Satellite-Terrestrial Integration: 5G Architectures for the Seamless Support of Mission Critical Services

Sara Nasirian[1], Giovanni Giambene[1], Sergio Barrachina-Muñoz[2], Josep Mangues-Bafalluy[2], Makhlouf Hadji[3], Miguel Ángel Vazquez[2], Lorenzo Santilli[4], Fotis Foukalas[5], Hamzeh Khalili[2]

[1]Department of Information Engineering and Mathematics, Università di Siena, Italy
[2]Centre Tecnològic de Telecomunicacions de Catalunya, Castelldefels, Spain
[3]Airbus Defense and Space, Paris, France
[4]Chief of the Innovation and Information Office, Telecom Italia S.p.A., Italy
[5]Dept. of Informatics & Telecommunications University of Thessaly, Lamia, Greece

Abstract— **This paper studies integration issues for the achievement of TN-NTN systems providing mission-critical (MCX) services in the EU 5G-GOVSATCOM project. The standardization of NTN as an integral part of 5G terrestrial systems is progressing, with important progress made in the last Releases 17 and 18. This important evolution, however, is not complete because the integration of the two networks is also needed at the protocol level. In the first part of this paper, we investigate MCX requirements. Then, two alternative architectures are considered with distinct core networks or with unified core networks, according to Release 20. The second part of this paper presents the issue of the seamless Vertical Handover between TN and NTN, proposing the adoption of a smart gateway server, a new block of the 5G system to manage a seamless switch from the two networks when some critical conditions are met. Using VPNs from UE to TN and TNT can facilitate the VHO procedure at the network layer, just avoiding a change of the IP address of the UE during the VHO that would cause a drop for the MCX end-to-end service. We believe this paper can provide useful insights into GOVSATCOM system design and the experimental phase of this project.**

Keywords—Fifth generation communication (5G), Non-terrestrial Networks, System Integration, Vertical Handover.

I. INTRODUCTION

Today, the integration of Terrestrial Networks (TNs) and Non-Terrestrial Networks (NTNs) is essential to address the increasing demand for ubiquitous and seamless connectivity and the drive to enhance network resilience and capacity [1]. This integration leads to a variety of network 3D architectures encompassing several aerial network components such as Geostationary Earth Orbit (GEO), Medium Earth Orbit (MEO), Low Earth Orbit (LEO) satellites, as well as High-Altitude Platform Stations (HAPS) and Unmanned Aerial Vehicles (UAVs), all beside the terrestrial infrastructure. 3GPP is defining the 5G standard, and recent releases, especially Releases 17 (year 2022) and 18 (year 2024), have addressed TN and NTN integration from architecture, physical layer, and protocol standpoints [2],[3].

In this scenario, the 5G-GOVSATCOM project of the Horizon framework supported by the European Union Agency for the Space Programme (EUSPA) aims to study architectural issues for integrating TN and NTN and demonstrate the feasibility of the seamless mobility between TN and NTN. The 5G-GOVSATCOM project belongs to the EU-GOVSATCOM cluster of projects whose aim is to support the development of a 5G NTN key infrastructure in Europe for crisis management and surveillance tasks [4]. The 5G-GOVSATCOM project aims to provide Mission Critical Services (MCX, also denoted as MCS) via a TN and NTN integrated system [5]. In the long term, the EU-GOVSATCOM system will be able to use the new IRIS^2 satellite constellation, the contract of which was signed with the EU in December 2024.

In the era of Public Safety and Disaster Relief (PPDR), MCX and related networks are crucial for ensuring safety, security, and seamless communications during critical situations. These specialized services and networks provide resilient and secure communication in devastated zones, enabling real-time coordination, situational awareness, and efficient resource allocation while safeguarding sensitive information. These MCX systems provide reliable connectivity in areas with limited coverage, where the conventional terrestrial infrastructure may be compromised. 3GPP Release 14 significantly expanded MCX by introducing Mission Critical Data (MC-Data, from a few kbps to a few Mbps) and Mission Critical Video (MC-Video, hundreds of kbps) capabilities, complementing the existing Mission Critical Push-to-Talk (MC-PTT, from 8 to 32 kbps) functionality.

