Beyond 5G Multi-Tenant Private Networks Integrating Cellular, Wi-Fi, and LiFi, Powered by Artificial Intelligence and Intent Based Policy

5G-CLARITY Deliverable D3.2

Design Refinements and Initial Evaluation of the Coexistence, Multi-Connectivity, Resource Management and Positioning Frameworks

Contractual Date of Delivery:	**April 30, 2021**
Actual Date of Delivery:	**May 31, 2021**
Editor(s):	**Tezcan Cogalan (IDCC)**
Author(s):	**Daniel Camps-Mur (I2CAT),** **Antonio Garcia, Kiran Chackravaram (ACC),** **Jorge Navarro-Ortiz, Juan Ramos-Munoz, Juan M. Lopez-Soler (UGR),** **Stefan Videv, Anil Yesilkaya, Hamada Alshaer (USTRATH),** **Srinivasan Raju, Rui Bian (PLF),** **Jesús Gutiérrez, Vladica Sark, Meysam Goodarzi (IHP),** **Mir Ghoraishi (GIGASYS)**
Work Package:	**WP3**
Target Dissemination Level:	**Public**

This document has been produced in the course of 5G-CLARITY Project. The research leading to these results received funding from the European Commission H2020 Programme under grant agreement No. H2020-871428. All information in this document is provided “as is", there is no guarantee that the information is fit for any particular purpose. The user thereof uses the information at its own risk and liability. For the avoidance of all doubts, the European Commission has no liability in respect of this document, which is merely representing the authors view.

Revision History

Revision	Date	Editor /Commentator	Description of Edits
0.1	25.01.2021	Tezcan Cogalan (IDCC)	D3.2 ToC and contributors
	03.02.2021	Mir Ghoraishi (GIGASYS)	Master doc and Sections templates
0.2	28.04.2021	Tezcan Cogalan (IDCC)	Individual section documents are merged to create the first full draft.
0.3	04.05.2021	Tezcan Cogalan (IDCC)	References are added/updated. Cross-reference to Figures/Tables are fixed. The list of acronyms is added. Some editorial comments are provided. Review up to Section 5 is finalized.
0.4	11.05.2021	Tezcan Cogalan (IDCC)	Measurements results for Section 2.1.1 and Section 4 are added. The list of acronyms is updated. Full review comments are provided – some comments are already addressed. Executive summary and conclusions are provided.
0.5	17.05.2021	Mir Ghoraishi (GIGASYS)	Full review
	24.05.2021	Tezcan Cogalan (IDCC)	Recent updates from Jorge (UGR) are consolidated with Mir's review.
0.6	26.05.2021	Tezcan Cogalan (IDCC)	Resolved comments are deleted. Some editorial changes.
1.0	30.05.2021	Jesus Gutierrez (IHP) Mir Ghoraishi (GIGASYS)	Final proof-reading and final version issue

Table of Contents

List of Figures

List of Tables

List of Acronyms

3GPP	third Generation Partnership Project
5GNR	5G New Radio
5GC	5G Core
AF	Application Function
AGC	Automatic Gain Control
AGV	Automated Guided Vehicle
AMC	Adaptive Modulation and Coding
AMF	Access and Mobility Management Function
AMPDU	Aggregated MAC PDU
AoA	Angle of Arrival
AoD	Angle of Departure
AP	Access Point
AQL	Adaptive Queue Limits
ARP	Address Resolution Protocol
ASK	Amplitude Shift Keying
ATSSS/AT3S	Access Traffic Steering Switching and Splitting
BALIA	Balanced Linked Adaptation
BLER	BLock Error Rate
BRF	Bayesian Recursive Filtering
BW	Bandwidth
CDF	Cumulative Distribution Function
CMOS	Complementary Metal Oxide Semiconductor
COTS	Commercial Off-The-Shelf
CP	Control Plane
CPE	Customer Premise Equipment
CPU	Central Processing Unit
CQI	Channel Quality Indicator
CSI	Channel State Information
CU	Central Unit
CUPS	Control and User Plane Separation
dB	decibel
DC	Direct Current
DCO-OFDM	DC-biased Optical OFDM
DHCP	Dynamic Host Configuration Protocol
DL	Downlink
DNM	Delay-based Network Maximization
DU	Distributed Unit
EAP	Extensible Authentication Protocol
eAT3S	5G-CLARITY enhanced A3TS
eMBB	enhanced Mobile Broadband
eMTC	enhanced Machine Type Communication
eNB	enhanced Node B
ERP	EAP Re-authentication Protocol
ESP	Encapsulated Security Payload
EU	Europe
FAPI	Femto-Application Platform Interface
FIFO	First Input-First Output
FOV	Field-of-view
FT	Fourier Transform
gNB	Next generation Node B
GPRS	General Packet Radio Service
GRE	Generic Routing Encapsulation
GTP/GTP-U	GPRS Tunnelling Protocol-User plane
HoL	Head-of-Line
IM/DD	Intensity Modulation/Direct Detection
ICMP	Internet Control Message Protocol
IP	Internet Protocol
IPSec	IP Security
IR	Infrared
ISAKMP	Internet Security Association and Key Management Protocol
KPI	Key Performance Indicator
LAN	Local Area Network
LD	Laser Diode
LED	Light Emitting Diode
LIA	Linked-Increases Algorithm
LLC	Logical Link Control
LoS	Line of Sight
LTE	Long Term Evolution
MAC	Medium Access Control
MA-PDU	Multi Access PDU
MIMO	Multiple Input Multiple Output
ML	Machine Learning
mMTC	massive Machine Type Communication
MNO	Mobile Network Operator
MPTCP	Multi-Path TCP
MCS	Modulation and Coding Schemes
MSE	Mean Squared Error
MVNO	Mobile Virtual Network Operators
N3AN	Non 3GPP Access Networks
N3IWF	Non-3GPP Interworking Function
NF	Network Function
NLoS	Non-Line-of-Sight
NLLSQ	Non-Linear Least Squares
NR	New Radio (5GNR)
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access
OLIA	Opportunistic Linked-Increases Algorithm
O-RAN	Open RAN

OTDOA	Observed Time Difference of Arrival
PCF	Policy Charging Function
PD	Photodiode
PDF	Probability Distribution Function
PDU	Protocol Data Unit
PF	Particle Filter
PFS	Proportional Fair Scheduler
PHY	Physical Layer
PMF	Probability Mass Function
PPP	Poisson point process
PPS	Pulse Per Second
PRB	Physical Resource Block
PRU	Physical Resource Unit
PWM	Pulse Width Modulation
QAM	Quadrature Amplitude Modulation
QoS	Quality of Service
RAN	Radio Access Network
RBG	Resource Block Group
RF	Radio Frequency
RIC	RAN Intelligence Controller
RMS	Root Mean Square
RoI	Region of Interest
RR	Round Robin
RRM	Radio Resource Management
RSRP	Reference Signal Received Power
RSRQ	Reference Signal Received Quality
RSS	Received Signal Strength
RSSI	Received Signal Strength Indicator
RT	Real-time
RTOA	Relative Time of Arrival
RTOF	Round time of Flight
RTS	Request To Send
RTT	Round Trip Time
RU	Radio Unit
Rx	Receiver
SA	Stand-Alone
SDN	Software Defined Networking
SDR	Software Defined Radio
SINR	Signal-to-Interference plus Noise Ratio
SLA	Service Level Agreement
SMF	Session Management Function
SRTT	Smoothed Round-Trip-Time
SSB	Single Sideband
SSID	Service Set Identifier
TCP	Transmission Control Protocol
TEID	Tunnel End Point Identity
TDoA	Time Difference of Arrival
TID	Traffic Identifier
TNGF	Trusted Non-3GPP Gateway Function
TNM	Throughput-based Network Maximization
ToA	Time of Arrival
ToF	Time of Flight
Tx	Transmitter
TTI	Transmission Time Interval
TWR	Two Way Ranging
UDP	User Datagram Protocol
UE	User Equipment
UK	United Kingdom
UL	Up Link
UP	User Plane
UPF	User Plane Function
URLLC	Ultra Reliable Low Latency Communications
US	United States
UTDOA	Uplink Time Difference of Arrival
VM	Virtual Machine
VL	Visible Light
VLAN	Virtual Local Area Network
VLC	Visible Light Communication
VLP	Visible Light Positioning
VPN	Virtual Private Network
WAT	Wireless Access Technology
WLAN	Wireless LAN
WRR	Weighted Round Robin

Executive Summary

This document, 5G-CLARITY D3.2, aims to provide evaluation results and refinements on the initially designed 5G-CLARITY user and control plane architecture that is introduced in 5G-CLARITY D3.1 [1]. This document is also aligned with the "network function and application stratum" that covers not only user- and control-plane but also application plane functionality as presented in 5G-CLARITY D2.2 [2].

In essence, 5G-CLARITY D3.2 provides the performance evaluations and refinements for:

- **Multi-WAT aggregation**: Including 5GNR CU/DU/RU integration, integration of Wi-Fi and LiFi networks as a single non-3GPP network, integration of 3GPP and non-3GPP wireless access technologies (WATs) and assignment of traffic flows via MPTCP;
- **5G-CLARITY eAT3S framework**: Including operational flows, initial enhanced access traffic steering, switching and splitting (enhanced AT3S / eAT3S) algorithm design and control plane aspects of the custom MPTCP scheduler;
- **Scheduling and resource management**: Including Wi-Fi and LiFi airtime-based schedulers and utility-based scheduler to manage different service types;
- **Positioning**: Including WAT-specific positioning scheme and its performance evaluations, as well as the fusion approach;
- **Integrated 5G/Wi-Fi/LiFi network performance evaluation**: Including possible access point (AP)/gNB deployment options, achievable communication bandwidths, technology-specific area-capacity achievements and integrated network area-capacity performance.

Details for the 5G-CLARITY multi-connectivity framework evaluation are presented in Section 2. The 5G-CLARITY multi-connectivity design includes, i) the multi-WAT aggregation, integrating 3GPP (5GNR) and non-3GPP (Wi-Fi and LiFi) access networks, and ii) an enhancement on the AT3S scheme to improve the (multiaccess based) multi-connectivity functionalities. The details of design and validation of these features for the 5G-CLARITY user- and control-plane are provided.

Section 3 delivers discussions on AP level and service level (traffic routing) resource scheduling techniques. Primarily, the corresponding telemetry and performance measurements are used to route the traffic across 3GPP/non-3GPP networks in near real-time (near-RT) using 5G-CLARITY eAT3S introduced to ensure quality-of-service (QoS), and as a following step, the AP level resource scheduling is performed by the gNB and/or Wi-Fi/LiFi AP. In this respect, a Linux-kernel based airtime management evaluation framework is discussed which can be used to segregate multi-WAT resources for a given 5G-CLARITY slice. Due to LiFi's different channel and link reliability characteristics, the airtime scheduling for the LiFi technology is specifically discussed and slicing the attocellular network resources is researched.

Section 4 is focused on 5G-CLARITY multi-WAT positioning solution. The associated technologies are 60 GHz mWave, sub-6 GHz, LiFi and Optical Camera Communications (OCC) based positioning. A localisation server obtains the position information from these WATs, and provides the position estimate, by fusing all the relevant data, to the entities requiring position services. Details of the overall architecture, each technology ranging/positioning scheme, and the fusion approach are provided.

The simulation architecture to evaluate the integration and performance of 5G-CLARITY multi-WAT scheme, including the corresponding user- and control-plane functionalities are presented in Section 5. Results for a dense deployment of multi-WAT AP/gNB in contrast to the generic scenario, using both conservative (based on the available technologies) and opportunistic (assuming greedy usage of available bandwidth), are presented. The achievable system area capacity in each scenario is discussed and the limiting factors are

introduced.

Overall, this document, 5G-CLARITY D3.2, presents the achievable KPIs of the main components of the 5G-CLARITY integrated 5G/Wi-Fi/LiFi network user- and control-plane architecture. The integration of these components and the evaluation of the overall 5G-CLARITY user- and control-plane will be reported in 5G-CLARITY D3.3.

1 Introduction

This document, 5G-CLARITY D3.2, provides refinements on the initially designed 5G-CLARITY control and user plane architecture reported in 5G-CLARITY D3.1 [1], and reports initial evaluation results of the control and user plane architecture. Therefore, this deliverable is aligned with the proposed control and user plane architecture in 5G-CLARITY D3.1 [1]. Moreover, this deliverable is also aligned with the 5G-CLARITY architectural principles presented in 5G-CLARITY D2.2 [2] on "Network Function and Application Stratum". The refinements and evaluations in this deliverable are on:

i) the performance of signalling interfaces between different wireless access technologies (WATs) that establishes coexistence/integration;

(ii) the multi-connectivity performance on achievable throughput, latency and reliability of a single user equipment (UE);

(iii) the performance of the UE-access point (AP) association algorithms, software defined networking (SDN) intelligence interface selection and smart resource allocation algorithms that enhances aggregated area-capacity; and,

(iv) the performance of the indoor localization performance that improves sub-meter positioning accuracy and reliability key performance indicators (KPIs) of the integrated 5G/Wi-Fi/LiFi network.

1.1 Scope of this document

5G-CLARITY D3.2 takes input from 5G-CLARITY D2.2 [2] and 5G-CLARITY D3.1 [1], and provides output to WP2 (T2.2), WP3, WP5 (T5.2) and WP6. A successful delivery of 5G-CLARITY D3.2 marks the acceptance criteria of the project's milestone M3.2: "*Achievable KPIs of the proposed user and control plane frameworks of the integrated 5G/Wi-Fi/LiFi network are extracted and available*".

In 5G-CLARITY D3.1 [1], the 5G-CLARITY control and user plane architecture has addressed subjects on (i) non-3GPP access to 5G core; (ii) spectrum access for non-public and public network coexistence; (iii) multiaccess based multi-connectivity; (iv) resource management as traffic routing on access network level and as resource scheduling on AP level; and (v) indoor positioning. In this deliverable, the focus is on subjects on non-3GPP access/integration, multiaccess based multi connectivity, resource management and indoor positioning; and the spectrum access for non-public and public network coexistence will be covered in the upcoming deliverable, 5G-CLARITY D3.3. The rationale behind this approach is to refine and evaluate the very core functionalities of 5G-CLARITY user and control plane architecture that significantly affect the progress in other work packages, such as 5G-CLARITY WP5 (T5.2).

In 5G-CLARITY D3.2, the refinements and evaluation on non-3GPP access and multiaccess based multi-connectivity subjects are covered within the same section, as the control and user plane functions for both subjects impact each other. Therefore, the combination of both subjects are refined and evaluated by means of multi-connectivity framework, control plane functions/algorithms and user plane functions/algorithms. The refinements and evaluation on resource management subject is covered in two different sections where in one Section 3 the algorithms are described in more details and in Section 5 a simulation platform that will be used to evaluate the performance of the integrated 5G/Wi-Fi/LiFi network and proposed resource management algorithms will be described and the evaluation results will be reported.

Any refinement identified in this 5G-CLARITY D3.2 will also be reflected in the upcoming deliverables, i.e., 5G-CLARITY D2.4 and 5G-CLARITY D3.3, where the final 5G-CLARITY system architecture and user and control plane will be provided.

1.2 Objectives of this document

The specific objectives of this 5G-CLARITY D3.2 are as follows (mapped to overall project objectives described in the project's description of work (DoW), and referenced below as **OBJ-TECH-*XX*)**:

- **OBJ-1:** Validation of the initially designed 5G/Wi-Fi/LiFi multi-connectivity framework (**OBJ-TECH-3**).
- **OBJ-2:** Evaluation of the initially designed integrated 5G/Wi-Fi/LiFi network control plane algorithms and signalling between control plane architectural components (**OBJ-TECH-3**).
- **OBJ-3:** Evaluation of the initially designed integrated 5G/Wi-Fi/LiFi network user plane algorithms and signalling between user plane architectural components (**OBJ-TECH-3**).
- **OBJ-4:** Evaluation of the simulation environment that validates the enhancements of the aggregated system area capacity via resource management algorithms (**OBJ-TECH-4**).
- **OBJ-5:** Evaluation of the positioning protocols and algorithms (**OBJ-TECH-5**).

1.3 Document structure

The rest of this document is structured as follows:

- Section 2 covers **OBJ-1**, **OBJ-2** and **OBJ-3**, and provides details on the evaluation methodologies of the integrated 5G/Wi-Fi/LiFi network multi-connectivity framework, user and control plane functions/algorithms. It also reports achievable KPIs of the multi-connectivity framework.
- Section 3 partly covers **OBJ-4** by providing details of the considered resource management algorithms along with their implementation specifications.
- Section 4 covers **OBJ-5** and describes not only technology-specific positioning algorithms but also the approach on fusing positioning information from all available technologies. KPIs of the 5G-CLARITY indoor positioning approach are reported in this section.
- Section 5 covers **OBJ-4** and describes the system level evaluation platform that is used to evaluate the integrated 5G/Wi-Fi/LiFi network capabilities by means of aggregated throughput and area spectral efficiency. The section reports the achievable aggregated area capacity of the integrated 5G/Wi-Fi/LiFi network and highlights 5G-CLARITY integrated network performance.
- Section 6 summarizes and concludes this deliverable.

2 Multi-Connectivity Framework

This section describes details of the evaluation procedures of the initially designed 5G-CLARITY multi-connectivity framework. In 5G-CLARITY D3.1 [1], this multi-connectivity framework is designed on top of a user and control plane architecture that adopts the control and user plane separation (CUPS) principles of 3GPP and the reference architecture and interfaces of the open radio access network (O-RAN). In this section, the initially designed multi-connectivity framework is considered under two main subjects namely, multi-WAT aggregation that covers integration of 3GPP and non-3GPP access networks; and multiaccess based multi-connectivity that is mainly focussed on access traffic steering, switching and splitting (ATSSS).

2.1 Multi-WAT aggregation

As noted in 5G-CLARITY D3.1 [1] and 5G-CLARITY D2.2 [2], the proposed network and application functions stratum supports virtualized 5G base station (gNB) functions such as distributed units (DUs) and centralized units (CUs). The integration of these virtualized DUs/CUs can be supported by leveraging O-RAN reference architecture and interfaces. It is also noted in 5G-CLARITY D3.1 [1] and 5G-CLARITY D2.2 [2] that 5G-CLARITY uses a 3GPP interworking/gateway function namely non-3GPP interworking function (N3IWF) or trusted non-3GPP gateway function (TNGF) to establish a direct connection for non-3GPP access networks to 5G core. In 5G-CLARITY, a non-3GPP network consists of Wi-Fi and LiFi networks. Therefore, a very first step for 5G-CLARITY multi-WAT aggregation (or integration) is to integrate Wi-Fi and LiFi networks in order to have a single non-3GPP access network. Then, the non-3GPP access network can be integrated to 5G core (5GC) network via N3IWF/TNGF to aggregate multi-WAT and establish an integrated 5G/Wi-Fi/LiFi access network.

In order to evaluate different elements of multi-WAT aggregation in more details, this section covers an integration process of 5GNR DU/CU (3GPP), an integration process of Wi-Fi and LiFi networks (non-3GPP), and an integration process of 3GPP and non-3GPP access networks separately.

2.1.1 5GNR CU/DU/RU integration

In Section 6.1 of 5G-CLARITY D3.1 [1] we introduced the use of an O-RAN Alliance aligned reference architecture shown in **Figure 2-1** to be used for the implementation of the control plane, management plane, user plane and intelligence plane functions to incorporate the use of 5GNR access technology in 5G-CLARITY multi-WAT architecture as per disaggregated open RAN paradigm.

Accelleran dRAX™ (the marketing name for Accelleran´s virtual RAN product line) near real-time (near-RT)-RAN intelligence controller (RIC) and CU control plane (CP) and CU user plane (UP) functions have been integrated together southbound with functions from a DU partner from the O-RAN ecosystem using 3GPP standardised F1 interface and northbound with 5GC functions based on Open5GS opensource project [1].

The DU function is composed of L2 functionality and upper L1 functionality using a femto-application platform interface (FAPI) interface bundled together that integrates southbound with a 5GNR radio unit (RU) using a fronthaul interface based on a 7.2x Fronthaul Interface as depicted in Figure 2-2.

The validation of the full end-to-end integrated system was done in several steps in Accelleran System and Target Integration laboratory using an x86 AMD based Ryzen 9 commercial off-the-shelf (COTS) server.

[1] https://open5gs.org/

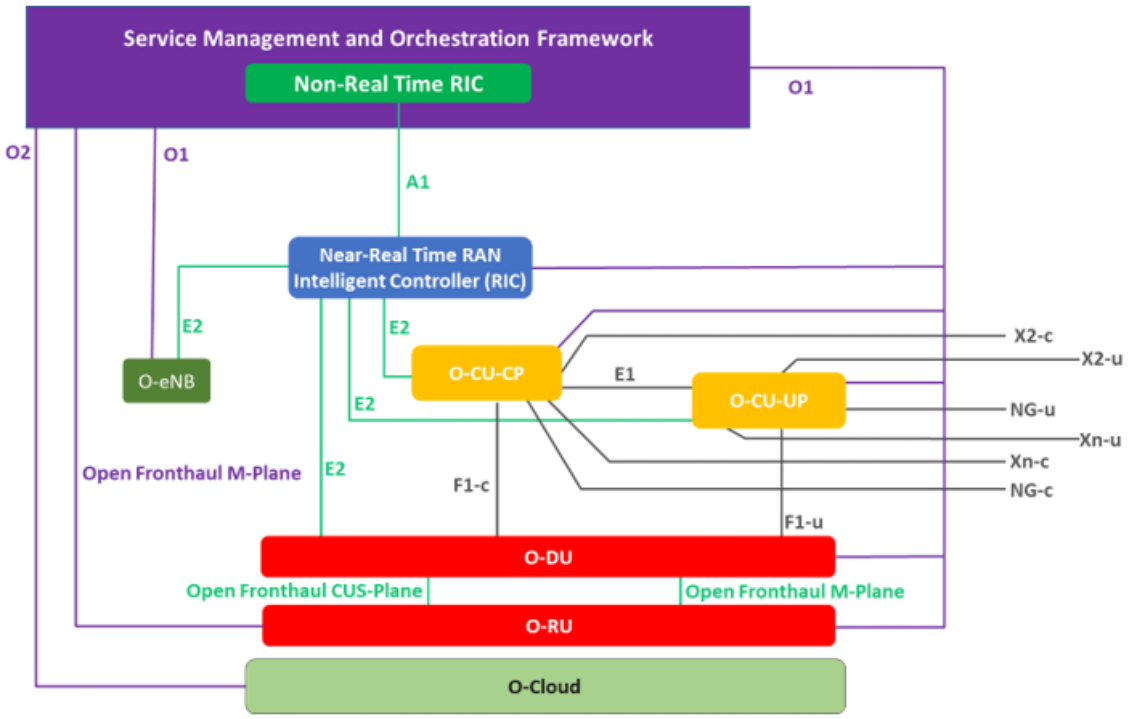


Figure 2-1: O-RAN overall logical architecture

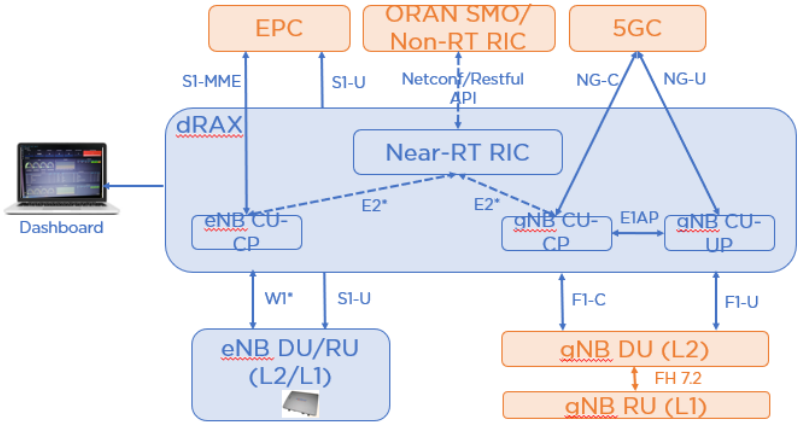


Figure 2-2:Accelleran dRAX disaggregated open RAN with third party 5GNR DU and RU

2.1.1.1 Phase 1: F1 and NG integration with virtual UE

During this phase, the Ryzen 9 server was used to host Open5GS, Accelleran dRAX near-RT RIC/CU-CP/CU-UP and third-party DU components tested against a virtual UE implementation as shown in **Figure 2-3**. This phase was done to first integrate the developed CU components together with the third-party DU over F1 using the virtual UE. This was done first using a commercial 5GC followed by a validation of the interoperability over NG with Open5GS.

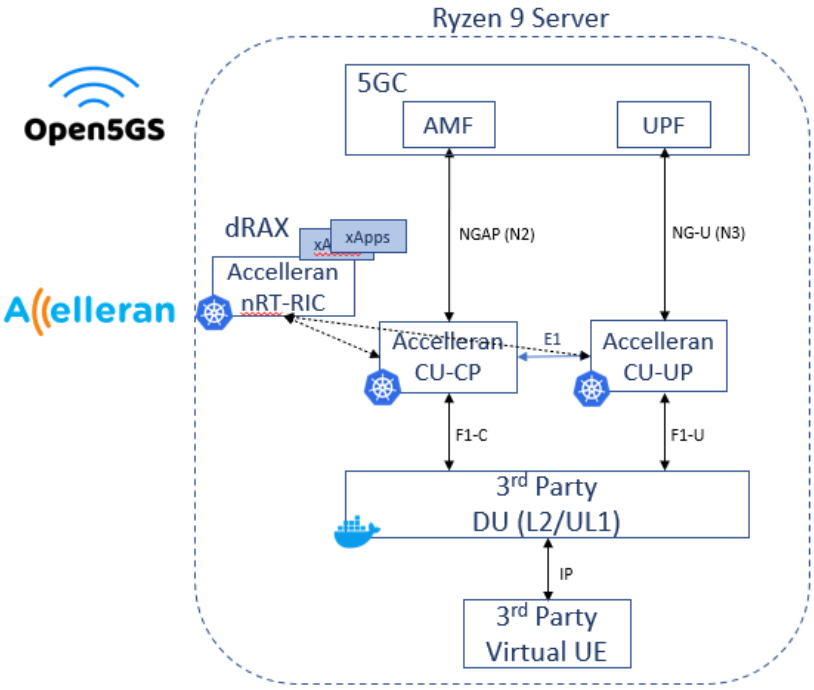


Figure 2-3: Phase 1 F1 and NG integration

2.1.1.2 Phase 2: NR-UU testing with Research SDR RU and commercial UE

This integration phase was based on the setup of Phase 1 where the DU component was integrated and connected with a National Instruments B210 software-defined radio (SDR) RU tested with OnePlus Nord smartphone as shown in Figure 2-4. This phase was done to enable integration and interoperability testing via B210 SDR RU with the OnePlus smartphone using NR-UU interface.

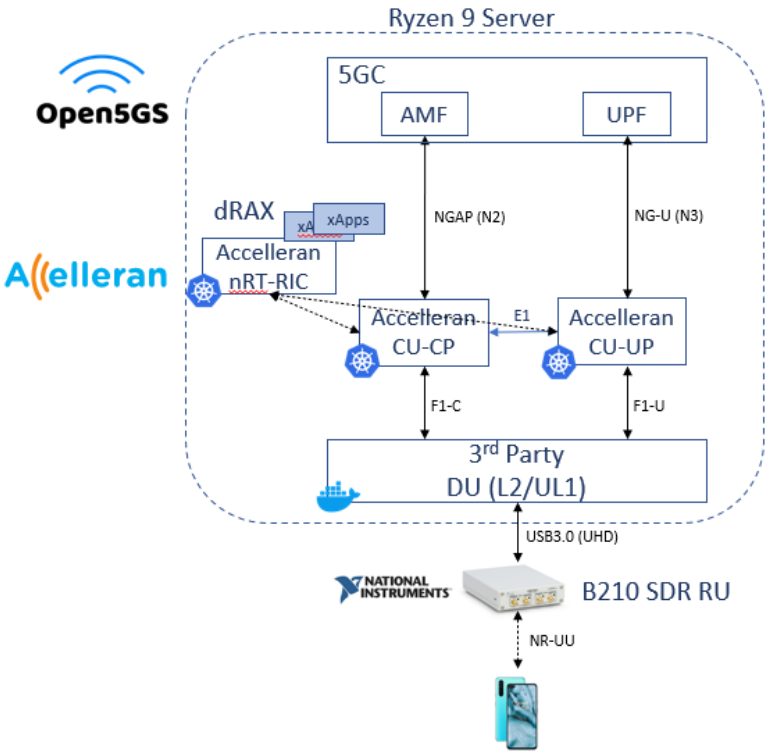


Figure 2-4: Phase 2 NR-UU integration and testing with Research RU and commercial UE

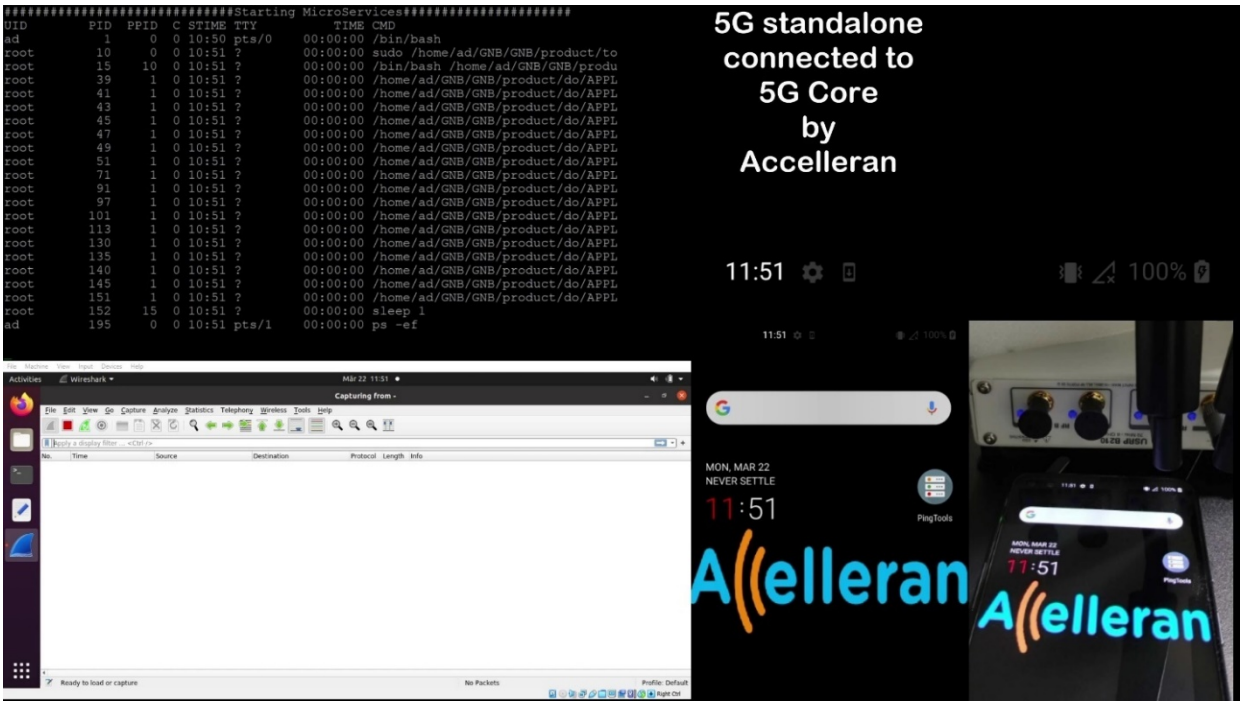


Figure 2-5: dRAX CU microservices and DU launch together with B210 RU and OnePlus Nord phone un-registered

During this phase, the focus was only on pure interoperability testing over the standardised interfaces (NG, F1, E1, NR-UU) and end-to-end connectivity via Ping tool without performing throughput measurements on the end-to-end system.

Figure 2-5 shows the launch of dRAX CU microservices and DU component together with an image of the B210 RU used and the OnePlus Nord phone in flight mode.

After the OnePlus phone is taken off flight mode the phone starts registering in the 5GC as shown in the Wireshark[2] trace window in **Figure 2-6** with initial F1-C and NGAP messages ongoing (control plane). Shortly after that a protocol data unit (PDU) session is established (user plane).

[2] https://www.wireshark.org/

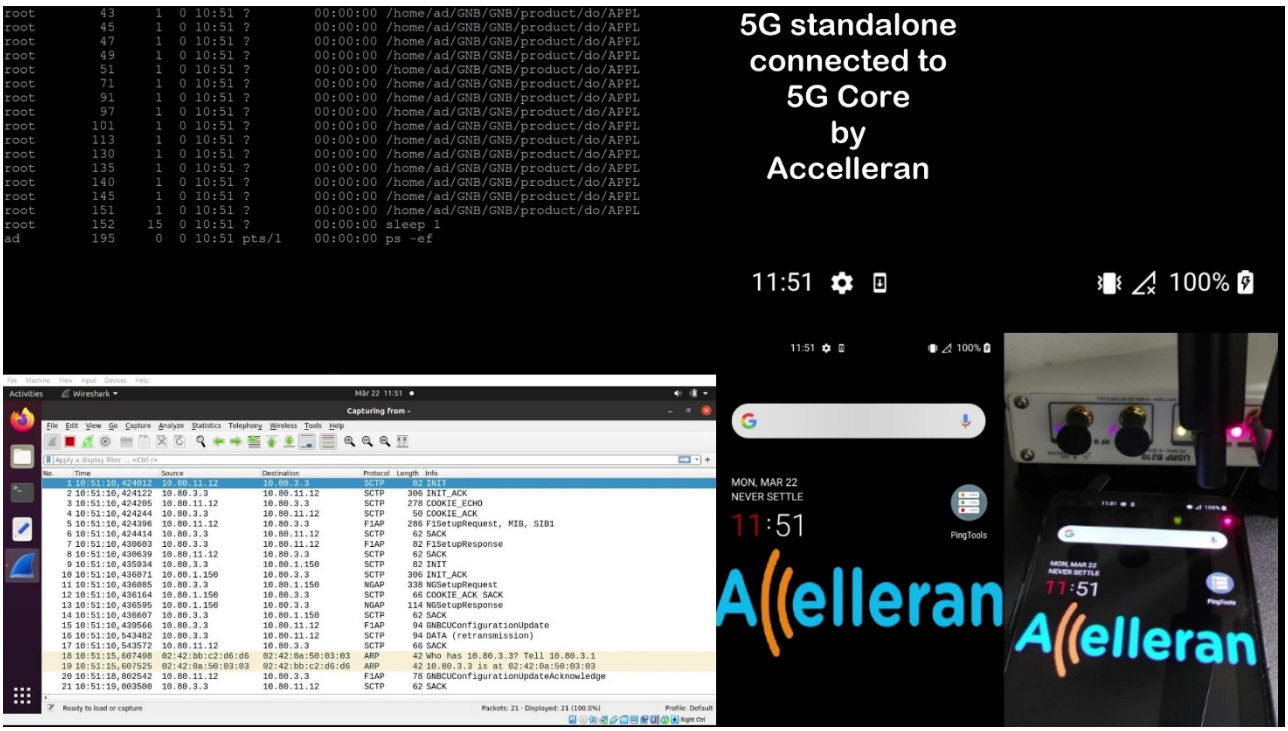


Figure 2-6: dRAX CU microservices and DU running with B210 RU and OnePlus Nord phone start registering

The PDU session establishes full user plane connectivity used by the Ping tool in the phone. Figure 2-7 shows the ping tool in the phone and the Wireshark traces showing the GPRS Tunnelling Protocol (GTP) user plane traffic corresponding to the ping application.

Figure 2-8 shows the details of a Wireshark call trace at CU level (mainly control plane on the F1-C and NGAP interface) up to the first GTP packet sent through for the user plane.

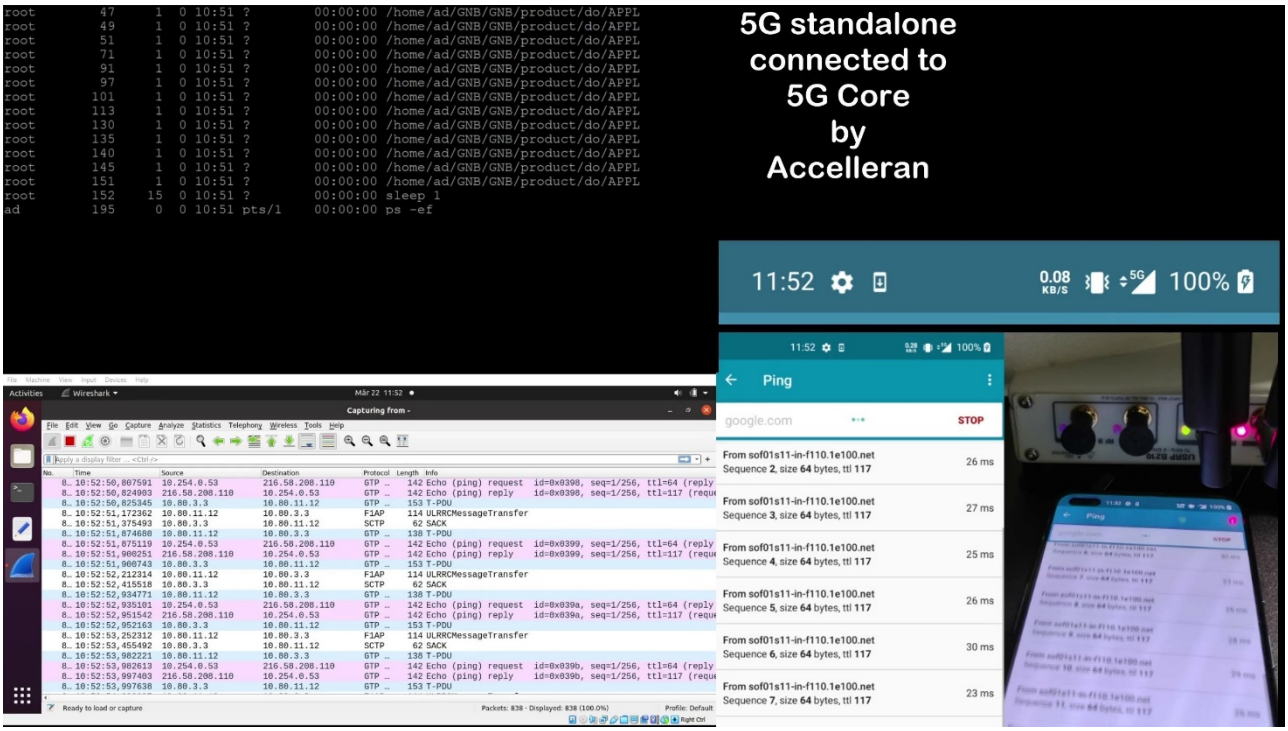


Figure 2-7: dRAX CU microservices and DU running with B210 RU and OnePlus Nord pinging data

No.	Time	Source	Destination	Protocol	Length	Info
5	0.000294	10.80.11.12	10.80.3.3	F1AP	286	F1SetupRequest, MIB, SIB1
7	0.006935	10.80.3.3	10.80.11.12	F1AP	82	F1SetupResponse
11	0.013931	10.80.3.3	10.80.1.150	NGAP	338	NGSetupRequest
13	0.014241	10.80.1.150	10.80.3.3	NGAP	114	NGSetupResponse
15	0.018912	10.80.3.3	10.80.11.12	F1AP	94	GNBCUConfigurationUpdate
32	8.353022	10.80.11.12	10.80.3.3	F1AP	78	GNBCUConfigurationUpdateAcknowledge
240	87.975353	10.80.11.12	10.80.3.3	F1AP/NR RRC	230	RRC Setup Request
241	87.984131	10.80.3.3	10.80.11.12	F1AP/NR RRC	226	SACK , RRC Setup
242	88.005330	10.80.11.12	10.80.3.3	F1AP/NR RRC/NAS-5GS/NAS…	218	SACK , RRC Setup Complete, Registration request, Registration request MAC=0x00000000
243	88.009457	10.80.3.3	10.80.1.150	NGAP/NAS-5GS/NAS-5GS	206	InitialUEMessage, Registration request, Registration request
244	88.009826	10.80.1.150	10.80.3.3	NGAP/NAS-5GS	106	DownlinkNASTransport, Identity request
245	88.012530	10.80.3.3	10.80.11.12	F1AP/NR RRC/NAS-5GS	122	SACK , DL Information Transfer, Identity request MAC=0x00000000
246	88.022837	10.80.11.12	10.80.3.3	F1AP/NR RRC/NAS-5GS	142	SACK , UL Information Transfer, Identity response MAC=0x00000000
247	88.024999	10.80.3.3	10.80.1.150	NGAP/NAS-5GS	146	UplinkNASTransport, Identity response
248	88.025787	10.80.1.150	10.80.3.3	NGAP/NAS-5GS	146	DownlinkNASTransport, Authentication request
249	88.027766	10.80.3.3	10.80.11.12	F1AP/NR RRC/NAS-5GS	162	SACK , DL Information Transfer, Authentication request MAC=0x00000000
250	88.097764	10.80.11.12	10.80.3.3	F1AP/NR RRC/NAS-5GS	146	SACK , UL Information Transfer, Authentication failure (Synch failure) MAC=0x00000000
251	88.103194	10.80.3.3	10.80.1.150	NGAP/NAS-5GS	150	UplinkNASTransport, Authentication failure (Synch failure)
252	88.103921	10.80.1.150	10.80.3.3	NGAP/NAS-5GS	146	DownlinkNASTransport, Authentication request
253	88.106090	10.80.3.3	10.80.11.12	F1AP/NR RRC/NAS-5GS	162	SACK , DL Information Transfer, Authentication request MAC=0x00000000
254	88.112833	10.80.11.12	10.80.3.3	F1AP/NR RRC/NAS-5GS	130	SACK , UL Information Transfer, Authentication failure (ngKSI already in use) MAC=0x00000000
255	88.114839	10.80.3.3	10.80.1.150	NGAP/NAS-5GS	134	UplinkNASTransport, Authentication failure (ngKSI already in use)
256	88.115064	10.80.1.150	10.80.3.3	NGAP/NAS-5GS	146	DownlinkNASTransport, Authentication request
257	88.117069	10.80.3.3	10.80.11.12	F1AP/NR RRC/NAS-5GS	162	SACK , DL Information Transfer, Authentication request MAC=0x00000000
260	88.245357	10.80.11.12	10.80.3.3	F1AP/NR RRC/NAS-5GS	130	UL Information Transfer, Authentication response MAC=0x00000000
261	88.248505	10.80.3.3	10.80.1.150	NGAP/NAS-5GS	150	UplinkNASTransport, Authentication response
262	88.249278	10.80.1.150	10.80.3.3	NGAP/NAS-5GS	122	DownlinkNASTransport, Security mode command
263	88.251462	10.80.3.3	10.80.11.12	F1AP/NR RRC/NAS-5GS	138	SACK , DL Information Transfer, Security mode command MAC=0x00000000
264	88.265313	10.80.11.12	10.80.3.3	F1AP/NR RRC/NAS-5GS	194	SACK , UL Information Transfer MAC=0x00000000
265	88.267529	10.80.3.3	10.80.1.150	NGAP/NAS-5GS	198	UplinkNASTransport
266	88.268030	10.80.1.150	10.80.3.3	NGAP/NAS-5GS	226	InitialContextSetupRequest
267	88.270876	10.80.3.3	10.80.11.12	F1AP	146	SACK , UEContextSetupRequest
268	88.271040	10.80.11.12	10.80.3.3	F1AP	230	SACK , UEContextSetupResponse
271	88.474551	10.80.11.12	10.80.3.3	F1AP/NR RRC	102	Security Mode Complete MAC=0x1b290a64
272	88.478932	10.80.3.3	10.80.11.12	F1AP/NR RRC	290	SACK MAC=0xa28329c3
273	88.502800	10.80.11.12	10.80.3.3	F1AP/NR RRC	118	SACK , Dedicated SIB Request MAC=0x4f7d995f
274	88.505546	10.80.3.3	10.80.11.12	F1AP	106	SACK , UEContextModificationRequest
275	88.505617	10.80.11.12	10.80.3.3	F1AP/NR RRC	130	SACK , RRC Reestablishment Complete MAC=0x16800a26
276	88.508263	10.80.3.3	10.80.1.150	NGAP/NAS-5GS	118	UplinkNASTransport
277	88.508675	10.80.1.150	10.80.3.3	NGAP/NAS-5GS	122	DownlinkNASTransport
278	88.511285	10.80.3.3	10.80.11.12	F1AP/NR RRC	138	SACK MAC=0x9884177b
279	88.511355	10.80.11.12	10.80.3.3	F1AP	206	SACK , UEContextModificationResponse
280	88.513351	10.80.3.3	10.80.1.150	NGAP	98	InitialContextSetupResponse
284	88.952838	10.80.11.12	10.80.3.3	F1AP/NR RRC	234	MAC=0x00f42a28
285	88.956556	10.80.3.3	10.80.1.150	NGAP/NAS-5GS	238	UplinkNASTransport
286	88.968895	10.80.1.150	10.80.3.3	NGAP/NAS-5GS	262	PDUSessionResourceSetupRequest
287	88.981147	10.80.3.3	10.80.11.12	F1AP	190	SACK , UEContextModificationRequest
288	88.981357	10.80.11.12	10.80.3.3	GTP	62	T-PDU
289	88.981420	10.80.11.12	10.80.3.3	F1AP	238	SACK , UEContextModificationResponse
290	88.990836	10.80.3.3	10.80.11.12	F1AP/NR RRC	334	SACK MAC=0x43fd0338
291	89.007784	10.80.11.12	10.80.3.3	F1AP/NR RRC	118	SACK , IAB Other Information[Malformed Packet]

Figure 2-8: Wireshark CU call trace details up to the first GTP packet sent

2.1.1.3 Phase 3: NR-UU testing with Commercial third-party RU and commercial UE

This phase (currently ongoing and to be part of a future 5G-CLARITY deliverable) consists on the replacement of the Research NI B210 RU with a commercial RU integrated end-to-end to be used in the 5G-CLARITY pilots as described in 5G-CLARITY D5.1 [3]. This phase will include results from the performance testing done end-to-end with this commercial RU.

2.1.2 Wi-Fi and LiFi integration validation

In Section 6.2.1 of 5G-CLARITY D3.1 [1] we introduced an SDN enabled L2 network integrating Wi-Fi and LiFi, which would allow UEs with Wi-Fi and/or LiFi interfaces to freely roam across Wi-Fi and LiFi AP while maintaining their IP address. Figure 2-9, from 5G-CLARITY D3.1 [1], depicts our envisioned integrate Wi-Fi/LiFi network.

Section 6.2 in 5G-CLARITY D3.1 [1] also included a detailed proposal explaining how such architecture could be realised, which is depicted in Figure 2-10.

