Optimizing 5G Network Slicing: An End-to-End Approach with Isolation Principles

Xhulio Limani*, Arno Troch*, Chieh-Chun Chen†, Chia-Yu Chang‡, Andreas Gavrielides*, Miguel Camelo*
Johann M. Marquez-Barja*, and Nina Slamnik-Kriještorac*
*University of Antwerp - imec, IDLab - Faculty of Applied Engineering, Belgium
†EURECOM, Sophia-Antipolis, France, ‡Nokia Bell Labs, Antwerp, Belgium

***Abstract*—5G Standalone (SA) networks introduce a range of new applications, including enhanced Mobile Broadband (eMBB), Ultra-Reliable Low-Latency Communication (URLLC), and massive Machine-Type Communications (mMTC). Each of these applications has distinct network requirements, which current commercial network architectures, such as 4G and 5G Non-Standalone (NSA), struggle to meet simultaneously due to their one-size-fits-all design. The 5G SA architecture addresses this challenge through Network Slicing, creating multiple isolated virtual networks on top a single physical infrastructure. Isolation between slices is crucial for performance, security, and reliability. Each slice owns virtual resources, based on the physical resources (e.g., CPU, memory, antennas, and network interfaces) shared by the overall infrastructure. To deploy Network Slicing, it is essential to understand the concept of isolation. The Third Generation Partnership Project (3GPP) is standardizing security for Network Slicing, focusing on authentication, authorization, and slice management. However, the standards do not clearly define the meaning of isolation and its implementation in the infrastructure layer.**

In this paper, we define and showcase a real-life Proof of Concept (PoC), which guarantees isolation between slices in 5G SA networks, for each network domain i.e., Radio Access Network (RAN), Transport Network (TN), and 5G Core (5GC) network. Furthermore, we describe the 5G SA architecture of the PoC, explaining the isolation concepts within the Network Slicing framework, how to implement isolation in each network domain, and how to evaluate it.

***Index Terms*—5G, Network Slicing, Isolation, O-RAN**

I. Introduction

5G Standalone (5G SA) networks are opening the doors to a multitude of new applications i.e., enhanced Mobile Broadband (eMBB), Ultra-Reliable Low-Latency Communication (URLLC), and massive Machine-Type Communications (mMTC) [1]. For instance, the eMBB applications, such as high-definition video streaming and Augmented Reality (AR), require a high throughput of more than 100 Mbps and moderate latency of less than 50 ms [1]. On the other hand, URLLC applications, such as remote surgery and Vehicle-to-Everything (V2X), require extremely low latency of less than 5 ms and high reliability of 99.9999% [1]. Finally, mMTC applications, such as smart cities and IoT ecosystems, require high connection densities of up to 1,000,000 devices per cell [1]. As a result, the network must be able to address simultaneously multiple requirements for different types of applications in terms of throughput, latency, and reliability. However, current cellular network architectures like 4G and 5G Non Standalone (5G NSA) cannot effectively differentiate these applications based on their network requirements. Therefore, the *one-size-fits-all* network model is becoming obsolete [2].

Network slicing, a pivotal technology introduced with 5G SA, enables the creation of multiple virtual networks on a single physical infrastructure, each tailored to satisfy precise network requirements i.e., URLLC, eMBB, and mMTC. Nonetheless, a critical challenge in implementing effective Network Slicing is ensuring isolation [2]. Isolation guarantees that each slice operates independently within the same physical network infrastructure, without interference from other slices. This means ensuring that the activities or failures of one slice do not compromise the performance, security, and reliability of another slice. For example, the high throughput required from AR applications should not affect the ultra-low latency required from vehicular applications.

Third Generation Partnership Project (3GPP) is actively standardizing security measures for Network Slicing, with critical features introduced from Release 15 to Release 17, and further studies ongoing for Release 18. Release 15 focused on security management, including User Equipment (UE) authorization, confidentiality and integrity protection of network slice identifiers, and network slice-specific Network Functions (NFs) authorization. Release 16 introduced Network Slice Specific Authentication and Authorization (NSSAA), adding an extra layer of security by ensuring only authenticated and authorized entities access certain slices. Release 17 focused on Application Function (AF) authorization with confidentiality protection of network slice identifiers, ensuring secure and authorized application interactions, further enhancing slice isolation and integrity.

While these releases primarily address authentication methods and slice management/orchestration, ensuring isolation at the lower layers of the network i.e., infrastructure, remains a fundamental challenge. The actual state-of-the-art also has its shortcomings in that regard [2]. Effective isolation at the infrastructure layer provides a solid foundation for deploying a secure, reliable, and efficient Network Slicing mechanism.