In the context of 5G-GOVSATCOM, we aim to empower MCX further in an integrated network scenario, enabling seamless connectivity across diverse geographical areas and ensuring that MCX services remain operational under various circumstances, from urban centers to remote disaster zones. The 5G-GOVSATCOM project will be demonstrated in two real-world scenarios. The first scenario, set in the maritime area near Barcelona, involves a rescue vessel from Open ARMS (OARMS), where the project will showcase telemedicine services. By leveraging a vertical handover between terrestrial and satellite networks, the trial will prove how 5G-GOVSATCOM technology can ensure continuous video transmission during critical medical operations at sea, using both the private 5G network of Port de Barcelona and satellite capacity from the SpainSat-NG satellite by HISDESAT. The second scenario in Italy aims to demonstrate 5G connectivity for crisis management during disaster or emergency situations, as supported by the Italian Red Cross (CRI) and the Italian mobile operator TIM.

The mobility problem addressed by the 5G-GOVSATCOM project concerns the Vertical Handover

(VHO) scheme so that User Equipment (UE) can have seamless mobility when moving between TN and NTN coverage [6]. Mobility issues are addressed by 3GPP RAN2 WG, whose scope is to study radio interface architecture and protocols, the specification of the Radio Resource Control (RRC) protocol, and the Radio Resource Management (RRM) procedures [7]. Mobility management has to be carried out both in inactive mode (cell selection/reselection) and active mode (handover). Release 17 has introduced new triggering events for Conditional Handover (CHO) procedures. Time-based or location-based CHOs are particularly adapted to satellite systems, where the trajectories and transit times are highly predictable. Moreover, System Information Block 19 (SIB 19) has been introduced in 3GPP TS 38.331. This SIB plays a key role in enabling the UE to connect with NTNs since it provides the UE with vital configurations such as satellite Ephemeris data, common timing advance parameters, k_{offset} (a time offset for timing recovery), validity duration for uplink (UL) synchronization epoch time, cell reference location, and cell stop time. The provision of additional information via SIB messages is under investigation in 3GPP RAN2 WG [8].

VHO is a procedure that entails changing radio communication from TN to NTN (or vice versa), thus affecting PHY and the link layer, including RRC control protocols. There is not much work on VHO for TN-NTN integrated networks and the 5G-GOVSATCOM project aims to fill this gap. An effective VHO scheme has to account for link degradation and dynamic network conditions while ensuring QoS and minimizing interruption time during transitions. However, VHO is more than this because it may entail the network change with the consequent potential change of IP addresses for the UE. This would cause a drop in the end-to-end connectivity with the MCX server. To avoid these problems, the 5G-GOVSATCOM project adopts IP tunneling with pre-set IP tunnels (VPN) for both TN and NTN. Therefore, at the IP layer, the network change just implies the change of the VPN to connect with the MCX server. The VHO process is seamless if the change from TN to NTN or vice versa is unnoticeable to the user (i.e., entails a short interruption time).

II. Related Works

The common functional architecture for MCX services, as defined in [9], forms the foundation for mission-critical communications in 5G. This architecture encompasses crucial aspects such as group management, configuration, and security. It serves as the blueprint for implementing MC-PTT, MC-Data, and MC-Video services. The work in [9] delves into the 3GPP standardization process for MCX in 5G. It traces the evolution of MCX features across different 3GPP releases, highlighting how the standards have progressed to meet the demanding requirements of MCX communications. Reference [10] emphasizes enabling technologies like network slicing and edge computing, which are crucial for meeting the stringent requirements of MCX services. This reference also identifies challenges and potential future research directions in the field. It specifies various access types supported in 5G for MCX services, outlines session connectivity requirements, and defines Quality of Service (QoS) parameters specifically for MC-PTT and MC-Video services. This technical detail is essential for ensuring that MCX services can operate effectively within the 5G ecosystem. Reference [11] offers valuable industry insights into the actual deployment of MCX services (e.g., MC-Data, MV-Video, MC-PTT) in 5G networks.

The key issue, which is relevant to the 5G-GOVSATCOM project, concerns the definition of system architecture and the validation of a VHO scheme between terrestrial and satellite mobile systems. This topic is not well covered in the current literature, where papers on VHO referred to old technologies or mobility between wireless LANs and 4G/5G. The work in [12] focuses on VHO between terrestrial and satellite networks for a mobile UE onboard an Unmanned Aerial Vehicle (UAV). The VHO decision is based on a combination of received signal strength, signal-to-noise ratio, and elevation angle. This paper concentrates on the VHO criterion suitable for the envisaged scenario but does not consider architecture implications and relevant 5G protocol issues as requested by the 5G-GOVSATCOM project.

