DOI: 10.1364/JOCN.551626

DEMONSTRATION OF A MULTI-TECHNOLOGY 6G TRANSPORT NETWORK INTEGRATING THZ AND OPTICAL NETWORK TECHNOLOGIES EMPOWERED BY FEDERATED LEARNING

M. ANASTASOPOULOS(1), A. TZANAKAKI(1), Y. JIAN(2), L. LOPACINSKI(2), J. GUTIÉRREZ(2), I. MESOGITI(3), E. THEODOROPOULOU(3), G. LYBEROPOULOS(3)

[1] National and Kapodistrian University of Athens, Department of Physics, Athens 15772, Greece
[2]IHP – Leibniz-Institut für innovative Mikroelektronik, DE
[3]OTE- Hellenic Telecommunications Organization S.A., GR
*atzanakaki@phys.uoa.gr

Received 12 Dec 2024; revised XX Month, XXXX; accepted XX Month XXXX; posted XX Month XXXX (Doc. ID XXXXX); published XX Month XXXX

This paper proposes a 6G architecture that adopts a multi-technology transport network integrating together THz links and optical network technologies to interconnect 6G Radio Access Network and Core domains. The proposed solution operates in a self-organized manner taking advantage of the Software Defined Networking (SDN) control of the optical network, while suitable SDN control is purposely developed for the THz solution. An end-to-end (E2E) Service Management and Orchestration (SMO) layer is adopted offering intelligence capabilities in service provisioning and resource allocation across the 6G infrastructure through the adoption of a Federated Learning (FL) scheme. The proposed architecture and developed control scheme are validated through an experimental demonstration. This to the best of the authors knowledge is the first experimental demonstration of an autonomously controlled 6G network implementation integrating optical network technologies and THz links in a common transport network. The developed intelligent management framework empowered by FL to jointly optimize THz and multi-vendor optical network equipment supporting 6G services is also experimentally showcased.

1. INTRODUCTION

The evolution from 5G to 6G is aiming at creating intelligent, ultra-fast, reliable/secure and connected networks that will revolutionize industries and empower digital transformation on a global scale. 6G will integrate multi-technology wired and wireless Transport Networks (TNs) operating in new spectrum bands providing enhanced connectivity for the Radio Access Network (RAN) and mobile core elements. In this context, high-capacity optical TNs combined with THz communication links providing high-bandwidth and flexibility transport connectivity will play a key role. THz links can be installed in indoor or outdoor environments to assist in the deployment of local/private 6G networks. In indoor environments, THz point to point links can provide fully wireless fronthaul/backhaul solutions interconnecting disaggregated RAN and core elements. In outdoor environments THz transceivers can be installed either at rooftops or can be mounted on poles providing point-to-point connectivity for 6G services. However, to enable seamless operations across the multiplicity of 6G technologies, domains and infrastructures these solutions need to be equipped with enhanced automation and real-time decision-making capabilities, through intelligent management and control frameworks [1].

Significant research efforts have been dedicated to optical/wireless network integration in support of mobile communication systems. However, these are mostly limited to communication links operating in sub-6 and mmWave frequency bands e.g. [1] and [3]. Although mmWave technology is a cost-effective solution that can accelerate the deployment of B5G networks, the 60 GHz band is expected to become crowded in the coming years, especially in urban environments. In addition, future cellular networks will demand continuously increasing bandwidths and data rates, notwithstanding the fact that unfavorable channel conditions and blockage are common effects impacting wireless communications at such carrier frequencies.

In this context, wireless THz links can complement mmWave links and wired optical networks providing a high-bandwidth alternative for transport networks, the necessary capacity for various mobile network deployment scenarios, as well as protection capabilities against possible outages. Recognizing the benefits of THz technologies and the advantages they can bring in 6G, the scientific community has considered wireless THz links as a key component supporting connectivity requirements for fronthaul [5], [6] and backhaul [7] services. Despite their huge potential, the majority of existing experimentation studies using THz technologies have been focusing on testing their physical layer performance under point-to-point scenarios [8], [9] without considering the specificities of 5G-related fronthaul and/or backhaul traffic streams. One of the first experimentations using THz links for the transmission of 5G network traffic has been reported in [10] utilizing a point-to-point THz link to transmit *NR air interface* signal streams.

Studies focusing on the integration of THz links with optical networks are mostly limited to data center networks [11], [12] whereas, to the best of our knowledge, experimentation activities instantiating fronthaul and/or backhaul services over multi-technology TN integrating optical and THz communication technologies have not been reported in the literature to date. This is attributed to the fact that in addition to development of the data plane interfaces interconnecting the different TN technology domains i.e., (THz wireless links, optical network, compute), the 5G Network Function system should be also supported by suitable control and management interfaces facilitating seamless

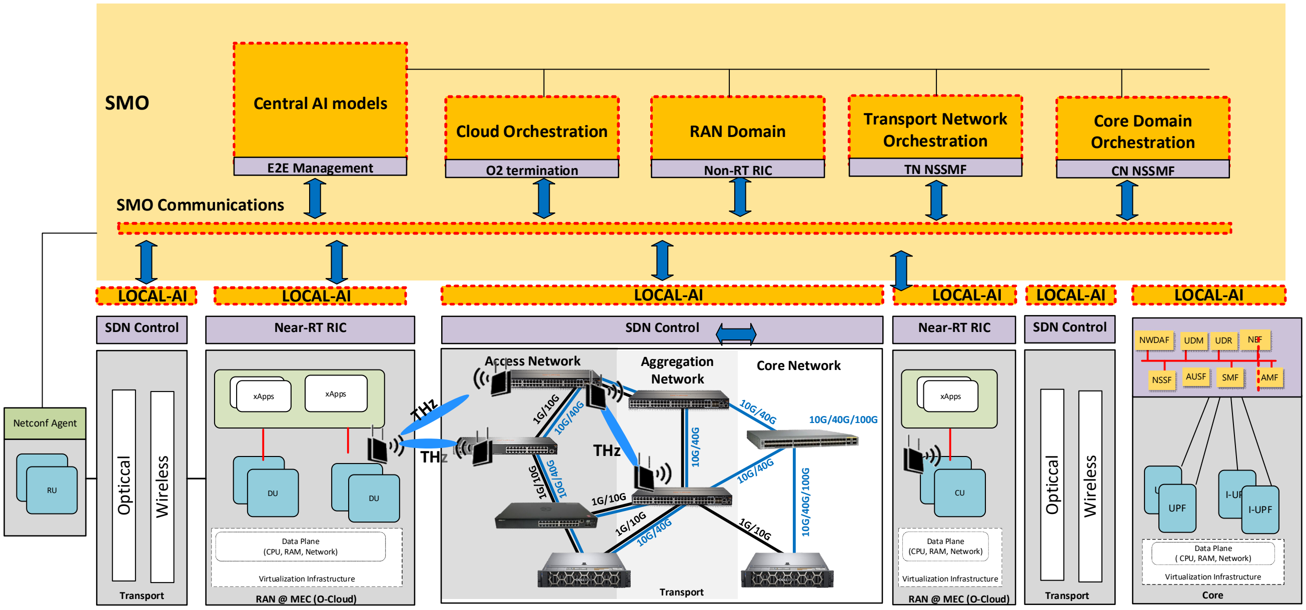

Fig. 1: AI-based 5G system Architecture: Multiple heterogeneous network segments are managed by local controllers.

interconnection of the RAN and Core Network (CN) functions. In practice, this can be achieved through the development of a hierarchical control and management framework comprising domain specific controllers for the RAN, the CN, the transport network, and compute elements. These controllers are supervised by a top-level Service Management and Orchestration (SMO) platform responsible to [14]:

1. orchestrate and coordinate RAN Network Function (NF) chaining,
2. configure the transport network ensuring RAN NF chaining connectivity under dynamic network conditions,
3. perform Topology and Inventory management,
4. manage end-to end service slices.