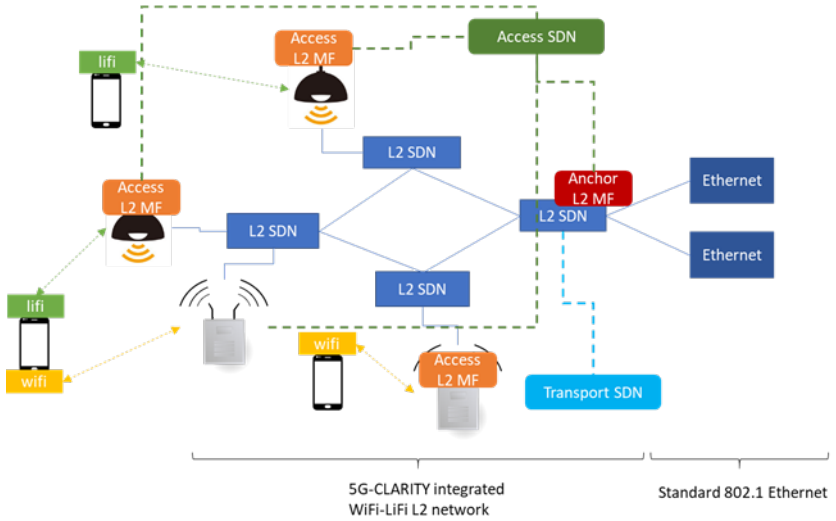


Figure 2-9. Envisioned integrated L2 Wi-Fi/LiFi network (5G-CLARITY D3.1 [1])

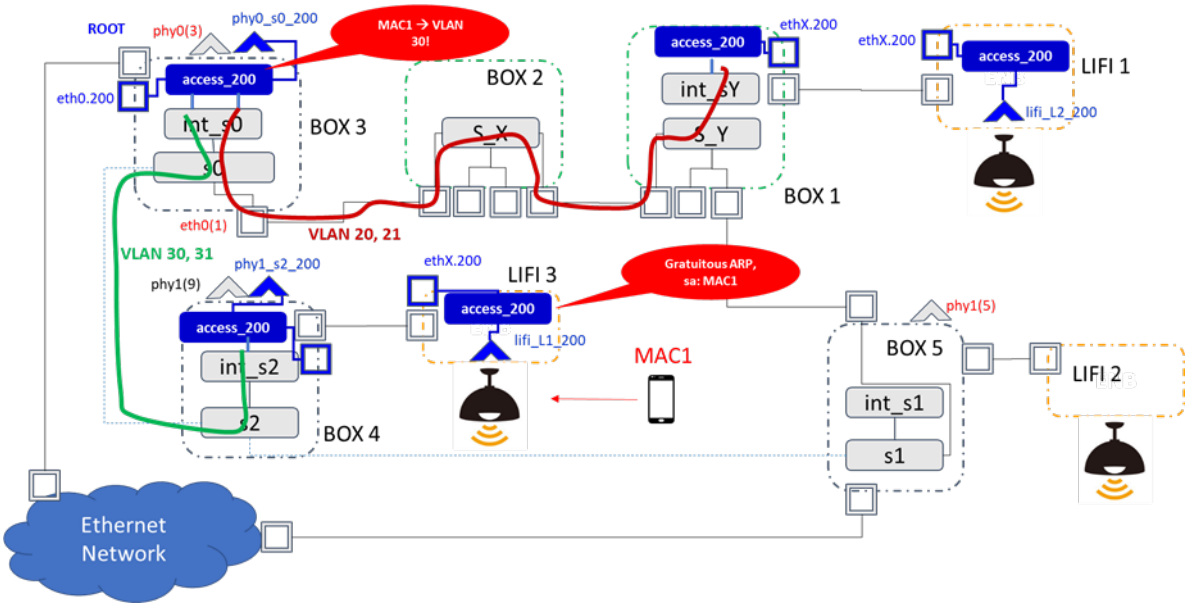


Figure 2-10. Mobility support in Wi-Fi-LiFi SDN L2 network (5G-CLARITY D3.1 [1])

Principles of the design outlined in Figure **2-10** are the following:

- LiFi APs connect using Ethernet to an access BOX, which integrated both SDN enabled Ethernet and Wi-Fi access interfaces. These integrated boxes are referred to as BOX1 to BOX5 in Figure 2-10.
- The SDN interfaces included in BOX1-BOX5 are based on software bridges implemented using Openvswitch[3], which supports OpenFlow [4] and OVSDB [5] SDN protocols. OVSDB allows to programmatically create new software bridges in these boxes, while OpenFlow controls how flows are forwarded within each bridge.
- The data plane element used to forward flows in our architecture is a virtual local area network (VLAN) tag injected by the Wi-Fi and LiFi APs. This VLAN tag is used in the integration bridge (i.e., int_s2 in BOX4 in Figure 2-10) to proactively inject forwarding paths that allow to reach all Wi-Fi and

[3] https://www.openvswitch.org/

LiFi APs in a given deployment. This VLAN based forwarding enables to support multi-tenancy where one could have multiple slices, representing each slice with a service set identifier (SSID) in the Wi-Fi and LiFi APs, which would then be linked to a given VLAN tag across the transport infrastructure.

- While VLAN tags allow to reach a given Wi-Fi or LiFi AP, individual user bindings, i.e., the point of attachment of a user represented by the IEEE medium access control (MAC) address of its Wi-Fi or LiFi interface is maintained using data plane learning, as in standard Ethernet switching. This is the goal of the "access bridges" represented in blue in Figure 2-10, where we would have one such access bridge for each slice in our infrastructure.
- Executing a handover requires then to update the forwarding path for a given user, i.e., the green or red paths in Figure 2-10. For this to happen an uplink packet needs to be sent by the user once it is attached to the new AP. This uplink packet will update the MAC-interface bindings in all traversed access bridges. Thus, our architecture supports only a break-before-make type of handover.

Next, we describe an evaluation in a laboratory testbed that allows to understand how this architecture works and to measure the handover interruption times in the Wi-Fi and LiFi interfaces.

2.1.2.1 Evaluating the 5G-CLARITY integrated Wi-Fi-LiFi L2 network

Figure 2-11 depicts a testbed that has been setup between I2CAT and PLF to validate the integrated Wi-Fi and LiFi SDN enabled L2 network. The testbed consists of:

- Two LiFi APs deployed in the PLF facilities in the UK. Since we require L2 connectivity between access BOXes and the LiFi APs, we use an L2 virtual private network (VPN) over the Internet to connect these APs with the I2CAT lab in Barcelona.
- In the I2CAT lab we have first two access BOXes, WIFI-1 and WIFI-2, acting as Wi-Fi AP and connecting via the L2 internet tunnel with the corresponding LiFi APs in the UK. These two access BOXes also have an Ethernet backhaul interface to connect to the access switch. These boxes implement the full SDN control plane based on Openvswitch described in 5G-CLARITY D3.1 [1].
- An access switch deployed at I2CAT connected to the two Wi-Fi boxes (WIFI-1 and WIFI-2) using two separate backhaul tunnels (identified over the wire with a per-tunnel VLAN). The access switch has an additional ethernet interface connecting to the core network, where the dynamic host configuration protocol (DHCP) and an *iperf* servers reside.

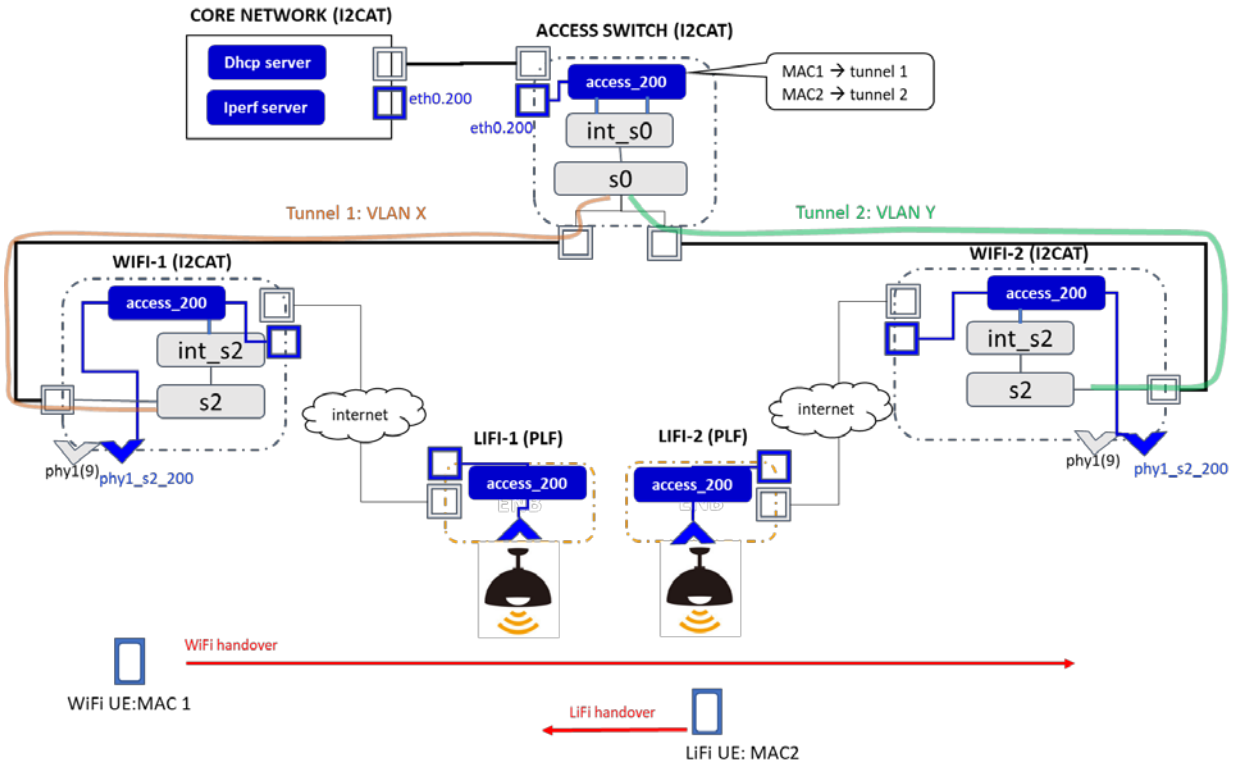


Figure 2-11. Testbed used for Wi-Fi-LiFi integrated L2 network

The deployed testbed enables us to perform the following tests:

- Test 1: Session continuity across Wi-Fi handover, where the Wi-Fi UE in Figure 2-11 executes a handover between WIFI-1 and WIFI-2.
- Test 2: Session continuity across LiFi handover, where the LiFi UE in Figure 2-11 executes a handover between LIFI-1 and LIFI-2.

We note that only horizontal handovers are considered in this section because vertical handovers are addressed within 5G-CLARITY through the multi-connectivity framework presented in Section 2.1.3.

2.1.2.1.1 Test 1: Measuring Wi-Fi handover performance

In this section the results of the Wi-Fi handover performance in the setup are presented as described in Figure 2-11. In this experiment nodes WIFI-1 and WIFI-2 are configured at channels 40 and 44 in the 5GHz band using a channel bandwidth of 20MHz. An Intel NUC is used as Wi-Fi UE. Given that the experiment is performed in a lab environment where the signal drop is not enough to trigger a handover naturally, the Wi-Fi handover is forced using the "*iw connect*" command in the Wi-Fi UE to force the reconnection from the AP WIFI-1 to the AP WIFI-2. Our goal in this experiment is to evaluate the disruption caused by the handover to an ongoing flow. Notice that in our architecture the UE device maintains the IP address while roaming through the combined Wi-Fi-LiFi network, so the only factors affecting the handover are: i) the radio link re-establishment time, which is a characteristic of the Wi-Fi air interface that we cannot modify; and ii) the backhaul reconfiguration time, namely the time it takes for packets coming from the core network to move from Tunnel 1 to Tunnel 2 in Figure 2-11. To generate traffic in our experiment we use the *iperf* application in both user datagram protocol (UDP) and transmission control protocol (TCP) mode, where the *iperf* server is located in the core network in Figure 2-11 and the client in the WIFI UE.

The evaluation starts by looking at UDP performance, which is easier to understand as packets flow only in one direction. Figure 2-12 depicts the goodput in bits per seconds (bps) over time for a downlink (core network → WIFI UE) and uplink (WIFI UE → core network) flows. It can be observed that in both cases the handover disruption lasts only for a couple of seconds, which is caused due to the radio link re-establishment time. In Figure 2-13 and Figure 2-14, we can see detailed *wireshark* traces of the DL and UL UDP flows during the handover time. Notice that in this architecture, in order to re-establish the path of packets through the backhaul once the WIFI UE connects to the new AP, we need to update the MAC bindings in the access switch (c.f. Figure 2-11). In the UL shown in Figure 2-13, we observe that once reconnecting to the new AP, the WIFI UE sends a logical link control (LLC) packet (time 29.016823 seconds) followed by an address resolution protocol (ARP) Request/Response. These L2 packets generated by the WIFI UE reach the access switch through Tunnel 2 and automatically update the MAC bindings in this switch (c.f. Figure 2-11). The DL is however more complex because the WIFI UE does not generate UL packets that can update the MAC bindings in the access switch. In this case, we implement a mechanism in the WIFI AP such that when a new device is connected to the AP (in our experiment to WIFI-2) a spontaneous ARP request is generated in the AP on behalf of the WIFI UE (using the UE's MAC in the source field of the ARP request) and transmitted through Tunnel 2. As a result, as soon as WIFI UE connects to the AP in node WIFI-2, the MAC bindings in the access switch are updated and the DL UDP packets start flowing through Tunnel 2 instead of Tunnel 1, resulting in the overall handover time observed in Figure 2-12, which we experimentally verify to be dominated by the radio reconnection time. Figure 2-13 and Figure 2-14 provide detailed packet traces of UDP UL and DL handover, respectively.

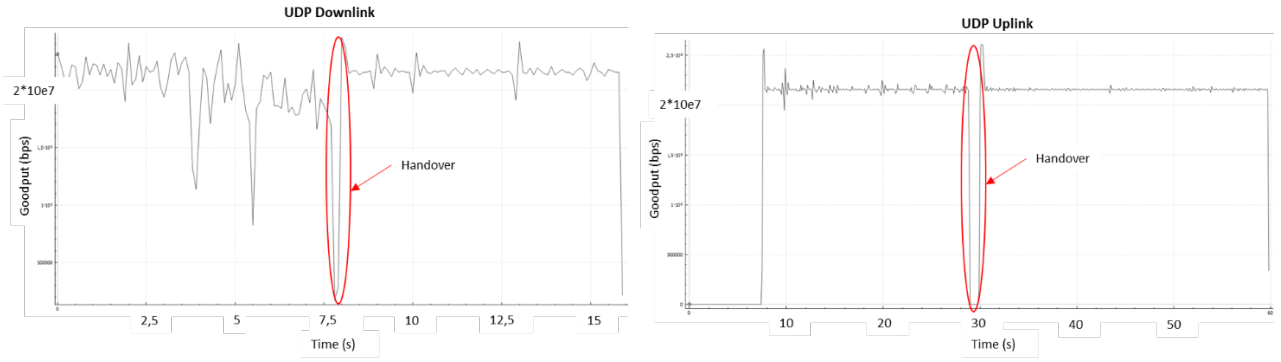


Figure 2-12. Wi-Fi handover impact on UDP flows, Downlink (left), and Uplink (right)

38196	28.959082	192.168.100.10	192.168.100.1	UDP	1512	52743 → 5001 Len=1470
38197	28.959869	192.168.100.10	192.168.100.1	UDP	1512	52743 → 5001 Len=1470
38198	29.016823	IntelCor_ff:82:82	Broadcast	XID	56	Basic Format; Type 1 LLC (Class I LLC); Window Size 0
38199	29.035106	fe80::223:15ff:feff:82…	ff02::16	ICMPv6	90	Multicast Listener Report Message v2
38200	29.686236	fe80::223:15ff:feff:82…	ff02::16	ICMPv6	90	Multicast Listener Report Message v2
38201	29.975694	IntelCor_ff:82:82	Broadcast	ARP	56	Who has 192.168.100.1? Tell 192.168.100.10
38202	29.975809	Gatework_f4:f5:4f	IntelCor_ff:82:82	ARP	42	192.168.100.1 is at 00:d0:12:f4:f5:4f
38203	29.980490	192.168.100.10	192.168.100.1	UDP	1512	52743 → 5001 Len=1470

Figure 2-13. Packet trace of the UDP UL handover. Default columns from wireshark

13044	7.806624	192.168.100.1	192.168.100.10	UDP	1512	43974 → 5001 Len=1470
13045	7.806624	192.168.100.1	192.168.100.10	UDP	1512	43974 → 5001 Len=1470
13046	7.875884	fe80::223:15ff:feff:82…	ff02::16	ICMPv6	90	Multicast Listener Report Message v2
13047	7.969438	192.168.100.1	192.168.100.10	UDP	1512	43974 → 5001 Len=1470
13048	7.971604	192.168.100.1	192.168.100.10	UDP	1512	43974 → 5001 Len=1470
13049	7.974115	192.168.100.1	192.168.100.10	UDP	1512	43974 → 5001 Len=1470

Figure 2-14. Packet trace of the UDP DL handover. Default columns from wireshark

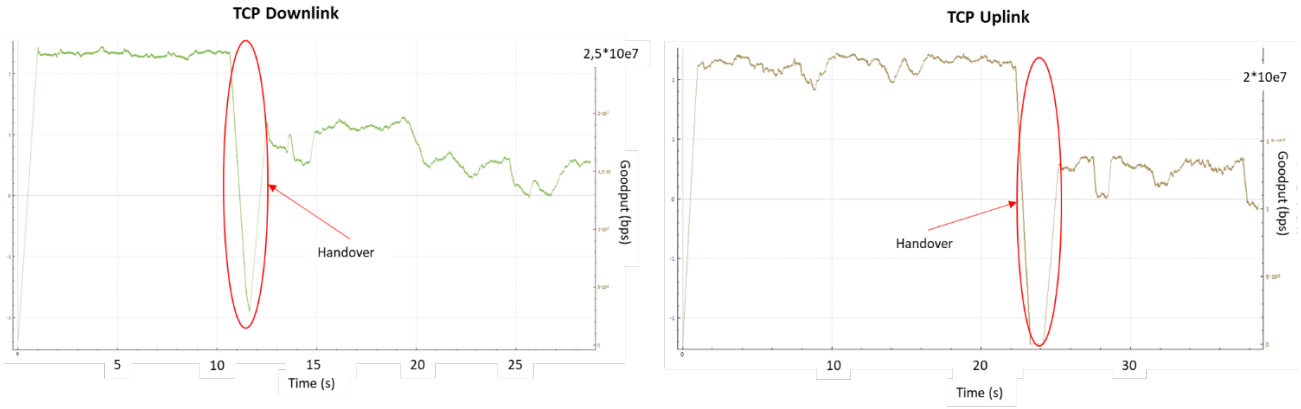


Figure 2-15. Wi-Fi handover impact on TCP flows, for DL (left), and UL (right)

To complete the analysis, we attach in Figure **2-15** the handover impact to DL and UL TCP flows. In the case of TCP, there is always a flow of ACK packets that makes the flow bidirectional. We can see in the figure how in both cases the handover interruption is again around 2 seconds, due to the quick update of the MAC bindings in the access switch. We note that in the TCP experiment, there is a performance difference of the TCP flow after moving to the WIFI-2 AP due to the radio conditions experienced in that AP during the time where the experiment was performed (e.g. due to external interferences from other Wi-Fi networks outside our control).

2.1.2.1.2 Test 2: Measuring LiFi handover performance

In Test 2 the LiFi handover performance in the setup is evaluated as depicted in Figure 2-11**.** Given that the experiment is performed in a laboratory environment where mobility is restricted, we emulate handover by first having the two LiFi APs turned on (c.f. Figure 2-11) and then turning off the LiFi-1 AP in order to force the pureLiFi dongle to connect to the LiFi-2 AP. Like in the Wi-Fi case there are two factors affecting handover delay, namely:

- The LiFi reconnection time, i.e. the time the dongle needs to establish a L2 connection with the target LiFi AP,
- The backhaul reconfiguration time, i.e. the time required for the bridge in the access switch (c.f. Figure 2-11) to update the MAC bindings. Given the network design the bindings are updated as soon as an uplink packet from the LiFi dongle is received from the target AP.

Figure 2-16 and Figure 2-17 depict respectively handover disruption caused to a UDP and TCP flows in DL and UL direction.

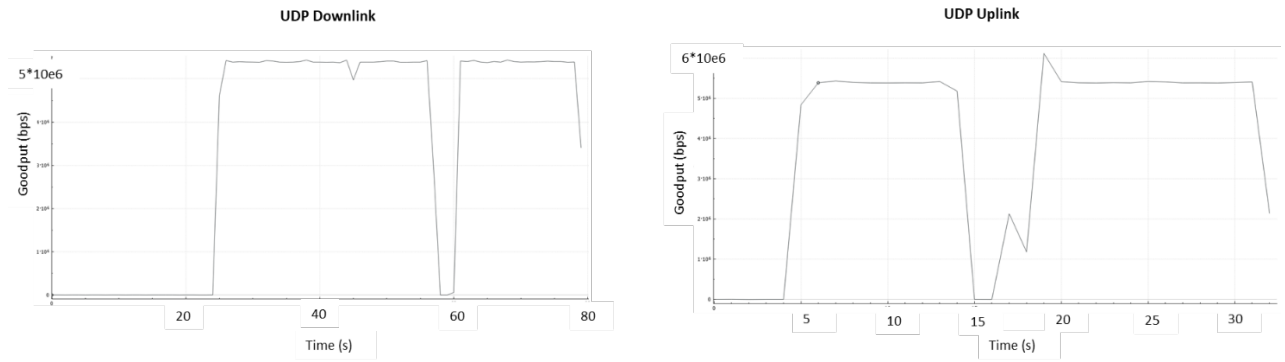


Figure 2-16. LiFi handover impact on UDP flows, DL (left) and UL (right)

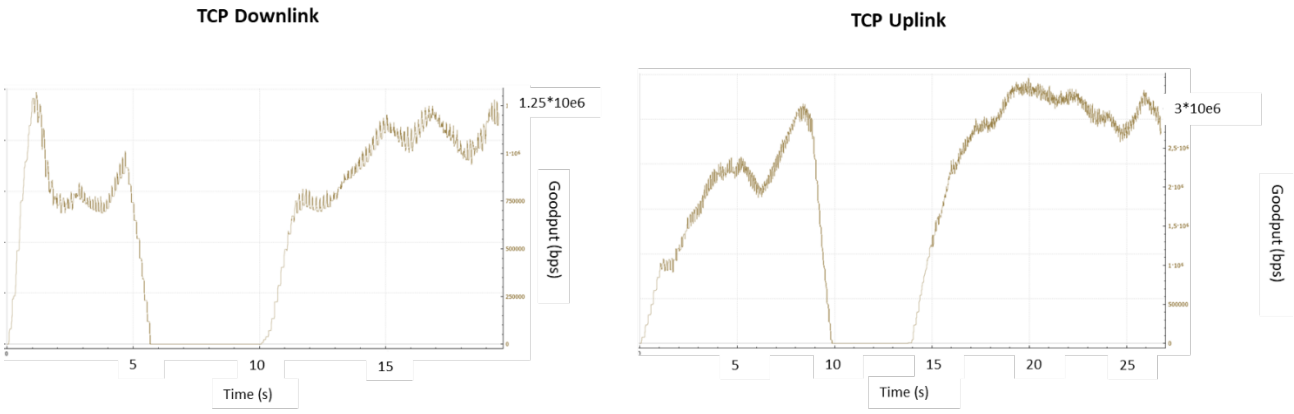


Figure 2-17. LiFi handover impact on TCP flows, for DL (left) and UL (right)

In the UDP case (Figure 2-16) we can see a disruption of around 3 seconds in both DL and DL, whereas the disruption grows to almost 5 seconds in the case of TCP (Figure 2-17). A longer disruption for TCP is expected due to the TCP slow start and retransmission mechanisms. Overall, we conclude that the bulk of the interruption time is due to the LiFi reconnection time inherent to the LiFi technology, whereas the backhaul reconfiguration time is kept minimum due to our network design.

2.1.3 5G-CLARITY UP architecture validation through a virtual testbed

In 5G-CLARITY D3.1, Section 6.2.3 [1], an initial design for the 5G-CLARITY UP architecture was proposed that integrates 5GNR, Wi-Fi and LiFi access networks with a common core network, as well as supporting the AT3S framework to aggregate the different access networks into a common data flow. The original high-level design is depicted in Figure 2-18. This design was further evolved and a virtual testbed, including all the required elements, is built which is used to provide a functional validation of this architecture.

The 5G-CLARITY virtual testbed is described in Figure **2-19** and is composed of three virtual machines (VMs) including the following elements:

- VM1: mptcpUe. This virtual machine represents a UE, in particular a 5G-CLARITY CPE that contains a Wi-Fi and LiFi interface. Since a gNB emulator (which can connect through the user plane function - UPF) is not available, the 5GNR interface is emulated using an Ethernet card directly connected to the data network (i.e., to the mptcpProxy VM). This VM deploys the multi-path TCP (MPTCP) kernel[4] 5.5 to enable the use of simultaneous paths. This VM includes three namespaces:
    - MPTCPns: the namespace providing an abstraction of the different access technologies available in the 5G-CLARITY CPE through three virtual IP interfaces, implemented using Linux *veth* pairs. This namespace deploys a layer 2 tunnel implemented using *openvpn*[5] that connects to a proxy counterpart on the network side (VM3).
    - WIFIns: the namespace representing the Wi-Fi interface of the 5G-CLARITY CPE. This namespace sets up the Internet Protocol Security (IPSEC) tunnel against the N3IWF function in the core network. Notice that IP forwarding is enabled in this namespace to route traffic coming from the Uens_mptcp namespace.
    - LIFIns: Is a namespace representing a LiFi interface. This namespace sets up the IPSEC tunnel against the N3IWF function in the core network. Notice that IP forwarding is enabled in this

[4] https://www.multipath-tcp.org/

[5] https://openvpn.net/

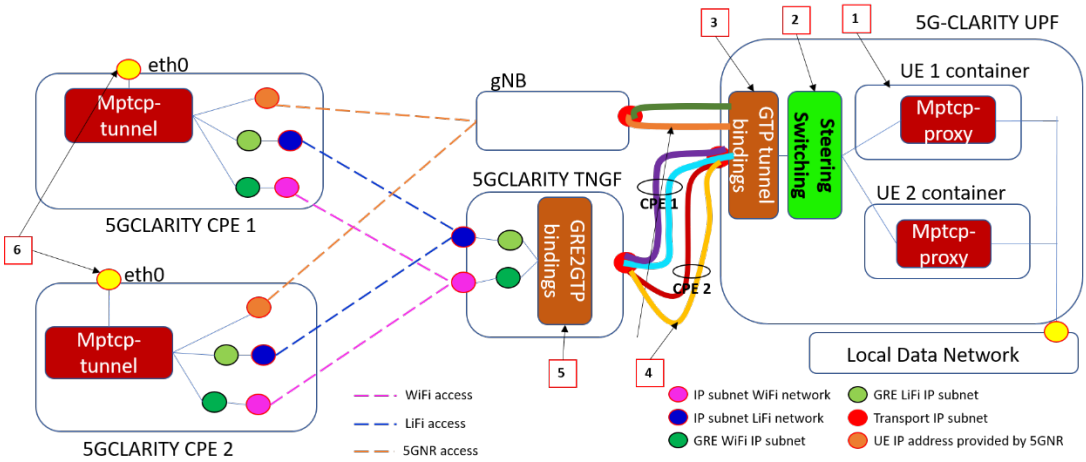


Figure 2-18. Initial UP architecture (5G-CLARITY D3.1 [1]).

namespace to route traffic coming from the Uens_mptcp namespace.

- 5GNR: As commented, we assume an Ethernet card as the 5GNR interface. For simplicity, this connection is not included as a namespace.

- VM2: free5gc. This virtual machine implements a full standalone (SA) 5G Core based on the *free5gc*[6] open-source implementation, which requires kernel 5.0.0-23. *Free5gc* provides support for interworking with non 3GPP technologies through the N3IWF function, which is however limited to a single non-3GPP UE. This was extended as part of our work to support connection of multiple UEs. The *free5gc* VM also contains a separate namespace running the UPF, whereas the rest of control plane functions (e.g., access and mobility management function - AMF, session management function - SMF) run in the host namespace. It is worth noting that the default SMF configuration provided by *free5gc* had to be adapted to allow the UPF to route traffic from the 10.0.1/24 subnet configured in the Uens_mptcp namespace depicted in Figure 2-19.

- VM3: mptcpProxy. This VM is used to terminate the *openvpn* over MPTCP connections coming from the 5G-CLARITY CPE (VM1) and thus also deployed the same MPTCP kernel as VM1.

Finally, we note that the virtual testbed described in Figure **2-19** models the different access technologies, i.e., 5GNR, Wi-Fi and LiFi, in a simple way using a virtual L2 network.

The described testbed addresses the main goals of the 5G-CLARITY UP architecture as follows:

- It supports a 5G SA deployment with a UPF deployed on the Edge compute cluster, represented in the testbed by VM2.
- It supports integration of Wi-Fi and LiFi devices through the N3IWF interworking functions.
- It supports smart traffic management in devices that have multiple wireless interfaces using MPTCP, which as we shall see next allows to define policies to control how traffic is balanced through the different wireless access networks. MPTCP is our implementation of the AT3S user plane function, which is also an option supported in 3GPP [6].
- It allows to bridge in a transparent manner L2 traffic entering the 5G-CLARITY CPE (VM1) with the application function sitting behind the core network (VM3) by means of a L2 tunnelling solution. This

[6] https://www.free5gc.org/

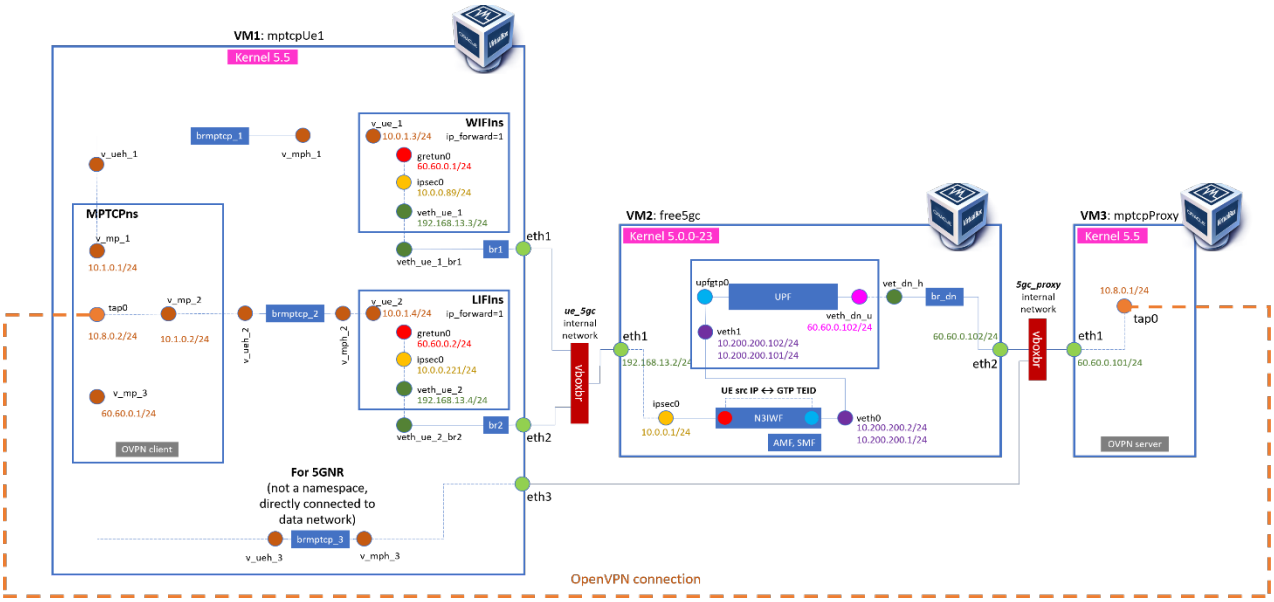


Figure 2-19. 5G-CLARITY virtual testbed architecture.

is a key feature to support 5G-CLARITY UC2.1 described in 5G-CLARITY D5.1 [3], where the 5G-CLARITY CPE is used to replace the Ethernet network connecting the BOSCH production lines to the MES servers.

Next, the proposed architecture is functionally validated, looking at the following aspects:

- The attachment procedure between the non-3GPP interfaces and the N3IWF function in the free5gc.
- The traffic flowing through the different interfaces.
- The connection setup process when we open the L2 tunnel between the 5G-CLARITY CPE and the MPTCP proxy function behind the core network.
- The functionality of the default MPTCP schedulers that allow to control how traffic flows through each wireless access interface. Notice that this evaluation does not include the optimized MPTCP scheduler developed in 5G-CLARITY that is described in Section 2.2.3.

2.1.3.1 Validation of non-3GPP connection process

The validation of the virtual testbed presented in the previous section is validated by looking at the connection setup between the non-3GPP interfaces, i.e. Uens_wifi and Uens_lifi in our testbed, and the N3IWF function in VM2 (free5gc). The 3GPP TS 23.501 [7] requires an IPSEC connection to be established between the non-3GPP UE and the N3IWF function, which we can see depicted in Figure 2-20.

Figure 2-20 depicts only the IPSEC setup between the Wi-Fi interface represented by Uens_wifi in Figure 2-19 and the N3IWF function; thus, we see communications between the transport IP addresses 192.168.13.3 and 192.168.13.2, which would correspond in a real deployment to the IP addresses of the Wi-Fi AP and Wi-Fi UE.

A first exchange of Internet Security Association and Key Management Protocol (ISAKMP) [8] are observed to setup the basic security association between the Wi-Fi terminal and the N3IWF function. After this step, starting at frame 47 in Figure 2-20, all signalling data is encrypted using the IPSEC Encapsulated Security Payload (ESP) protocol. The encrypted signalling includes the Extensible Authentication Protocol (EAP) messages transmitted by the Wi-Fi UE to authenticate against the AMF.

35	9.318272	192.168.13.3	192.168.13.2	ISAKMP	642	IKE_SA_INIT MID=00 Initiator Request
36	9.354858	192.168.13.2	192.168.13.3	ISAKMP	642	IKE_SA_INIT MID=00 Responder Response
37	9.366908	192.168.13.3	192.168.13.2	ISAKMP	214	IKE_AUTH MID=01 Initiator Request
38	9.374035	192.168.13.2	192.168.13.3	ISAKMP	1446	IKE_AUTH MID=01 Responder Response
39	9.379093	192.168.13.3	192.168.13.2	ISAKMP	198	IKE_AUTH MID=02 Initiator Request
40	9.905146	192.168.13.2	192.168.13.3	ISAKMP	[illegible]	IKE_AUTH MID=02 Responder Response
41	9.906861	192.168.13.3	192.168.13.2	ISAKMP	Initial security association for control plane	IKE_AUTH MID=03 Initiator Request
42	9.929470	192.168.13.2	192.168.13.3	ISAKMP		IKE_AUTH MID=03 Responder Response
43	9.931807	192.168.13.3	192.168.13.2	ISAKMP	182	IKE_AUTH MID=04 Initiator Request
44	11.487107	192.168.13.2	192.168.13.3	ISAKMP	118	IKE_AUTH MID=04 Responder Response
45	11.491610	192.168.13.3	192.168.13.2	ISAKMP	134	IKE_AUTH MID=05 Initiator Request
46	11.521732	192.168.13.2	192.168.13.3	ISAKMP	294	IKE_AUTH MID=05 Responder Response
47	11.575804	192.168.13.3	192.168.13.2	ESP	150	ESP (SPI=0x01000000)
48	12.588460	192.168.13.3	192.168.13.2	ESP	150	ESP (SPI=0x01000000)
49	12.588519	192.168.13.2	192.168.13.3	ESP	134	ESP (SPI=0x01000000)
50	12.588739	192.168.13.3	192.168.13.2	ESP	134	ESP (SPI=0x01000000)
51	12.681973	192.168.13.2	192.168.13.3	ESP	182	ESP (SPI=0x01000000)
52	12.682424	192.168.13.3	192.168.13.2	ESP	134	ESP (SPI=0x01000000)
53	12.682441	192.168.13.3	192.168.13.2	ESP	134	ESP (SPI=0x01000000)
54	12.682526	192.168.13.2	192.168.13.3	ESP	134	ESP (SPI=0x01000000)
55	13.183308	192.168.13.3	192.168.13.2	ESP	166	ESP (SPI=0x01000000)
56	13.183379	192.168.13.2	192.168.13.3	ESP	134	ESP (SPI=0x01000000)
58	14.159536	192.168.13.2	192.168.13.3	ISAKMP	Child security association for default PDU session	CREATE_CHILD_SA MID=05 Initiator Request
59	14.161652	192.168.13.3	192.168.13.2	ISAKMP		CREATE_CHILD_SA MID=05 Responder Response
61	14.193938	192.168.13.3	192.168.13.2	ESP	[illegible]	ESP (SPI=0x9a26e71f)
64	17.195345	192.168.13.3	192.168.13.2	ESP	138	ESP (SPI=0x9a26e71f)
65	18.196452	192.168.13.3	192.168.13.2	ESP	138	ESP (SPI=0x9a26e71f)

Figure 2-20. IPSEC tunnel setup between Wi-Fi/LiFi interface and N3IWF function.

To be able to transmit data traffic in addition to signalling, a child security association is established again using ISAKMP, after which user plane traffic is encapsulated in a different ESP tunnel. In case the Wi-Fi UE would want to transmit multiple simultaneous PDU sessions, a separate child security association would be setup for each PDU session.

2.1.3.2 N3IWF functional validation

The user plane functionality of the N3IWF function in VM2 (free5gc) is being validated next, which needs to bind the IP traffic from different Wi-Fi or LiFi UEs to GTP tunnels towards the UPF function residing in the UPFns namespace in Figure 2-19**.**

Figure 2-21 depicts a capture of user plane traffic before (interface ipsec0 in Figure 2-19) and after (interface veth0 in Figure 2-19) the N3IWF function. The depicted traffic comes from the Wi-Fi UE (above) and LiFi UE (below). Looking at the traffic before the N3IWF function, which is the traffic embedded in the ESP tunnel described in the previous section, we can see that traffic is IP in IP encapsulated using GRE, where the outer IP header are the endpoints of the IPSEC tunnel, and the inner header corresponds to the UE IP address space (60.60.0/24).

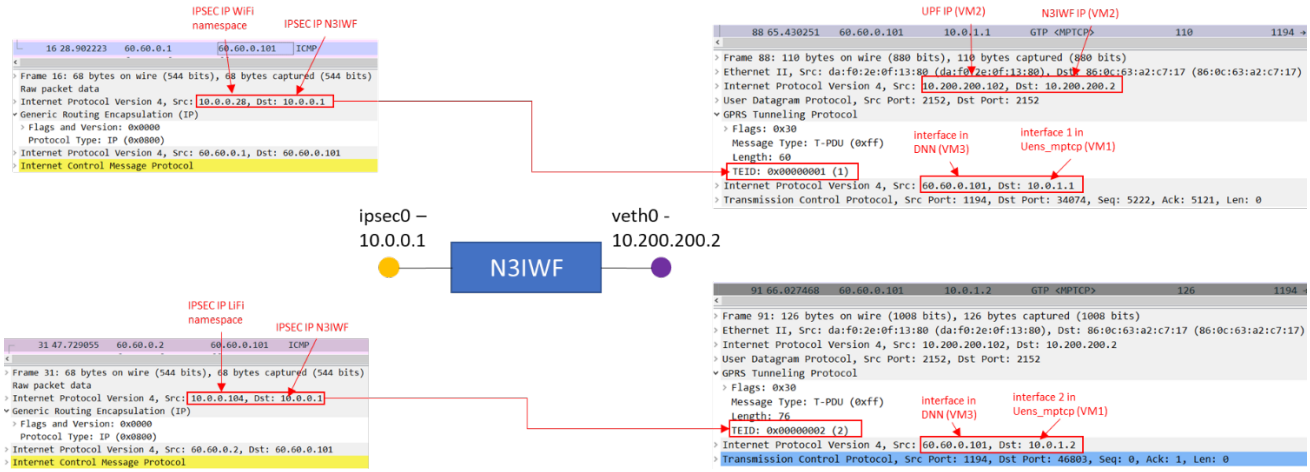


Figure 2-21. N3IWF user plane validation for Wi-Fi UE (above), and LiFi UE (below).

Looking at the traffic after the N3IWF function, we see as expected that it is GTP encapsulated. Notice that in the 3GPP Release 16 architecture, the N3IWF function acts as a gNB towards the UPF and therefore sets up a separate GTP tunnel for each associated non-3GPP UE. We can indeed see how the Wi-Fi UE traffic is mapped to a GTP tunnel with Tunnel End Point ID (TEID) equal 1 and the LiFi UE is mapped to a GTP tunnel with TEID equal 2.

2.1.3.3 L2 tunnel over MPTCP connection setup

In the proposed architecture, the 5G-CLARITY CPE sets up a L2 VPN towards the MPTCP proxy sitting behind the core. The reason for this approach is that our aim is to be able to transport transparently L2 traffic over the multiple WATs. However, our choice of AT3S user plane function, i.e., MPTCP, resides at layer 4, and therefore we need a way to transport L2 traffic over the layer 4 MPTCP connections. Establishing a L2 over MPTCP tunnel is the solution to this problem. At the time of this deliverable, we have only experimented with *openvpn*, although other more performing approaches could be considered in the future.

It is started by disabling encryption within *openvpn*, as traffic is already encrypted over the non-3GPP network using the IPSEC connection, as described above. The interest is therefore validating how *openvpn* is able to use the multiple WATs.

An unmodified version of *openvpn* is used selecting TCP as transport. However, we are using the MPTCP kernel 4.19, which overwrites the default TCP implementation in the kernel to force all TCP connections to try to open an MPTCP connection by default, reverting back to TCP if the server does not support MPTCP. In our case though, we make sure that the *openvpn* server supports MPTCP because it is collocated with the MPTCP proxy in VM3 (c.f. Figure 2-19), which also has the MPTCP kernel 4.19.

MPTCP can be configured by a component known as MPTCP *path_manager* [9] to decide how many IP paths should be attempted to be used in each connection. In our implementation, we select the *fullmesh path_manager* available in the kernel, which attempts to use all the IP interfaces available, which in our testbed are two, i.e. interface v_mp_1 and v_mp_2 in the MPTCPns in VM1: mptcpUe in Figure 2-19.

Figure 2-22 depicts what happens when the 5G-CLARITY CPE establishes the *openvpn* tunnel. We can see that two TCP connections are established with the *openvpn* server. The first connection, in the left, comes from IP address 10.0.1.1 (corresponding to interface v_mp_1 in Figure 2-19), and includes an option in the SYN packet declaring this connection as MTPCP capable. Subsequently, we see a second connection, right side, coming from IP address 10.0.1.2 (corresponding to interface v_mp_2 in Figure 2-19), with an option in the SYN packet declaring a “MPTCP Join Connection” intention.

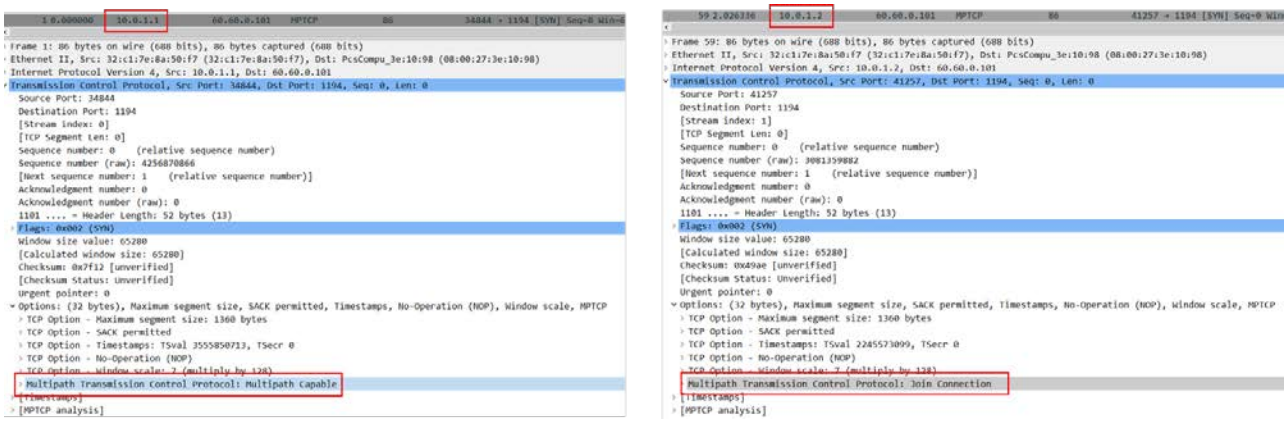


Figure 2-22. Establishment of MPTCP connection from v_mp_1 (left) and v_mp_2 (right)

After this exchange, the two individual TCP connections are bundled within a single MPTCP connection between the client and the server that allows to balance traffic across the different connections.

The key MPTCP component that controls how traffic is balanced across the connections is the MPTCP scheduler that will be studied in the next section.

2.1.3.4 Function validation of default MPTCP schedulers over virtual testbed

The standard MPTCP kernel supports three types of schedulers that we will study in this section:

- *Default*: This scheduler transmit data through the IP path of lower round-trip-time (RTT), where RTT is obtained through an estimation maintained by the kernel module. Thus, using this scheduler, a path will be prioritized until buffers start to fill up, RTT grows, and then packets start to be transmitted over other paths. This scheduler aims to select the fastest path to send next packet.
- *Redundant*: Using this scheduler each packet will be transmitted over all the active IP paths, disregarding the latency of each path. This scheduler aims to maximizing reliability.
- *Round-robin (RR)*: Using this scheduler, the number of packets *n* can be configured, such that *n* consecutive packets are transmitted across each of the active IP paths.

A variation of these schedulers that has been developed to fit the 5G-CLARITY goals is described in Section 2.2.3.

2.1.3.4.1 Default MPTCP scheduler

The traffic flows between the *openvpn* client and server are firstly validated in the virtual testbed using the default MPTCP scheduler.

Figure 2-23 depicts the results, first showing above the capture of a set of Internet control message protocol (ICMP) request/responses generated using the PING tool, and below a subset of the corresponding *wireshark* capture. Looking at the *wireshark* capture, we can see the following behaviour:

1. 5G-CLARITY CPE sends the ICMP request through the 10.0.1.2 interface, corresponding to the LiFi namespace in our virtual testbed.
2. An MPTCP acknowledgement is generated towards the same interface.
3. The MPTCP proxy sends the ICMP response towards the 10.0.1.1 interface, corresponding to the Wi-Fi namespace.
4. An MPTCP acknowledgment is generated towards the proxy.

