HORIZON 2020

Deliverable ID: D5.4

Preparation date: 30.10.20

Milestone: Final proposed

Title:

Final DLL Layer Model and Simulator

Terahertz based Ultra High Bandwidth Wireless Access Networks

Editor/Lead beneficiary (name/partner):
Noureddine Boujnah / WIT

Internally reviewed by (name/partner):
Sean Ahearne/ DELL EMC
Luis Pessoa/INESC TEC

Approved by:
TBD

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

Revisions

Version	Date	Author	Organisation	Details
0.1	01.04.20	Noureddine Boujnah	TSSG	In progress
02	01.06.20	Noureddine Boujnah	TSSG	In progress
0.3	15.06.20	Noureddine Boujnah	TSSG	In progress
0.4	01.07.20	Noureddine Boujnah	TSSG	Ready for internal review
0.5		Noureddine Boujnah	TSSG	Revised based on reviewers comments
0.6	25.08.20	Saim	TSSG	Revised
0.7	30.10.20	Noureddinee	TSSG	Final version

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

Table of Contents

List of figures

Executive summary:

D5.4 is the final data link layer and simulator deliverable, it provides complementary study on network topology and access technique, network performance are evaluated theoretically, then proposed method are integrated in NS-3 simulator

1.1 Summary

WP5 is the work package of TERAPOD project related to communication protocols design and simulation, which consists of three main tasks: physical layer functionalities implementation and simulation, data link layer modelling and simulation, and finally THz wireless integration in a data center and network configuration to support high data rate THz communications within a data centre.. This document represents the final data link layer model and simulation deliverable of the project. It describes top of rack (TOR) network topologies within the data centre and DLL models, algorithms and simulations.

This document extends the work performed in deliverable D5.3 and addresses new networking challenges by proposing new networking topologies, moreover, node discovery, access techniques and synchronisation techniques are modelled for WDCN, network performances are evaluated theoretically and using NS-3.

1.2 Structure of this document

This document is organized as follow:

- Sections 1 introduces the document and set the background for work package 5.
- Section 2 gives an overview of network models and topologies.
- Section 3 describes the initial access procedures.
- Section 4 describes three medium access techniques.
- Section 5 describes the simulator architecture, implementation and performance evaluation.
- Section 6 concludes the document and proposes further work.

1.3 Relationships with other deliverables

The deliverable D5.4 is the final deliverable of work package 5 dealing with data link layer. The work presented in this document relates to the following deliverables:

- D2.2 – Final requirements and scenario specifications: requirements for point to point scenario.
- D5.3 – Initial data link layer model and simulator
- D5.6 – Final network layer model and simulator: some parts of section 2 are repeated through all deliverables of WP5 such as: initial physical layer model and simulator as well as initial networking layer model and simulator.

1.4 Contributors

The following partners have contributed to this deliverable:

- TSSG - Noureddine Boujnah
- TSSG -Saim Ghafoor

1.5 Acronyms and abbreviations

Abbreviation/Acronym	Description
ACK/NACK	Acknowledgment/non-acknowledgement
AF	Amplify and Forward
BER	Bit Error Rate

CID	Cluster ID
CRC	Cyclic Redundancy Check
CTS	Clear to send
DF	Decode and Forward
DLL	Data Link Layer
FEC	Forward Error Control
FER	Frame Error Rate
HPBW	Half-Power Beam Width
MAC	Medium Access Control
ND	Node discovery
NID	Node ID
P2P	Point to Point
PHY	Physical Layer
RTS	Ready to send
SID	Sector ID
SNR/SINR	Signal to Noise ratio/Signal to Interference plus Noise Ratio
SYN	Synchronisation
TOR	Top-of-Rack
TCP	Transport Control Protocol
UDP	User Datagram Protocol
WDCN	Wireless Data Centre Network
N_{ret}	Number of re-transmission
$\dot{\Omega}$	Number of turning rounds per second
C_n^k	Combination notation given by: $C_n^k = \frac{n!}{k!(n-k)!}$
a_0	The reference beam angular speed
τ_p	The total processing time
R	The data rate fixed to 100 Gbps
d_{max}	The maximum transmission distance fixed to 2 metres
B_{max}	The maximum frame length
L_{ack}	Acknowledgement frame length
$(x)_+$	x, if $x > 0$, and 0 if $x \leq 0$
P (x)	Probability of x
$E_{i,j}$	Set of neighbors of node (i,j)
N_{sec}	The number of sectors per node.
(p,q)	Two co-prime numbers for nodes discovery
L_{block}	Block size, a frame with length L consists of many blocks
$P_B(T_n)$	Probability that one node has data in the buffer at time T_n
N_{cycle}	The number of transmission timeslots within a cycle
M_{blocks}	The number of transmitted blocks per sector per cycle
N_{TX}	The number of cycles during transmission period
λ_p	Packet arrival rate
λ_a	Frame arrival rate to the TX buffer
T	Total transmission period
T_{stop}	Beam's stopping duration
$\Delta\phi_{3dB}$	Antenna beam width
G_{max}	Maximum antenna gain

1.6 Communication Layers functionalities, network metrics and simulation tools

Parts of the work of WP5 are described in deliverables D5.1, D5.3 and D5.5. Activities on the physical layer focuses on designing new functionalities such as modulation and coding schemes for THz channels, channel models based on results from WP4, and device characteristics implemented in simulators based on results from WP3. Activities on the data link layer focus on assessment of point to point THz links within a data centre and proposing new DLL networking models and functionalities, such as node discovery, synchronization and access and finally integration into the NS-3 simulator. The last activity of WP5 is configuration and design of networking layer algorithms to support THz communication and to interface with the DLL and physical layer.

The three layers are inter-dependent, and some functionality in each layer can be extended to other layers such as resource allocation, node discovery and synchronization. The performance of NET and DLL layer depend on device technologies and physical layer functionalities such as modulation and coding.

Table 1 depicts the main functionalities at each layer, key performance indicators and tools expected to be used by WP5 to build THz simulators.

Table 1 Summary of layer description

	Physical layer	**Data Link Layer**	**Network Layer**
Models and functionalities	• Modulation and coding schemes • Waveforms • Channel modelsRF impairments modeling	• Framing, block aggregation and buffering • Error control and re-tranmsission • Transmission modes • Medium access • Nodes discovery • Synchronisation • Relaying • Scheduling and resource allocation	• Packet format • DC Topology • DC components • Network interfaces • Routing algorithms • SDN configuration and control plane • Signalling • Logical adressing • Fault detection
Metrics	• Error probability • PHY layer delay • BER • SNR/SINR	• Throughput • Latency • Frame Error Rate • Number of re-transmissions • DLL buffer status • Channel utilization	• Packet arrival rate • NL throughput • Load /congestion • Packet loss • Delay • NL buffer status • Recovery time
Tools	• SiMoNe	• MATLAB • Ns-3 [1]	• Mininet [2] • CORE [3] • DCT2Gen[7]

2 Network models and topologies

In this section, different network models and topologies are described which are considered for the modelling and implementation. We start by describing network requirements and a summary of results from point to point (P2P) links modelling and simulation previously presented in D5.3, afterwards a layered model and network topologies are proposed for the wireless data centre network (WDCN), and finally initial and medium access procedures are theoretically investigated with the preliminary presentation of simulation results.

Nodes in data centre are arranged in a rectangular fashion, therefore each node can communicate to one or many neighbouring nodes and we propose 3 possible setting of node connectivity and evaluate each one theoretically.

2.1 Network environment and requirements

Datacentre networks consist of racks arranged in a rectangular fashion and interconnected using optical fibre and copper cables via switches. Each inter-rack link can reach 100 Gbps. The TERAPOD project aims to reduce wired connections and augment or replace them by using a THz link to achieve network requirements wirelessly using the THz band around 300 GHz described in deliverables D2.1 and D2.3. Additional factors that steer WDCN design are: data centre geometry, presence of obstacles and reflectors, and indoor environmental parameters at 300 GHz such as temperature and humidity. Measurement and characterization of channels can be found in deliverables D4.3 and D4.4.

THz nodes will be placed on top of racks (TOR nodes), and equipped with THz devices proposed by WP3 and demonstrated in WP6, THz transceivers will be interconnected to the existing optical interface to carry DC traffic. The inter-rack distance varies between 1.4m and 2 m, and link throughput can reach 10 Gbps. We assume in this study that traffic is identically distributed between all nodes of the network.

Through this deliverable, we will propose a couple of technical solutions and demonstrate that it is possible to meet the following requirements:

- Achieving a wireless throughput per node between 10 Gbps and 100 Gbps: reaching high data throughput is enabled by: short transmission time and time-based scheduling, fast network discovery and synchronization, error control via block aggregation and retransmission, and centralized time based fast beam switching medium access.
- Achieving low latency by using fast beam switching and steering, time slot allocation techniques and fast scheduling, short transmission time and node relaying.
- Data link layer should include functionalities such as node discovery, synchronization, error control, medium access techniques and scheduling.
- Topology adaptability and autonomous network: topology can affect network throughput and latency, the network should be able to change seamlessly its topology to improve quality of service, and changing network parameters can be monitored by SDN. For instance, increasing the number of node’s neighbors based on data load on nodes to increase throughput and reduce latency.

WDCN is constrained by the actual geometry of data centre, the proposed models should consider arrangement and distance between racks for nodes placement, however it is still possible to come up with different network layouts and propose improved solutions for a WDCN. Theoretical modeling and simulation using both MATLAB and NS-3 will be considered.

2.2 Point-to-point WDCN

In the previous deliverable D5.3, a point-to-point wireless link was modelled and described where different functionalities can be integrated in WDCN such as framing and frame splitting, error control and re-transmission. A P2P wireless link can be fully integrated to augment an optical fibre link to boost data rate or to replace an existing wired link [4]. Through improvements in the physical layer functionalities, communications range can be extended and higher throughput can be reached using higher modulation schemes, while BER can be further enhanced using new channel coding schemes (physical layer link level simulations are described in the deliverable D5.2). Basic functionalities and results for point-to-point physical layer can be found in [5].

A good understanding of the THz channel within the data centre environment will push the final design of the WDCN and improve the global WDCN key performance indicators.

The WDCN link quality depends on device technology proposed in the TERAPOD project including emitters, receivers, mixers and antennas. New THz devices can reach an output power of 1 mW, multiple THz channels can be used and noise can be considerably reduced, antenna gain can reach 27 dBi at the transmitting and receiving side and new antenna capabilities can be integrated such as beam sweeping techniques.

We will present a summary of the previous results for P2P links, in terms of P2P models and simulations which will be extended for the WDCN design to provide new improvements.

- Frame error rate (FER): due to channel impairments, frames can be corrupted and lost; the frame error rate is an indicator of frame quality that depends on two parameters, the bit error rate (BER) and the frame length, according to:

$$P_{\mathrm{f}} = 1 - (1-p)^B \quad \text{(Eq. 1)}$$

where B is the frame length and p is the bit error rate (BER). The IEEE 802.15.3d standard [5] shows that a frame length exceeding 15000 bytes can be used for P2P THz links, moreover lower values of BER can be achieved($<10^{-6}$). Since FER depends on BER and frame length, techniques such as error control and re-transmission can be integrated in DLL layer to reduce frame loss and improve useful throughput. FER can also be improved by using sophisticated channel coding techniques.

- Packet to frame conversion equation: the frame arrival rate to the DLL buffer is linked to the packed arrival rate and average packet length by the following equation:

$$\lambda_{\mathrm{a}} = \frac{\bar{L}_p}{B - B_{\mathrm{overhead}}} \lambda_{\mathrm{p}} \quad \text{(Eq. 2)}$$

  where, λ_{p} is the packet arrival rate from the network layer, $\bar{L}_{\mathrm{p}}$ is the average packet length and $B_{overhead}$ is the length of frame overhead data which contains cyclic redundancy check (CRC) data and frame preamble.

- Results on mean number of re-transmissions and useful throughput for P2P link.

  The instantaneous useful throughput is given by:

$$R(l) = \frac{B}{lT} \quad \text{(Eq. 3)}$$

  where l is the number of frame re-transmissions. The mean useful throughput is given by:

$$\bar{R} = \sum_{l=1}^{+\infty} R(l) P_{\mathrm{r}}(l) = \frac{-B^2(1-p)^B \log(1-p)}{T(1-(1-p)^B)} \quad \text{(Eq. 4)}$$

  And $\bar{R}$ can be expressed also as a function of the mean number of re-transmissions $\bar{l}$:

$$\bar{R} = \frac{B}{T} \frac{\log(\bar{l})}{\bar{l}-1} \quad \text{(Eq. 5)}$$

  The mean number of re-transmissions can be expressed as:

$$\begin{cases} \bar{l} = \left(\frac{1}{1-p}\right)^B & N_{\mathrm{ret}} = \infty \\ \bar{l} = \frac{(N_{ret}-1)P_e(p,B)^{N_{\mathrm{ret}}} - N_{ret}P_e(p,B)^{N_{\mathrm{ret}}-1} + 1}{1-P_e(p,B)} & N_{\mathrm{ret}} < \infty \end{cases} \quad \text{(Eq. 6)}$$

  When FEC is considered, the frame error rate can be reduced, depending on the channel coding used at the physical layer. If η is the maximum number of corrected bits for a frame with length L, the frame error rate becomes:

$$P_f = 1 - \sum_{k=B-\eta}^{B} C_B^k \, p^{B-k}(1-p)^k \quad \text{(Eq. 7)}$$

and the mean number of retransmission becomes:

$$\bar{l}(p,B) = \frac{1}{\sum_{k=M}^{L} C_L^k p^{B-k}(1-p)^k} \quad \text{(Eq. 8)}$$

The useful throughput is then expressed as:

$$\bar{R} = \frac{-B\left(\sum_{k=B-\eta}^{B} C_B^k p^{B-k}(1-p)^k\right) \log\left(\sum_{k=B-\eta}^{B} C_B^k p^{B-k}(1-p)^k\right)}{T\left(1-\sum_{k=B-\eta}^{B} C_B^k p^{B-k}(1-p)^k\right)} \quad \text{(Eq. 9)}$$

- Frame arrival rate with block aggregation and re-transmission:

to optimize link performances, frames are split into blocks with size L_{block}, and only corrupted blocks will be re-transmitted. The frame arrival rate can then be calculated as:

$$\lambda_{\text{a}} = P_{\text{f}}\mu + \frac{\bar{L}_{\text{p}}}{K L_{\text{block}}}\lambda_{\text{p}} \quad \text{(Eq. 10)}$$

This result will be proved mathematically for the general use case in section 3. Figure 1 shows the evolution of the total frame arrival rate based on DC traffic, BER and frame length. To avoid buffer load and high frame latency, the frame arrival rate λ_a should be less than the service rate μ. We show that it is possible to optimize the frame length to reduce the impact of re-transmissions on buffer load and delay.