Regarding local domain controllers, the management and control of integrated wireless and optical 5G transport networks mainly rely on multi-domain Software Defined Networking (SDN) solutions interacting with the transport network elements and the overlay SMO platform [13]. Responsible for the management of the RAN segment is the Radio Intelligent Controller (RIC) that controls the Radio Resource Management (RRM) decisions for the RAN. RIC comprises two main components, a) the near-real-time RIC (near-RT RIC) a logical function that enables near-real-time control and optimization of O-RAN elements, b) a non-real-time RIC (Non-RT RIC) that is a logical function that enables non-real-time control and optimization of RAN elements and resources. It exploits Artificial Intelligence (AI) / Machine Learning (ML) techniques in the workflow including model training and updates, and policy-based guidance of applications/features [14]. In the CN the main function supporting optimal resource management decisions is the Network Data Analytics Function (NWDAF).

The resource management policies applied to the different domains (cloud infrastructure, RAN NFs) can be calculated adopting either centralized [21] or distributed decision-making frameworks [22], while the overall network management decisions are supported by the SMO that collects data from the different network elements and performs analytics to support lifecycle management of the network. However, the centralized nature of the controllers adopted to facilitate the required control may raise security and trust problems [23]. A possible solution that has been recently proposed to address this is Federated Learning (FL) that can be used to split ML training among multiple entities without exposing sensitive data [24] improving also scalability and efficiency [25].

This paper proposes a 6G system architecture that utilizes a heterogeneous transport network combing together THz links and optical network technologies to interconnect the RAN and the CN domains, where optical switching plays a key role. The proposed solution operates in a self-organized manner taking advantage of the SDN control of the optical transport network and an end-to-end (E2E) SMO layer with intelligence capabilities offered through the adoption of an FL scheme. Based on these, network data are collected and stored locally (within a trusted domain e.g. near RT RIC) and are used to generate a global model applied for E2E service provisioning. To the best of the authors knowledge this is the first experimental demonstration of an autonomously controlled 6G network implementation integrating optical network technologies and THz links in a common transport network. The developed intelligent management framework is empowered by FL to jointly optimize THz and multi-vendor optical networks supporting 6G services.

The remainder of this paper is organized as follows: Section 2 provides an overview of the developed system architecture including the SMO, the management of the integrated THz and optical transport network infrastructure interconnecting the RAN NFs with the CN and, finally, the blocks used to provide intelligence in decision making. A detailed description of the designed and fabricated THz node along with their control plane is provided in Section 3, while a description of the proposed distributed decision-making framework based on FL is given in Section 4. Section 5 experimentally evaluates the proposed FL based decision framework over the multi-technology transport network whereas Section 6 concludes the paper.

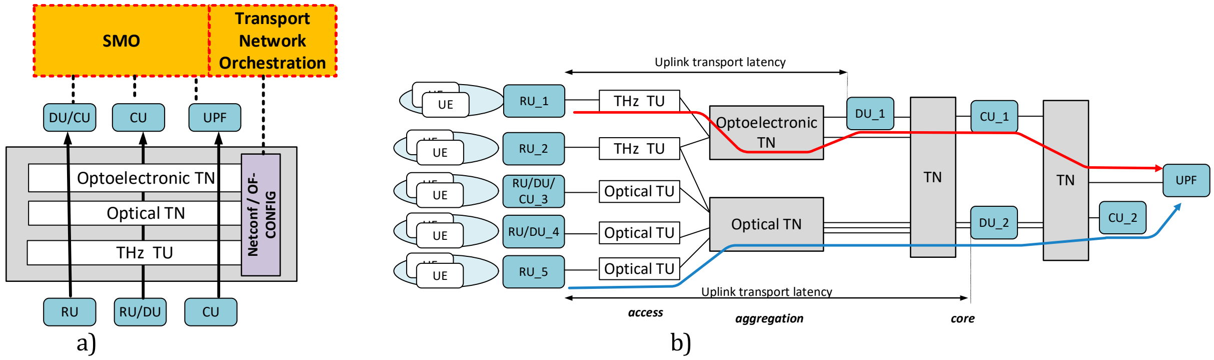

Fig. 2 a) SDN controlled transport network, b) traffic aggregation over multi-technology transport.

2. System Architecture

A. Overall system description

The present study considers an Open-RAN architecture aligned with O-RAN [14] and 3GPP specifications [15]. As shown in Fig. 1, the architecture consists of the RAN virtualized NFs i.e. the Distributed Units (DUs) and the Centralized Units (CUs), the cloud infrastructure[1] hosting the RAN NFs, the 3GPP CN and the SMO managing the overall platform. Connectivity between the different components of the architecture is provided through an SDN controlled multi-technology transport network infrastructure supporting fronthaul (FH) i.e. connectivity between the Remote Unit (RU) and DU, midhaul (F1) i.e. connectivity between the DU and CU and backhaul (BH) services i.e. connectivity between the CU and the CN. The architecture supports hierarchical control loops managing the RAN segment operating at near real time (response time in the range 10ms-1s) and non-real time (response time > 1s)) time scales. Near real time decisions are taken by the Near RT RIC, whereas non-real time decisions are taken by the Non-RT RIC hosted at the SMO.

The Near-RT RIC enables near real-time control and optimization of services and resources of the DU and CU nodes connected using the E2 interfaces. The Near-RT RIC hosts one or more xApps that use E2 interface to provide value added services through the collection of near real-time information e.g., on a User Equipment (UE) and/or a Cell basis. Based on the available data, the Near-RT RIC generates the RAN analytics information that can be exposed to external users via Y1 interface. The Non-RT RIC is a functionality internal to the SMO used to support intelligent RAN optimization by providing policy-based guidance, ML model management and enrichment information to the Near-RT RIC function [16]. Interaction between the Non-RT RIC and the Near RT RIC is provided through the A1 interface. Non-RT RIC can use data analytics and AI/ML training/inference to determine the RAN optimization actions. Non-RT RIC collects statistics and measurements from the various interfaces it exposes to the underlying physical resources leveraging SMO capabilities. These interfaces include O2 to the O-Cloud infrastructure, O1 to the RAN NFs, Open FH Management (M)-plane for the RUs etc. More specifically, configuration and management of RAN NFs, including Fault, Configuration, Accounting, Performance and Security (FCAPS), and their Operation Administration and Maintenance (OAM) are supported through the O1 and the Open FH Management-plane (M-plane) interfaces that are implemented using the NETCONF-protocol.

According to the O-RAN Alliance specifications, the SMO can be used to manage and orchestrate the multi-vendor disaggregated O-RAN Network Functions (i.e., Near-RT RIC, CU, DU, RU) and the O-Cloud infrastructure. However, additional domain orchestration functionalities have to be integrated into the SMO to manage the core and transport network domains. For example, based on the E2E service connectivity requirements the corresponding E2E orchestrator communicates with the transport network orchestrators to configure the requested network slices. Therefore, the transport network orchestrator converts high level (abstracted) connectivity requirements into a model that can be actioned by the SDN controller of the transport network. The SDN controller then converts the network models into device specific configurations [17].

In the present solution, the transport network controller is used to manage an integrated opto-electronic/wireless infrastructure interconnecting the RAN with the MEC elements where core functions are placed. The overall system is organized in a hierarchical manner offering RAN connectivity (access network) and collecting and aggregating transport traffic (aggregation network) from various cells to a central location. Both the access and aggregation transport network segments are implemented through multi-vendor SDN-controlled optoelectronic switches with different capabilities (number of ports, capacity per port and latency). The access network is equipped with low energy consuming switching nodes having limited number of input ports and relatively small capacity (1/10GbE/port, with SFP, SFP+ and RJ45 transceivers). The aggregation network segment comprises high-end switches with higher capacity and density (10G/40G/100GbE/ports) offering much higher energy efficiency (lower bit/Joule) under high utilization compared to low-end switches. The optical network is complemented with THz nodes used to provide high-speed connectivity between the RAN NFs and transport network nodes or between transport nodes. To integrate the THz nodes with the optical transport network, the OpenFlow (OF) Configuration Protocol has been used allowing centralized SDN controllers (such as RYU and OpenDaylight), to manage the whole transport network.