In this paper we provide a Proof of Concept (PoC) of 5G Network Slicing with isolation principles, ensuring that each slice operates independently and securely, in terms of i) performance, ii) security, and iii) dependability. We start by describing the main components of the 5G SA architecture, followed by an explanation of the overall Network Slicing architecture, highlighting the importance of isolation for all the domains involved i.e., Radio Access Network (RAN), Transport Network (TN), and 5G Core (5GC). Furthermore, we explain the practical implications of isolation in terms of performance, security, and reliability. We implement and validate isolation in our 5G SA real-world testbed, demonstrating how isolation can be achieved across different slices in each network domain.

II. 5G SA architecture background

The 5G SA architecture consists of three main components: i) UE with a Universal Subscriber Identity Module (USIM), ii)

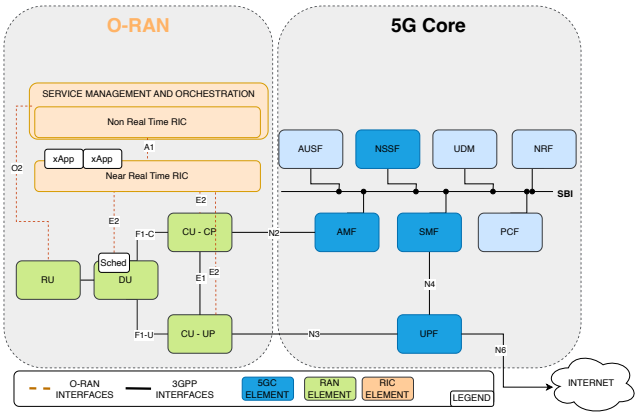

Fig. 1: 5G Standalone Architecture with O-RAN paradigm.

the RAN, which connects UEs to the network, and iii) the 5GC, the central unit managing authentication, session, mobility, and network policy control [3]. At its core, 5G SA employs a cloud-native infrastructure to enhance the modularity and flexibility of the 5G network. The deployment of such infrastructure is microservices-based [4], involving virtualization, container-based deployments of Network Function Virtualization (NFV), and orchestration techniques [4].

Such a flexible and modular approach enables i) dynamic network configuration of resource allocation e.g., CPU, memory, and radio resources, crucial for enabling Network Slicing, and ii) the portability of NFs across different physical locations and different hardware. To fully leverage the flexibility and modularity of the 5G SA network, the communication between the 5GC and RAN needs to be facilitated through open and standardized communication channels. The Open Radio Access Network (O-RAN) paradigm standardizes and facilitates the integration of the different segments of the 5G architecture by considering open interfaces between RAN and 5GC. In this section, we describe the main components of the 5G SA architecture illustrated in Figure 1.

1) ***RAN****:* In practice, the O-RAN paradigm involves the separation of RAN software and hardware, resulting in de-segregation into Radio Unit (RU), Central Unit (CU), and Distributed Unit (DU) [5]. The CU handles higher-level RAN functions, such as session management, mobility management, encryption, and interfaces with the 5GC through the N2 and N3 interfaces. Furthermore, O-RAN introduces two types of RAN Intelligent Controllers (RICs): the Non-Real-Time RIC, which oversees the high-level orchestration of the RAN, and the Near-Real-Time RIC, which manages precise control policies, including slicing, scheduling, and load balancing.

Moreover, the O-RAN architecture integrates xApps [5], software applications that operate on the Near-Real-Time RIC and are used by developers to interact with the RAN to deploy applications e.g., dynamic control and optimization of RAN resources. The flexibility and modularity offered by the O-RAN architecture enable Network Slicing through the customized allocation of radio resources e.g., Resource Blocks (RBs), per slice, enabling customized scheduling mechanisms to guarantee network requirements for different types of applications e.g., URLLC, eMBB and mMTC. However, within a Network Slicing context, it is important to guarantee isolation between the slices, allocating to each slice a dedicated pool of radio resources, as we further discuss in Section III.

2) ***5G Core Architecture****:* The 5GC employs a Service-Based Architecture (SBA) with NFs as key components (Table I). These NFs are software entities responsible for networking tasks e.g., authentication, routing, and forwarding. Enabled by NFV, the NFs operate on virtual machines and containers, like Docker[1], and communicate through a Service Based Interface (SBI), which facilitates Application Programmable Interface (API)-based interactions. The Key components of the 5GC are shown in Figure 1 as the block in dark blue.