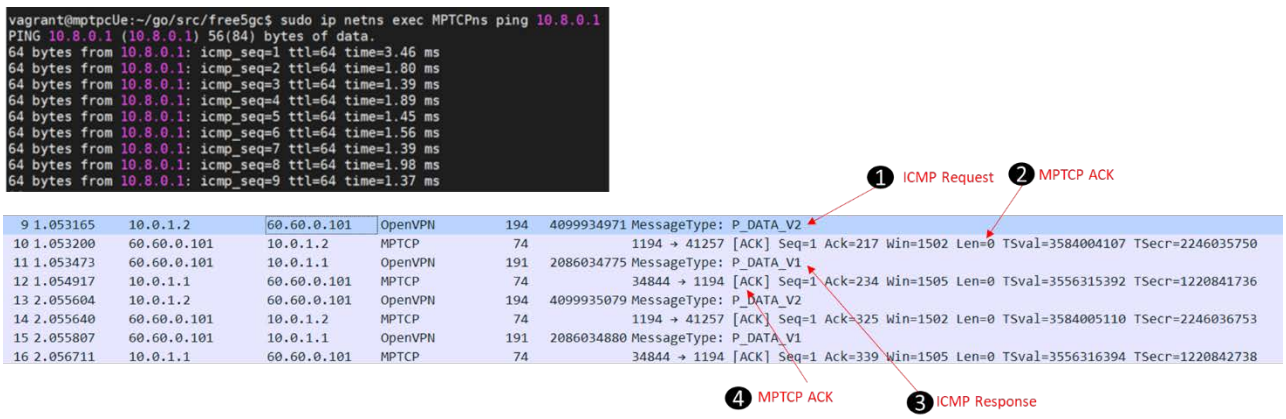


Figure 2-23. OpenVPN tunnel over MPTCP connection validation

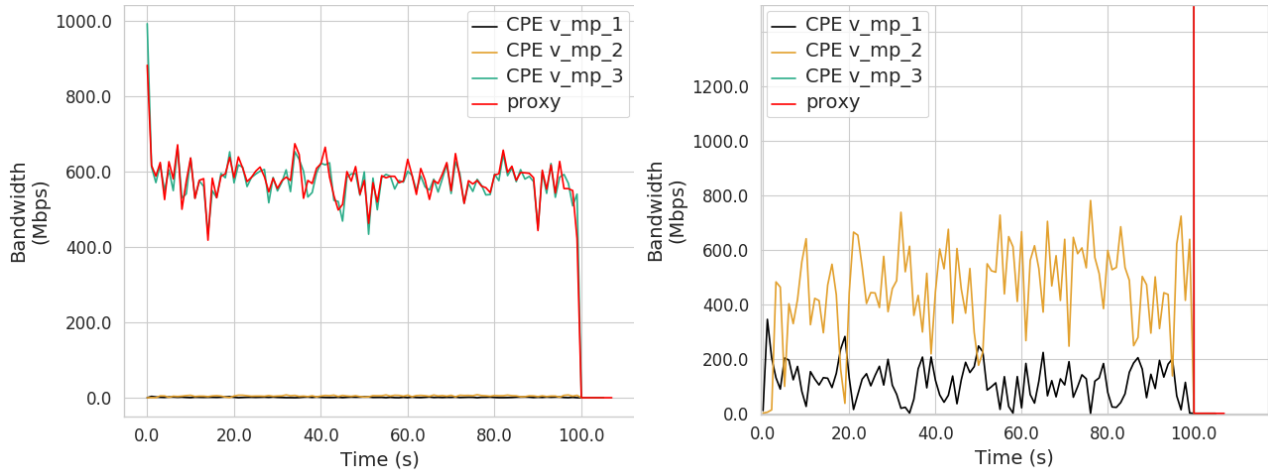


Figure 2-24: MPTCP default scheduler using *free5gc*. Outgoing and incoming traffic for each interface of CPE and proxy, respectively (left). Zoomed plot to show performance on paths 1 and 2 (right)

It is observed that how the choice of paths is performed independently by each MPTCP sender (client and server) and corresponds to the RTT statistics obtained by the kernel. We assume that certain hysteresis is used in this estimation since we observe that client and server stick to the paths they choose when the network is not congested.

This scheduler is also validated by testing the data rates achieved by the different paths. Assuming the testbed shown in Figure 2-19 (which connects the UE to the proxy through *free5gc*), the left part of Figure 2-24 shows the data rates for each path (v_mp_X, with X=1,2,3). The right part of the figure shows a zoomed portion (up to 15 Mbps) in order to highlight the difference (near two orders of magnitude) between the data rates of the first two paths of the CPE (v_mp_1 and v_mp_2) and the third path (v_mp_3).

In this figure, v_mp_1 and v_mp_2 are the paths through the Wi-Fi and LiFi interfaces, respectively, i.e., whose data are sent through the N3IWF entity in the *free5gc* VM. v_mp_3 is the path that goes through the 5GNR interface. However, since we do not have a gNB emulator to connect through the UPF entity, this path goes directly to the data network (i.e., it is directly connected to the mptcpProxy VM). As shown in the figure, the *iperf* test sends information through all the three paths. Due to the low performance of the *free5gc* executed on a VM, the throughput of the third path is much higher (near 600 Mbps whereas it is lower than 8 Mbps for the first and second paths, see the right figure). Data rates fluctuate due to the instantaneous RTT in each interface, which may vary depending on the congestion control mechanism.

In the case of this scheduler, the different paths can be in three different states: ON, OFF and BACKUP. These states can be very useful for different purposes. For example, all paths on the CPE may be employed to achieve the maximum possible data rate. Or some of them may be left in BACKUP state so they are only used when one path is inactive, e.g., due to bad coverage, in order to increase the reliability of the connection. Figure 2-25 shows that we can change in real-time the states of the different paths. The figure is also zoomed to see the behaviour of the first and second paths.

As shown, when path 3 is changed to BACKUP state (around t = 25 s), it stops sending data until the other two paths are changed to OFF (around t = 45 s). Then, the third path stops again as soon as one or two of the other paths are back to ON state (around t = 55).

Due to the low performance of *free5gc* executed on a VM, which affects these results, we also evaluate the default scheduler using one “client” VM (called mptcpUe1) and a “server” VM (called mptcpUe2), both supporting MPTCP. This testbed is depicted on Figure 2-26.

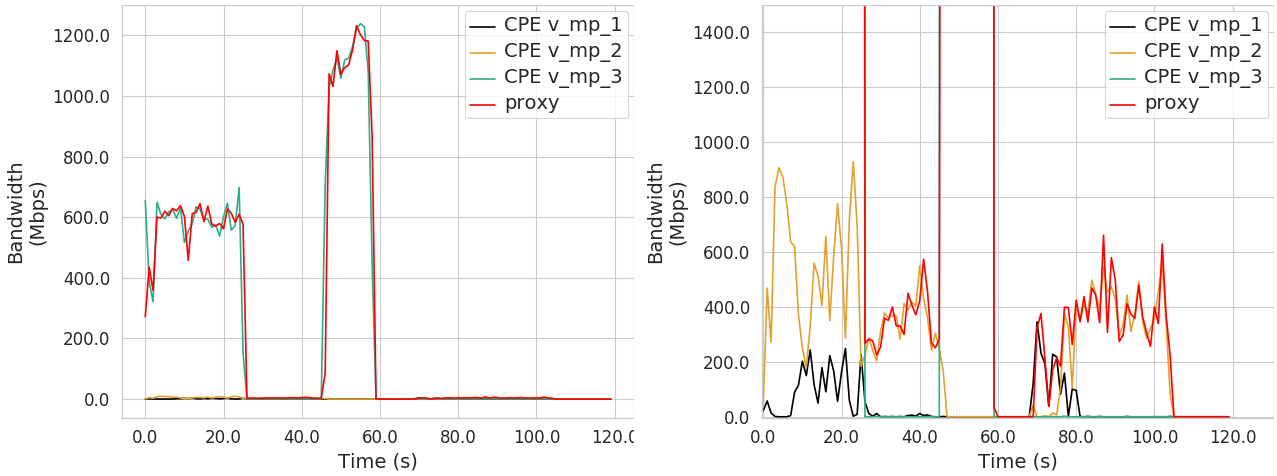


Figure 2-25: MPTCP default scheduler, using *free5gc*, changing the state (ON, OFF, BACKUP) of the interfaces. Outgoing and incoming traffic for each interface of CPE and proxy, respectively (left). Zoomed plot to show performance on paths 1 and 2 (right)

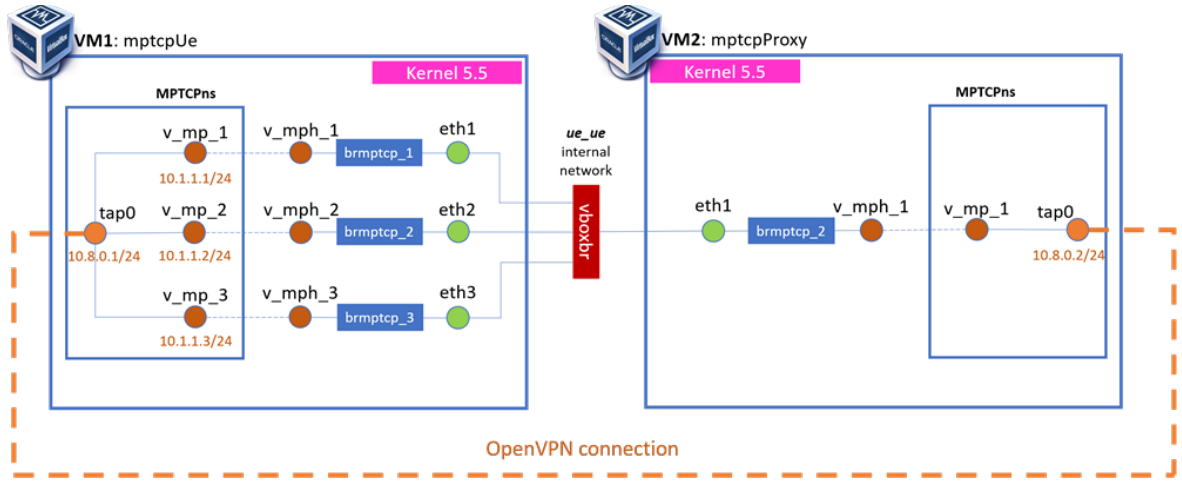


Figure 2-26: Testbed to validate MPTCP schedulers without free5gc.

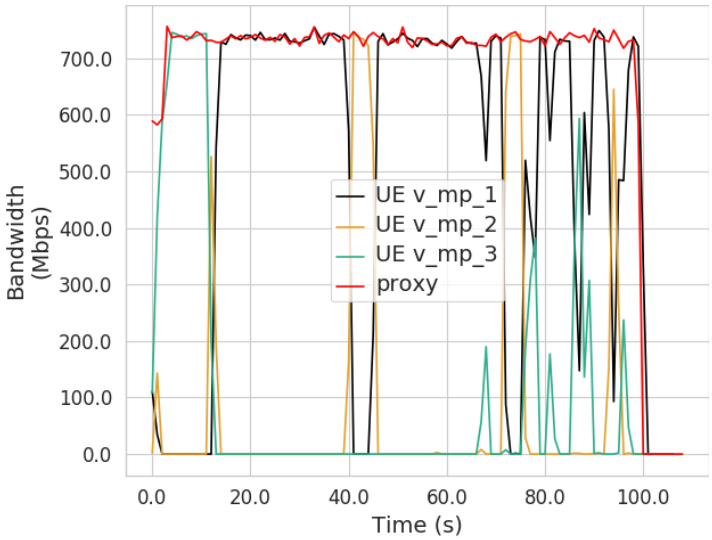


Figure 2-27: MPTCP default scheduler without *free5gc*. Outgoing and incoming traffic for each interface of CPE and proxy, respectively.

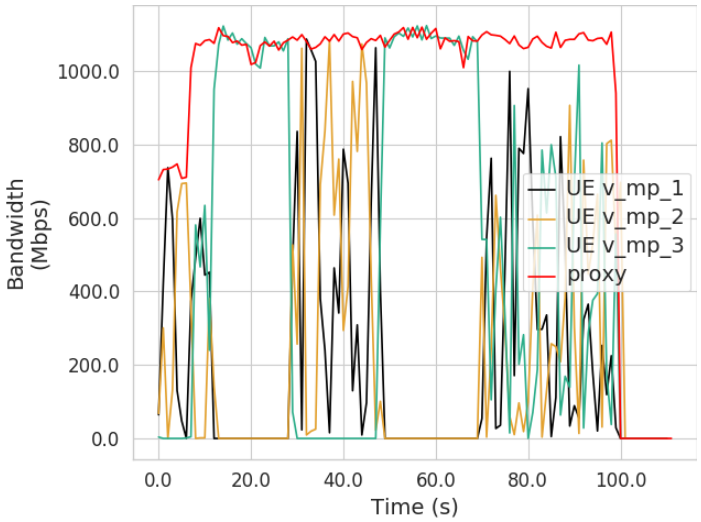


Figure 2-28: MPTCP default scheduler, without *free5gc*, changing the state (ON, OFF, BACKUP) of the interfaces. Outgoing and incoming traffic for each interface of CPE and proxy, respectively.

As depicted, the client has 3 paths (v_mp_X, X=1,2,3) and the server has 1 path. Both are connected through a switch (implemented by VirtualBox).

As shown in Figure 2-27, data rates are much higher (> 1 Gbps) than using *free5gc*. The throughput of the different paths is also very unstable since the instantaneous RTT for each interface varies depending on the TCP congestion control mechanism. This causes that, typically, one of the interfaces becomes dominant for certain periods.

Similarly, we validate the proper functioning of the ON/OFF/BACKUP states. As explained in Figure 2-28, the path on BACKUP state becomes dormant until the other two paths become inactive (due to bad coverage, connectivity problems or, as in this case, set to OFF state).

As before, the path on BACKUP state stops working until it is needed, i.e., when the other paths are inactive (due to bad coverage, connectivity problems or, as in this case, set to OFF state). In this example, path 3 changes to BACKUP, and paths 1 and 2 change to OFF at t = 15 s. Then, paths 1 and 2 change to ON at t = 25 s, changing to OFF again at t = 45 s. Finally, they all return to ON at around t = 70 s.

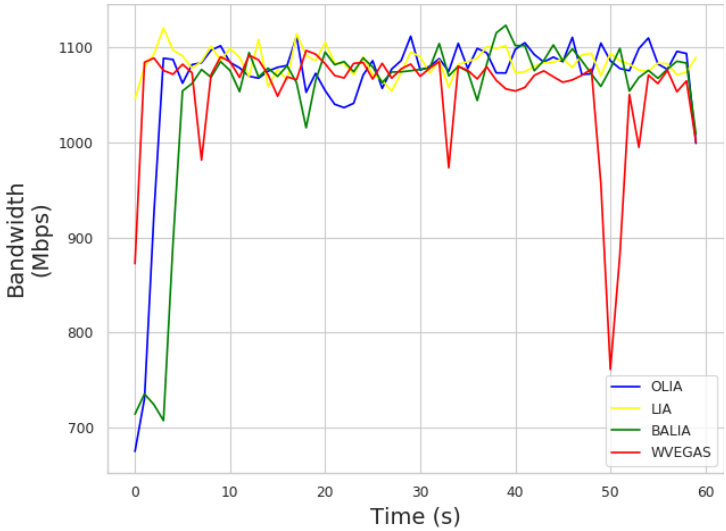


Figure 2-29: MPTCP default scheduler, without free5gc. Data rates for different TCP congestion control algorithms.

We also tested how different congestion control algorithms may have an impact on the performance of the default scheduler in this scenario. Previous results employed OLIA (Opportunistic Linked-Increases Algorithm). Figure 2-29 summarizes the results for LIA (Linked-Increases Algorithm), BALIA (Balanced Linked Adaptation) and WVEGAS (Weighted VEGAS).

As shown, in this scenario the different TCP flavours have similar performance.

2.1.3.4.2 Redundant MPTCP scheduler

This scheduler sends the same information through all the different paths. This may be convenient for ultra-reliable low latency communication (URLLC) communications, since it increases the reliability (due to multiple transmission of the same packets) and reduces the latency (the packet from the path with lower delay is received by the application).

Figure 2-30 shows the data rates from the different paths in the scenario with *free5gc* (left) along with a zoomed portion due to the differences between paths 1, 2 and 3. Similar to the experiments with the default scheduler, the data rates highly vary between paths 1-2 and path 3. As before, this may be due to the low performance of *free5gc* being executed on a VM. Path 3 achieves much higher throughput since it goes directly to the data network.

This scheduler was also evaluated on the simple scenario without free5gc. The results are shown in Figure 2-31. As before, the performance on the three paths (UE v_mp_1, UE v_mp_2 and UE v_mp_3) is similar since the paths have the same characteristics. Compared to the default scheduler, the throughput (measured by the application, i.e., *iperf*) is reduced, since the application receives the same data from all the paths.

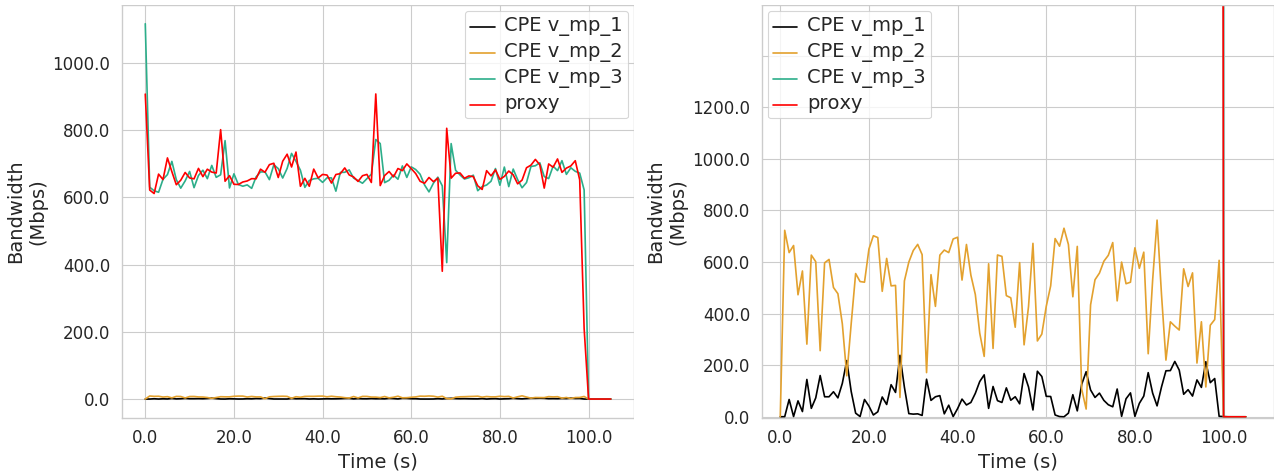


Figure 2-30: MPTCP redundant scheduler using *free5gc*. Outgoing and incoming traffic for each interface of CPE and proxy, respectively (left). Zoomed plot to show performance on paths 1 and 2 (right).

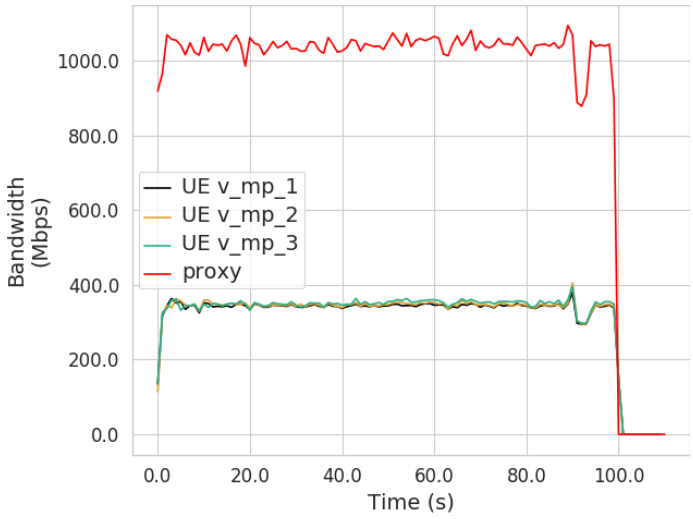


Figure 2-31: MPTCP redundant scheduler without *free5gc*. Outgoing and incoming traffic for each interface of CPE and proxy, respectively.

2.1.3.4.3 Round-robin MPTCP scheduler

The RR scheduler for MPTCP was also validated. This scheduler has a tuneable parameter, *cwnd_limited*, which specifies whether the scheduler tries to fill the congestion window on all paths (true) or prefers to leave open space in the congestion window to achieve real round-robin (false), even if the paths have very different capacities.

Figure 2-32 shows the performance of this scheduler in the scenario with free5gc for both values of the *cwnd_limited*- parameter. Since the number of packets sent from each interface is affected by the packets sent by other interfaces, the data rates are very unstable due to the low performance of the *free5gc* platform even for the third path.

Figure 2-32 presents the performance of MPTCP using the Round-Robin scheduler with *cwnd_limited*=TRUE (left) and FALSE (right) on the simple scenario. As shown in Figure 2-32(left), in the case in which the congestion window is allowed to be filled (*cwnd_limited*=TRUE), the data rate of the third path is much higher than the data rates of the other two paths.

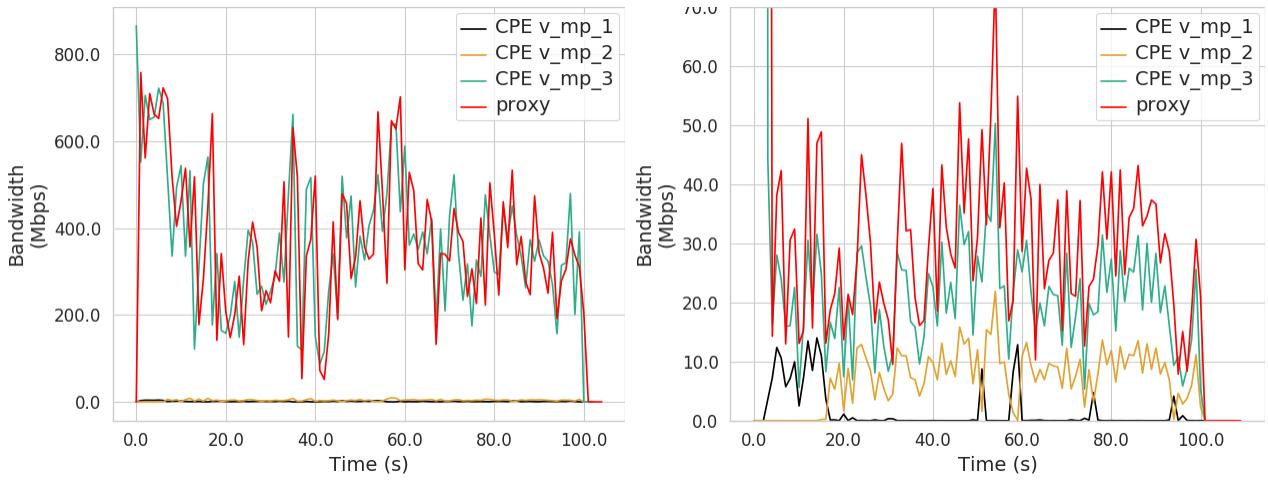


Figure 2-32: MPTCP Round-Robin scheduler with *cwnd_limited*=TRUE (left) and FALSE (right) using *free5gc*. Outgoing and incoming traffic for each interface of CPE and proxy, respectively. Data rates for the round-robin scheduler, (left) *cwnd_limited*=TRUE; and (right) *cwnd_limited*=FALSE (scenario with *free5gc*).

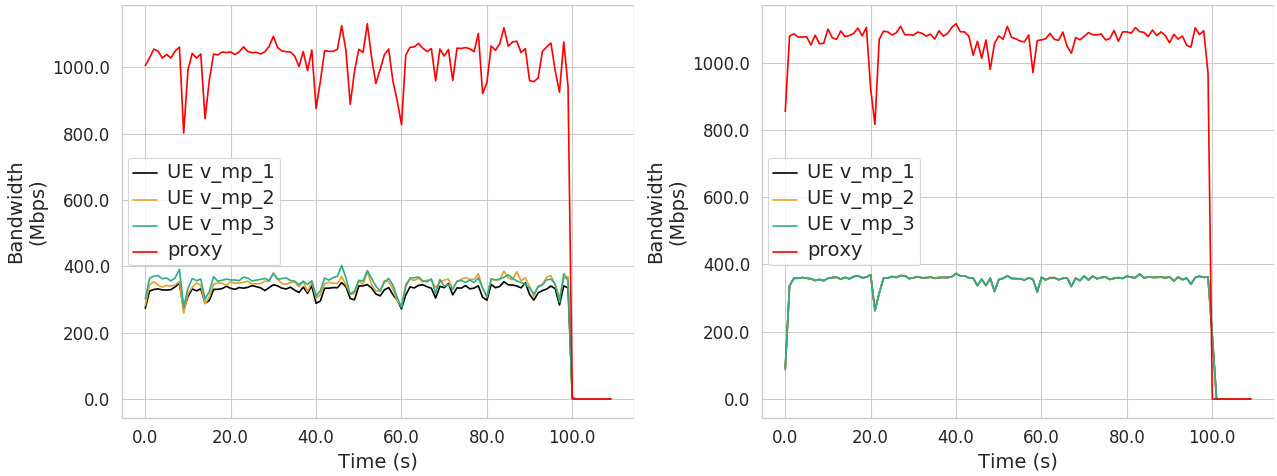


Figure 2-33: MPTCP Round-Robin scheduler with *cwnd_limited*=TRUE (left) and FALSE (right) without *free5gc*. Outgoing and incoming traffic for each interface of CPE and proxy, respectively.

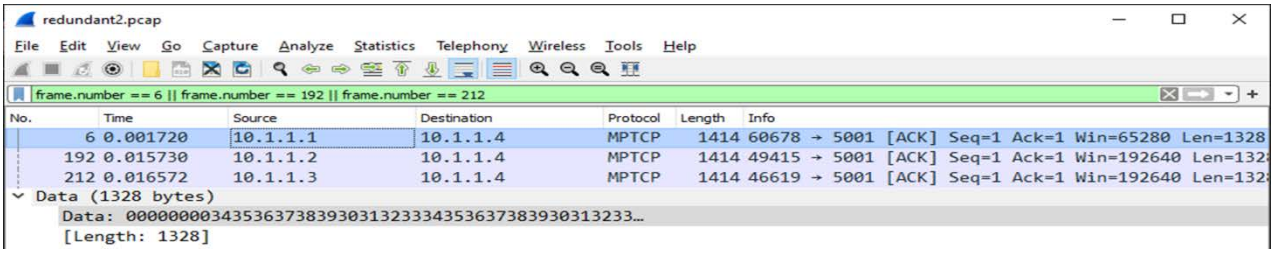


Figure 2-34: First packet of the different flows, corresponding to the three paths.

However, when the congestion windows are not allowed to be filled (real round-robin), the data rates of the paths become more similar. Path 3 continues achieving the best data rate, but probably due to the low performance of *free5gc*.

When the test is performed on the simple scenario (without *free5gc*), the data rates of the different paths are very similar (see Figure 2-33) since they have the same characteristics, irrespective of whether *cwnd_limited* is true (left figure) or false (right figure). However, the instantaneous data rates in the second case are almost identical since a true round-robin scheduler is employed.

It was checked if the information sent through all the different paths are the same. Figure 2-34 shows the first packet for each flow, which contains the same information on the data field.

2.2 5G-CLARITY eAT3S

In 5G-CLARITY D3.1, Section 6.2.3 [1], it is noted that 5G-CLARITY multiaccess based multi-connectivity solution will be based on an enhanced version of 3GPP AT3S framework, named enhanced AT3S (eAT3S). The idea of eAT3S is to incorporate near-RT RAN control of the AT3S policies behind the UPF by utilising multi-WAT telemetry in a xApp within the O-RAN RIC framework. The motivation behind introducing such an enhancement is to mitigate any increase on latency in case policy control function (PCF) and SMF are deployed separately such as having SMF within the private premises and PCF inside the mobile network operator (MNO). The same problem also exists in case SMF is also deployed inside the MNO and UPF is deployed inside the private premises. Such an enhancement also enables efficient use of available access networks by taking into account near RT telemetry or predictions based on the collected telemetry.

2.2.1 5G-CLARITY eAT3S operational flows

In the standard 3GPP AT3S framework, predefined AT3S rules are used to steer the access traffic onto 3GPP

and non-3GPP networks. This can be considered as the depicted flow from AT3S Rule #2 on the left-hand side of Figure 2-35, which is also included in 5G-CLARITY D3.1 [1]. However, in 5G-CLARITY eAT3S framework the idea is to provide flexibility on these predefined rules in order to reflect near-RT performance measurements. Moreover, as 5G-CLARITY non-3GPP network is composed of the integrated Wi-Fi-LiFi network, a further traffic steering capability can be applied to Wi-Fi and LiFi networks. Accordingly, the traffic assigned onto non-3GPP access network in an ATSSS rule can be further steered among Wi-Fi and LiFi networks with different weights. For an initial algorithm development, an equal distribution of non-3GPP traffic to Wi-Fi and LiFi networks is considered in this deliverable. The updated eAT3S rules may also take into account any machine learning (ML)-based prediction on the access link status, user mobility, etc.

To enable access traffic steering rule update based on near-RT telemetry, 5G-CLARITY uses eAT3S xApp that resides in O-RAN near-RT RIC. The eAT3S xApp gathers near-RT telemetry data from 5GNR, Wi-Fi and LiFi access nodes along with ATSSS path performance measurements from the user plane, and modifies the traffic steering rules in near-RT (RT-ATSSS rule).

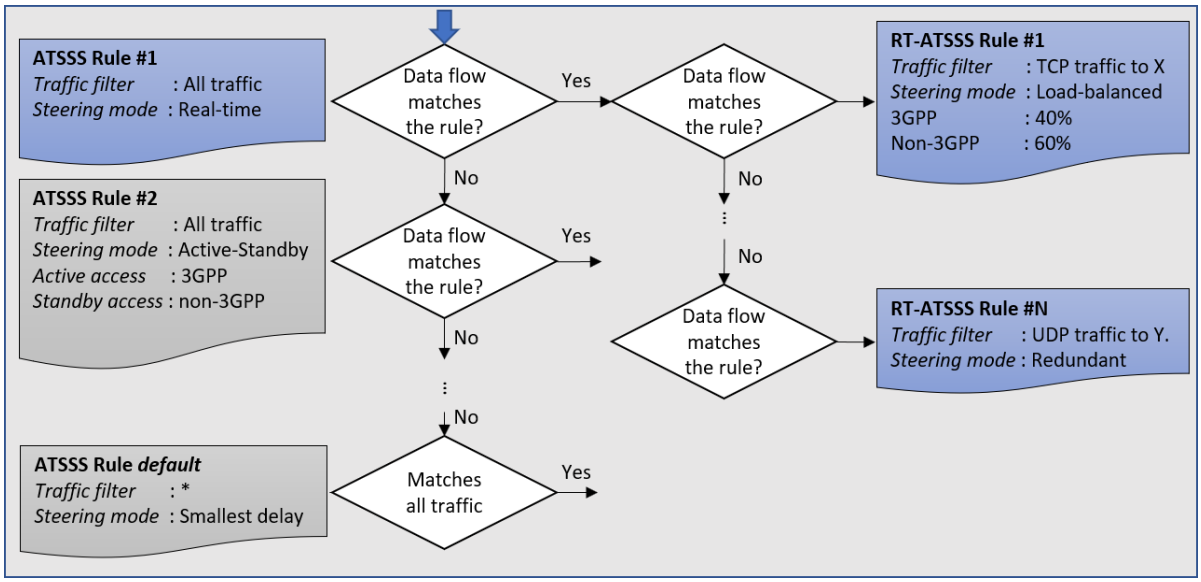


Figure 2-35: Flowchart of traffic distribution based on 5G-CLARITY eAT3S rules [1]. Gray boxes represent what is available in the conventional 3GPP ATSSS and blue boxes represent 5G-CLARITY eAT3S contributions.

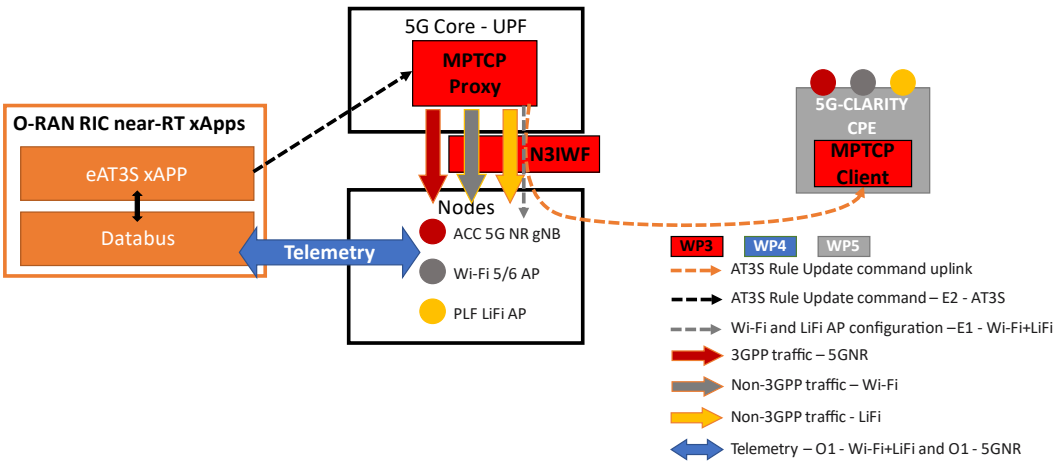


Figure 2-36: 5G-CLARITY eAT3S xApp environment. The details of the shown 5G-CLARITY interfaces can be found in [1]. Red, blue and grey highlighted blocks represent 5G-CLARITY WP3 (in addition to eAT3S xApp), WP4 and WP5 related work items, respectively.

As the eAT3S xApp resides in O-RAN near-RT RIC platform, the eAT3S user plane function can be considered as behind the UPF. Therefore, any update on the traffic steering rules is conveyed to the 5G-CLARITY eAT3S user plane function aka custom MPTCP scheduler.

Figure **2-36** depicts the considered 5G-CLARITY eAT3S xApp environment. For O-RAN RIC platform, Accelleran O-RAN dRAX RIC near-RT platform is used to deploy the eAT3S xApp. The platform provides necessary functional blocks for an xApp deployment [10]. In this deliverable, the 5G-CLARITY eAT3S xApp algorithm development is reported. Integration of the eAT3S xApp algorithm to Accelleran O-RAN dRAX RIC near-RT platform and relevant interfaces to push access traffic steering rule updates to the 5G-CLARITY custom MPTCP scheduler as described in Section 2.2.3 will be reported in 5G-CLARITY D3.3.

2.2.2 5G-CLARITY eAT3S control plane: eAT3S algorithm development

The implemented 5G-CLARITY multiaccess solution will be based on 3GPP (5GNR) and non-3GPP (Wi-Fi and LiFi) integration. Accordingly, the telemetry and performance measurement values will be used to configure the weight of different access traffic flows. In this deliverable, the traffic routing is realised by thresholding with network service KPIs and WAT-specific telemetry data/performance measurements.

2.2.2.1 Input parameters

As noted, 5G-CLARITY eAT3S uses WAT-specific telemetry data/performance measurements and network service KPIs to configure ATSSS rules. In 3GPP ATSSS framework [6], the access measurement report procedure considers certain threshold levels and a timer value to trigger measurements. More specifically, for 5G access network, certain threshold values for reference signal received power (RSRP) and reference signal received quality (RSRQ) levels are used to trigger RSRP and RSRQ measurements, and to assess the signal level and quality. For non-3GPP access network, more specifically Wi-Fi network, the threshold and measurement parameters are noted as, (i) channel utilization; (ii) downlink backhaul available bandwidth; (iii) uplink backhaul available bandwidth; and (iv) average beacon received signal strength indicator (RSSI).

In 5G-CLARITY eAT3S, in addition to the noted threshold values and measurement parameters, physical resource block (PRB) utilization, e.g., the average number of PRBs used for data traffic and total available PRBs, and signal-to-interference-plus-noise ratio (SINR) of users accessing to 5G network are continuously collected at the RIC, as depicted in **Figure 2-36**. Moreover, the measurement parameters considered for a Wi-Fi network are extended for LiFi access networks. Accordingly, two sets of non-3GPP access network threshold values and measurement parameters are collected at the RIC and used in the eAT3S framework.

As the purpose of 5G-CLARITY eAT3S is not only to optimize near-RT access network performance but also incorporate any network preference on the traffic steering rules/weight factors, network service KPIs as well as network operator policies are used as input for the eAT3S rule configuration algorithm. As the proposed eAT3S framework reacts on near-RT changes on the link performance, it can be expected that the configured eAT3S rules may diverge from the operator policies. However, in a long run, the proposed eAT3S algorithm takes into account the network operator preference and converges to the operator policies.

2.2.2.2 Access traffic steering modes and rules

The steering modes described in 3GPP ATSSS framework are, (i) active-standby, (ii) smallest delay, (iii) load-balancing, and (iv) priority-based. When either the active-standby or priority-based steering mode is selected, a priority information element is used to indicate at which condition 3GPP or non-3GPP access network is used to transmit the data flows. However, when the load-balancing steering mode is selected, a weight factor is used to indicate the proportion of the traffic to be forwarded to 3GPP and non-3GPP access networks.

In Figure 2-37, "active-standby" steering mode is depicted. As can be seen from the figure, the preferred access network, either 3GPP or non-3GPP, could be utilized "actively" for the traffic, where the other access will be in "standby" and will be activated in the case of unavailability. Once the initial "active" access becomes suitable/available for traffic, the traffic will be transferred to the initially assigned "active" access. In 5G-CLARITY eAT3S framework, when the "active" access is chosen as non-3GPP network, both Wi-Fi and LiFi networks are used to transmit data flows. As long as one of the non-3GPP access networks is considered as available (measurement parameters are above the threshold), the "standby" 5G network is not going to be activated.

Figure 2-38 depicts the "load-balancing" steering mode where each WAT receives a certain portion of the traffic in load-balance steering. As depicted in the figure, γ_G, γ_W and γ_L , ($\gamma_G + \gamma_W + \gamma_L = 100\%$) denotes the load balancing weight factors for the 5GNR, Wi-Fi and LiFi networks in eAT3S, respectively. It is also important to note that in load-balance steering mode, the traffic split is done at the access-level. Therefore, within the non-3GPP access another low-level allocation among Wi-Fi and LiFi WATs could also be realized in a flexible level based on the favourable conditions e.g., channel or noise, provided that $\gamma_{\text{Non-3GPP}} = 100\% - \gamma_G$. In this steering mode, the multiaccess PDU (MA-PDU) session can provide bandwidth aggregation and traffic load assignment onto each available access network capabilities.

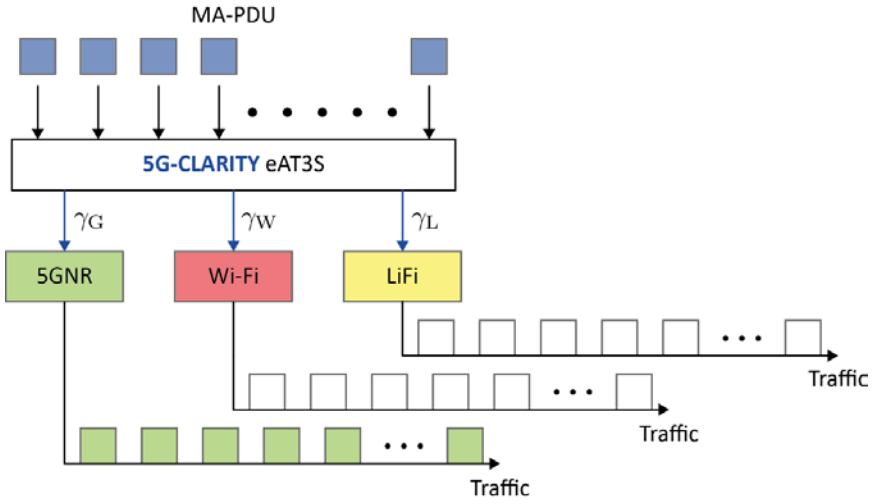


Figure 2-37: AT3S active-standby steering mode with active access of 3GPP.

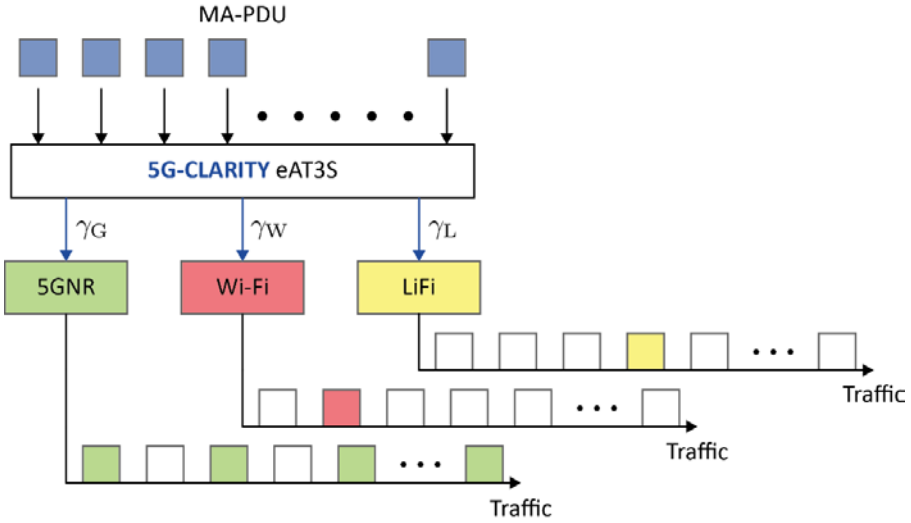


Figure 2-38: AT3S load-balancing steering mode with γ_G=60% (3GPP) and $\gamma_{\text{Non-3GPP}}$=40% (non-3GPP).

As previously depicted in Figure 2-35, an ATSSS rule includes (i) a precedence value; (ii) a traffic descriptor; (iii) a steering mode; and (iv) a steering function [6]. The precedence value identifies the priority level of the rule. The traffic descriptor is used to identify service data flows that may include application identities, IP descriptors and non-IP descriptors. The steering mode indicates the chosen steering mode (active-standby, smallest delay, load-balancing and priority-based). The steering function indicates the type of steering function that can be a lower layer steering or higher layer steering to steer the traffic of the matching data flows. 5G-CLARITY eAT3S introduces a new AT3S rule which is defined as "real-time" steering mode rule. This rule is considered as the highest priority rule, Rule #1, and is adaptive to network status. For example, if there is a congestion on non-3GPP network (the channel utilization parameter is above the network load threshold) and the load balancing mode is used to enforce 80% for non-3GPP network, the eAT3S algorithm modifies/overwrites the weight factors. Once the non-3GPP network load reaches an acceptable level (the channel utilization parameter is below the network load threshold), the eAT3S algorithm pushes higher weight factors for non-3GPP network in order to converge to the operator-defined weight factors in the long term. Table 2-1 presents the basis of the proposed eAT3S algorithm on determining load-balancing weight factors. A similar approach can also be applied to other steering modes where in the active-standby and priority-based steering modes, only one of the three WATs can be chosen to transmit data flows.

Table 2-1: eAT3S Algorithm

1:	Access network-specific telemetry data, performance measurements, measurement threshold values/parameters and network operator preference on the active access/priority access/load-balancing weights (e.g., γ_G=60% (3GPP) and $\gamma_{\text{Non-3GPP}}$=40%) are available.
2:	A list of active users *U* is obtained.
3:	**while** $U \neq \{\}$ **do**
4:	Consider/obtain a condition based on a specific measurement/telemetry parameter such as SINR, path performance measurements (RSRP, RSRQ, etc.), gNB/AP resource utilization, etc. for all available access technologies (5G/Wi-Fi/LiFi)
5:	**if** The considered condition is satisfied for all access technologies **then**
6:	Apply operator preferences for steering mode/load-balancing weight factors
7:	**else if** The considered condition is not satisfied for only 5G **then**
8:	Steer traffic flows to Wi-Fi and LiFi networks according to their condition satisfaction level (e.g., higher the satisfaction level, higher the weight factor)
9:	**else if** The considered condition is not satisfied for only Wi-Fi **then**
10:	Steer traffic flows to 5G and LiFi networks according to their condition satisfaction level (e.g., higher the satisfaction level, higher the weight factor)
11:	**else if** The considered condition is not satisfied for only LiFi **then**
12:	Steer traffic flows to 5G and Wi-Fi networks according to their condition satisfaction level (e.g., higher the satisfaction level, higher the weight factor)
13:	**else if** The considered condition is satisfied for only one of the access networks **then**
14:	Steer all traffic flows to the access network that satisfies the condition (e.g., the weight factor for this access network becomes 100%)
13:	**else if** The considered condition is not satisfied for all access technologies **then**
14:	There is no traffic flow assignment for this user (e.g., the weight factor for all access networks becomes 0%)
15:	**end if**
16:	Go to the next user in the list *U*
17:	**end while**

2.2.2.3 Initial results

In order to validate the proposed eAT3S algorithm in Table 2-1, an indoor scenario, where 5GNR gNB, Wi-Fi and LiFi APs are deployed, is simulated. The simulation platform is detailed in Section 5. The simulated scenario to validate the proposed eAT3S algorithm considers a mobile user and 100 other static users which is aligned with UC2.1 in 5G-CLARITY D5.1 [3]. The mobile user can be considered as an automated guided vehicle (AGV) and 100 other static users can be considered as data demanding machineries such as switches, programmable logic controllers, etc., inside a factory. The performance of the proposed eAT3S algorithm is validated by the achieved user data rate and WAT-specific load distributions. Accordingly, SINR is considered as the condition to check whether WATs can satisfy the given SINR threshold that can also reflect achievable user data rate. Hence, the condition satisfaction level noted in Table 2-1 provides an insight on the amount of achievable user data rate from a specific WAT which can be used to decide the weight factor. In order to clarify this further, once one of the WATs cannot satisfy the SINR threshold value of 5 dB, the other two WATs share the overall traffic either equally or based on the ratio between their achieved SINR values. In algorithm validation, two different AT3S rules that reflect the operator preference on load assignment to 3GPP and non-3GPP networks are considered to compare how the load weight factors are distributed in different conditions. The first AT3S rule assigns γ_G = 70% and $\gamma_{\text{Non-3GPP}}$ = 30%, whereas the second AT3S rule assigns the opposite, γ_G= 30% and $\gamma_{\text{Non-3GPP}}$ = 70%. It is assumed that LiFi APs are deployed along the path of the mobile AGV to keep some of the static users intentionally uncovered. In other words, not all three WATs are always available for all users. Therefore, it can be expected that the load distribution among non-3GPP networks Wi-Fi and LiFi will not be equal.

Figure 2-39 shows the achieved data rate performance of AGV and the WAT-specific load distribution performance of AGV and overall system for the considered two AT3S rules. It can be observed that when all WATs considered as available (their signal level/SINR is above the threshold), the proposed algorithm achieves the network operator-preferred load weight factor assignment. When one of the WATs is not always available (the signal level/SINR of LiFi is below the threshold for some users), the eAT3S algorithm allocates more data flows to Wi-Fi in a way to converge to the network operator-preferred load assignment.

