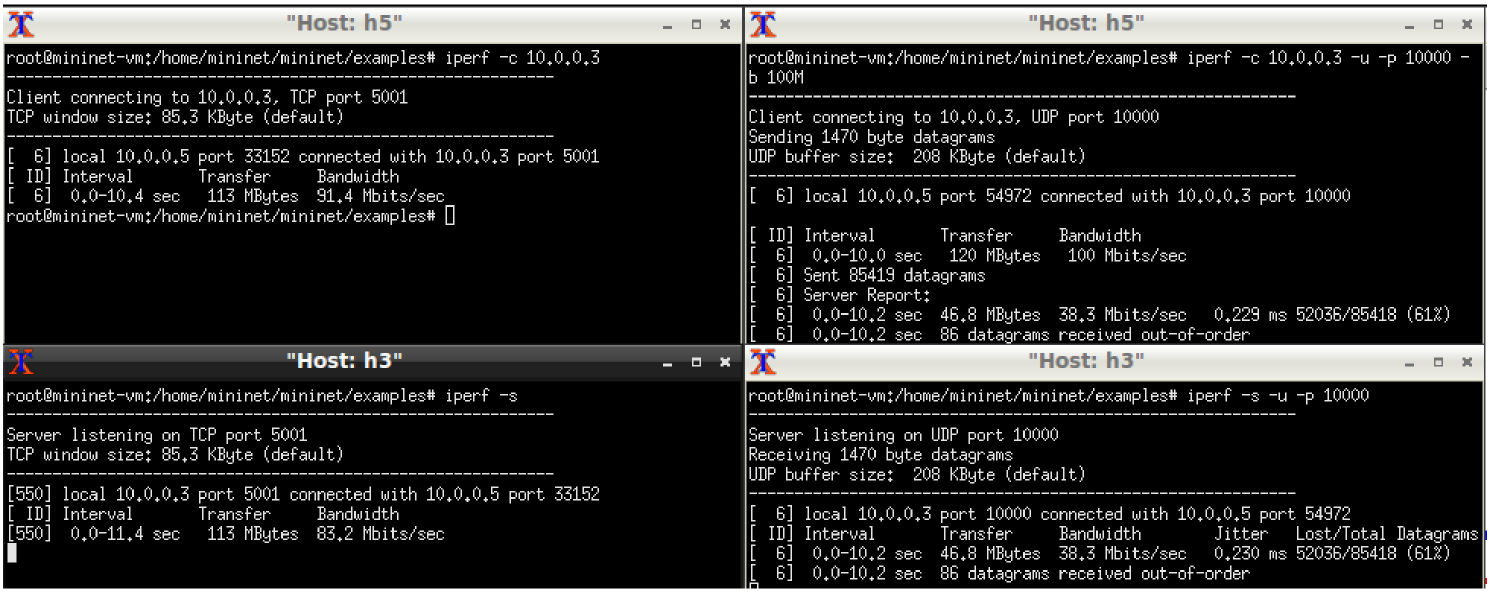


Figure 3.4: Communication between H3 and H5 with simultaneous TCP and UDP connections

5. The bandwidth recorded for the TCP stream was 83.2 Mbps. The bandwidth recoded for the UDP stream was 38.3 Mbps with a packet loss ratio of 61%.
6. We can confirm the aggregate bandwidth to host H3 exceeded 100 Mbps with packet capture analysis. The below graph (Figure 3.5) shows a maximum bitrate of approx. $1.3x10^8$ bits per second (130 Mbps), confirming both links were effectively utilized between switches.

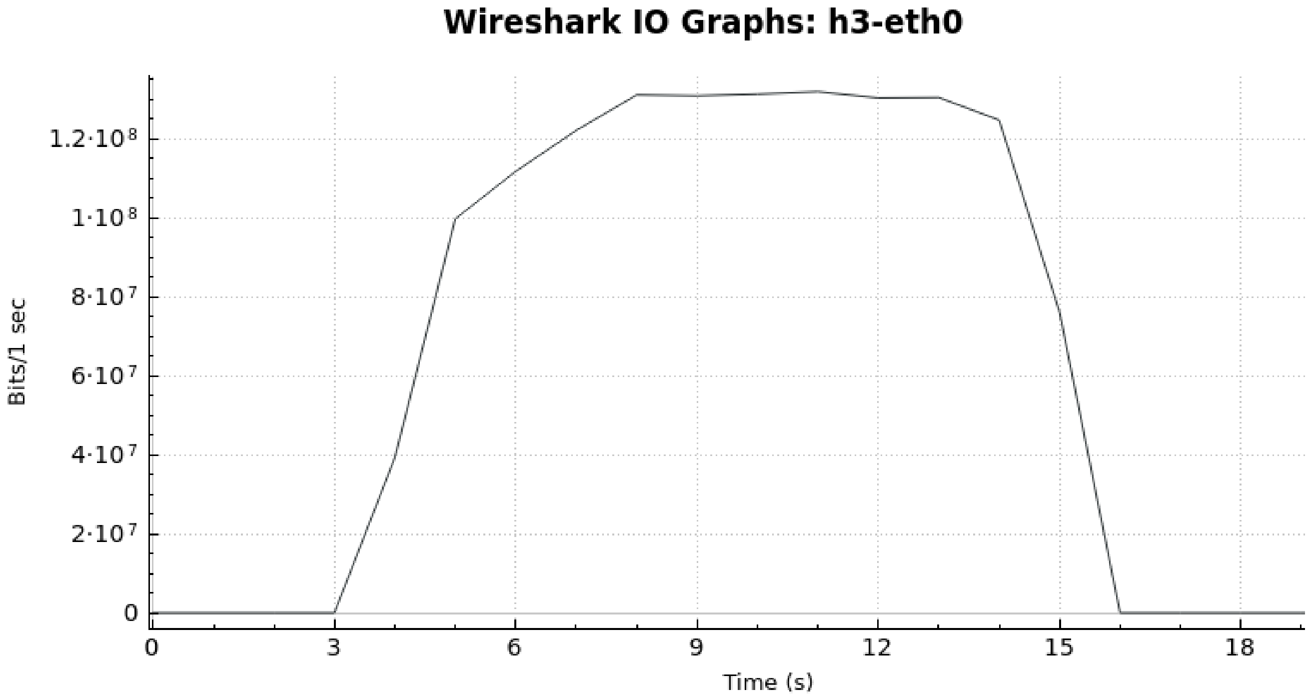


Figure 3.5 Throughput graph of total traffic received by H3 (~130Mbps)

