IEEE Open Journal of the
Communications Society

Received 22 May 2024; accepted 17 June 2024. Date of publication 24 June 2024; date of current version 5 September 2024.

Digital Object Identifier 10.1109/OJCOMS.2024.3418495

A 5G NTN Emulation Platform for VNF Orchestration: Design, Development, and Evaluation

**FRANCISCO MURO[1], EDUARDO BAENA[1], TOMASO DE COLA[2] (Member, IEEE),
SERGIO FORTES[1] (Senior Member, IEEE), AND RAQUEL BARCO[1]**

[1]Telecommunication Research Institute, Escuela Tecnica Superior de Ingeniería de Telecomunicación, University of Málaga, 29010 Málaga, Spain

[2]German Aerospace Center, 82234 Weßling, Germany

CORRESPONDING AUTHOR: S. FORTES (e-mail: sfr@ic.uma.es)

This work was supported in part by the Ministerio de Asuntos Económicos y Transformación Digital y la Unión Europea - NextGenerationEU within the Framework "Recuperación, Transformación y Resiliencia y el Mecanismo de Recuperación y Resiliencia" under the Project MAORI.

ABSTRACT The integration of Non-Terrestrial Networks (NTN) with 5G represents a monumental leap in wireless network capabilities, significantly enhancing capacity, range, and reliability. This fusion fosters technological innovation and global connectivity. Crucial to this development is the softwarization of 5G networks through Virtual Network Functions (VNFs) within the OpenRAN paradigm, which disaggregates traditional radio access networks into Central and Distributed Units (CU and DU), accelerating deployment and facilitating upgrades. However, incorporating edge computing in Low Earth Orbit (LEO) satellites into this virtualized infrastructure presents arising management complexities, particularly in the interaction and resource sharing among VNFs, including satellite components. The absence of advanced emulators capable of simulating these environments represents a significant hurdle. As a response, this study introduces a groundbreaking emulation framework based on the Open Air Interface platform, tailored to reflect OpenRAN's disaggregated network architecture. The contribution of this work encompasses the design and implementation of a 5G-NTN platform, laying on Kubernetes for VNF lifecycle management automation, enabling a boost in multi-domain networks operational efficiency. The emulation of the NTN environment is performed at network level, abstracting the setup from the radio link implementation. The platform's capabilities in VNF deployment and resource optimization are validated by extensive performance metrics, with a full slice instantiation taking approximately 76.30 seconds, highlighting its contribution in emulating 5G-NTN orchestrated systems. The presented innovation lays the groundwork for future research in efficient resource management within complex multi-domain network environments, enabling the exploration of diverse VNF deployment scenarios and facilitating networks to adapt dynamically to advanced mobile communication demands.

INDEX TERMS 5G, non-terrestrial networks, network slicing, network orchestration, network management, VNF allocation optimization, O-RAN.

I. INTRODUCTION

IN RECENT years, cellular network development has steadily progressed towards more agile, efficient, and resilient deployments. This evolution is driven by the growing data traffic, the need for more flexible services, and the pursuit of increased efficiency in capital (CAPEX) and operational (OPEX) expenditures. As such, virtualization technologies have emerged as a pivotal attribute for the future of telecommunications networks. These technologies facilitate cost reductions by enabling formerly hardware-specific functions, such as routers and switches, to be executed on general-use hardware. This shift is part of a broader trend towards Network Function Virtualization (NFV), which simplifies the creation and quick configuration of new Virtual Machines (VMs) or containers to accommodate emerging services or to alleviate the strain on existing ones.

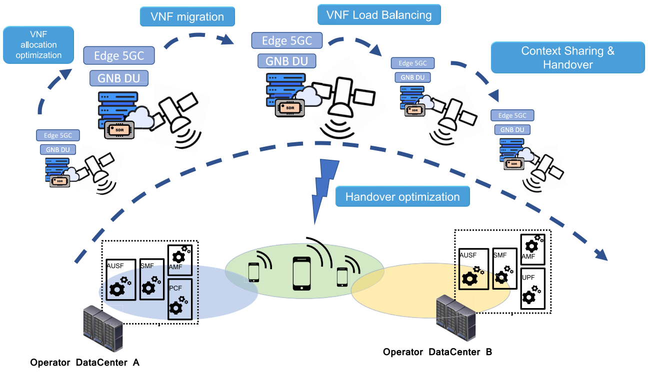

FIGURE 1. 5G NTN Framework.

From this standpoint, one of the applications of NFV technologies in new-generation mobile networks is Network Slicing (NS) [1], [2], [3]. In this setup, a virtualized network infrastructure can be divided into multiple logical networks or slices [4], [5], [6], [7], each tailored to different types of services. Each slice must manage virtual resources without interference, thus providing the service with the perception of exclusive infrastructure use. The primary challenge in enabling NS lies in achieving two objectives: guaranteeing independence between slices (functional isolation) to meet service requirements and ensuring the flexibility and adaptability of shared resources among different slices.

Looking beyond 5G (b5G/6G), integrating mobile and Non-Terrestrial-Networks (NTN) has emerged as a critical research area. This represents a key step towards achieving global coverage, enhanced capacity, resilience, and enabling new use cases, such as Mobile Edge Computing (MEC) based services and applications for massive Internet of Things (mIoT), especially from the Low Earth Orbit (LEO) and High Altitude Platforms (HAPS) standpoint.

Regarding this integration, two prevailing approaches frame the current discussion. The traditional model of transparent satellites, also known as bent pipe payload, utilizes NTN for backhauling 5G network links. Here, satellites serve as long-distance communication links, connecting remote or less accessible areas with the centralized network infrastructure. On the other hand, as a more contemporary trend, the regenerative payload approach corresponds to a profound paradigm shift in which satellites can process and transmit New Radio (NR) signals to and from user devices. This approach requires the satellite to have onboard processing capabilities robust enough to implement gNodeB (gNB) and other requisite Virtual Network Functions (VNF), to fulfill the architecture requirements. Furthermore, the dynamic nature of VNFs on moving aerial platforms introduces new challenges. These include optimizing VNF allocation in response to the constantly changing satellite positions, managing VNF migration efficiently, and enhancing handover mechanisms to maintain seamless connectivity, among other challenges (Figure 1).

Emerging as a key trend for the next generation of mobile networks, Open Radio Access Network (OpenRAN), as defined by the O-RAN Alliance, intends to revolutionize traditional wireless networking [8]. OpenRAN intends to disaggregate the Radio Access Network (RAN) into components with open and interoperable interfaces, facilitating the use of hardware and software from a diverse array of suppliers. This reshaping of the RAN industry towards open, intelligent, virtualized and fully interoperable solutions, position OpenRAN as a critical enabler for the integration of 5G and NTN networks and a significant driver of innovation and progress for the next generation of mobile networks.

The integration of 5G and Non-Terrestrial Networks (NTN) unveils a set of challenges particularly in emulating satellite communication delays which are crucial for effective integration. Past efforts, like the ESA project 5G-GOA [9] and the ongoing 5G-LEO project [10], aimed at addressing these issues but faced hurdles especially in supporting 5G Standalone (SA) configurations and realistic delay emulation. Additionally, the lack of support for the split of Distributed Unit (DU) and Centralized Unit (CU) for the gNB, and the absence of provision for 5G SA gNB Software Defined Radio (SDR) based implementations highlighted a significant gap in the existing state of the art.

In response to these challenges, this work introduces a novel container-based platform utilizing OpenAirInterface (OAI) software on top of Kubernetes, designed to emulate the 5G SA NTN characteristics accurately. This platform serves as a robust groundwork for the design, development, and testing of dynamic orchestration algorithms, marking a substantial stride towards the realistic emulation and effective management of 5G SA NTN environments, thus contributing significantly to the ongoing efforts in advancing 5G NTN integration. The contributions of this article are articulated as follows:

1) Novel Design and Implementation of a Cloud Native 5G NTN Platform: this work unveils a novel design of a virtualized 5G NTN platform, being supported by cutting edge tools for the implementation and orchestration, including a containerized disaggregation of the Open Air Interface 5G functions and using Kubernetes for the orchestration. This makes it the first contribution to provide a 5G SA NTN network based on a Cloud Native architecture. Notably, the split implementation of the Open Air Interface for both DU and CU, in addition to radio simulation, is emphasized, allowing for the optimization of the distribution of RAN VNFs.
2) Validation of the 5G NTN characteristics: the presented implementation is followed by a rigorous validation process to verify the operability and performance of the platform, specifically in the implemented satellite characteristics emulation.
3) Foundational Framework for Dynamic Virtual Resource Orchestration in 5G NTN: This contribution

establishes a baseline platform for future research in satellite-terrestrial network integration. It focuses on enabling the development of advanced algorithms for dynamic orchestration of virtual resources in 5G Non-Terrestrial Networks (NTN), setting the stage for significant advancements in this field.