[1] In the terminology of O-RAN this is known as O-Cloud

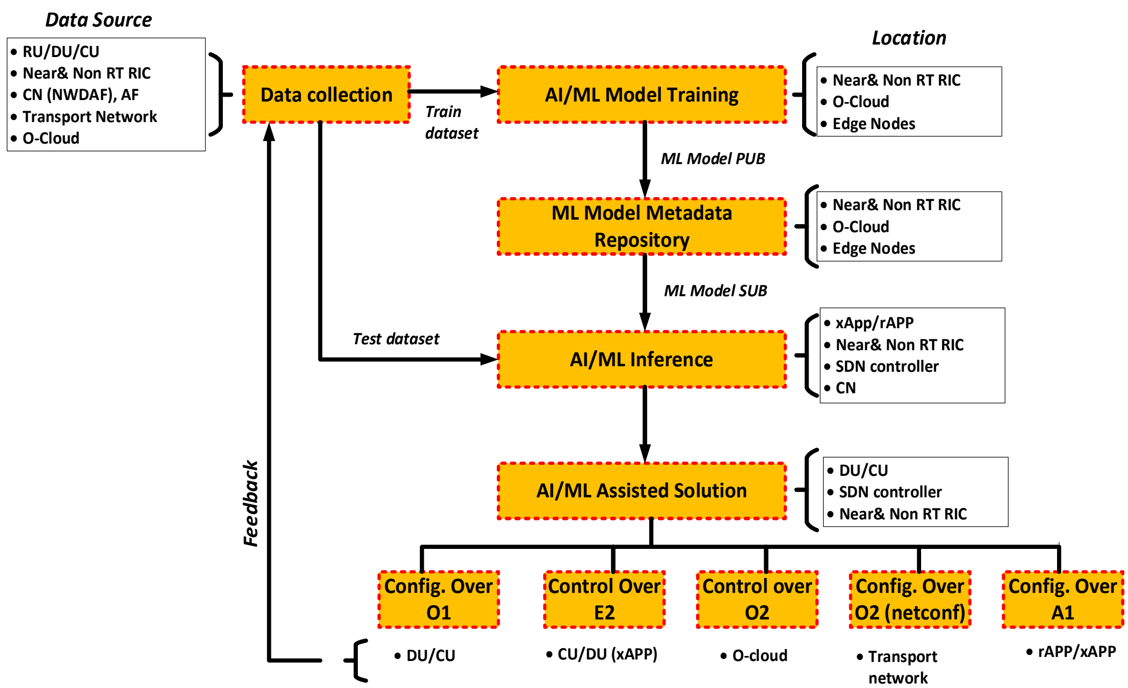

Fig. 3: SMO Intelligence Management: Supported AI/ML workflows and candidate locations where data, ML inference and training can be hosted.

B. Multi-technology Transport Network

A multi-technology transport network is considered that is operating in a cooperative manner supporting traffic connectivity requirements between RAN NFs (RU-DU, DU-CU) and between RAN NFs and the CN (CU-User Plane Function (UPF)) as shown in Fig. 2 (a) [18]. The TN nodes can operate in a complementary and cooperative manner addressing a wide range of transport network connectivity options (i.e. point-to-multi-point and multi-point to multipoint) utilizing optical and opto-electronic switches. Wireless Transport network Units (TU) primarily used for point-to-point connectivity operating in the THz spectrum band are also considered. These technologies support capacities in the range of 1Gbps up to 100Gbps and key features, such as synchronization.
Low cost wired (i.e. point-to-point optical) or wireless (THz) TUs can be installed close to the RUs offering connectivity to the radio transmission points, collecting and aggregating transport traffic from various cell sites to a central location (hub site), where centralized processing is hosted. The overall architecture addresses different disaggregated RAN deployment options and protocols. Optoelectronic and all optical TNs with higher capacity per port and more sophisticated network connectivity capabilities (i.e. mesh connectivity) are used to collect and aggregate traffic from multiple TUs. This traffic can be further aggregated through the transport core network that is responsible to collect traffic from the aggregation network and forward the corresponding traffic flows deeper into the 6G CN thus fulfilling the requirements of the requested end-to-end service slices.

To meet the protocol requirements of the requested service flows, the TN orchestrator interacts with the SMO (RIC) in order to identify: a) how traffic aggregation and network resource sharing can be performed and b) where the corresponding services will be terminated. Therefore, through continuous interaction with the RIC the optoelectronic switches take decisions related to the output ports that the incoming traffic will be forwarded and the scheduling policies that will be applied at the corresponding virtual output queues. For example, as shown in Fig. 2 b) due to buffering and scheduling latencies introduced by the Optoelectronic TN the FH connection of RU_1 will be terminated at DU_1 in order to satisfy the

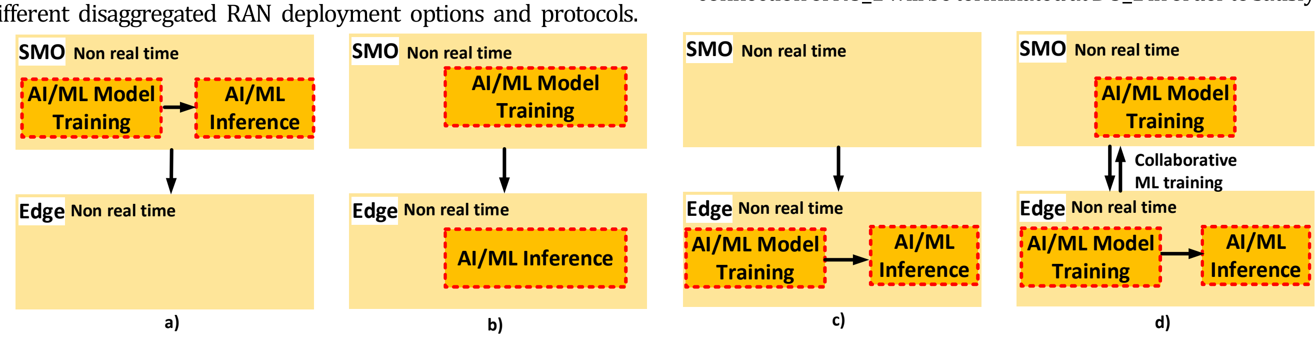

Fig. 4: AI/ML model deployment options under a) Option 1: Offline training and inference, b) Option 2: Offline training and online inference, c)Option 3: Online training and online inference, d) Option 4: Distributed training and online inference.

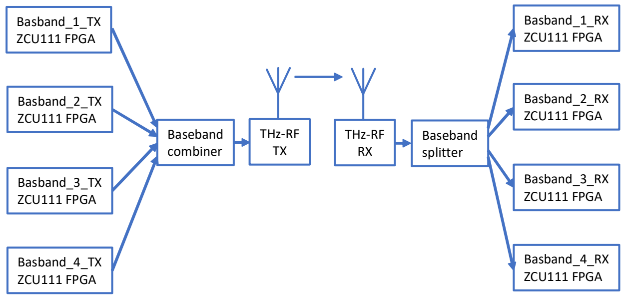

Fig. 5 Baseband processing with analogue combining.

corresponding uplink transport latency constraints. On the other hand, the all-optical TN can transparently forward traffic streams without introducing additional delays associated with scheduling policies and buffering and therefore enables the TN orchestrator to direct and terminate connections deeper into the transport network. This results in higher degree of aggregation of the transported traffic flows, thus achieving increased overall resource efficiency.

C. Data and network intelligence

AI/ML techniques are core in optimizing the operation of the integrated infrastructure. Intelligent operations for the RAN, CN and transport network domains are facilitated through the various controllers installed at the different segments of the system. For example, intelligent RAN optimization is supported by the Non-RT RIC through the provisioning of policy-based guidance, ML model management and enrichment information offered to the Near-RT RIC function. Intelligent radio resource management can be also performed within a non-real-time interval (i.e., greater than 1 second). The Non-RT RIC can determine the RAN optimization actions through data analytics and AI/ML training/inference leveraging SMO services. These services include data collection and provisioning of the O-RAN nodes through O1 and O2 interfaces. Non-RT RIC can provide AI/ML capabilities through rApps. Similarly, near-real-time control and optimization of RAN elements and resources is provided through the Near-Real-Time RAN Intelligent Controller adopting also AI/ML techniques. These are achieved through high granularity data collection and other actions supported by the E2 interface. Overall, the Near-RT RIC functionalities are supported through xApps [23].
The SMO intelligence management processes are shown in Fig. 3. In addition to RAN NFs, the SMO can also aggregate data from the UEs, the CN, the application function and the transport network. The collected data can be either regular monitoring information or enriched statistics that can then be utilized for ML model training and inference. Data, ML training and inference can be hosted at different locations depending on the available physical resources and the specifications of the services. The trained ML model is stored in an ML repository which can be then selected for inference. The output of the inference can be used by an Actor (i.e., DU, CU, Non/near RT RIC, SDN controller, CN controller) to provide the "AI/ML assisted Solution" optimizing the performance of a specific network segment. The relevant actions (i.e., network configurations, management policies, scheduling and routing decisions etc) are applied to the targeted entity using appropriate SMO interfaces (O1, O2, E2, A1).