3.4.4 Scenario 3.1: Multi Hop Routing

Topology

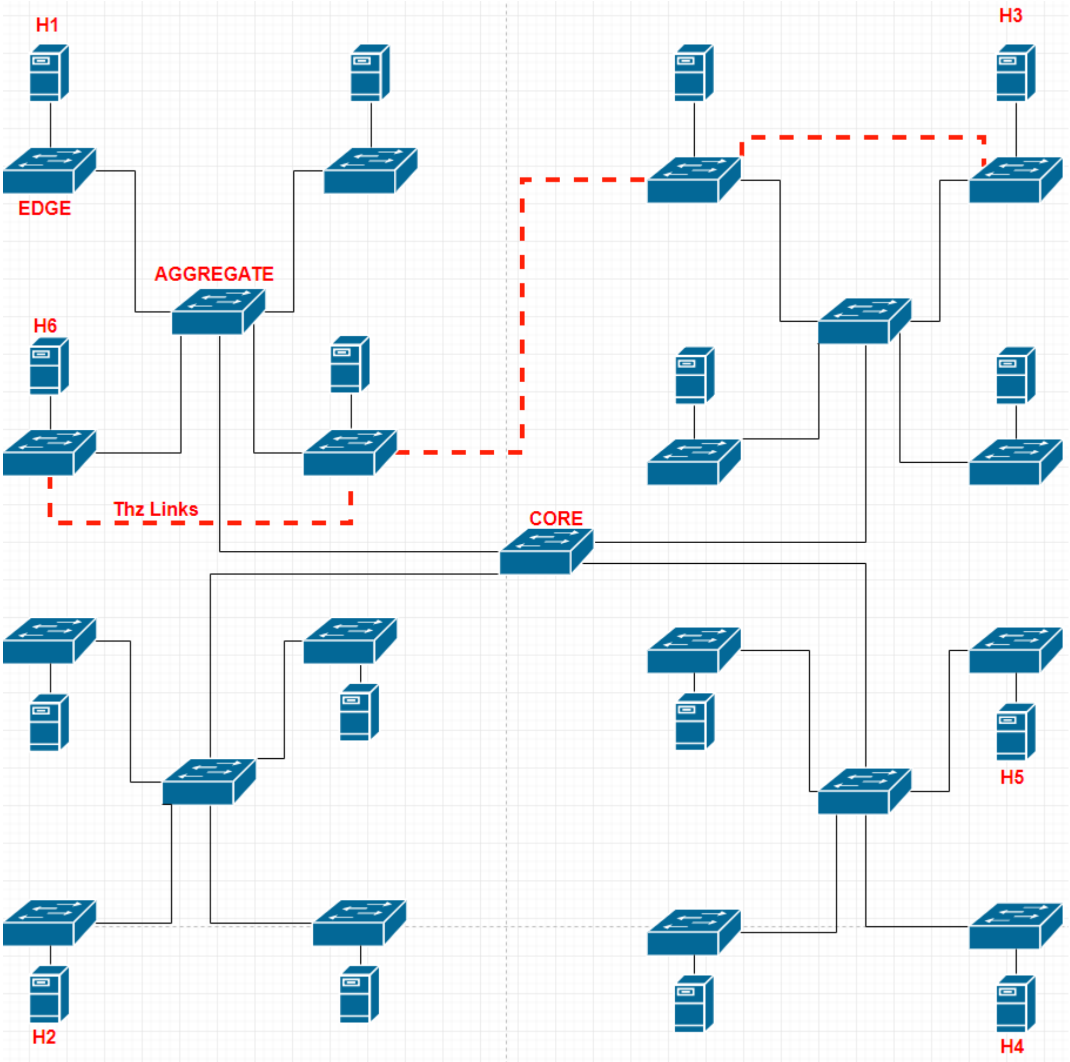


Figure 3.5: Network topology for multi-hop scenarios with three wireless links

The topology for this scenario, illustrated in Figure 3.6, represents a collapsed FAT tree model that could potentially be found in a live datacentre. It is comprised of a 4x4 grid of Top of Rack switches (edge), with 4 aggregation switches, followed by a core switch. This comprises a total of 21 switches, and 16 hosts. An SDN controller (not pictured) is controlling all the switches via the OpenFlow protocol. Multiple point-to-point THz wireless links have been put in place to create a connection between host H6 and H3.

Event

1. This event begins with the THz wireless links unconfigured (i.e. not active for data transmission).
2. A constant high bandwidth UDP stream will flow between H1 and H4, causing the link from the aggregate switch to the core to become congested.
3. H6 and H3 will attempt to communicate using the TCP protocol (simulating a VM migration).
4. Once the TCP connection terminates, H6 will then also begin a high bandwidth UDP stream to H3, with the packet loss and jitter characteristics of both UDP streams measured.
5. The SDN controller will then configure the THz wireless links to allow H3 to communicate with H6 without using the core link, to avoid network congestion.

6. H3 and H6 will communicate using TCP once again, with the bandwidth results measured.
7. Once the TCP connection terminates, H6 will then also begin a high bandwidth UDP stream to H3, with the packet loss and jitter characteristics of both UDP streams measured once more

Link Characteristics

1. Optical links:
    - Aggregate to Core Links: 100Mbps
2. THz Wireless Links
    - 100Mbps

Timeline

1. A UDP stream of 100Mbps will flow between H1 and H4 for the duration of the TCP stream.
2. A TCP stream will flow for 13.3 seconds from H6 (10.0.0.6) to H3 (10.0.0.3)
3. In this capture, the TCP stream begins at packet No. 4 at time 13.71 seconds. The TCP stream completes at packet No. 1795 at time 28.30 seconds. The bandwidth recorded for the TCP stream was 1.5 Mbps.

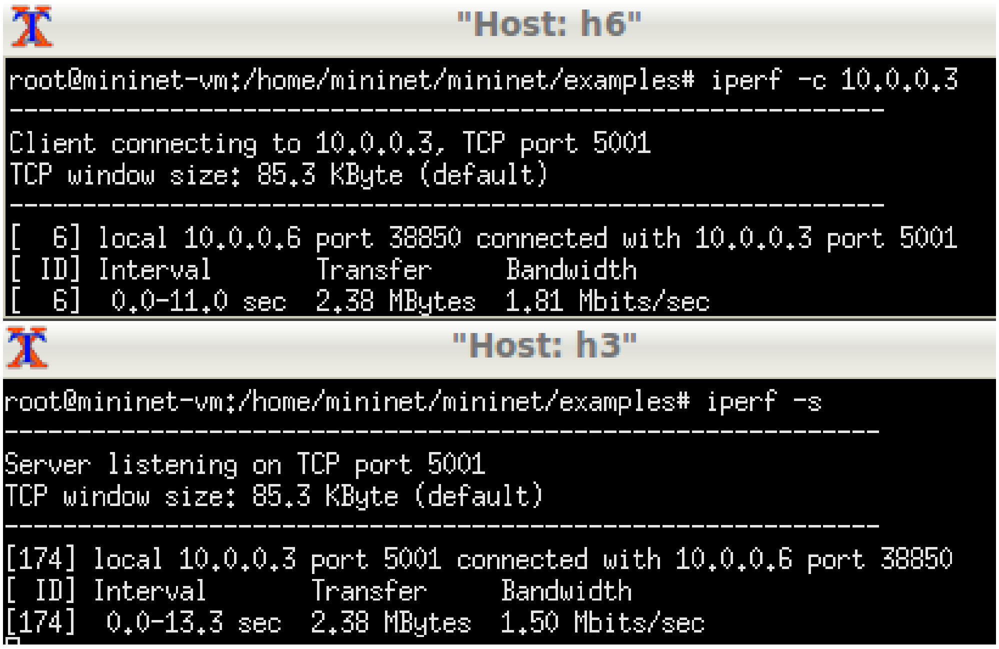


Figure 3.6: TCP Performance between H6 and H3 during Core traffic overload

4. The UDP stream from H1 to H4 was then terminated,
5. Two simultaneous UDP streams were then created from H1 to H4, and H6 to H3.
6. Both streams were measured over 10 seconds in 1 second intervals to determine their jitter and packet loss characteristics.
7. The UDP stream between H1 and H4 completed with an average bandwidth of 80.19 Mbps, jitter of 2.42 milliseconds, and an average packet loss of 11.91% over 10 seconds.