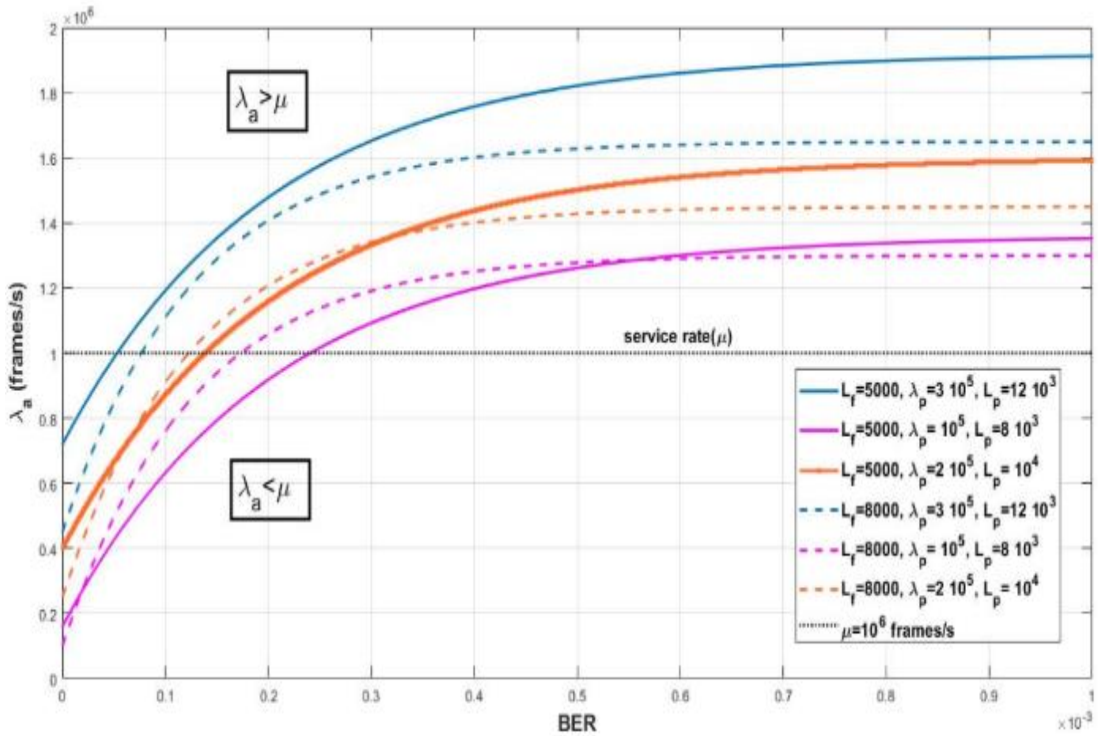

Figure 1 Evolution of frame arrival rate as a function of packet arrival rate, average packet length, frame length and BER.

Additional simulation results from NS-3 for P2P WDCN will be provided in section 5.

2.3 Layered model

Figure 2 shows an architecture overview of the different data link layer functionalities and interfaces with network and physical layers.

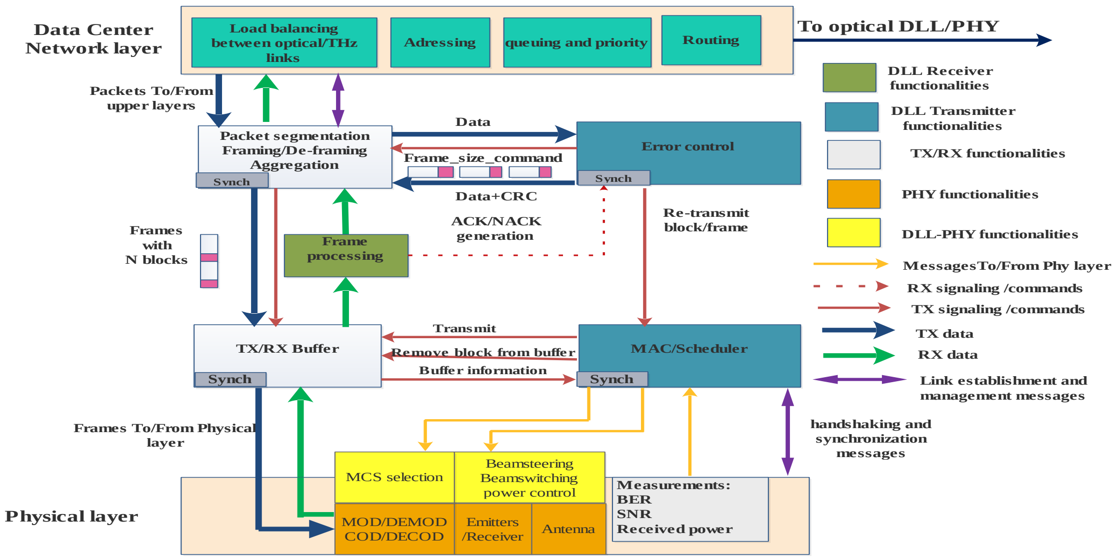

Figure 6 DLL protocol model and interfaces

The diagram in Figure 2 represents the data link layer for WDCN. The transport control protocol (TCP) and user datagram protocol (UDP) packets with fixed and variable length arrive to the link layer from the network layer, and are encapsulated into frames and blocks with regular shape, while CRC bits are added to each chunk of data (block/frame) to protect data. The error protection module is responsible for CRC computation and decision if a frame or block must be re-transmitted based on received acknowledgment. Frames include all necessary data for transmission: destination id, source id, CRC field, and other required information. The buffer should be designed for THz WDCN communications according to the expected high frame arrival rate and short service time, while frame and service rates are selected to meet throughput requirements and also to avoid buffer over-load and delay increase.

Packets are received from networking layer and converted into frames,

frame consists of blocks, with error control fields, which should wait for transmission in the transceiver's memory until a transmission order is received during a scheduled timeslot. The medium access control (MAC) scheduler module checks if the buffer contains frames or is empty, then decides about if transmission is taken based on information received during synchronization period, additional configurations are sent to the physical layer before the transmission starts. The DLL is linked to the physical layer, the physical layer reports on channel information, link quality, bit error rate, and received power, for instance. The information received from the physical layer helps with selecting the number of blocks per frame, and re-transmission of blocks/frames. The data link layer sends commands for beam steering, and senses periodically the wireless link to decide about frame length, re-transmission and resource allocation. The model proposed in figure 1 along with networking mechanisms at DLL can be used for future hardware implementation of THz WDCN.

2.4 Nodes arrangement within WDCN

Supposing nodes are placed on top of data centre racks, we propose two types of arrangement; in the first one, the node is placed on a central position of TOR, and on the second one, the node's positions form an horizontal hexagonal grid. Placing the nodes at the TOR centre can increase the blockage problem. Communication between distant nodes is not possible, and therefore the number of neigh-

bours will be reduced. Establishing more connectivity can improve link management and reduce networking operations such as nodes synchronisation and relaying, however it can affect link throughput, moreover, synchronisation and relaying may increase network latency.
In Figure 3 and Figure 4 we present two node placement techniques: the horizontal rectangular network grid and the horizontal hexagonal network grid, respectively. The selection of an arrangement technique of nodes will impact the network design cost, networking operations, and quality of service and it will be based on network performance metrics such as throughput and delay.

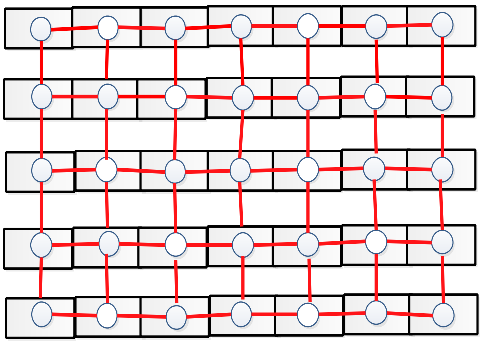

Figure 3 Rectangular grid: nodes placed on central position of top of racks

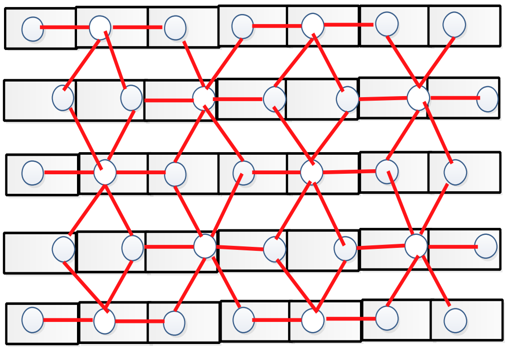

Figure 4 Hexagonal grid arrangement of nodes

Due to the rack's geometrical arrangement and physical properties of THz waves, a direct line of sight between two nodes is not possible for distant nodes. To enable communications between all nodes in the network, we need to select a network topology that organizes and monitors wireless communications at the data link layer. Using an autonomous ad-hoc network requires that all nodes possess the same functionalities and advanced coordination procedures should be implemented in each node to decide about the communications paths and the number of transmission timeslots required for each node. The sectorized-clustered-centralized network will be more suitable for such a network, where nodes are split into sectors, and each group of sectors from one cluster monitored by one central node with additional functionalities to perform switching and synchronisation.

2.5 Cluster-based central network

One of the most advantages of using wireless communication is to increase node connectivity and enhance network flexibility. Nodes on top-of-rack can exchange networking information and establish a reliable link with their neighbours. Two possible approaches can be deployed for THz WDCN, the first one uses cluster-based central network and the second on based on ad-hoc network, in the latest technique, nodes are autonomous however, this approach includes complicated networking operations and requires frequent synchronization between nodes and sectors leading to low data throughput and more interference, moreover, a technical solution based on placing reflectors to assist link establishment can be used, however is not considered in TERAPOD. A study on network performance for ad-hoc WDCN can be found in [15]. In a cluster-based central network, communication is assisted by a central node for each cluster, the cluster can be synchronized during the Synchronise (SYN) period. The central node will be also responsible for forwarding frames if the destination is outside the cluster. Due to the data centre geometry, racks are arranged in a rectangular fashion and communication between two TOR nodes can be blocked. In our model, we assume each TOR node can establish a limited number of possible connections with its neighbours, the number of possible connections depends on the number of sectors per node and antenna model. Each cluster is identified by its CID(cluster Identifier), and each frame preamble contains the source CID and destination CID to be used for switching.

2.6 Node connectivity

We propose 3 possible connectivity models associated to the proposed TOR topologies: 4-nodes, 6-nodes and 8-nodes connectivity. The neighbouring list of a central node (i,j) is designated by $E_{i,j}$, it includes all nodes with sectors with a direct link with the central node (i,j), a node can belong to more than one cluster. A nodes connectivity is constrained by DC parameters such as inter-rack distance, and presence of reflectors and blockers. Moreover, transceiver capabilities and THz channel can affect the network quality, we will study the network performance of the proposed 3 node connectivity models in the theoretical part, and we will implement only one topology :

- 4- Neighbors connection: Each node has 4 possible connections, a cluster consists of 5 nodes with a central node with additional capabilities to organize transmission and reception, the central node is also aware of the routing path to decide about the next hop within the cluster. Nodes are distributed with equal distance with the central node. This topology is implemented and simulated in ns-3 (5 nodes).
- 8-neighbors connection: Each cluster consists of 9 nodes, a central node can establish 8 connections, this connection scheme suits well for the rectangular scenario, the transmission time is equally distributed between sectors belonging to the cluster.
- 6-neighbors connection: each node can manage 6 possible connections, it can be considered in the hexagonal arrangement of TOR nodes.