The Access and Mobility Management Function (AMF), a NF part of the control plane, manages user registration, handovers, and authentication over the N2 interface using the Next-Generation Application Protocol (NGAP)[2]. The Session Management Function (SMF), another control plane NF, manages session contexts, coordinates session setup with the AMF, and manages the data plane session via the N4 interface. The SMF is responsible for the allocation of Internet Protocol (IP) addresses to the UEs and the coordination of session setup with other NFs. By isolating the SMF, the integrity of these sessions is maintained, ensuring that the operations and performance of one slice do not affect the rest of the network. The User Plane Function (UPF) is in charge of the data plane. The UPF handles data routing and policy enforcement, deep packet inspection, charging data collection, and interfacing directly with the gNodeB (gNB) via the N3 interface. The Network Slice Selection Function (NSSF) handles (i) network slice selection and (ii) access to slices based on the configurations of the UE, by coordinating with the AMF and the SMF through the N22 interface. The 5GC uses Slice/Service Type (SST), Session Description (SD), Data Network Name (DNN), and 5G QoS Identifier (5QI) to manage slice descriptions and slice configuration. These parameters are crucial for creating isolated network slices, ensuring that each slice operates independently and securely.

Network Slicing within the 5GC is enabled by deploying and configuring, multiple NFs of the same function. For instance, it is possible to build one SMF and one UPF for a dedicated slice. These NFs need to be configured with the correct values of SST, SD, and DNN. Furthermore, each of these NFs i.e., SMF and UPF, must be isolated to ensure that operations in one slice do not affect another slice, as described in Section III.

III. Concept and Principles of Isolation

In this section we focus on Network Slicing, i) defining the concept of isolation, iii) discussing where isolation needs to be applied, and iv) providing an overview of what means isolation in practice, thereby highlighting its impact on performance, security, and dependability.

The Network Slicing architecture, illustrated in Figure 2, comprises three functional layers i.e., the Service Layer, Network Slice Layer, and Resource Layer [6]. The Service Layer is where applications run and interact with the network, using APIs, to ensure they meet specific Service Level Agreements (SLAs). The Network Slice Layer creates and manages slices based on the network requirements requests coming from the Service Layer. Each slice is built on top of a dedicated pool of virtual resources e.g., Virtual Machines (VMs), vCPU, vRAM, RBs coming from the Resource Layer.

The Resource Layer manages network resources, including connectivity, processing, and storage. It is composed of three

[1]Docker: https://www.docker.com/resources/what-container/

[2]NGAP protocol: https://docs.magmaindia.org/Free5gc_5gCore/amf/amf.html

TABLE I: Network Functions and Configuration settings.

Architecture Segment	Name Element	Interface	Name Paramenter	Value	Scope
5GC	AMF	N2,N11	PLM_ID	Num	Identifies mobile networks globally.
	SMF	N4,N11	DNN	String	Specifies the name of the network to which the device connects
			5QI	1-90	Defines the specific QoS characteristics of data traffic
	UPF	N4,N3,N6	DEV	String	Interface where the data traffic pass
	NSSF	SBI	SST	URLLC=1 eMBB=2 mMTC=3 V2X=4	Identifies the type of service the slice is intended to support
			SD	24 bit (optional)	Distinguishes between multiple network slices that share the same SST
RAN	DU	E2	RBs	12 sub-carriers per RB	Small units that divide the radio frequency (spectrum) and are used to transmit data.
	CU	E2,E1, F1C,F2C	RRM Policy Ratio	Dedicated Ratio	The amount of resources that are dedicated to a slice and cannot be used by other slices
				Min Ratio	The minimum guaranteed resources that a slice will always have available.
				Max Ratio	The maximum limit of resources that a slice can use,if resources are available

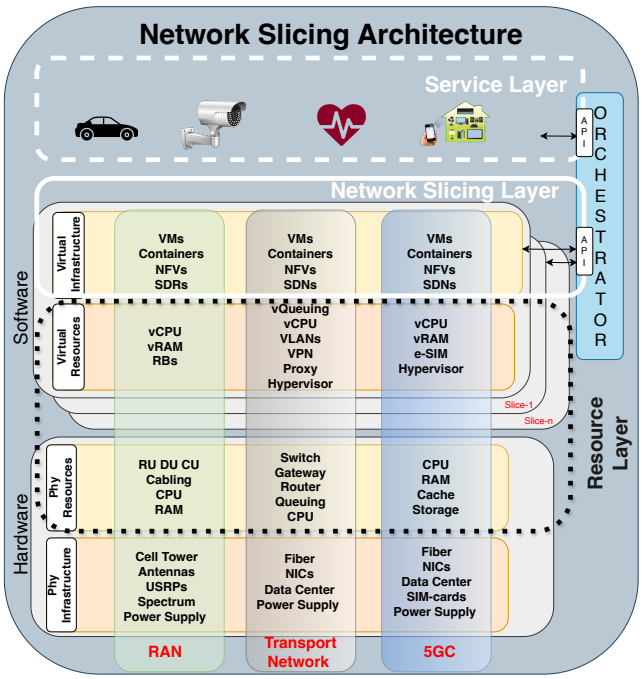

Fig. 2: Network Slicing Architecture.

domains: the RAN, the TN, and the 5GC. Each domain contains static resources, such as hardware (e.g., antennas, routers, and servers), and dynamic resources, such as software (e.g., vCPU, vRAM, VM and NFs).