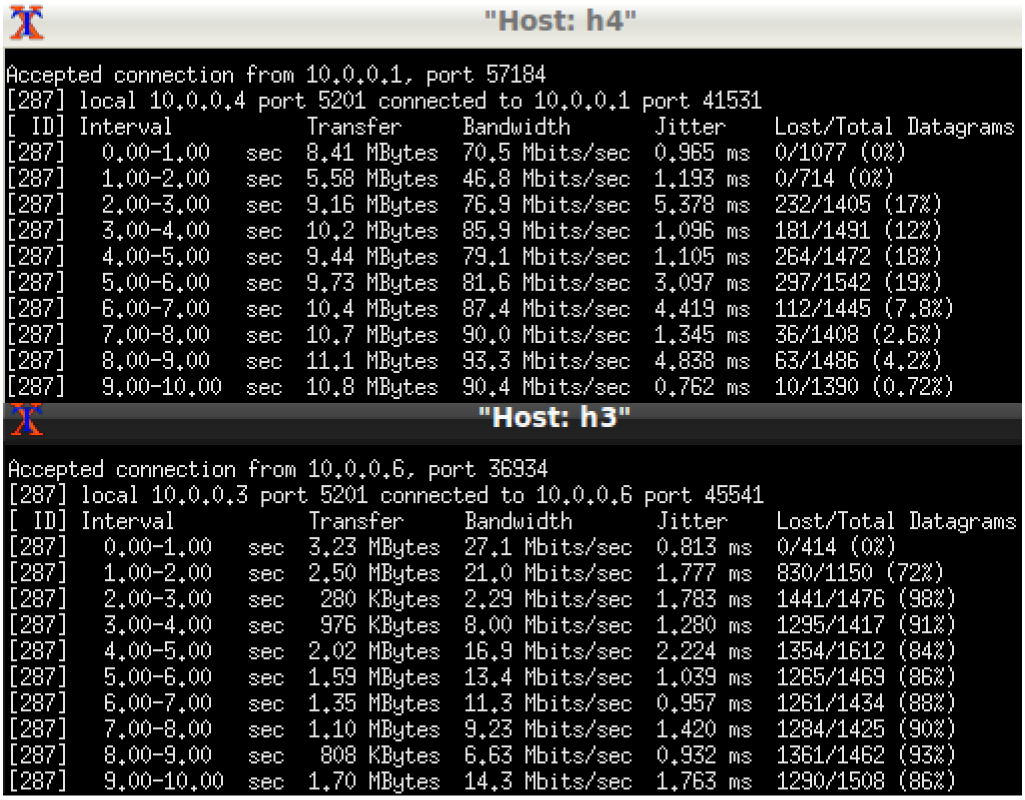


Figure 3.7: UDP Performance between H6 and H3 during Core traffic overload

8. The UDP stream between H6 and H3 completed with an average bandwidth of 13.02 Mbps, jitter of 1.40 milliseconds and an average packet loss of 78.8% over 10 seconds.
9. The SDN controller was then configured to utilise multiple THz wireless links to allow a second network path from H6 to H3.
10. A UDP stream of 100Mbps will flow between H1 and H4 for the duration of the TCP stream once more.
11. A TCP stream will flow for 11.4 seconds from H6 (10.0.0.6) to H3 (10.0.0.3)
12. In this capture, the TCP stream begins at packet No. 4 at time 12.06 seconds. The TCP stream completes at packet No. 11974 at time 23.50 seconds. The bandwidth recorded for the TCP stream was 89.9 Mbps.

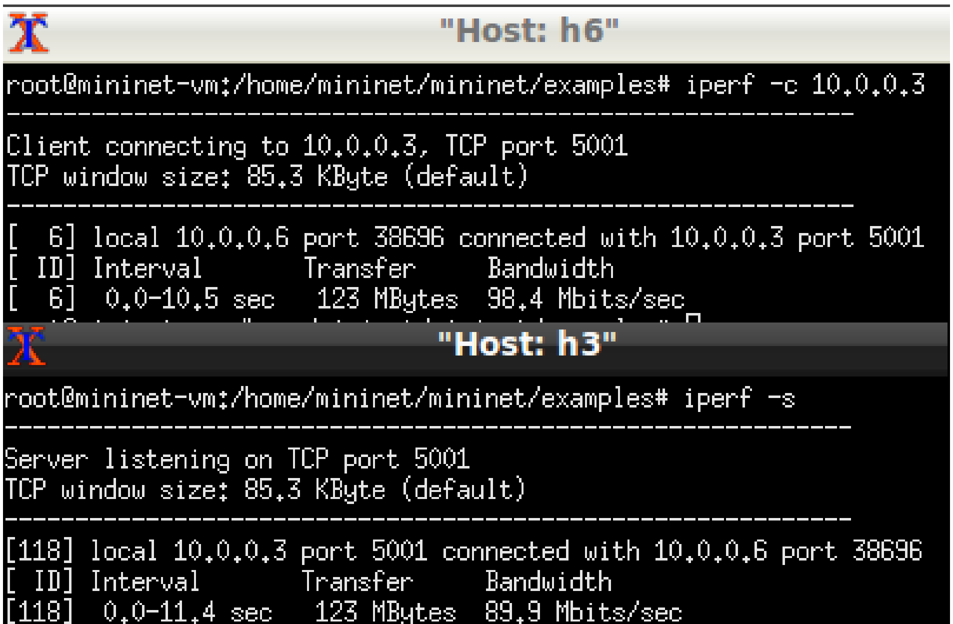


Figure 3.8: TCP performance between H6 and H3 using THz routing

13. The UDP stream from H1 to H4 was then terminated,
14. Two simultaneous UDP streams were then created from H1 to H4, and H6 to H3.
15. Both streams were measured over 10 seconds in 1 second intervals to determine their jitter and packet loss characteristics.
16. The UDP stream between H1 and H4 completed with an average bandwidth 93.78 Mbps, jitter of 0.549 milliseconds, and an average packet loss of 0.07% over 10 seconds.
17. The UDP stream between H6 and H3 completed with an average bandwidth 93.69 Mbps, jitter of 0.57 milliseconds, and an average packet loss of 0% over 10 seconds.
18. This shows an overall packet loss reduction of 45.36% for both flows to almost zero when compared to traffic offloading, and an average 241.38% decrease in jitter characteristics.

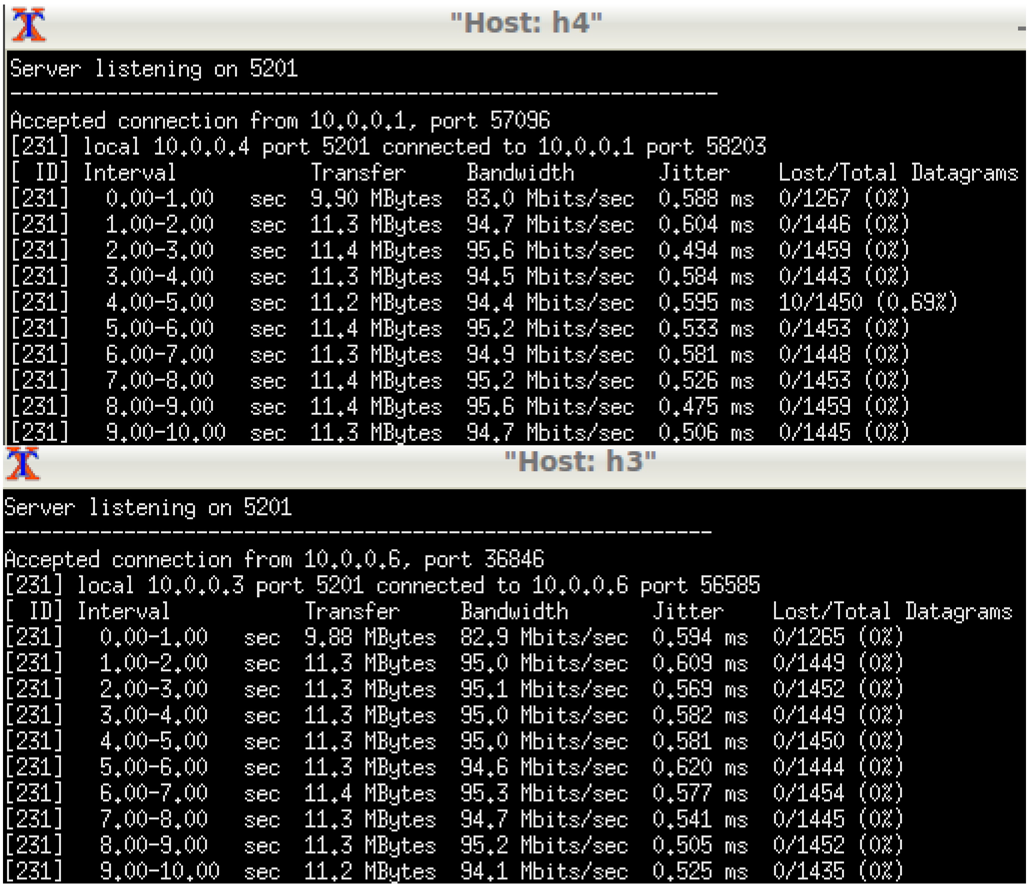


Figure 3.9: UDP performance between H6 and H3 using Thz routing

3.4.5 Scenario 3.2: Multi-hop realistic performance evaluation

Topology

- **Same as 3.1**

Event

- **Same as 3.1**

Link Characteristics

1. Optical links:
   - Aggregate to Core Links: 10Gbps
2. THz wireless links:

In this scenario, the THz wireless links will be given an estimated performance based on real-world measurements of existing hardware. This scenario will use three THz wireless links in the same topology layout as the previous scenario. In this scenario, a distance between each link in metres will be given. This value is based on the physical measurements taken of Dell EMC's datacentre. The links are modelled from left to right are:

- Link 1: 1.909m
- Link 2: 4.489m
- Link 3: 2.112m

Using these values, we can calculate the signal-to-noise ratio (SNR) per bit. This value determines if the bits transmitted are still visible at the receiving end of the wireless link. It can be calculated in conjunction with the following known hardware values:

	Value	Unit	Comments
UTC power	-15	dBm	TERAPOD UTC @ 300 GHz
Tx Lens antenna gain	≈43	dBi	Based on [17]
Transmission loss	95	dB	4.5 m @ 300 GHz (absorption negligible)
Rx Lens antenna gain	≈43	dBi	Based on [17]
Received power	-24	dBm	
IF power	-32	dBm	8 dB conversion loss
IF equivalent noise floor	-164	dBm	NF=10 dBm
E_b/N_0	32	dB	SNR per bit supports data rate: 10 Gbit/s

Table 2: THz Hardware Specifications

SNR per bit (E_b/N_0)

Data rate: 10 Gbit/s

E_b/N_0=IF power – (IF equivalent noise floor + 10*$\log_{10}$(Bit rate))

An E_b/N_0 of 32dB means that the bit-error rates (BER) required by the 10 Gigabit Ethernet standard of 10^{-9} or better are supported at this distance using all common data transmission schemes[17]. It can be inferred that it also supports this bitrate at shorter distances.

Timeline

1. A UDP stream of 10Gbps will flow between H1 and H4 for the duration of the TCP stream.
2. A TCP stream will flow for 10.04 seconds from H6 (10.0.0.6) to H3 (10.0.0.3)
3. In this capture, the TCP stream begins at packet No. 2 at time 1.95 seconds. The TCP stream completes at packet No. 138920 at time 12.11 seconds. The bandwidth recorded for the TCP stream was 2.11 Gbps.

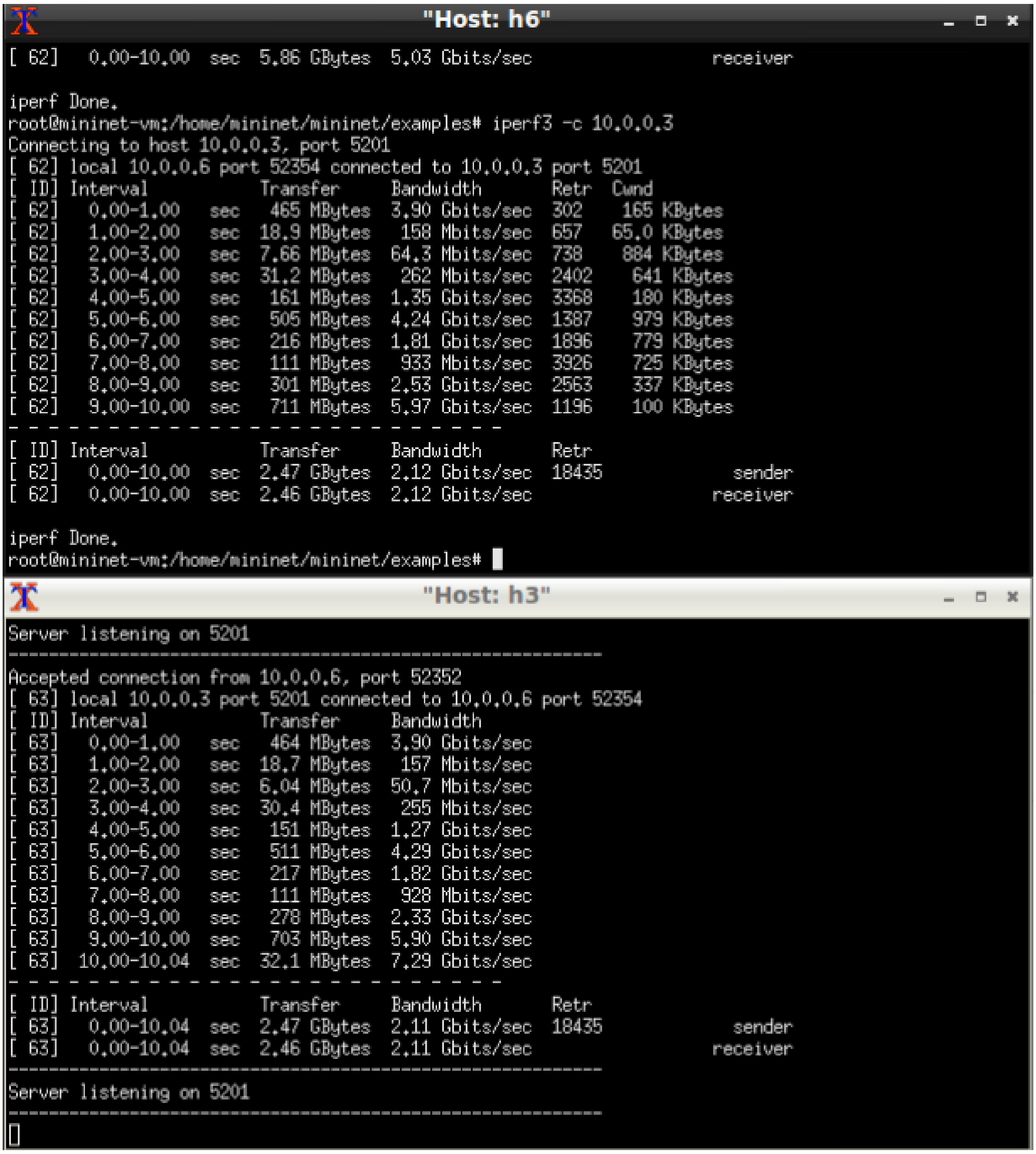
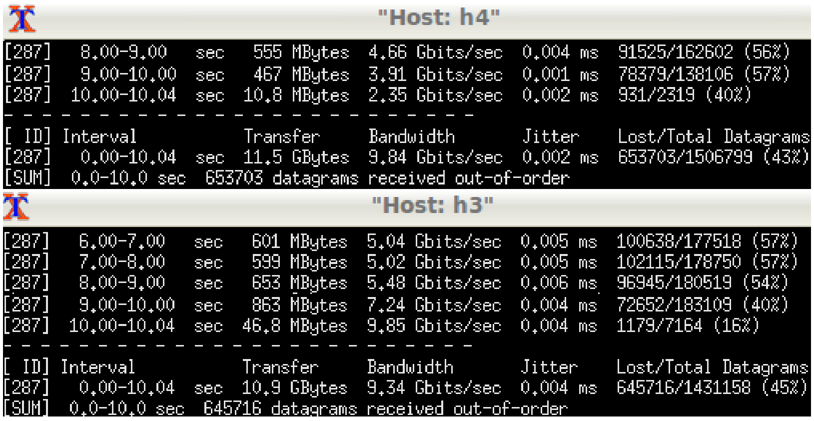