The paper is structured into distinct sections that build upon each other to provide a comprehensive understanding of the proposed 5G NTN platform. Starting with Section II, it delves into the evolution of 5G RAN and Core, alongside the challenges and limitations inherent in integrating 5G and NTN. Following this, Section III section is articulated, laying the foundation on a reference architecture, as well as depicting the requirements and implementation details of the Cloud Native 5G NTN emulation Platform. In the subsequent Section IV, the paper assesses the features of the proposed framework, including the OAI Radio Simulation and the emulation of the NTN characteristics together with the limitations of the implemented environment. This leads to Section V where the key findings of the paper are summarized, providing insights into the contributions and implications of the presented 5G NTN emulation environment.

II. STATE OF THE ART

The 3rd Generation Partnership Project (3GPP) introduced the Fifth Generation Mobile Telephony, known as 5G or 5GS, with Release 15. The 5GS functionality was frozen in June 2018 and fully detailed by September 2019. The defined system goes beyond the air interface and embodies a full suite of protocols and network interfaces, embracing all aspects of mobile operations including call and session control, mobility management and service provisioning, thus facilitating seamless interoperability among various vendors and operators.

The evolution from LTE to the 5GS system architecture was split into two primary deployment options. The first option, referred to as the Non-StandAlone (NSA) architecture, leverages the existing LTE and EPC infrastructure Core Network (4G Radio and 4G Core respectively), coupling them with the 5G Radio Access Network (RAN) and its New Radio (NR) interface. This arrangement does not necessitate any network replacement and exclusively supports 4G services while simultaneously taking advantage of the 5G New Radio's enhanced capabilities, such as lower latency and higher bandwidth ranges. The NSA architecture embodies a fusion of past and future technologies. The second option is the StandAlone (SA) architecture, where the NR is directly connected to the 5G Core Network (CN). It is exclusively within this configuration that the complete set of 5G Phase 1 defined use cases could be performed.

To further enhance the user experience and deliver robust capabilities, 5G includes the use of specialized technologies, such as Network Function Virtualization and Network Slicing. While the basic functionality of the 5GS was defined in Release 15, later releases scope new enhancements such as Edge Computing, which is used to reduce response time, crucial for time-sensitive applications. Another example of these enhancements is the integration of the 5GS with Non-Terrestrial Networks, which are leveraged to provide universal network coverage, ensuring connectivity in even the most remote regions [11].

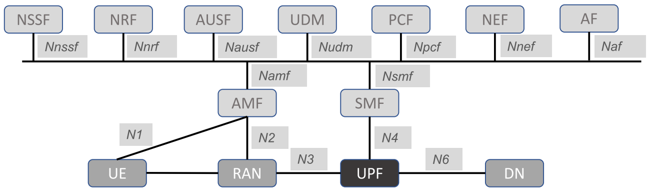

FIGURE 2. **5GS Basic Reference Architecture.**

Figure 2 shows the basic 5GS architecture with the main Network Functions (NFs) defined in Release 15. The 5G System architecture is ingeniously designed to boost data connectivity and service flexibility. Key to this architecture is the separation of User Plane (UP) functions from Control Plane (CP) functions, as well as the minimal dependence between the Access Network (AN) and the Core Network (CN), enabling independent scalability, flexibility, and evolution of network deployments. Moreover, the architecture supports "stateless" Network Functions, where compute and storage resources are distinctly separated. This innovative approach further enhances the network's adaptability and efficiency in managing heterogeneous service demands.

In the 5G architecture [12], several key functions are instrumental in ensuring efficient network operations (Figure 2). Foremost among these are the Access and Mobility Management Function (AMF), User Plane Function (UPF), and Session Management Function (SMF). The AMF is critical for NAS signaling termination, registration, and connection management, alongside access authentication and security context management, reflecting some of the Mobility Management Entity (MME) functionalities from the Evolved Packet Core (EPC) era. The SMF takes charge of session establishment, modification, release, IP address allocation, and traffic steering configuration, encapsulating some functions of the MME and Packet Gateway (PGW) from the EPC era. Meanwhile, the UPF handles packet routing and forwarding, Quality of Service (QoS) management, and acts as an anchor point for intra- and inter-RAT mobility, echoing the SGW and PGW functionalities from the EPC world.

The progress on Non-Terrestrial Networks (NTN) in Release 17 introduces fresh network designs to the 3GPP specifications, relying on high-altitude platforms (HAPS), Low Earth orbit (LEO) and geosynchronous orbit satellites (GEO). The goal of NTNs is to enhance TNs by providing global coverage, particularly in areas lacking TN infrastructure, providing seamless connectivity for society and industry verticals. However, the integration of 5G with NTNs introduces complex challenges in managing and orchestrating platforms that have not yet been resolved.

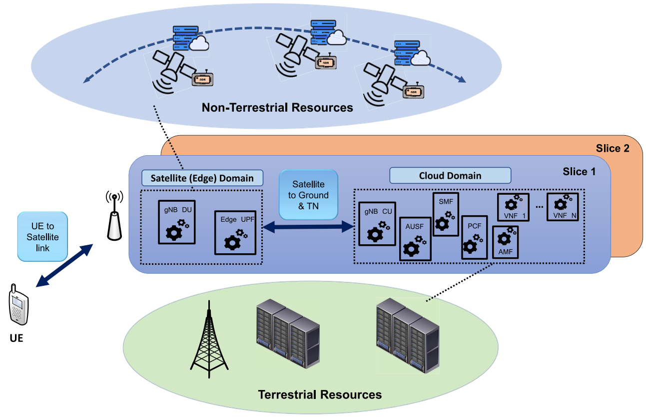

FIGURE 3. 5G NTN Domains Architecture.

Additionally, the potential of these platforms has not been fully exploited. Therefore, this research focuses on overcoming these obstacles to maximize the benefits of this technology.

Key to this architecture is the capability of Network Slicing. As presented in Figure 3, Network Slicing intents to allow the network defining different isolated logical architectures, or “slices”, on the same physical infrastructure, which proves to be fundamental for the divergent requirements demanded by the upcoming services. Consequently, each slice can be optimized in charge of satisfying a specific set of Quality of Service (QoS) deployed in a 5G network. Additionally, this also permits variability in security/isolation degrees, exposure, and self-management. This complexity has led to a surge of studies evaluating NS functionalities through simulators and testbeds composed of multiple virtual network elements, as evidenced by works such as [13] and [14].

A. 5G RAN EVOLUTION

The evolution of the 5G Radio Access Network (RAN) architecture represents a significant departure from the traditional approach of 4G LTE networks. In LTE, the base station, or eNodeB (eNB), was a single entity responsible for all the control and user plane functions. This monolithic design, while functional, lacked the flexibility to adapt to the diverse requirements of 5G applications.

With the advent of 5G, the base station, now termed gNodeB (gNB), has been split into two main components: the Centralized Unit (CU) and the Distributed Unit (DU). The CU handles the control plane functions, such as connection setup, mobility management, and QoS control. The DU, on the other hand, manages the user plane functions, such as data transmission and reception. This split architecture allows for more flexible and efficient network deployments, as DUs can be placed closer to the user equipment, reducing latency, while CUs can be centralized for efficient resource utilization. This split also plays an important role in the integration of 5G and NTN, primarily due to the inherent trade-off in hosting VNFs on satellites, what is mainly limited by the limited onboard capacity, presenting a challenge in finding the optimal allocation of 5G VNFs [15].