Since slices are created with resources from the same physical infrastructure, multiple slices operate simultaneously using shared resources. For example, the URLLC slice serving mission-critical services, such as vehicular applications, shares the same infrastructure as the eMBB one dedicated to AR applications. However, despite this sharing, each slice must operate independently as if it is a separate physical network.

Hence, network slices cannot interfere with each other in terms of i) performance, ii) security, and iii) dependability, further discussed in Section III-B1, Section III-B2, and Section III-B3.

A. Network Slice isolation domains

Ensuring isolation in the lower levels i.e., Resource Layer, provides a solid foundation for building a secure, reliable, and efficient Network Slicing mechanism. Each network domain i.e., RAN, TN, and the 5GC, plays a pivotal role in the overall functionality and efficiency of the Network Slicing architecture.

*1) **Isolation in the RAN**:* The RAN domain is a critical aspect of cellular networks, as it often acts as a traffic bottleneck due to limited radio resources. In the RAN, the medium access is managed through a combination of scheduling and resource allocation mechanisms [5]. The UEs request radio resources based on i) the Quality of Service (QoS) needed by specific type of application i.e., URLLC, eMBB, mMTC, and ii) the amount of traffic the UE generates or consumes. In response, the gNB dynamically allocates radio resources to UEs in the form of Physical Resource Blocks (PRBs). These allocations are typically made in both time and frequency domains, allowing flexible and efficient use of the available radio resource of the RAN.

To enable Network Slicing, 5QI values are used to manage and prioritize traffic based on the QoS requirements of the applications. Each 5QI value corresponds to a specific set of QoS characteristics, such as latency, reliability, and packet error rate. However, while 5QI helps in defining QoS profiles, it does not inherently enforce resource isolation between slices. When multiple UEs configured with different values of 5QI generate or consume data traffic, they compete for the same pool of radio resources provided by the gNB. When accommodating a UE request, network slices should not access radio resources that belong to another slice. Proper isolation in the RAN domain ensures that each slice receives its pool of radio resources without interference from other slices, preventing performance degradation and maintaining service quality, as further described in Section III-B1.

*2) **Isolation in the 5GC**:* On the 5GC side, network slices are created by linking together different NFs e.g., one SMF and one UPF. Each slice is identified by the SST and SD value pair, which define i) the logical network and ii) the type of slice i.e, eMBB, URLLC, and mMTC. Each couple of SST and SD can be identified by the UEs using the related DNN. To ensure proper isolation, both physical and virtual resources in the 5GC infrastructure must be allocated to each slice. This includes allocating specific vCPUs, memory, storage, and network interfaces to each slice.

For instance, each SMF can be configured to give different pools of IP addresses per DNN (so per slice). That means that is possible to define different pools of IP addresses per UPF. Conceptually, having a dedicated UPF for each slice, with a specific pool of IPs addresses, means having distinct networks. However, if these UPFs are not properly isolated from one another, then despite having separate logical networks, they would function as a single network. For example, having two UPFs in the same VM, with *IP_forwarding* enabled, would allow UEs from one slice-A to access the data of slice-B, similar to how users on one network could access data on another network, leading to significant security and privacy issues. In Section IV we describe the practical implementation of it within our PoC.

TABLE II: Isolation Mechanisms and Implications in Different Domains.

Domain	Isolation Mechanism	Performance Implications	Security Implications	Dependability Implications
RAN	Dedicated PRBs, isolated scheduling	Ensures consistent network requirements for throughput, jitter, latency, and packet loss without interference	Prevents direct communication between UEs of different slices	Maintains service quality, prevents performance degradation, and ensures reliable communication
5GC	Separate VMs for control and data plane functions	Maintains session integrity, avoids resource contention, ensures stable throughput and low latency	Isolates data traffic by allocating dedicated virtual resources, preventing data leakage between slices. Ensures secure session management and authentication processes are isolated per slice.	Prevents impact of failures on other slices, maintains reliability and operational continuity
TN	Separate fiber links, VLANs, MPLS	Ensures data transport efficiency, avoids bottlenecks, and maintains optimal throughput and low latency	Isolates data flows, prevents unauthorized access, ensures secure data transport	Maintains optimal performance, prevents resource monopolization, and ensures consistent data delivery

*3) **Isolation in the TN**:* The TN plays a crucial role in ensuring efficient and reliable data transport between the RAN and the 5GC elements, for both the control plane and the data plane. Furthermore, the TN is in charge of establishing the links between the elements of the 5GC domain, described in Section II-2, and the elements of the RAN domain described in Section II-1.