Figure 3.10: TCP Performance between hosts during core overload

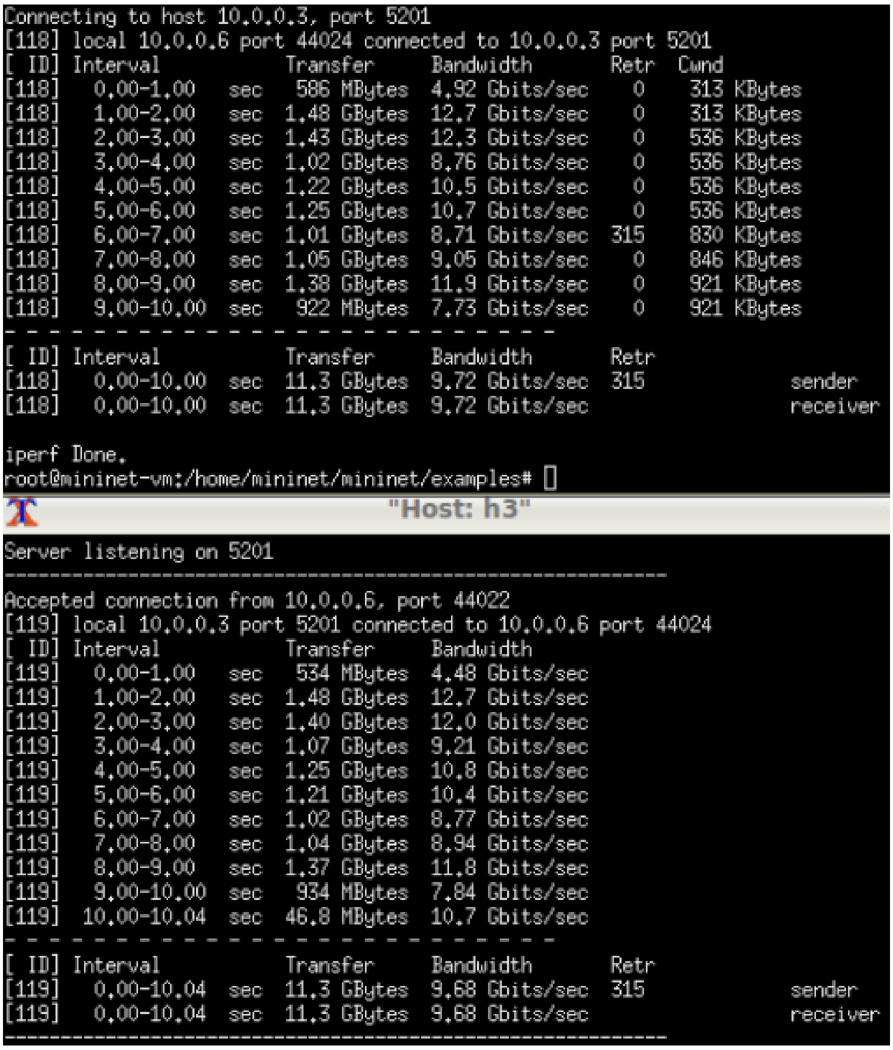
4. The UDP stream from H1 to H4 was then terminated,
5. Two simultaneous UDP streams were then created from H1 to H4, and H6 to H3.
6. Both streams were measured over 10 seconds to determine their packet loss characteristics.
7. The UDP stream between H1 and H4 completed with an average bandwidth of 9.84 Gbps, and an average packet loss of 43% over 10 seconds.
8. The UDP stream between H6 and H3 completed with an average bandwidth of 9.34 Gbps, and an average packet loss of 45% over 10 seconds.



```
"Host: h4"
[287]   8.00-9.00   sec   555 MBytes  4.66 Gbits/sec  0.004 ms  91525/162602 (56%)
[287]   9.00-10.00  sec   467 MBytes  3.91 Gbits/sec  0.001 ms  78379/138106 (57%)
[287]  10.00-10.04  sec  10.8 MBytes  2.35 Gbits/sec  0.002 ms  931/2319 (40%)
- - - - - - - - - - - - - - - - - - - - - - - - -
[ ID] Interval           Transfer     Bandwidth       Jitter    Lost/Total Datagrams
[287]   0.00-10.04  sec  11.5 GBytes  9.84 Gbits/sec  0.002 ms  653703/1506799 (43%)
[SUM]  0.0-10.0 sec  653703 datagrams received out-of-order
"Host: h3"
[287]   6.00-7.00   sec   601 MBytes  5.04 Gbits/sec  0.005 ms  100638/177518 (57%)
[287]   7.00-8.00   sec   599 MBytes  5.02 Gbits/sec  0.005 ms  102115/178750 (57%)
[287]   8.00-9.00   sec   653 MBytes  5.48 Gbits/sec  0.006 ms  96945/180519 (54%)
[287]   9.00-10.00  sec   863 MBytes  7.24 Gbits/sec  0.004 ms  72652/183109 (40%)
[287]  10.00-10.04  sec  46.8 MBytes  9.85 Gbits/sec  0.004 ms  1179/7164 (16%)
- - - - - - - - - - - - - - - - - - - - - - - - -
[ ID] Interval           Transfer     Bandwidth       Jitter    Lost/Total Datagrams
[287]   0.00-10.04  sec  10.9 GBytes  9.34 Gbits/sec  0.004 ms  645716/1431158 (45%)
[SUM]  0.0-10.0 sec  645716 datagrams received out-of-order
```

Figure 3.11: UDP Performance between hosts during overload

9. The SDN controller was then configured to utilise multiple THz wireless links to allow a second network path from H6 to H3.
10. A UDP stream of 10 Gbps will flow between H1 and H4 for the duration of the TCP stream once more.
11. A TCP stream will flow for 10.04 seconds from H6 (10.0.0.6) to H3 (10.0.0.3)
12. In this capture, the TCP stream begins at packet No. 3 at time 5.44 seconds. The TCP stream completes at packet No. 311236 at time 15.84 seconds. The bandwidth recorded for the TCP stream was 9.68 Gbps.



```
Connecting to host 10.0.0.3, port 5201
[118] local 10.0.0.6 port 44024 connected to 10.0.0.3 port 5201
[ ID] Interval           Transfer     Bandwidth       Retr  Cwnd
[118]   0.00-1.00   sec   586 MBytes  4.92 Gbits/sec    0    313 KBytes
[118]   1.00-2.00   sec  1.48 GBytes  12.7 Gbits/sec    0    313 KBytes
[118]   2.00-3.00   sec  1.43 GBytes  12.3 Gbits/sec    0    536 KBytes
[118]   3.00-4.00   sec  1.02 GBytes  8.76 Gbits/sec    0    536 KBytes
[118]   4.00-5.00   sec  1.22 GBytes  10.5 Gbits/sec    0    536 KBytes
[118]   5.00-6.00   sec  1.25 GBytes  10.7 Gbits/sec    0    536 KBytes
[118]   6.00-7.00   sec  1.01 GBytes  8.71 Gbits/sec  315    830 KBytes
[118]   7.00-8.00   sec  1.05 GBytes  9.05 Gbits/sec    0    846 KBytes
[118]   8.00-9.00   sec  1.38 GBytes  11.9 Gbits/sec    0    921 KBytes
[118]   9.00-10.00  sec   922 MBytes  7.73 Gbits/sec    0    921 KBytes
- - - - - - - - - - - - - - - - - - - - - - - - -
[ ID] Interval           Transfer     Bandwidth       Retr
[118]   0.00-10.00  sec  11.3 GBytes  9.72 Gbits/sec  315             sender
[118]   0.00-10.00  sec  11.3 GBytes  9.72 Gbits/sec                  receiver

iperf Done.
root@mininet-vm:/home/mininet/mininet/examples# []
"Host: h3"
Server listening on 5201
-----------------------------------------------------------
Accepted connection from 10.0.0.6, port 44022
[119] local 10.0.0.3 port 5201 connected to 10.0.0.6 port 44024
[ ID] Interval           Transfer     Bandwidth
[119]   0.00-1.00   sec   534 MBytes  4.48 Gbits/sec
[119]   1.00-2.00   sec  1.48 GBytes  12.7 Gbits/sec
[119]   2.00-3.00   sec  1.40 GBytes  12.0 Gbits/sec
[119]   3.00-4.00   sec  1.07 GBytes  9.21 Gbits/sec
[119]   4.00-5.00   sec  1.25 GBytes  10.8 Gbits/sec
[119]   5.00-6.00   sec  1.21 GBytes  10.4 Gbits/sec
[119]   6.00-7.00   sec  1.02 GBytes  8.77 Gbits/sec
[119]   7.00-8.00   sec  1.04 GBytes  8.94 Gbits/sec
[119]   8.00-9.00   sec  1.37 GBytes  11.8 Gbits/sec
[119]   9.00-10.00  sec   934 MBytes  7.84 Gbits/sec
[119]  10.00-10.04  sec  46.8 MBytes  10.7 Gbits/sec
- - - - - - - - - - - - - - - - - - - - - - - - -
[ ID] Interval           Transfer     Bandwidth       Retr
[119]   0.00-10.04  sec  11.3 GBytes  9.68 Gbits/sec  315             sender
[119]   0.00-10.04  sec  11.3 GBytes  9.68 Gbits/sec                  receiver
-----------------------------------------------------------
```