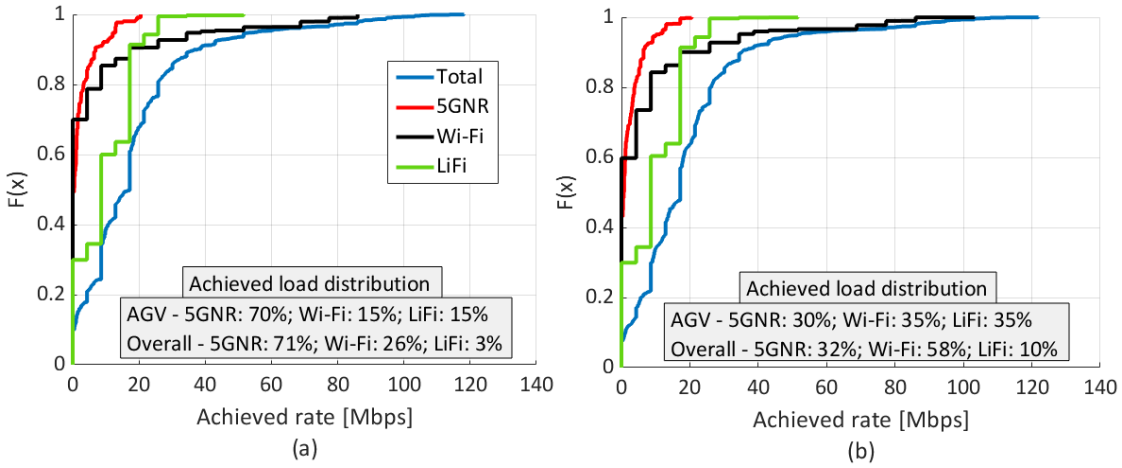


Figure 2-39: Performance of the eAT3S algorithm (a) γ_G=70% and $\gamma_{\text{Non-3GPP}}$=30%, (b) γ_G=30% and $\gamma_{\text{Non-3GPP}}$=70%.

2.2.3 5G-CLARITY eAT3S user plane function: custom MPTCP scheduler

Given that a requirement of the eAT3S framework is to be able to control the amount of packets transmitted through each interface, in this section we investigate how to extend the MPTCP user plane function described in Section 2.2.1 to support this capability. The MPTCP RR scheduler explained in Section 2.1.3.4.3 transmit the same number of packets for each subflow. To allow different throughput for each subflow, the

MPTCP RR scheduler is modified so that a weight to MPTCP subflows can be assigned. At this moment, the rule to specify the intended subflow is by its source or destination IP address. The weights refer to the number of TCP segments that a subflow will emit each round.

Moreover, several tools are implemented to interact with the Weighted RR (WRR) scheduler, so the weights can be assigned or checked during a MPTCP transmission. These tools allow to get some monitoring parameters for the MPTCP subflows such as their congestion window sizes or the estimated smoothed RTTs (SRTT).

2.2.3.1 Weighted Round Robin algorithm

The weighted round robin (WRR) scheduler implements an algorithm in which, for each round, it selects a subflow which has not sent yet the number of segments specified as a weight. The basis of the algorithm is presented in Table 2-2.

Table 2-2: Weighted Round Robin Pseudocode

dst_address_list: List of destination address with a weight. src_address_list: List of source address with a weight. mptcp_subflow_list: list of subflows for a MPTCP socket. n_segments: number of TCP segments to be sent through a subflow in the active round. subflow.sent: number of segments sent during the active round for a specific subflow. It has to be lower or equal than n_segments.

function send_next_segment(segment):
segment_sent=false
while !segment_sent:
iteration=0
while iteration < size(mptcp_subflow_list) && !segment_sent:
subflow=mptcp_subflow_ip[iteration]
n_segments=1
 	
if subflow.ip_src_address in src_address_list:
n_segments=weight[subflow.ip_src_address]
endif
if subflow.ip_dst_address in dst_address_list:
n_segments=weight[subflow.ip_dst_address]
endif
 	
if subflow.sent<n_segments:
send_segment(subflow,segment)
subflow.sent= subflow.sent + 1
segment_sent=true

```
                else
                        subflows_full=subflows_full+1
                endif

                iteration=interation+1
        endwhile

        if subflows_full==iterations:
                for i=0 to size(mptcp_subflow_list)-1:
                        subflow.sent = 0
                endfor
        endif
endwhile
```

2.2.3.2 Configuration of the WRR scheduler parameters

There are several ways of interacting with the scheduler:

- **Using the sysctl command**. The configuration of the WRR scheduler can be set in the /etc/sysctl.conf file. This is a static configuration. However, it can be modified at runtime via the sysctl –w <param_name> command line tool. It can be checked with sysctl <param_name> command line. Figure 2-40 shows the list of MPTCP parameters implemented in the Linux kernel 5.5.0.
- **Accessing via the /proc/sys/net/mptcp files**. The previous tool accesses the kernel parameters that are exposed as part of the /proc filesystem. More precisely, the following files are available:
    - mptcp_wrr_li_weights: readable/writable parameter. The content is a list of pairs of source IP address to be weighted and weight. These weights should be higher or equal than 1. The IP addresses are represented as a 32-bit, network byte order integer. Current implementation defines only 4 different IP addresses.
    - mptcp_wrr_ri_weights: readable/writable parameter. Same than the previous parameter but specifying destination IP addresses and their weights.
    - mptcp_wrr_cwnd: readable parameter. List of pairs of IP addresses and their corresponding congestion window size.
    - mptcp_wrr_srtt: readable parameter. List of pairs of IP addresses and their corresponding SRTT estimated. The SRTT values are expressed in microseconds.
- **Using a python library**. We have implemented a first version of a helper library of functions which perform the assignment and the lookup of the previous parameters. It uses system calls through the *sysctl* command line tool. It can be optimized by accessing directly to the /proc filesystem parameters' kernel files. These functions can be called from python scripts.
- **Using a python command line tool**. To ease the test of the common tasks, we have created a command line python tool which uses the previous python library.

Using a REST API. This implementation is in progress. We are using Flask as a REST framework, to

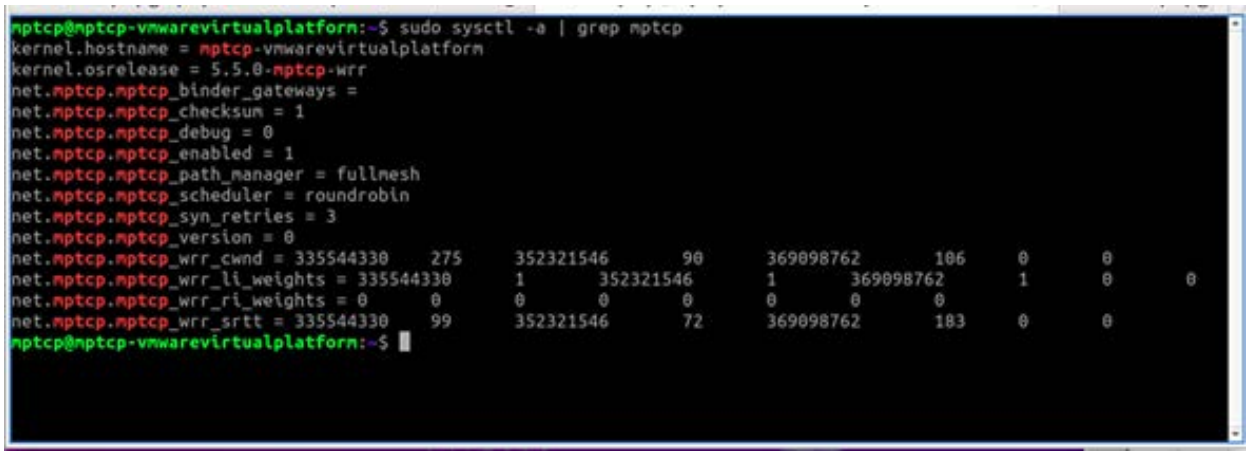


Figure 2-40: Variables of MPTCP WRR implementation, exposed as kernel parameters.

perform the previous functions that will be callable via GET/POST HTTP requests. The request parameters and results will be encoded as JSON messages. This API is the interface that the eAT3S xApp will use to control the packet allocation.

2.2.3.3 WRR Validation

To validate the WRR implementation, the topology shown in Figure 2-41 is configured in the Common Open Research Emulator (CORE)[7] network emulator. This emulator uses Linux Containers to encapsulate the network nodes (routers, switches, hosts, etc). This approach allows to deploy large networks setups. However, every node uses the same kernel, although has its own kernel namespaces. For this test, this setup permits a rapid prototyping.

The setup shown in Figure 2-41 represents a CPE with three interfaces connected to a proxy, with only one interface. All the interfaces belong to the 10.0.0.0/8 subnet. *iperf* is used to generate 60Mbps from the CPE to the Proxy. *Ifstat* is used to monitor the throughput of the CPE's and proxy's interfaces. The *cwnd_limited* parameter is set to false.

The test consists in changing the weights' assignment of the CPE's interfaces each 20 seconds. Figure 2-42 shows the results of this test. The curves represent the CPE's outgoing traffic, when a (w0, w1, w2) vector of weights are assigned to the interfaces. Concretely, w0, w1 and w2 are assigned to the eth0, eth1 and eth2 interfaces, respectively.

The results evidence the correct scheduler behaviour, as the throughput is distributed proportionally to their weights. More specifically, eth0 sends the w0/(w0+w1+w2) fraction of the 60Mbps traffic sent from the CPE, eth1 sends the w1/(w0+w1+w2) fraction, and eth2 sends the w2/(w0+w1+w2) fraction.

Our next steps will consist of: (i) investigating ways to support multiple CPEs from a single MPTCP proxy, (ii) extending the way to assign weights to subflows specifying source and/or destination ports, (iii) developing the API, and (iv) demonstrate how the eAT3S xApp can drive packet steering based on real-time telemetry.

[7] http://coreemu.github.io/core/

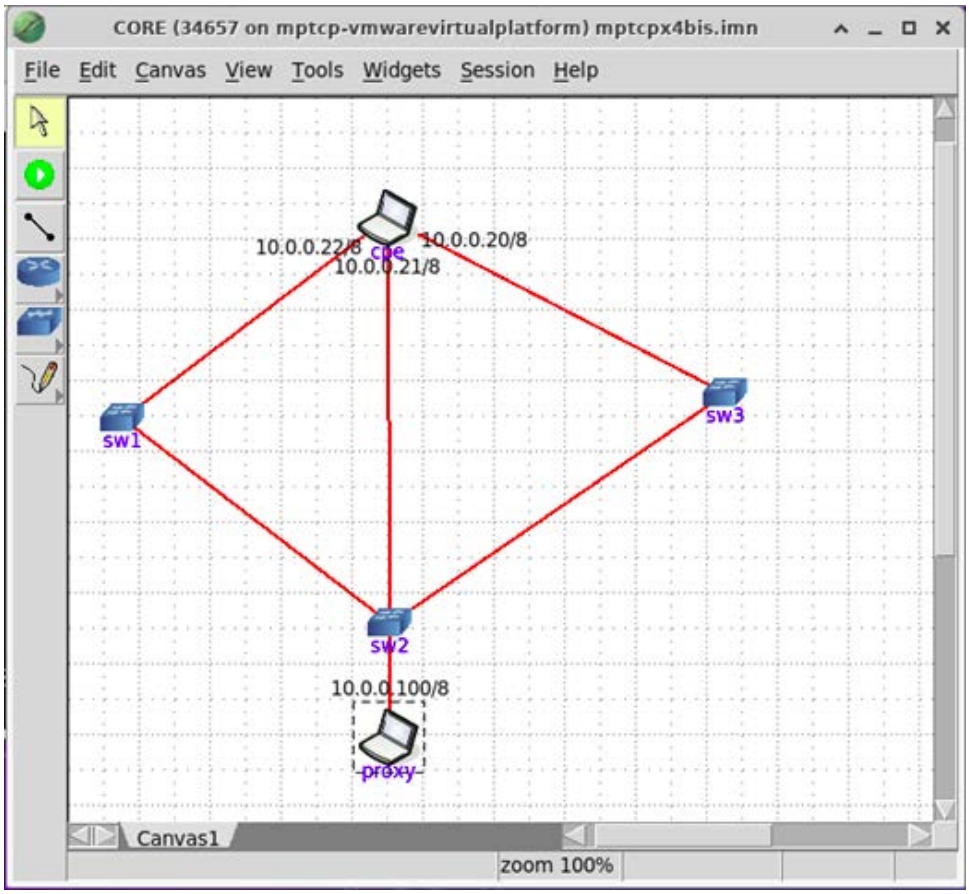


Figure 2-41: Experimental topology implemented in CORE-EMU environment, featuring a CPE with three network interfaces, and a MPTCP proxy with only one interface.

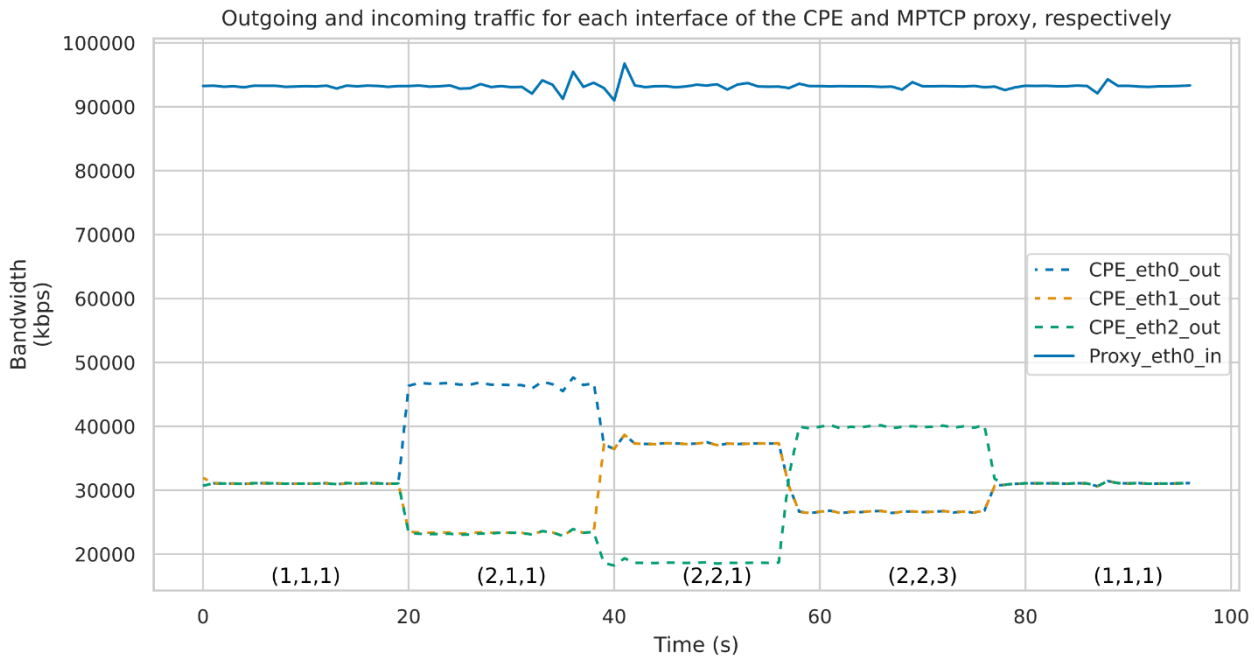


Figure 2-42: Throughput for each CPE interface with different weights during the experiment. The weights are shown below the curve.

3 Scheduling and Resource Management

Unlike the legacy cellular networks, where the utilization of vendor/service operator-specific preferences and predefined resource management policies are adopted, the definition of the resource and service has been changed significantly in 5GNR systems. Accordingly, three main service categories in 5G, i.e., enhanced mobile broadband (eMBB), URLLC, and the massive machine type communication (mMTC) are defined along with the flexible frame structure and network slicing concept. The adequate allocation of the resources is a long-lasting problem for both wired and wireless networks and the service flexibility and efficiency emphasized in 5GNR will play the key role to overcome the mentioned challenges. In 5G-CLARITY, as introduced in [1], the resource management will be considered as a two-stage process, i.e., i) traffic/packet routing, and ii) gNB/AP-level resource scheduling, as depicted in Figure 3-1.

The ubiquitous traffic/packet routing in 5G is a UPF-level AT3S problem. The related telemetry and performance measurements will be used to route traffic flows among 3GPP and non-3GPP networks in near-RT by eAT3S to ensure the quality of service (QoS). Once a traffic flow is routed to 5G, Wi-Fi or LiFi network, the AP level resource scheduling will be taken care of by gNB and/or Wi-Fi AP and/or LiFi AP. Furthermore, the user access is controlled by AP level functions to support both DL and UL flows. In order to emphasize the functional advantages of 5GNR networks comprehensively, both the AP level and service level resource scheduling techniques will be investigated in the following subsections.

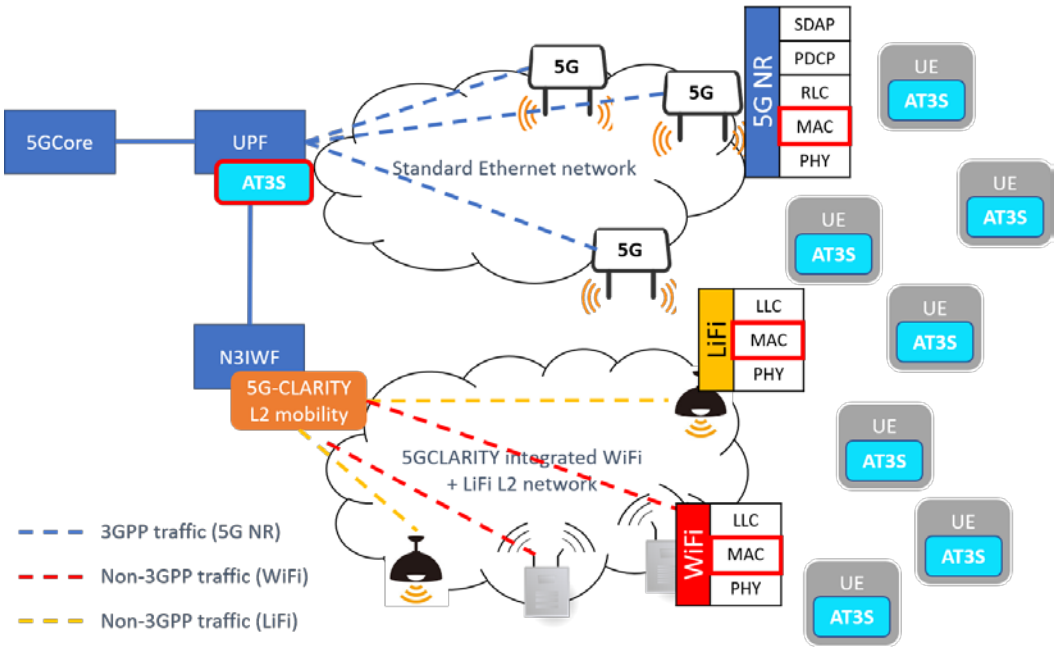


Figure 3-1: High-level network architecture of 5G-CLARITY [1]

3.1 Wi-Fi airtime-based scheduling to support 5G-CLARITY wireless quotas

An evaluation of the airtime management capabilities available in the Linux kernel is presented here, which can be used to control airtime allocations on a per-SSID basis. This airtime management framework is therefore aligned with the wireless quota concept presented in 5G-CLARITY D4.2 [11], which can be used to segregate multi-WAT resources for a given 5G-CLARITY slice.

Let us start by describing the Linux wireless architecture described in Figure 3-2, consisting of the following three layers that are available in the Linux kernel to control any Wi-Fi vendor chipset:

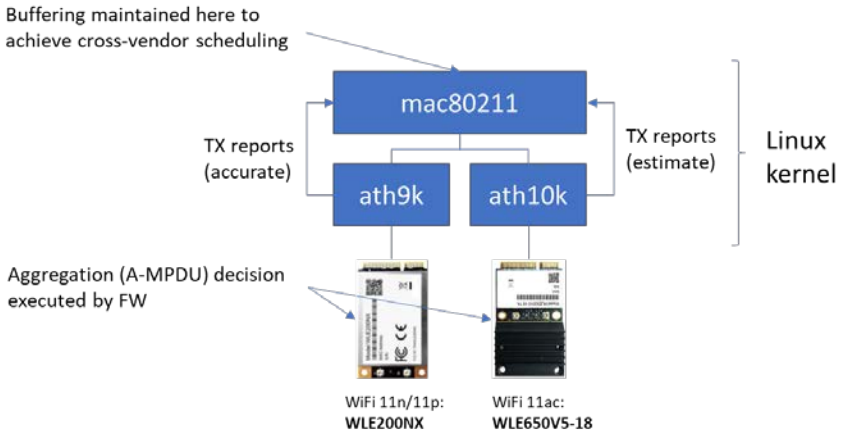


Figure 3-2. Kernel modules in charge of airtime scheduling

- mac80211, it is a kernel module common to all Wi-Fi chipsets that can operate in Linux. This module handles the wireless interface that is part of the network subsystem, and implements the higher layers of the Wi-Fi MAC, e.g., frame creation, beaconing, etc. mac80211 includes a queuing system that buffers frames on a per station and per traffic identifier (TID) basis. In the latest versions of the kernel, an airtime-based scheduler is available in mac80211 [12] that decides how to dequeue packets from these queues based on an airtime quantum. Notice though that this airtime-based scheduler needs to understand the actual time consumed by each transmission.
- Below mac80211 there are vendor specific drivers. These are the drivers that manage the interface with the hardware module at the bottom of Figure 3-2 while offering a common API towards mac80211. In Figure 3-2 we can see the ath9k and ath10k drivers that support the integration of the 802.11n and 802.11ac chipsets from Qualcomm Atheros. A key feature required from these drivers regarding airtime scheduling is the ability to report the actual airtime consumed by a transmission. This is done through transmission reports (TX reports). Understanding the actual transmission time in the kernel is hard because there are some transmission-related decisions, like Aggregated MAC Protocol Data Unit (AMPDU) aggregation, which are left to the hardware module, since it would be too slow to take these decisions in software. In the case of ath9k, this module is responsible for the aggregation logic, and therefore ath9k can deliver an accurate airtime consumption report to mac80211. However, ath10k leaves the aggregation logic to the hardware module, and thus in this case only an estimate of the consumed airtime can be delivered to mac80211.
- Below the vendor dependent modules, we have actual hardware modules, which handle all real-time related issues, such as the CSMA/CA protocol and in the case of ath10k the aggregation logic.

In the next sections, the airtime allocation capabilities available in the Linux kernel which can be used to support the concept of wireless quotas used in the 5G-CLARITY slicing concept will be evaluated.

3.1.1 Ath9k evaluation

We start by evaluating airtime scheduling in ath9k with the following setup. We use Linux kernel 5.4 and Ubuntu 18.04 as operative system. We setup one AP configured in a 40MHz channel, radiating two SSIDs, each one representing one slice. The hostapd software[8] is used to configure the operation of the APs. Two stations connect to the AP, one for each slice. In this setup we build a software layer that allows us to control

[8] https://en.wikipedia.org/wiki/Hostapd

the airtime resources on a per-SSID basis, defining a minimum percentage of the airtime utilisation guaranteed for each SSID. Within an SSID, airtime is shared equally among the connected stations.

To evaluate the behaviour of the airtime allocation policies we use a Prometheus exporter for hostapd[9] , which reports TX/RX bytes and airtime per associated station. Thus, we connect the Prometheus server to Grafana[10] to generate the monitoring dashboards depicted in Figure 3-3, which include airtime per slice (upper plot) and throughput per slice (lower plot).

Looking at Figure 3-3, starting from the left-hand side, we can see a first period where the stations of each slice transmit independently. First, the station from slice 1 (green) transmits until time 11:09, achieving 100% of channel utilisation (upper graph) and around 220 Mbps (lower graph). Second, between 11:09 and 11:10, the station from slice 2 (yellow) transmits achieving 100% airtime utilisation (upper graph) and around 160Mbps (lower graph). The reason for different throughput performances is the experienced radio channel conditions that result in station 2 selecting a lower modulation and coding scheme (MCS). Notice that regardless of the configured airtime percentages, when transmitting in isolation, all slices can grab all the resources of the air interface. Third, starting at around 11:10:20, the two slices start to operate concurrently while we progressively adjust the airtime weights configured for each slice, which are indicated in white percentage values in Figure 3-3. Looking at the airtime utilisation (upper plot) we can see how the airtime allocation smoothly follows the configured weights, while maintaining the aggregated utilisation at 100%. Updating the airtime weights is also proven to be very quick with the new policy becoming effective immediately. Overall, ath9k based airtime scheduling is able to enforce airtime allocation policies in real-time while maintaining an overall aggregate throughput around 200 Mbps. The only negative note on the reported results is that we observed that when ratios above 80:20 were configured, the resulting allocations were not enforced.

Based on the results reported in Figure 3-3, it is concluded that Wi-Fi based airtime allocation based on 802.11n and ath9k, is a suitable solution to support the concept of wireless quotas defined by the 5G-CLARITY slicing concept.

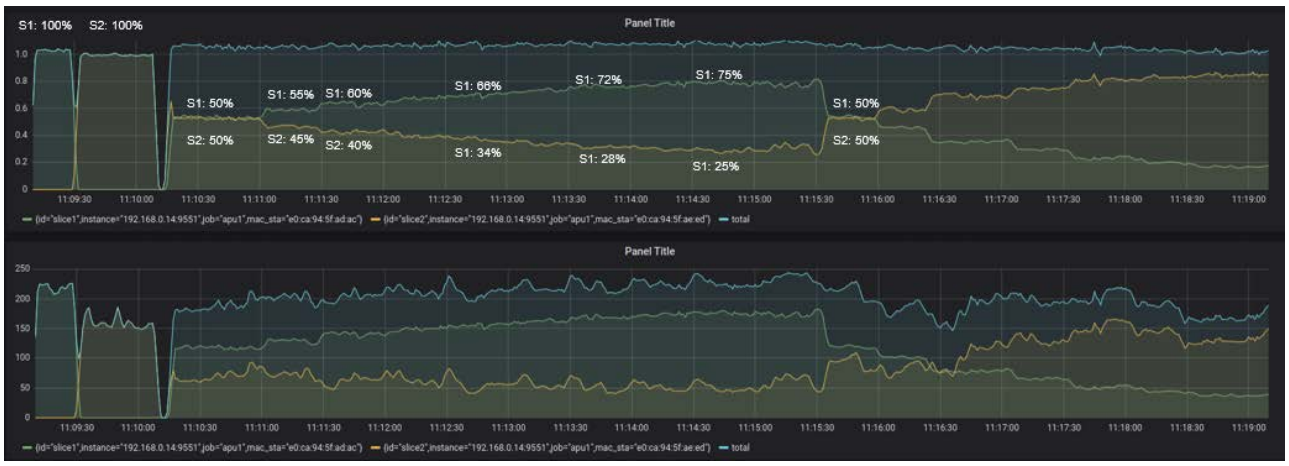


Figure 3-3. Evaluating airtime scheduling with 802.11n/ath9k 40MHz channel – airtime per slice (upper) and throughput per slice (lower).

[9] https://github.com/Fundacio-i2CAT/hostapd_prometheus_exporter

[10] https://grafana.com/

3.1.2 Ath10k evaluation

In this section a similar evaluation but based on 802.11ac technology and the ath10k driver is performed. There is a fundamental difference between this technology and 802.11n/ath9k and is that in this case aggregation is performed directly in the hardware, which results in mac80211 not obtaining reliable information about the airtime consumed by a particular transmission. This, of course, will impact the precision achieved by the airtime scheduler.

Another observation is that reducing the number of packets available in the hardware buffers had a direct impact on throughput, as it impacted the ability of the chipsets to build larger aggregates, which results in lower throughput. Therefore, airtime scheduling in 802.11ac/ath10k is implemented using a different strategy, called Adaptive Queue Limits (AQL). AQL maintains an estimate of the airtime required to transmit the packets buffer in the hardware buffers, for each destination, i.e., a pair of MAC address and TID. Based on this pending airtime estimate three configurable thresholds are defined:

- Low limit (default 5ms): If the estimate pending airtime inside the hardware queues for a particular MAC+TID is below this threshold packets are pushed down to the hardware disregarding any airtime-based scheduling. The reason for this is that a minimum of number of packets should be always available to the hardware buffers, otherwise aggregation is not used and channel capacity decreases.
- High limit (default 12ms): If pending airtime inside the hardware queues for a particular MAC+TID is above this threshold, packets are held in the mac80211 buffers.
- Interface Threshold (default 24ms): If the total pending airtime for a given interface, considering all MAC+TID pairs, is above this threshold, packets are kept in the mac80211 buffers.

In addition to the previous AQL algorithm, mac80211 still maintains an airtime-based scheduler that applies to the packets that are kept in the mac80211 queues, which are dequeued in the same way as ath9k.

Figure 3-4 depicts the results of an experiment where we analyse the ability to enforce an airtime allocation policy using the ath10k scheduling implementation. The same setup described in Section 3.1.1 is used. Starting from the left, we start evaluating the AQL algorithm with the default values for the AQL parameters. With this configuration, we observe that the ath10k is not able to enforce any airtime allocation policy. The reason is that the value of “Low Limit” set to 5ms allows always to push packets down to the hardware and so no airtime policy can be enforced in mac80211.

Figure 3-4. Evaluating airtime scheduling with 802.11ac/ath10k 80MHz channel

It is observed in Figure 3-4 that under this configuration ('default AQL') both stations in each slice achieve around 300 Mbps[11]. Given that the default AQL does not allow to enforce airtime policies, we change the default AQL queue limits to "Low Limit=0", "High Limit=Threshold=12ms". This means that once more than 12ms are available in the hardware queues, mac80211 starts to buffer packets and enforce the airtime scheduling policy. The results applying this configuration are depicted in Figure 3-4 under the 'modified AQL' area. The first thing we observe in this case, is that when active in isolation the station from slice 1 can achieve a lower throughput than in the default AQL configuration, i.e., slightly above 200 Mbps between 12:21:10 and 12:21:20. The reason is that the hardware queues can maintain only up to 12ms of data, which if a slow MCS is used, result in fewer packets and hence less efficient AMPDU aggregates. On the other hand, we can observe how in this case when the two slices are active concurrently, starting from 12:21:30, airtime policies can be enforced. We can see how starting from 12:21:30 the airtime available to each slice responds to the configured airtime weights. However, while following the trend set by the airtime weights, the net allocations are not as clean as they were with ath9k. The reason for this is two-fold. First, the 12ms of buffering available in the hardware queues scape the control of the airtime policy. Second, ath10k needs to maintain an estimate of the actual airtime allocation due to missing information reported by the driver.

We conclude that airtime-based allocation as required by the 5G-CLARITY slicing concept can also be applied for 802.11ac/ath10k, but this is less efficient than in the case of 802.11n/ath9k, due to the current driver implementation limitations. We note though that the shortcomings available in 802.11ac/ath10k could be compensated by the ML algorithms developed in WP4 that could be trained to discover a non-linear relation between the weights configured in the APs and the resulting slice shares.

3.2 LiFi airtime scheduling

The LiFi channel characteristics is different from a typical radio frequency (RF) channel. Stationary LiFi devices generally experience constant throughput, while the UE with mobility and a random orientation would possibly experience drops in its throughputs. Thus, it is a challenge to achieve an effective and fair slicing of LiFi resources, independent of capacity variability and availability of LiFi AP resources [13]. Considering the hardware architecture of the pureLiFi's LiFi-XC a uniformly assigned average channel-time scheduling algorithm is used to allocate users. As depicted in Figure 3-5, a counter is associated with each user and during packet transmission to a user, central processing unit (CPU) ticks are accumulated with the user-specific counter.

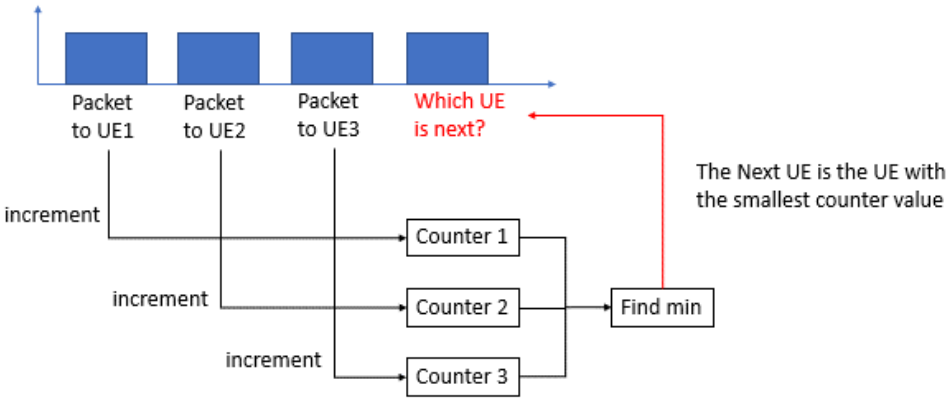


Figure 3-5: LiFi-XC scheduling process

[11] Slice 1 in the upper plot has a utilization above 100% due to a misconfiguration in the Grafana dashboard

When there is a new packet arrival, the counter values are compared and the user with the minimum counter value is granted for transmission.

In 5G-CLARITY, the virtualization scheme proposed in [13] will also be studied. The proposed scheme enforces dynamic sharing of LiFi DL and UL resource slices among service providers and their respective user groups. It runs based on simple abstractions for the MAC and physical resources of LiFi APs to reliably support an arbitrary number of slices. An airtime control abstraction function selects a specific physical transmission rate for each supported slice, while guaranteeing a minimum service rate for active user devices. The LiFi airtime scheduler will be working with the utility-based scheduler, which employs a utility function that maximizes the system utility regarding throughput, delay, or other pre-defined criteria. More detail of the utility-based scheduler is presented in Section 3.3. An example of the implemented scheduler in simulation is given below:

i) Each UE generates a data flow with a number of packets; each packet has a random packet size,

ii) By using the utility scheduler, a utility function, that better characterize the packet sizes or better suits the use case requirement, is selected,

iii) Select the UE with packet length that maximizes the system utility regarding the selected utility function.

Initial simulations were carried out with the above presented scheduling utility functions. The first considered function is the one used in LiFi-XC scheduler, which is based on uniform average channel-time across UEs. Another utility function being considered is to maximise packet size that in the queue from different users.

In the simulation, 3 UEs are considered while for each UE there are 100 data packets with random sizes to be transmitted. The scheduling, buffering and transmission processes for all packets of each UE is called a 'Task'. Figure 3-6 shows the task completion rate for 3 UEs regarding: 1) ratio of transmitted packets to full packet number – 100, in this case; and 2) ratio of transmitted data size to full task data size. It is shown that with the equal-time based scheduler, users are scheduled fairly regarding the airtime. In case of maximum packet size policy, the user with larger data packets would be selected earlier, which makes the task to be completed sooner. However, this also means that the user with heavy workload is more likely to consume the whole resources for longer. This can also be observed in Figure 3-7 where the UE allocation is shown.

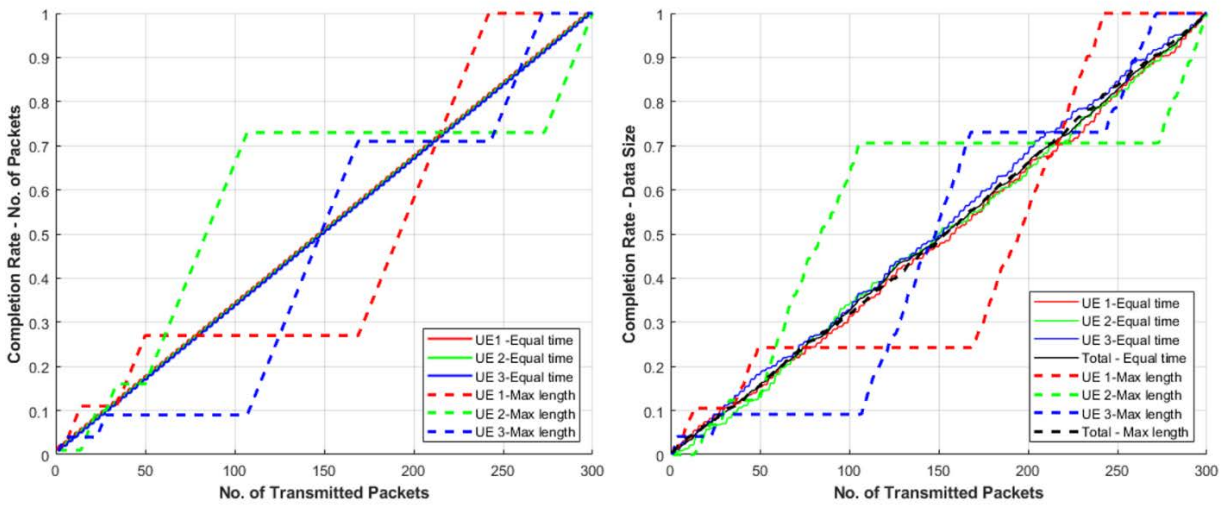


Figure 3-6: Task completion rate for equal time-based and max packet size-based scheduler

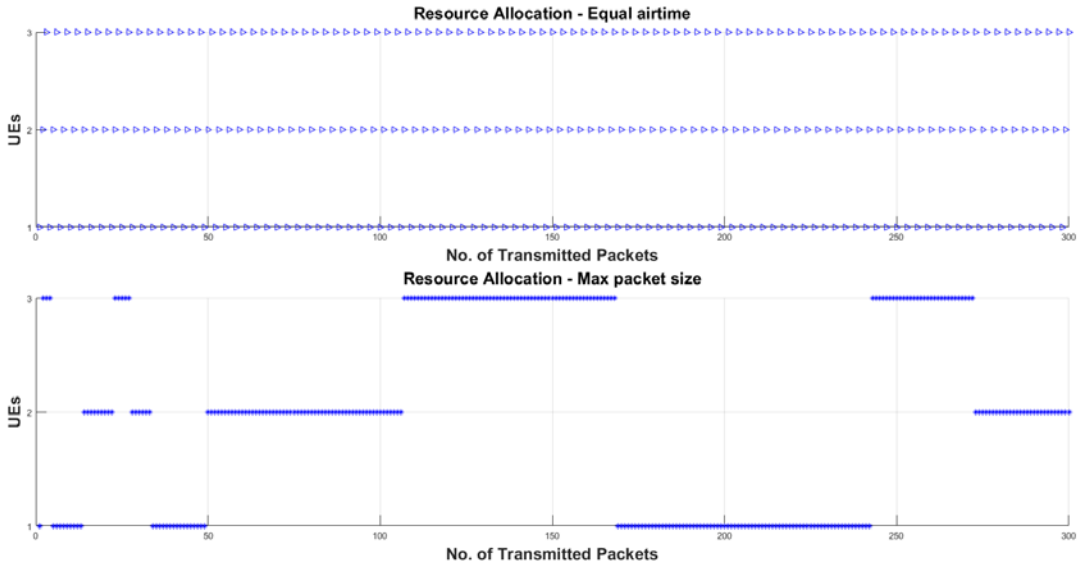


Figure 3-7: Resource allocation for equal-time based (top) and max packet-size based (bottom) scheduler

Other utility-functions could be defined as well, such as for optimising transmission latency, the whole throughput, etc. More details on these utility functions are presented in Section 3.3.

3.3 Utility-based scheduler

A utility-based scheduler handles users according to their network resource utility. Users broadly receive delay and throughout-sensitive services from their mobile virtual network operators (MVNOs). A LiFi attocellular network, shown in Figure 3-8 , is dynamically shared among three MVNOs to provide delay and throughput-sensitive services to users. A LiFi network slice is dynamically instantiated for MVNOs by allocating a resource quota in each LiFi AP attocell.

Sigmoidal delay and throughput utility functions are introduced to maximize the total sliced LiFi network utility with respect to users' throughput and the head-of-line (HoL) packet delays. The optimal throughput and delay values are mapped to resource utility values by using the delay and throughput utility functions. Then, users are scheduled according to their resource utility values. Alternatively, a utility weight is calculated for each user per resource based on marginal delay and throughput utility functions, which are derived as the first order derivative of the proposed utility functions. Users are scheduled according to priority wights which are calculated as the multiplication of their resources' weights and instantaneous data rates. Resource utility policies-based heuristic scheduler algorithm is introduced as a practical utility-based scheduler to schedule users to the resource quotas of the different LiFi network slices. This is explained, including the practical considerations, in Section 3.3.2.

3.3.1 Description

As part of slicing 5GNR/Wi-Fi/LiFi integrated network resources, this research work focuses on slicing LiFi attocellular network resources. LiFi AP DL channel resources are sliced using a utility-based MAC scheduler. MVNOs provide plethora of services to end-users by using different network slices. This work considers some main service types envisioned for 5G, namely: URLLC, mMTC, and eMBB, as shown in Figure 3-8.

Network slices are assigned resource quotas in LiFi AP attocells to meet their resource guarantees or data rate targets [13]. Resource scheduling among the users subscribed to the different MVNOs can be based on different criteria such as: traffic load, data rate, bandwidth, power, interference, pre-defined contract,

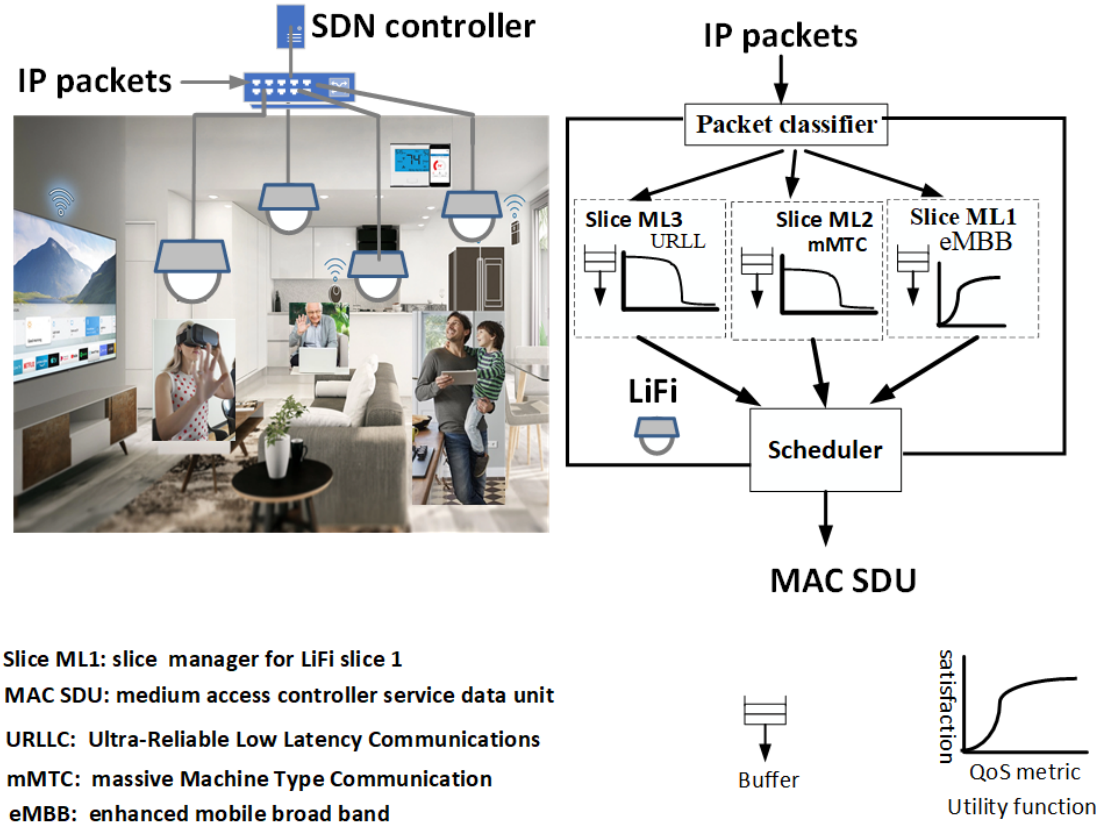


Figure 3-8 : SDN-enabled multi-tenant LiFi attocellular network.

channel condition or a combination of these. The challenge here is in ensuring the resource allocation to meet the different network slices constraints, while meeting the QoS requirements of MVNOs and their users. Note that the scheduler performance must not be measured in terms of network-centric quantities such as delay, packet loss, utilization, or energy (power) efficiency, but rather should be evaluated solely in terms of the degree to which the network satisfies the service requirements of user applications.

A utility-based scheduler employs a utility function, $U(.)$, that maps the QoS metrics (e.g., delay and throughput) to resource utility values. These are used by the scheduler to schedule users to network resources, considering different objectives such as the total network utility maximization.

A network slice can be abstracted by a virtual buffer in each LiFi attocell AP, where it is instantiated. It accommodates the traffic packets generated from the service flows destined to the users subscribed to the MVNO having the network slice [14]. It is assumed that each MVNO has a single network slice. A slice buffer is managed according to the first input-first output (FIFO) policy, though the slice manager can change this policy to support network slice resource customization and service differentiation.

The set of network performance metrics, such as average user throughput and minimum latency is denoted by M, where for each metric $m \in M$ a generic utility function is modelled as follows:

$$U_s\left(f(x_s), x_{s_{req}}\right), \quad s \in S$$

where $x_{s_{req}}$ is a requirement of slice $s \in S$ for metric m; and $f(x_s)$ is a function of the total value of metric m resulting from the allocated resources to all users receiving services from the slice s. For example, in a time-slotted system model, x_s can be expressed for the total throughput of slice s, as follows:

$$\boldsymbol{x_s = \sum_{t \in T} \sum_{k \in K_s} \gamma_{sk}[t]} \qquad (1)$$

where $\gamma_{sk}[t]$ denotes the maximum achievable rate of user k subscribed to MVNO (i.e., slice) s at time slot. Note that the parameter s is used interchangeably to denote for a MVNO and their slice.

It is assumed that a perfect channel knowledge is available at the scheduler. It is also assumed that the eMBB and mMTC services are more sensitive to throughput, whereas the URLLC service is sensitive to delay (latency). Two policies for dynamic resource allocation to network slices are proposed to maximize the total sliced LiFi attocellular network utility with respect to the throughput and delay of supported services, i.e., throughput-based network maximization (TNM) and delay-based network maximization (DNM) policies. The main objective is to maximize the network utility over the LiFi AP DL channel resources while meeting the target minimum and maximum rates, and resource guarantee for MVNOs in each attocell.

The LiFi AP DL channel is consisted of N sub-carriers, each with bandwidth b Hz. The status of network resources in each attocell is updated every T subframe, where there are TN resources available in each LiFi attocell. The achievable data rate of user k subscribed to MVNO s, on sub-carrier n, in attocell a, at time slot, t according to a binary variable $z_{sk}^{[a,t,n]} \in \{0,1\}$, is denoted as $\gamma_{sk}^{[a,t,n]}$, which is given by Shannon theorem as follows:

$$\boldsymbol{\gamma_{sk}^{[a,t,n]} = \frac{b}{2T}\log_2\left(1 + \mathrm{SINR}_{sk}^{[a,t,n]}\right)} \qquad (2)$$

where SINR denotes the signal-to-interference-plus-noise ratio of user k subscribed to MVNO s. A resource n is allocated to a single user k in an attocell a at a time slot t, according to a binary variable $z_{sk}^{[a,t,n]} \in \{0,1\}$. This is equal 1 if resource $[a,t,n]$ is allocated to user k, otherwise, it is equal 0. The objective of maximizing the network resources utility is to find a scheduling policy that maximizes the total sliced network utility, given by

$$\boldsymbol{\sum_{a\in A}\sum_{s\in S}\omega_s^a U_s^a\left(x_s^{[a,t,n]}\right), \qquad \forall t,n} \qquad (3)$$

where the utility function is non-decreasing and concave to balance the efficiency and fairness of the resource allocation.

$$\boldsymbol{x_s^{[a,t,n]} = \sum_{k\in\mathcal{K}_s^a}\sum_{t=1}^{T}\sum_{n=1}^{N} z_{sk}^{[a,t,n]}\gamma_{sk}^{[a,t,n]}} \qquad (4)$$

where $\mathcal{K}_s^a$ denotes the set of users associated to AP a and subscribed to MVNO s. Note that, ω_s^a determines the values of slice-specific parameters in each attocell a, which enables the MVNOs to weight their utility delay and throughput components. This enables dynamic resource scheduling mechanisms to provide service differentiation and prioritization through setting the slice-specific parameters according to the network slices tenants.