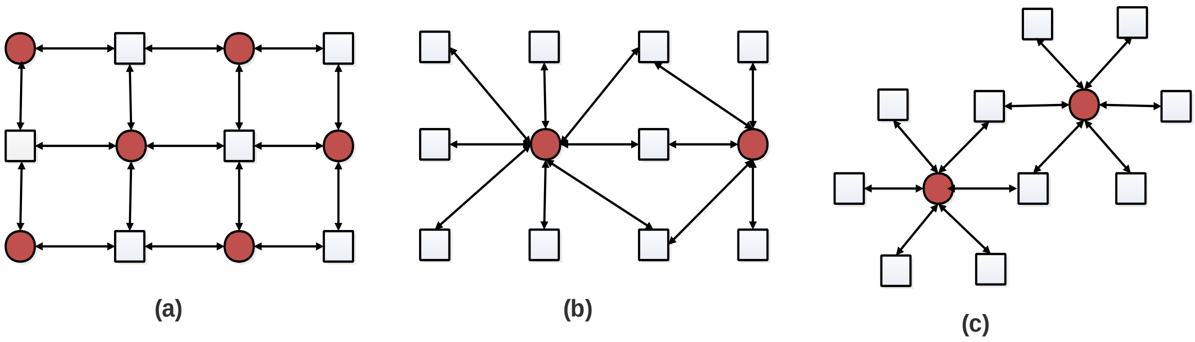

Figure 7 (a) 4-nodes connectivity, (b) 8-nodes connectivity, (c) 6-nodes connectivity (hexagonal cluster)

The selection of a node's connectivity affects the network quality of service in ways such as data throughput, delay and frame error rate. For example, the cycle period increases if node connectivity increases, therefore the mean delay increases.
To reduce network operations for a cluster based central network, nodes are arranged in small clusters, each cluster will be monitored by one central node with additional functionalities. A central node geometrically placed at the cluster centre plays the role of node coordinator and it is linked directly to all other nodes.
Neighbours of the central node can belong to other clusters; therefore, each node can transmit to more than one central node (coordinator), each node consists of independent sectors and each sector belongs to a cluster. the traffic generated by one node is split evenly between its sectors.
Moreover, we assume a dependency between sectors in the same node, where only one sector can transmit or receive during a transmission period to avoid inter-sector interference. To activate all sectors simultaneously a deep investigation on inter-sector interference will be carried out in the future.
Within a cluster, each sector is assigned a constant timeslot. All sectors in the network are synchronized during the synchronisation phase.

2.7 Node and antenna model

Antenna technology developed by TERAPOD and previous research works at the THz band will impact the topology design and communication protocols, antenna arrays can improve antenna gain and coverage by focusing the total energy toward one steerable direction. For WDCN, nodes should be split into sectors, each sector will be equipped with a transmitter, receiver and antenna.

- Antenna model: to optimize the link budget, directional antennas are required at the transmission and receiver side, beams for THz communication are tight and steerable to serve many nodes. In this deliverable, we consider a phased array with miniaturized antenna elements spaced with half-wavelength at the operating frequency of 300 GHz, the number of antenna elements can be 4, 8 or 16, antenna parameters can be further optimized to reduce side lobe magnitude and improve the total SINR of the network. Previous works consider a 1-by-4 phased array on chip-antenna, at 300 GHz and using 20 GHz of bandwidth, showing the capability of beam forming and beam steering, the proposed architecture can reach a gain of 15.8 dBi and angular gain of 24^0 [8], a phased array of 4 elements achieving a gain of 20.7 dBi, a HPBW equal to 10.3^0 and beam steering range of 40^0 is proposed in [9]. Antenna design and optimization is considered in WP3.

- Node Sectorization: the idea of node sectorization is to split the node into neighbouring sectors with or without overlapping regions, each sector is equipped with a one-directional antenna, the angular sweeping range, proposed by TERAPOD project, can reach 90^o. Therefore, a 4-nodes connectivity network can use four directional antennas per node, for 8-nodes and 6-nodes connectivity networks, four steerable antennas per node can be used. Each node to communicate to a neighboring node needs to send its data through one sector after beam alignment, both beam switching and steering can be used to transmit. Sectors are physically independent, however all sectors within a node are synchronized at DLL level.

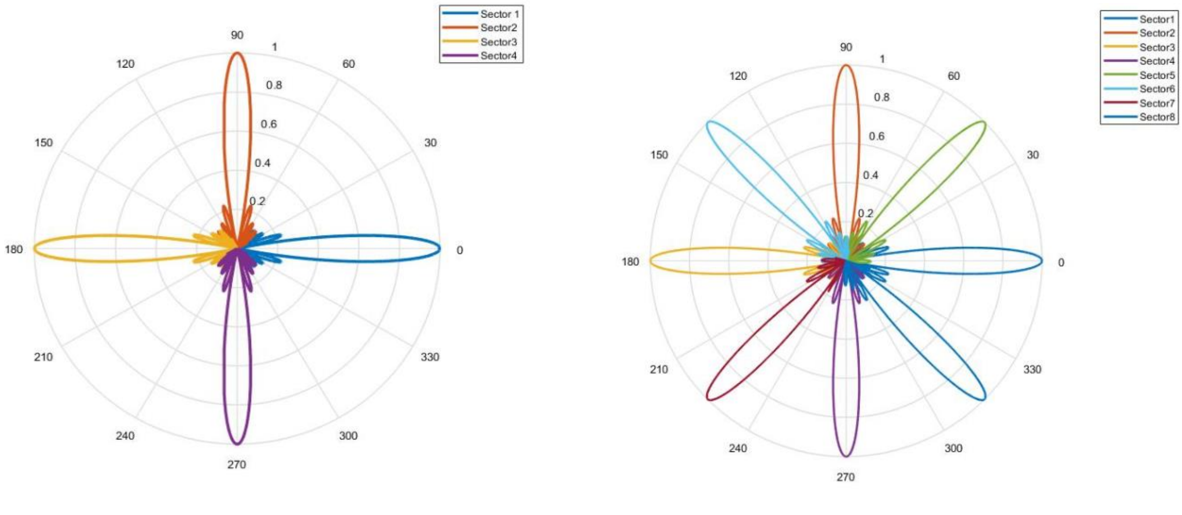

Figure 8 (left): Radiation patterns of 4-sectors(Central node) , (right): Radiation patterns of 8-sectors node(Central node).

The antenna range proposed by TERAPOD lies between -45^o and 45^o, then the node needs to be sectorized to cover a total angular space of 360^0. The number of sectors for each node depends on the number of its neighbours and it will impact on network performances, for the rectangular model it can be 4 or more sectors, for the hexagonal model, it can be 3 or 6 sectors, the number of sectors is limited also by the communication range. We propose two sectorization techniques:

- Overlapping sectors: overlapping between sectors happens when the angular separation between neighbouring sectors is low, it leads to mutual interference between sectors when are simultaneously activated (point to multipoint connection), overlapping between sectors depends on antenna pattern, interference can be triggered by antenna side lobes. To reduce inter-sector interferences, we need to tune antenna parameters or to use different sub-band.
- Non-overlapping sectors: when the inter-sector angle within a node is high, then sectors can be activated simultaneously (point to multipoint scenario), when using phased array, increasing the number of elements per antenna can reduce side lobe effects on neighbouring sectors. This model is considered in the proposed simulator using cosine pattern.

Figure 6 shows 2 use cases of node sectorization which can be considered in a data centre. Beam steering and beam switching as modelled in the previous paragraph show each node consists of a number of sectors considering the arrangement of its neighbouring nodes. A sector is equipped with one antenna with or without beam steering capabilities. In section 4, we will use an existing antenna model with radiation pattern similar to phased array with HPBW equal to 30^0, implementation of phased array in the simulator will be considered in future work.

2.8 Nodes relaying

Two possible relaying techniques can be proposed if a direct link is not possible to establish between transmitter and receiver, the first one is amplify and forward, the second is decode and forward. Each algorithm has an impact on the target QoS

- Amplify and forward technique (AF): if the destination node does not belong to the neighbours of the transmitter, the signal received by the relay node is amplified and forwarded to the next relay until it reaches the final destination node; the signal will be received at the final destination with high SNR.
- Decode and forward technique (DF): each received frame is decoded at a neighbour node, if there are no errors detected, it will be forwarded to the next node with a new header, and erroneous frames are retransmitted. This technique requires high transceiver performance.

Relaying is required when a signal cannot reach its destination from source due to blockage or distance effects, the relay improves signal quality by boosting its amplitude(AF) or regenerating the signal using the decode and forward technique(DF). Using a centralized clustered network, each forwarded packet contains the destination node, the node controller forwards it to the next node based on preliminary static table, this operation can also be performed at higher layer, by the SDN controller. During the network operation controllers are aware about traffic load at each node; the relaying path should contain nodes with low load to avoid delays. Relaying introduces additional delays to the WDCN, paths with high number of relays will increase the network latency, deploying an 8-connectivity network can help reduce the number of required relays in the network compared to a 4-connectivity network, however data throughput will decrease.The distance between the source node and the distance depends on network topology.
To facilitate frame switching two additional fields should be added to identify the source and destination, which are SID(Sector Identifier) and NID(Node identifier), relaying is performed if destination NID is different from source NID. In our proposed topology, one source SID can be associated with only one destination SID for the 4-nodes and 6-nodes neighbors. And one sector can be associated with one or two sectors for 8-nodes connectivity.

2.9 Beam turning procedure

Beam turning or sweeping is required to change beam orientation and sends/receives signals to/from different directions, using both node sectorization and directional steerable antennas it is possible to seamlessly perform sweeping with different angular speeds. The angular speed can be adaptively selected for each access phase, starting from node discovery to network synchronisation and finally to the transmission phase. For the discovery period, node parameters are agnostic to the new node entering the network for the first time, beam turning speed should be selected to achieve fast discovery, during node discovery, node position and timing information will be exchanged, during the synchronization period, nodes exchange their transmission information and parameters to other nodes in the network. During the transmission period, beam turning speed will be reduced as the data to transmit increases and transmission requires a stable link.

- Clockwise equal beam turning speed: if two nodes are using the same turning speed, then the phase difference should be equal to π to guarantee the beam alignment.
- Different beam turning speeds: to increase the meeting probability between two nodes, different speeds should be used, let $\varphi_i(t)$ and $\varphi_j(t)$ equal the beam angles associated to nodes i and j respectively, let t_n and t_{n+1} equal two consecutive beam meeting times, such the following equations hold:
-

$$\begin{cases} \varphi_i(t_{n+1}) - \varphi_i(t_n) = 2k\pi, & k \in \mathbb{Z} \\ \varphi_j(t_{n+1}) - \varphi_j(t_n) = 2l\pi, & l \in \mathbb{Z} \end{cases} \quad \text{(Eq. 11)}$$

And for $i \neq j$,

$$\varphi_j(t_n) = \varphi_i(t_n) + (2m+1)\pi,\ m \in \mathbb{Z} \quad \text{(Eq.12)}$$

Assuming that beam angle is a linear function of time, then Eq. 11 and Eq.12 lead to the following equation:

$$\frac{a_i}{a_j} \in \mathbb{Q} \quad \text{(Eq. 13)}$$

Where a_i and a_j are the beam angular speeds for node *i* and node *j*, respectively. We can prove the following:

- $a_i = p_i a_0$ and $a_j = p_j a_0$, where p_i and p_j two co-prime numbers and a_0 is a reference speed.
- The alignment probability is equal to $\frac{1}{|p_i - p_j|}$
- When beams are turning in opposite directions, the alignment probability becomes $\frac{1}{|p_i + p_j|}$
- The number of beam stops is equal to $|p_i - p_j|$ if beams are turning in the same direction and equal to $|p_i + p_j|$ if turning is performed in opposite directions.

3 Network initial access

The initial access consists of two phases: node discovery and network synchronization, the node discovery procedure (ND) happens rarely for WDCN as node positions are well known, ND aims at discovering the network by a new node and discovering of new nodes by the network, the second networking procedure is node synchronization during which information related to transmission, timing and available resources are exchanged between nodes.
Medium access is in the phase for data transmission and reception after network resource allocation and link preparation. In this deliverable we will select time-based beam allocation techniques for medium access. We start by network modeling of the three networking phases, then results will be presented for throughput and network delay.

3.1 Nodes discovery procedure

The node discovery phase occurs rarely for a data centre scenario, as nodes are fixed and information on node's positions can be cached, each node uses beam turning to send and receive ND frames, an ND frame includes node ID, position and timing information. A new technique based on fine node discovery and prime numbers is proposed to increase ND efficiency. Information of the discovered node will be shared by all nodes in the network, then the discovered node shares the information of its network. The following procedures are used during ND:

- Handshaking modes: handshaking is used to establish a reliable link between two nodes, it can be used in ND procedure to exchange transmission parameters
- Beam turning and stopping time: in order to maintain link stability, a stopping time during node transmission is required, increasing the beam stopping times to increase link reliability, however, the delay can increase.
- The number of stops during one round (one cycle): in order to increase the beam meeting probability, the number of beams stops should be increased

The stopping time and the received power are a function of distance between transmitter and receiver, as distance is unknown during the discovery phase, device's capabilities such as antenna gain, and output power are important factors to discover new nodes along with steering parameters.

During the network access phase, nodes steer their beams with a constant speed $\dot{\Omega}$, a stopping time is required to send and receive discovery frames and acknowledgements, it includes propagation, transmission and processing time. The sector stopping time should be designed based on maximal distance and maximal frame length according to the inequality:

$$T_{stop} > 2\frac{d_{\max}}{c} + \frac{L_{\mathrm{ND}}+L_{\mathrm{ack}}}{R} + \tau_p \qquad \text{(Eq.14)}$$

And under the following constraint:

$$N_{\mathrm{stop}} T_{\mathrm{stop}} = \frac{1}{\dot{\Omega}} \qquad \text{(Eq.15)}$$

N_{stop} is the number of stopping times during one turning round at a specific angular direction, it can be different from number of neighbors discussed previously, L_{ND} is the frame size used for node discovery, increasing N_{stop} improves discovery accuracy. The transmission and reception of acknowledgement occurs during T_{stop}. During T_{stop} information related to node parameters, timing and node capabilities are exchanged.