Figure 3.12: TCP performance during core overload using THz routes

13. The UDP stream from H1 to H4 was then terminated,
14. Two simultaneous UDP streams were then created from H1 to H4, and H6 to H3.
15. Both streams were measured over 10 seconds in 1 second intervals to determine their jitter and packet loss characteristics.
16. The UDP stream between H1 and H4 completed with an average bandwidth 4.57 Gbps, and an average packet loss of 0.063% over 10 seconds.
17. The UDP stream between H6 and H3 completed with an average bandwidth 6.85 Gbps, and an average packet loss of 0.3% over 10 seconds.

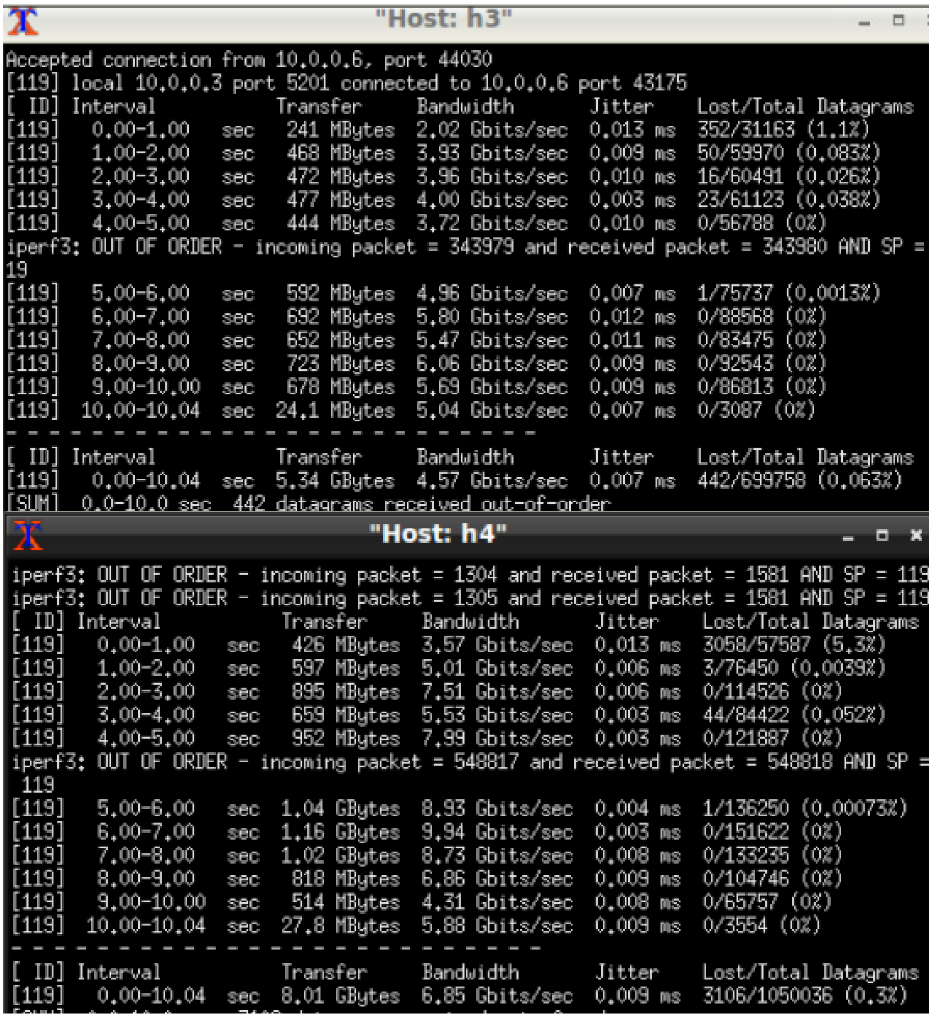


Figure 3.13: UDP performance using THz offloading

Note: UDP performance in this scenario is below 10 Gbps due to CPU limitations within the emulator. TCP packets were typically at their maximum permissible size of 64 Kilobytes. UDP packets have a maximum size of 8 Kilobytes. This increases the number of UDP packets that need to be processed per second, causing a loss of bandwidth due to the CPU being unable to process that many individual packets at that rate. In this case, we show that the THz links are being utilized for the UDP demonstration via the packet loss percentage value.

4 Data-Link Layer Model

4.1 Overview

The Data Link Layer (DLL) is the layer between the networking layer of the datacentre and the physical layer, this layer deals with high data rate traffic with specific requirements.

Fig 4.1 shows the data link layer architecture using blocks and highlighting the main interfaces of the DLL with the physical and networking layer. The main functionalities of the DLL are:

- Medium Access Control (MAC) techniques and transmission scheduling: In a network topology, nodes should efficiently use the shared medium. We assume a balanced traffic usage between nodes within the datacentre for this deliverable, with nodes having equal opportunity for medium access.
- Nodes discovery and synchronization: For a distributed architecture, a node should collect information from other nodes in the network. Timing for node synchronization is very important and reference times should be shared between all nodes which are, for example, the starting time of discovery procedure, and link establishment. We describe in this deliverable a technique for node discovery using forward backward message passing.
- Error Control: Techniques to protect frames from errors. It includes messages and techniques to insert protection bits such as CRC (Cyclic Redundancy Check) field, and command messages to reduce or increase frame length and detailed reports of frame losses.
- Frame construction: Networking layer is agnostic about DLL and physical layer; the DLL should consider fluctuations in transmission by selecting the appropriate frame length in order to reduce losses and strengthen the transmission link. Techniques used for frame construction are packet segmentation into short blocks and block aggregation.
- Buffering: A generated frame waits at the DLL buffer for future transmission or re-transmission. Buffer capacity depends on transceiver technology, the DLL should avoid buffer overload and increased queue times.

Interfaces between data link layer and other layers include:

- Incoming and outgoing data: Packets from/to network and frames from/to physical layer.
- Incoming and outgoing control message: Messages during the link establishment and during the communication phase, message exchange should be optimized to reduce link establishment time and message overhead.
- Incoming and outgoing measurement reports: Measurements are very important for DLL decisions, and reflecting the status of each layer.
- Antenna module: The DLL is responsible for link establishment. If the antenna is directional then DLL should send a command to physical layer in order to steer the beam toward the destination node.

The aim for DLL is to adapt incoming packets from the datacentre created in chapter 3 to the THz medium.

4.2 DLL modules and functionalities

4.2.1 General architecture

The block diagram of the data link layer is described figure 4.1. It includes DLL modules, links between modules, and exchanged messages and flows. Interfaces with the networking and physical layer are also considered.