Recent studies have highlighted the role of Artificial Intelligence (AI) and edge computing in optimizing the VHO process [13]. AI-based approaches, such as Reinforcement Learning (RL), enable predictive and adaptive handover decisions by analyzing real-time network conditions and user mobility patterns. For instance, deep RL techniques have been employed to minimize handover latency and optimize QoS in hybrid 5G networks, demonstrating improved handover efficiency compared to traditional algorithms [14]. Moreover, edge computing plays a vital role in reducing the latency associated with handover decisions. By processing data closer to the user, edge nodes can provide rapid, context-aware handover decisions [15]. The integration of AI with edge computing further enables distributed decision-making for handovers, improving the adaptability and robustness of TN-NTN transitions.

Some 3GPP RAN2 WG documents have been achieved under Thales' guidance, focusing on enhancing mobility management and QoS in TN-NTN integrated systems [8], [16]. These studies include propagation delays and dynamic beam configurations, which are critical for ensuring seamless handover and service continuity between TN and NTN. Thales' work includes the development of advanced specifications for optimizing random access procedures, network selection mechanisms between TN and NTN, and feeder link handover [17].

The SAT5G project of the 5G-PPP framework has focused on integrating satellite communications within 5G terrestrial systems, with a primary emphasis on boosting Enhanced Mobile Broadband (eMBB) capabilities [18]. SAT5G showcased the integration of network virtualization and softwarization with a GEO satellite system. However, this project has not addressed the study of handovers from terrestrial to satellite networks. Instead, our paper concentrates on resource management and network layers, providing innovative considerations for supporting VHO procedures for 5G TN and NTN interoperable systems.

III. MCX Services in the Context of GOVSATCOM

Real-time video streaming, remote telemedicine, and coordinated emergency response efforts can be facilitated by combining TN with NTN to ensure uninterrupted connectivity and enable more effective disaster response and public safety efforts worldwide. Some features of MCX services for GOVSATCOM in terms of radio transmissions can be summarized as follows: (*i*) Up to 500 MHz in X-band; (*ii*) X-band spot beams with less than 0.8º beamwidth (GEO case); (*iii*) X-band is used for both feeder and user links. GOVSATCOM has been tested in the X band for a wide 220 MHz bandwidth with 16-QAM to achieve a high data rate of 800 Mbps. In such a communication system, the occupied frequency bandwidth can range from 125 MHz out of the full-width allocation of 375 MHz with different modulation types. It was observed that a high-speed downlink (DL) of 505 Mbps is achieved with only 125 MHz bandwidth, which is one-third of the full 375 MHz bandwidth [19]. If we consider the 5G NR NTN systems, we can observe that in the case of low-frequency bands such as the S-band with the DL/UL carrier frequencies both at 2 GHz, the system bandwidth is 30 MHz, and the Sub-Carrier Spacing (SCS) is 15 kHz. In the case of the Ka-band with DL/UL carrier frequencies at 20-30 GHz, the system bandwidth is 400 MHz, and SCS is 60 kHz [20].

We consider that real-time data exchanges between MCX clients and servers occur using TCP or UDP protocols. MC-Video and MC-Data over a 5G-GOVSATCOM telecom system have several requirements from 3GPP as follows [21],[22]: The MC-Video service should provide video compression ratios and quality at least comparable to H.264 video codec implementations. An MC-Video UE may support a video resolution from 320x240 at 10 frames per second (fps) up to and beyond 1280x720 at 30 fps. The MC-Video service should enable high-definition rendering of the video information with quality at least corresponding to video format 1080p (1920x1080 pixels HDTV resolution). The transmitting and/or receiving MC-Video UEs are moving at speeds from 0 km/h to 160 km/h. The MC-Video service shall ensure that the end-to-end delay from transmitting MC-Video UE to receiving MC-Video UE or console for urgent real-time video transmissions shall be no more than 1 s. Given the MCX requirements above, the following possible solutions are suggested for the 5G-GOVSATCOM system [23],[24]: (*i*) To use small slot transmission duration; (*ii*) To shorten the time of OFDM symbols by increasing SCS; (*iii*) To enhance, if possible, the power control command's scheduling and mapping timing to guarantee a quicker reaction time. We conclude that we can use SCS of 30 kHz with 100 MHz bandwidth or SCS of 60 kHz with 200 MHz bandwidth to run MCX services in the 5G-GOVSATCOM system.