When data flows from different UEs belonging to different slices are sent to the data network, the data passes first through the RAN, then they are forwarded to the related data plane (UPF) via the TN. Hence, the data flows through shared TN elements, e.g., switches, network interfaces, and fiber. The TN elements have limited resources, such as bandwidth and processing capacity, which can cause a bottleneck in the entire 5G network. For example, one slice might handle AR applications, which require high bandwidth and low latency. Another slice might handle vehicular communication data, requiring low latency but different throughput. The high data throughput generated by video traffic can occupy a significant portion of the resources of the TN. This occupation of resources can degrade the performance of the vehicular communication slice by causing delays, packet loss, or high jitter.

Moreover, slices often require distinct network policies to meet their specific needs in terms of performance, security, and privacy. Implementing different network policies within the same TN infrastructure is essential to accommodate the unique demands of each slice. Isolation ensures that each slice operates independently, maintaining its policies without interference with other slices. Isolation can be achieved through physical means, such as dedicating specific fibers or switches to each slice, or through logical means, such as using Virtual Local Area Networks (VLANs), Multiprotocol Label Switching (MPLS), or NFV techniques to create isolated virtual network segments. These methods allocate distinct resources and manage traffic flows separately, preventing one slice from monopolizing the TN resources and degrading the performance of others.

B. Impact of isolation

To better understand the implication of isolation in Network Slicing, and how to evaluate the isolation, we analyzed the concept of isolation from three principles i.e., performance, security, and dependability.

*1) **Performance**:* Isolating slices, in terms of performance, means that the traffic flow within a specific slice maintains consistent network requirements in terms of latency, throughput, jitter, and packet loss, regardless of the data flow present in the other slices. Isolation ensures that the high demand of resources from one slice does not degrade the performance of another slice. Such isolation must be present from the RAN to the 5GC, till the moment the data leave the last hop within the 5G SA network.

*2) **Security**:* In the context of network slicing, security is a critical measure to safeguard against attacks, disruptions, and data leaks. For that purpose, isolation between slices is crucial for preserving i) confidentiality, ii) integrity, and iii) availability of data and services, within each slice.

In practice, it means that UEs that belong to different slices cannot communicate directly with each other, even if they are connected to the same network infrastructure e.g., two UEs connected to the same gNB that belong to different slices, cannot directly reach (e.g., ping) each other (taking in account the layers up to L3). This strict isolation ensures that potential attacks from the outer boundaries of a certain slice, cannot propagate within the slice.

For example, a vehicular application within a Smart City that runs within the URLLC slice, between vehicles and Road Side Units (RSUs), requires strict confidentiality and data integrity. The eMBB slice used by tourists in the city for AR applications, is less sensitive but still requires security to prevent misuse. Isolation in terms of security ensures that vehicular data remains confidential and protected from potential threats originating from tourists using AR applications in the city.

*3) **Dependability**:* In terms of dependability, isolating slices ensures that any failure in terms of performance or security in one slice does not impact the operation and reliability of other slices. Dependability also involves applying network policies in a manner that such policies remain confined to their respective slices. The deployment of isolated slices is essential for maintaining the overall stability and robustness of the entire network, ensuring that each slice performs independently from the other slices and the rest of the network.

For instance, within a smart city scenario, vehicular communication requires extremely high reliability and minimal latency to ensure safe and efficient operations. The slice used for AR applications requires consistent performance but can tolerate slight delays. Dependability ensures that a failure in the eMBB slice does not affect the URLLC slice.

IV. 5G SYSTEM WITH NETWORK SLICING AND ISOLATION PRINCIPLES

Based on the concepts and principles of *isolation* discussed in the previous section, in this section we illustrate our 5G SA deployment, enabling isolation in Network Slicing. Our PoC involves real devices and open-source solutions. The main components of our testbed are illustrated in Table III. Our 5G SA testbed is designed to be O-RAN oriented. The testbed runs open-source solutions software to deploy a 5G SA network. This setup integrates Open Air Interface (OAI)[3] for the RAN functionalities and Open5GS[4] for the 5GC, with FlexRIC[5] serving as the RIC to facilitate advanced radio network management. To conduct over-the-air transmission experiments within our real-world testbed, we obtained the appropriate spectrum licenses, which include 50 MHz within the 5G NR band 77. For the 5GC, our infrastructure leverages three separate VMs

[3]OAI: https://gitlab.eurecom.fr/oai/openairinterface5g

[4]OPEN5GS: https://open5gs.org/

[5]FLEXRIC; https://gitlab.eurecom.fr/mosaic5g/flexric

TABLE III: Testbed Components Overview.