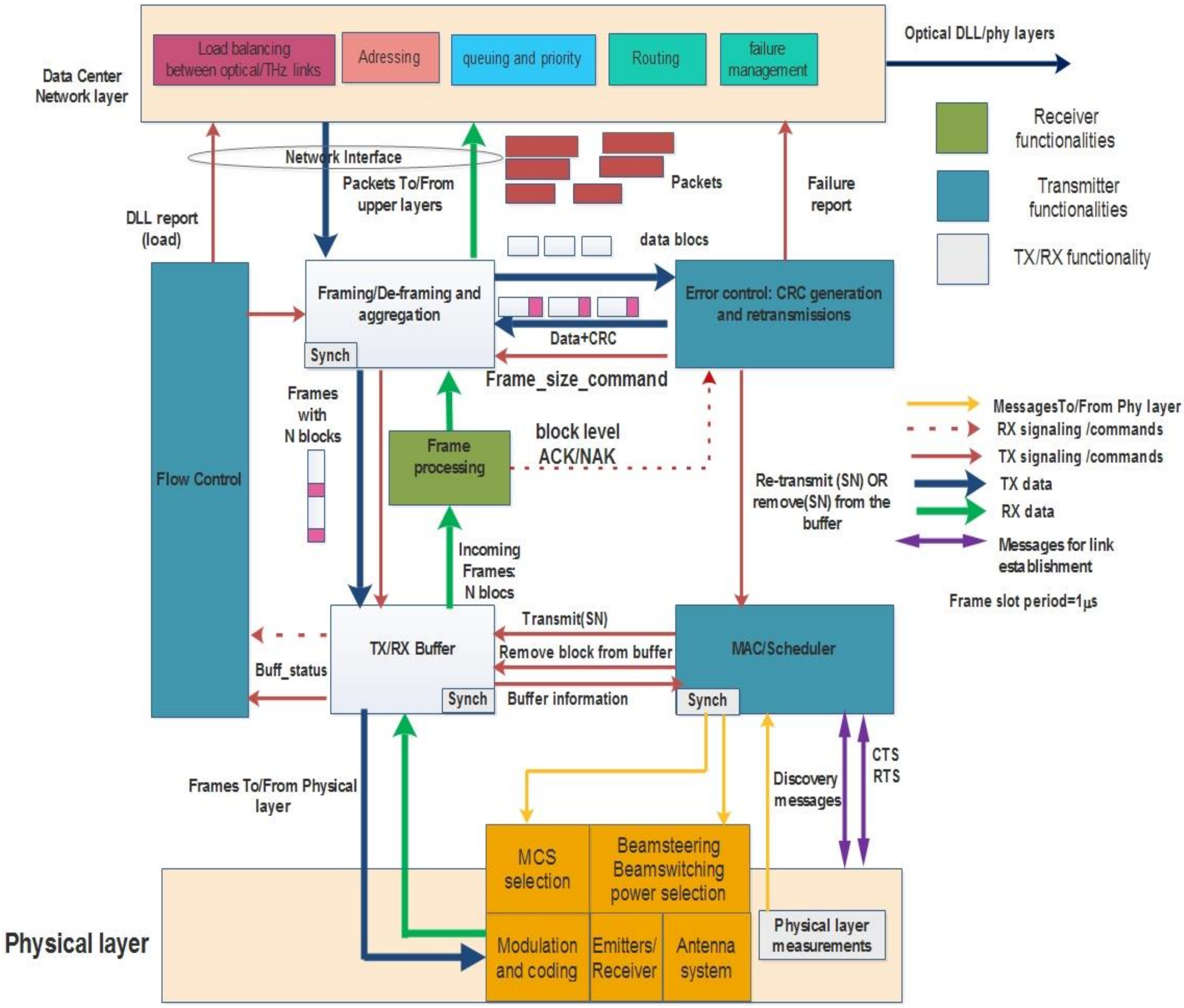


Figure 4.1: General DLL architecture: Main modules, interface and message flows

4.2.2 Network discovery strategy

We assume a distributed architecture of nodes for this deliverable, where each node updates its information about the network and schedules its transmission. Before transmission starts, nodes gather information from the network to construct their connection table and calculate their resource requirements in a coordinated manner.

A network discovery phase is required for each node to update its information regarding the network topology and if new changes in node configurations occur.

4.2.3 Framing module

Based on data collected from MININET simulation we can model the packets arriving to the DLL using the Poisson process. Two parameters should be calculated from packet capture files: mean-interarrival rate and mean packet size.

The packet source and destination addresses should be also extracted by the DLL to generate MAC addresses.

The framing module is responsible for frame creation using an adaptive size based on packets received from the network layer.

- Frame structure:
    - Frame header
        - Frame sequence number SN: a unique number to identify frame
        - Frame synchronization marker: special sequence for frame detection
        - Source address
        - Source position
        - Destination address
        - Number of blocks per frame: for frame aggregation, a frame can carry more than one block with its CRC subfield
    - Block unit
        - CRC per block
        - Block length
        - Each block contains data from a packet and a header containing the packet reference used at the receiver for packet reassembly procedure.
    - Frame
        - Frame length
        - Frame rate entering the buffer:

The frame size should be selected to reduce Frame Error Rate (FER) and increase throughput, the FER can be calculated based on acknowledgement messages received.

4.2.4 Transmitter and receiver Buffer:

As described in figure 4.1, the buffer receives blocks and frames from the framing module for transmission and from the physical layer for processing and conversion into packets.

Frames are labelled with a sequence number SN, the possible statuses of a frame in the DLL are:

- Transmit if the frame is at the head of the line (HOL).
- Frame waiting in the queue.
- Discarded from the buffer list if the maximum number of re-transmissions is reached, or the frame was successfully transmitted (acknowledgement received).
- Re-transmission if the transmission of its copy failed.

Status' of the buffer are:

- Empty: No frames waiting in the queue.
- Frames are in the queue and the load is below the threshold.
- Buffer loaded: Flow control should detect the overload status and stop sending packets to DLL.

The buffer is loaded if:

- Packet arrival rate increases.
- Frame Error Rate increases, and error control is activated.

To avoid the buffer overload status, the total frame arrival rate including the retransmission rate, should remain below the frame service rate.

The useful throughput, considering frame retransmission and block aggregation, is calculated in deliverable D5.3[18]:

4.2.5 Error control module:

This module is responsible for CRC block insertion, retransmission decisions and DLL failure report generation, considered by the networking layer.

4.2.6 Medium access and scheduler

Random access to channels is not possible for a THz wireless system using a directional antenna. Beams between two nodes should be synchronized to exchange link establishment message and data traffic. Before the transmission period starts, each node should know when it should transmit.

The medium access and scheduler modules are linked to the buffer module. For transmission, a command is sent to the MAC/scheduler module to start a network discovery process and link establishment procedure. In the next phase, communication will start considering if the buffer is not-empty.

The medium access and scheduler module is the heart of the DLL, communication efficiency relies on messages exchanged with the physical layer and with other modules in the DLL.

Based on the number of possible active links, each node will be assigned a service time duration.

4.2.7 Antenna and device management module

The antenna and device management module interfaces the DLL to the physical layer. For a directional antenna system, transmission is possible only after completing the antenna steering procedure.

We will assume that the scanning antenna range is limited to 90^0. To establish links to its neighbours, a node should contain more than one sector.

During the discovery phase, two nodes willing to establish a link with each other exchange their positions. Positions are defined using a polar system.

Exchanging reference time between nodes is still a challenge and time consuming, as nodes can discover each other with a low probability, and initial information should be well protected but also must be recognized.

4.2.8 Nodes synchronization:

It is difficult to synchronize all modules of the DLL with networking and physical layer. To synchronize between nodes, a reference time should be set and shared between nodes. The time to start any transmission phase should be agreed between all nodes in the network.

The definition of the reference time can be performed at the design phase. The time unit is nanoseconds, the duration of each phase should be also known to each node and updated if changes happen in the network.

Nodes in the network should store the information about the reference time to use it during the node discovery phase.

Based on the signal received during the discovery phase, it is possible to estimate its position by determination of the time of alignment, which is then considered as a reference time.

4.2.9 Measurements:

Collecting measurements periodically before and during a communication is required to maintain the link quality. Gathered measurements are:

- Received power
- Received SNR
- Channel state information

Measurements can be embedded in feedback messages or concatenated with ACK/NACK messages. A data compression algorithm can be implemented to reduce information overhead.