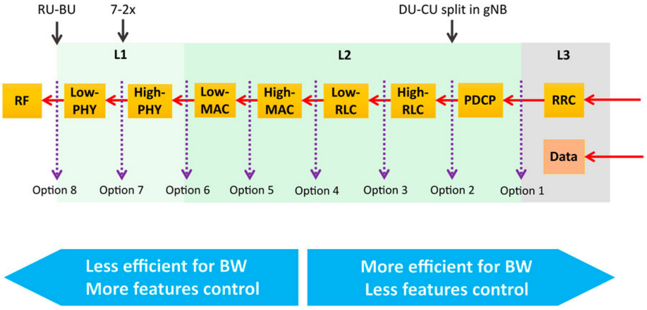

FIGURE 4. 5G RAN split options [16].

Partners of 3GPP have conducted research and subsequently defined various options for functional division between the CU and DU. These options range from a high-layer RAN split to a low-layer split, as depicted in Figure 4.

Building upon the foundations of NFV and SDN, one of the key trends for the next generation of mobile networks is OpenRAN. Defined by the O-RAN Alliance, OpenRAN is a type of radio access network architecture that disaggregates the traditional RAN in open and interoperable interfaces ([8], [17]) enabling the use of hardware and software components from multiple vendors. OpenRAN is built on the principles of Commercial Off-The-Shelf (COTS) hardware and software-defined networking (SDN).

Despite being in its early stages, the OpenRAN concept has already sparked significant research and development, as it provides a clear path towards fully programmable, intelligent and multi-vendor RAN. The O-RAN architecture is mainly based on two pillars: 1) The disaggregation of RAN into VNFs and Physical Network Functions (PNF); 2) The existence of a RAN Intelligence Controller (RIC) in charge of running central RAN applications, responsible for controlling and optimizing RAN functions. The RIC entity enables the onboarding of third-party applications which may rely on data-driven and advanced ML/AI-based tools to improve resource management capabilities.

To facilitate this evolution, O-RAN network architecture has been designed following the principles of virtualization and cloud services. This has given rise to the concepts of Cloud-Native telecom Applications (CNAs) and Cloud-Native Network Functions (CNFs). These concepts refer to the decomposition of software into smaller pieces or microservices, resulting in an application or a function being structured as a collection of loosely coupled stateless and stateful backing services [18]. Adopting a Cloud-Native architecture in O-RAN components would unlock several benefits, including adaptability, efficiency and performance, in which multiple works have conducted their research for the exploitation of the O-RAN capabilities [19], [20], [21], [22].

In parallel to the O-RAN-defined architecture, OAI is an open source initiative that provides a 3GPP compliant reference implementation of key elements of 4G and 5G

Radio Access Network (RAN) and core network that run on general purpose computing platforms together with COTS SDR cards. OAI offers an implementation of RAN, UE, and 5G Core systems. This allows for the development, testing, and validation of orchestration for the 5G elements. These tools, which have been utilized in many contributions towards 5G and beyond technologies [23], provide the capability of simulating the RF channel or implementing the radio protocol stack on an SDR device, thereby enabling a real system implementation.

Other projects like GNU Radio, srsRAN, and OpenLTE offer similar LTE/5G system implementations [24]. However, OAI's ability to split the gNB into CU and DU, essential for a Cloud Native distributed scenario, positions it as a leading project for the 5G NTN setup in this research. In this sense, the present work further extends the OAI distributed RAN to a Kubernetes orchestrated environment enabling its automated lifecycle management.

B. 5G CORE EVOLUTION

At the same time, the advancement of the 5G Core Network (CN) is a crucial research area for the evolution towards b5G/6G networks. Recent projects like open5GS, free5GC, open5GCore, and Open Air Interface (OAI) have significantly contributed to the implementation of 3GPP-Compliant Non-Standalone (NSA) and Standalone (SA) 5G Core networks. These initiatives are essential in facilitating the transition to advanced network technologies and enhancing network management, orchestration, performance, and efficiency.

The development and utilization of virtualization techniques represent another major challenge within the cellular network landscape. In this sense, Container-based virtualization, also known as Operating System (OS) level virtualization, has become increasingly important in cellular networks for improved performance, efficiency, and adaptability. This technology allows applications and their dependencies to be packaged in containers, offering a more efficient use of infrastructure compared to traditional virtual machines. Containers share the same kernel without running a complete OS (contrary to the traditional approach of VMs), making them lighter and enhancing memory, CPU, storage performance, and portability [25]. This makes containerization a key paradigm for the upcoming generations of mobile networks, addressing the growing demands for flexibility, agility, cost efficiency, resiliency, and high availability.

Several open-source 5G network implementations, such as OAI-5G CN, open5GS, and Towards5GS, have been extended to virtualized scenarios, including both virtual machine-based and container-based deployments. These extensions aim to enable the complete 5G system (RAN and SA 5G Core) to run on top of an orchestrator, adapting to the evolving needs of 5G networks.

Kubernetes, an open-source platform, has emerged as the standard for orchestrating containers in virtualized network environments. Its ability to automate deployment, scaling, and management of containerized applications makes it an integral part of the Network Function Virtualization (NFV) and Network Slicing (NS) paradigm. Kubernetes simplifies complex management tasks and has underpinned several proofs of concept, demonstrating the efficacy of Cloud-Native scenarios in mobile networks [26], [27], [28]. Notably, research in [29] introduced a CNF-scaler in Kubernetes, enhancing the scalability of 5G network services by adjusting CNFs' capacity based on their workload, validated using NRF and AUSF functions in the 5G Core. However, earlier studies have not fully addressed the complexities associated with VNF allocation across distributed environments, where factors such as latency and capacity imply careful trade-offs. In the context of these distributed multi-domain environments, NTN emerge as a pivotal enabler of edge computing capabilities within 5G networks. The precise allocation of VNFs in such settings represents a significant challenge, what is addressed as the main focus of our implementation, particularly in optimizing VNF placement across a complex network architecture.

C. INTEGRATION OF 5G AND NON-TERRESTRIAL NETWORKS: ACTIVITIES AND CHALLENGES

The 3rd Generation Partnership Project (3GPP) has significantly contributed to the evolution of mobile telecommunications, especially in the context of Non-Terrestrial Networks (NTN) in 5G technology. Their comprehensive studies, particularly in Technical Reports [30] (Release 15) and [31] (Release 17), have been fundamental in understanding the use cases, challenges, and necessary adaptations for integrating NTNs with 5G systems. These reports focus on aspects like propagation delay, frequency allocation, and the technicalities of the New Radio (NR) protocol in the NTN environment. Release 17 especially addresses the complexities arising from the satellite's mobility and orbital height, such as the path loss in GEO satellites and the Doppler offset in LEO satellites. These insights have led to the preliminary set of specifications for integrating NTN with NR, NB-IoT, and LTE-M [32].

However, the journey towards effective integration of 5G with NTN encounters several roadblocks, particularly in accurately emulating satellite communication delays. Past initiatives, including the ESA's 5G-GOA and the ongoing 5G-LEO project, have attempted to tackle these issues but faced challenges in fully supporting 5G SA configurations and in realistically simulating the significant delays characteristic of satellite communications. In this context, these studies aimed to incorporate NTN delays by modifying the software in charge of simulating OAI's radio link between gNB and UE. However, this method encountered limitations due to its reliance on a CPU-bounded sample rate; the sent and received samples are not processed at a fixed rate but instead fluctuate based on CPU availability. Consequently, introducing a latency through delaying a certain number of samples via software does not directly correspond to real-time, leading to challenges in mapping these delays to actual clocks and resulting in synchronization issues between connection endpoints. In contrast, the current study introduces a novel approach by implementing delay at the network level, effectively abstracting the delay implementation from the radio link simulation, thus decoupling it from CPU performance constraints.

Moreover, these efforts have not adequately addressed the implementation complexities of an orchestrated environment, while enabling the allocation of VNFs is understood critical to ensure service continuity and guarantee QoS in a 5G NTN environment. These limitations highlight a gap in existing research and underscore the need for emulation tools that integrate NTN conditions emulation with a framework to dynamically update VNF distribution to satisfy the changing network conditions.