Component	Software	Role	Slices	Characteristics
5GC	Open5GS	Control Plane (VM3)	-	16GB of RAM, Intel Xeon E5-2620 v4-4 cores at 2.10GHz, 120GB of storage space.
		UPF (VM1)	URLLC	16GB of RAM, Intel Xeon E5-2620 v4-4 cores at 2.10GHz, 120GB of storage space.
		UPF (VM2)	eMBB	140GB storage, Intel Xeon Silver - 4 cores 2.40GHz, 8GB RAM.
RAN	OAI	gNB	eMBB-URLLC	Intel i7-11700K - 8 cores, 64GB RAM, NVIDIA RTX 3060 GPUs, USRP B210, dual 10 GB SFP
RIC	FlexRIC	RAN controller (VM3)	-	Same specifications as the CP VM for 5GC.
UEs	Real equipment	users	eMBB	Intel NUC connected to a Quectel RM500Q 5G module.
			URLLC	Same as eMBB UEs.
MGEN Server	-	Traffic generator/receiver	-	Remote VM used to generate/receive data traffic, synchronized with nodes with an error of 0.003 ms.

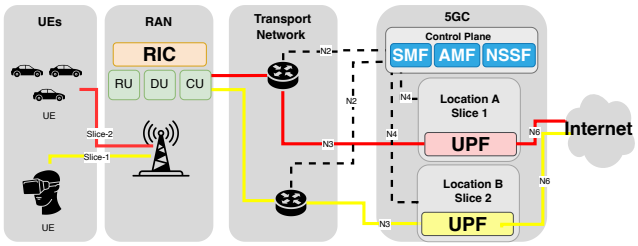

Fig. 3: Architecture of the Proof of Concept.

to distinctly support different types of slices and control plane elements related to the 5GC and the RIC. Importantly, our UE comprises real devices in the testbed environment. Each UE can be connected to one slice at a time.

Our architecture, illustrated in Figure 3, is designed to support flexible and modular implementations, adhering to SBA principles by leveraging Software Defined Networking (SDN) and NFV. We separate the control plane functions from data plane functions, creating a distributed 5G SA architecture. The key components of the different domains involved on the 5G SA network i.e., RAN, TN, and 5GC, are strategically distributed to i) enable total isolation between slices in all three domains, combining isolation of physical and virtual resources, and ii) optimize and isolate the resources of the Resource Layer to meet the specific requirements for different slices i.e., eMBB, URLLC, and mMTC. Such architecture guarantees isolation between the slices in terms of performance, security, and dependability, as discussed in Section III.

*1) **RAN Domain**:* Network Slicing is enabled using 5QI values to manage and prioritize traffic based on the QoS requirements of different applications/services. However, as discussed in Section III-A1, 5QI values alone are insufficient for ensuring proper isolation. Dedicated pools of radio resources must be allocated to each slice.

In our PoC, isolation is achieved through the dynamic allocation of PRBs among different slices. Our RAN setup leverages the O-RAN paradigm, which incorporates a RIC and xApps. These xApps guide the RIC on PRB allocation within the DU based on SST and SD values. The PRB allocation process involves defining dedicated, minimum, and maximum quotas for each slice, as defined in the 3GPP TS 28.541 for Release 16 [7]. This dynamic allocation ensures that each slice receives its own dedicated pool of available radio resources, preventing the utilization of radio resources allocated to other slices.

*2) **5GC Domain**:* Our 5GC deployment uses a distributed architecture to ensure robust isolation. The control plane elements e.g., AMF, SMF, and NSSF, are deployed on separate VMs. This separation simplifies management and orchestration within the 5G SA network. On the other hand, the UPF for each slice is deployed on its own VM, with each VM hosted in a different data center.

In the context of Network Slicing, the UPF has a critical role. It serves as the entry and exit point for traffic within the 5G SA network, handling routing and forwarding operations between the RAN and external data networks. Deploying a separate UPF for each slice across distinct VMs and data centers ensures that i) data plane traffic remains isolated, and ii) the network policies applied for one slice, are not applied to the rest of the 5G SA network. This guarantees that one slice i) does not use the pool of resources of other slices, and ii) has independent network policies. To maintain robust isolation, each UPF is provisioned with dedicated vCPU cores, memory banks, storage devices, and network interfaces.