The optimization objective, described in the Equation (5) below, is to find the optimal resources z and their assigned power p while maximizing the total sliced LiFi network utility. However, as it is explained in the next section 3.3.2, the solution of this optimization problem focuses on finding the optimal resources allocation (data rates) to each MVNO for a given fixed equal power allocation to subcarriers. To this end, the decision variable of power allocation, p, is removed.

$$\boldsymbol{\max_z \sum_{a\in A}\sum_{s\in S}\omega_s^a U_s^a\left(x_s^{[a,t,n]}\right)} \qquad (5)$$

s.t.

$$\sum_{s=1}^{S}\sum_{k\in\mathcal{K}_s^a} z_{sk}^{a,t,n} \le 1, \quad \forall a,t,n \tag{6}$$

$$-\sum_{k\in\mathcal{K}_s^a}\sum_{t=1}^{T}\sum_{n=1}^{N} z_{sk}^{[a,t,n]}\gamma_{sk}^{[a,t,n]} \le -\gamma_{s,\min}^{a} \quad \forall a,s \tag{7}$$

$$\sum_{k\in\mathcal{K}_s^a}\sum_{t=1}^{T}\sum_{n=1}^{N} z_{sk}^{[a,t,n]}\gamma_{sk}^{[a,t,n]} \le \gamma_{s,\max}^{a} \quad \forall a,s \tag{8}$$

$$z_{sk}^{a,t,n} \in \{0,1\}, \quad \forall a,t,n,k,s \tag{9}$$

$$p_{a,t,n} \ge 0, \quad \forall n \tag{10}$$

The optimization objective Eq. (5) maximizes the network aggregate utility of all MVNOs. Constraint Eq. (6) restricts user association to a single attocell and subscription to a single MVNO per time t. Constraint Eq. (7) and constraint Eq. (8) impose lower and upper bounds on the total data rate achieved by MVNO in each attocell, respectively. Constraint Eq. (9) requires a unique resource allocation to only one user subscribed to MVNO in each attocell. Constraint Eq. (10) requires that the power allocation to each subcarrier should not be negative in each attocell.

The optimization problem Eq. (5-10) is an NP-hard problem, as it is a mixed integer non-linear optimization problem. Furthermore, it cannot be optimally solved within a reasonable time window T, which can be around 100 ms for LiFi systems. For practical considerations, this problem will be solved heuristically to obtain the resource allocation to users for a given power allocation to subcarriers, discussed in Section 3.3.2.

3.3.2 Practical considerations

The utility-based scheduler mechanism is dictated by two main factors, i.e., the utility functions defined for the different slice types, and how the tenants weigh them together by setting ω_s^a. Sigmoidal utility functions, $U_s^a\left(x_s^{[a,t,n]}\right)$, are proposed to map the QoS metrics, delay, and throughput, to network utility values, as shown in Figure 3-9. These utility functions are used in the formulated optimization problems for TNM and DNM policies to optimally allocate resources to throughput and delay sensitive users, respectively.

$$U_s^a\left(x_s^{[a,t,n]}\right) = \frac{e^{v\phi\left(x_s^{[a,t,n]} - x_{s_{req}}^{[a,t,n]}\right)}}{1+e^{v\phi\left(x_s^{[a,t,n]} - x_{s_{req}}^{[a,t,n]}\right)}}, \quad \forall s,a \tag{11}$$

where

$$\phi = \frac{\log\left(\frac{1-\xi}{\xi}\right)}{\eta x_{s_{req}}^{[a,t,n]}}$$

where $x_s^{[a,t,n]}$ represents the measured service performance metric, and $x_{s_{req}}^{[a,t,n]}$ denotes the QoS requirement of network slice s (service). v is set to 1 and -1 to account for the increasing throughput and decreasing delay utility functions, respectively, as shown in Figure 3-9. ϕ is a non-negative parameter that determines the function slope. ξ and η are constants which are set to 0.01, 0.5 [15], respectively. The utility-based weight marginal function is derived to calculate the priority weight of user $k \in \mathcal{K}_s^a$, ω_{sk}^a, as follows:

$$\omega_{sk}^{a,m} = \frac{\partial U_s^a}{\partial x_s^{[a,t,n]}} = \frac{\nu\phi e^{\nu\phi\left(x_s^{[a,t,n]} - x_{s_{req}}^{[a,t,n]}\right)}}{1+e^{\nu\phi\left(x_s^{[a,t,n]} - x_{s_{req}}^{[a,t,n]}\right)}} - \frac{\nu\phi e^{2\nu\phi\left(x_s^{[a,t,n]} - x_{s_{req}}^{[a,t,n]}\right)}}{\left(1+e^{\nu\phi\left(x_s^{[a,t,n]} - x_{s\backslash textrmreq}^{[a,t,n]}\right)}\right)^2} \quad (12)$$

The mMTC service generates small size packets that require a minimum-bit rate to support communications among many devices. The eMBB service generates large size packets that require a high throughput connection. Both mMTC and eMBB are sensitive to throughput but less to delay. The URLLC service generates a low rate of small size packets but requires stringent latency requirements and increasing throughput. Figure 3-9 shows the utility functions of the delay and throughput performance metrics of URLLC, mMTC, and eMBB services.

The aggregate throughput and average delay of service running on a slice s are measured in Mbps and ms, respectively, as shown in Figure 3-9. The increase in the throughput of users associated to MVNO, drives higher the aggregate network utility of their MVNO. However, the increase in the delay of services running on the slice s, drives lower the aggregate network utility of their MVNO s. A MVNO becomes satisfied, when their measured aggregate throughput approaches and exceeds the throughput requirement in their service level agreement (SLA). Based on the utility functions, the optimization problem aims to maximize the total network utility, as it gives a higher priority to users with higher utility in the resource allocation process. However, based on the marginal utility functions, the optimization problem aims to maximize the MVNOs satisfaction, as it gives a higher priority to users with lower utility in the resource allocation process, as shown in Figure 3-9.

A heuristic utility-based subcarrier assignment algorithm is proposed to solve the optimization problem Eq. (5 –10) for obtaining an optimal resource allocation to users, assuming equal power allocation to subcarriers. It is simple non-iterative heuristic algorithm, which follows a greedy approach in the allocation of subcarriers to the different users of MVNOs. MVNOs are allocated the subcarriers offering the best network utility until their data rate targets are satisfied. The resource allocation process is parameterized with four parameters, i.e. slice utility weight, user utility weight, minimum and maximum data rate targets of MVNOs. The user utility weight can be calculated based on Eq. (12) and Eq. (13), which manages the allocation of sub-carriers to the different users of MVNOs based on maximizing the total network slices utility or users' satisfaction. The excess resources are allocated to MVNOs offering the highest data rate-network utility, as explained in Algorithm shown in Figure 3-10.

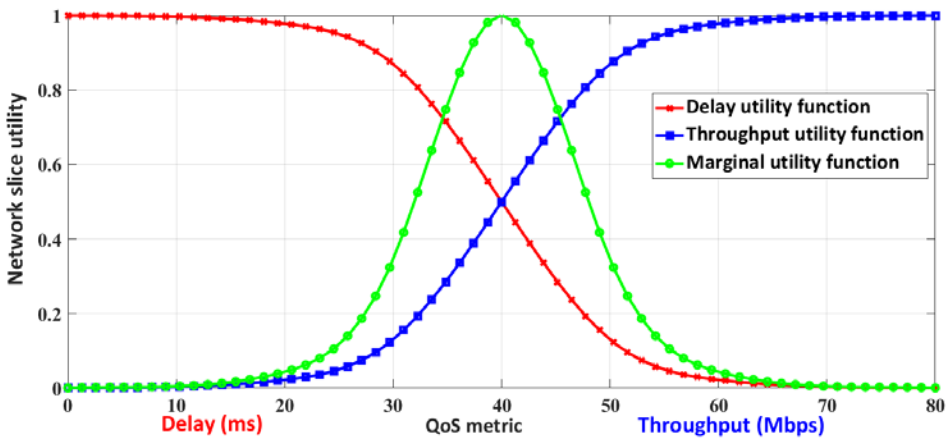


Figure 3-9: Throughput, delay, and marginal utility functions.

Initialize:
$\mathbf{p}_{a,t,n} = \frac{P_{a,max}}{N}, \quad \forall a, t, n$
$\gamma_s^a = 0, \quad \forall s, a$: this sets the data rate of MVNOs to zero in attocell a
$U_a^s = \{1, 2, \cdots, \mathcal{S}\}$: Set of unsatisfied MVNOs in attocell a
Output: 1.
for $a \leftarrow 1$ *to* A **do**
  for $t \leftarrow 1$ *to* T **do**
    for $n \leftarrow 1$ *to* N **do**
      Solve the optimization problem:
      if $\gamma_s^a \leq \gamma_{s,\max}^a$ **then**
        $k_{a,s} \leftarrow \arg\max_{k \in \mathcal{K}_s^a} \omega_s \gamma_{sk}^{a,t,n}$
        if $\gamma_s^a \geq \gamma_{s,\min}^a$ **then**
          $U_a^s = U_a^s - \{s\}$
        else
          $k_{a,s} \leftarrow [a, t, n]$
          $\gamma_s^a = \gamma_s^a + \gamma_{sk}^{a,t,n}$
      else
        $U_a^s = U_a^s - \{s\}$

Figure 3-10: Utility-based subcarrier assignment algorithm

A discrete time simulation environment was developed in MATLAB to evaluate the performance of the proposed heuristic utility-based resource scheduling policies algorithm, shown in Figure 3-10, in the sliced LiFi attocellular network shown in Figure 3-8. This is composed of four LiFi APs, which is shared dynamically among three MVNOs offering URRLC, mMTC, and eMBB services. It provides services to users uniformly distributed in a room of size 5 x 5 x 3 (L x W x H). The network simulation parameters are summarized in Table 3-1. The users remain active but stationary during the simulation time. The traffic packets of URLLC and mMTC services are generated with data rates following a Poisson distribution with λ taken to be 64 kbps - 0.4 Mbps. Whereas the traffic packets of eMBB services are generated following a Pareto distribution with a scale parameter taken to be 0.5 - 5 Mbps.

The simulation time is divided into 80 sub-windows, each one of length equal to 100 time slots. The scheduler schedules users every time slot of length 1 ms. However, the scheduling policies are evaluated every sub-window, to check as it would satisfy SLA requirements of MVNOs. The slice-specific parameter ω_s^a accounts for the throughput utility weight ω_s^{th} and the delay utility weight ω_s^d, such as $\omega_s^a = \{\omega_s^{th}, \omega_s^d\}$. The service guarantees of URLLC, mMTC and eMBB are represented in a tuple [x x x] that expresses the percentage of their maximum data rate or resources to be guaranteed during a simulation sub-window time. For example, a service tuple [1 1 1] indicates that the URLLC, mMTC, eMBB services should receive fully their maximum data rate or resource guarantees. Whereas a service tuple [0.1 0.2 0.5] indicates that the URLLC, mMTC, eMBB services should receive a percentage of their data rate or resource guarantees during the simulation sub-window time. These are set (programmed) by the slice-specific throughput and delay utility weight parameters.

A number of scenarios are conducted to study the impact of changing the slice-throughput utility weight parameter ω_s^{th} on the aggregate throughput of eMBB and mMTC services, while fixing the slice-delay utility weight parameter ω_s^d, as shown in Figure 3-11 and Figure 3-12.

Table 3-1 LiFi Attocellular Network Simulation Parameters

Parameter	Description	Value
$\Phi_{1/2}$	LED half power semi angle	60°
η_w	Reflectivity factor of wall	0.8
g_f	Gain of optical filter	1
A_R	Physical area of photo detector (PD)	$10\ mm^2$
η_f	Reflectivity factor of floor	0.8
N_0	Noise power spectral density	$10^{-21}\ A^2/Hz$
B	Modulated bandwidth	20 MHz
η_c	Reflectivity factor of ceiling	0.8
Ψ_c	Receiver field of view (FOV)	90°
P_t	Transmission power	9 W
ζ	Refractive index	1.5
n_r	PD's orientation vector	[0, 0, 1]
R_{PD}	PD responsivity	0.5 A/W
n_t	AP's orientation vector	[0, 0, -1]
$\lvert\mathcal{N}\rvert$	DCO-OFDM sub-carriers	512
$\lvert\mathcal{A}\rvert$	Number of APs	4
$\lvert\mathcal{S}\rvert$	Number of slices (MVNOs)	3

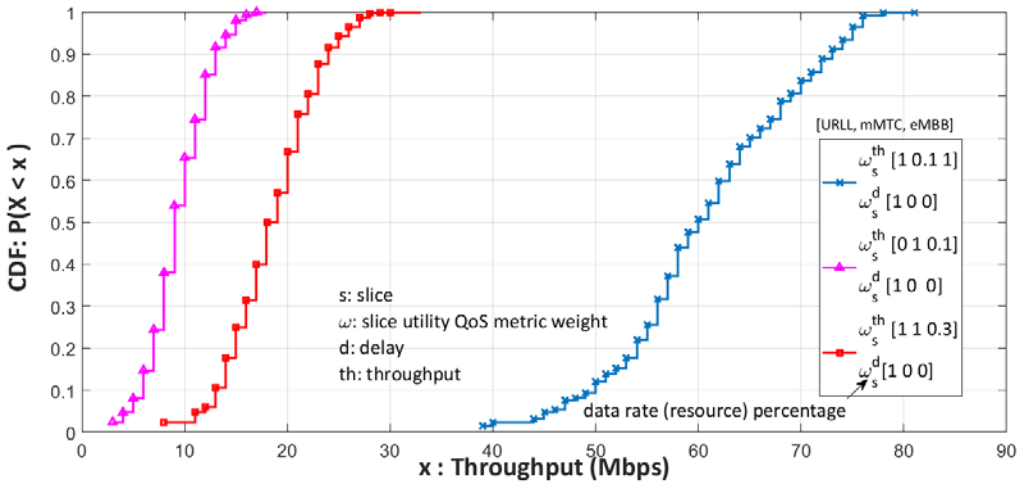


Figure 3-11: CDF of network slice aggregate throughput providing eMBB service

The URLLC service slice ω_s^d is set to 1, which makes the scheduler assigns more bandwidth to meet its delay requirements. In the meanwhile, the scheduler balances the allocation of the available resources among the eMBB and mMTC slices, prioritizing the slice with better network utility. Note that the scheduler guarantees a minimum data rates to all slices, resulting in a minimum network utility value. The performance of the eMBB slice does not experience significant degradation due to changes in the ω_s^{th} of the mMTC, as shown in Figure 3-11. The eMBB users get higher priority than the mMTC users when the minimum data rate of mMTC users is guaranteed.

Therefore, only when the mMTC slice requires higher priority than the eMBB slice, as shown Figure 3-12, the mMTC users can achieve higher throughput. The channel conditions of users do not affect the scheduling priority of delay-sensitive services. Also, for throughput sensitive slices, the increase in utility can be caused by good channel conditions of the users, where the utility values can also be above the minimum utility value, as shown in Figure 3-12.

It is observed that the tuning of ω_s^{th} in the utility function allows the competing eMBB and mMTC slices to get higher number of resources. On the other side, however, the competition among these slices does not affect the resource distribution for the delay sensitive URLLC service.

Three scenarios are conducted to study the impact of changing the slice-delay weight parameter ω_s^d on the aggregate average delay utility of the different network slices, while fixing their slice-throughput utility weight parameter ω_s^{th}, as shown in Figure 3-13. It is observed that when MVNOs decrease the value of their slice delay utility weight ω_s^d from 1 to 0.5, 0.2, 0.1, their network utility values also decrease.

This conforms with the average delay utility function, when the average delay of slice increases, the slice network utility decreases. To this end, users need to be scheduled immediately as soon as their packets arrive at the slice buffer, irrespective of the channel conditions. These packets are assigned with higher priority and therefore they are allocated more resources to be delivered on time. The higher ω_s^d is, the higher the percentage of delay rate or resources are required to guarantee the delay requirement. So, the URLLC service requires a higher throughput (bandwidth) to meet its rigorous delay requirement, which may be already used by mMTC and eMBB services.

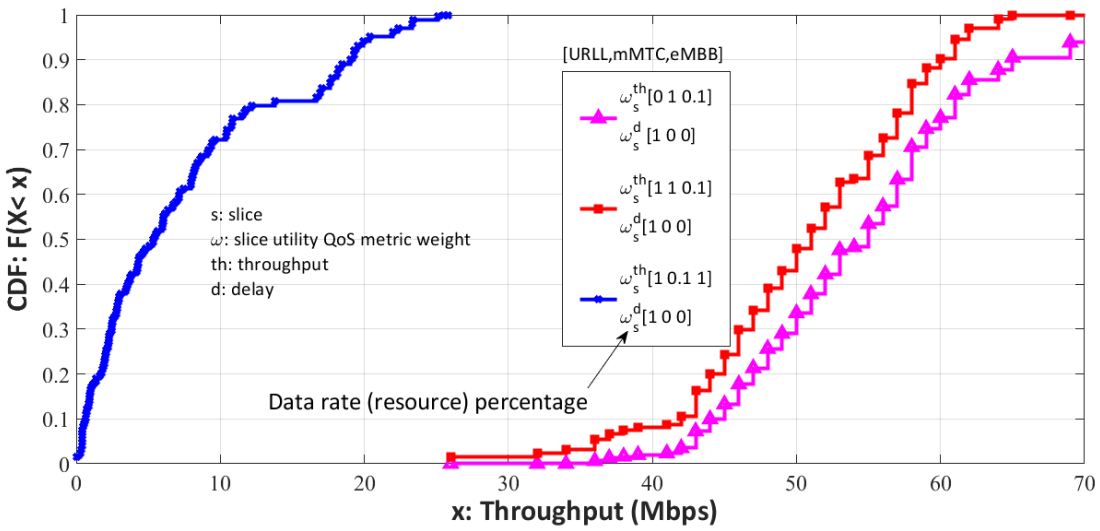


Figure 3-12: CDF of network slice aggregate throughput providing mMTC service.

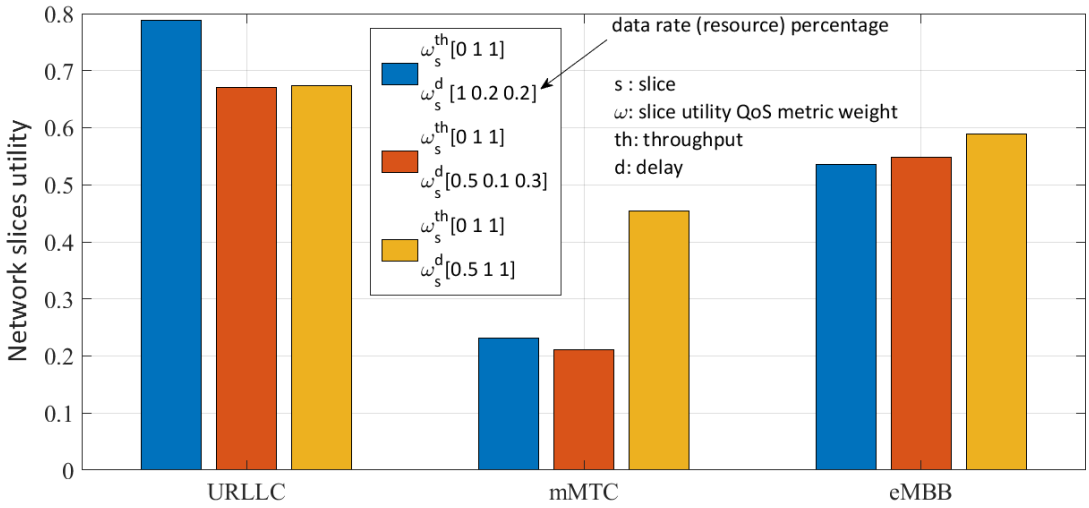


Figure 3-13 Aggregate average delay utilities of network slices under varying their delay weight parameters

4 Positioning

In 5G-CLARITY, different WATs are used to perform positioning of the UEs. Having multiple different WATs for positioning brings new advantages including better coverage and better signal quality, while at the same time improving positioning precision and accuracy. Additionally, all of these WATs can be used simultaneously, which would enable different data fusion approaches for further improvement of the positioning precision.

There are four different WATs proposed for positioning within the 5G-CLARITY. Namely, these technologies include 60 GHz mmWave, a sub-6 GHz positioning system, LiFi and OCC [1]. The 60 GHz mmWave technology uses devices developed at IHP [16]. The main advantage with this technology is the large channel bandwidth (approximately 2 GHz) which enables centimetre level positioning precision. The sub-6 GHz positioning technology is a proprietary technology developed at IHP. This technology is implemented on a SDR and all of the positioning algorithms are implemented in software. The LiFi technology is supplied by pureLiFi and it is used for data communication as well as for positioning. Finally, the OCC, or Visible Light Positioning (VLP), positioning technology is developed by I2CAT.

In order to merge all of these technologies, a localization server, developed by IHP is used. This localization server connects to all of the WATs in order to obtain the position of UEs and provides a unified interface, independent of the underlying technology, towards the entities requiring positioning services. It also implements different algorithms including data fusion algorithms. More details can be found in [1].

In 5G-CLARITY, mainly range based methods are used for positioning. These methods include measuring distances between the AP and known positions, commonly called anchor points, and between the AP and UE. The distances can be estimated by measuring the time-of-flight (ToF) of the radio wave to travel a given distance, or by measuring the received signal strength and assuming a known propagation loss.

In order to estimate the position of the UE in range-based methods, trilateration is commonly used. In trilateration, at least three APs with known locations in the system are used in order to perform 2-dimensional (2D) localization. Once the UE has estimated the distance from the corresponding APs, circles with the radii of the calculated distance can be drawn, as shown in Figure 4-1(a). Given the coordinates of the anchor points, i.e., (x_1, y_1), (x_2, y_2), and (x_3, y_3), the process can be described as a system of equations, as follows [17]:

$$\begin{cases} (x - x_1)^2 + (y - y_1)^2 = d_1^2 \\ (x - x_2)^2 + (y - y_2)^2 = d_2^2 \\ (x - x_3)^2 + (y - y_3)^2 = d_3^2 \end{cases}$$

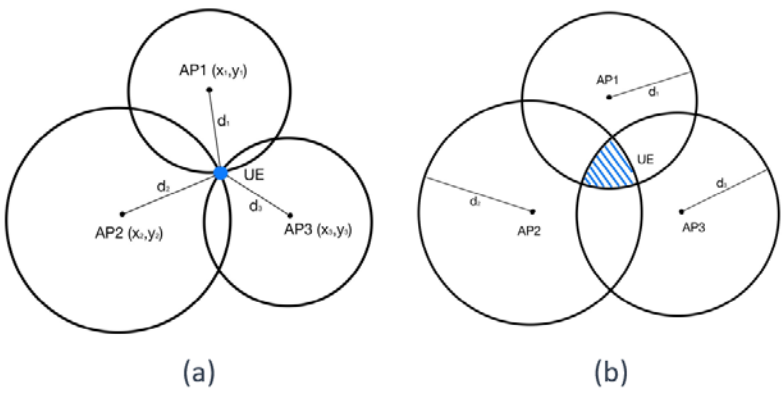


Figure 4-1: Trilateration, (a) intersect in one point, (b) intersect within an area.

Considering the quantization and measurement errors in most cases the drawn circles will appear like the ones shown in Figure 4-1(b).

This means that the previous system of equations cannot be solved, since the three circles are not intersecting in a single point. In order to find the position of the UE in this case, usually a least squares method is used. This would provide a solution having the minimum-square-error with respect of the measured distances. Since the system of equations is nonlinear, a non-linear least squares (NLLSQ) method is used. In order to solve the NLLSQ, numerical methods as Gauss-Newton, Levenberg–Marquardt algorithm, QR decomposition, etc. are usually used.

The architecture of the localization system including in the localization server as well as its functionalities is given in 5G-CLARITY D3.1 [1].

4.1 60 GHz mm-Wave

The 60 GHz ISM band offers channels with bandwidths of 2 GHz. The large channel bandwidth is crucial for very precise ranging, and, therefore, very precise localization. Hence, 5G-CLARITY has chosen this band to enable centimeter level precision localization.

In order to use the 60 GHz band, a custom modem supporting direct access to the baseband signal are needed. For this purpose, modems described in [16] are used. The 60 GHz mmWave modems use a COTS analog frontend with a phase antenna array supporting beam steering. The modems also include an FPGA which can be configured for different operation. Additionally, they can be used in a SDR mode, where the signal samples are processed on a general-purpose host computer. This mode is also used for positioning purposes in 5G-CLARITY.

The initial testing was performed in a laboratory. Only ranging was tested at this time. The ranging test scenario is shown in Figure 4-2. Two 60 GHz mmWave devices are connected on a general-purpose computer which control the ranging process. They are connected using the 1 Gbps Ethernet.

The ranging test are performed using two mmWave devices, positioned on a known distance and multiple distance estimations are performed. The true distance is measured with a high precision laser distance meter model BOSCH GLM 80 Professional. This distance between the nodes is increased gradually, by a given step, and for each distance, multiple distance estimations are performed. Multiple distance estimates are needed in order to calculate the sample variance of the estimated distances.

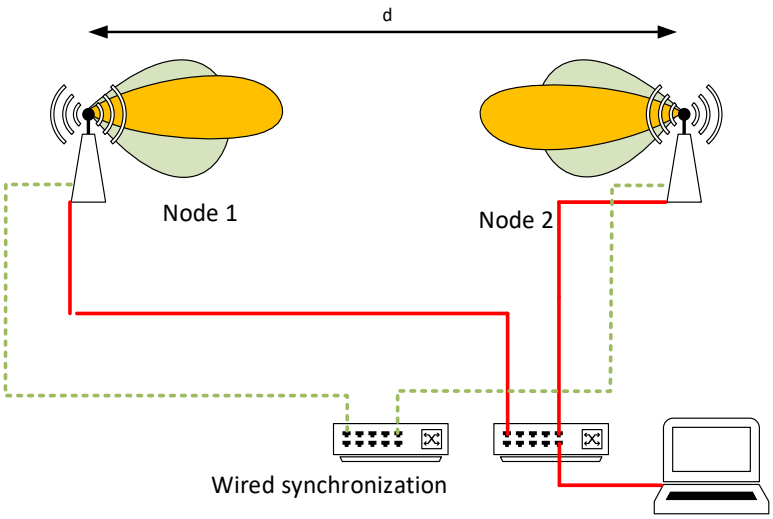


Figure 4-2: Ranging test scenario

The following KPIs are evaluated using this experiment:

- Sample variance, i.e., ranging precision as a function of the distance between the mmWave nodes
- Distance (range) estimation rate, i.e., number of distance estimations per second

In order to perform positioning using the mmWave nodes, at least 2 of them must be used as an anchor nodes with known positions. Since 2 anchor nodes are not sufficient for 3 dimensional (3D) or 2 dimensional (2D) localization in general case, the following constraints are introduced:

- The UE is assumed moving in 2D space on a floor which can be represented as a single plane. This assumption is aligned with realistic movement of the AGV in the BOSCH factory.
- The UE is always on the same side with respect to the line described with the 2 mmWave anchor points.

The localization scenario is shown in Figure 4-3. According to this scenario in order to perform positioning, the distance between each of the anchor nodes and the UE is measured and then the position is estimated in the intersection of the two circles described by the two measured distances (i.e., radii). The second constraint introduced previously, guarantees that only one of the intersections is realistic.

At this moment, only a ranging experiment was performed in laboratory conditions. The positioning experiments would be performed later within the 5G-CLARITY project.

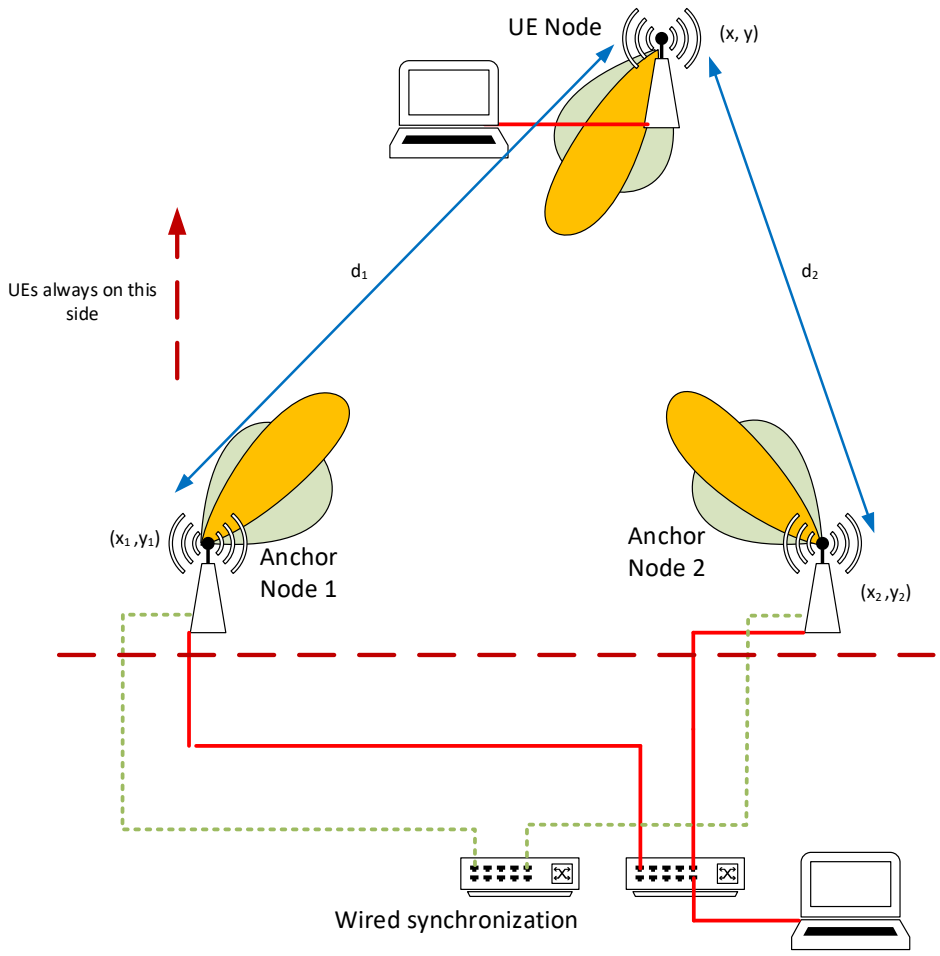


Figure 4-3: Positioning test architecture

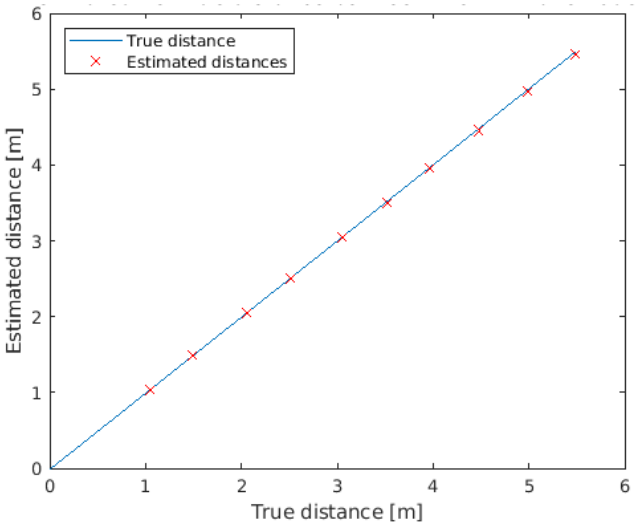


Figure 4-4: Comparison of the true and the distance estimated using the 60 GHz mmWave system

The experiment was performed by placing the 2 mmWave nodes facing each other and increasing the distance between them. The transmit power was kept at a constant level. The dynamic range of the receiver is sufficient for all of the distances used in this experiment. The distance between the nodes was changed starting from approximately 1 meter up to 5.5 meters using a step size of approximately 0.5 meters. The main limiting factor for using larger distances, was the synchronization cable, used between the nodes. At this moment it was not possible to use longer cable due to the driver circuitry used in the mmWave nodes. In the final UC2.2 demonstrations [3], a radio-based synchronization will be used, eliminating any need for synchronization cables.

In Figure 4-4, the estimated distances (red) are compared to the true distances. The blue line is the line which represents where the ideal estimates should be. As can be noted, the estimated distances are lying on this line.

In order to precisely quantify the distance estimation error, the root mean square (RMS) was calculated. In Figure 4-5, the RMS error of the estimated distances is given. As can be noted, the maximal RMS error is approximately 2.2 centimeters for a distance of 1.5 meters. For the rest of the distances the RMS error is below 1.5 centimeters. This can be further improved buy actively controlling the transmit power as well as using an automatic gain control (AGC) at the receiver, which was not the case in this experiment.

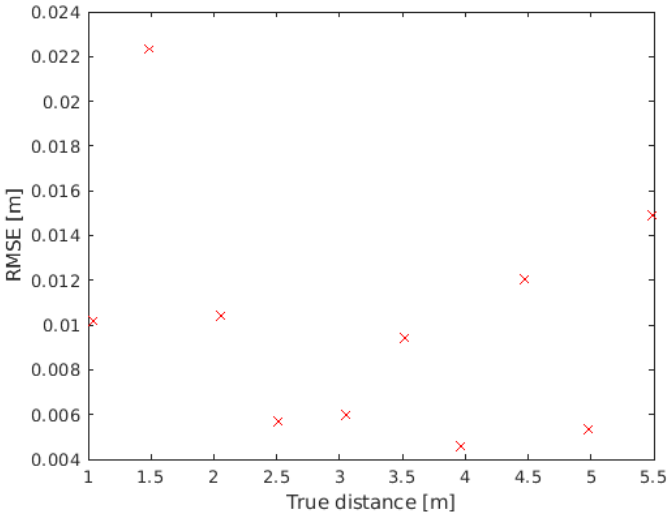


Figure 4-5: RMS error of the estimated distances

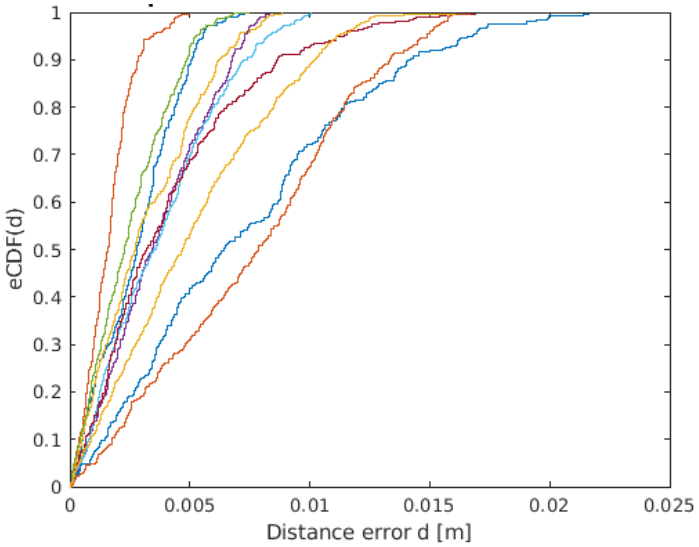


Figure 4-6: Empirical CDF functions of the distance estimates

Additionally, the empirical cumulative distribution function (CDF) was calculated for each of the distances in the experiment. In Figure 4-6, the empirical CDFs are given. It can be noted that for most of the distances, the CDFs show that the estimates lie below 1 centimeter.

4.2 Sub-6 GHz

The sub 6 GHz localization is performed using a SDR from Ettus research[12]. The model used for the initial testing is USRP N321. The channel bandwidth used is 160 MHz and the carrier frequency is 5.25 GHz. The radios can be used in all of the frequency bands starting from 100 MHz till 6 GHz. Nevertheless, for this experiment they were used in the 5 GHz ISM band since no additional licencing is required and it is assumed that this band is not heavily used by other devices. The ISM band in the 2.4 GHz can be also used, but with limited channel bandwidth, of maximum 80 MHz, and higher interference from the other devices (WiFi, Bluetooth etc.) transmitting in this band.

The transmit power was less than 15 dBm due to the limited output power of the used SDRs as well as the additionally introduced back-off in order to minimize nonlinear distortion introduced by the power amplifier.

For the localization in the sub 6 GHz bands, DL time-difference-of-arrival (DL-TDoA) is used. This requires that the anchor nodes are synchronized. Two possibilities for synchronization are available. One is to use 1 pulse per second (PPS) and 10 MHz reference signals or to use White Rabbit protocol for synchronization [18].

A minimum of 4 anchor nodes with known positions are needed for 2D positioning using this approach. With a minimum of 5 anchor nodes, a 3D positioning can be performed. In case more than the minimum number of anchor nodes installed, these can be used for improving the positioning precision. For the initial testing performed at IHP, 6 anchor nodes were used.

The test scenario for sub 6 GHz localization is shown in Figure 4-7. In the figure, 4 anchor nodes and one UE node are shown for the sake of simplicity, but in the real testing a total of 6 anchor nodes are used. The number of anchor nodes can be arbitrarily extended, depending on the size of the area that needs to be covered as well as the required precision to be achieved.

[12] https://www.ettus.com/

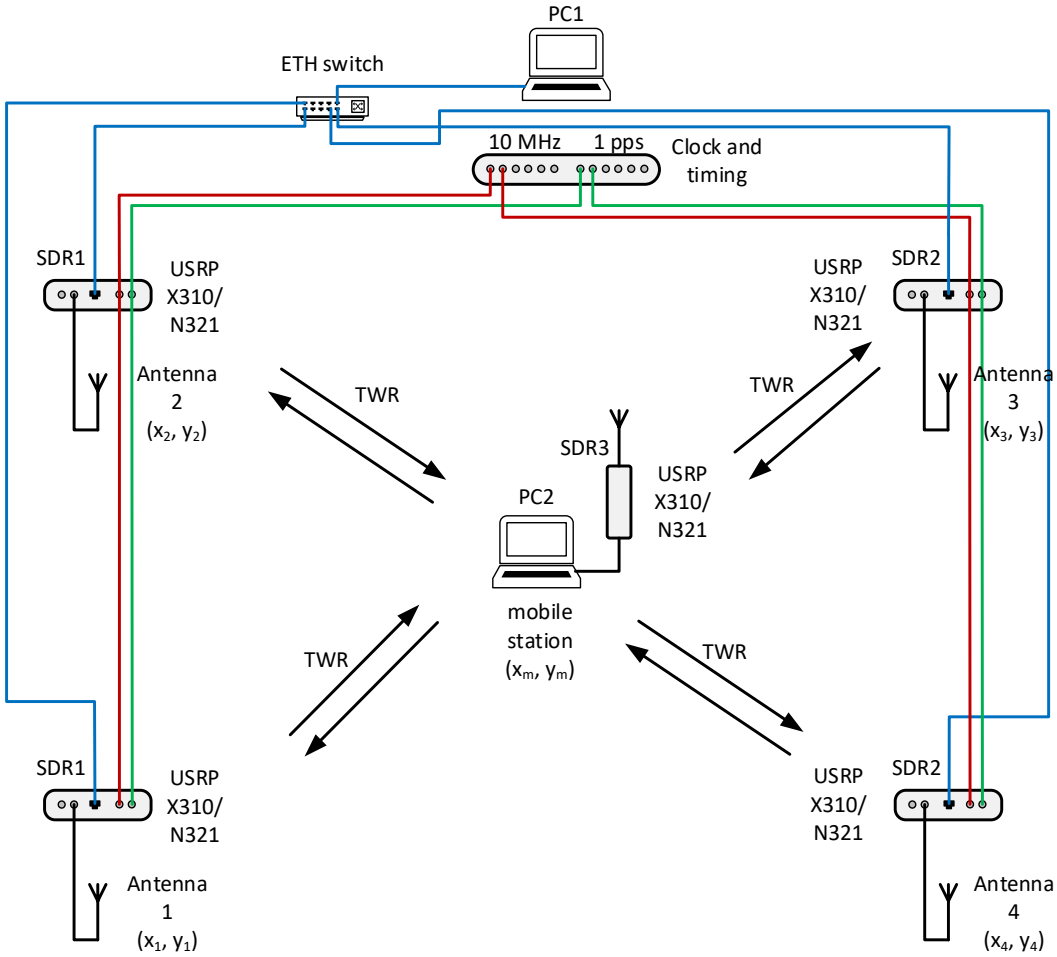


Figure 4-7: Sub 6 GHz positioning architecture

All of the nodes, anchors and UE, are SDRs. The anchor nodes are controlled from a single general purpose personal computer, running a C/C++ code which communicates with the SDRs, prepares the frames to be transmitted and schedules them for transmission using the SDRs. For DL-TDoA, the frames are sent using the Anchor nodes and are received by the UE node. These frames are further processed by the UE, and the position is estimated.

In this scenario only a 2D positioning is of interest and, therefore, it is assumed that the UE node is moving along a single plane and the height is fixed to 1.6 meters. The UE node is placed on a set of different known positions and for these positions, multiple position estimates are performed. The anchor nodes are also placed on the same height of 1.6 meters. The UE uses an omnidirectional antenna having a gain of 4.6 dBi, while the anchor nodes use patch antennas with an antenna gain of 7 dBi, pointed towards the area of interest.

All of the position estimates were performed while the UE was not moving. The following KPIs are evaluated with these tests:

- Sample variance, i.e., precision, of the obtained position estimates
- Position estimation rate

The evaluation is performed in a large room at IHP which floorplan is shown in Figure 4-8. The positions of the anchor nodes are shown with red circles and the coordinate system origin is shown with green arrows. The UE was placed in total of 8 positions.

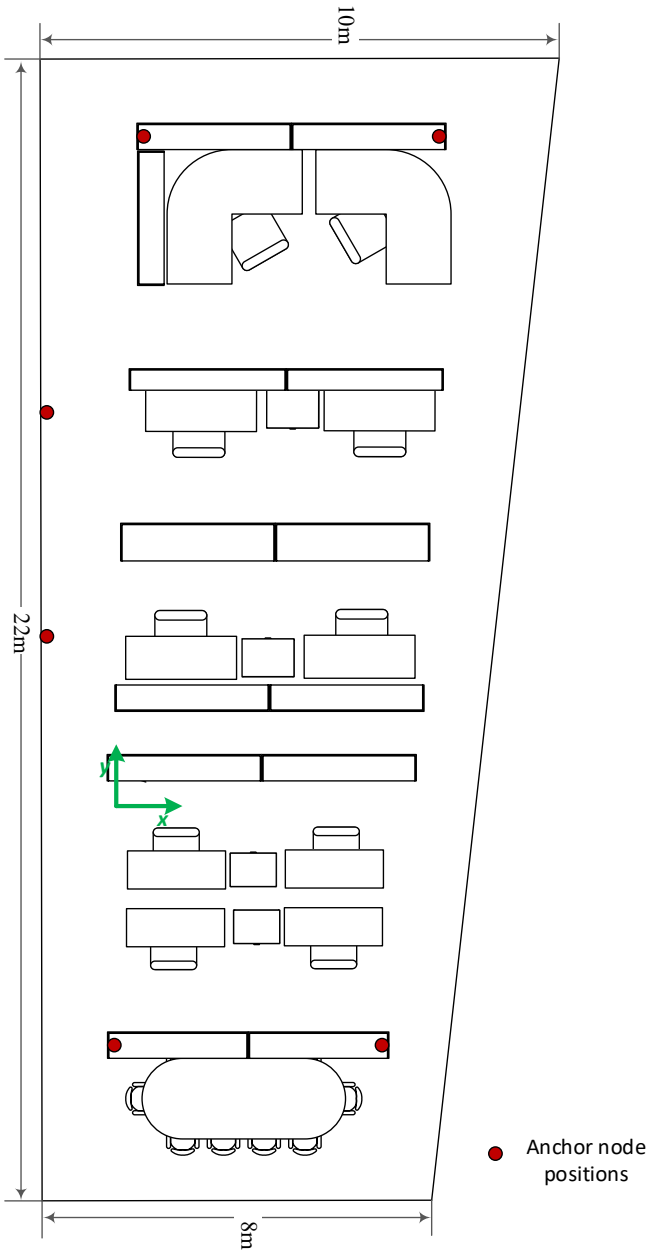


Figure 4-8: Area used for testing of the sub-6 GHz localization system

The coordinates of the anchor nodes with respect to the origin of the coordinate system are given in Table 4-1, while the positions of the anchor nodes at which the position estimates were obtained are given in Table 4-2.

In order to perform the DL-TDoA, the anchor nodes transmit frames consisted of a preamble and a Gold sequence [19]. The preamble is used to estimate the carrier frequency offset and to perform correction if needed. In this case this was not needed because the used SDRs have an extremely high accuracy clock source. Different gold sequences were used for each of the anchor nodes in order to be able to distinguish between them. The anchor nodes transmit sequentially, in time slots of 40 microseconds. Each frame duration is around 13.5 microseconds. The received power at the received, from each of the anchor nodes is shown in Figure 4-9. The transmissions are repeated each 4 milliseconds. This is the maximum repetition rate due to the limited bandwidth of the gigabit Ethernet used to transfer the samples from the host computer to the SDRs.

Table 4-1: Positions of the Anchor Nodes

Anchor Node No.	X [m]	Y [m]
1	4.4712	12.7628
2	-0.1653	12.9995
3	-3.1835	9.1530
4	-2.9826	5.8054
5	-0.4409	-3.3310
6	3.1369	-3.7051

Table 4-2: UE Positions During the Experiment

UE Position No.	X [m]	Y [m]
1	0.3715	5.1620
2	3.4095	5.1258
3	4.3654	0.8717
4	2.1204	0.7023
5	1.6587	2.0055
6	4.7950	3.4510
7	2.9239	6.3403
8	2.4962	9.4902

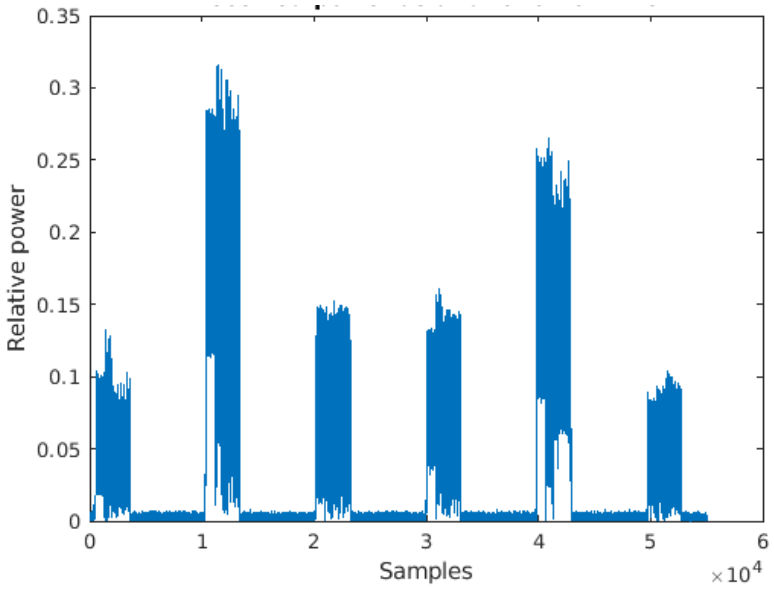


Figure 4-9: Received power as a function of time (transmissions from 6 anchor nodes)

This means that every 4 milliseconds (250 position estimates per second) a position estimate can be obtained by the UE. Of course, this is usually limited by the computational capability of the UEs. Having more anchor points, and a single gigabit Ethernet connection for all of the anchor nodes, would additionally reduce the repetition rate as well as the number of possible position estimates per second at the UE side. This can be increased by using 10 Gbit Ethernet connection or a separate host computer for each of the SDRs. Nevertheless, the repetition rate should not be extremely high, since this would introduce huge overhead on the wireless medium.