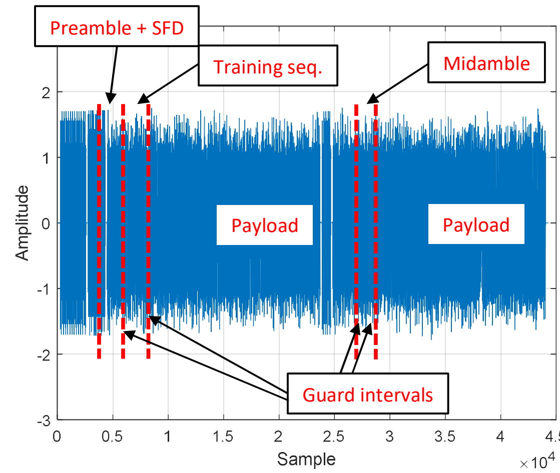

Fig. 6 An example frame generated by the implemented transmitter.

Finally, depending on the location of the ML training and inference, the time scale where decisions need to be taken, data availability and computational complexity, multiple ML deployment options can be considered [19]:

1. Option 1: Offline training and inference. In this case the ML workflow is used to support non-real time decisions (with response time >1s) and therefore, ML training and inference can be hosted in the SMO and/or in the domain specific orchestrators (i.e., Non-RT RIC for RAN, transport network orchestrator, CN orchestrator etc as shown in the top layer of Fig. 1)
2. Option 2: Offline training and online inference. In this deployment option ML training is performed at the SMO whereas inference is typically running close to the edge (i.e., in the near RT RIC). This option is suitable in scenarios requiring fast decisions times and global knowledge of the status of the network (data collected from all network entities are stared in centralized entity). Due to the large volume of training dataset, the SMO needs to be supported by significant computational resources.
3. Option 3: Online training and online inference where both ML training and inference are performed at the edge using limited (local) information. In this scenario data used for inference are stored locally (i.e., in the Near RT RIC) mitigating privacy risks.
4. Option 4: Distributed training and online inference. This approach, also known as federated learning, combines the benefits of Option 2 and 3 as it allows multiple AI/ML entities to train an ML model collaboratively, whereas inference is performed at the edge. Instead of gathering training data into a central server, FL keeps the training data decentralized to mitigate privacy risks. AI/ML models are trained locally in distributed entities, and the local models are aggregated by a central server that orchestrates the FL.

Although the proposed approach supports all ML deployment options, the management framework used in the present study adopts and implements Option 4 using FL to jointly optimize THz and multi-vendor optical networks. To apply the relevant policies this framework relied on the NETCONF Orchestration Application Programming Interface (API), appropriately extended to support the THz nodes.

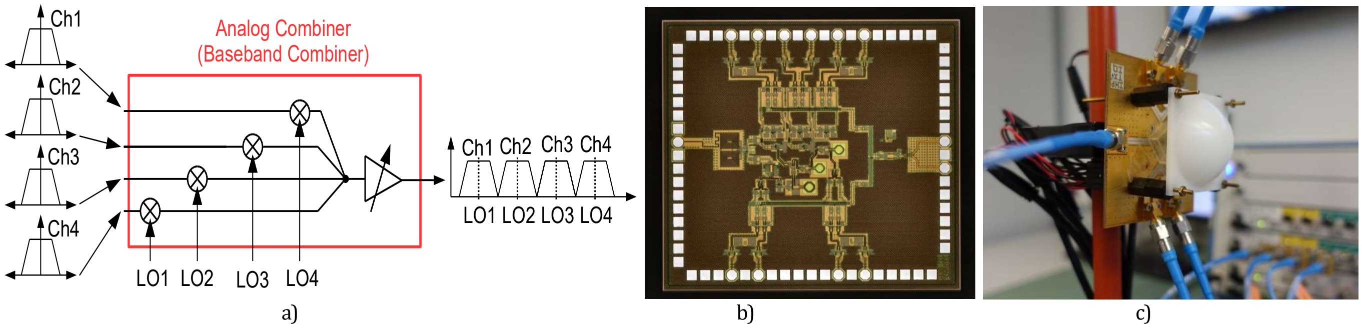

Fig. 7 a) An Analogue baseband combiner and b) IC-Die photography of the analogue BB-splitter, c) 240 GHz RF frontend used for baseband processing validation.

The FL-based optimization framework is used to support optimized E2E service provisioning by identifying the optimal match action rules applied to the transport network by the SDN controller. Specifically, using actual mobile network traffic statistics (provided by OTE), prediction of resource requirements per domain is performed through FL. The predicted values are sent to the central orchestrator responsible to calculate optimal E2E optimization policies.

3. THz Node Development and Control

A. THz Node Development

The THz transport targets communication at 240 GHz and comprises a modular single carrier (SC) (multi)-baseband processor, the cores of which run in parallel in a single (or multiple) FPGA platforms. This is considered as one of the main innovations in the development of THz communications as using the modular baseband processing engine, the large bandwidth of THz systems is divided into sub-channels that are processed by a set of SC baseband processors. This approach relaxes the demands for Analog to Digital Conversion (ADC) and Digital to Analog Conversion (DAC) bandwidth, as well as the timing for baseband processing.

The targeted baseband unit supports THz communications and allows splitting the computation and ADC/DAC processing effort among parallel cores (see Fig. 5), by the dedicated baseband (BB) signal combiner realized in 130 nm SiGe technology. To simplify the processing of a very wide THz band, SC transmission is employed. A modular baseband approach is designed, where the bandwidth of the THz node is split into channels, and each channel is subsequently processed by one FPGA baseband core running in parallel (for a maximum of up to four cores). The independent IQ baseband signals are upconverted to different intermediate frequencies before being combined in the electrical domain by an analogue channel combiner. At its output, a single real signal is interfaced with a wireless transceiver. In contrast to other typical Orthogonal frequency division multiplexing (OFDM) basebands, the proposed baseband core uses SC modulation. This approach reduces baseband complexity, as well as power consumption and implementation area.

The baseband transmitter (Baseband Tx) model consists of a timing control subsystem, a preamble generator, a training sequence generator, a data modulator, and a parallel root-raised-cosine finite impulse response (FIR) filter. The FIR filter with a length of 31 taps and $\alpha = 0.2$ is used as a matched filter at the transmitter's output.

The transmitted frames consist of the preamble with the start of frame delimiter (SFD), channel estimation training sequence, the data part (payload), and optional midambles (Fig. 6). Between the preamble and the payload, a guard interval is inserted. The midambles are used for continuous synchronization of frequency, time, and phase, which are especially desired for extended data frames. A base time unit used in the implementation is the transmission length of Golay-64 sequence employed in the preamble. The transmission length of all frame parts (preamble, guards, data part, training sequence, and midambles) is a multiple of the Golay-64 sequence. For the initial synchronization of frequency, time, and phase, ten to fourteen Golay-64 sequences are needed. The minimum guard interval length is 1 x Golay-64 length, and for the midamble, four Golay-64 sequences are usually used. Moreover, the training sequence length can be adjusted in the following steps: 0, 64, 128, 256, 512, and 1024 bits for I and Q branches. The preamble, training sequence, and midambles are separated from each other and form the payload by automatically inserted guard intervals. Our implementation supports BPSK, QPSK, and 16-QAM modulation. The length of the data part is adjustable in multiples of 64 bits. The optimal frame structure, including the length of the training sequences, guard intervals, and other slots, can be configured by the SDN controller.