The condition on angular speed and the number of beam stops is given by:

$$\frac{1}{N_{stop}\dot{\Omega}} > 2\frac{d_{\max}}{c} + \frac{L_{\mathrm{ND}}+L_{\mathrm{ack}}}{R} + \tau_p \qquad \text{(Eq.16)}$$

Increasing N_{stop} leads to reducing angular beam speed.
Meeting time scheduling: as beams are tight, a scheduling time should be set between neighbors TOR. Based on results in the previous paragraph, the stopping time and number of stops are given by the following equations:

$$\begin{cases} N_{\mathrm{stop}} = |p_i - p_j| \quad \textit{turning in same direction} \\ N_{\mathrm{stop}} = |p_i + p_j| \quad \textit{turining in opposite direction} \\ T_{\mathrm{stop}} = \frac{2\pi}{N_{\mathrm{stop}} a_0} \end{cases} \qquad \text{(Eq.17)}$$

The reference angular speed should verify:

$$a_0 < \frac{2\pi}{N_{stop}\left(2\frac{d_{max}}{c}+\frac{L_{ND}+L_{ack}}{R}+\tau_p\right)} \qquad \text{(Eq. 18)}$$

Theoretically, nodes within a range of d_{max} can be discovered in one round, using a lower reference speed will increase the discovery probability as the distance between new nodes and a node in the network is unknown, however, low angular speed will increase the discovery period.
Examples:

- $p_i = 5, p_j = 17$ are two co-prime numbers, when beams are turning in the same direction, N_{stop}=12
- $p_i = 5, p_j = 7$ are two co-prime numbers, when beams are turning in opposite direction, N_{stop}=12

3.2 Network synchronisation

Synchronization between nodes is required to exchange information related to node transmission time and buffer status of nodes before a transmission period starts. Between two consecutive synchronization times, nodes with data in the buffer can transmit and receive, nodes without data in the buffer during the last synchronization period can transmit in the next transmission period.
Synchronization procedures can be simulated for different topologies.
During SYN period nodes exchange frames with size L_{SYN} which corresponds to a synchronization time T_{SYN} to exchange data in both ways between two neighbour nodes, information is carried from node to node until all nodes in the network are synchronized. Any node in the network receiving new information, forwards it to its next neighbour one time.

- Synchronization for rectangular arrangement: first the horizontal synchronization of nodes is performed then then vertical synchronization to exchange networking information between rows
- Cluster based synchronization: the procedure starts with intra cluster synchronization, then inter cluster synchronization

Using a simulation of a 5X7 TOR network, it is possible to optimize the synchronization period. In figure 7 we show the variation of the number of synchronized nodes as function of time, we prove that the first procedure can achieve the lowest synchronization period T_{opt}=(5+7-1) T_{SYN}=11T_{SYN}.

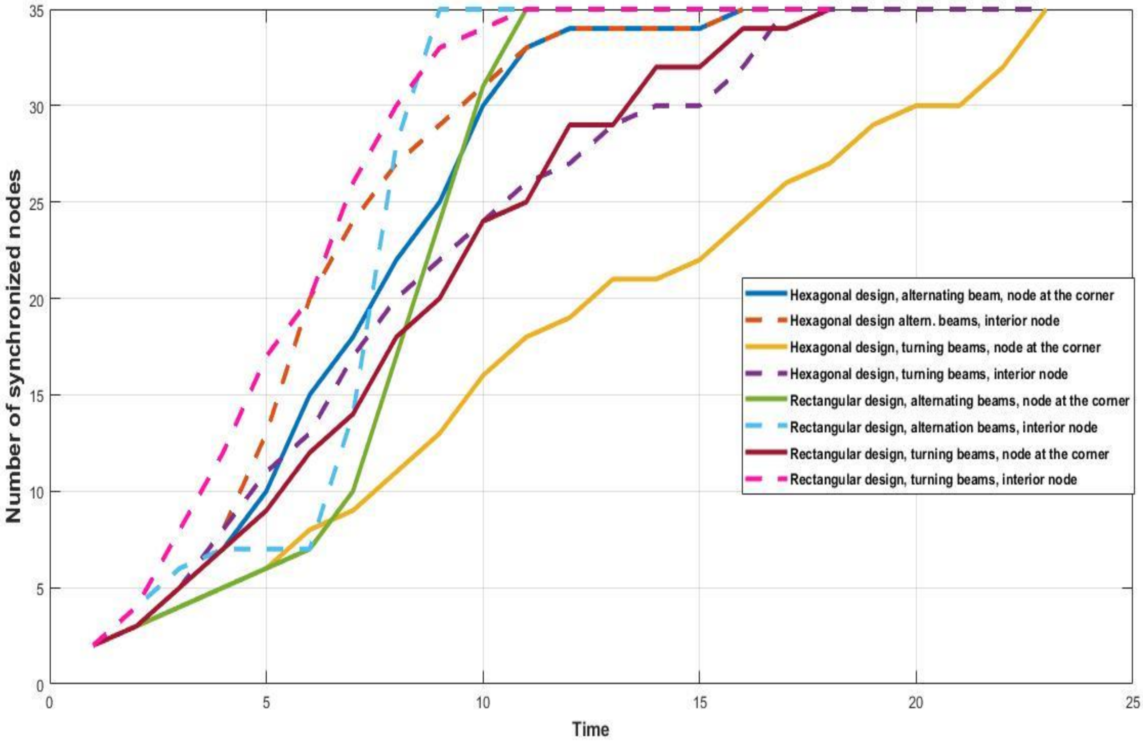

Figure 9 Synchronisation period evaluation for different topologies

4 Medium access

Previous protocols were proposed for THz medium access using multiple band handshaking for beam alignment, for instance TAB-MAC protocol [6], for a data centre use case where nodes are fixed, then it is possible use a standalone technique to reduce the time needed for the initial access phase.
The cluster consists of one central node with beam switching capabilities and sectors with fixed orientation toward the central node. During the synchronization procedure all nodes are informed about their time for transmission and reception. The medium access can be performed in three ways:

- Random access during the TX/RX time associated to each sector
- Using Request/Clear to send (RTS/CTS) procedure to avoid collision between TX and RX, RTS and CTS can be initiated by central node or the edge node.
- Fixed allocation of transmission and reception time for each sector

Two techniques are implemented in the simulator, the last one is studied theoretically, and it will be integrated in the future in the simulator.
Only the cluster central node is performing beam switching and steering, other nodes are synchronized to the central node, each node can transmit to many central nodes, each node sector can be involved only in transmission and reception in one cluster. Only one sector per node can be activated simultaneously.

4.1 Time-based beam switching with random access

In this scheme, the transmitter node directly sends data to the receiver node. The receiver, upon the reception of the data frame sends an ACK message back to the transmitter, which confirms the reception. Figure 8 shows the MAC scheme flow for random data transmission where collisions can occur.

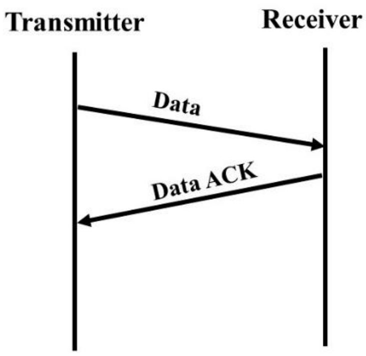

Figure 10: MAC scheme without RTS CTS for data transmission.

For the WDCN, time multiplexing coupled with a beam switching technique is used**.** During one data transmission period, each node transmits its data via each one of the sectors. Each sector transmits one frame with size B bits during one fixed time duration T if it has data to transmit (non-empty buffer), where the transmitted frame contains M_{blocks} blocks. The central node switches its beam to align with all nodes in the cluster.
A collision can happen if the two nodes (the cluster central node and the edge node) have data in their buffer to transmit, such that the collision probability depends on frame arrival rate to each sector and the time between two consecutive transmission periods. A transmission cycle contains N_{cycle} transmit periods for point-to-point transmission (timeslots), and N_{cycle} transmissions can be multiplexed during a period of T_{cycle}, the total transmission period consists of M cycles,. Both beam steering and beam switching based medium access methods can be applied.

Let P_B represent the probability that one sector has data to transmit after the transmission period T, assuming that initially there is no data in the buffer sector, then this initial probability, at time $T_0 = T_{\mathrm{cycle}}$, is given by:

$$P_B(T_0) = 2e^{-\lambda T}(1 - e^{-\lambda T}) \quad \text{(Eq.19)}$$

Where T_{cycle} is the transmission period during one round and λ is the frame arrival rate to the buffer (of each sector). Then the initial throughput per cluster during a cycle period is:

$$R_s(T_0) = \frac{B}{T_{\mathrm{cycle}}} \sum_{k=0}^{N_{\mathrm{cycle}}} kC_{N_{\mathrm{cycle}}}^{k} {P_B}^{k} (1 - P_B)^{N_{\mathrm{cycle}}-k} = 2e^{-\lambda T_{\mathrm{cycle}}}(1 - e^{-\lambda T_{\mathrm{cycle}}}) \frac{N_{\mathrm{cycle}}B}{T_{\mathrm{cycle}}} \quad \text{(Eq.20)}$$

We assume that in a data centre scenario, all nodes have always data to transmit, therefore the arrival rate is very high. In the beginning of each transmission period, we suppose the number of blocks in the buffer is $N_b(T_n)$, and the amount of transmitted data blocks at $T_{n+1} = T_n + T_{\mathrm{cycle}}$ is equal to $(N_b(T_n) - M_{\mathrm{blocks}})_+$, and the number of frames at time T_{n+1} evolves as:

$$N_b(T_{n+1}) = m(T_{\mathrm{cycle}}) + (N_b(T_n) - M_{\mathrm{blocks}})_+ \quad \text{(Eq.21)}$$

Where $m(T)$ is the number of frame arrivals during T, which is assumed to be a Poisson process with parameter T and λ and mean λT, therefore the probability of an empty buffer at T_{n+1} is:

$$P_B(N_b(T_{n+1}) = 0) = P(N_b(T_n) \leq M_{\mathrm{blocks}})e^{-\lambda T_{\mathrm{cycle}}} \quad \text{(Eq22)}$$

Eq. 22 can be used when M_{cycle} blocks are multiplexed in one timeslot period.
From Eq. 21 we can calculate the probability of non-empty buffer $P(N_b(T_{n+1}) = k, k \neq 0)$ as:

$$P(N_b(T_{n+1}) = k) = P(m(T) + (N_b(T_n) - M_{\mathrm{blocks}})_+ = k) \quad \text{(Eq.23)}$$

Hence,

$$P(N_b(T_{n+1}) = k, k \neq 0) = \sum_{j=0}^{k-1} P(m = j)P(N_b(T_{n-1}) = M_{\mathrm{blocks}} + k - j) + P(m = k) \sum_{l=0}^{M_{\mathrm{cycle}}} P(N_b(T_{n-1}) = l) \quad \text{(Eq.24)}$$

And,

$$P(N_b(T_{n+1}) = 0) = P(m = 0) \sum_{l=0}^{M_{\mathrm{blocks}}} P(N_b(T_{n-1}) = l) \quad \text{(Eq.25)}$$

Then, $P(N_b(T_{n+1}) \leq M_{\mathrm{blocks}})$ can be iteratively formulated as:

$$P(N_b(T_{n+1}) \leq M_{\mathrm{blocks}}) = P(m \leq M_{\mathrm{blocks}}) P(N_b(T_n) \leq M_{\mathrm{blocks}}) + \sum_{k=1}^{M_{\mathrm{blocks}}} \sum_{j=0}^{k-1} P(m = j)P(N_b(T_n) = M_{\mathrm{blocks}} + k - j) \quad \text{(Eq.26)}$$

Eq. 25 can be simplified as:

$$P(N_b(T_{n+1}) \leq M_{\mathrm{blocks}}) = P(m \leq M_{\mathrm{blocks}}) P(N_b(T_n) \leq M_{\mathrm{blocks}}) + \sum_{j=1}^{M_{\mathrm{cycle}}} P(N_b(T_n) = M_{\mathrm{blocks}} + j) \sum_{l=0}^{M_{\mathrm{cycle}}-j} P(m = l) \quad \text{(Eq.27)}$$

Assuming all nodes in the cluster have the same frame arrival rate equal to $\lambda_{\mathrm{node}} = N_{\mathrm{sector}}\lambda_{\mathrm{sector}}$, and the service rate is $\frac{1}{T_{\mathrm{cycle}}}$, the service rate decreases if the number of per node connections increases. For example, for a 4-nodes connectivity case the frame arrival rate per node is: $\lambda_{\mathrm{node}} = 4\lambda$.
The probability that M_{blocks} blocks are transmitted, during T_{cycle}, without collision is given by:

$$P(\text{TX without collision}) = 2P(N_b(T_{n+1}) = 0)(1 - P(N_b(T_{n+1}) = 0)) \quad \text{(Eq.28)}$$

Using equation Eq. 27, the instantaneous throughput per cluster is then given by:

$$R_s(T_{n+1}) = 2P(N_b(T_n) \leq M_{\text{blocks}})e^{-\lambda M_{\text{cycle}}T_{\text{cycle}}}(1 - P(N_b(T_n) \leq M_{\text{blocks}})e^{-\lambda T_{\text{cycle}}})\frac{N_{\text{cycle}}B}{T_{\text{cycle}}}$$
(Eq.29)

Transmission of frames with an adaptive length can improve the performance of the wireless link, however deriving the analytical form of data throughput when using block aggregation is a complicated task, thus we will assume that the number of blocks per frame is equal to 1, it means M_{blocks}=1, the instantaneous data throughput becomes:

$$R_s(T_{n+1}) = 2P(N_b(T_n) \leq 1)(1 - P(N_b(T_n) \leq 1)e^{-\lambda T_{\text{cycle}}})e^{-\lambda T_{\text{cycle}}}\frac{N_{\text{cycle}}B}{T_{\text{cycle}}} \quad \text{(Eq. 30)}$$

At the steady state and for $M_{\text{blocks}} = 1$, equation Eq. 25 becomes:

$$P(N_b = 0) = (P(N_b = 0) + P(N_b = 1))\, e^{\lambda T_{\text{cycle}}} \quad \text{(Eq. 31)}$$

The steady state probability of empty buffer and the probability of one frame in the buffer are given by [7]:

$$\begin{cases} P(N_b = 0) = 1 - \lambda T_{\text{cycle}} \\ P(N_b = 1) = (1 - \lambda T_{\text{cycle}})(e^{\lambda T_{\text{cycle}}} - 1) \end{cases} \quad \text{(Eq. 32)}$$

Eq.31 will be used to calculate the probability of an empty buffer and also non-empty buffer probability.