The initial testing of the system was performed by placing the UE SDR on 8 different positions and receiving

the positioning signals. The received signals are saved in files for post-processing. The resulting positioning estimates are shown in Figure 4-10. The anchor point positions are shown with black solid dots, while the red solid points show the position of the UE during the testing. The clouds of dots around the red points are the estimated positions. As can be noted, the clouds are well positioned around the true positions. A few main reasons for higher position estimate errors were identified. Some of them are systematic errors or bugs in the current software and can be corrected. This would further improve the positioning precision. Namely, the system was not thoroughly calibrated for the initial testing. This calibration should take into account the delays of the signals trough the synchronization cables, as well as through the SDRs, in order to improve the positioning precision. Additionally, it can be noted that the position estimates for each of the UE test positions, are grouped in smaller clouds. This issue was investigated, and the reason was found. The main problem is that the time slots were defined to be 40 microseconds, while the used sample rate was 245.76 MHz. With this ratio between the slot length and the sample rate, the slot length of 40 microseconds corresponds to 9830.4 samples. It is not possible by the SDR to introduce a non-integer number of samples and, therefore, it would perform rounding. Depending on the time passed from the start of the system, the rounding would introduce different errors in the timing. These errors would cycle trough a few different values and would group the position estimates in a few smaller clouds around the true position. In order to avoid this kind of error, the time-slot duration at which the anchor nodes transmit should be tied to integer number of samples, e.g. 9830 samples in this case, which corresponds to 39.998 microseconds. With these corrections, the positioning precision should improve.

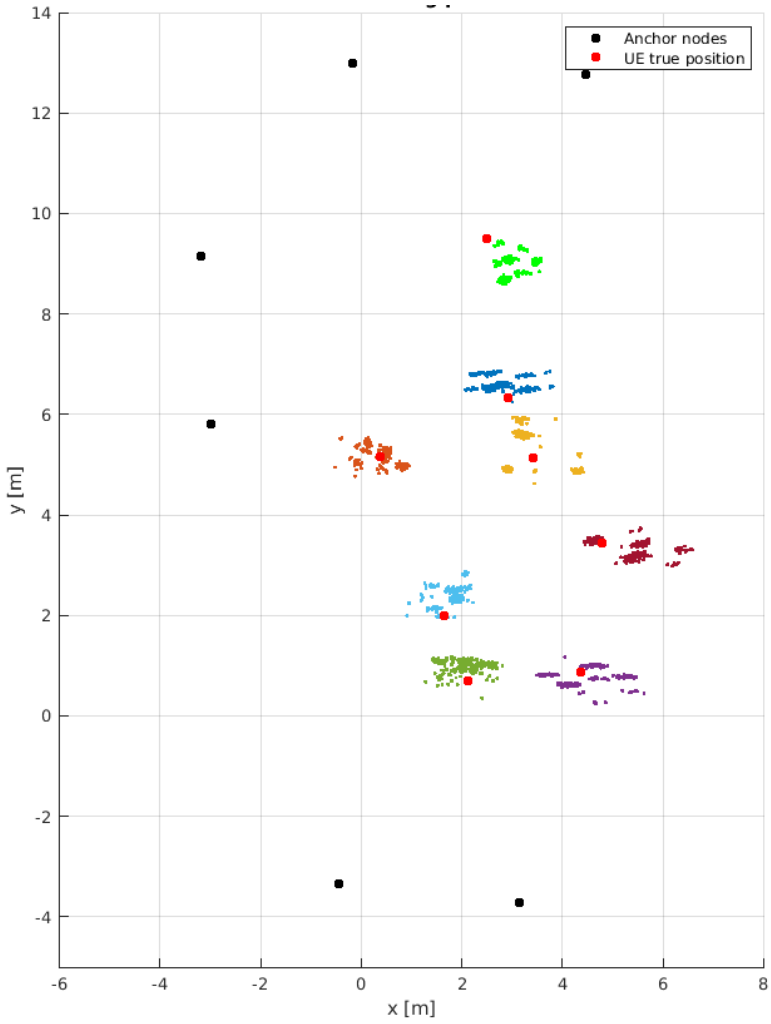


Figure 4-10: Initial testing position estimates

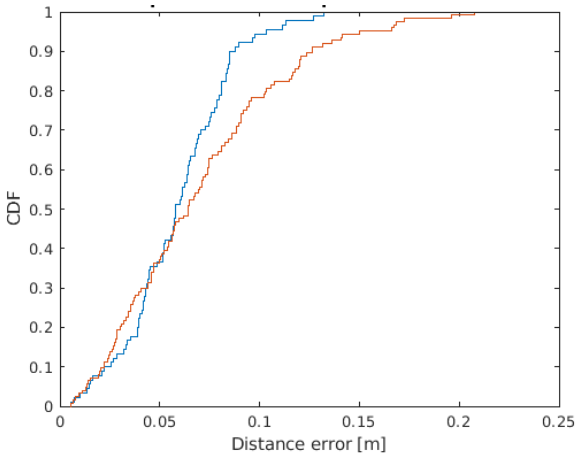


Figure 4-11: Empirical CDF of the position estimates, excluding the systematic errors

In order to estimate the positioning precision, excluding the systematic error due to the non-integer number of samples representing a single time slot, an empirical CDF of some of the smaller clouds, i.e., groups of position estimates, are being calculated. In Figure 4-11 the CDFs for two position estimates, excluding the systematic errors, are shown. As can be noted, almost all of the errors are lower than 20 centimetres. This is mainly due to the large channel bandwidth, i.e., 160 MHz, as well as the larger number of anchor nodes, i.e. 6 instead of the minimum 4 in this case.

4.3 LiFi

The LiFi localization system is based on received signal strength (RSS) technique, as introduced in 5G-CLARITY D3.1 [1], which relies on the signal power measured at the receiver. Two approaches, i.e. fingerprinting and trilateration, are considered for the LiFi positioning system.

Fingerprinting consists of two steps, (i) off-line survey, and (ii) real-time positioning. During the process of off-line survey, the RSS info of the signal at different locations is collected and recorded in a database. Then the on-line positioning process is completed by comparing the received information with the recorded data. An initial experimental measurement will be carried out with a LiFi system set up, to acquire the preliminary results for developing the fingerprinting-based algorithm.

The implementation will be based on pureLiFi LiFi-XC [20]. The number of deployed APs depends on the environment and the required positioning accuracy of the LiFi subsystem. The type of deployed LED luminaires also depends on the demonstration space to guarantee sufficient area is covered by each LiFi AP. The RSSI can be easily read from the LiFi-XC dongle in different operating systems. The acquired values can be in dBm or in percentage. Then, applying these numbers to the positioning algorithms an estimation of position could be acquired.

LiFi channels have quite different characteristics compared to RF channels. An advantage that the whole positioning system can benefit from LiFi is that an LiFi AP usually has much smaller coverage area comparing to other RF techniques, i.e., several meters in diameter under the lighting bulb. The positions of all the LiFi APs are known, thus by simply checking which AP the UE is connected-to the system can directly locate the user within a small area without any calculation or estimation. Then, by reading the RSSIs from the LiFi receiver, or using other positioning techniques that introduced in Sections 4.1, 4.2 and 4.4, a more accurate UE location can be acquired. However, the optical front-end devices usually have limited field-of-view, which makes the user device orientation significantly affect the RSS value and this should be considered in the algorithm. In an early implementation, we assume the LiFi dongle (UE) is constantly facing towards the ceiling.

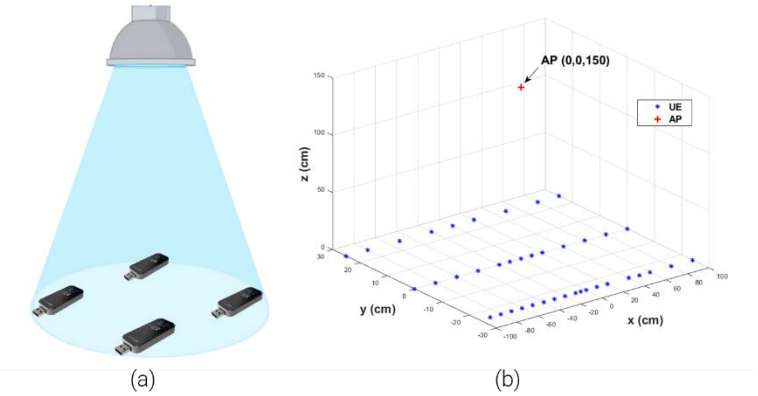


Figure 4-12: RSSI measurement using LiFi dongle (UE), (a) stationary measurement, (b) coordinates of AP and UE locations

The initial test is set up as shown in Figure 4-12(a). With one functional LiFi AP, the UE dongle is placed 1.5 m under the AP across the whole lighting area. Parameters such as coordinates, link distance and RSSI are recorded. As a result, the RSSI distribution within one LiFi AP coverage area can be acquired. This should be repeated several times to get the mean value and variance of the collected data. Initial results are shown in Figure 4-13. It can be seen that within the 1 metre diameter area under the AP, the RSSI stays at the maximum while its value drops by moving away from the centre. One limitation observed for the current measurement is the resolution of measured RSSI values. This is being looked at for possible improvements.

The following test uses the same setup, whereas the UE dongle moves within the lighting area following a fixed route as shown in Figure 4-14(a),(b). The real-time RSSI data and the timing info are recorded. Initial results are shown in Figure 4-14(c), including the measured RSSI and the calculated link distance. A negative correlation between the RSSI values and the link distance can be observed. Again, the low resolution of RSSI values is a limitation. The initial test results show that a basic positioning function could be obtained from the LiFi AP, simply by checking which AP is connected and what RSSI level is acquired. In addition, the negative corelation between the RSSI and the link distance has been observed, which could be further analysed and exploited for implementing the position estimation algorithm. The current low resolution of RSSI values will be investigated in the software, to achieve higher resolution which is necessary for an accurate positioning performance. The development of using RSS values with trilateration approach, which could potentially provide more accuracy, is ongoing.

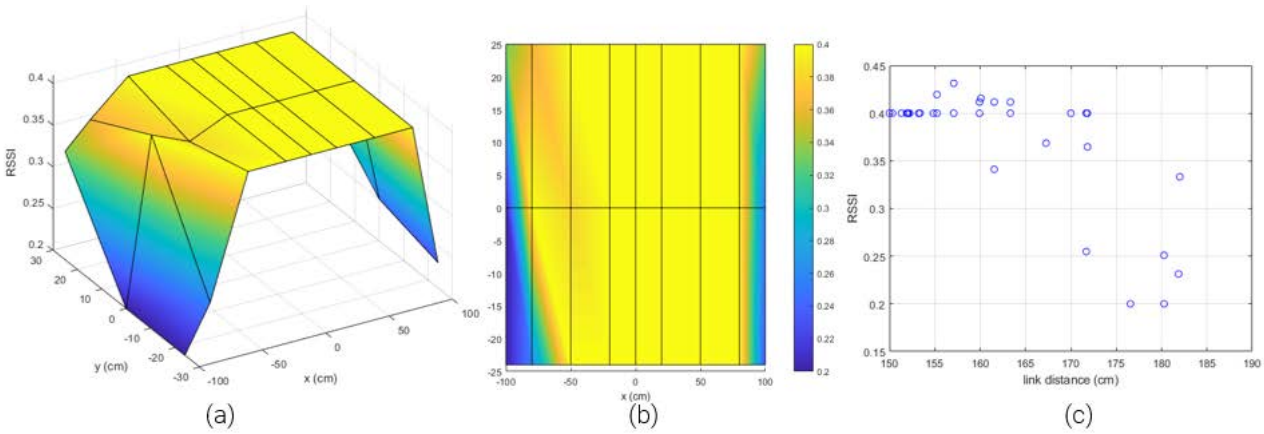


Figure 4-13: RSSI measurement under one AP, (a) 3D plot, (b) 2D plot, (c) RSSI vs link distance

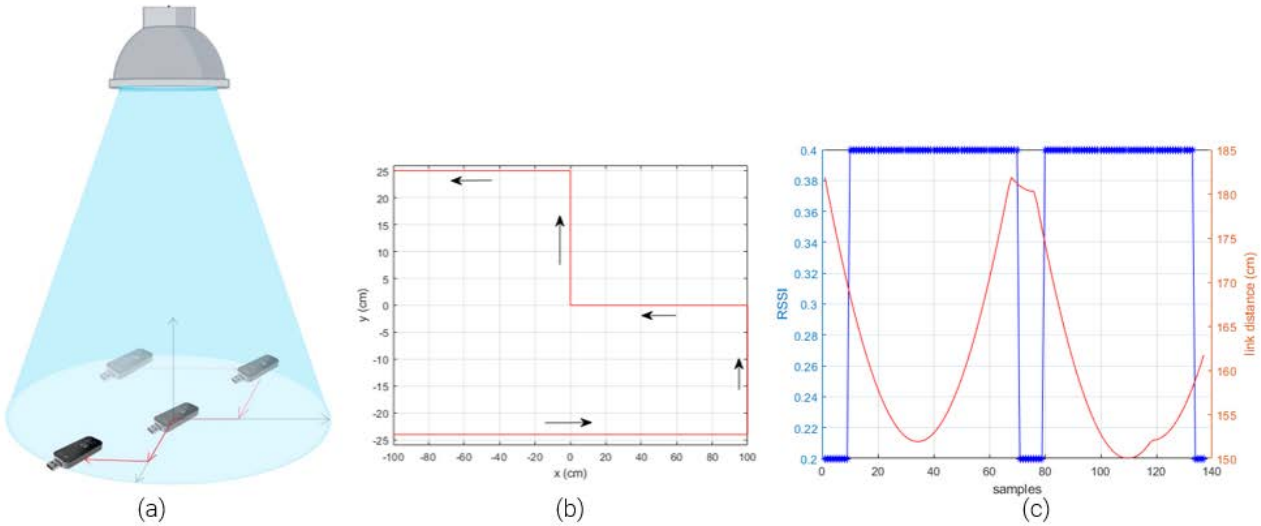


Figure 4-14: RSSI measurement with user movement, (a) set-up, (b) user moving route, (c) measured results.

4.4 OCC

In 5G-CLARITY D3.1 [1] the 5G-CLARITY Visible Light Positioning (VLP) solution was introduced, which allows a CMOS camera receiver to obtain cm-level accuracy in an environment where light fixtures transmit a known identifier. In this section detailed design of this solution is provided, and an experimental performance evaluation obtained in lab environment at i2CAT is discussed.

4.4.1 Design of 5G-CLARITY VLP solution

4.4.1.1 Modulation

This solution uses LED type transmitters, which allows switching (modulating) at a high speed. Thus, the light can switch between the two states (ON and OFF) fast enough so that the naked eye cannot perceive it. More sophisticated modulations are possible for higher data rates, as used in IEEE 802.11bb, which is not needed for light identification and therefore not used in this design. The structure of the VLP transmitter is depicted in Figure 4-15.

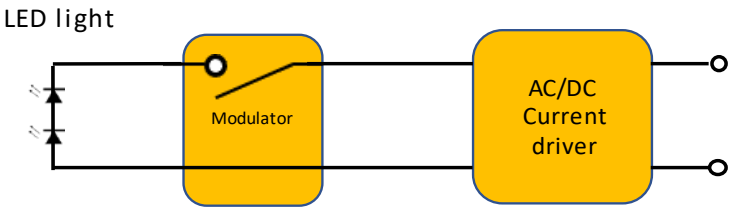


Figure 4-15: VLP modulator and LED driver

A complementary metal oxide semiconductor (CMOS) camera is used as the receiver. This type of cameras is present in almost all digital devices. These cameras have the peculiarity of not capturing the entire image all at once. Instead, the capture takes some time because the image acquisition is done in pixel columns, starting from one side and moving to the other. This mechanism is known as rolling shutter[13] and allows to

[13] https://en.wikipedia.org/wiki/Rolling_shutter

capture illumination variation during the time image is taken. Using a camera provides further advantages such as space multiplexing, since information transmitted by more than one light can be received with a single frame.

Different modulation schemes can be implemented in VLP. The simplest one is on-off keying (OOK), but Pulse Width Modulation (PWM) or Amplitude Shift Keying (ASK) are also possible. Introducing a modulation means that light fixtures will adjust their luminosity according to the encoded signal. The main issue when integrating a modulation with a commercial LED luminaire is ensuring that the current driver used for the light can support a load that varies with time. In some commercial drivers introducing current variation can lead to instabilities. In general, reducing the time the light is off and increasing the switching rate helps in overcoming this problem. The usage of a current driver ready to support dimming by means of a PWM scheme is another way to address the problem. If ASK modulation used a variable current source is needed, which is slightly more complex than the used light drivers and implies changing them if we want to incorporate VLP on an existing illumination infrastructure.

In case only a 2-level modulation is used (e.g., on and off), the modulator is simply a current switch implemented with a transistor. The modulation rate should be high enough to not be noticeable to the human eye, but low enough to be noticeable to the CMOS camera. For this purpose, we use switching rates is between 4 kHz and 10 kHz.

A drawback of simple OOK is that it can result in luminosity variations that can be noticeable depending on the transmitted bit pattern, independently of the transmission rate. For example, while a pattern 01010101010 can be perceived as a constant illumination level, the periodic sequence 00000111110 might result in light variations perceived by the human eye. Significant light variation may also be problematic. For example, 11111011111 offers a high luminosity while 0000010000 results in a poor illumination. In order to minimize these two problems, two approaches are possible.

- Manchester coding introduces a change of state to indicate a 0 or 1. This type of coding has the advantage of keeping the same illumination level in all the light fixtures since it transmits the same number of ones and zeros regardless of the information bits. However, only 50% of the illumination power of the LED fixture is effectively used.
- PWM consists of encoding a number of bits depending on the length of the on period. The off period is used as a separation and can be very short. The advantage of PWM is that it offers high luminosity (by keeping the on period longer), we can work with low switching frequencies (since the off period is short and barely noticeable to the human eye). PWM also avoid instability issues in most commercial LED drivers.

For the above reasons PWM is used in the design.

4.4.1.2 Required camera adjustment

A CMOS receiver needs to be properly tuned in order to be able to decode the data transmitted by the LED luminaire.

First, exposure time needs to be set low, as it is needed to scan the image fast (to cope with the modulation speed in the transmitter). Having a short exposure comes with the following advantages:

- The image does not saturate from excess of illumination. Given that the camera will be pointing towards the lights, having this setting set to low assures that the sensor does not saturate so we can see the fixtures well defined in the image.

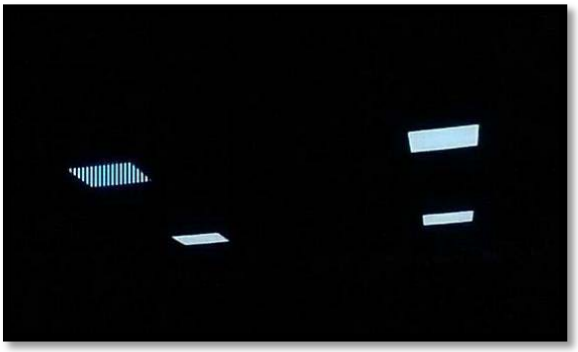


Figure 4-16: LED panels observed with a CMOS camera with correct exposure and ISO configuration. The top-left light is transmitting a code

- Each column of the image can contain only one modulation symbol. If we set the exposure time too high, the LED could transmit more than one modulation symbol while scanning one column, introducing aliasing.

Second, a drawback of using a low exposure is that the image will appear dark, sometimes too dark to visualize the transmission of the lights. By setting the ISO setting to a high value, it is possible to ensure that the lights will be visible in the image.

Figure 4-16 depicts an example image captured with a correctly configured CMOS receiver, where four LED panels are observed and the top-left one is transmitting information.

4.4.1.3 Landmark detection

Once an image like the one in Figure 4-16 is obtained, the next step is to obtain the 2D positions of the LED fixtures on the image. A set of tasks are required to achieve this depicted in Figure 4-17.

4.4.1.3.1 Resizing

To obtain a good positioning accuracy a high-resolution camera is needed, which however makes image processing more demanding, resulting in a reduction of the frames per second that can be treated for location, and therefore on the rate of locations per second that can be retrieved. To speed up processing, we apply an initial resizing of the image. This resizing does not impact the obtained accuracy because at this point of the process we are not interested in the information transmitted by the LEDs but only on detecting the regions of interest. A copy of the original image (high resolution) is kept decoding the transmitted information later on.

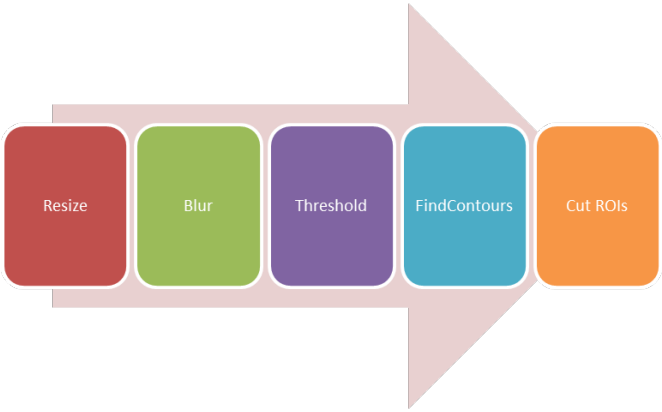


Figure 4-17: Required image processing steps

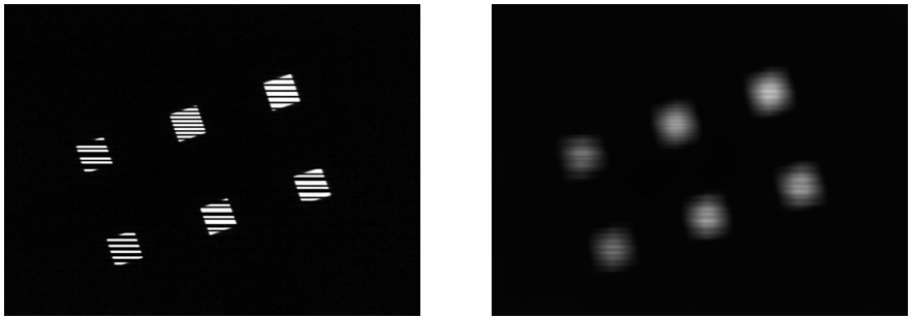


Figure 4-18: Image perfectly focused (left) and blurred image (right)

4.4.1.3.2 Blurring

Blurring helps to eliminate noise from other sources such as reflections. As we can see in Figure 4-18, blurring also turns the LED in the image more compact and easier to separate from the rest.

4.4.1.3.3 Thresholding

Used to separate the regions of interest from the rest of the image. It checks if the intensity of each pixel is above or below a given value. If the level is above the threshold the pixel will be painted it white, otherwise it will be painted it black. The threshold value is not difficult to set as it suffices to be somewhere in the middle between the black and white regions. This method will create a binary image highlighting the places where the lights are present.

After thresholding we apply a concatenated erosion and dilation operators that help eliminate any spurious noise from the image.

4.4.1.3.4 Finding contours

With the image result of the threshold detection, it is possible to identify the contour of the lights (the perimeter of the area where there is the change from white to black). This contour is applied to identify the area of the light where the codes are defined. Once the contour is detected from the low-resolution image this is applied to the high-resolution copy of the image. The contour allows to identify the part of the high-resolution image where the code can be extracted. Figure 4-19 depicts the overall image processing pipeline.

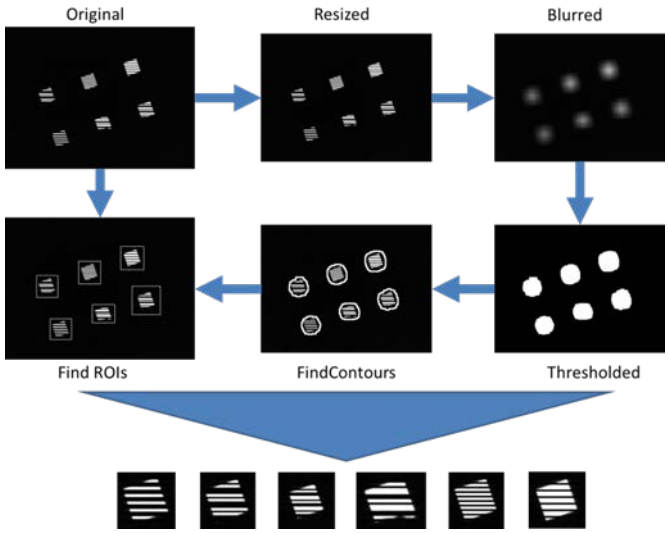


Figure 4-19: Sequential treatment of an image from the original to the detection of the ROI where the codes can be detected.

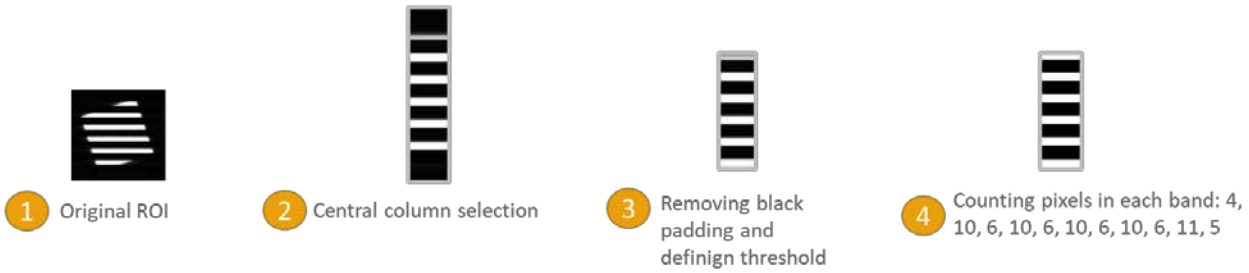


Figure 4-20: Steps in demodulation from RoI

4.4.1.4 Encoding within the Region of Interest (RoI)

After obtaining the regions of interest (RoI), the code transmitted by each LED fixture is detected. Several steps required for this are described in Figure 4-20. As we can see in step 1 in Figure 4-20 the receiver and LED may not have the same orientation, which can introduce errors in the decoding process. To avoid these errors, we select the central column of the RoI and only use this column for decoding (step 2 in Figure 4-20). In step 3, we eliminate the black padding above and under the code, as it does not pertain to the transmission itself. We then apply an Otsu threshold[14] to the resulting image to separate the white and black regions. Finally, in step 4, the number of pixels in each white or dark "gap" is obtained comparing the pixel value against the threshold. At this stage, the modulation bits can be retrieved by applying a ratio between bits and pixels, which is fixed for a given device and transmitter. In the case of Figure 4-20 $\frac{nº\ bits}{nº\ pixels} = \frac{1}{5}$, which results in a demodulated bit-sequence of "1, 0, 0, 1, 0, 0, 1, 0, 0, 1, 0, 0, 1, 0, 0,1". An additional processing is applied to this sequence that is to remove the first and last white gaps, which could be altered by the size of the light fixture, resulting in a final sequence of "0, 0, 1, 0, 0, 1, 0, 0, 1, 0, 0, 1, 0, 0".

4.4.1.5 LED code space design

The goal of our VLP system is to be able to identify a set of landmarks, i.e., the LED luminaires, in the 2D space of the image. These landmarks is then matched to the location in 3D space of the actual light fixture. To do the mapping between the landmark in the 2D figure and the light fixture in the database, we need to decode the light ID transmitted within the RoI, as discussed in Section 4.4.1.4. We discuss now how we design the light IDs.

As the transmission of the light ID is done continuously, we need a method to identify the light ID from a raw sequence of bits (as the one obtained in Section 4.4.1.4). The most common approach is to define pattern of bits that works as a preamble to separate consecutive codes. This implies adding additional bits and avoiding the presence of the preamble bits inside the light ID codewords. Another problem with preambles is that they may not be correctly received in a single frame, which turns the whole capture unusable.

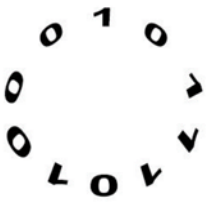


Figure 4-21: Acyclic code design

[14] https://en.wikipedia.org/wiki/Otsu%27s_method

Instead, we use another approach based on acyclic codes, which are codes with a unique bit-pattern that cannot appear in any other sequence of codes. All the possible bit sequences in of length L (C) = 10 in Figure 4-21 belong to the same code: 0010111010, 0101110100, 1011101000, 0111010001, 110100101, 101001011, 0100010111, 100010110 and 0001011101. Our encoding approach still allows to support dimming if several codes with different code intensities are allocated to the same light.

Notice that using cyclic codes obviously reduces the space of different light IDs. This can be mitigated by reusing light IDs across different light fixtures and use context (e.g., codes of surrounding lights) to disambiguate them.

4.4.1.6 Positioning algorithm

At this point, we have all the information we need to run the positioning algorithm. We have the 3D points read from the lights, the 2D projections on the camera and we know how the camera alters the real world, as depicted in Figure 4-22.

The problem we need to solve is finding the relation between 2D and 3D coordinates. This problem is well known as solvePnP in the state of the art. Finding the pose of the camera relative to the object consists of minimizing the reprojection error. Having 3D-2D point correspondences means we can project the 3D points on the image and take the error as the distance between the projections and the real 3D points. That way we can extract the relation between the world and the camera frame. The cost function is depicted in Figure 4-23 [21].

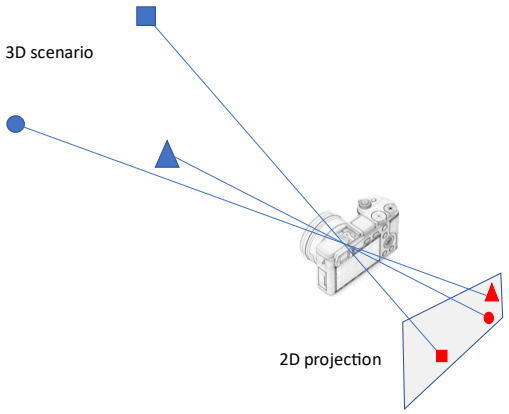


Figure 4-22: Correspondence between 3D and 2D locations of landmarks

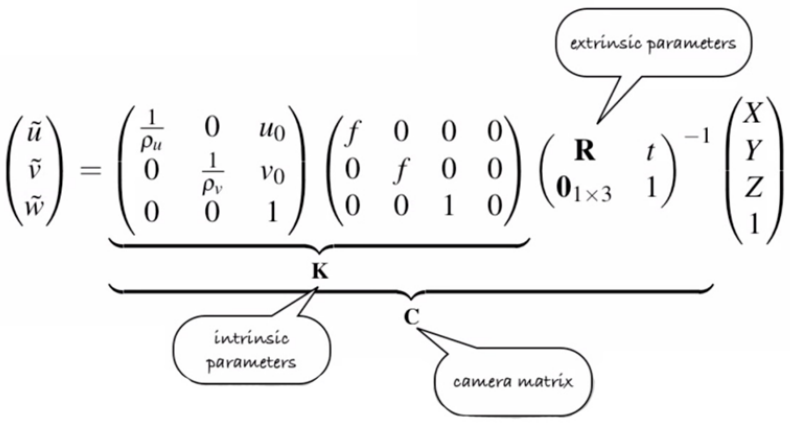


Figure 4-23: Cost function to be minimized in solvePnP [21]

To solve this problem, we use the iterative solution available in *OpenCV*[15], which consists in the following steps:

i) Initialize algorithm with hypothesis (e.g., x = 0, y = 0, z = 0). We can also use any other algorithm that can help initialize these values (such as homograph estimation).

ii) Project 3D points obtained from the lights to the image plane.

iii) Get the error between the real projections and the obtained projections.

iv) Calculate new extrinsic parameters.

v) Repeat until convergence.

The same algorithm can also be used with restrictions. For instance, if we know an object will only move on the ground plane, we can fix Z = height, pitch = 0 and roll = 0. This will make the calculations faster and more precise (e.g., when we want to calculate the position of a moving vehicle).

4.4.2 Evaluation of 5G-CLARITY VLP solution

4.4.2.1 Experimental scenario description

We have built a testbed in order to evaluate our VLP system. It includes an array of 6 LED fixtures located on the ceiling as a rectangle of 0.6 x 1.2 m. On the floor, at 2.4 m on the Z axis, we have positioned a grid of the same size containing 91 known locations separated 10 cm from each other, where we do the measurements. As a receiver, we have used a Google Nexus 5 and a Raspberry Pi 3.

4.4.2.2 Accuracy in static conditions

We evaluate the accuracy of our VLP solution under static conditions, over 90 points in a grid pattern of a 60 x 120 cm^2 area located 240 cm below the LED transmitters, where we perform 50 localization measurements in each point. We also investigate the improvement provided by applying a Kalman filter.

Figure 4-24 (left) shows the CDF of the localization error measured in this set of experiments. As shown in the figure, using 6 lights and a Kalman filter, 96% of the computed localizations suffer an error below 10 cm. The average accuracy with and without the filter is 4.9 cm and 6.8 cm, respectively. As the number of LED landmarks decreases, accuracy decreases as well. For example, for 5 light transmitters, only 78% and 63% of the localizations were obtained within decimetre accuracy (with and without Kalman filter, respectively). While accuracy is still relatively high for 4 lights, using 3 lights leads to significant accuracy degradation.

In order to analyse the spatial distribution of the localization error, we plot such error on a heat map. We carry out this study in the 6-light scenario. As shown in Figure 4-25, the localization error is relatively distributed in space, and does not specifically depend on the location.

We have also measured the orientation accuracy of the system. To this end, we have performed 1186 individual orientation error measurements. Figure 4-24 (right) illustrates the orientation CDF for 6 lights. The average orientation error in that case is $2.87°$, which is similar to (slightly better than) that of other systems such as Luxapose. Orientation error increases as the number of lights decreases.

[15] https://learnopencv.com/tag/solvepnp/

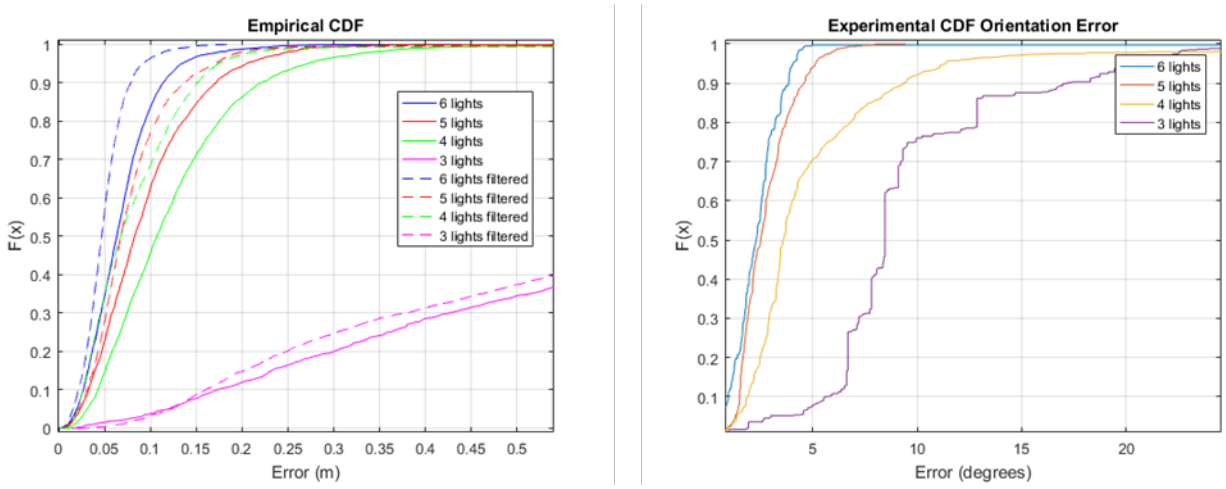


Figure 4-24: CDF of the measured localization error(left), and CDF of the measured orientation error (right).

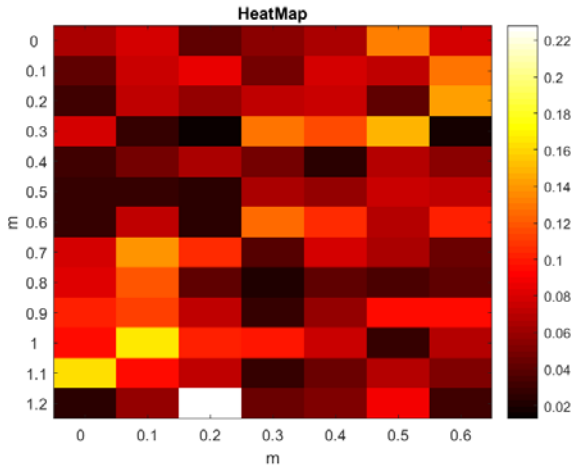


Figure 4-25: Localization error heat map

4.5 Fusion approach

Perhaps the most basic and popular data fusion technique is Bayesian Recursive Filtering (BRF), where the prior knowledge is combined with the measurement data to estimate a parameter [22] Although there exist multiple especial cases of BRFs, e.g., Kalman Filtering (KF), we stick to the most general case for the purpose of this work. This enables us, among other things, to deal with any type of distributions (of prediction or measurement data). In particular, we consider two concrete cases where, (i) the distributions of prediction and measurement are both Gaussian, and (ii) there is no condition on the distribution. The latter is realized in practice with the aid of Particle Filters (PF). We denote the first case by G-BRF (stands for Gaussian-BRF) while the latter is represented by PF-BRF. Figure 4-26 describes the procedure of parameter estimation in BRF where *x* denotes the parameter to be estimated and *z* represents the measurement.

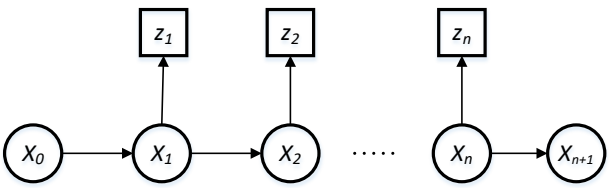


Figure 4-26: BRF procedure to estimate parameter *x*

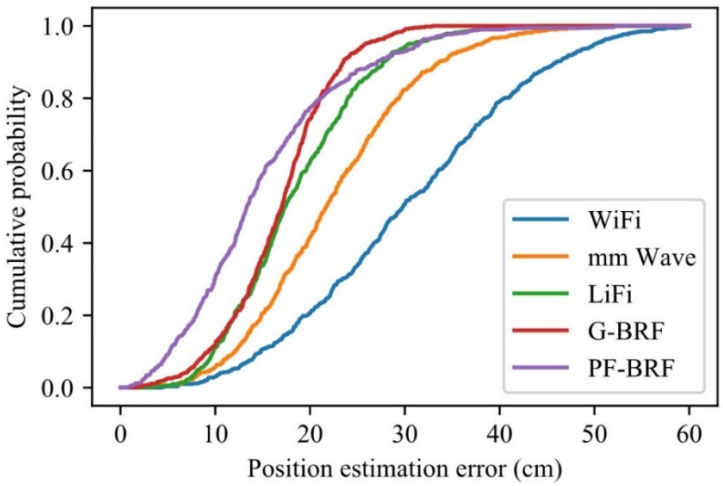


Figure 4-27: Position estimation accuracy of different technologies.

In the context of 5G-CLARITY project, the BRF fusion approach is expected to be employed at the core of the localization server where it will receive the location-related data, e.g., ranging measurements, from different technologies and returns an estimation of the position of user. Traditionally, the accuracy of position estimation is a measure to evaluate the performance of an estimator of such a kind. Therefore, we rely on the CDF of the position estimation error to compare the accuracy of position estimation conducted by different wireless technologies with that of the fusion algorithm.

In particular we consider a scenario where the user travels 20 m with constant velocity of 50 cm/s along the *x* axis. It then turns randomly to left or right and moves for another 20 m along the *y* axis until it again turns randomly and continues its journey for 20 m along the x axis (altogether 60 m). Moreover, we assume that multiple distance estimations are available from different technologies, each subject to a degree of uncertainty. Further assumption is that a sensor on the user device returns a noisy measurement of its velocity. The following primary results obtained when applying G-BRF and PF-BRF.

As shown in Figure 4-27, the accuracy of estimation increases when the distance measurements from different technologies are fused using BRF approaches. In particular, when estimating, G-BRF gives higher weights to the measurements with higher certainty, thereby improving the precision of the position estimation. PF-BRF takes another step forward and estimates the posterior distribution using the particles (instead of a single Gaussian function). Such an approach can highly improve the accuracy of the position estimation. However, PF-BRF requires a great deal of parameter tuning to deliver superior performance. This has not been carried out in the simulation setup generating results presented in Figure 4-27.

5 System Level Evaluation

To meet the challenge of a 1000x mobile system capacity increase, multiple solutions such as interference coordination, advanced transmitters (TXs) / receivers (RXs), tighter 3GPP and non-3GPP internetworking, network densification, interference management etc. have been identified as the key enablers. Specifically, a very high system area capacity target of 500 Mbps/m^2 via user and control plane functionalities has been envisaged in 5G-CLARITY D3.1 [1]. To achieve this objective, mobility, multi-connectivity protocols, SDN controllers, and network agents for eMBB and URLLC will be utilized during the lifetime of the 5G-CLARITY project. In this deliverable, integration of multi-WATs, specifically 5GNR, Wi-Fi and LiFi, is going to be investigated by using computer simulations. In the following subsections, the methodology to implement novel user and control plane functional blocks of the project will be presented. Moreover, the functional blocks will also be validated by our proposed real time traffic simulator to obtain proof-of-concept results.

5.1 Simulation platform architecture

The real-time traffic simulator developed for 5G-CLARITY user and control plane functionalities will be detailed in this subsection. As stated in 5G-CLARITY D3.1 [1], the aim is to design a user and control plane framework for indoor private network applications under the 5GNR, Wi-Fi and LiFi coexistence. Accordingly, several evaluation targets arise to show the proposed system capabilities. Firstly, the investigation of area capacity achieved with multi-WAT (5GNR/Wi-Fi/LiFi) implementation is an important indicator to meet the aims of IMT-2020 and beyond. Secondly, the KPI gains obtained by the multi-connectivity framework, compared to conventional networks, will be able to clearly show the benefits of the proposed 5GNR, Wi-Fi and LiFi coexistence. Lastly, the end-to-end traffic performance gains obtained via the utilization of the proposed eAT3S, compared to conventional systems, will also emphasize the benefits of 5G-CLARITY user and control plane functionalities. Therefore, the real-time simulator environment was developed to evaluate the targets as well as to monitor and compare the achievable performance values. In this deliverable, both the first and second targets will be addressed without loss of generality. It is important to note that the remaining target will be carried over in 5G-CLARITY D3.3 and 5G-CLARITY D2.3.

The real-time realistic traffic simulation platform for the multi-WAT indoor private network application simulations was developed in MATLAB. Thus, the simulator will have the ability to utilize some of MATLAB's built-in 4G LTE and 5G Toolbox functionalities. The operation principle of the simulator package is detailed by the high-level block diagram in Figure 5-1. As can be seen from the figure, the real-time traffic simulator employs custom MATLAB functions to mimic the UE, 3D geometry conditions and access network, and core network behaviour. Accordingly, the room dimensions, location of the 5GNR/Wi-Fi/LiFi APs, and further details on each AP (including the operation frequency) are defined by the "environment parameters" configuration file. Furthermore, third party solutions such as Python based *owcsimpy* for physical (PHY) layer channel modelling [23], OpticStudio raytracing based realistic channel impulse response generation [24], SDN controller solutions along with in-house functionalities etc., are supported. As an output, the telemetry data of the traffic generated by mobile terminals are recorded and further processed to assess the network performance.

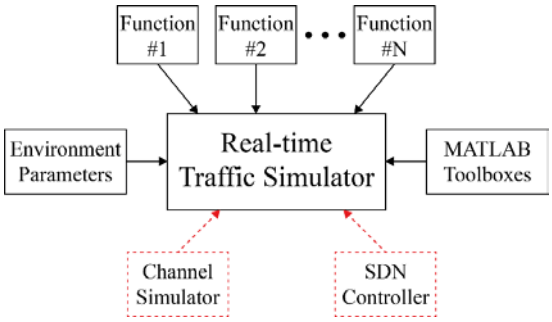


Figure 5-1: The high-level architecture of the simulation platform. An example in-house and third party functionalities are depicted by solid black and dashed red lines, respectively.

5.2 Simulation platform functions and capability

In this section, details of the developed platform that can be used to simulate the performance of (i) 5GNR, Wi-Fi and LiFi access networks individually; and (ii) an integrated 5GNR/Wi-Fi/LiFi based multi-connectivity framework is provided.

The high-level architecture of the traffic simulator, which can satisfy real-time data delivery and realize L1 and L2 of the OSI protocol stack is given in Figure 5-2. The simulation platform can mimic the operation of a multi-WAT 5GNR/Wi-Fi/LiFi deployment scenario by generating a data request at the UE to the provision of the data by multiaccess traffic routing in each time instant. As can be seen from the figure, the simulator runs two main loops. Firstly, the transmit time interval (TTI) loop is utilized to capture the channel and resource-based changes within the system at both the user and network sides during the transmission of a PRB. Secondly, the Monte-Carlo loop is employed to capture the user requests.

The operation of the simulator begins with the physical geometry initialization, which consists of the generation of simulation environment dimensions, shape and locating the APs (5GNR/Wi-Fi/LiFi) and UEs. Then, the initial assignment of each UE to an AP, by picking the highest channel-gain yielding AP, is executed. The size of the contents, the content types that will be requested by the UE and the popularity, which follows Zipf distribution [25], are generated. The probability distribution function (PDF) of downloading-content popularity, which follows Zipf distribution is given by

$$P(x) = \frac{x^{-(\rho+1)}}{\zeta(\rho+1)},$$

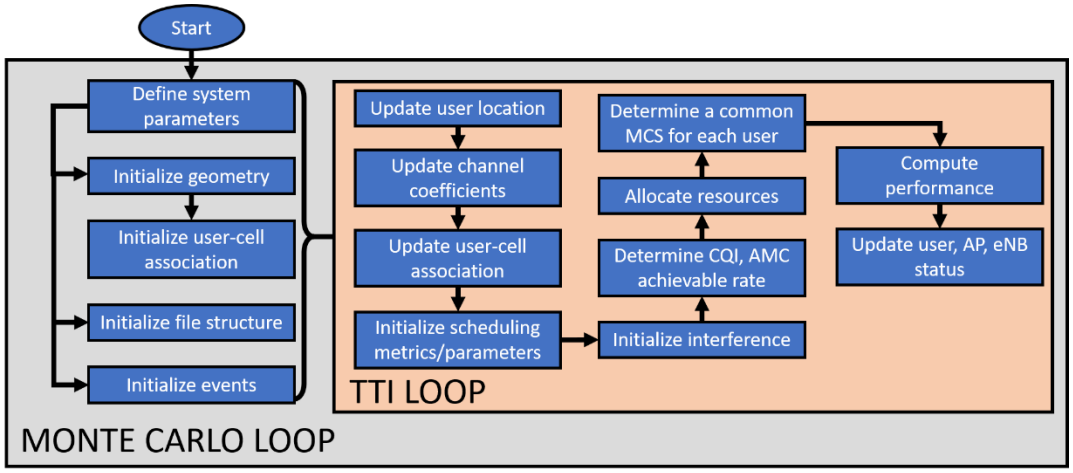


Figure 5-2: The building blocks for multi-WAT traffic simulator.

where ρ and $\zeta(\cdot)$ are the positive adjustment parameter and Riemann zeta function, respectively.