4.3 Inputs

- Collected .pcap files form MININET for 100Mbps and 10 Gbps links
- Packet length for TCP and UDP traffic
- Inter-arrival time between packets
- Source and destination address

4.4 Outputs

- Frame lengths formulation
- Buffer model
- Throughput model

4.5 Data Formats

- .mat from MATLAB tool
- txt files from ns-3
- .mat from MATLAB tool
- txt files from ns-3

5 Conclusion/Further work

Based on the results gathered for this deliverable, we can conclude that THz links can replace or augmenting wired links in a datacentre environment. The initial point-to-point nature of these links allows them to be seamlessly placed in an existing datacentre and to inter-operate with wired links in a Software-Defined Network.

From Data Link layer point of view and inspired by data collected at networking layer, it is possible to create frames using packet segmentation into blocks and framing, we proposed a model and formulate the net DLL throughput while considering error control.

Future THz links will be capable of point-to-multipoint communication. This is an important feature that will allow THz links to dynamically reconfigure their topology based on network conditions and will enable far greater network flexibility than is possible using wired optical links. An example of this can be seen in figure 5.1, where an SDN controller is deciding what is the best possible path between switch A and B. This decision is not based solely on distance or number of hops, it will require more advanced parameters such as current load on each switch, interference levels, existing links, and possibly more from both higher and lower layers in order to make an informed decision. The implementation of this feature in SDN would be very advantageous and will be the subject of future research.

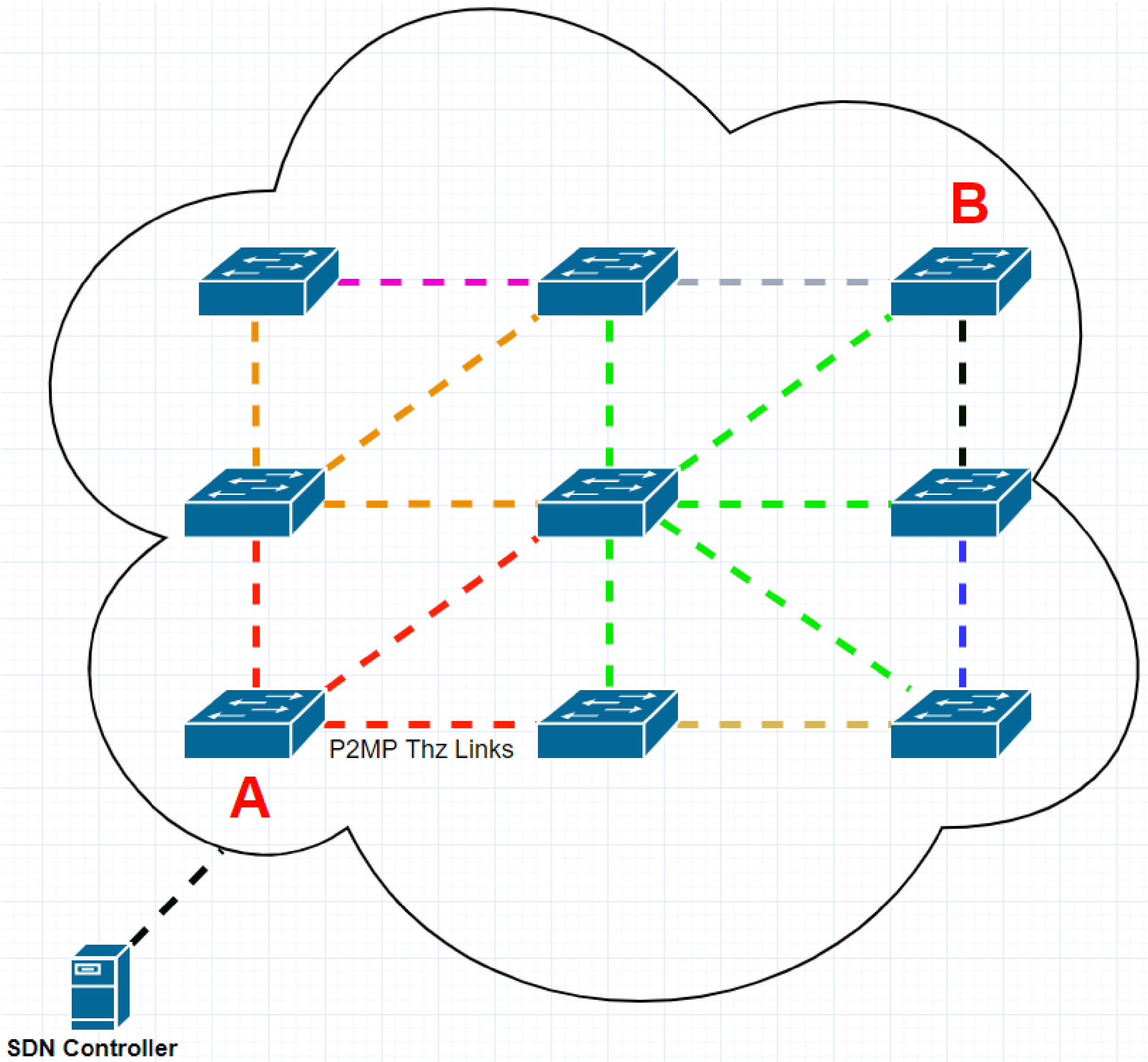


Figure 5.1: SDN Controller deciding best route from A to B using THz Links.

References

[1] McKeown, Nick. "*Software-defined networking.*" INFOCOM keynote talk 17.2: 30-32. (2009)

[2] Habib, , et al. "*Routing Techniques in Data Center Networks*". 507-532. 10.1007/978-1-4939-2092-1_16. (2015)

[3] Wang, Ting, et al. "*Rethinking the Data Center. Networking: Architecture, Network Protocols, and Resource Sharing."* IEEE Access. 2. 1481-1496. 10.1109/ACCESS.2014.2383439. (2014)

[4] Zhou, J., and H. Runjing. "*High-speed SFP+ signal integrity simulation and measurement*." Jour. of Com. and Comp 11: 371-377. (2014)

[5] Narasimha, Adithyaram, et al. "*A 40-Gb/s QSFP optoelectronic transceiver in a 0.13 µm CMOS silicon-on-insulator technology*." Optical Fiber Communication Conference. Optical Society of America, (2008)

[6] Nasu, Hideyuki, et al. "*28-Gb/s× 4-channel solderable optical transceiver module for QSFP28*." 2016 IEEE CPMT Symposium Japan (ICSJ). IEEE. (2016)

[7] Seeds, Alwyn J., et al. "*Terahertz photonics for wireless communications*." Journal of Lightwave Technology 33.3: 579-587. (2015)

[8] Shams, H., et al. "*100 Gb/s multicarrier THz wireless transmission system with high frequency stability based on a gain-switched laser comb source.*" IEEE Photonics Journal 7.3: 1-11. (2015)

[9] Levin, Dan, et al. "*Incremental SDN deployment in enterprise networks*." ACM SIGCOMM Computer Communication Review. Vol. 43. No. 4. ACM. (2013)

[10] Wang, Kun, et al. "*An SDN-based architecture for next-generation wireless networks.*" IEEE Wireless Communications 24.1: 25-31. (2017)

[11] Shu, Zhaogang, et al. *"Traffic Engineering in Software-Defined Networking: Measurement and Management."* IEEE Access. 4. 1-1. 10.1109/ACCESS.2016.2582748. (2016)

[12] Mininet, *Network emulation tool for prototyping Software Defined Networks*, URL: https://github.com/mininet/mininet

[13] Wireshark, *Tool to capture network traffic and perform traffic analytics*, URL: https://www.wireshark.org/

[14] iPerf3, *Tool to perform traffic generation and network performance analysis*, URL: https://iperf.fr/

[15] OpenDayLight, *Open-Source SDN controller*, URL: https://www.opendaylight.org/

[16] NS-3, *Discrete event network simulator*, URL: https://www.nsnam.org/

[17] S. Koenig, , et al. "*Wireless sub-THz communication system with high data rate*" Nat. Photonics, vol. 7, no. 12, pp. 977–981. (2013)

[18] N. Boujnah, et al. "*Initial Data Link Level model and simulator*" TERAPOD Deliverable D5.3 (2018)