The Baseband combining functionality is realized by a set of analogue mixers working with independent local oscillators, LO1-LO4, as shown in Fig. 7 (a). In our case, the LO signals are positioned at 3.75 GHz, 7.5 GHz, 10 GHz, and 15 GHz. In the end, the signal is added and amplified. The baseband splitter at the receive side performs the reverse operation. The fabricated BB-splitter chip is shown in Fig. 7 (b).

Finally, for the THz RF frontend the system shown in Fig. 7 (c) has been used operating in the 240GHz frequency band with 35GHz bandwidth and 110 Gbps data rate for 16 QAM and 80Gbps for QPSK modulation. More information on the 240 GHz frontend can be found in [27].

B. SDN control

To integrate the THz node with the overall system architecture a set of Northbound and Southbound Interfaces (NBIs & SBIs) have been developed. These provide the necessary abstraction required for the integration of the THz node with the optical network controllers and SMO. The overall architecture of the SDN controller developed is shown in Fig. 8. The SDN-capable THz controller is able to expose to the orchestrator a set of actions and notifications including: a) its

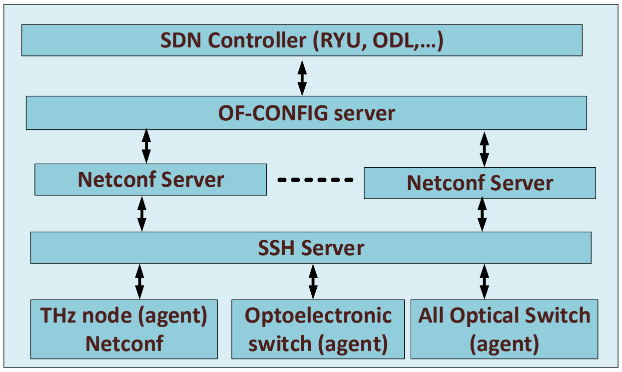

Fig. 8: SDN Control of the Integrated Transport Network.

availability for the provisioning of connectivity services, b) details of the internal PHY layer configuration parameters, c) the option to reconfigure the system and change its parameters, d) termination of connectivity services, etc. The system is also able to expose the THz nodes actual network topology graph through the NBI of the SDN Controller. To implement the corresponding functions, the NETCONF Orchestration API has been extended to support the THz nodes discussed in Subsection 3.A.

As shown in Fig. 8, the NETCONF architecture consists of two elements: the server (Agent) and the client (Driver). The server (Agent) is responsible to maintain information on the managed THz devices and responds to the client-initiated requests. When receiving a request from a NETCONF client, the NETCONF server

- parses the request, applies the relevant actions to the THz node and sends a reply to the client.
- If a fault or another type of event occurs on a managed device, the NETCONF server reports an alarm or event, enabling the client to learn the status of the managed device.

The NETCONF client provides the following functions:

- Management of network devices using NETCONF.
- Sending Remote Procedure Call (RPC) requests to the NETCONF enabled THz devices to query or modify various parameter values.
- Learning the status of a managed THz device through the alarms and events that the NETCONF server of the managed device sents.

Communication between the NETCONF client and server is achieved through the RPC mechanism, following the establishment of a secure connection-oriented session between them.

The process of establishing and terminating a NETCONF session is as follows:

1. The THz SDN controller (client) establishes a Secure Shell (SSH) connection with a THz node. Then a NETCONF session is established between the two entities after authentication and authorization are complete.
2. The SDN controller and the THz node advertise their capabilities.
3. The Controller sends one or more RPC requests to the THz node. The following lists some examples of requests:
   - Query the configuration data or status of the device
   - Edit the configuration parameters of the device
   - Shutdown the device
4. The Controller terminates the NETCONF session.
5. The SSH connection is terminated.

The implemented NETCONF agent for the THz node provides a small set of low-level operations to manage the device configurations and retrieve device state information. In the current

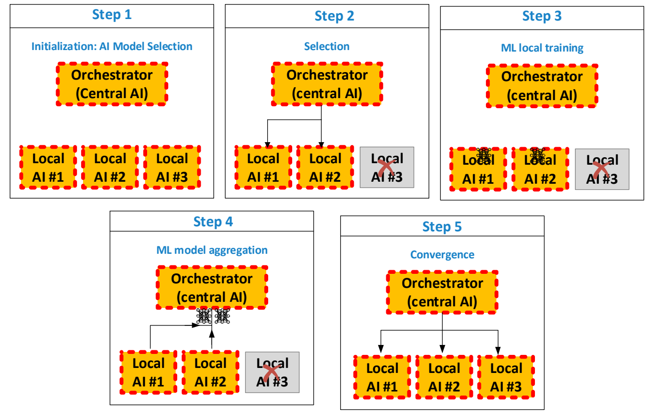

Fig. 9: AI model update using federated learning..

implementation, the basic operations supported are related to retrieval, change, copy, and deletion of the THz node configuration parameters. This allows to control a variety of THz node parameters (modulation format, frame length, guard length, etc.). At hardware level, the THz node parameters are stored in the registers of the FPGA. An example for a THz node configuration with eight registers and register length 8 bits ($b_7 b_6 b_5 b_4 b_3 b_2 b_1 b_0$) is shown in Table 1. The registers can be accessed through their virtual address (Register Addr 0x00, 0x01,..., 0x07) and the value of each bit or their combination is associated with a specific THz node operational parameter.

REG Add	b_7	...	b_3	b_2	b_1	b_0	RST
0x00			Ch3	Ch.2	Ch.1	Ch.0	0x0F
0x01					Modulation		0x02
0x02	Frame format and settings						0x00
0x03							0x00
0x04							0x00
0x05							0x00
0x06							0x00
0x07							0x00

Table 1: Register Address and configuration parameters of the THz Node

For example, in the Register Address 0x01 the two most right bits ($b_1 b_0$) are used to indicate the modulation format supported by the system. Hence, $b_1 b_0 = 00$ indicate that the system has been configured to operate with BPSK modulation format $b_1 b_0 = 01$ indicates the QPSK modulation whereas $b_1 b_0 = 10$ QAM16 modulation whereas $b_1 b_0 = 11$ is reserved for future extensions. In other registers (i.e., Register Addr. 0x02) other parameters that are used to configure the THz node, such as the frame format, are stored. At the current implementation stage, the modulation format, frame length, guard length, symbol length, training sequence length and preamble can be defined.

The low-level operations can be executed either directly through the command line interface or indirectly using the NETCONG agent. Directly a user connected to the FPGA platform of the THz node through a serial connection can read and modify the configuration of the system.

Once the low configuration parameters of the system have been retrieved, the NETCONF agent encodes all available information creating a NETCONF compliant XML message. The NETCONF agent can then respond to a set of commands issued by the NETCONF client [26]. These include:

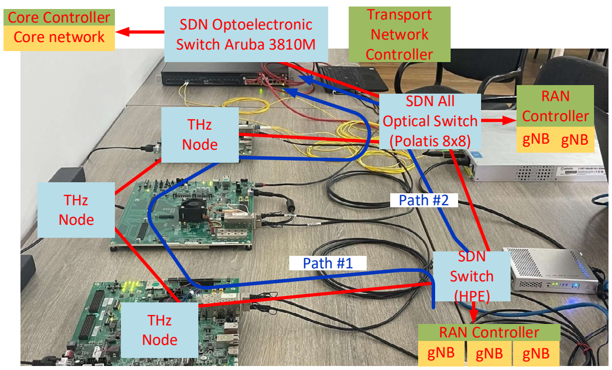

Fig. 10: Experimental demonstration of multi-technology transport network interconnecting core and gNB, b) GPU load (top) and power consumption (bottom) over time for a FL process with 10Mbps and 1Gpbs uplink channels (time in minutes).