After this, UE content request events are generated based on either the full-buffer traffic or Poisson distribution. It is important to note that the full-buffer traffic assumes that all the resources are allocated to a single user. Furthermore, the probability mass function (PMF) is given as follows [25, 26]:

$$f(k;\lambda) = \Pr(X = k) = \frac{\lambda^k e^{-\lambda}}{k!},$$

where e is Euler's number, k is the number of occurrences, and λ is both the expected value and variance of the random variable X. In parallel to the initialization, UE channel coefficients are altered based on the terminal's location, where the channel-gain per subcarrier/PRB is defined by the combination of pathloss and fading characteristics. Once the changes in the channel-gain coefficients are determined from one transmission instant to another, the updated channel coefficients are used in the re-initialization of UE-AP association. For a given TTI, the scheduling process is initiated once there is one or more active UE(s) making a request for a service.

The main vision of 5G-CLARITY is to design a network where the mid and highband spectrum is harnessed along with the optical bands, which consists of visible light (VL) and infrared (IR) bands. To conform to this vision, three WATs, namely 5GNR, Wi-Fi and LiFi are chosen as the main components of the RAN. For 3GPP, two centre frequency values, i.e., i) 3.5 GHz (3.4-3.8 GHz) mid-band spectrum, and ii) 26 GHz (24.25-27.5 GHz) high band mmWave spectrum, which is identified as the "pioneer" band for 5G across Europe, are adopted. For non-3GPP Wi-Fi, IEEE 802.11ac standard is adopted as the main wireless local area network (WLAN) solution. Therefore, the centre frequency of the operation is 5 GHz (5.17-5.33 GHz). Similarly, for non-3GPP LiFi, the 299.79-374.74 THz IR band is adopted to comply with the IEEE 802.11bb standard. It is important to emphasize that LiFi has both the illumination and communication functionalities simultaneously, which brings a significant power efficiency gain due to the dual operation. Hence, the VL band spectrum is also considered in our simulation scenarios to bring the complete advantages of optical domain. For both the 3.5 GHz 5GNR and 5 GHz Wi-Fi, the geometry-based stochastic channel modelling approach, known as the WINNER model [27], is used for parameterisation of indoor propagation characteristics. Accordingly, the pathloss model (in dB) could be given as follows [28]:

$$\mathrm{PL}_1(f, \boldsymbol{r}_\mathrm{S}, \boldsymbol{r}_\mathrm{R}) = A\log_{10}(d) + C\log_{10}\left(\frac{f}{5\times10^9}\right) + B + \chi_1,$$

where f denotes the operation frequency, in Hz, and d is distance, in metres, between the transmit AP and mobile terminal in meters, respectively. The distance between the transmit AP and receive mobile terminal could easily be calculated by $d = \|\boldsymbol{r}_\mathrm{S} - \boldsymbol{r}_\mathrm{R}\|_2$ where the position vectors of the source (S) and receiver (R) denoted by $\boldsymbol{r}_\mathrm{S}$ and $\boldsymbol{r}_\mathrm{R}$, respectively. The parameters A, B and C represent the fitting parameters, i.e., the pathloss exponent, the intercept and pathloss frequency dependency factor, respectively. Moreover, χ_1 is a normal distributed random variable with zero-mean and variance of σ_1^2 whose value is chosen based on the scenario. According to the WINNER model, 13 different propagation scenarios are considered. Since our goal in this report is to simulate the indoor private network applications, "the A1 – indoor propagation scenario" and the related parameters are adopted. In A1, the use case is defined as an indoor office application with both line-of-sight (LoS)/non-LoS (NLoS). Furthermore, the operating frequency and mobile terminal speed ranges are given by $2 \leq f \leq 6$ GHz and $0 \leq v \leq 5$ Km/h, respectively. The set of parameters for the scenario A1 is given in Table 5-1.

Table 5-1: Parameters for PL_1

Parameter	Value
A	18.7
B	46.8
C	20
σ_1^2	3 dB

Table 5-2: Parameters for PL_2

Parameter	Value
n	1.8
σ_2^2	2.03 dB

Similarly, for 26 GHz mmWave 5GNR APs, the empirical large scale path loss model given by [29] is used,

$$\mathrm{PL}_2(f, \boldsymbol{r}_\mathrm{S}, \boldsymbol{r}_\mathrm{R}) = -147.6 + 20\log_{10}(df) + 10n\log_{10}(d) + \chi_2$$

where the operation frequencies set for the above expression is $f \in \{26, 32, 39\}$ GHz. Moreover, the element χ_2 is also a normal distributed random variable with zero mean and variance of σ^2. For the considered 26 GHz band, which is best suited for our application, the parameters will be as shown in Table 5-2.

Lastly, for LiFi, it has been reported in [30] that the reflected light from the walls, ceiling, and floor, becomes negligible for most users if a dominant LoS path exists. As the LiFi AP density in our considered scenario is sufficient to provide a dominant LoS path to the UE even when the UE moves, the reflections are ignored. It is notable that the blockage effects are omitted without loss of generality. Hence, the LoS-only channel assumption (Eq. (10) in [31]) will be sufficient to obtain realistic results. The LoS-based pathloss model could be obtained by some simple algebra as follows:

$$\mathrm{PL}_3(L_\mathrm{S}, A_\mathrm{R}, \Psi_\mathrm{R}, \boldsymbol{n}_\mathrm{S}, \boldsymbol{n}_\mathrm{R}, \boldsymbol{r}_\mathrm{S}, \boldsymbol{r}_\mathrm{R}) = 7.9818 + 10\log_{10}\left(\frac{d^{(L_\mathrm{S}+3)}}{(L_\mathrm{S}+1)A_\mathrm{R}}\right) - 10L_\mathrm{S}\log_{10}(\boldsymbol{n}_\mathrm{S}\cdot(\boldsymbol{r}_\mathrm{R}-\boldsymbol{r}_\mathrm{S}))$$
$$-10\log_{10}(\widehat{\boldsymbol{n}_\mathrm{R}}\cdot(\boldsymbol{r}_\mathrm{S}-\boldsymbol{r}_\mathrm{R})) - 10\log_{10}\left(\Lambda\left(\arccos\left(\frac{\widehat{\boldsymbol{n}_\mathrm{R}}\cdot(\boldsymbol{r}_\mathrm{S}-\boldsymbol{r}_\mathrm{R})}{d}\right)\right)\right)$$

where the parameters $\boldsymbol{n}_\mathrm{S}$ and $\boldsymbol{n}_\mathrm{R}$ denote the normal vector of the source (S) and receiver (R)respectively. Moreover, the Lambertian emission order and the effective area of the S and R are also denoted by L_S and A_R, respectively. The parameter L_S is defined by $L_\mathrm{S} = -\ln(2)/\ln(\cos(\Phi_{1/2}))$ where $\Phi_{1/2}$ is the half power semi-angle of the S. Furthermore, the indicator function, $\Lambda(\cdot)$, which decides whether the incident beam is within the field-of-view (FoV) of the receiver or not is defined as follows:

$$\Lambda(x) = \begin{cases} 1, & \text{if } |x| \le \Psi_\mathrm{R} \\ 0, & \text{otherwise} \end{cases}$$

The parameter $\Psi_{1/2}$ is the semi-angle of FoV for the R. It is important to note that a random polar orientation angle must be taken in account at R, which emerges from the way the mobile terminal is held by the user. Therefore, the normal vector obtained after the random orientation effect ($\widehat{\boldsymbol{n}_\mathrm{R}}$) could be represented by

$$\widehat{\boldsymbol{n}_\mathrm{R}} = \boldsymbol{G}_{xy}(\theta_z)\,\boldsymbol{G}_{xz}(\theta_y)\,\boldsymbol{n}_\mathrm{R}$$

where the 3×3 real valued $\boldsymbol{G}_{xy}(\theta_z)$ and $\boldsymbol{G}_{xz}(\theta_y)$ represent the rotation matrices in xy and xz-planes, respectively. Moreover, the parameters θ_z and θ_y are the z and y axis polar rotation angles as shown on the Figure 5-3.

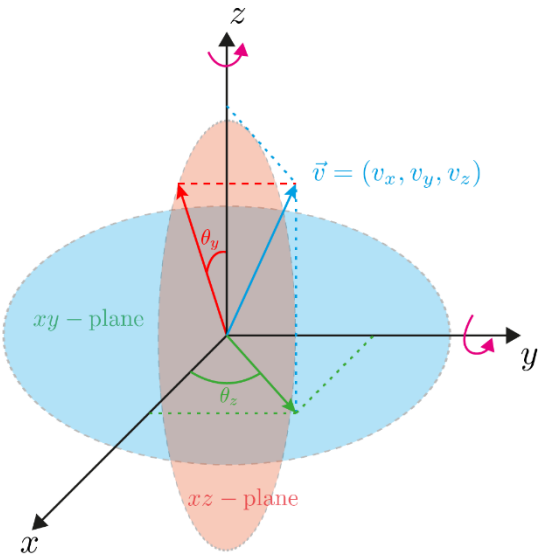


Figure 5-3: The random orientation model for the mobile LiFi terminal.

Based on the empirical data, it is observed that the polar angle, $\theta_y \in [0, \pi/2]$, could be approximated as a Gaussian distributed random variable [32] when the mobile user is walking within the room with the mean μ_3 and variance σ_3^2. Similarly, the azimuth angle, $\theta_z \in [-\pi, \pi]$, closely follows a uniform distribution when the mobile user is engaged in a walking activity. Hence, the adopted set of values for PL_3 expression is given in Table 5-3.

After the AP-UE level initialization of the scheduling parameters, the interference from inter-WAT and intra-WAT are calculated. The multi-WAT framework simulator is designed to support the current standards, e.g., 802.11ac/ax for Wi-Fi, 802.11bb for LiFi and 3GPP for UMTS, LTE and 5GNR with the eAT3S and the adaptive modulation and coding (AMC) capability for the transmission of content. Thus, in the DL direction, the eNB/gNB selects the appropriate MCS based on the channel quality indicator (CQI) of the UE, which is determined by the SINR and the UE's front-end characteristics. Based on the adopted MCS per user, the achievable spectral efficiency will be determined by using the modulation order and coding rate. Finally, the performance of each user is evaluated by the aggregate throughput. The aggregate throughput values of each UE are used to inform the APs to update their policies and this entire flow is repeated as long as the Monte Carlo loop size. In the following subsections, further details on the adopted WATs and the aggregation in the simulation environment are provided, and the obtained capacity results discussed.

Table 5-3: Parameters for PL_3

Parameter	Value
L_S	1
A_R	1 cm^2
Ψ_R	85°
$\boldsymbol{n}_S$	$[0 \quad 0 \quad -1]^T$
$\boldsymbol{n}_R$	$[0 \quad 0 \quad 1]^T$
μ_3	29.67°
σ_3^2	8.91 dB

5.3 Simulation scenario description

In 5G-CLARITY, the cellular networks are envisaged to employ the key technologies like mmWave and LiFi, to harness the dense network structure along with the bandwidth abundance in the higher portion of the spectrum. Due to the intrinsic nature of the mmWave and the optical band, the size of the cells must be significantly reduced compared to the previous generation legacy systems. In 3G and 4G cellular networks, the traffic capacity was enhanced by macro- and microcells, respectively. However, in 5G, the emphasis is on the pico, femto and attocells, where the very small cells formed by LiFi APs make up the LiFi attocellular structure. Another key enabler technology, namely massive multiple-input-multiple-output (MIMO), requires the transmit powers to decrease by around 10-20 times to keep the transmit power per AP the same. Therefore, the small cell and ultra-dense network concepts emerge as one of the core characteristics in 5G. Accordingly, two important parameters, i.e. interference and coverage, must be taken care of. To accurately capture these parameters and other characteristics of indoor private 5G networks with 5GNR, Wi-Fi and LiFi technologies, a rectangular prism shaped structure is generated with the dimensions of $140 \times 70 \times 4\ \mathrm{m}^3$. It is important to note that the typical geometry could represent anything from corporate office headquarters, hospitals, museums, universities, and indoor stadiums. The floor area of the simulation scenario becomes 9800 m^2 as depicted in Figure 5-4(a), (b), (c) and (d), which will be used in system area capacity calculations. To be able to reach or exceed the 500 Mbps/m^2 target area capacity objective, various parameters, such as the number of 5GNR/Wi-Fi/LiFi APs, locations of the APs, operation frequencies, bandwidth etc., needs to be considered for both the coverage and interference management purposes. The indoor 5GNR, Wi-Fi and LiFi applications for residential room, office, and hospital applications could be realized by hexagonal, square lattice grid and Poisson point process (PPP) AP deployments [33]. The PPP deployment offers a mathematically tractable tool to achieve better cell design in terms of interference management. However, the PPP based LiFi AP deployment is impractical as the lighting structure in such indoor environments consists of fixed light aperture sockets and wiring. The hexagonal deployment is widely adopted by cellular communication networks due to its geometric simplicity and ready-to-use frequency reuse support. However, the illumination is another concern for LiFi deployment, where there are UK and European set regulations on the uniformity and minimum levels of in working environments [34]. Therefore, a square lattice grid LiFi AP deployment is adopted in system level simulations.

In the considered generic application scenario, three different WATs; LTE/5GNR at 3.5 GHz and/or 60 GHz, Wi-Fi at 5GHz, to comply with 802.11ac, and LiFi for 800-1000nm band, aligned with 802.11bb, are deployed. For the LiFi AP deployment, sufficient coverage within such a large area requires a careful estimation. Accordingly, a basic PL formula, which encounters only LoS channel contributions for optical wireless communication devised from [35], is employed. As the intersection between the illumination cone of a LiFi AP and the floor surface creates a shape of a disk, the optical cell boundary can be defined as the circle where the maximum optical power gets halved as follows:

$$\frac{h^{(0)}(S,R)}{\max\{h^{(0)}(S,R)\}} = \frac{1}{2}, \qquad \forall S,R$$

where the parameter $h^{(0)}(S,R)$ represents the optical channel DC gain between the source (S) and receiver/terminal (R), which could be defined by [31]

$$h^{(0)}(S,R) = \frac{(L_S+1)A_R}{2\pi d^{(L_S+3)}}\left(\boldsymbol{n}_S \cdot (\boldsymbol{r}_R - \boldsymbol{r}_S)\right)^{L_S}(\boldsymbol{n}_R \cdot (\boldsymbol{r}_S - \boldsymbol{r}_R))\,\Lambda\left(\arccos\left(\frac{(\boldsymbol{n}_R \cdot (\boldsymbol{r}_S - \boldsymbol{r}_R)}{d}\right)\right)$$

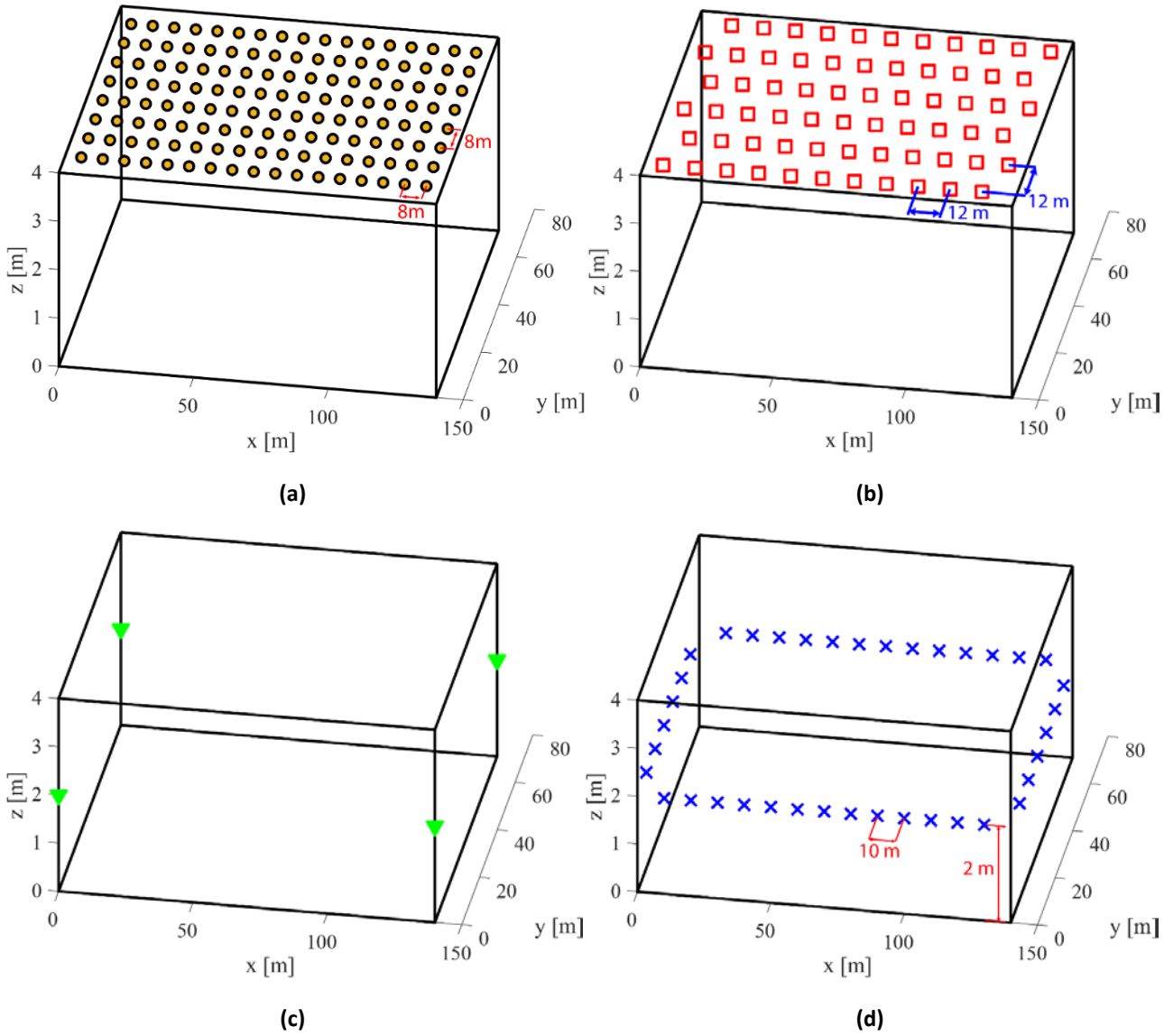


Figure 5-4 Generic indoor private network deployment scenario, where (a) LiFi APs, (b) Wi-Fi APs, (c) 5GNR-I (3.5 GHz) gNBs, and (d) 5GNR-II (26 GHz) gNBs are depicted.

It is important to note that random orientation effect at the mobile terminal is omitted in LiFi optical cell boundary calculations, where the AP and UE are assumed to be directed towards each other in downward and upward directions, respectively without loss of generality. After the combination of (10) in [35] and the above equation along with basic algebra, the optical cell radius expression could be obtained by

$$r = (H_{\text{LAP}} - H_{\text{UE}})\sqrt{4^{\frac{1}{L_S+3}} - 1}$$

where the height of the LiFi AP and distance of the mobile terminal from the floor are given by H_{LAP} and H_{UE}, respectively. The entire set of simulation parameters used in our scenario and their detailed descriptions are provided in Table 5-4.

As far as the coverage of the LiFi AP is concerned, the coverage radius of an optical cell could be calculated by using the expression (4) and Table 5-4 as $r \approx 2.06$ m. For the sake of simplicity in our analyses, a cell radius of 2 m is adopted for the LiFi AP deployment, which is also depicted in Figure 5-5.

Table 5-4: Simulation Parameters

Parameter	Description	Value
n_{LAP}	Number of LiFi APs	136 (generic) 351 (dense)
n_{WAP}	Number of Wi-Fi APs	66 (generic) 102 (dense)
n_{GAP1}	Number of 5GNR-I APs	4 (generic) 8 (dense)
n_{GAP2}	Number of 5GNR-II APs	38 (generic) 66 (dense)
x_{LAP}	Locations of LiFi APs in x-axis	6:8:134 m (generic) 5:5:135 m (dense)
y_{LAP}	Locations of LiFi APs in y-axis	7:8:63 m (generic) 5:5:65 m (dense)
z_{LAP}	Locations of LiFi APs in z-axis	4 m
z_{WAP}	Locations of Wi-Fi APs in z-axis	4 m
z_{GAP1}	Locations of 5GNR-I APs in z-axis	2 m
z_{GAP2}	Locations of 5GNR-II APs in z-axis	2 m
A	Area of the PD	1 cm^2
$\Phi_{1/2}$	Semi-angle of half power of the LED	60°
$\Psi_{1/2}$	FoV of the PD	85°
H_{LAP}	Height of the LiFi AP	4 m
H_{UE}	Height of the mobile terminal	1.44 m

Similarly, for Wi-Fi APs, the rule of thumb in Wi-Fi deployment practice is to have approximately one AP per in a 150 m^{2}area [36], [37]. Therefore, the coverage of the Wi-Fi APs is assumed to be confined within a square with the edge value of $a \approx 12.25$ m, where, for calculation simplicity, the effective value is assumed to be 12 m (refer to Figure 5-5). For 26 GHz mmWave 5GNR gNBs, it has been reported in [38], [39], that the shadow fading and the LoS probability are correlated in the spatial domain with the correlation distance value of 10 m, where this value is adopted as the gNB separation distance in our simulations.

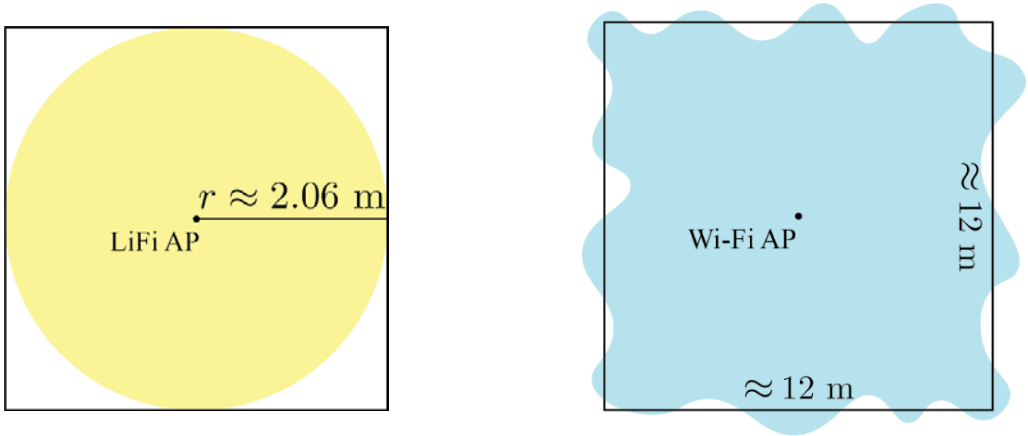


Figure 5-5 Coverage area assumptions for LiFi (left) and Wi-Fi (right) APs.

Lastly, for the 3.5 GHz band 5GNR gNBs, the APs are assumed to be located at the corners of the considered geometry without loss of generality. It is important to note that the deployment of the LiFi APs, Wi-Fi APs, and the 5GNR gNBs are distributed symmetrically within the given geometry to cover the entire structure, as well as to achieve sufficient data rate to reach the area capacity objective.

It is also important to emphasize that one of the main goals of our simulations is to investigate the achievable network capacity with respect to the AP density. Therefore, two deployment scenarios, being generic (with the rule of thumb cellular overlap ratio) and dense (with less APs separation in between) are considered. As depicted in the figures, a total of 136 LiFi APs, 66 Wi-Fi APs, 4 5GNR-I gNBs (3.5 GHz) and 38 mmWave 5GNR-II gNBs (26 GHz) are deployed in the generic scenario.

Unlike the generic deployment scenario, the network densification is utilized in the dense scenario to address the potentially high number of users and resource demand. Accordingly, the separation distance for LiFi, Wi-Fi and 5GNR-II (26 GHz) APs is reduced by 37.5%. Whereas, for 5GNR-I (3.5 GHz), this reduction was 50% in both x and y axes, for the sake of deployment simplicity [33]. Therefore, a total of 351 LiFi APs, 102 Wi-Fi APs, 8 5GNR-I gNBs (3.5 GHz) and 66 5GNR-II gNBs (26 GHz) are deployed on the dense deployment scenario. The geometrical details of the dense deployment for the adopted WATs are provided in Figure 5-6(a), (b), (c) and (d).

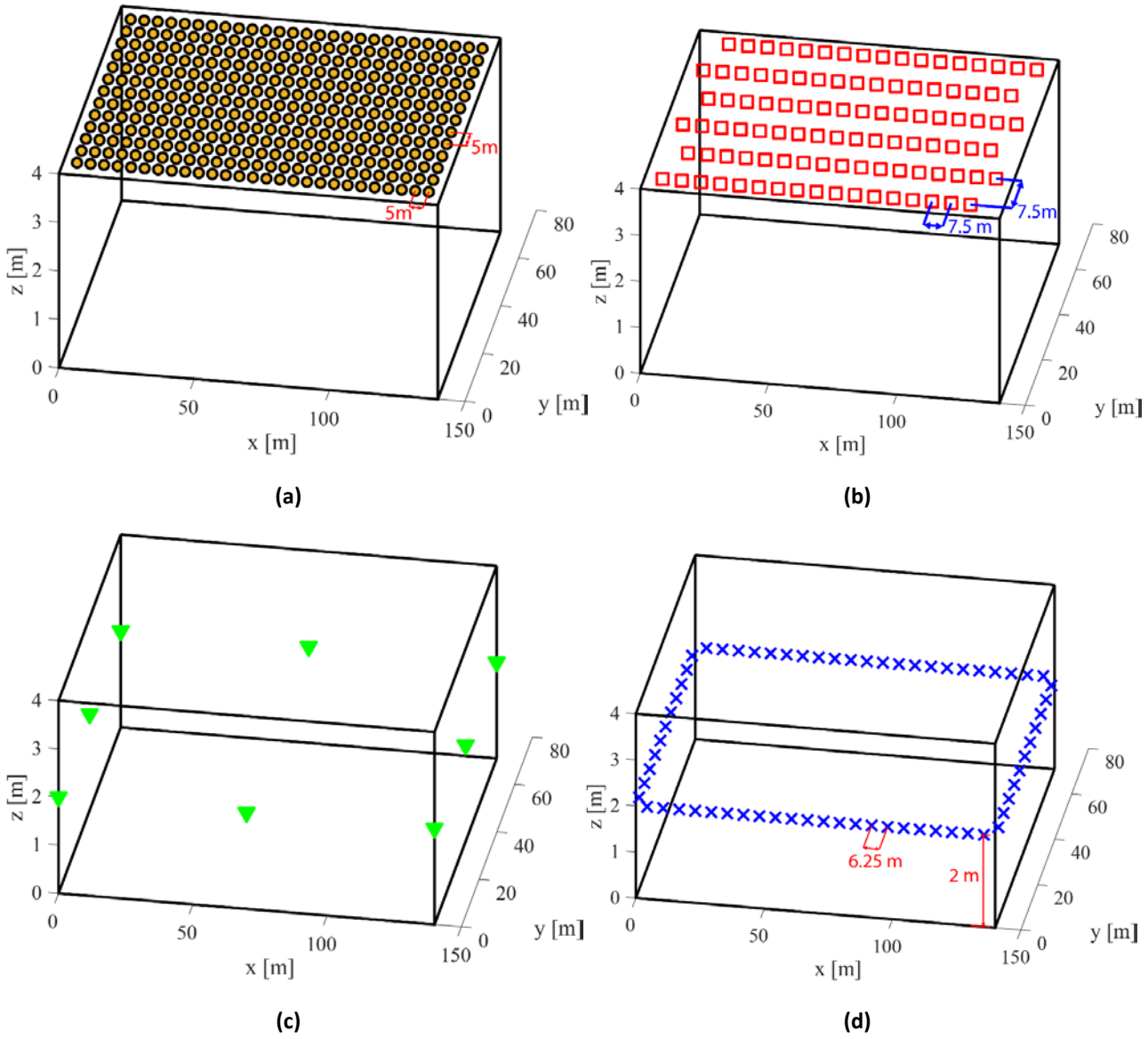


Figure 5-6 Dense indoor private network deployment scenario, (a) LiFi APs, (b) Wi-Fi APs, (c) 5GNR-I (3.5 GHz) gNBs, and (d) 5GNR-II (26 GHz) gNBs.

5.3.1 5G/Wi-Fi/LiFi WATs

In the scope of the 5G-CLARITY multi-WAT framework, the coexistence of hybrid indoor RF and LiFi networks is prioritized to enhance the aggregate area capacity of the system with the aid of 5GNR, Wi-Fi and LiFi networks. In the simulation platform, 3.5 GHz mid-band and 26 GHz high band 5GNR APs are chosen, where both legacy 4G and NR millimetre wave bands could also be supported depending on the scenario. For Wi-Fi, the 5GHz band for 802.11ac standard is adopted. Lastly, for LiFi transmission, the 802.11bb structure is chosen. In the PHY layer, all three systems use orthogonal frequency division multiplexing (OFDM) in both the UL and DL directions via single-input-single-output (SISO) to comply with the standards.

In Wi-Fi, an OFDM symbol consists of 256 subcarriers, with only 242 of them being available per physical resource unit (PRU), as a 80 MHz available bandwidth (4×20 MHz) is chosen. Similarly, in LiFi, there are 242 available subcarriers. Unlike typical RF-based communications systems, LiFi operates based on the intensity-modulation-direct-detection (IM/DD) transmission technique. In IM/DD, the transmit signals are encoded onto instantaneous transmit optical intensity values, which are not visible to the human eye due to

the frequent changes ($\gg 100$ Hz) in the light intensity. At the receiver, simple off-the-shelf photodiodes (PDs) can detect the subtle changes in the received optical power and recover the transmitted information signal. Since the information bearing optical intensity value is intrinsically restricted to be real and positive valued,

$$I(t) \geq 0, \forall t \in \mathbb{R},$$

where t denotes the time, LiFi will require modulation-level alteration to comply with its RF counterparts. Hence, the consequences of using the IM/DD transmission method are detailed as follows:

- Unlike conventional RF transmission, where the information is carried by the changes in electrical field, the information is carried by the optical intensity in LiFi systems. The change of domains, electrical-to-optical and vice versa, creates a fundamental difference for LiFi as the capacity and error performance are directly dependent on the *electrical signal power*, whereas the eye safety and transmission power constraints are imposed on the *optical signal power*. It is also important to note that for a signal, $x(t)$, the electrical and optical domain power values are calculated by $\mathrm{E}\{|x(t)|^2\}$ and $\mathrm{E}\{|x(t)|\}$, respectively.
- The information bearing signal cannot be modulated by using its frequency and phase properties.
- The transmitted signal must always be in the baseband, which also means that the single sideband (SSB) transmission cannot be realized by LiFi. Hence, the effective bandwidth usage of the LiFi becomes halved compared to conventional RF systems as depicted in Figure 5-7.

The most straightforward way to obtain a real-valued sample with OFDM-based systems in the current LiFi literature is with frequency domain Hermitian symmetry, in other words, the conjugate symmetry property of the Fourier Transform (FT) [40]. Accordingly, the even and odd symmetric real and imaginary parts, respectively, of a frequency (time) domain signal will ensure a real and bipolar valued output in the time (frequency) domain. Again, IM/DD based incoherent LiFi systems are not able to transmit bipolar signals. Another simple solution for the bipolarity issue is to introduce a direct current (DC) bias to the time domain OFDM symbol, which is known as the DC-biased optical OFDM (DCO-OFDM) in the LiFi literature [41]. The rest of the transmission is relatively straightforward. The time domain unipolar samples are double clipped to ensure the signal limits are within the dynamic range of the LED. Then, a suitable cyclic prefix is added to the frame and a digital-to-analog conversion is imposed on the signal. As the final step at the Tx, the LED transfers real and unipolar valued time electrical domain time samples into the optical domain. At the Rx, all the Tx operations are reversed to obtain the transmit information. It is important to note that, due to the Hermitian symmetry requirement in LiFi systems, half of the subcarriers are not bearing useful information. Therefore, in the simulator, an OFDM frame with 512 total subcarriers are chosen in the LiFi system to ensure 242 available subcarriers in the chosen bandwidth.

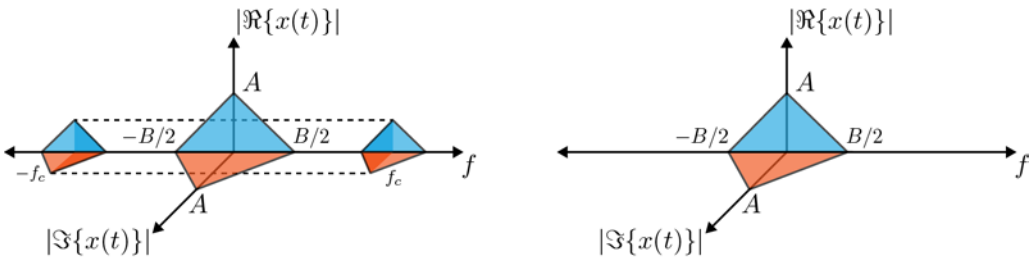


Figure 5-7 The frequency domain representation of information bearing signal for conventional RF (left) and IM/DD-based incoherent LiFi (right)

The available bands for the adopted WATs are constrained based on local regulations and the operators. Thus, the Ofcom regulations for the UK bands for Wi-Fi, 5GNR-I (3.5 GHz) and 5GNR-II (26 GHz) are adopted in the simulations. For LiFi, the 20 MHz channel, adopted as a practical value by both the academia and industry, is utilized [30] [42]. It is important to note that the optical band, which includes both the VL and IR bands, is approximately 2600 times wider than the entire RF spectrum [43], however a typical channel bandwidth value is several kilohertz up to tens of megahertz for a practical system implementation in LiFi. The main limiting factor behind this is the slow response rate of the transmit LED opto-electronic characteristics. The most efficient production method for the white LEDs is using blue LED chips as the base and yellow phosphor coating to obtain white light. Nonetheless, the yellow colour conversion reduces the frequency responsivity of the white LEDs in the region of $1-2$ MHz [43]. It has been reported in [44] that filtering the phosphor component could yield a bandwidth up to 20 MHz for an off-the-shelf LED. It is also important to note that the current record for the laboratory achieved bandwidth with a single laser diode (LD) is 6.8 GHz [45], which indicates the volatility of the practical limitations. Moreover, in the 5GNR mmWave band, the actual band is fractioned into small chunks to serve multiple MNOs. Therefore, two bandwidth values; (i) conservative, which considers the practical limitations and regulations, and (ii) opportunistic, which assumes the greedy usage of the available bandwidth, were adopted to capture the practical constrains in our simulations. Accordingly, in the conservative set for LiFi, Wi-Fi, 5GNR-I (3.5 GHz) and 5GNR-II (26 GHz) are assumed to be allocated 20, 80, and 800 MHz, respectively. On the contrary, the opportunistic set for LiFi, Wi-Fi, 5GNR-I (3.5 GHz) and 5GNR-II (26 GHz) assumes the bandwidth values of 6.8 GHz, 320 MHz, 400 MHz and 1.792 GHz, respectively. The full set of operation frequencies and adopted bandwidth sets are given in Table 5-5.

Table 5-5: Adopted WAT Operation Frequencies

Parameter	Description	Values	Conservative Values	Opportunistic Values
λ_{LAP}	Operation wavelength region of LiFi APs	337 THz (299.79-374.74 THz), where the visible band is optional		
f_{WAP}	Operation frequency of Wi-Fi APs	5 GHz (5.15-5.725 GHz)		
f_{GAP1}	Operation frequency of type-I 5GNR APs	3.5 GHz (3.4-3.8 GHz)		
f_{GAP2}	Operation frequency of type-II 5GNR APs	26 GHz (24.5-26.5 GHz)		
B_{LAP}	Bandwidth of the LiFi APs		20 MHz	6.8 GHz [45]
B_{WAP}	Bandwidth of the Wi-Fi APs		80 MHz	320 MHz
B_{GAP1}	Bandwidth of the type-I 5GNR APs		80 MHz	400 MHz
B_{GAP2}	Bandwidth of the type-II 5GNR APs		800 MHz	1.792 GHz [46]

5.3.2 Multi-WAT aggregation and resource scheduling

To capture the multi-WAT aggregation at the mobile terminal, we simulate the case that a UE/CPE could connect to 5GNR, Wi-Fi and LiFi networks at the same time. Accordingly, three wireless WATs, both radio and optical, could be nested in a single device. Depending on the weights associated with each WAT, the mobile UE/CPE will be able to decode the transmit information by using the associated front-end structure for the UL/DL direction. Similarly, the mobile UE/CPE could be able to choose the suitable access technology, based on the availability, signal power, number of active users in a certain WAT, etc., for the UL direction to transmit information. This practically means that three WAT interfaces must be implemented within the same mobile UE/CPE, where UE/CPE could also request a WAT switch or weight revision to meet the system key performance indicators.

The resource scheduling policy is a vital aspect of radio resource management (RRM) for the designed 5G network as interference becomes the biggest issue as the density of the network increases. Moreover, delay, jitter, fairness, and loss rate performance could easily be enhanced with the utilization of an appropriate scheduling algorithm. Unlike wired communication systems, the RF based Wi-Fi, LTE and 5GNR networks need to take the location-based channel conditions into consideration. In addition, in 5G-CLARITY, propagation and channel characteristics of optical signals need to be carefully included into these calculations, since LiFi is utilized as the key enabler. Accordingly, multiple resource allocation algorithms are realized in the proposed simulator to coordinate both the intra and inter cell interferences, which opportunistically utilizes the network density as the resource. In the simulator, full buffer traffic, round robin (RR) and frequency-selective proportional fair schedulers (PFS) are implemented within the real-time traffic simulation environment to enhance the performance of multi-WAT networks. It important to emphasize that the PFS algorithm is a widely adopted technique as it provides a good compromise between the maximum throughput and user fairness. PFS achieves this task by exploiting the diversity and game-theoretic equilibrium in a fading channel environment. The resource scheduling is achieved by the following steps in the simulation environment. Firstly, SINR per UE is computed for all three WATs, namely 5GNR, Wi-Fi and LiFi. Then, each UE is associated with a CQI based on SINR, which is also one-to-one mapped with the AMC lookup table to determine modulation order and coding rate values. The PRB grouping and UE-PRB assignment is handled by the chosen scheduling function for each UE, with the achievable rate per allocated PRB group possibly being downgraded to provide 10% block error ratio (BLER) to realize common MCS assignment. Lastly, the performance for each UE is evaluated by using the KPIs detailed in the following section.

5.3.3 KPIs used for evaluation

The KPIs can provide an important set of data points and patterns relating to the network health, availability, performance, and efficiency to network service providers and system designers. Accordingly, in our capacity simulations, the following KPIs are adopted:

- Achievable capacity, for each WAT independently and aggregated value.
- System area traffic capacity, in b/s/m^2, for each WAT independently and aggregated value.

5.4 System level simulation results

In this section, the achievable system capacity for two deployment scenarios, generic and dense, are obtained and presented for two sets of bandwidth values, namely the conservative and opportunistic approach. Note that the capacity performance simulations are conducted for a mobile terminal.

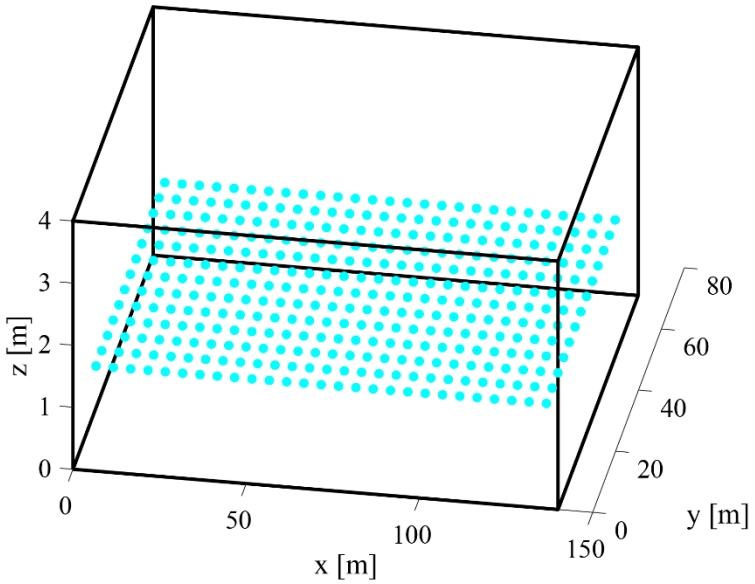


Figure 5-8 Mobile UE locations considered in our simulations.

Accordingly, the UE is assumed to move within a plane at the height of 1.44 m from the floor to capture any user and a handheld device movement characteristic. The UE locations under interest are depicted in Figure 5-8, where the mobile terminal is located uniformly inside the room with x and y dimension ranges of [5 135] m and [5 65] m, respectively, and a separation of 5 m at the height of 1.44 m from the ground floor. Hence, the area traffic capacity results are obtained for 351 data points in our simulations.

The reported achievable data rate values in the real-life applications are vastly limited by the practical limitations, which could be superseded over the years with advancements in technology. Therefore, in our simulations, we will be considering the Shannon-Hartley upper bound formula to project the system limits. It is also important to note that, due to the very high AP density of both generic and dense deployment scenarios, the given network becomes interference limited rather than noise limited. Consequently, the high AP density leads to a system where the noise power is negligible compared to the high received interference power. Hence, the signal-to-interference ratio (SIR) metric could be used to calculate the achievable capacity value without loss of generality. The Shannon-Hartley system capacity expression for the chosen WAT with the $\boldsymbol{i}^{\text{th}}$ PL model, where $\boldsymbol{i} \in \{\mathbf{1}, \mathbf{2}, \mathbf{3}\}$, could be given as follows:

$$\boldsymbol{C_i(x,y) = B\,log_2(1 + SIR_i(x,y))}$$

where the parameter $\boldsymbol{B}$ denotes the modulation bandwidth for a given WAT, and the SIR for a UE whose location could be presented by $(\boldsymbol{x}, \boldsymbol{y})$ when only the $\boldsymbol{i}^{\text{th}}$ WAT is considered can also be calculated by:

$$\text{SIR}_i(x,y) = \frac{P^i_{\text{sig}}(x,y)}{P^i_{\text{int}}(x,y)} = \frac{P^{(S\to R)}_{\text{r},i}}{\sum_{(S'\neq S)} P^{(S'\to R)}_{\text{r},i}} = \frac{10^{-\text{PL}_i(S,R)/10}}{\sum_{(S'\neq S)} 10^{-\text{PL}_i(S',R)/10}}$$

where the received signal power at the receiver, R, which is emerged due to the source, S, of the i^{th} WAT is given by $P^{(S\to R)}_{\text{r},i}$. Similarly, the received interference power from the source S' of the i^{th} WAT is denoted by $P^{(S'\to R)}_{\text{r},i}$. It is important to note that the source, S, and receiver, R, sets are different for each WAT, where the indexes are dropped for the sake of readability. The received signal power is calculated by finding the AP, which yields the lowest PL value, within a given WAT APs set. This AP is referred to as the "connected AP" in the simulations. The rest of the APs within are considered as interferer APs. Please also note from the above expression that the transmit power of all the APs among various WATs is assumed to be 1 W without loss of generality. As LiFi operates with an IM/DD based incoherent transmission mode, the effective bandwidth

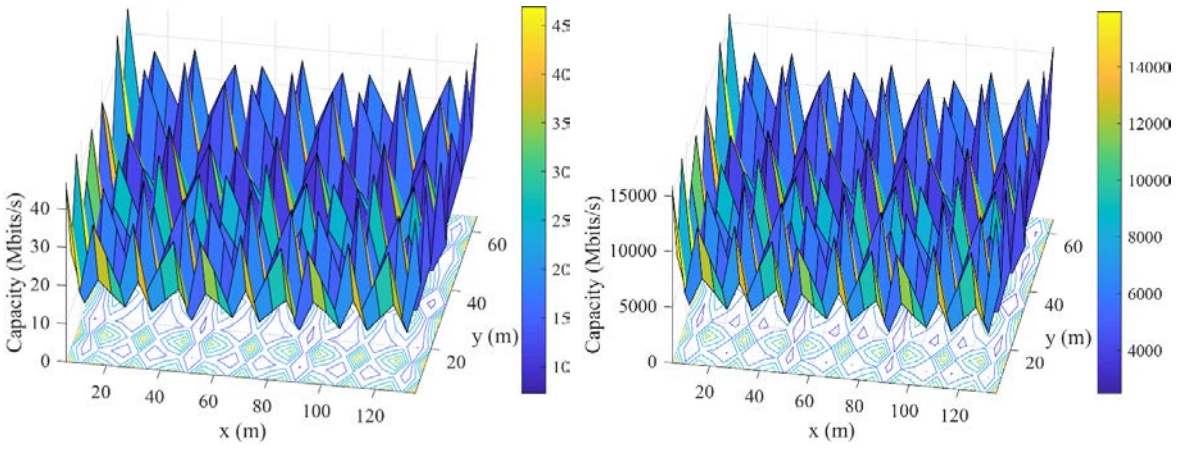
value for LiFi systems becomes, $B_{\text{eff}} = B_{\text{LiFi}}/2$ as reported in [47]. Whereas the effective bandwidth for all the other WATs is equal to the available bandwidth values. Then, the aggregate system area spectral efficiency value could be calculated by,

$$C_A = \frac{\bar{C}}{A_{\text{total}}} = \frac{\sum_{\forall k}\sum_{\forall x}\sum_{\forall y} C_k(x,y)}{n_x n_y A_{\text{total}}} \text{ bps/m}^2$$

where the number of mobile terminal data points in x and y axes are denoted by n_x and n_y, respectively. Moreover, A_{total} is the total area for the UE mobility plane as depicted in Figure 5-8.