Addressing these challenges, this study introduces an innovative container-based platform utilizing OAI software. This platform is designed to closely emulate the characteristics of 5G SA NTN. Hosted on Kubernetes, it offers a robust framework for managing the lifecycle of VNFs across a distributed setup. This development is pivotal, marking a significant step towards realistic emulation and effective management of 5G SA NTN environments. It not only contributes to overcoming the current limitations in the field but also sets a solid foundation for future advancements in the integration of 5G and NTN.

III. 5G-NTN PLATFORM ARCHITECTURE

A. REQUIREMENTS OF THE 5G-NTN INFRASTRUCTURE

This research outlines multifaceted requirements for the 5G Non-Terrestrial Network (5G-NTN) infrastructure, addressing the unique aspects of Low Earth Orbit (LEO) satellite constellations and the complexities of 5G systems. The requirements encompass a foundation on a reference architecture, a virtualized, disaggregated, and distributed 5G setup, complete orchestration of VNFs' lifecycle, a scalable cloud environment supported by ground infrastructure, and the emulation of 5G-NTN characteristics.

- *Foundation on a Reference Architecture:* Leveraging LEO constellations for reduced propagation delay and enabling edge computing. This advantage comes with a trade-off: their coverage area is smaller than in the case of Medium Earth Orbit (MEO) and Geostationary Earth Orbit (GEO), necessitating a larger constellation for global coverage and service continuity. Despite this challenge, initiatives such as Starlink have successfully launched and operate a substantial number of satellites. Thus Starlink serves as the reference architecture for the present study, allowing a better understanding of the NTN domain.
- *The 5G Setup: Virtualized, Disaggregated, and Distributed:* Emphasizing disaggregation of VNFs to enhance system responsiveness and flexibility, with a focus on addressing the limited onboard capacity of satellites through ground-based infrastructures.
- *Complete Orchestration of VNFs' Lifecycle:* Utilizing virtualization technologies for isolating and allocating network functions, requiring an effective orchestrator like Kubernetes to manage VNFs across satellite and ground systems.
- *Scalable Cloud Environment in Ground Infrastructure:* Envisioning a cloud environment to augment onboard satellite resources, focusing on the deployment, management, and optimization of VNF allocation.
- *Emulation of 5G-NTN Characteristics:* Setting up an infrastructure to mimic 5G-NTN behavior in terms of delay, throughput, and packet loss, facilitating the allocation of functions and resources in the distributed segments of the 5G-NTN environment.

B. IMPLEMENTED PLATFORM DESCRIPTION

The entirety of this setup operates on a VM with an allocation of 70 cores (AMD EPYC 7713P 64-Core) and 172 GB RAM. This VM runs on top of a Kernel-based Virtual Machine (KVM) hypervisor. Installed on this virtual machine is Kubernetes, an open-source container orchestration platform, in charge of the management of the containerized applications, ensuring scalability, fault tolerance, and streamlined operations.

Several hypervisors are in contention when deliberating on the optimal foundation for a virtualized environment, such as *VMWare vSphere / ESXi*, *KVM*, *Red Hat Enterprise Virtualization*, and *Xen*. On the one hand, *VMWare vSphere / ESXi* and *Red Hat Enterprise Virtualization* offer a polished, enterprise-grade virtualization platform, with a potentially high cost. *Xen*, being an open-source type-1 hypervisor provides high performance but may require a complex setup. In contrast, *KVM* appears as an open-source offering integration into the Linux kernel, simplifying management and offering robust performance. Given the balance of flexibility, performance, and ease of management, *KVM* has emerged as the preferred solution for virtualization-based testing purposes, and so for the foundation of this work.

Regarding the selection of the Kubernetes orchestrator solution, when considering a lightweight Kubernetes deployment, four notable solutions were assessed: *Minikube*, *K3s*, *MicroK8s*, and *kubeadm*. *Minikube* offers exceptional ease-of-use, predominantly for testing purposes, but lacks in supporting production-grade deployments. *K3s*, presents a production-ready lightweight environment but requires a rather complicated multi-node setup. *kubeadm* provides a conventional way to initialize clusters, giving users a more manual and granular control. Meanwhile, Canonical's *MicroK8s* strikes a good balance between ease of installation and robust scalability, designed for both experimental and production setups.

To this end, *MicroK8s* emerged as the most favorable choice for this work. Its scalable nature ensures rapid deployment and optimal infrastructure management. Regarding network management, *Calico* implements the Kubernetes Container Network Interface (CNI) as a plugin

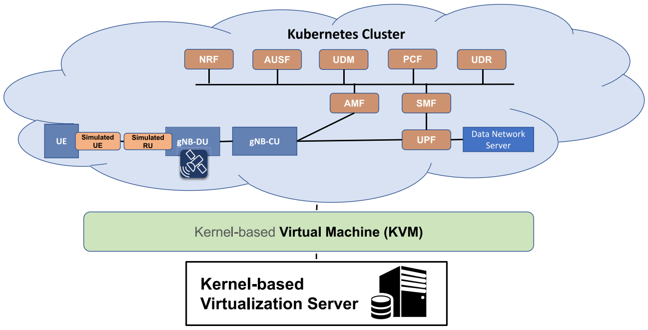

FIGURE 5. 5G-NTN Setup Architecture.

and provides agents for Kubernetes to provide networking for containers and pods. In order to accommodate multiple interfaces in containers, the *Multus* plugin was employed. Additionally, the *microk8s* deployment has been configured with a *host path* plugin to enable local storage tailored to the requirements of the 5GS storage.

While the use of a single virtual machine (VM) for the entire setup might appear unconventional for production-grade deployments given its inherent limitations in scalability and resilience, the context and goals of this project must be considered. With Kubernetes acting as the orchestrator, the practical difference between employing one VM or multiple VMs resides in the resilience of the setup to the failure of any of the Nodes (either in charge of performing the workload or the Control Plane of the cluster), what is negligible for the specific objectives of this study. The intentional selection of *microk8s* as our Kubernetes solution ensures an efficient and streamlined path for potential scaling to accommodate more VMs in future extensions of the cluster. By contrast, using a single VM simplifies the management, configuration, and potential migration of the testing platform, in comparison to the use of a multiple VMs Kubernetes cluster.

Regarding the 5GS, the proposed setup (Figure 5) leverages the Open Air Interface (OAI) implementation. One aspect of this implementation is the disaggregation of the 5G Standalone (SA) Core into various VNFs. To this concern, the selected UPF from OAI relies on the gtp5g kernel module to perform the GTP tasks inherent to the 5G Core, such as packet encapsulation and decapsulation for data forwarding. Furthermore, the Radio Access Network (RAN) is disaggregated into three main elements as per split 7-2: Radio Unit (RU), Distributed Unit (DU), and Centralized Unit (CU). Each of these VNFs operates in a distinct container, thereby enabling individual management, monitoring, and scaling of each VNF. This modular architecture provides flexibility in allocating each VNF based on targeted network performance, Quality of Service (QoS) requirements for specific services, and the capacity constraints inherent in a multi-domain environment. For the 5G-NTN setup, the plan is to consistently position the RU and DU onboard the satellite. The deployment location of the remaining VNFs will be strategically determined, considering whether they should be onboard or not, based on the specific demands and capabilities of the network.

The OAI RAN implementation offers two different modalities for operating the RU:

- *Software Defined Radio (SDR) hardware:* This modality permits the deployment of both the UE and the gNB using tangible hardware. This modality has not been finally validated by the OAI project for its use in disaggregated gNB, so its use is currently intended for the monolithic approach.
- *Radio simulation:* With this modality, the radio interface can be simulated using a TCP connection, allowing the modality of the disaggregated DU and CU for the gNB. The Radio Simulation is executed in both the UE and the gNB DU containers. This was the selected modality for supporting the present implementation.

Lastly, a Data Network (DN) server is incorporated into the architecture to facilitate end-to-end (E2E) testing within the infrastructure.