IV. 5G-GOVSATCOM Architectural Options For MCX Support

According to the software-defined radio paradigm, the 5G building blocks are divided between the user plane and the control plane (see Figure 1). In what follows, we identify the 5G Core Network (CN) as the control plane of the 5G system. The main blocks considered in this project, besides the UE, are detailed below.

- **User Plane Function (UPF)**: The role of the UPF (user plane) is to provide packet-based routing/forwarding, header manipulation, QoS differentiation for the traffic flows, billing/caching, and policy support. The UPF can also be onboard the satellites to allow local data switching onboard the satellite; this can be particularly useful for reducing the latency experienced by UE-satellite-UE communications for MCX services [25].
- **Access and Mobility Management Function (AMF)**: It belongs to the control plane. It handles the registration of the UE to the 5G network. It authenticates the UE and authorizes its access to the 5G services. In addition, it tracks the location of the UE and manages its mobility within the 5G network.
- **Policy Control Function (PCF)**: It belongs to the control plane. It supports the unified policy framework that governs the network behavior by providing policy rules for control plane functions (such as network slicing, roaming, and mobility management) and subscription information for the policy decisions.
- **Session Management Function (SMF)**: It is the CN element (control plane) responsible for managing the sessions between the UEs and the network. SMF interacts with UPF, AMF, and PCF to establish and manage sessions. SMF is responsible for supporting slicing, QoS control, charging, and enforcement of the policies received from PCF.

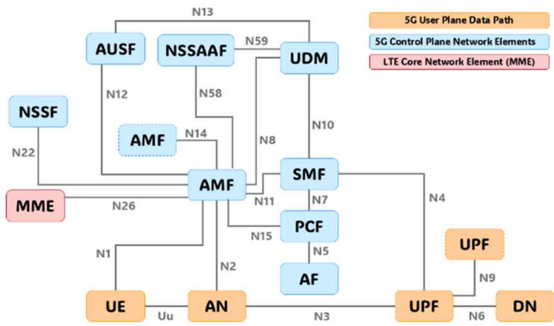

Fig. 1. 5G system building blocks: The cyan blocks belong to the control plane; the orange blocks are of the user plane.

We describe alternative architectures for the 5G-GOVSATCOM system below. We can have two distinct interoperable CNs with the possibility of exchange signaling between them or a unified CN where TN and NTN are like two distinct Radio Access Technologies (RATs) of the same system.

A. Baseline integrated scenario with TN and NTN having distinct core networks

This architecture is shown in Figure 2 with interfaces according to the 3GPP standard and new Gx interfaces. The UE can connect to both TN and NTN even simultaneously (multi-connectivity option). The 5G-GOVSATCOM project envisages an additional building block, the Smart Gateway (SGW) server, an intermediate router between the UPF and the external 5G-GOVSATCOM MCX server. The SGW server has the task of facilitating the joint use of TN and NTN with functions for QoS-based traffic steering (traffic routing enforcement). The SGW is divided into two interoperable blocks: the *SGW client* (SGW-C) on the UE side and the *SGW server* (SGW-S) on the network side, which is connected to the UPFs of both TN and NTN and manages access to MCX. The SGW-S is like a smart router deciding which traffic goes via TN and which traffic goes via NTN, depending on latency conditions. Gx interfaces are the new interfaces needed because of the SGW-S and SGW-C. Instead, the SGW-C is a smart router coordinating the VHO process on the UE side. A signaling protocol keeps SGW-S and SGW-C aligned.

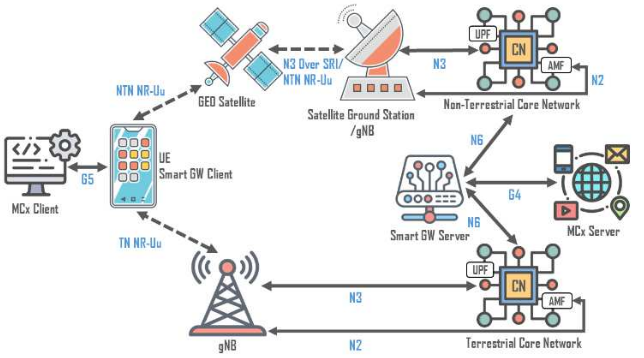

Fig. 2. Baseline integrated systems (3GPP Release 17 onwards) for mission-critical services: use of distinct core networks (TN/NTN). In this architecture, we assume that TN and NTN CNs can interoperate via the SGW-S that connects with the MCX service provider. The SGW-S can decide where to route the traffic (giving priority to MCX one) via TN and NTN, depending on its QoS and the network status (latency). The SGW-C is a router that manages the VHO process.