Note that condition $\lambda < \frac{1}{T_{\text{cycle}}}$ should be true to converge to the steady state, and the useful data throughput at the steady state for 4-nodes connectivity converges to:

$$R_{\text{cluster}} = 2\lambda(1 - \lambda T_{\text{cycle}})N_{\text{cycle}}B \quad \text{(Eq.33)}$$

In fact, the useful throughput depends also on synchronisation period T_S, respectively. The new expression of the useful throughput for 4-nodes connectivity is then given by:

$$R_{\text{cluster}} = 2\lambda(1 - \lambda T_{\text{cycle}})\frac{N_{\text{TX}}T_{\text{cycle}}}{N_{\text{TX}}T_{\text{cycle}} + T_S}N_{\text{cycle}}B \quad \text{(Eq. 34)}$$

Similarly, for 6-nodes connectivity, we assume the traffic arrival rate per sector is λ and 2λ for the central node and edge node, respectively. The number of neighbours nodes is 6 for the central node and 3 for the edge node, therefore the cluster throughput is given by:

$$R_{\text{cluster}} = 2\lambda(3 - 4\lambda T_{cycle})\frac{N_{\text{TX}}T_{\text{cycle}}}{N_{\text{TX}}T_{\text{cycle}} + T_S}N_{\text{cycle}}B \quad \text{(Eq.35)}$$

In order to calculate the throughput for 8-nodes connectivity, we assume that the number of neighbours of the central node is N_{cycle} =8, that 4 edge nodes have only 2 neighbours and 4 edge nodes have 4 neighbours, and the frame arrival rate per sector are λ for central node and, 4λ and 2λ for edge nodes with two neighbours and 4 neighbours respectively, then the total throughput per cluster is:

$$R_{\text{cluster}} = \frac{\lambda}{2}(5 - 6\lambda T_{\text{cycle}})\frac{N_{\text{TX}}T_{\text{cycle}}}{N_{\text{TX}}T_{\text{cycle}} + T_S}N_{\text{cycle}}B \quad \text{(Eq.36)}$$

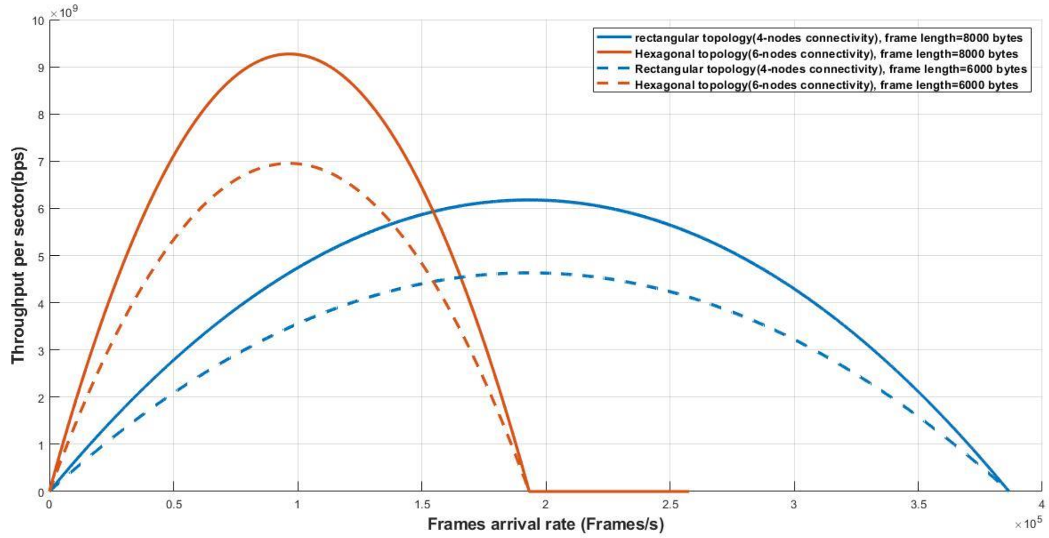

Figure 11 Throughput per sector for: 4-nodes connectivity, and 6 nodes connectivity topologies, for time based beam switching random access as function of frame arrival rate.

For each topology, the throughput per cluster is given by:

$$R_{\text{sector}} = \frac{R_{\text{cluster}}}{2N_{\text{cycle}}} \qquad \text{(Eq.36)}$$

In figure 9, we plot the throughput per sector for two different topologies: 4-nodes connectivity and hexagonal topology with 6-nodes connectivity. The hexagonal topology can reach a maximum throughput per sector of 9.3 Gbps and 7 Gbps using frame length equal to 8000 bytes and 4000 bytes respectively at a lower frame arrival rate. Whereas, for 4-node connectivity a maximum mean throughput of 6.1 Gbps and 4.8 Gbps per sector is achieved using a frame length of 8000 bytes and 4000 bytes respectively for a higher frame arrival rate. The hexagonal topology is convenient for low frame arrival rate scenarios and the second one is more suitable for a high frame arrival rate.

4.2 Medium access with RTS/CTS:

In this scheme, node exchange RTS/CTS frames before starting the data transmission. When the time to send data arrive for the transmitter node, itt first sends an RTS frame to the receiver. The receiver node on reception of the RTS packet, sends back the CTS packet for confirmation. The data transmission starts only when the transmitter receives a CTS packet. After receiving the data packet, the receiver sends the acknowledgment (ACK) for the data packet and its sequence number. The procedure is shown in Figure 10.

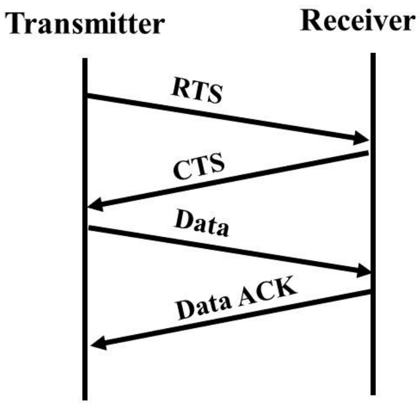

Figure 12 MAC scheme with RTS/CTS mechanism for data transmission.

The technique using RTS/CTS reduces the risqué of collision between central nodes and other nodes, however it adds additional time to transmit within a cluster, hence throughput can be reduced. The time dedicated to each sector to transmit is higher than the previous method; however, the transmission success rate will be improved.

The risk of collision can still happen if the two nodes, via their sector, intend to transmit during their time slot, and the transmission will be delayed for the next time slot. The delay can be also increase if the two sectors have data to transmit.

4.3 Medium access using fixed slot allocation

In the random method, each sector transmits or receives during its associated time slot; a collision can occur if the two nodes have data to transmit. Therefore, more time slots per cluster are required.

By using a fixed time allocation for transmission and reception for each sector, data throughput can be improved using fair allocation between sectors: during the transmission period, each sector has the same period of transmission and reception.

Using N_{cycle} timeslots per cycle will lead to low throughput for high frame arrival rate due to collisions. The solution is to separate between the TX and RX periods, such that each sector can fully use one timeslot for transmission during one cycle period without any risk of collision.

The throughput per sector for 4-nodes-connectivity is given by:

$$R_{\text{sector}} = \lambda \frac{N_{\text{TX}} T_{\text{cycle}}}{N_{\text{TX}} T_{\text{cycle}} + T_{\text{S}}} B \qquad \text{(Eq.37)}$$

For 8-nodes connectivity the throughput per sector is given by:

$$R_{\text{sector}} = 2\lambda \frac{N_{\text{TX}} T_{\text{cycle}}}{N_{\text{TX}} T_{\text{cycle}} + T_{\text{S}}} B \qquad \text{(Eq.38)}$$

And for 6-nodes connectivity, the throughput is:

$$R_{\text{sector}} = \frac{3}{2}\lambda \frac{N_{\text{TX}} T_{\text{cycle}}}{N_{\text{TX}} T_{\text{cycle}} + T_{\text{S}}} B \qquad \text{(Eq.39)}$$

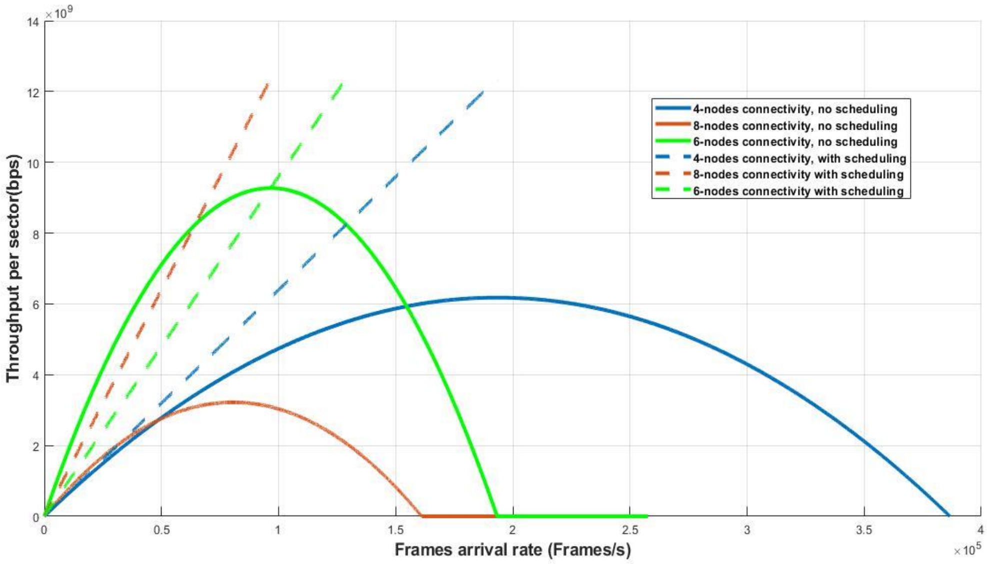

Figure 13 throughput per sector for: 4-nodes connectivity, 8 nodes connectivity and 6 nodes connectivity topologies, for both random access and transmissions with slot allocation as a function of frame arrival rate.

Figure 11 depicts the throughput per sector for 4-nodes connectivity, 6-nodes connectivity and 8-nodes connectivity topologies for both random and fixed slot scenario as a function of frame arrival rate per sector, the frame length is B=8000 bytes, the timeslot duration calculated according to maximum distance and frame length and is equal to 0.64668 μs (based on Eq.14), the distance between transmitter and receiver is 2 m and we assume the processing time is equal to zero. For fixed slot scenario, the hexagonal topology can reach the highest throughput for low frame arrival rate; the 4-nodes connectivity throughput can reach 6.2 Gbps per sector for higher frame arrival rate. The throughput in the fixed slot scenario, is higher than the throughput in random scenario, where collisions can happen, in scheduled scenario each sector can use one timeslot to transmit during a cycle period. The throughput is a linear function of frame arrival rate and constrained by: $\lambda_{\text{sector}} T_{\text{cycle}} < 1$. The 4-nodes connectivity topology can reach 12 Gbps per sector for a frame arrival rate of 20000 Frames/s. table 2 shows the numerical value of T_{cycle} for each use case.
The average throughput per node is the sum of throughputs per sector.