IV. 5G-NTN FRAMEWORK FEATURES ASSESSMENT

In addressing the emulation of 5G Non-Terrestrial Network (5G-NTN) characteristics, this study incorporates NetEm [33] to simulate the unique network conditions typical of the satellite constellations. This approach is informed by the expansive deployment of Starlink, which serves as our reference model for the emulation process. Our aim is to capture the essence of satellite network dynamics, particularly focusing on throughput, latency, and packet loss, without delving into excessive technical details.

The emulation strategy is designed to reflect realistic conditions, including managing packet loss rates up to approximately 4%, ensuring that our emulation closely aligns with the actual performance of satellite networks like Starlink. By focusing on replicating end-to-end (E2E) performance characteristics, the emulation targets capturing the typical latency and packet loss profiles observed in these networks.

A critical part of our emulation methodology involves correlating the induced delays and round-trip times (RTT) between the User Equipment (UE) and the gNB Distributed Unit (DU), based on empirical data. This correlation is essential for fine-tuning our emulation to reflect real-world satellite communication dynamics. The process is systematically documented and illustrated, ensuring clarity and accuracy in the emulation setup.

Alongside the traffic shaping performed by NetEm, it is crucial to classify the network traffic to apply varying characteristics to distinct connections. To achieve this, the Linux tool ‘Traffic Control’ (or ‘tc’) is employed. ‘tc’ facilitates traffic classification based on parameters like IP source, destination, and protocol, among others. This utility enables us to distinguish between the UE-to-satellite link characteristics and those intrinsic to the satellite-to-ground station connection.

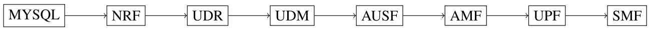

FIGURE 6. Deployment order of the 5G Core VNFs.

TABLE 1. Approximate duration for various 5g system deployment tasks.

Task	Approximate Duration (seconds)
UDR + MySQL DB Instantiation	42.22 (5.13 + 37.09)
NRF Instantiation	3.98
UDM Instantiation	3.97
AUSF Instantiation	3.97
AMF Instantiation	3.05
SPGWU Instantiation	4.02
SMF Instantiation	7.04
gNB CU Instantiation	4.03
gNB DU Instantiation	4.03
Slice Removal	12.00 (1.41 + 10.58)
Complete Slice Instantiation	76.30
Total Time for Slice Complete Redeployment	89.80

Pertaining to the assumption that the edge computing component of the network will be located onboard the satellite, it is bound to the resource constraints of the satellite capacity. Conversely, the cloud computing domain, which is accessed via the ground station and through the terrestrial Transport Network (TN), as presented in Figure 3, is assumed to have an "unlimited" capacity but has the trade-off of higher latency.

Regarding the 5GS, for a static deployment of the 5G Network, the functions instantiation occurs in a sequenced manner. The VNFs have a specific deployment order, visually represented in Figure 6.

After the launch of the 5G Core, the gNB is instantiated, which is followed by the activation of the UE. The prompt initialization of the UE instigates the Control Plane procedures for registration. To thoroughly evaluate these processes, it is imperative to pre-adjust the conditions of the pod interfaces after their instantiation to incorporate the NTN characteristics, what is typically addressed through the use of Kubernetes initContainers.

In the realm of 5G networks, understanding the temporal nuances of instantiation is of paramount significance. The agility of a network, particularly its adaptability to fluctuating network conditions and varying service demands, is heavily contingent on the efficiency of its instantiation tasks. As delineated in Table 1, we provide a meticulous evaluation of the durations involved in each deployment task associated with distinct VNFs. Such insights are not just an exposure of system behavior under deployment but also set the basis for later decision-making during the design and development of SON (Self-Organizing Network) algorithms for the orchestration of the 5G NTN network.

Within the context of Self-Organizing Networks (SON), the granularity of data analysis and the timeframe for the resolution of the optimization decisions are closely tied to the virtual network architecture. In our current setup, a full network slice takes approximately 76 seconds to deploy. In contrast, the minimal deployment time for a single VNF — specifically the Access and Mobility Management Function (AMF) — is around 3 seconds. Within the framework of O-RAN architecture, tasks demanding quicker responses are generally positioned closer to the Data Unit (DU), as exemplified by the near-RT RIC. Conversely, the SMO entity has a less stringent timing requirement and handles optimization tasks related to orchestration.

Given these considerations and based on the O-RAN reference architecture, we concluded that real-time data analysis is not mandatory for our current setup. We can afford to adopt a coarser level of granularity in our analysis. This facilitates understanding of the network environment, enabling informed decisions for re-deploying VNFs, creating slices, or re-optimizing current allocations.

A. OAI RADIO SIMULATION

One of the key features of this setup is the Radio Interface Simulator provided by OAI, which removes the necessity for SDR hardware. This simulation operates by establishing a TCP connection between the UE and the gNB DU, upon which all traffic is transmitted. However, it is crucial to be aware that the radio simulation possesses several limitations that must be considered for satellite link emulation. The physical simulators provided by OAI permit the execution of specific tests on individual transport channels or coding schemes. Nevertheless, these operations are not conducted in real-time, thereby preventing meaningful real-time throughput testing, among other aspects.

Upon a detailed inspection of the platform's operation, several characteristics can be identified:

- The simulator does not separate transmissions into actual Transmission Time Intervals (TTIs) as prescribed in 3GPP standards. Instead, it simulates these "virtual TTIs" (vTTIs) by periodically sending a fixed amount of data, thereby mimicking the Resource Blocks (RBs) of a 5G TTI. However, the transmission of this fixed amount of data, and consequently the duration of these vTTIs, is influenced by the underlying infrastructure's throughput (link between the UE and gNB DU container facilitated via TCP simulation), as shown by Equation 1.

$$\text{TTI}_{\text{Duration}} = \frac{\text{TTI}_{\text{Data Size}}}{\text{Throughput}_{\text{UE-gNB DU Link}}} \tag{1}$$

- User Plane (UP) traffic from the UE is allocated within a vTTI. If the scheduler is unable to allocate it within the current vTTI, it will be delayed until the entire transmission of the present one. Thus, the delay in UE's UP traffic is contingent upon the duration of the vTTIs and thus, conditioned by UE-gNB DU link Throughput.
- In the context of radio simulation using TCP, it becomes crucial to understand the throughput limitations imposed by the protocol. While the theoretical upper bound for TCP throughput is defined by the window size and Round Trip Time (RTT), real-world TCP behavior is often influenced by other factors like packet

losses. These losses, leading to retransmissions, affect the throughput. An estimation of TCP throughput in practical scenarios can be achieved through the Mathis model [34] as shown by Equation (2):

$$Thr = \frac{MSS}{RTT} \times \frac{C}{\sqrt{p}} \tag{2}$$

where Thr represents TCP throughput, MSS is the Maximum Segment Size, RTT indicates the round-trip time, C is a proportionality constant, and p denotes the packet loss rate. As identified by the TCP equation, the increment in RTT implied by the NTN characteristics, the achievable throughput under TCP is significantly reduced. This reduction becomes more pronounced when packet loss is factored in, highlighting the challenges in TCP-based data transmission in NTN environments.

Given these points, enforcing limitations on the delay, throughput, and packet loss for the link between UE and gNB DU leads to subsequent consequences. The relationships among these variables are expressed using derivative notation, where the sign of the derivative indicates the direction of the relationship.

- Throughput limitations will increase the duration of the vTTIs, resulting in a longer transmission time for the UE's UP data (Equation (3)).

$$\frac{d(\text{vTTI Duration})}{d(\text{TCP Thr.})} < 0, \quad \frac{d(\text{UP Latency})}{d(\text{vTTI Duration})} > 0 \tag{3}$$

- A delay limitation will increase the RTT between these entities, consequently reducing throughput, lengthening the duration of the vTTIs, and thus extending the transmission time for the UE's UP data (Equation (4)).

$$\frac{d(\text{TCP Thr.})}{d(\text{UE-DU RTT})} < 0, \quad \frac{d(\text{vTTI Duration})}{d(\text{TCP Thr.})} < 0, \quad \frac{d(\text{UP Latency})}{d(\text{vTTI Duration})} > 0 \tag{4}$$

- Packet loss on this link will directly affect the TCP connection, leading to retransmissions at the TCP layer, limiting the connection's effective throughput (or goodput) and amplifying the delay in UE's transmissions (Equation (5)).

$$\frac{d(\text{Eff. Thr.})}{d(\text{P.Losses})} < 0, \quad \frac{d(\text{vTTI Duration})}{d(\text{Eff. Thr.})} < 0, \quad \frac{d(\text{UP Latency})}{d(\text{vTTI Duration})} > 0 \tag{5}$$

While the radio simulation presents a distinct advantage by obviating the need for actual SDR hardware in realizing radio interfaces, it inherently introduces certain constraints. For the present work, it becomes a necessity to recognize and consider these limitations when emulating NTN characteristics to ensure realistic modeling. The subsection below outlines how the limitations were overcome and the methodology used to achieve the emulation of NTN characteristics in the 5G environment.