This architecture enables a client-server-based VHO scheme for mission-critical services by dynamically selecting routes between TN and NTN 5G networks. The approach adopted is based on dynamic routing mechanisms (inspired by the virtual routers described in [26]) located on both SGW-S and SGW-C. These blocks offer a generic framework that could consider not only network-related parameters but also application (APP)-related and 5G CN-related criteria to ensure optimal route selection. The example illustrated in Figure 3 establishes an architectural foundation to enable seamless VHO between TN and NTN using site-to-site VPNs, ensuring IP address privacy across a vertical server, vertical client, and their associated UEs.

The SGW-S, located on the MCX server (vertical provider) side, allows for the selection of routes between the TN and NTN networks. It ensures that the application server (MCX server) can securely communicate with the MCX client by abstracting the complexity of the underlying networks. A crucial aspect of the SGW-S is that it leverages site-to-site VPNs to create secure tunnels through NTN or TN, enabling seamless transmission of data among them (VHO) without revealing the client's IP address. This is particularly relevant since UEs typically operate as clients and are not designed to listen to incoming connections as servers do. By abstracting the UE IP behind the SGW-C, the architecture enhances security while preserving communication flexibility and the possibility of performing VHO as a rerouting among two alternative VPNs between SGW-C and SGW-S endpoints. In our architecture, we envisage using a static IP address for the MCX client, which is always the same regardless of the selected route (TN or NTN). The adoption of VPNs has the drawback of increased latency because of encapsulation, encryption, decryption, and decapsulation. This is an extra issue for the support of MCX services that can be addressed by using reserved slices. The system can be conceptualized with two distinct slices, one dedicated to MCX for TN and the other to MCX for NTN. Each slice encompasses its respective gNBs and UPFs, with VPNs integrated as a critical component of the data plane instantiation. Orchestration can provide an additional layer built on top of the proposed system to enable an intelligent algorithm to dynamically select the optimal path (either TN or NTN) based on real-time network conditions, service requirements, and resource availability within the slice.

SGW-C resides on the user side, enabling dynamic routing between the TN and NTN 5G networks and instantiates the VHO process based on measurements of Signal-to-Interference-and-Noise Ratio (SINR), Reference Signal Received Power (RSRP), Reference Signal Received Quality (RSRQ), Round Trip Time (RTT), and any other parameter we need to account for the selection of the network (e.g., packet loss); more details are provided in Section V.C. The SGW-C is the interface between the client device (running a mission-critical application, such as an MCX client) and the networks, both TN and NTN. The SGW-C can be regarded as a 2-UE device, where one UE interfaces with the 5G TN and the other with the 5G NTN. The client's IP address, which is assigned beyond the SGW-C, remains the same throughout the VHO process. This IP address is crucial to maintaining reliable connections during the handover between networks, ensuring that the client's connectivity with the server remains unaffected by the network change. The SGW-C utilizes the IP addresses assigned by the TN and NTN 5G networks as relay node addresses. These addresses are dynamically selected based on network availability and other VHO decision-making criteria, but they remain hidden from the application layer, thus simplifying the communication process.

Dynamic routing is achieved through the cooperation of the SGWs on both sides, allowing the MCX server and MCX client to communicate using the same private IP addresses while the actual routing decisions are abstracted by the SGW infrastructure (i.e., transparent to the MCX application on both sides). The key to this flexibility is the use of VPN tunnels between the SGW-C and SGW-S, featuring:

- **Site-to-Site VPNs**: These tunnels ensure secure, encrypted communication between the server and client without exposing the underlying IP addresses of the UEs.
- **Relay Node IPs**: IP addresses assigned to the UEs by the TN and NTN 5G networks are used as relay addresses between the client and server. These IP addresses are employed only as intermediate points for routing without being exposed to the application layers.

If the 5G CN control plane is accessible to third-party applications, the SGW-S can be an Application Function (AF) interfacing with the PCF to enforce QoS priorities in mission-critical scenarios requiring VHOs. For instance, in the event of congestion or poor coverage within the TN, the SGW-S can seamlessly transition the connectivity to an NTN to sustain service continuity aligned with the QoS parameters defined by the PCF. Furthermore, during emergency conditions (such as natural disasters causing severe TN congestion), the SGW-S can collaborate with the PCF to prioritize MCX services. The encryption overhead introduced by VPN tunnels can be kept at small values (compatible with MCX services) using optimized cryptographic algorithms and hardware acceleration in the SGW.