- *get*: Retrieves all or part of the information about the running configuration and device state
- *get-config*: Retrieves all or part of the configuration details of the THz node
- *edit-config*: Submits all or part of a configuration to a target configuration THz node
- *copy-config*: Creates or replaces a target configuration for a THz node with the information from another configuration file
- *delete-config*: Deletes a target configuration for a THz node, but only if it's not running
- *lock*: Locks a target configuration for a THz node, unless a lock already exists on any part of that configuration file.
- *unlock*: unLocks a target configuration datastore, unless a lock already exists on any part of that datastore
- *close-session*: Requests the NETCONF server to terminate an open session

The NETCONF agent is also able to set up a secure connection with the NETCONF server exposing/receiving all physical parameters of the node in a standardized XML scheme. The THz node has been programmed to operate as an OF "Capable-Switch" and, therefore, being compatible with standard OF controllers. To achieve this, the OF-CONFIG protocol has been appropriately configured to operate on top of the NETCONF protocol as shown in Fig. 8. To make the THz node compatible with standard openflow controllers (i.e., RYU and Open Day Light) the XML schema of the THz node has been modified in order to address the meet the basic requirements of OF-CONFIG protocol.

4. FL-based management

The O-RAN compliant architecture allows both non-real time controllers hosted in the SMO and near real time domain specific controllers to support data collection, training and inference of ML functions. This feature unlocks the option to implement a decision support framework with distributed training and online inference capabilities for the ML models allowing each network segment to be managed autonomously through AI-assisted solutions. As shown in Fig. 3, the AI/ML assisted solutions can apply policies and control strategies taking decisions for a variety of tasks including management of radio and compute resources, prediction of mobile network traffic, etc. Information from local domain managers can be combined in a network-wide context. This capability can be exploited by the SMO to globally optimize the overall system performance.

To implement "network intelligence" functionality, we adopt a distributed approach (Fig. 1), where the main decisions are taken in a hierarchical manner. The main parameters of interest for each domain (i.e., resource utilization, energy consumption) are monitored through node exporters while the collected measurements are stored in local databases (DBs). Information stored at each local DB is accessible only by the local (domain) controllers. Based on the local datasets, each domain controller individually trains an ML model that for this study performs network traffic prediction. Then the locally trained ML models in the form of prescriptive analytics are transmitted to a centralized server that aggregates the received models creating a new global ML model.

The role of the central server aggregating the ML model is served by the SMO that continuously interacts with the local controllers. The global model is transmitted back to the domain controllers which based on their local datasets create an updated ML model that is transmitted back to the SMO. The process is repeated for several rounds until certain prediction accuracy is achieved. The predicted network traffic is used as a system resource requirement estimate.

The procedure of federated learning among real time and non-real time entities hosted at the local controllers and the SMO, respectively, can be summarized into following steps:

<u>Step 1: Predictive analytics at the local controllers</u>: In the first step, a set of local AI models are trained estimating traffic demands for each RAN NF. The local models are created adopting a federated learning technique where for each RAN NF their associated AI models are trained using local data samples. The hyperparameters (e.g. the weights and biases) of the trained models are exchanged between the local nodes on a regular basis generating a global model which is shared by all elements of the system. Sharing of the ML models among multiple RAN NFs is performed using the SMO that is responsible to orchestrate the different steps of the learning process and coordinate the operation of the whole system. The server is also responsible for the selection of the near real time controllers that will participate in the training process and for the extraction of the central (global) AI models. The learning procedure adopted is summarized below (Fig. 9):

- **Initialization**: The servers identify the ML model (e.g. LSTM, CNN, MLP) that will be trained on the local gNBs. The ML model and its hyperparameters (i.e., number of neurons, layers etc) are common for all RAN NFs involved in the learning process.
- **Client selection**: Following this, the SMO selects a random subset of RAN NF controllers that will train their ML models using their local datasets. The RAN NFs that have not been selected wait for the next federated round.
- **Training process**: The training process is initiated and the ML models for each RAN NF are calculated. The hyperparameters are stored in an appropriate format (i.e., Open Neural Network Exchange -ONNX)
- **Reporting**: each selected gNB sends its local model to the server for aggregation. The central server aggregates the

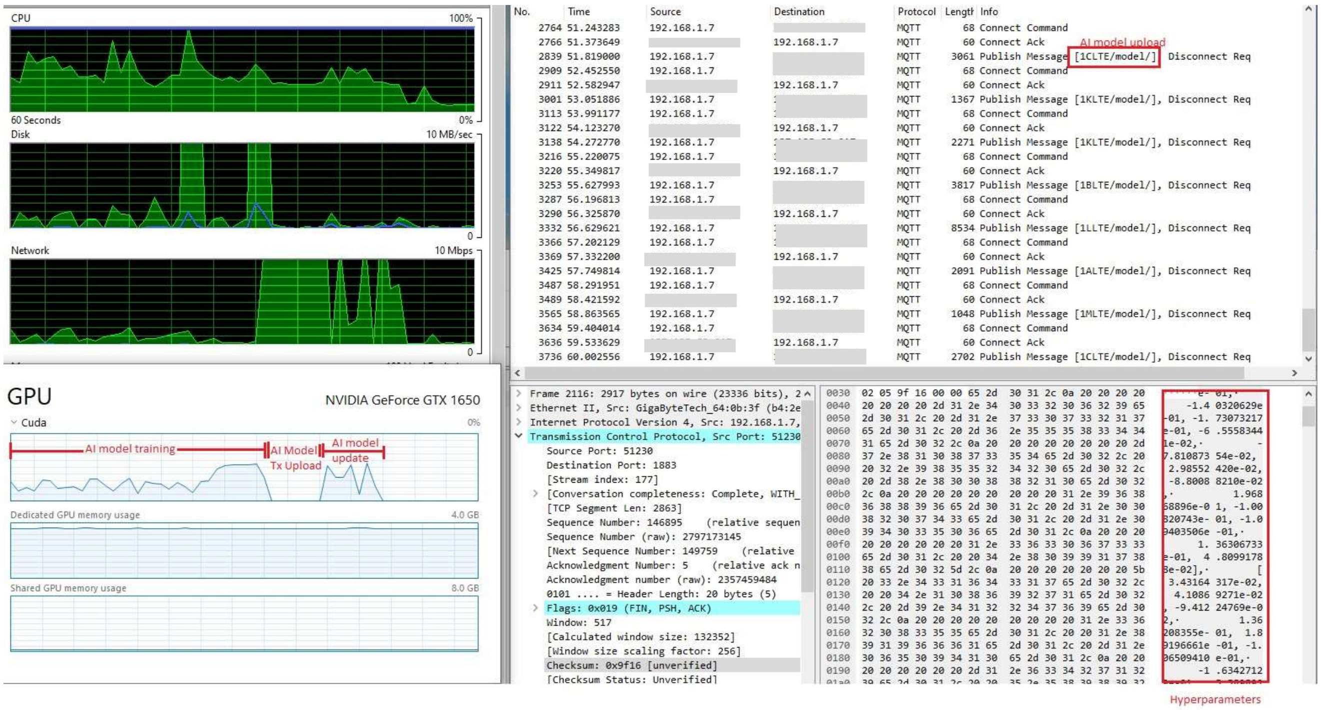

Fig. 11: Top left: CPU/DISK/Network utilization, bottom left: GPU used for AI model training, right: wireshark traces of the 6 Antennas publishing to the broker the trained AI models hyperparameters.

received models and sends back the model updates to the gNBs. The next federated round is started returning to the client selection phase. Model exchange among elements is performed using standard publish/subscribe messaging protocols (i.e., MQTT).

- **Termination**: once a pre-defined termination criterion is met (e.g., a maximum number of iterations is reached or the model accuracy is greater than a threshold) the central server aggregates the updates and finalizes the global model.

The predicted network traffic can be used to estimate the resource requirements for the NFs.

Step 2: Orchestrator-Triggered Reconfiguration: In the second stage, the output of the predictive analytics performed at local level scheme are exposed to the E2E orchestrator that prescribes optimal decisions across all layers of the system. This is achieved applying standard optimization techniques having the objective to identify optimal strategies (optical transport routing paths) minimizing a cost function (i.e., number of links/modes used. The relevant problem can be written as a constraint Optimization problem of the form: $min\ \mathcal{E}(\mathbf{x})\,, s.t.\ \ \mathrm{c}(\mathbf{x}) \leq \Lambda$ .where $\mathcal{E}(\mathbf{x})$, is the cost function under routing strategy $\mathbf{x}$ and $\mathrm{c}(\mathbf{x}) = [c_1(\mathbf{x}), \ldots, c_m(\mathbf{x})]$, $\Lambda = [\lambda_1, \ldots, \lambda_m]$ is the set of inequality constraints that need to be satisfied for the system to operate [28].