B. EMULATION OF THE NTN CHARACTERISTICS

As delineated in before, the Starlink satellite constellation, owing to its expansive deployment, has been identified as the reference architecture for this study. The primary objective is to emulate the intricate behavior exhibited by the particular architectural paradigm of the 5G NTN integration.

Considering the aforementioned limitations, the imposition of constraints, especially those influenced by NetEm conditions, requires meticulous calibration to ensure the reliability of the emulation. For the current configuration, direct throughput limitations are avoided. The OAI configuration tailored for this experiment is estimated to accommodate a maximum of approximately *100 Mbps* in downlink throughput and *10 Mbps* for uplink. However, the eventual throughput will be primarily influenced by the RTT conditions dictated by the satellite dynamics due to the TCP simulation.

In terms of packet loss, empirical analysis of the infrastructure demonstrate that the setup can tolerate up to 4% packet loss in satellite links, a parameter that warrants careful consideration in satellite link emulation.

One significant caveat to note is that the delay introduced by NetEm does not directly mirror the actually measured network conditions. Indeed, the actual delay encountered tends to significantly surpass the delay set via NetEm, stemming from the associated decrease in throughput and subsequent augmentation in vTTI duration.

Given this constraint, since a direct simulation of the delay from UE to satellite and subsequently from the satellite to ground is unattainable, the research pivots to capture the end-to-end (E2E) attributes of the Satellite constellation. The insights presented in [35] offer a comprehensive understanding of Starlink's user-perceived performance, evaluated across various reference benchmarks. Based on this foundational analysis, the E2E performance attributes of Starlink can be discerned as:

- A median latency hovering around *50 ms*, punctuated by intermittent peaks reaching approximately *70 ms*, with a recurring period spanning *15 ms* to *30 ms*, which can be attributed to satellite handovers.
- An average packet loss of 1.5%, manifested in sporadic bursts of variable duration.

Incorporating these constraints and being cognizant of the impediments introduced by the TCP connection for the present setup, the satellite constellation's delay emulation is orchestrated as follows:

1) A comprehensive series of tests was conducted to establish a correlation between the delay resulting from NetEm and the actual RTT reported from the User Equipment (UE) to the gNB DU. Illustrated in Figure 7 are the findings from these tests. To establish the RTT metrics for various delays caused by NetEm, the *Ping* tool with a default packet size of 64 bytes was used on the link between the UE and gNB DU. Highlighting

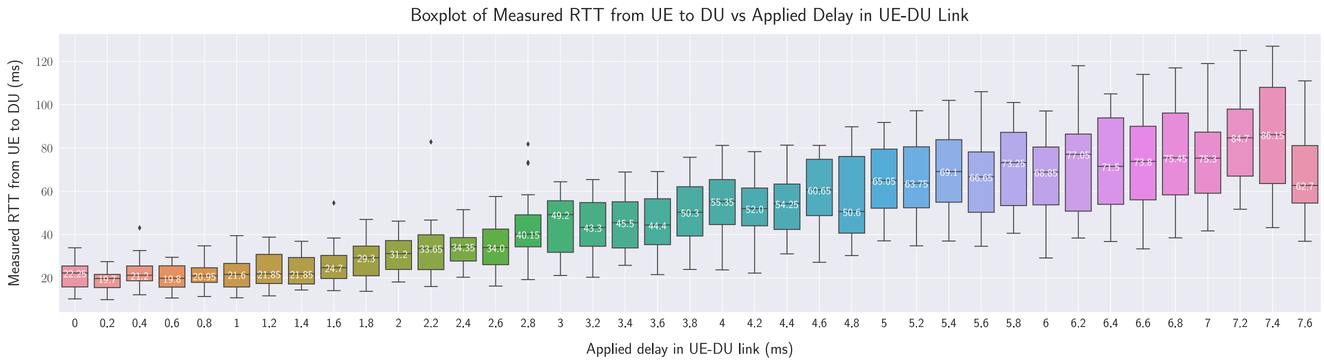

FIGURE 7. **Delay mapping from NetEm to tests measurements.**

the median RTT, each boxplot in the illustration depicts the measured RTTs distribution.

2) Drawing upon the empirical data from Starlink's performance metrics presented in [35], the NetEm-induced delay is judiciously adjusted to mimic a comparable E2E behavior.

Incorporating the aforementioned considerations, a meticulous mapping of the delay has been executed to emulate the end-to-end (E2E) latency intrinsic to the NTN constellation. This mapping facilitates a realistic representation of the latency dynamics, which is crucial for the realistic emulation of satellite-terrestrial network interactions. Alongside the delay mapping, packet loss has been methodically induced utilizing a Markov Chain model via NetEm to simulate loss bursts. This model, delineated into two distinct states - the Good State and the Bad State, aids in mirroring the probabilistic nature of packet loss in real network scenarios. Each state in this model represents different packet loss rates, and the transition between these states is orchestrated based on realistic network dynamics. The mathematical representation of this model is expressed as:

$$\text{Mean Packet Loss} = P_g \times L_g + P_b \times L_b \tag{6}$$

Herein, P_g and P_b denote the probabilities of the system being in the Good State and Bad State, respectively. Similarly, L_g and L_b represent the packet loss rates associated with each state. This model leads to an approximate mean packet loss of 1.5% as depicted in [35]. It is important to note that the throughput limitation has not been conclusively ascertained due to the constraints inherent to the TCP-based radio simulation.

The comprehensive approach, encompassing both the delay mapping and the packet loss modeling, enhances the emulation fidelity of the NTN environment, rendering it a more accurate reflection of real-world satellite communication dynamics. The verification tests conducted thereafter affirm the efficacy of these models. The obtained distribution of E2E delays from these tests is elucidated in Figure 8, where the interquartile range (IQR), spanning the 25th to 75th percentiles, is highlighted. Through this rigorous methodology, a significant stride towards a realistic emulation of 5G NTN environments has been achieved, laying a robust foundation for the design, development, and testing of dynamic orchestration algorithms essential for the effective operation of 5G NTN networks.

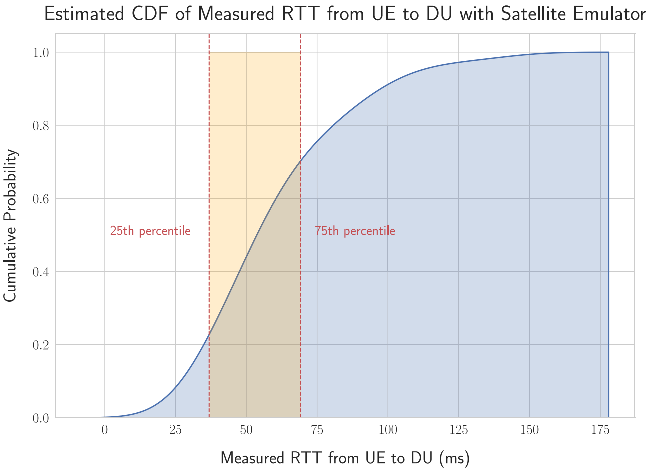

FIGURE 8. **Implemented 5G NTN measured delay.**

Furthermore, as previously delineated, the inherent constraints of satellite constellations dictate a limitation in onboard satellite capacity. Consequently, judicious selection and allocation of onboarded VNFs become crucial, ensuring that network performance is optimized in line with prevailing network conditions and service demands. Specifically, when deliberating upon resource allocation for various VNFs, understanding the unique requirements of each entity is imperative. The evaluation underscores the necessity of discerning the minimum requisites enabling VNFs to competently manage Control and User Plane tasks, especially during the UE registration procedure. Table 2 depicts the deployment outcomes for multiple VNFs under stringent resource constraints experimentally determined (0.25 cores of CPU and 250MiB of Memory). Notably, during these tests, resources for a single function were selectively constrained, enabling an assessment of network deployment and UE registration outcomes while independently limiting each of the VNFs. Of the functions assessed, the gNB DU

TABLE 2. Deployment result for isolated VNF resource constraints.