B. *Advanced integration scenario with TN and NTN sharing the same core network*

TN and NTN are expected to converge into a single unified CN, thus achieving the highest level of network integration planned from Release 20 onwards. This unified CN would seamlessly manage traffic from both terrestrial and satellite access networks.

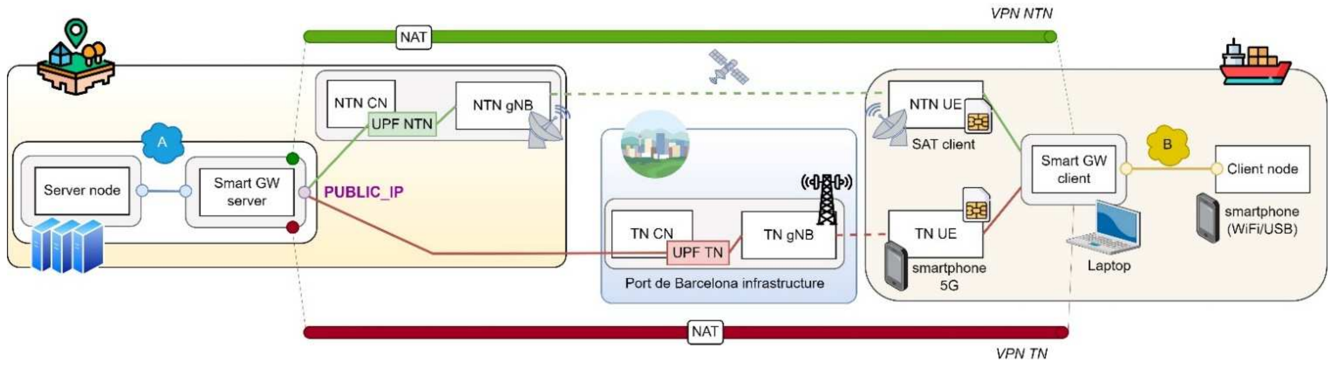

Fig. 3. VPNs are adopted to interconnect TN and NTN for better mobility support between these networks.

In this scenario, the SGWs can continue to play a key role, offering enhanced capabilities to support MCX services that go beyond current 3GPP-defined functionalities, especially in the fields of traffic engineering, policy enforcement, and QoS differentiation. In a fully unified CN, a UE experiences seamless mobility between TN and NTN, being two RATs of the same system; see Figure 4.

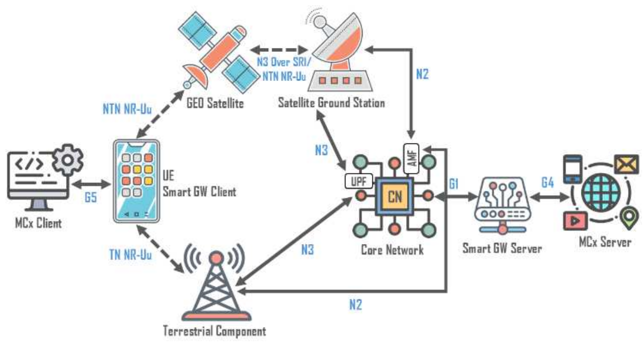

Fig. 4. System integration with a unified core network: TN and NTN share the same CN from 3GPP Release 20 onwards. SGW-S connects with the MCX service provider.

V. Preliminary Study on VHO for 5G-GOVSATCOM

A. Basic assumptions for VHO

The handover cases in 3GPP are categorized as "events" denoted with capital letters A or B and a number: from A1 to A5 and from B1 to B2 [27],[28]. Particularly relevant to our study are event A2 (it triggers when the serving cell becomes worse than a threshold; considering the measurements of SINR, RSRP, or RSRQ), event A3 (it triggers when a neighboring cell becomes better than the serving cell; it is still based on SINR, RSRP or RSRQ), and event B2 (it triggers when the serving cell level – SINR, RSRP or RSRQ – become worse than a certain threshold while a neighboring cell becomes better than a certain threshold). We can use the data in Table I for typical threshold values of SINR, RSRP, and RSRQ.