Table 2 The cycle period duration for each use case

	with and without RTS/CTS	fixed slot allocation
4-nodes connectivity	2.5867 μs	5.1734 μs
6-nodes connectivity	3.8801 μs	7.7602 μs
8-nodes connectivity	5.1734 μs	10.3469 μs

4.4 Medium access with re-transmission:

Re-transmission of corrupted frames can improves the wireless link, however, it affects the buffer load by increasing the frame arrival rate, the re-transmission affects also the end to end delay, the number of re-transmission can be reduced by using new coding schemes at the physical layer to boost error correction. The frame mean-arrival rate per sector when considering re-transmission depends on arrival rate of new frames coming from higher layer and non-received frames, the total frame arrival rate is given by the following equation:

$$\lambda_{\mathrm{a}} = \lambda + \frac{(1-(1-p)^{B})}{T_{\mathrm{cycle}}} \qquad \text{(Eq.40)}$$

Proof

Let m the number of frame arriving to the queue during one cycle period T_{cycle} and s the number of frame to be re-transmitted during one cycle, using Little's formula we can write:

$$\lambda_{a} = \frac{\sum_{k=1}^{\infty} kP(s+m=k)}{T_{\mathrm{cycle}}} \qquad \text{(Eq.41)}$$

The number of re-transmitted frames during one cycle entering to the buffer is $s \in \{0,1\}$, and:

$$\begin{cases} P(s=0) = (1-p)^{B} \\ P(s=1) = 1-(1-p)^{B} \end{cases} \qquad \text{(Eq.42)}$$

p is the bit error rate (BER). m is following a Poisson distribution of parameter λ, then:

$$P(s+m=k) = \frac{(\lambda T_{\mathrm{cycle}})^{k-1}}{(k-1)!}\left[\frac{\lambda T_{\mathrm{cycle}}}{k}(1-p)^{B} + 1-(1-p)^{B}\right]e^{-\lambda T_{\mathrm{cycle}}} \qquad \text{(Eq.43)}$$

$$\sum_{k=1}^{\infty} kP(s+m=k) = \lambda T_{\mathrm{cycle}} + 1-(1-p)^{B} \qquad \text{(Eq.44)}$$

Therefore, $\lambda_{a} = \lambda + \frac{(1-(1-p)^{B})}{T_{\mathrm{cycle}}}$.

The medium access data throughput when frame re-transmission is activated is given by:

- Useful throughput for 4-nodes connectivity
  - Random access:

    $$R_{\mathrm{sector}} = \lambda_{\mathrm{a}}(1-\lambda_{\mathrm{a}}T_{\mathrm{cycle}})\frac{N_{TX}T_{cycle}}{N_{TX}T_{cycle}+T_{S}}\frac{\log(\bar{l})}{\bar{l}-1}B \qquad \text{(Eq.45)}$$

    Where, $\bar{l}$ is the average number of retransmissions given by Eq.6
  - Fixed slot allocation

    $$R_{\mathrm{sector}} = \lambda_{\mathrm{a}}\frac{N_{\mathrm{TX}}T_{\mathrm{cycle}}}{N_{\mathrm{TX}}T_{\mathrm{cycle}}+T_{\mathrm{S}}}\frac{\log(\bar{l})}{\bar{l}-1}B \qquad \text{(Eq.46)}$$
- Useful throughput for 8-nodes connectivity
  - Random access:

    $$R_{\mathrm{sector}} = \lambda_{a}(3-4\lambda_{a}T_{\mathrm{cycle}})\frac{N_{\mathrm{TX}}T_{\mathrm{cycle}}}{N_{\mathrm{TX}}T_{\mathrm{cycle}}+T_{S}}\frac{\log(\bar{l})}{\bar{l}-1}B \qquad \text{(Eq.47)}$$
  - Fixed slot allocation

    $$R_{\mathrm{sector}} = 2\lambda_{a}\frac{N_{\mathrm{TX}}T_{\mathrm{cycle}}}{N_{\mathrm{TX}}T_{\mathrm{cycle}}+T_{S}}\frac{\log(\bar{l})}{\bar{l}-1}B \qquad \text{(Eq.48)}$$
- Useful throughput for 6-nodes connectivity
  - Random access:

    $$R_{sector} = \frac{\lambda_{a}}{4}(5-6\lambda_{a}T_{\mathrm{cycle}})\frac{N_{TX}T_{\mathrm{cycle}}}{N_{TX}T_{\mathrm{cycle}}+T_{S}}\frac{\log(\bar{l})}{\bar{l}-1}B \qquad \text{(Eq.49)}$$
  - Fixed slot allocation

    $$R_{sector} = \frac{3}{2}\lambda_{a}\frac{N_{\mathrm{TX}}T_{\mathrm{cycle}}}{N_{\mathrm{TX}}T_{\mathrm{cycle}}+T_{\mathrm{S}}}\frac{\log(\bar{l})}{\bar{l}-1}B \qquad \text{(Eq.51)}$$

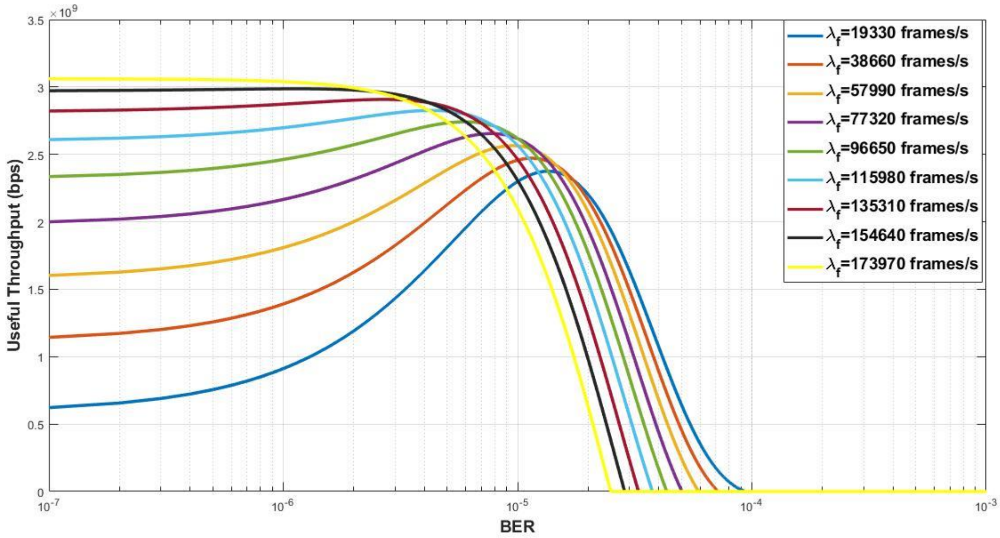

Figure 14 Useful throughput peer node sector as function of bit error probability for 4-nodes connectivity, the frame length is B=8000 bytes

Figure12 shows the sector throughput for 4-nodes connectivity for different frame arrival rate and BER,

4.5 Latency:

Latency in the WDCN network depends on the following factors:

- The service time with re-transmission $\bar{\tau}$
- The queuing time depending on frame inter-arrival time and the service time
- Average number of relays r between the source and destination

Using the Pollaczek-Khintchine theorem [10], the waiting time for a deterministic service time and exponential frame inter-arrival time the average waiting time of a frame for one link is given by:

$$\begin{cases} W = \bar{\tau} + \frac{\lambda\bar{\tau}^2}{2(1-\lambda\bar{\tau})} \\ \bar{\tau} = \bar{l}T_{\text{cycle}} \end{cases} \qquad \text{(Eq.52)}$$

The total average delay is:

$$D = rW \qquad \text{(Eq.53)}$$

The delay increases with the mean number of re-transmissions, and the total number of relays.
The mean number of re-transmissions increases with collisions and bit error probability; moreover, the number of relays decreases if the number of connectivity per node increases.

5 Network simulator

In this section, the network simulator for Terahertz band communication is described and results for analysing the performance of different medium access techniques and neighbour discovery are presented.

5.1 Simulator architecture

An NS3 based network simulator for Terahertz band communication is used to simulate the Centralized network architecture for Data Centre Network. The simulator mainly contains the channel, directional antenna and communication layers modules for Terahertz communication. First, the basics of the simulator architecture are described followed by the extensions. The details of this extension are described in [11].

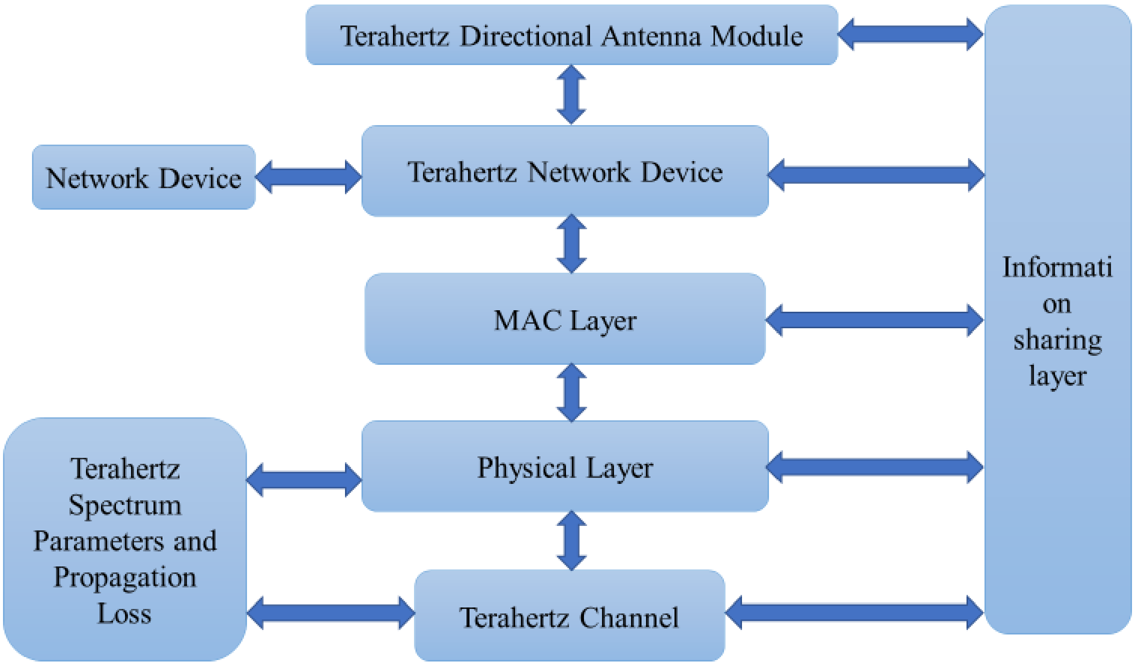

Figure 15 Simulator architecture and building blocks.

The general structure of simulator is shown in Figure 13. It contains the Node Objects, where each node is associated with the network device which is similar to Network Interface Cards (NIC) implementing the physical and MAC layer protocols. The spectrum module is used to generate carrier waveforms for macroscale scenarios. The channel module uses the propagation loss to calculate the spreading loss and absorption loss of Terahertz bands. The Network Device controls the Directional Antenna Module. A brief description of each module is presented here.

a. THz Network Device Module

It is the base class inherited from the class NetDevice of NS3 which can be used to implement different MAC protocols. It connects the Channel, Physical and MAC modules together with the directional antenna module.

b. THz Channel Module

It provides the channel to be used for communication protocols. It considers the waveform and propagation loss behaviour of the Terahertz channel. It is then obtaining the antenna gain of the transmitter and receiver devices from the directional antenna module to obtain the antenna directions, followed by the calculation of propagation loss. These values then passed to the

THz spectrum propagation loss to calculate the received power by using the received power spectral density. The THz channel module then passes the packet to the physical layer with the received power. It is designed to distinguish between the transmitter and receiver antenna to use the correct antenna gain and received signal power.

c. THz Spectrum Waveform and Propagation Loss module:

This module allows the user to select the preferred frequency window, central frequency and the bandwidth for a window. The generated transmitted signal is then used by the channel module and the calculation of propagation loss. The Terahertz channel frequency response H_c is used as:

$$H_c(f, d) = \left(\frac{c}{4\pi f d}\right) \exp\left(-\frac{k_{abs}(f)d}{2}\right) \quad \text{(Eq.54)}$$

Where c refers the speed of light and k_{abs} is the molecular absorption coefficient computed using the HITRAN database [12][13]. This module can also calculate the path loss which then is applied on the transmitted signal on each sub-band to achieve the receive signal and the received signal power.

d. Terahertz Directional Antenna module:

In this module, the cosine antenna module is used, whose gain is implemented in NS-3 [14]:

$$g(\varphi) = \sqrt{G_{max}} Cos^{\alpha}\left(\frac{\varphi-\varphi_0}{2}\right) \quad \text{(Eq.55)}$$

Where G_{max} is the antenna gain, φ represents the inclination angle, φ_0 represents the azimuthal orientation and the exponent α is used for the 3dB beam width as,

$$\alpha = \frac{-3}{20\log_{10}\left(\cos\left(\frac{\Delta\varphi_{3dB}}{4}\right)\right)} \quad \text{(Eq.56)}$$

$\Delta\varphi_{3dB}$ is the antenna beam width, the beam turning functionality is also implemented in the directional antenna module with different parameters like initial angel, turning speed, maximum gain and beam width. The feedback from this module can be used in the channel module. The transmitter and receiver antenna behave differently where the transmitter always points to the receiver and the receiver periodically sweeps the entire 360-degree space.

The antenna pattern implemented in ns-3 can approximate the phased array pattern mentioned In the theoretical part of this document using curve fitting, the current module will be updated with phased array model in the future

e. Terahertz Physical Layer module:

In this module, mainly the time duration is considered for the packet and if the receiver is able to receive the signal with sufficient power by comparing with the SINR threshold.

f. Terahertz MAC layer module:

In this module, the transmission modes like transmitter and receiver are used and switching between them. These parameters are calculated using the values from the directional antenna module. In transmission mode, the node always points the direction to the receiver node, whereas, in receiver mode the node sweeps the 360-degree space. The NAV is implemented which takes value from the data transmission duration from the packet header. The carrier sensing decisions are made based on the NAV value. The transmission happens when the channel is idle. Zero way and two way handshakes are implemented at MAC layer.