Limited VNF	Deployment Success
Network Repository Function (NRF)	Yes
Unified Data Repository (UDR)	Yes
Unified Data Mgmt. (UDM)	Yes
Authentication Server Function (AUSF)	Yes
Access and Mobility Mgmt. Function (AMF)	Yes
User Plane Function (UPF)	Yes
Session Mgmt. Function (SMF)	Yes
gNB Centralized Unit (gNB-CU)	Yes
gNB Distributed Unit (gNB-DU)	No

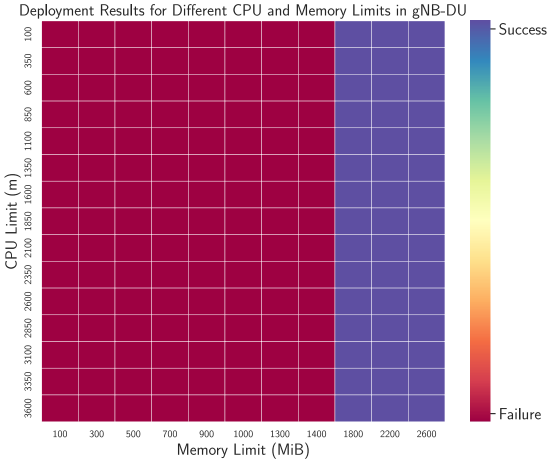

FIGURE 9. Resources constrains results for gNB DU assessment.

emerged as the most susceptible to resource limitations, underscoring its sensitivity.

Therefore, the vulnerability of the gNB DU to resource constraints should be a pivotal consideration in VNF allocation and network dimensioning strategies. Recognizing its significance, an in-depth evaluation was conducted to discern the sensitivity of the gNB DU against various resource constraints. Figure 9 elucidates the findings from tests with diverse resource limitations for the gNB DU VNF. The results starkly highlight Memory Allocation as the primary bottleneck; successful tests were only achieved with memory allocations of 1600MiB or more for the DU. Contrarily, CPU constraints did not present any discernible impediments for the outlined operations. This finding plays a pivotal role in guiding network capacity planning and VNF allocation strategies. The susceptibility of the DU to memory limitations demands its strategic allocation to computing nodes having expandable memory, safeguarding against potential operational disruptions due to escalating network demands. This approach ensures the network's adaptability and continuity by maintaining sufficient memory allocation margins. On the contrary, the CPU's role, while pivotal in determining network throughput and efficiency, does not directly threaten VNF stability, and sudden increasement in CPU demands might be handled through effective admission control.

The insights given into resource allocation are critical for informed network capacity planning and VNF allocation strategies, focusing on enhancing User Plane performance through optimal resource utilization. Moreover, this study manifests the capability of the implemented platform for VNF lifecycle management, allowing to orchestrate the distributed resources to understand the importance of an optimal network orchestration. This contribution lays the groundwork for a deeper understanding of the orchestration impact on User Plane performance, aiming to finely tune VNF allocations for optimal network operations.

V. CONCLUSION

In this research, the integration of 5G and Non-Terrestrial Networks was examined, focusing on the specific characteristics of the 5G Radio Access Network and Core Network. Leveraging the Starlink constellation as a model, a comprehensive, virtualized 5G NTN setup has been designed, utilizing Open Air Interface software. This setup distinguished itself with its disaggregated RAN architecture and employed Kubernetes for efficient Virtual Network Function lifecycle management, showcasing adaptability and dynamic capabilities.

Throughout the testing phase, several challenges have been encountered and addressed. One of the primary hurdles were the limitations posed by the TCP-based radio simulation. While this approach eliminated the need for physical Software-Defined Radio hardware, it introduced complexities in accurately emulating the radio interface. The inability of the simulation to mimic real-time Transmission Time Intervals accurately and its sensitivity to underlying infrastructure throughput required a nuanced approach to replicate actual NTN conditions authentically.

Another significant finding has been the sensitivity of the Distributed Unit to memory constraints. Experiments revealed that the operational stability of the Distributed Unit is critically dependent on sufficient memory allocation, a key insight for network capacity planning and Virtual Network Function allocation. This underscores the need for scalable memory resources in nodes hosting the Distributed Unit.

Future research directions include transitioning from an emulation-based framework to a real-world environment. This transition, involving the use of actual Software-Defined Radio hardware, as well as realistic external constraints in the cloud-domain, is expected to introduce new complexities and depth to the network's behavior. The application of real Software-Defined Radio hardware will offer a more accurate replication of NTN characteristics, enhancing understanding of 5G NTN integration.

Further, the potential application of Artificial Intelligence in optimizing Self-Organizing Network algorithms presents a promising area for exploration. Artificial Intelligence integration can offer intelligent, adaptive solutions for real-time network optimization, particularly in managing the challenges identified in the study. The development and implementation of Self-Organizing Network algorithms within an Open RAN architecture, considering Artificial Intelligence methodologies, will be a primary focus moving forward, marking a significant leap towards more efficient and autonomous network management in the 5G NTN domain.

ACKNOWLEDGMENT

This work has been conducted in the context of joint M.S. Thesis work between DLR and UMA.

REFERENCES

[1] K. Abbas, M. Afaq, T. Ahmed Khan, A. Rafiq, and W.-C. Song, “Slicing the core network and radio access network domains through intent-based networking for 5G networks,”‘*Electronics*, vol. 9, no. 10, p. 1710, 2020. [Online]. Available: https://www.mdpi.com/2079-9292/9/10/1710

[2] H. Basilier, J. Lemark, A. Centonza, and T. Asberg. “Applied network slicing scenarios in 5G.” 2021. [Online]. Available: https://www.ericsson.com/491069/assets/local/reports-papers/ericsson-technology-review/docs/2021/applied-network-slicing-scenarios-in-5g.pdf

[3] *Study on Key Quality Indicators (KQIs) for Service Experience*, 3GPP Standard 32.862, Mar. 2016. [Online]. Available: http://www.3gpp.org/ftp/Specs/archive/32_series/32.862/32862-e00.zip

[4] B. Blanco et al., “Technology pillars in the architecture of future 5G mobile networks: NFV, MEC and SDN,” *Comput. Stand. Interfaces*, vol. 54, pp. 216–228, Nov. 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0920548916302446

[5] R. Botez, J. Costa-Requena, I.-A. Ivanciu, V. Strautiu, and V. Dobrota, “SDN-based network slicing mechanism for a scalable 4G/5G core network: A Kubernetes approach,” *Sensors*, vol. 21, no. 11, p. 3773, 2021. [Online]. Available: https://www.mdpi.com/1424-8220/21/11/3773

[6] W. da Silva Coelho, A. Benhamiche, N. Perrot, and S. Secci, “Function splitting, isolation, and placement trade-offs in network slicing,” *IEEE Trans. Netw. Service Manag.*, vol. 19, no. 2, pp. 1920–1936, Jun. 2022.

[7] A. A. Khan, M. Abolhasan, W. Ni, J. Lipman, and A. Jamalipour, “An end-to-end (E2E) network slicing framework for 5G vehicular Ad-Hoc networks,” *IEEE Trans. Veh. Technol.*, vol. 70, no. 7, pp. 7103–7112, Jul. 2021.

[8] W.-C. Yang, “Interoperability testing tool for operations and maintenance interfaces of 5G open RAN base station,” in *Proc. 23rd Asia-Pacific Netw. Oper. Manage. Symp. (APNOMS)*, 2022, pp. 1–4.