*3) **TN Domain**:* In our PoC, isolation in the TN domain is realized by distributing separate fiber links to the data plane of each slice. This physical separation ensures that data traffic between the RAN and 5GC remains isolated across different slices. Each UPF is connected via dedicated connections, which mitigates potential bottlenecks and interference associated with shared network resources. This setup guarantees that each slice operates independently, maintaining optimal performance without interference from other slices. By managing traffic flows separately, we prevent resource monopolization and performance degradation, thereby ensuring that each slice meets its specific network requirements.

V. Validation of Isolation Principles

To validate the configuration and the design of our 5G SA setup discussed in Section IV, we conducted real-world experiments using two UEs. We use a dedicated VM for control plane operations (VM3), and a separate VM for the data plane of each slice (VM1 and VM2 in Table III) Using that configuration, we manage the 5GC more easily since the control plane is centralized in a unique VM. Furthermore, to obtain the one-way latency, first, we synchronized the nodes with an error of 0.003 ms, and then we developed an in-house solution, based on the log from MGEN[6].

In this experiment, we consider a smart city scenario where mission-critical applications i.e., vehicular applications, are performing within the URLLC slice. Simultaneously, tourists use AR applications on their mobile devices, which require high-bandwidth downlink connections provided by the eMBB slice. Without proper isolation, the bandwidth demand of AR applications can disrupt the URLLC slice, potentially causing accidents and traffic congestion.

To validate our PoC, discussed in Section IV, we considered three dimensions, discussed in Section III i.e., performance, security, and dependability. In our experiment, we consider an aggregated throughput in both slices. We generated approximately 50 Mbps of UDP traffic for slice 1 (AR application) and 25 Mbps for slice 2 (vehicular application) using MGEN. The AR application (slice 1) should not affect the performance of the vehicular application (slice 2). Additionally, errors or disturbances in slice 1 should not affect slice 2.

It is important to remember that in the design of our network, the TN and 5GC domains are static, while the resource allocation in the RAN domain is dynamic.

Figure 4 shows the results in the Service Layer, which means that performance is measured on the user side, when the data arrives in the downlink direction. That means that the flows traverse all the network domains, before arriving at the UE. In this way, we are able to observe the isolation affecting the overall end-to-end network.

[6]MGEN: https://www.nrl.navy.mil/Our-Work/Areas-of-Research/Information-Technology/NCS/MGEN/

In Figure 4 the graph on the top represents the achieved throughput, while the one on the bottom shows the one-way latency. Each graph includes three axes: i) the x-axis for timestamp, ii) the y-axis for latency or throughput, and iii) the secondary y-axis for PRBs allocation.

In the first segment of the graph, until timestamp *14:55:00*, the traffic in both slices is stable in terms of throughput and latency. The average throughput of slice 1 is 49.20 Mbps with a standard deviation of 0.95 Mbps and a latency of 10 ms with a standard deviation of 5.76 ms. The average throughput of Slice 2 is 24 Mbps with a standard deviation of 0.02 Mbps and a latency of 8.04 ms with a standard deviation of 0.11 ms. In that segment, the PRBs allocation varies due to the UEs demand, and slight variations of the channel condition. AR applications require high throughput, demanding more radio resources to consume and generate data. On the other hand, vehicular applications need low latency and low throughput. To ensure the reliability and security of vehicular applications, it is necessary to guarantee a dedicated and isolated pool of radio resources for slice 2. This prevents slice 1 from compromising the network requirements of AR applications at the expense of slice 2, thereby maintaining the integrity and performance of slice 2.

From timestamp 14:55:00, we change the network configuration in the RAN domain to i) reduce the pool of radio resources of slice 1 and see if it invades the pool of resources of slice 2, and i) we allocate more radio resources to slice 2 to guarantee required network performance for mission-critical applications such as vehicular. The graph in Figure 4 shows the difference in radio resource allocation, from timestamp 14:55:00.

The pool of resources for slice 1 has 30% of the total PRBs, while slice 2 has 70%. Consequently, the average throughput of slice 1 is 36.19 Mbps with a standard deviation of 5.33 Mbps, and a latency of 2 seconds with a standard deviation of 374 ms. On the other hand, Slice 2 maintained stable performance, showing an average throughput of 24.6 Mbps with a standard deviation of 0.01 Mbps, and latency of 8.43 ms with a standard deviation of 0.12 ms. Despite the reduction of resources for slice 1, the throughput and latency of slice 2 remain stable, demonstrating that the slices are isolated in terms of performance, as discussed in Section III-B1. On the other hand, that means that slice 1 does not have access to the pool of resources allocated to slice 2. In terms of security, this means that potential breaches or attacks on the resources in slice 1, remain confined within the slice, not allowing attacks to penetrate slice 2. By observing the allocation of radio resources along the secondary y-axis, we can see that the network can adaptively manage resources while maintaining the isolation necessary to prevent interference between slices. This ensures that disturbances or errors in one slice do not propagate to others, maintaining continuous operation, as discussed in Section III-B3. This confirms the effectiveness of our network configuration in ensuring isolation in all network domains. Any errors and disturbances in one slice do not propagate to the others, maintaining overall system performance and reliability.