5.2 Packet flow

A node contains all the communication layers of a communication protocol stack. The application module generates the packet using the Poisson distribution. The generated packet is then passed to the transport layer using standard UDP sockets. After which the Network device is called which uses Enqueue () method in the MAC layer protocol. The enqueue method the prepare the packet structure with information like enqueued time, destination address, sequence number for the ongoing transmissions. This information is used when the ACK packet is received to calculate the throughput. Depending on the basic MAC protocol used as with RTS CTS or without, the packet is directed towards the physical layer also depending on the channel availability. It also checks is the receiver directional beam is facing towards the transmitter. The packets are transferred to physical layer by calling SendPacket () method. The physical layer then pass the packet to the channel layer using SendPacket () method at channel module. The channel calculates the received power for the node and schedules the packet reception after the propagation time which then calls the receive method of the physical layer. After the transmission time, the received packet done method is scheduled for the channel and physical layer module. The receive packet done method at physical layer checks for the collisions and calculates the SINR every time. The packet is dropped, if it is collided, otherwise the packet is passed back to the MAC layer. MAC layer confirms the destination address to finally receive the packet and schedules the ACK packet for the corresponding source.

5.3 Simulator extensions:

The simulator is extended as follows to implement the Data Centre scenario,

- Fixed node positions are implemented using a grid allocation for the centralized and Adhoc network architecture.
- The simulation parameters are updated according to the TERAPOD project outcomes.
- Devices and Antenna modules adapted to TERAPOD parameters
- Beam switching from sector to sector is implemented.
- Neighbor discovery mechanism is implemented.
- 4-nodes connectivity with central node is implemented and simulated.
- Medium access techniques with and without RTC/CTS and error control mechanism (re-transmission) is analyzed.

5.4 Evaluation parameters:

Following are the performance evaluation metrics used to evaluate the above network.
Throughput: The throughput is measured as follows,

$$\text{Throughput} = \frac{\text{Size of packet}}{\text{Time to receive a packet}}$$

Delay: The delay is measured as the difference of the time when the packet is first enqueue till the time it is successfully received at the receiver.
$Delay = Packet\ enqueue\ time - Packet\ reception\ time$**Packet delivery ratio (PDR):** It is measured as the ratio of the total number of packets received to the total number of packets sent in the network.

$$\text{PDR} = \frac{\text{Total number of packets received}}{\text{Total number of packets sent}}$$

Packets drops: It is the total number of packets dropped during the whole simulation time.
Number of retransmissions: It is the number of times packet is retransmitted in case the transmitter does not receive the ACK message from the received.

Average Neighbour Discovery Time: It is the average time when the first beacon is sent in the network till the last node receives the ACK packet from the central node.

5.5 Simulation results from NS-3

6.5.1 Simulation parameters:

Table 3 Parameters simulation set

Parameters	**Values**
Number of nodes	5 (Four transmitters and one receiver)
Frequency	300 GHz
Bandwidth	12.96 GHz
Number of bands	1
Data Rate	100 Gbps
Tx Power	-20 dBm
Tx/Rx antenna Gain	17.27 dBi
SINR Threshold	10 dB
Maximum frame Size	15000 bytes
Beam width	30 degree
Distances	1.4 m, 4.5 m
Noise level	1e-14 W
Simulation time	100 milliseconds
Mean Inter-arrival Time	1, 25, 50, 100, 500, 1000 µs
Average number of retransmissions	5
Beam stopping time	1 µs

Two scenarios are implemented in NS3 to analyse the throughput in different MAC schemes, a point to point and a 5 nodes cluster scenario. The simulations parameters are monitored at each reception for both with and without RTS CTS and with and without retransmissions. Initially, a point to point scenario is implemented in which both transmitter and receiver antenna directions are fixed toward each other. A five-node cluster scenario is presented later.

6.5.2 Point to point scenario results:

In this scenario, two nodes are used as a transmitter transmitterand receiver. The antenna of both transmitter transmitterand receiver are facing towards each other where both nodes antennas beams are facing towards each other. The mean inter-arrival time is fixed to 25 microsecond and simulation runs for 1 second time. In Figure 14, the average throughput is shown using two MAC strategies with and without RTS CTS handshake. The average throughput for with RTS CTS is less than the average throughput of without RTS CTS. The RTS CTS mechanism confirms the channel status and beam alignment before transmitting the packets which causes some delays, as also shown in Figure 15. The packet transmission time increases due to RTS CTS packets exchange and due to which the next packets waiting time in the queue increases. This overall affects the throughput of the network. Whereas, in without RTS CTS mechanism, the packet is sent without exchange the handshake packets, due to which the packet waiting time in the queue is less. This overall improves the throughput and delay for without RTS CTS mechanism. The beams are fixed which can cause extra delay in the packet delivery. Due to the fixed beam directions the delay is considerably low with high packet delivery ratio, as shown in Figures 14 and 15.

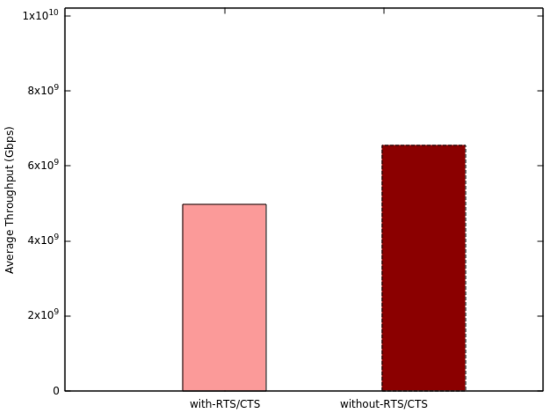

Figure 16: Average throughput for point to point scenario.

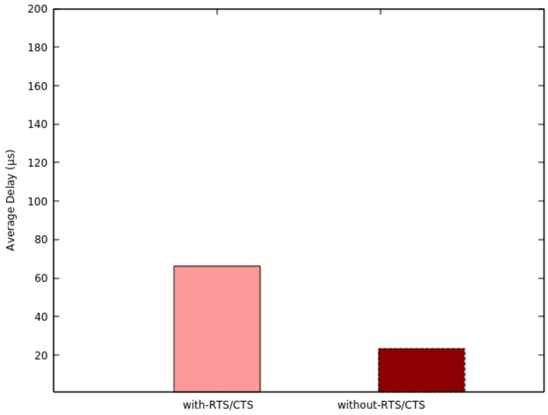

Figure 17: Average delay for point to point scenario.

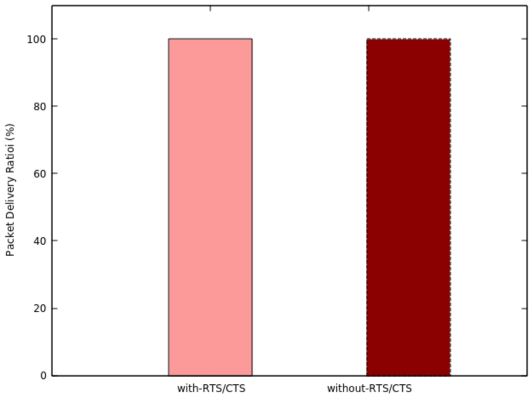

Figure 18: Packet delivery ratio for point to point scenario.

6.5.3 Five nodes cluster scenario results:

In five nodes scenario, the node positions are fixed. One node is fixed to work as a receiver or access point, whereas the other four nodes are fixed as transmitters. The transmitters beam always points towards the receiver or access point node. The receiver uses the beam switching and stays at each sector for 1 microsecond time before moving to the next sector. This time is sufficient to complete one transmission cycle including the handshake and data transmission. Each node is at the distance of 2 m from the receiver node. Due to lack of resources, 100 milliseconds windows is simulated for each different configuration and average is calculated at each successful reception of the packets at the receiver node. The simulations are repeated with and without retransmission for both the with and without RTS CTS handshake mechanisms. The number of retransmissions is set to 5. The averages of simulation parameters are taken for different mean inter arrival times from 1 to 1000 microseconds to analyse the throughput, delay and packet delivery ratio.

Analysis of throughput based on different mean interarrival times.

The traffic in a datacentre is random which can be high at times and low at other times. Due to this reason different mean arrival times are simulated in NS3 for a cluster architecture which ranges from 1 microsecond to 1000 microsecond. The low mean inter-arrival times means that the traffic is high and very frequently generated, and due to beam switching each node has less chance to deliver the data packet successfully to the access point. This waiting time in queue increases the delay as shown in Figure 16 and so can result in low throughput, as can be seen in Figure 17 for 1 microsecond mean inter-arrival time.

When inter-arrival time is higher the throughput is observed as high up to 9 Gbps for without retransmission scenario and up to 6 Gbps for with re-transmission scenario. The re-transmission improves the packet delivery as shown in Figure 17, in which the delivery ratio for with transmission is higher than without retransmission scenario. However, the delay is considerably low for higher mean inter-arrival times as less than 100 microseconds, as shown in Figure 18. The average number of packets received, and loss are shown in Figures 20 and 21. The retransmissions incidents are higher for without RTS CTS strategy, as shown in Figure 22 due to random packet transmission without handshake process.

With decreasing mean inter-arrival time the throughput also decreases and the delay increases, as shown in the Figures 17 and 18. The less delay is observed until 100 microsecond inter-arrival time, after which the delay tends to increase with delivery ratio and throughput tends to decrease. The best configuration was observed for with RTS CTS with and without retransmissions where higher throughput, less delay and considerably higher delivery ratio is observed.

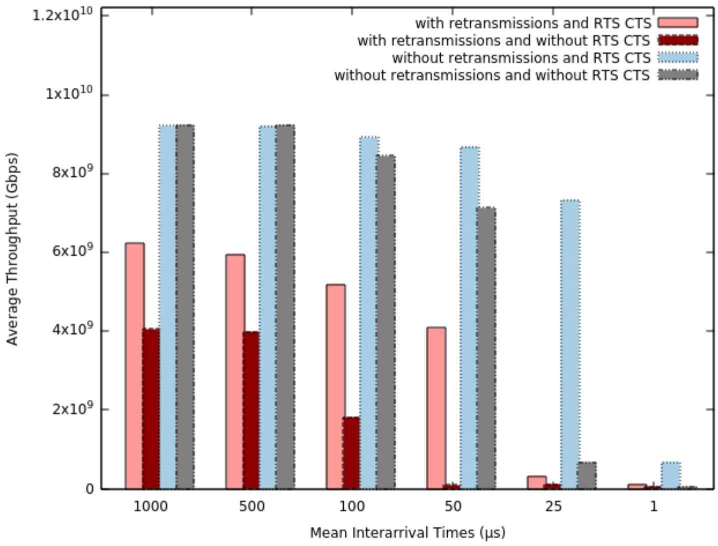

Figure 19: Average throughput per node for 5 nodes cluster.

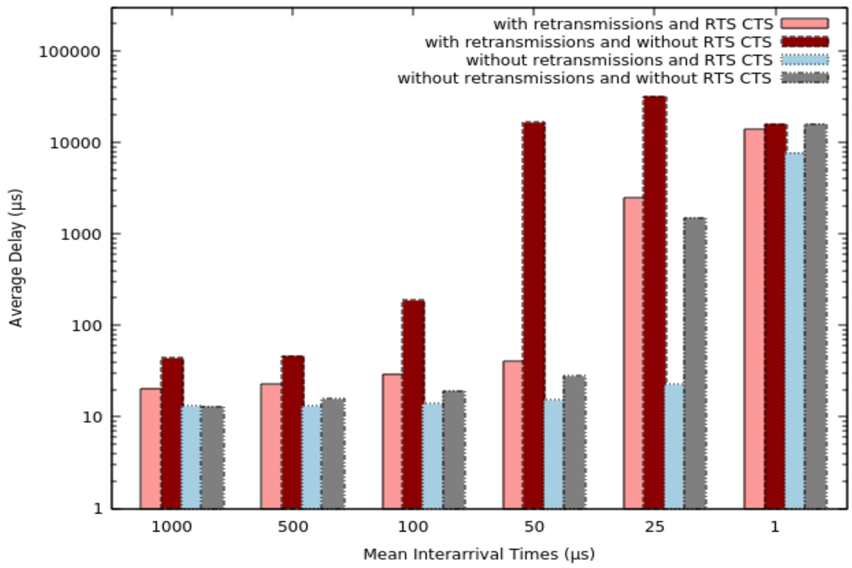

Figure 20: Average delay for 5 nodes cluster.