5. Experimental Validation

To experimentally validate the FL empowered 6G Transport Network we have considered a testbed (Fig. 10), comprising a combination of THz and multi-vendor optical nodes. For the THz nodes, only the baseband layer processing units implemented in FPGAs are depicted. For the optical nodes, both high and low port count optoelectronic switches have been used. The testbed comprised an 8x8 all-optical switch supporting the Standard Commands for Programmable Instruments (SCPI) control interface extended to support the OF protocol. RAN NFs are interconnected with the core network using either THz (with 1.6 Gbps capacity) or optical links (1 Gbps and 10 Gbps capacity). UEs attached to the gNBs generate traffic based on real measurements provided by OTE.

In our experiments we have considered in total 6 gNBs serving UE traffic requests. For the monitoring and AI training processes two scenarios have been examined:

- The conventional centralized approach where **all** monitoring information is collected at a single location. The full dataset is then used to train a global AI model that predicts network traffic and computational resources for the whole system.
- The federated learning approach where data traffic statistics from each RAN NF are stored locally (i.e., at a gNB/RAN NF

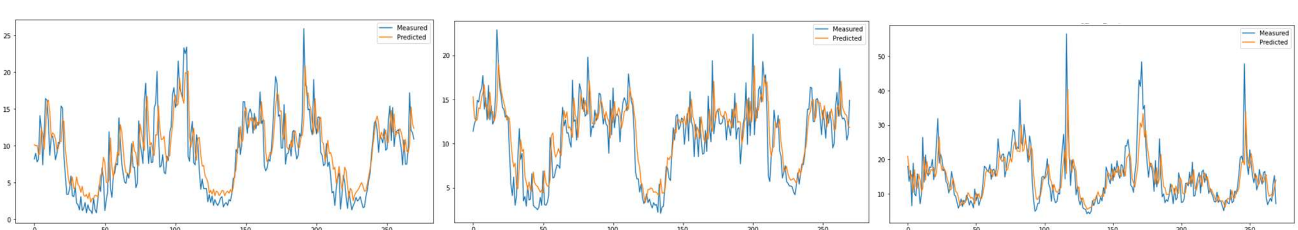

Fig. 12: Measured and predicted resource blocks time series for three different gNBs.

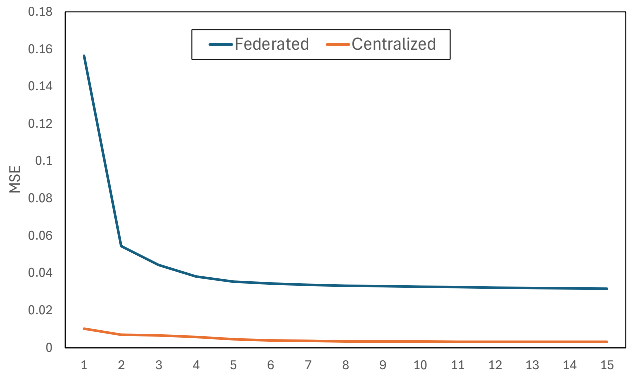

Fig. 14: Accuracy: Mean Square Error (MS) as a function of the number of epochs for the federated learning and the centralized learning approach.

level). These statistics are then used to train local ML models

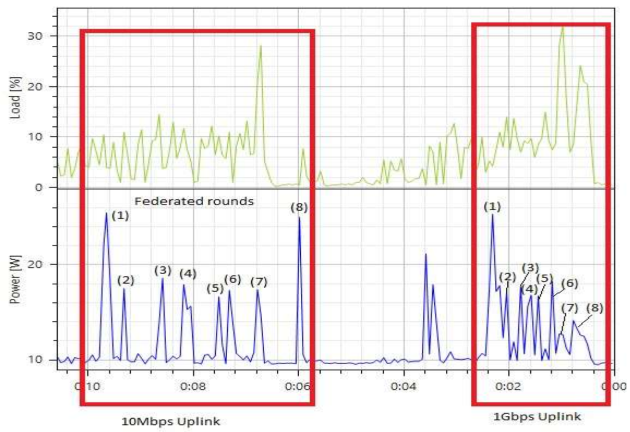

Fig. 15: GPU load (top) and power consumption (bottom) over time for a FL process with 10Mbps and 1Gpbs uplink channels (time in minutes).

that are aggregated by a centralized controller, creating a new global model that is then sent back to the domain controllers. The whole process runs for several federation rounds until convergence is achieved.

For both scenarios the AI models can be trained and executed using either hardware acceleration (GPU) or a standard x86 CPU. Fig. 11. provides a summary of the exchanged information and the protocols involved in the federated learning process, as well as the network, CPU and GPU utilization. In this set of results, the multilayer perceptron (MLP) neural network model has been used for training. Using hardware acceleration, during the training process the AI models are offloaded to the GPU leading to an average utilization of the GPU in the order of 30%. Once the training process has been completed, the trained ML models are uploaded to the centralized server using publish/subscribe messaging protocol (MQTT). The bottom right part of Fig. 11 shows the metadata of the trained ML models. Exchange of information in the federated learning models results to an increase of the network utilization as shown in top left part of Fig. 11.

Once the training process has been completed and all entities have received the global AI model, each gNB predicts the traffic (in terms of resource blocks) for the upcoming time periods. Typical prediction results for 3 gNBs are shown in Fig. 12 while a comparison of the mean square error (MSE) between the centralized and the federated learning approach is shown in Fig. 14.

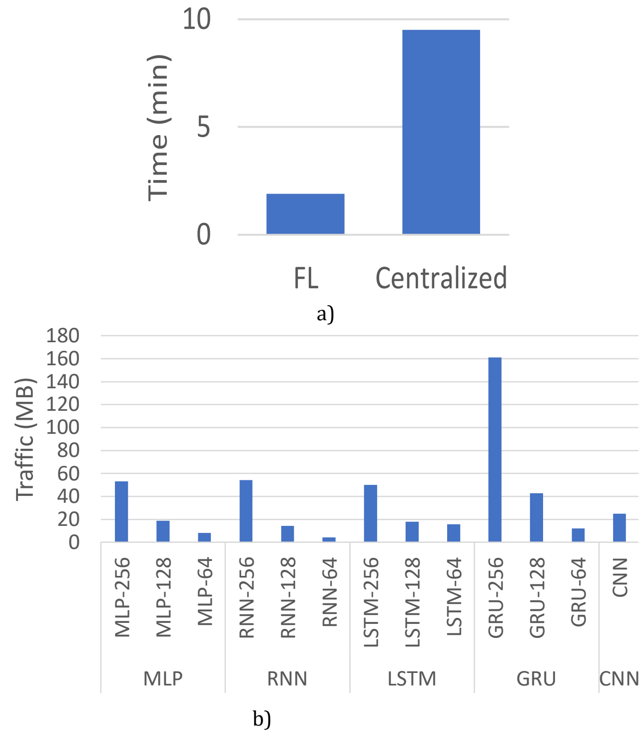

Fig. 13: a) Training time for FL vs centralized (MLP, 64) AI, b) Total volume of exchanged information for 10 federation rounds.

In both scenarios increase of the number of epochs increases accuracy. However, the centralized approach provides higher prediction accuracy levels compared to the federated learning.

Fig. 15 compares the GPU processing load and power consumption over time for different uplink bandwidths. The system has been configured to terminate after 8 federated rounds. For low channel bandwidth (10 Mbps uplink) the interarrival time of the trained ML models exchanged by gNBs increases, causing also increase in the time needed to execute the whole FL process (approx. 4 mins). As the bandwidth increases, the total time to complete the process reduces by 50%.

In Fig. 13 (a), the FL model is compared with the centralized learning model. In both cases, the ML models are offloaded to GPUs. Due to that, FL models split the processing workload in multiple devices, lower training times are required. Fig. 13 (b) shows the total volume of data (metadata) exchanged for different ML implementations and 10 FL rounds.