VI. Conclusion

In this work, we explain and demonstrate the feasibility and importance of achieving isolation in real-world 5G SA networks with Network Slicing. Isolation is crucial for maintaining the performance, security, and dependability of slices. Through the implementation of a PoC, we validate the effectiveness of our isolation design with real-world experiments, considering

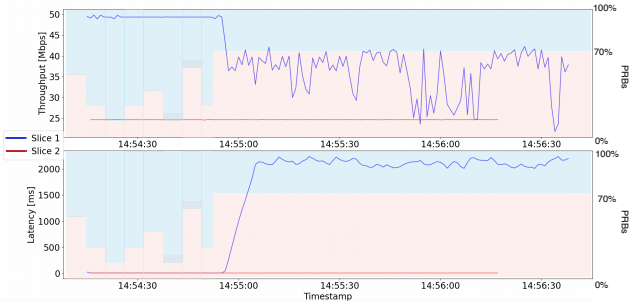

Fig. 4: Throughput and one-way latency behavior. In the background is shown the allocation of PRBs.

all network domains, i.e., RAN, TN, and 5GC. Our tests confirm that the deployment of isolation within our PoC ensures that enhanced eMBB applications, such as Augmented Reality, do not compromise URLLC applications, such as vehicular applications. This result is achieved through: i) dynamic allocation of radio resources, ii) rigorous configuration of network functions, and iii) dedicated communication paths in the TN. A key learning from our study is the critical role of network architecture design in achieving isolation. Designing each slice in a separate VM, and in separate data centers, ensures robust isolation. However, our network design presents limitations in terms of flexibility due to the static configuration of the TN and 5GC domains. The isolation in the TN domain is based on full physical isolation, without considering virtual resources. Additionally, we did not include specific security experiments, such as DDoS attacks or data breaches, to assess their potential propagation across slices. In future work, we aim to address these limitations by exploring more dynamic configurations for the TN and 5GC domains, and by integrating virtual resources to enhance flexibility and efficiency.

Acknowledgement

This work has been performed in the framework of the Flemish Government through FWO SBO project MOZAIK S003321N.

References

[1] ETSI, "LTE; Service requirements for the 5G system (3GPP TS 22.261 version 16.14.0 Release 16)," Technical Specification TS 122 261 V16.14.0, European Telecommunications Standards Institute, 2021. [Online] Available: https://www.etsi.org/deliver/etsi_ts/122200_122299/122261/16.14.00_60/ts_122261v161400p.pdf.

[2] C. De Alwis, P. Porambage, K. Dev, T. R. Gadekallu, and M. Liyanage, "A survey on network slicing security: Attacks, challenges, solutions and research directions," *IEEE Communications Surveys Tutorials*, vol. 26, no. 1, pp. 534–570, 2024.

[3] 5GPPP Architecture Working Group, "View on 5G Architecture," tech. rep., 5G Infrastructure Public Private Partnership (5G PPP), 2021. [Online] Available: https://5g-ppp.eu/wp-content/uploads/2020/02/5G-PPP-5G-Architecture-White-Paper_final.pdf.

[4] R. Mijumbi, J. Serrat, J.-L. Gorricho, N. Bouten, F. De Turck, and R. Boutaba, "Network function virtualization: State-of-the-art and research challenges," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 1, pp. 236–262, 2016. doi: http://dx.doi.org/10.1109/COMST.2015.2477041.

[5] O-RAN Alliance, "Use Cases and Deployment Scenarios," white paper, O-RAN Alliance, 2 2020. doi: https://mediastorage.o-ran.org/white-papers/O-RAN.WG1.Use-Cases-and-Deployment-Scenarios-White-Paper-2020-02.pdf .

[6] 3rd Generation Partnership Project (3GPP), "3GPP TS 28.530 V16.6.0 (2023-03) Technical Specification: Management and orchestration; Concepts, use cases and requirements (Release 16)," Technical Specification, 3rd Generation Partnership Project (3GPP), Mar. 2023.

[7] 3rd Generation Partnership Project (3GPP), "3GPP TS 28.541 V16.5.0 (2020-06) Technical Specification: Management and orchestration; 5G Network Resource Model (NRM); Stage 2 and Stage 3 (Release 16)," Technical Specification, 3rd Generation Partnership Project (3GPP), June 2020. [Online] Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3400.