Protocols and related signaling behaviors are based on control timers that are triggered in case of failures. In particular, timer T304 is used for handover procedures. It starts with the reception by the UE of an RRC_Reconfiguration message, including reconfiguration with synchronization to the new cell. Timer T304 can be set in a range of values from 50 ms to 2000 ms. T304 ends upon the successful completion of the random access in the new cell. If this timer expires, the UE starts the reconnection to the source cell, but this may imply a Radio Link Failure (RLF). In the case of VHO from TN to a satellite network, depending on the satellite altitude, we should consider the T304 timer value to be 500, 1000, or even 2000 ms.

TABLE I. Main 5G PHY Measurements - Reference Values

		RSRP (dBm)	RSRQ (dB)	SINR (dB)
RF conditions	Excellent	≥ −80	≥ −10	≥ 20
	Good	−80 to −90	−10 to −15	13 to 20
	Mid cell	−90 to −100	−15 to −20	0 to 13
	Cell edge	≤ −100	< −20	≤ 0

We adopt the classical 3GPP approach of the *break-before-make* handover, which implies a disconnection from the old cell of coverage before connecting to the new one. This process entails a time that is called *handover interruption time*, during which there is no data transmission on either side. During this interruption time, data can be stored at the gNB and the UE and signaling messages are exchanged between TN and NTN and vice versa. Moreover, if the NTN is based on a LEO satellite constellation, we consider the *earth-moving cell* case: the satellite keeps a constant orientation for the onboard multibeam antenna that points down in the direction of the sub-satellite point. Spot-beam footprints move on the ground according to the satellite ground track speed, which is quite high, about 7.8 km/s.

B. Measurements to support VHO decision

In our VHO case, we assume that the SWG-C can take measurements of the signal levels of TN and NTN (we refer here to downlink SINR, RSRP, or RSRQ measurements). The SWG-C and the related UE can be configured to perform PHY downlink measurements (for TN and NTN) using the received Synchronization Signal Block (SSB) with a periodicity of 5, 10, 20, 40, and 160 ms. The interval for subsequent measurements T_{PHY} is set according to the SSB-based RRM Measurement Timing Configuration window (SMTC) specified by 3GPP [29]. A typical interval for these measurements is every 40 ms. The VHO decision is based on a filtered trace with L1 (short-time fading removal) and L3 filtering (reduction of fluctuations due to shadowing).

Moreover, SGW-C can send periodic ping requests via both TN and NTN (whenever they are available) to the SGW-S to keep updated the RTT measurements: RTT_{TN} and RTT_{NTN}. Let T_{ping} denote the periodicity of these ping requests. The RTT values affect the VHO process behavior. The RTT values depend on the network congestion, the time of the day, the source-destination pair, and the use of VPNs. For instance, Figure 5 provides exemplary histograms of the RTT values for both a terrestrial mobile network [30] and LEO Starlink [31], using, in both cases, close source-destination pairs in Italy. These results were achieved by repeated pings every second (T_{ping} = 1 s) for 30 minutes. As for the TN mobile access, OOKLA reports in Italy a median RTT value of 53 ms with a jitter of 9 ms [30]; our measurements in Figure 5 are consistent with these results. The TN RTT is typically lower than LEO satellite RTT, but there can be some cases where the TN latency is higher.

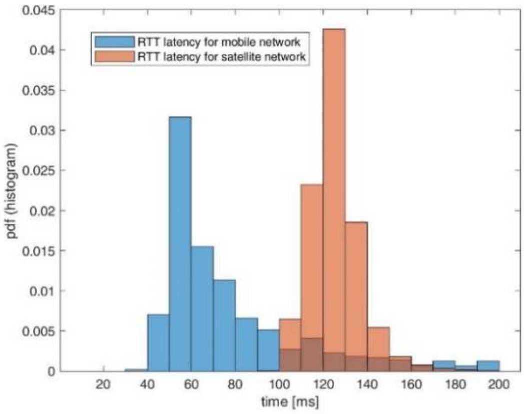

Fig. 5. TN vs. NTN (Starlink LEO satellite system) exemplary RTT distributions from repeated ping measurements for close source-destination pairs in Italy. These RTT values are exemplary of the data that can be collected by SGW-C via ping requests to the SGW-S via both TN and NTN.

Figure 6 shows the autocorrelation function of the RTT values for both TN and LEO satellite cases: we can identify the time interval on which the RTT behaviors are correlated. These results are obtained using the same set of results as in Figure 5. These autocorrelation results can help us set the frequency at which SGW-C sends ping requests to SGW-S via both TN and NTN so as not to send unnecessary ping requests. According to these results, the pings can be sent every T_{ping} of 20 s via TN and NTN.