Finally, Fig. 16 shows the total predicted traffic in the network and the resources used (in terms of number of active links) by the system to accommodate gNB connectivity requirements. The optimization scheme used by the orchestrator tries to minimize the total number of active links. Under low demands, traffic is redirected through the optical path (path # 2, active links=2). As the traffic increases above 1 Gbps, the controller forwards traffic through the THz path (path #1, active links = 5). When the traffic exceeds the capacity of the THz links (1.6 Gbps), both paths are utilized to support network demands.

6. Conclusions

This paper proposed a 6G architecture that adopts a multi-technology transport network that integrates together THz links and optical network technologies to interconnect 6G RAN and CN domains. The proposed solution adopts SDN

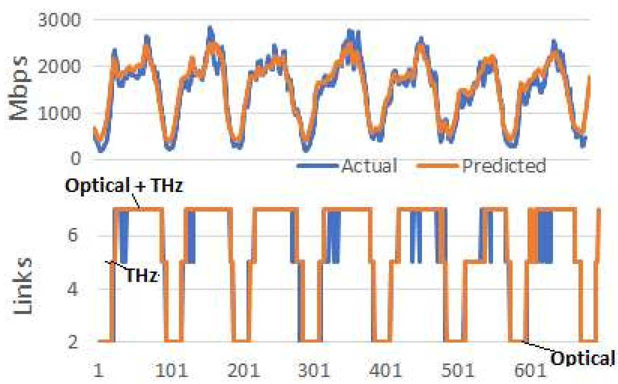

Fig. 16: (top) Total Predicted traffic using FL, (bottom) Number of active transport links.

for the control of the both the optical network and the THz technologies in order to enable operation of the infrastructure in a self-organized manner. The control required for the THz solution has been developed extending the SDN control of the optical network with suitable interfaces. E2E service provisioning and resource allocation across the 6G infrastructure is achieved through an E2E SMO layer. The E2E SMO offers intelligence capabilities through the adoption of an FL scheme and a previously developed optimization framework. The proposed architecture and the implemented control scheme were validated through an experimental demonstration supported by a 5G test-bed. To the best of the authors knowledge this was the first experimental demonstration of an autonomously controlled 6G network implementation integrating optical network technologies and THz links in a common transport network. The experimental demonstration showcased also the developed intelligent management framework empowered by FL to jointly optimize THz and multi-vendor optical network equipment supporting 6G services.

Acknowledgments. The authors would like to acknowledge the financial support of the EU funded projects 6G-SENSES and ECO-eNET.

References

1. M. Anastasopoulos et. al., A 6G Transport Network converging THz and Optical network technologies empowered by Federated Learning techniques, ECOC 2024
2. A. Tzanakaki, M. P. Anastasopoulos, and D. Simeonidou, "Converged Optical, Wireless, and Data Center Network Infrastructures for 5G Services," J. Opt. Commun. Netw. 11, A111-A122 (2019)
3. I. Mesogiti, G. Lyberopoulos, et al. Macroscopic and microscopic techno-economic analyses highlighting aspects of 5G transport network deployments. Photon Netw Commun 40, 256–268(2020).
4. 5G-COMPLETE, D2.1
5. D. Townend, R. Husbands, S. D. Walker and A. Sutton, "Challenges and Opportunities in Wireless Fronthaul," in *IEEE Access*, vol. 11, pp. 106607-106619, 2023.
6. D. Townend, S. D. Walker, A. Tukmanov and A. Sutton, "Cell Site Densification Using mmWave and sub-THz Line-of-Sight Wireless Fronthaul: A Deployment Feasibility Study," in *IEEE Transactions on Antennas and Propagation*, doi: 10.1109/TAP.2024.3468454.
7. T. Kürner *et al*., "THz Communications and the Demonstration in the ThoR–Backhaul Link," in *IEEE Transactions on Terahertz Science and Technology*, vol. 14, no. 5, pp. 554-567, Sept. 2024
8. C. Castro, R. Elschner, T. Merkle, C. Schubert and R. Freund, "Experimental Demonstrations of High-Capacity THz-Wireless Transmission Systems for Beyond 5G," in *IEEE Communications Magazine*, vol. 58, no. 11, pp. 41-47, November 2020
9. O. Stiewe *et al*., "Experimental Demonstration of a High Capacity THz-Wireless Dual Link Transmission System," *2023 53rd European Microwave Conference (EuMC)*, Berlin, Germany, 2023, pp. 468-471
10. R. Mundlamuri, S. Badran, R. Gangula, F. Kaltenberger, J. M. Jornet and T. Melodia, "5G Over Terahertz Using OpenAirInterface," 2024 19th WONS, Chamonix, France, 2024, pp. 29-32
11. S. Ahearne *et al*., "Integrating THz Wireless Communication Links in a Data Centre Network," *in proc. of IEEE 5GWF* 2019
12. H. Elayan, O. Amin, R. M. Shubair and M. -S. Alouini, "Terahertz communication: The opportunities of wireless technology beyond 5G," 2018 International Conference on Advanced Communication Technologies and Networking (CommNet), Marrakech, Morocco, 2018
13. A. Manolopoulos, V. Alevizaki, M. Anastasopoulos A. Tzanakaki, "An AI-Assisted Framework for Lifecycle Management of Beyond 5G Services," in *IEEE Access*, vol. 12, pp. 179449-179463, 2024
14. O-RAN.WG1.OAD-R003-v12.00, O-RAN Architecture Description, Technical R003, June 2023
15. 3GPP TS 38.300, : NR; NR and NG-RAN Overall description; Stage-2, Ver. 18.3.0, 26-09-2024
16. O-RAN TS: "A1 interface: General Aspects and Principles".
17. O-RAN.WG9.XTRP-MGT.0-R004-v09.00, O-RAN Management interfaces for Transport Network Elements 9.0, Oct. 2024
18. O-RAN.WG4.CTI-TCP.0-R003-v04.00, O-RAN Cooperative Transport Interface Transport Control Plane Specification 4.0, June 2023
19. O-RAN.WG2.AIML-v01.03, O-RAN AI/ML workflow description and requirements 1.03
20. K. Katsalis et al., "Implementation experience in multi-domain SDN: Challenges, consolidation and future directions", Computer Networks Volume 129, Part 1, 24 December 2017, Pages 142-158
21. A. -I. Manolopoulos, M. P. Anastasopoulos, V. -M. Alevizaki and A. Tzanakaki, "Optimal Service Provisioning in Mobile 5G and Beyond Systems," *IEEE Trans. on Services Computing*, vol. 16, no. 4, pp. 2841-2854, 1 July-Aug. 2023
22. V. -M. Alevizaki, M. Anastasopoulos, A. -I. Manolopoulos and A. Tzanakaki, "Distributed Service Provisioning for Disaggregated 6G Network Infrastructures," *IEEE Transactions on Network and Service Management*, vol. 20, no. 1, pp. 120-137, March 2023
23. O-RAN.WG11.TR.AIML-Security.O-R004-v02.00.02
24. M. Furdek, C. Natalino, A. Di Giglio and M. Schiano, "Optical network security management: requirements, architecture, and efficient machine learning models for detection of evolving threats [Invited]," in *Journal of Optical Communications and Networking*, vol. 13, no. 2, pp. A144-A155, February 2021
25. Elaine Wong, Sourav Mondal, and Lihua Ruan, "Machine learning enhanced next-generation optical access networks—challenges and emerging solutions [Invited Tutorial]," J. Opt. Commun. Netw. 15, A49-A62 (2023)
26. IETF, Network Configuration Protocol (NETCONF), RFC 6241, June 2011
27. WORTECS Deliverable 3.2, "Common Analog and Digital Baseband Design for Flexible Radio and Optical Transceiver", November 2018
28. P. Georgiadis, M. Anastasopoulos, A. -. Manolopoulos, V. -. Alevizaki, N. Nikaein, and A. Tzanakaki, "Demonstration of Energy Efficient Optimization in Beyond 5G Systems supported by Optical Transport Networks," in *Optical Fiber Communication Conference (OFC) 2023*, Technical Digest Series (Optica Publishing Group, 2023), paper W4F.4.