5.4.1.1 Generic deployment scenario

In this subsection, the capacity simulation results for LiFi, Wi-Fi, 5GNR-I and 5GNR-II will be given for the generic deployment scenario with both the conservative and opportunistic bandwidth values. The capacity results for the LiFi network for each data point associated with the mobile terminal locations are provided for the conservative bandwidth in Figure 5-9(a). As can be seen from the figure, the $40 - 45$ Mbps of spikes correspond to the closest points to the LiFi APs, where the receiver SIR is maximum. The troughs are the locations where the interference power is more comparable to the received signal power. It is also important to note that, although the results for each UE location point are obtained after 10000 Monte Carlo trials, both peaks and troughs are not identical among each other due to the random orientation effects. Accordingly, the mean achievable capacity for LiFi becomes 20.92 Mbps with the maximum and minimum values of 46.91 Mbps and 7.3 Mbps, respectively. Also, the standard deviation of the LiFi capacity also becomes approximately 10 dB. The area capacity contribution, which comes solely from the LiFi APs for the conservative approach, becomes 19.8 Mbps/m^2. This value is calculated by the interpolation of the capacity simulation results, which are obtained by the measurement squares with the area of 5×5 m^2, into a finer grid. The capacity value within each unitary square, with the area of 1×1 m^2, is assumed to be fixed that could be calculated at the centre of each square. Then, the interpolated value for each unitary square is summed, and the result is divided by the total surface area of the measurement surface, as depicted in Figure 5-8. Similarly, the capacity results for a LiFi network under the opportunistic bandwidth values are given in Figure 5-9(b). As can be seen from the figure, a similar capacity pattern also emerges in the bandwidth opportunistic scenario as expected. The mean, maximum and minimum values for the achieved capacity becomes 7.12 Gbps, 16 Gbps and 2.48 Gbps, respectively, with the standard deviation of 35.30 dB. Furthermore, the area traffic capacity in the opportunistic scenario becomes 6.74 Gbps/m^2.

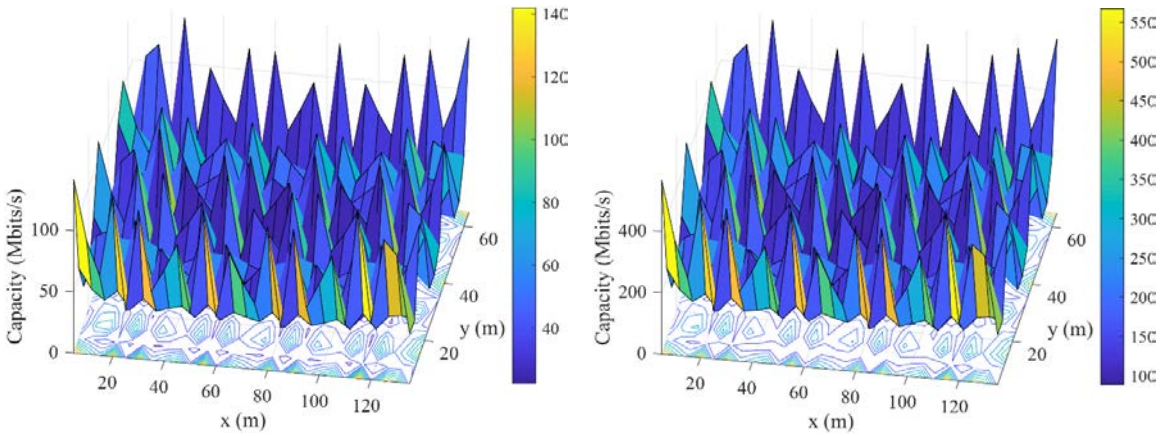


(a) **(b)**

Figure 5-9. The capacity for LiFi APs in the generic scenario, (a) conservative, and (b) opportunistic approaches.

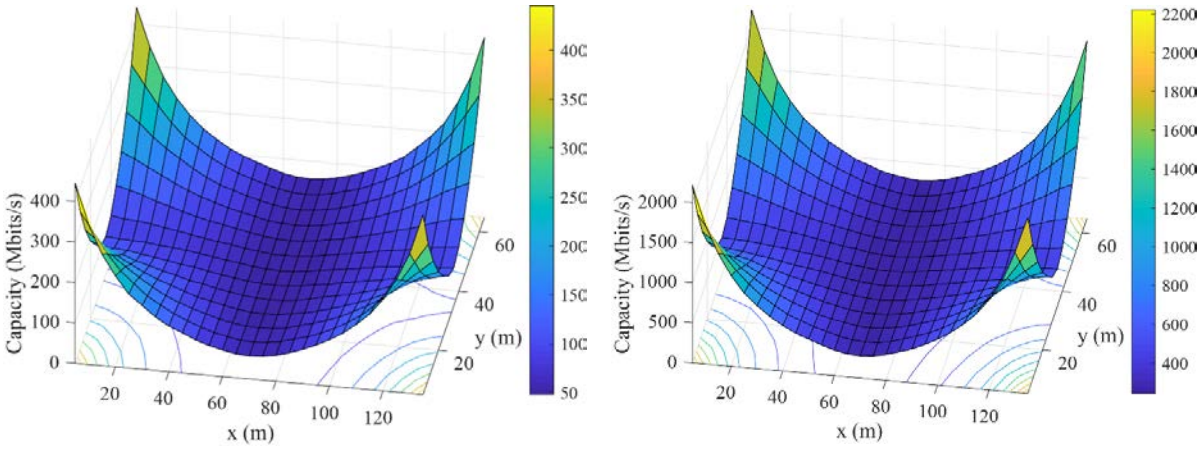
In Figure 5-10(a) and (b), the capacity calculations are given for the 5GHz Wi-Fi deployment when the conservative and opportunistic bandwidth values are adopted, respectively. As can be seen from Figure 5-16(a), the spiky nature of the capacity results resembles the LiFi results as the AP density of both LiFi and Wi-Fi are significantly higher compared to other systems. The numerical relationship between the number of APs for each WAT could be represented as follows: $N_{\text{LAP}} > N_{\text{WAP}} \gg N_{\text{GAP1}} \gg N_{\text{GAP2}}$, where the parameter N represents the number of APs for the associated WAT. The main difference between the LiFi and Wi-Fi results is the magnitude of the peaks in the Wi-Fi plots, which are around $120 - 140$ Mbps, which is almost more than three times that of the LiFi results. The main reason behind this is the abundance of the bandwidth in Wi-Fi compared to LiFi in the conservative scenario. The mean, maximum and minimum values for the achievable capacity values in a Wi-Fi deployment becomes 51.28 Mbps, 141.89 Mbps and 22.46 Mbps, respectively, with the standard deviation of 14.31 dB. Also, the area traffic capacity for Wi-Fi is calculated to be 48.18 Mbps/m^2 for the given conservative bandwidth value. The opportunistic results for a Wi-Fi deployment are presented in Figure 5-10(b), where the pattern stayed the same with the increase in magnitudes values. Accordingly, the opportunistic approach in Wi-Fi yields the mean, maximum and minimum capacity values of 205.23 Mbps, 567.28 Mbps and 89.14 Mbps, respectively. Moreover, the standard deviation in the capacity simulations also becomes 20.34 dB**.** The area traffic capacity value for the opportunistic scenario is given by 192.83 Mbps/m^2.

In Figure 5-11(a) and (b), 3.5 GHz (mid-band) 5GNR-I capacity simulation results for conservative and opportunistic scenarios are given, respectively. As depicted in Figure 5-11(a) the peaks of the 5GNR-I occur at the corners of the room as, which is parallel with the AP deployment. The mean, maximum and minimum capacity values could be given by 117.09 Mbps, 445.4 Mbps and 48.14 Mbps, respectively. The main reason behind the relatively high mean capacity value is the lack of inter-cell-interference caused by the large separation among 5GNR-I APs. Also, the standard deviation of the obtained capacity data points becomes 18.71 dB. Thus, the obtained values yield the area traffic capacity of 108.64 Mbps/m^2 for the conservative bandwidth scenario. In Figure 5-11(b), the same simulation is repeated for the bandwidth opportunistic approach. Hence, the mean, maximum and minimum capacity values of 584.94 Mbps, 2.22 Gbps and 242.3 Mbps are obtained, respectively, along with the standard deviation of 25.68 dB. Moreover, the area traffic capacity value of 542.82 Mbps/$\text{m}^2$is achieved by the 5GNR-I APs for the given configuration.



(a) **(b)**

Figure 5-10. The capacity for Wi-Fi APs in the generic scenario, (a) conservative, and (b) opportunistic approaches

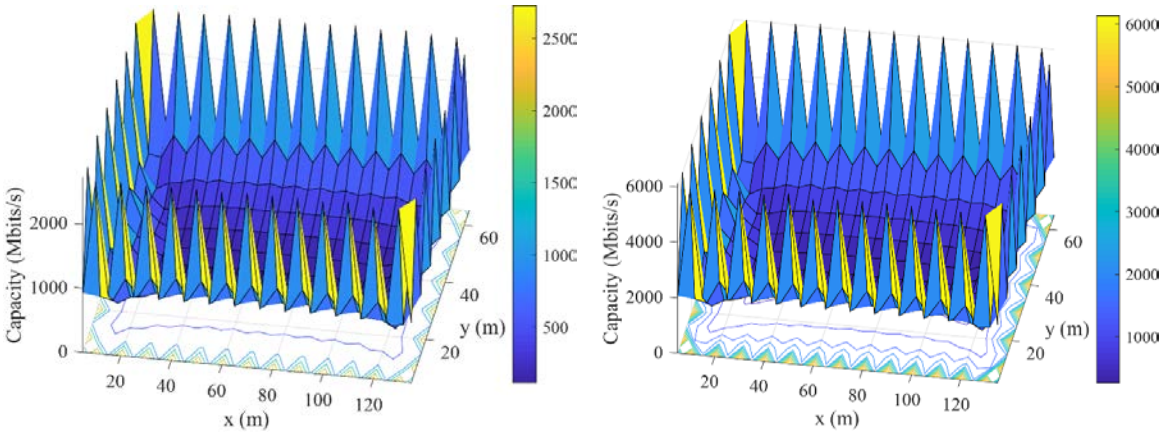


(a) **(b)**

Figure 5-11. The capacity for 5GNR-I (3.5 GHz) APs in the generic scenario, (a) conservative, and (b) opportunistic approaches

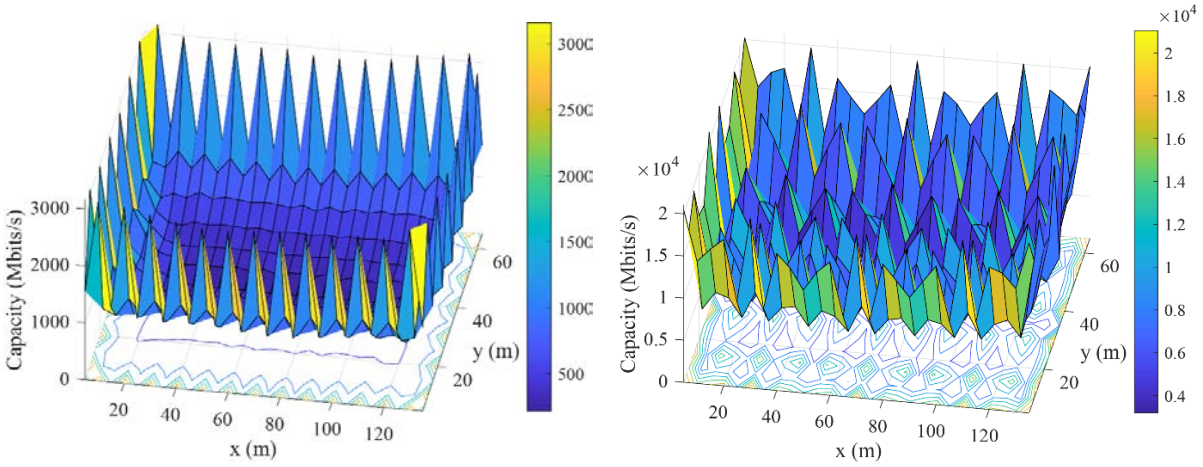
The capacity results for the high band mmWave 5GNR-II APs are given for conservative and opportunistic bandwidth values in Figure 5-12(a) and (b), respectively. As can be seen from Figure 5-18(a), the peaks are concentrated around the edges of the room as the mmWave gNBs are placed to the side walls. The maximum capacity value peaks as high as 2.5 Gbps as the inter-cell-interference is significantly lower in the 26 GHz band due to very high PL compared to 3.5 and 5 GHz systems. Accordingly, the mean, maximum and minimum capacity values for the conservative bandwidth scenario becomes 756.42 Mbps, 2.73 Gbps and 117.43 Mbps, respectively, with the standard deviation of 28.63 dB. Lastly, the area traffic capacity of only 5GNR-II APs is given by 627.33 Mbps/m^2 for the given configuration. Similarly, Figure 5-12(b) depicts the capacity simulation results when the opportunistic bandwidth value is adopted. In the given scenario, the average capacity of 1.70 Gbps is obtained with the maximum and minimum values of 6.13 Gbps and 263.31 Mbps, respectively. Moreover, the standard deviation of the obtained data is 32.14 dB. Hence, these obtained values yield the achievable area traffic capacity of 1.4 Gbps/m^2.

Finally, the aggregated capacity maps using all the WATs, is provided for conservative and opportunistic bandwidth values in Figure 5-13(a) and (b), respectively. It is important to note from Figure 5-13(a) that the aggregate capacity pattern for the conservative scenario is significantly dominated by the mmWave 5GNR capacity values due to the significantly higher contribution of 26 GHz band.



(a) **(b)**

Figure 5-12. The capacity for 5GNR (26 GHz) APs in the generic scenario, (a) conservative and (b) opportunistic approaches.



(a) (b)

Figure 5-13. The aggregate capacity (LiFi/Wi-Fi/5GNR-I (3.5 GHz)/5GNR-II (26 GHz)) APs in the generic scenario, (a) conservative, and (b) opportunistic approaches.

The average aggregated capacity of 945.71 Mbps is achieved for the given scenario with the maximum and minimum values of 3.16 Gbps and 219.65 Mbps, respectively. The standard deviation of the aggregate capacity also becomes 28.89 dB. Note that the area traffic capacity of 803.95 Mbps/m^2 is obtained with the conservative bandwidth values selection, which is more than 60% higher than the 500 Mbps/m^2 target. In Figure 5-13(b), the results are obtained for the opportunistic bandwidth values set. As can be seen from figure, the aggregate capacity is influenced by the LiFi APs, unlike the conservative configuration, as the highest bandwidth among all WATs are obtained for LiFi in the bandwidth opportunistic case. Accordingly, the average capacity in the given scenario becomes 9.6 Gbps with the maximum and minimum values of 21 Gbps and 3.19 Gbps, respectively, along with the standard deviation of 36.37 dB. It also important to emphasize that the average capacity in the given opportunistic approach becomes 8.88 Gbps/m^2, which exceeds the objective more than 1675%.

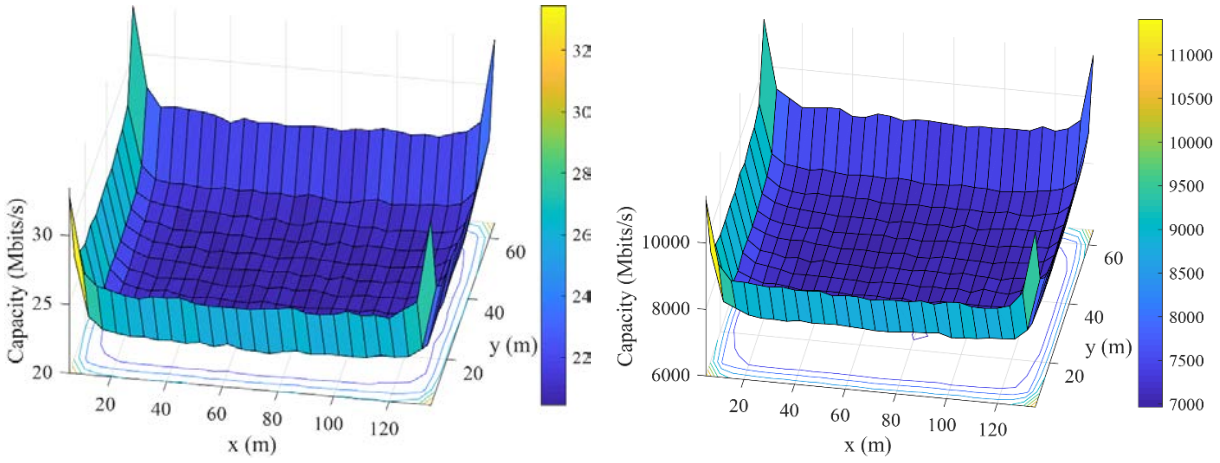
5.4.1.2 Dense deployment scenario

In this subsection, the effect of the network densification on the achievable capacity results are investigated by the Monte Carlo capacity simulations. The capacity plots will be presented for each WAT as well as the aggregate value. The capacity plots of the dense deployment with the conservative and opportunistic bandwidth values are plotted in Figure 5-14(a) and (b), respectively. As can be seen from Figure 5-14(a), the peak values only occur at the corner of the room, whereas the rest of the measurement plane yields capacity values around 20 Mbps. Accordingly, the average achievable capacity value for the given scenario becomes 22.36 Mbps with maximum and minimum values of 33.45 Mbps and 20.45 Mbps, respectively. Moreover, the standard deviation of the capacity simulation results is 3.82 dB, which yields a significantly smooth curvature compared to the generic deployment scenario. Thus, the area traffic capacity for the given configuration also becomes 21.81 Mbps/m^2. Approximately 10% increase in area traffic capacity compared to the more than 158% increase in the number of APs, proves the capacity saturation, which is induced by the enhanced interference as a result of AP densification. In Figure 5-14(b), the same simulation has been repeated for the opportunistic bandwidth values, where the average capacity for LiFi network became 7.6 Gbps with the maximum, minimum and standard deviation values of 11.41 Gbps, 6.97 Gbps and 29.15 dB.

The capacity investigation of Wi-Fi in the dense deployment scenario is given with conservative and opportunistic bandwidth values in Figure 5-15(a) and (b), respectively. As can be seen from Figure 5-15(a) the smoothness of the figure has increased compared to the generic deployment results. The reason for this

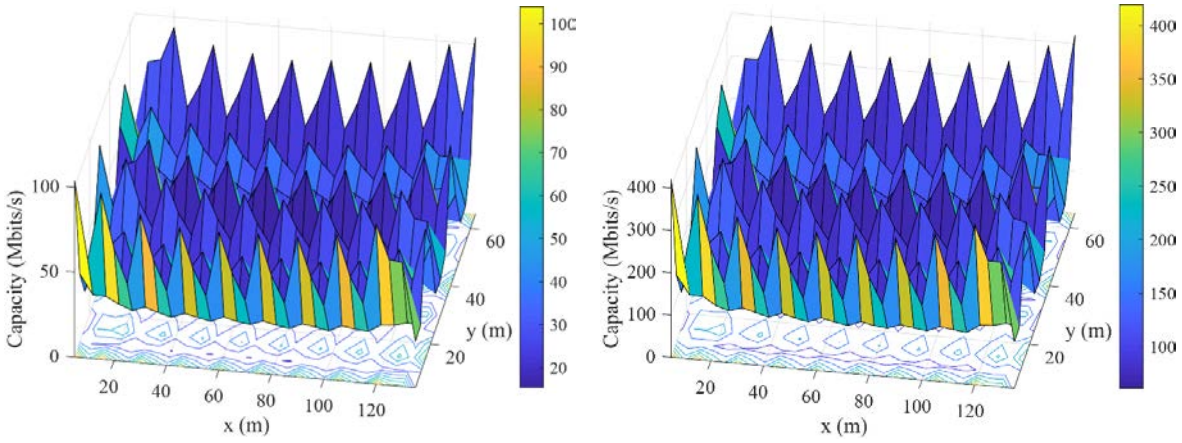
is the enhanced density of the APs, which yields more homogeneously distributed received signal power on the UE mobility surface. The average capacity in the bandwidth conservative approach is 38.85 Mbps with the maximum, minimum and standard deviation values of 103.93 Mbps, 15.37 Mbps and 12.76 dB, respectively. The network densifications yield the area traffic capacity of 36.38 Mbps/ m^2 when the conservative bandwidth value is adopted. It is important to emphasize that the area capacity value reduced more than 24% compared to the generic deployment scenario. Again, the densification related inter-cell-interference is the main reason behind the capacity drop. The result for the opportunistic bandwidth scenario is given in Figure 5-15(b). Accordingly, the average achievable capacity value becomes 155.46 Mbps, where the maximum and minimum values and standard deviation of the capacity simulations results are given by 419.23 Mbps, 61.33 Mbps and 18.79 dB, respectively. The achievable area capacity for the dense Wi-Fi deployment with the opportunistic bandwidth value becomes 145.57 Mbps/m^2.

In Figure 5-16(a) and (b), the capacity simulation results for 3.5 GHz 5GNR under the dense deployment scenario are given for the conservative and opportunistic bandwidth values, respectively. As depicted in Figure 5-16(a), the average, maximum, minimum capacity values for the conservative bandwidth scenario becomes 91.03, 367.6 and 35.47 Mbps, respectively, with the standard deviation value of 17.82 dB.



(a) **(b)**

Figure 5-14. The capacity for LiFi APs in the dense scenario, (a) conservative and (b) opportunistic approaches

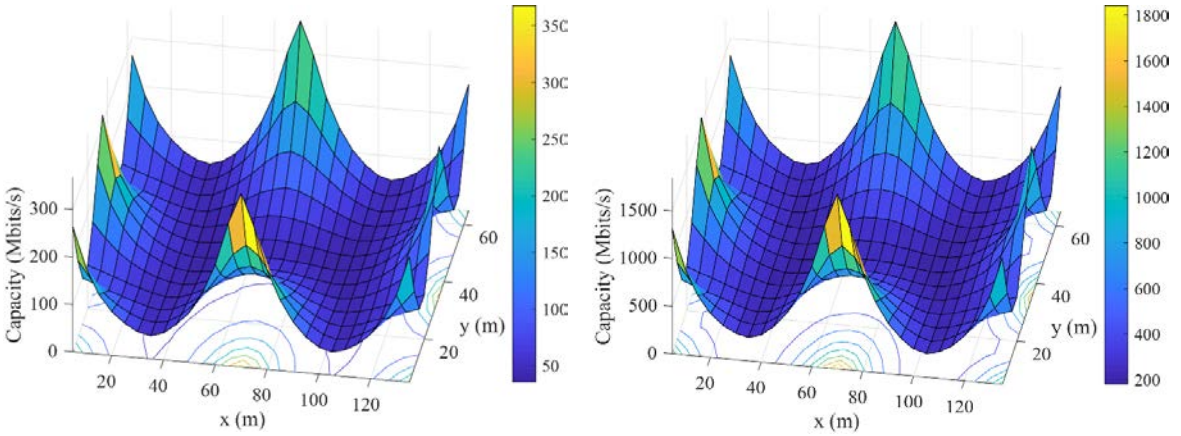


(a) **(b)**

Figure 5-15. The capacity for Wi-Fi APs in the dense scenario, (a) conservative, and (b) opportunistic approaches

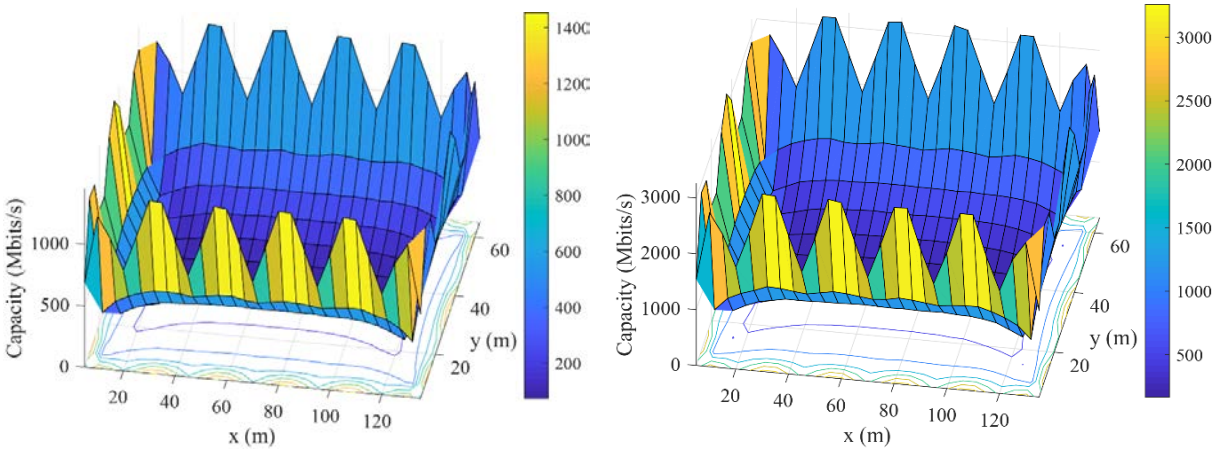
The area traffic capacity in the given configuration also becomes 83.47 Mbps/m^2, which is more than 23% less compared to the generic scenario. Similarly, in Figure 5-16(b), the same simulation is repeated for the opportunistic bandwidth scenario. Accordingly, the average, maximum and minimum capacity values in the given scenario becomes 455.5 Mbps, 1.84 Gbps and 179.4 Mbps, respectively. The standard deviation value of the achievable capacity also becomes 24.81 dB. Furthermore, the area traffic capacity of 417.63 Mbps/m^2, which is also more than 23% less compared to the generic scenario is also obtained. Consequently, we can deduce that the inter-cell-interference yields more than a 23% drop in the area traffic capacity of Wi-Fi deployment when the cell densification is employed.

The dense deployment capacity simulation results for the 26 GHz mmWave AP for the conservative and opportunistic bandwidth sets adopted are given in Figure 5-17(a) and (b), respectively. The mean, maximum and minimum values of the obtained capacity under the conservative bandwidth set becomes 468.8 Mbps, 1.45 Gbps and 74.51 Mbps, respectively, with a standard deviation of 25.87 dB. Moreover, the achievable area traffic capacity for the given scenario becomes 388 Mbps/m^2, which is a reduction more than 38% compared to the generic deployment scenario.



(a) (b)

Figure 5-16. The capacity for 5GNR-I (3.5 GHz) APs in the dense scenario, (a) conservative, and (b) opportunistic approaches.



(a) (b)

Figure 5-17. The capacity for 5GNR-II (26 GHz) APs in the dense scenario, (a) conservative and (b) opportunistic approaches.

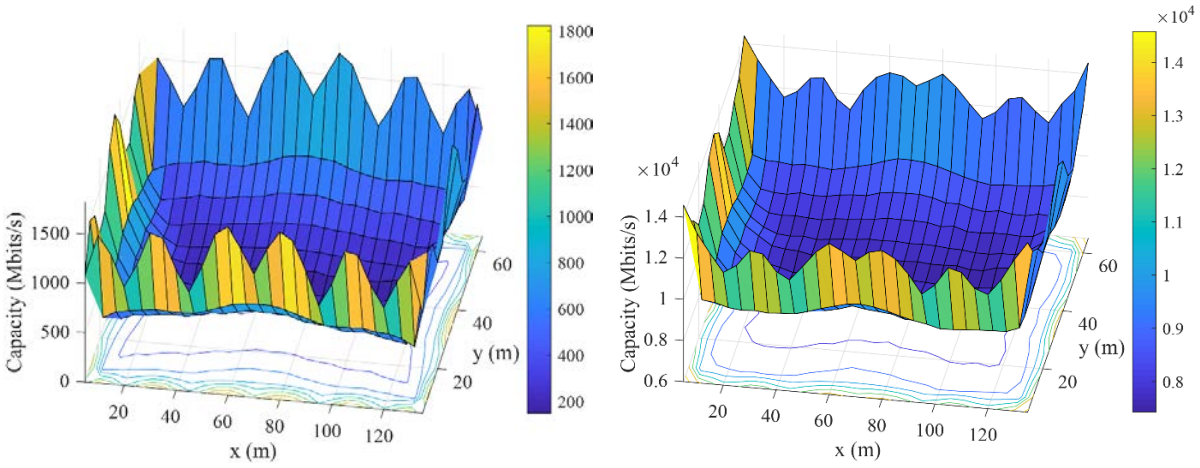


Figure 5-18. The aggregate capacity results for LiFi, Wi-Fi, 5GNR-I (3.5 GHz) and 5GNR-II (26 GHz) APs in the dense scenario with: (a) conservative and (b) opportunistic approaches.

In Figure 5-17, the average, maximum and minimum capacity values are given by 1.05 Gbps, 3.26 Gbps and 166.95 Mbps, respectively, with a standard deviation value of 29.38 dB for the opportunistic bandwidth value. The mentioned results in a given scenario yield the area traffic capacity value of 869.44 Mbps/m^2, which is also more than 38% less compared to the generic scenario. Therefore, we can infer from the results that the area traffic capacity values for the network densification in 26 GHz 5GNR-II yields around 40% loss due to the enhanced interference.

In Figure 5-18(a) and (b), the aggregate capacity simulation results are presented for the dense AP deployment scenario under conservative and opportunistic bandwidth values, respectively. As can be seen from Figure 5-18(a), the mean, maximum and minimum aggregate capacity results become 621.04 Mbps, 1.82 Gbps and 150.4 Mbps, respectively, with a standard deviation of 26.37 dB. Furthermore, the aggregate area traffic value also becomes 529.65 Mbps/m^2, which is more than 34% less than that of the generic scenario. The same simulation scenario with the opportunistic bandwidth value is given in Figure 5-18(b). Accordingly, the average, maximum and minimum capacity values in the given configuration becomes 9.27, 14.59 and 7.42 Gbps, respectively, with a standard deviation of 32.69 dB. Hence, the aggregate area traffic capacity of 8.88 Gbps/m^{2}is achieved in the dense opportunistic scenario, which is less than a 1% reduction compared to the generic scenario. Consequently, the initial objective is exceeded by approximately 6% and 1670% for the dense AP deployment when the conservative and opportunistic bandwidth sets are utilized, respectively.

Finally, the area capacity results are summarized in Table 5-6, where μ_C, $C_{\max}$, $C_{\min}$, σ_C denotes the average, minimum, maximum, standard deviation values of the obtained capacity simulation results, respectively. The achievable area traffic capacity values are also given by parameter C. As can be inferred from the results provided in table, the area traffic capacity in generic deployment is reduced with the cellular densification for Wi-Fi, 5GNR-I, 5GNR-II and aggregated results due to the higher relative increase in interference power compared to the signal power. On the contrary, this is not the case for LiFi since the pathloss of the optical spectra is significantly higher compared to operation frequency RF based WATs, which yields a larger margin for network densification. It is also important to note that the generic deployment parameters, such as cell radius, AP separation and location values are obtained from recommended/typical application scenarios, which are already empirically optimized in terms of the coverage and interference trade-off. Therefore, the further densification on top of this empirically optimized generic RF deployment is not capable of providing higher area traffic capacity values.

Table 5-6: Summary of the Simulated Area Capacity Results.

	Generic Deployment		Dense Deployment	
WAT	**Conservative Bandwidth**	**Opportunistic Bandwidth**	**Conservative Bandwidth**	**Opportunistic Bandwidth**
LiFi	$C = 19.8$ Mbps/m^2 $\mu_C = 20.92$ Mbps $C_{\max} = 46.91$ Mbps $C_{\min} = 7.3$ Mbps $\sigma_C = 10$ dB	$C = 6.74$ Gbps/m^2 $\mu_C = 7.12$ Gbps $C_{\max} = 15.94$ Gbps $C_{\min} = 2.48$ Gbps $\sigma_C = 35.30$ dB	$C = 21.81$ Mbps/m^2 $\mu_C = 22.36$ Mbps $C_{\max} = 33.45$ Mbps $C_{\min} = 20.45$ Mbps $\sigma_C = 3.82$ dB	$C = 7.42$ Gbps/m^2 $\mu_C = 7.6$ Gbps $C_{\max} = 11.41$ Gbps $C_{\min} = 6.97$ Gbps $\sigma_C = 29.15$ dB
Wi-Fi	$C = 48.18$ Mbps/m^2 $\mu_C = 51.28$ Mbps $C_{\max} = 141.89$ Mbps $C_{\min} = 22.46$ Mbps $\sigma_C = 14.31$ dB	$C = 192.83$ Mbps/m^2 $\mu_C = 205.23$ Mbps $C_{\max} = 567.28$ Mbps $C_{\min} = 89.14$ Mbps $\sigma_C = 20.34$ dB	$C = 36.38$ Mbps/m^2 $\mu_C = 38.85$ Mbps $C_{\max} = 103.93$ Mbps $C_{\min} = 15.37$ Mbps $\sigma_C = 12.76$ dB	$C = 145.57$ Mbps/m^2 $\mu_C = 155.46$ Mbps $C_{\max} = 419.23$ Mbps $C_{\min} = 61.33$ Mbps $\sigma_C = 18.79$ dB
5GNR-I	$C = 108.64$ Mbps/m^2 $\mu_C = 117.09$ Mbps $C_{\max} = 445.4$ Mbps $C_{\min} = 48.14$ Mbps $\sigma_C = 18.71$ dB	$C = 542.82$ Mbps/m^2 $\mu_C = 584.94$ Mbps $C_{\max} = 2.22$ Gbps $C_{\min} = 242.3$ Mbps $\sigma_C = 25.68$ dB	$C = 83.47$ Mbps/m^2 $\mu_C = 91.03$ Mbps $C_{\max} = 367.6$ Mbps $C_{\min} = 35.47$ Mbps $\sigma_C = 17.82$ dB	$C = 417.63$ Mbps/m^2 $\mu_C = 455.5$ Mbps $C_{\max} = 1.84$ Gbps $C_{\min} = 179.4$ Mbps $\sigma_C = 24.81$ dB
5GNR-II	$C = 627.33$ Mbps/m^2 $\mu_C = 756.42$ Mbps $C_{\max} = 2.73$ Gbps $C_{\min} = 117.43$ Mbps $\sigma_C = 28.63$ dB	$C = 1.41$ Gbps/m^2 $\mu_C = 1.70$ Gbps $C_{\max} = 6.14$ Gbps $C_{\min} = 263.31$ Mbps $\sigma_C = 32.14$ dB	$C = 388$ Mbps/m^2 $\mu_C = 468.8$ Mbps $C_{\max} = 1.45$ Gbps $C_{\min} = 74.51$ Mbps $\sigma_C = 25.87$ dB	$C = 869.44$ Mbps/m^2 $\mu_C = 1.05$ Gbps $C_{\max} = 3.26$ Gbps $C_{\min} = 166.95$ Mbps $\sigma_C = 29.38$ dB
Aggregate	$C = 803.95$ Mbps/m^2 $\mu_C = 945.71$ Mbps $C_{\max} = 3.16$ Gbps $C_{\min} = 219.65$ Mbps $\sigma_C = 28.89$ dB	$C = 8.88$ Gbps/m^2 $\mu_C = 9.6$ Gbps $C_{\max} = 21.1$ Gbps $C_{\min} = 3.19$ Gbps $\sigma_C = 36.37$ dB	$C = 529.65$ Mbps/m^2 $\mu_C = 621.04$ Mbps $C_{\max} = 1.82$ Gbps $C_{\min} = 150.4$ Mbps $\sigma_C = 26.37$ dB	$C = 8.85$ Gbps/m^2 $\mu_C = 9.27$ Gbps $C_{\max} = 14.59$ Gbps $C_{\min} = 7.42$ Gbps $\sigma_C = 32.69$ dB

6 Conclusions

This deliverable, 5G-CLARITY D3.2, reported the KPI evaluations of the initial design of 5G-CLARITY user and control plane architecture for the integrated 5G/Wi-Fi/LiFi network. The performance evaluations were carried out based on the objectives of the three tasks defined in 5G-CLARITY WP3, namely T3.2, T3.3 and T3.4. The objectives of the noted WP3 tasks are mapped to specific objectives in deliverable as OBJ-1 to OBJ-5 (refer to Section 1.2).

Regarding OBJ-1, OBJ-2 and OBJ-3 of this deliverable, which focus on validating the initially designed 5G/Wi-Fi/LiFi multi-connectivity framework user and control plane algorithms as well as signalling between the architectural components (5G-CLARITY OBJ-TECH-3), evaluations on 3GPP (5GNR DU/CU/RU) and non-3GPP (integration of Wi-Fi and LiFi networks) networks are provided. More specifically, several virtual testbeds, lab-based testbeds and simulation environments are described to evaluate, (i) the integration of 5GNR CU/DU/RU; (ii) the handover performance of the integrated Wi-Fi/LiFi L2 network; (iii) the attachment procedure between the non-3GPP interfaces and N3IWF; (iv) the traffic flows through different interfaces via MPTCP schedulers; (v) the custom-made MPTCP scheduler; and (vi) the performance of the 5G-CLARITY eAT3S framework.

Regarding OBJ-4 of this deliverable, works on evaluating the simulation environment that validates the enhancements of the aggregated system area capacity via resource management algorithms (5G-CLARITY OBJ-TECH-4), Wi-Fi airtime-based scheduler, LiFi airtime-based scheduler and a service utility-based scheduler have been studied in Section 3. The performance evaluation results of these schedulers have been reported along with some practical considerations. On the other hand, the aggregated system area capacity and possible network deployment options have been evaluated in Section 5 where the system level simulation platform has been described in detail. The system area capacity performance has been obtained for each individual WAT as well as the integrated 5G/Wi-Fi/LiFi network deployment based on today's practical considerations and possible future achievements obtained from academic research.

Regarding OBJ-5 of this deliverable, which focuses on evaluation the positioning technologies and algorithms (5G-CLARITY OBJ-TECH-5). The positioning precision performance of different technologies namely, mm-Wave, sub-6 GHz, LiFi and OCC (i.e. VLP) has been studied in Section 4. The technologies were evaluated in laboratory conditions (mmWave, sub-6 GHz and VLP) and in simulation (LiFi). This study has been followed up with a fusion approach that integrates the positioning information from the noted technologies to improve the positioning accuracy.

The next deliverable of WP3, 5G-CLARITY D3.3, will consider the achieved KPIs for the integration of individual components of the user and control plane architecture, and provide performance evaluations for the overall 5G-CLARITY user and control plane architecture. Moreover, the task associated to the next deliverable, T3.1, will also be studied in more detail, with evaluation results being reported in 5G-CLARITY D3.3.

7 Bibliography

[1] 5G-CLARITY Deliverable D3.1, "State-of-the-art review and initial design of the integrated 5GNR/Wi-Fi/LiFi network frameworks on coexistence, multi-connectivity, resource management and positioning," September 2020. [Online]. Available: https://www.5gclarity.com/wp-content/uploads/2020/09/5G-CLARITY_D3.1.pdf.

[2] 5G-CLARITY Deliverable D2.2, "5G-CLARITY Primary System Architecture," October 2020. [Online]. Available: https://www.5gclarity.com/wp-content/uploads/2020/12/5G-CLARITY_D22.pdf.

[3] 5G-CLARITY Deliverable 5.1, "Specification of Use Cases and Demonstration Plan," February 2021. [Online]. Available: https://www.5gclarity.com/wp-content/uploads/2021/02/5G-CLARITY_D51.pdf.

[4] N. McKeown, T. Anderson, H. Balakrishnan, G. Parulkar, L. Peterson, J. Rexford and J. Turner, "OpenFlow: enabling innovation in campus networks," *ACM SIGCOMM Computer Communication Review,* vol. 38, no. 2, pp. 69--74, 2008.

[5] B. Pfaff and B. Davie, "The open vswitch database management protocol," Internet Requests for Comments, RFC Editor, RFC, 7047, 2013.

[6] 3GPP TR 23.793, *Study on Access Traffic Steering, Switch and Splitting Support in the 5G System Architecture (Release 16),* 2018.

[7] 3GPP TS 23.501, *System Architecture for the 5G System; Stage 2.*

[8] D. Maughan, M. Schertler, M. Schneider and J. Turner, "Internet security association and key management protocol (ISAKMP)," 1998.

[9] M. Scharf and A. Ford, "MPTCP Application Considerations, draft-scharf-mptcp-api-02," IETF77, Mar..

[10] Accelleran, "dRAX RIC xApps Development Guide," [Online]. Available: https://cutt.ly/drax-xapp-dev-guide.

[11] 5G-CLARITY Deliverable D4.2, *Validation of 5G-CLARITY SDN/NFV Platform, Interface Design with 5G Service Platform, and Initial Evaluation of ML Algorithms.*

[12] M. Vipin and S. Srikanth, "Analysis of open source drivers for IEEE 802.11 WLANs," in *International Conference on Wireless Communication and Sensor Computing (ICWCSC)*, 2010.

[13] H. Alshaer and H. Haas, "Bidirectional LiFi attocell access point slicing scheme," *IEEE Transactions on Network and Service Management,* vol. 15, no. 3, pp. 909--922, September 2018.

[14] A. Lieto, I. Malanchin and A. Capone, "Enabling Dynamic Resource Sharing for Slice Customization in 5G Networks," in *Global Communications Conference*, 2018.

[15] E. Rodrigues, F. R. M. Lima, T. F. Maciel and F. R. P. Cavalcanti, "Maximization of user satisfaction in OFDMA systems using utility-based resource allocation," *Wireless Communications And Mobile Computing,* vol. 16, pp. 376-392, 2016.

[16] M. Petri and M. Ehrig, "A SoC-based SDR Platform for Ultra-High Data Rate Broadband Communication, Radar and Localization Systems," in *11th Wireless Days Conference (WD 2019)*, 2019.

[17] Y. Zhuang, L. Hua, L. Qi, J. Yang, P. Cao, Y. Cao, Y. Wu, J. Thompson and H. Haas, "A survey of positioning systems using visible LED lights," *IEEE Communications Surveys & Tutorials,* vol. 20, no. 3, pp. 1963--1988, 2018.

[18] J. Serrano, M. Lipinski, T. Wlostowski, E. Gousiou, E. v. d. Bij, M. Cattin and G. Daniluk, "The white rabbit project," 2013.

[19] R. Gold, "Optimal binary sequences for spread spectrum multiplexing," *IEEE Transactions on Information Theory (Correspondence),* Vols. IT-13, no. 4, p. 619–621, 1967.

[20] pureLiFi, "LiFi-XC datasheet," 2017. [Online]. Available: https://purelifi.com/wp-content/uploads/2017/12/LiFi-XC-Data-sheet-Snapshot.pdf. [Accessed May 2021].

[21] M. Vizjia and M. Jankovic, "Pose estimation & camera movement tracking," 2018. [Online]. Available: http://tnt.etf.bg.ac.rs/~mv/MV1718_v4.pdf.

[22] A. L. Barker, D. E. Brown and W. N. Martin, "Bayesian estimation and the Kalman filter," *Computers \& Mathematics with Applications,* 1995.

[23] A. A. Purwita, "OWCsim-Py," [Online]. Available: https://github.com/ardimasp/owcsimpy.

[24] "Zemax - OpticStudio," [Online]. Available: https://www.zemax.com/products/opticstudio.

[25] T. Yamakami, "A Zipf-Like Distribution of Popularity and Hits in the Mobile Web Pages with Short Life Time," in *2006 Seventh International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT)*, Taipei, Taiwan, 2006.

[26] M. Cha, H. Kwak, P. Rodriguez, Y. Ahn and S. Moon, "Analyzing the Video Popularity Characteristics of Large-Scale User Generated Content Systems," *IEEE/ACM Transactions on Networking,* pp. 1357-1370, 2009.

[27] "WINNER II Channel Models," in *Radio Technologies and Concepts for IMT-Advanced*, 2009, pp. 40-92.

[28] "Chapter 3 - WINNER II Channel Models," in *Radio Technologies and Concepts for IMT-Advanced*, Germany, Wiley, 2009, pp. 39-93.

[29] M. Khalily, M. Ghoraishi, S. Taheri, S. Payami and R. Tafazolli, "Millimeter-wave directional path loss models in the 26 GHz, 32 GHz, and 39 GHz bands for small cell 5G cellular system," London, UK, 2018.

[30] C. Chen, D. A. Basnayaka and H. Haas, "Downlink Performance of Optical Attocell Networks," *Journal of Lightwave Technology,* pp. 137-156, 2016.

[31] J. M. Kahn and J. R. Barry, "Wireless infrared communications," *Proceedings of the IEEE,* pp. 265-298, 1997.

[32] M. D. Soltani, A. A. Purwita, Z. Zeng, H. Haas and M. Safari, "Modeling the Random Orientation of Mobile Devices: Measurement, Analysis and LiFi Use Case," *IEEE Transactions on Communications,* pp. 2157-2172, 2019.

[33] J. G. Andrews, F. Baccelli and R. K. Ganti, "A Tractable Approach to Coverage and Rate in Cellular Networks," *IEEE Transactions on Communications,* vol. 59, no. 11, pp. 3122-3134, 2011.

[34] "Lighting at work," Great Britain Health and Safety Executive, 1997.

[35] J. M. Kahn and J. R. Barry, "Wireless infrared communications," *Proceedings of the IEEE,* vol. 85, no. 2, pp. 265-298, 1997.

[36] "How To Estimate Number of Access Points Needed For WiFi Projects," AccessAgility, [Online]. Available: https://www.accessagility.com/blog/how-to-estimate-number-of-access-points-needed. [Accessed 06 05 2021].

[37] I. Cisco Systems, "Wi-Fi Location-Based Services 4.1 Design Guide," Americas Headquarters, 2008.

[38] K. B. S. Manosha, K. Hiltunen, M. Matinmikko-Blue and M. Latva-Aho, "Performance Comparison of Alternative Indoor 5G Micro-Operator Deployments in 3.6-GHz and 26-GHz Bands," *IEEE Transactions on Cognitive Communications and Networking,,* vol. 5, no. 4, pp. 886-899, 2019.

[39] "Study on channel model for frequencies from 0.5 GHz to 100 GHz (ver-sion 14.1.1.)," 3rd Gener. Partnership Project, Sophia Antipolis, France, Rep. TR 38.901, 2017.

[40] J. G. Proakis and M. Salehi, Digital Communications, McGraw-Hill Education, 2007.

[41] J. B. Carruthers and J. M. Kahn, "Multiple-subcarrier modulation for nondirected wireless infrared communication," *IEEE Journal on Selected Areas in Communications,* vol. 14, no. 3, pp. 538-546, 1996.

[42] H. L. M. e. al, "100-Mb/s NRZ Visible Light Communications Using a Postequalized White LED," *IEEE Photonics Technology Letters,* vol. 21, no. 15, pp. 1063-1065, 2009.

[43] Harald Haas, "LiFi is a paradigm-shifting 5G technology," *Reviews in Physics,* vol. 3, pp. 26-31, 2018.

[44] H. Chun and e. al, "Visible Light Communication Using a Blue GaN $\mu $ LED and Fluorescent Polymer Color Converter," *IEEE Photonics Technology Letters,* vol. 26, no. 20, pp. 2035-2038, 2014.

[45] C. Lee, C. Shen, C. Cozzan, R. Farrell, J. Speck, S. Nakamura, B. Ooi and S. DenBaars, "Gigabit-per-second white light-based visible light communication using near-ultraviolet laser diode and red-, green-, and blue-emitting phosphors," *Opt. Express,* pp. 17480-17487, 2017.

[46] OFCOM, "OfW48 UK Frequency Allocations for Fixed (Point-to-Point) Wireless Services and Scanning Telemetry," 2020.

[47] H. H. e. al., "Introduction to indoor networking concepts and challenges in LiFi," *IEEE/OSA Journal of Optical Communications and Networking,* vol. 12, no. 2, pp. A190-A203, 2020.