[9] (Eur. Space Agency, Paris, France). *5G-GOA—5G Enabled Ground Segment Technologies and Over the Air Demonstrator via Geo Satellite*. Mar. 2021. [Online]. Available: https://artes.esa.int/projects/5g-goa

[10] (Eur. Space Agency, Paris, France). *5G-LEO—OpenAirInterface*^{TM} *Extension for 5G Satellite Links*. Dec. 2021. [Online]. Available: https://connectivity.esa.int/projects/5gleo

[11] A. Sultan. “5G system overview.” Accessed: Jul. 14. 2023. [Online]. Available: https://www.3gpp.org/technologies/5g-system-overview

[12] M. Dryjanski. “5G core network functions.” 2018. [Online]. Available: https://www.grandmetric.com/5g-core-network-functions/

[13] T. Irshad, “Design and implementation of a testbed for network slicing,” M.S. thesis, School Electr. Eng., Aalto Univ., Espoo, Finland, 2018. [Online]. Available: http://urn.fi/URN:NBN:fi:aalto-201802231619

[14] X. Foukas, M. K. Marina, and K. Kontovasilis, “Orion: RAN slicing for a flexible and cost-effective multi-service mobile network architecture,” in *Proc. 23rd Annu. Int. Conf. Mobile Comput. Netw.*, 2017, pp. 127–140. [Online]. Available: https://doi.org/10.1145/3117811.3117831

[15] H. Gupta, A. Antony Franklin, M. Kumar, and B. R. Tamma, “Traffic-aware dynamic functional split for 5G cloud radio access networks,” in *Proc. IEEE 8th Int. Conf. Netw. Softw. (NetSoft)*, 2022, pp. 297–301.

[16] “Study on new radio access technology: Radio access architecture and interfaces,” 3GPP, Sophia Antipolis, France, Rep. 38.801, 2016.

[17] *O-RAN Architecture Description 9.0, Release R003*, O-RAN Alliance, Alfter, Germany, Jun. 2023.

[18] *The Cloud-Native Transformation: A Guide to Cloud-Native Design and Operations Principles*, Ericsson, Stockholm, Sweden, 2022.

[19] S. Mondal and M. Ruffini, “Fairness guaranteed and auction-based x-haul and cloud resource allocation in multi-tenant O-RANs,” *IEEE Trans. Commun.*, vol. 71, no. 6, pp. 3452–3468, Jun. 2023.

[20] B. Agarwal, M. A. Togou, M. Ruffini, and G.-M. Muntean, “QoE-driven optimization in 5G O-RAN-enabled HetNets for enhanced video service quality,” *IEEE Commun. Mag.*, vol. 61, no. 1, pp. 56–62, Jun. 2023.

[21] H. Zhou, M. Erol-Kantarci, and V. Poor, “Knowledge transfer and reuse: A case study of AI-enabled resource management in RAN slicing,” *IEEE Wireless Commun.*, vol. 30, no. 5, pp. 160–169, Oct. 2023.

[22] F. Pase et al., “Distributed resource allocation for URLLC in IIoT scenarios: A multi-armed bandit approach,” in *Proc. IEEE Globecom Workshops (GC Wkshps)*, 2022, pp. 383–388.

[23] F. Pase et al., “Distributed resource allocation for URLLC in IIoT scenarios: A multi-armed bandit approach,” in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Rio de Janeiro, Brazil, 2022, pp. 383–388, doi: 10.1109/GCWkshps56602.2022.10008671.

[24] S. Kumar et al., “OpenAirInterface as a platform for 5G-NTN research and experimentation,” in *Proc. IEEE Future Netw. World Forum (FNWF)*, 2022, pp. 500–506.

[25] S. Ramanathan, K. Kondepu, M. Razo, M. Tacca, L. Valcarenghi, and A. Fumagalli, “Live migration of virtual machine and container based mobile core network components: A comprehensive study,” *IEEE Access*, vol. 9, pp. 105082–105100, 2021.

[26] C. Novaes, C. Nahum, I. Trindade, D. Cederholm, G. Patra, and A. Klautau, “Virtualized C-RAN orchestration with docker, Kubernetes and OpenAirInterface,” 2020, *arXiv:2001.08992*.

[27] O. Arouk and N. Nikaein, “5G cloud-native: Network management & automation,” in *Proc. IEEE/IFIP Netw. Oper. Manage. Symp.*, 2020, pp. 1–2.

[28] S. Barrachina-Muñoz, M. Payaró, and J. Mangues-Bafalluy, “Cloud-native 5G experimental platform with over-the-air transmissions and end-to-end monitoring,” in *Proc. 13th Int. Symp. Commun. Syst., Netw. Digit. Signal Process. (CSNDSP)*, 2022, pp. 692–697.

[29] A. Khichane, I. Fajjari, N. Aitsaadi, and M. Gueroui, “Cloud native 5G: An efficient orchestration of cloud native 5G system,” in *Proc. IEEE/IFIP Netw. Oper. Manage. Symp.*, 2022, pp. 1–9.

[30] “Study on new radio (NR) to support non-terrestrial technical networks (release 15),” 3GPP, Sophia Antipolis, France, Rep. 38.811, 2020.

[31] “Solutions for NR to support non-terrestrial networks (NTN) (release 16),” 3GPP, Sophia Antipolis, France, Rep. 38.821, 2023.

[32] I. Rahman et al., “5G evolution toward 5G advanced: An overview of 3GPP releases 17 and 18,” *Ericsson Technol. Rev.*, vol. 2021, no. 14, pp. 2–12, Oct. 2021, doi: 10.23919/ETR.2021.9904665.

[33] (Linux Found., San Francisco, CA, USA). *tc-netem(8)—Linux Man Page*. 2023. [Online]. Available: https://man7.org/linux/man-pages/man8/tc-netem.8.html

[34] M. Mathis, J. Semke, J. Mahdavi, and T. Ott, “The macroscopic behavior of the TCP congestion avoidance algorithm,” *SIGCOMM Comput. Commun. Rev.*, vol. 27, no. 3, pp. 67–82, Jul. 1997. [Online]. Available: https://doi.org/10.1145/263932.264023

[35] F. Michel, M. Trevisan, D. Giordano, and O. Bonaventure, “A first look at starlink performance,” in *Proc. 22nd ACM Internet Meas. Conf.*, 2022, pp. 130–136. [Online]. Available: https://doi.org/10.1145/3517745.3561416

FRANCISCO MURO received the degree in telematics engineering from the University of Málaga, Spain, in 2021, where he works as a Research Assistant, focusing on performance monitoring and optimization in virtualized wireless networks.

EDUARDO BAENA received the M.Sc. degree in telecommunication engineering from the Universidad de Granada, Spain, in 2010. He is currently pursuing the Ph.D. degree with the University of Málaga, Spain, where he works as a Lecturer and the Researcher with the Department of Communications Engineering since 2017. He is focus on unlicensed band applications and advanced network management schemes. He has held various industry positions in several companies, including operators, service providers, and manufacturers.

TOMASO DE COLA (Member, IEEE) received the Laurea degree (Hons.) in telecommunication engineering, the Qualification degree in professional engineer, and the Ph.D. degree in electronic and computer engineering, robotics and telecommunications, from the University of Genoa, Italy, in 2001, 2002, and 2010, respectively, where he has worked as a Scientist Researcher with the Italian Consortium of Telecommunications from 2002 to 2007. Since 2008, he has been with German Aerospace Center, where he is involved in different European Projects focusing on different aspects of DVB standards, CCSDS protocols, and testbed design.

SERGIO FORTES (Senior Member, IEEE) received the M.Sc. and Ph.D. degrees in telecommunication engineering from the University of Málaga, Spain. He began his career in the field of satellite communications, holding positions in European Space Agencies where he participated in various research and consultant activities on broadband and aeronautical satellite communications. In 2012, he joined the University of Málaga, where his research is focused on self-organizing networks for cellular communications.

RAQUEL BARCO received the M.Sc. and Ph.D. degrees in telecommunication engineering from the University of Málaga, Spain. In 2000, she joined the University of Málaga, where she is currently a Full Professor. She has worked in projects with major mobile communications operators and vendors and is an author of more than 100 high-impact papers.