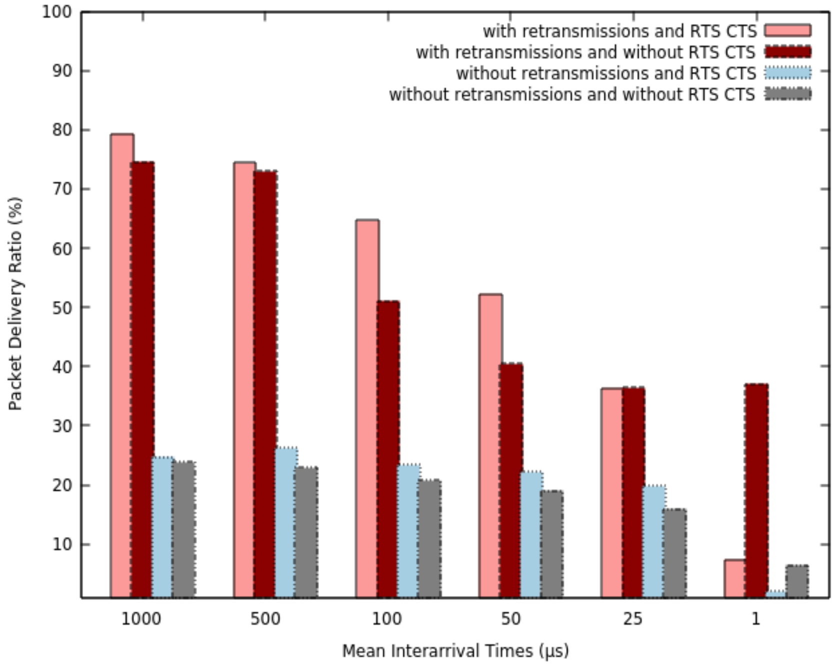

Figure 21: Packet delivery ratio for 5 nodes cluster.

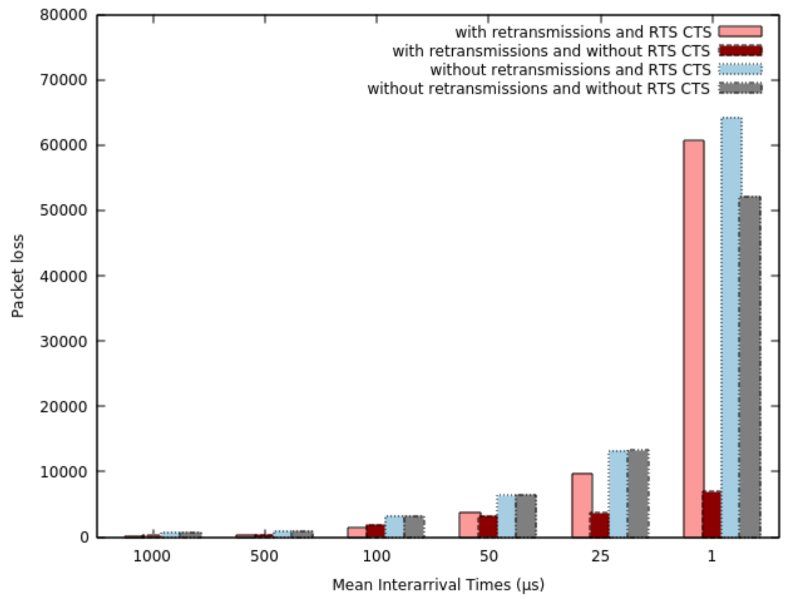

Figure 22: Average number of packets loss in 5 nodes cluster.

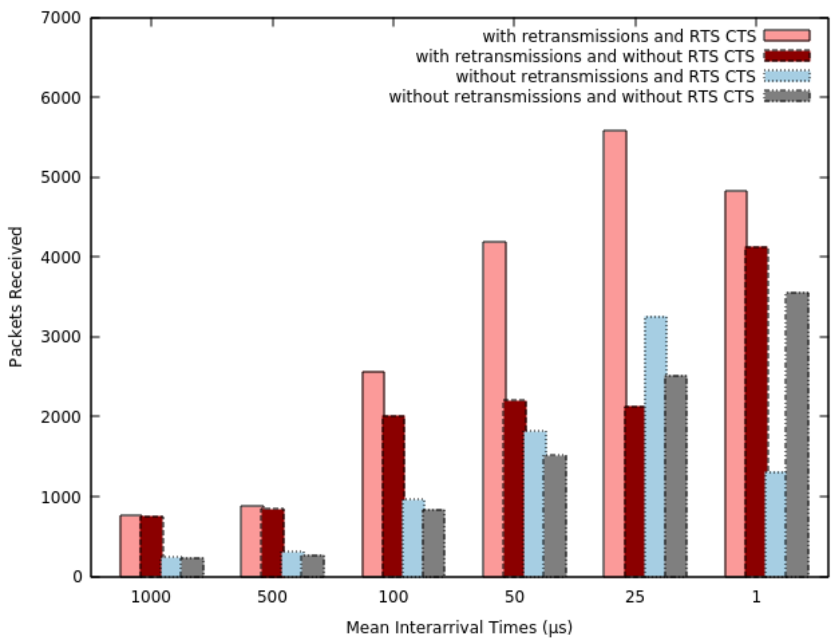

Figure 23: Average number of packets received in 5 nodes cluster.

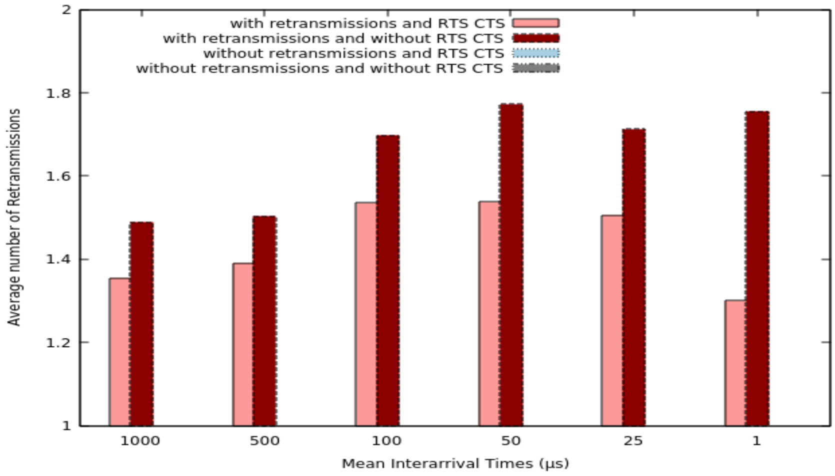

Figure 24: Average number of retransmission of packets.

6.5.4 Node discovery results:

A centralised node architecture is considered for the neighbour discovery scenario. The simulation was repeated for different number of nodes, where the nodes placement was fixed randomly in each simulation run. The transmission range was fixed as 2 meters. In this basic scenario, the central node is performing as receiver and other nodes behave as a transmitter node. The central node knows the total number of one hop nodes and stops the neighbour discovery and synchronisation process. For synchronisation the packet header includes the fields of source ID, antenna beam width and angle and sector repeating time.
In the neighbour discovery process the transmitter node sends a beacon message periodically in a particular direction towards the central node. The central node performs beam switching and moves from sector to sector to receive the beacon and send an acknowledgment for the beacon. The process stops when each node receives an acknowledgment of its discovery beacon.
The results for average discovery delay are shown in Figure 23. The time is almost linearly increasing with the increase in the number of nodes.

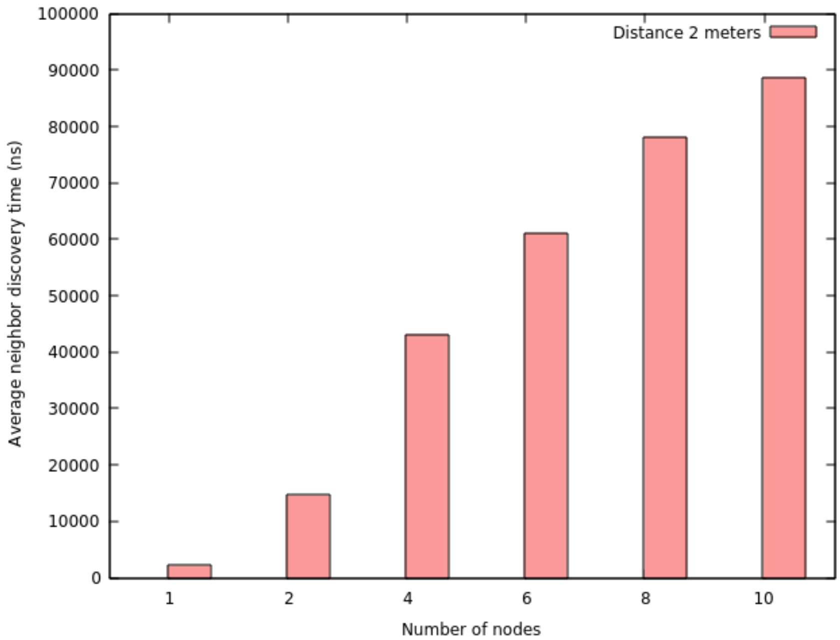

Figure 25: Average Neighbour Discovery Time for different nodes for 2-meter distance.

Currently, a two-way message exchange-based neighbour discovery mechanism is implemented. In future, three-way beacon handshake process for neighbour discovery will be implemented with beam switching functionality at both transmitter and receivers. A suitable strategy will also be proposed to minimize the overall neighbour discovery time.

6 Conclusion/Further work

The deliverable D5.4 gives a detailed description of data link layer of THz WDCN using mathematical modelling and simulation. Based on previous studies provided in D5.3 and new improvements, we propose the DLL layered models and possible node arrangements and connectivity. A theoretical study on the achieved throughput per sector shows that it is possible to meet WDCN requirements and reach high data rate by using short transmission times, re-transmissions and scheduling.

The DLL performance mainly depends up on the physical layer techniques and the antenna technology in which fast beam switching time and high antenna gain are required to reduce the discovery time and to meet the link budget requirements and reduce errors. Cross layer functionalities can be implemented at Network and Link layers for efficient network management and resource allocation.

The DLL modelling part includes preliminary results showing that a network using 4-nodes connectivity can achieve higher data throughput per sector and per node for high frame arrival rate. The delay in the network depends on frame arrival rate, transmission period and the average number of relays in the network.

The second part of this deliverable is devoted to DLL functionalities integration including node discovery and medium access and simulation using NS-3. The results from 5-nodes network using beam switching and TERAPOD parameters shows that it is possible to reach 10 Gbps The proposed architecture supports additional improvements to support new medium access and functionalities.

We conclude that since the WDCN has static topology, the time-based medium access approach is suitable to achieve ultra-high data rate and acceptable latency. Additional DLL layer design and parameter optimizations can be performed to reduce frame loss and reduce transmission overhead. Antenna and THz devices parameters can also be tuned to improve DLL performance.

As future work, multiple carrier communication can be considered, where each sector node can be assigned one or multiple sub-bands, the aim is to reduce interferences and increase system capacity. It is also possible to propose neew topologies where reflectors can be installed on roof or walls to increase node visibility and reduce relays node which can improve system delay

References

[1] ns-3 discrete-event network simulator for research and educational use - https://www.nsnam.org/
[2] Mininet network emulator, creating virtual networks on a single machine - http://mininet.org/
[3] CORE, Common Open Research Emulalator for network emulation on multiple machines - https://www.nrl.navy.mil/itd/ncs/products/core
[4] Sean Ahearne, Niamh O'Mahony, Noureddine Boujnah, Saim Ghafoor, Alan Davy, Luis Gonzalez Guerrero, Cyril Renaud "Integrating THz Wireless Communication Links in a Data Centre Network" accepted, IEEE 5G World Forum, Dresden, Germany, 30 September to 2 October 2019.
[5] IEEE 802.15.3d IEEE Standard for High Data Rate, Wireless Multi-Media Networks, Amendment 2: 100 Gb/s Wireless, Switched Point-to-Point Physical Layer
[6] X.W.Yao, J.Jornet "TAB-MAC: Assisted beamforming MAC protocol for Terahertz communication networks" Nano communication Network(Elsiever) Journal, vol. 9, pp.36-42,2016
[7] Philip Wette, Holger Karl, DCT2Gen: A traffic generator for data centres, Computer Communications, Volume 80, 2016, Pages 45-58, ISSN 0140-3664,
[8] Nakagawa, Kenji (2005). "On the Series Expansion for the Stationary Probabilities of an M/D/1 queue". Journal of the Operations Research Society of Japan. 48 (2): 111–122. doi:10.15807/jorsj.48.111.
[9] X. Deng, Y. Li, J. Li, C. Liu, W. Wu and Y. Xiong, "A 320-GHz 1$\times$4 Fully Integrated Phased Array Transmitter Using 0.13-μm SiGe BiCMOS Technology," in IEEE Transactions on Terahertz Science and Technology, vol. 5, no. 6, pp. 930-940, Nov. 2015, doi: 10.1109/TTHZ.2015.2477604.
[10] S. Rey, T. Merkle, A. Tessmann and T. Kürner, "A phased array antenna with horn elements for 300 GHz communications," *2016 International Symposium on Antennas and Propagation (ISAP)*, Okinawa, 2016, pp. 122-123.
[11] Cooper, Robert B. (1981). Introduction to Queuing Theory. Elsevier Science Publishing Co. p. 189. ISBN 0-444-00379-7.
[12] Jornet J.M., Akyildiz I.F.Channel modeling and capacity analysis of electromagnetic wireless nanonetworks in the terahertz band, IEEE Trans. Wireless Commun., 10 (10) (2011), pp. 3211-3221
[13] Zahed Hossain, Qing Xia, Josep Miquel Jornet, "TeraSim: An extension to simulate Terahertz Band Communication Networks", Elsevier Software Impacts, Volume 1, 2019.
[14] L. Chujian, Cosine Antenna Module, 2003. https://www.nsnam.org/doxygen/classns3_1_1_cosine_antenna_model.html
[15] Noureddine Boujnah, Saim Ghafoor, Alan Davy "Modeling and Link Quality Assessment of THz Network Within Data Centre" 2019 European Conference on Networks and Communications (EuCNC), Valencia, Spain