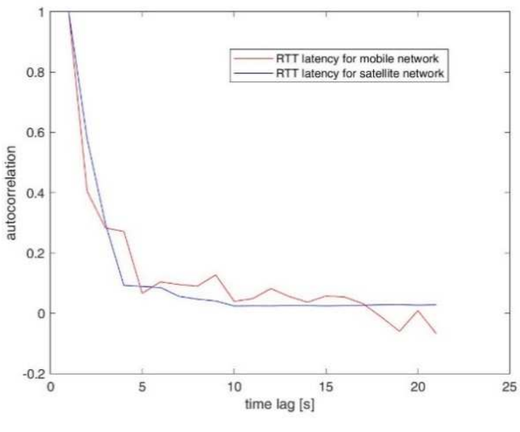

Fig. 6. Autocorrelation functions of TN and NTN RTT values to determine the frequency for the transmission of ping requests.

The following Figure 7 provides an example of the RTT traces, showing that the RTT of the TN can occasionally be above the RTT of the LEO satellite network. In the example below referring to normal operating conditions, the duration of this event is short enough not to motivate the need for a VHO. The situation can be different in the context of an emergency when the terrestrial network can be congested so that a VHO is needed to steer the traffic via satellite.

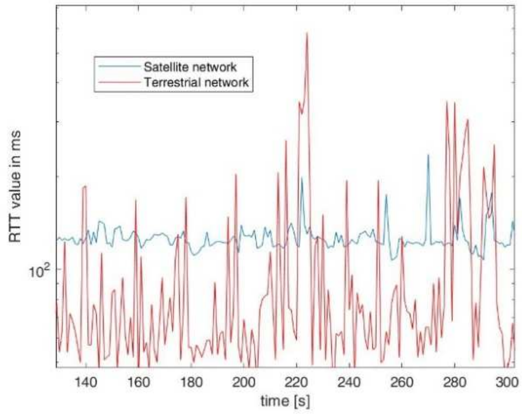

Fig. 7. Example of RTT traces for TN and NTN to show the possible occurrence of larger delays via TN with today's systems.

C. VHO decision criterion

In the following study, we refer to the VHO from TN to NTN (the opposite case has some similarities, but it will be the subject of further investigations). In our VHO case, since the UE is leaving the TN coverage, the UE realizes that it is moving away from the TN coverage when the SINR value goes below a certain $SINR_{thereshold}$ value (outage threshold) plus a HandOver Margin (HOM). This condition has to persist for a certain time TTT (Time-to-Trigger) to avoid that small signal fluctuations may cause VHOs back and forth between TN and NTN (the co-called *ping pong effect*). If these conditions are met, we consider that a VHO event type E1 is triggered at SGW-C. On the other hand, having a better RTT value on NTN rather than TN for a certain TTT can trigger a VHO event type E2 at SGW-C. More details will be provided later in this section.

We consider that the UE is equipped with GPS to know its updated position and speed information. The UE position can be used to apply a CHO scheme for VHO (this aspect is not considered in this paper). The UE speed can also be used to select adapted HOM and TTT values.

Current NTN systems based on LEO satellites are designed with very high minimum elevation angles (e.g., 30° for Startlink) to guarantee short interruption times. The SGW-C issues periodic ping requests via both TN and NTN VPNs to the SGW-S to continuously monitor the availability of these networks. Therefore, when the UE needs to perform a VHO, it already knows about the NTN availability.

We assume that to support MCX services, the UE has to be (re-)connected with TN as soon as it is available, referring to the common case with $RTT_{TN} < RTT_{NTN}$. This means that the UE should use the NTN coverage only when it leaves the TN coverage and should reconnect from NTN to TN as soon as it reenters under the TN coverage. This approach is needed because, typically, NTN resources are scarcer than TN ones and entail larger RTT values, which have a negative effect on MCX services.

3GPP adopts mobile-assisted HO schemes where the UE periodically reports its measurements to the gNB that takes handover decisions. In the VHO context, such an approach does not seem appropriate since the reporting scheme could delay the VHO decision. Therefore, we consider that a VHO decided on the UE side (actually, by the SGW-C) is a better approach to reduce the latency in deciding the handover based on both the physical layer measurements at the UE (e.g., SINR, position), and the network layer latency at the SGW-C (i.e., RTT). To perform a VHO, the AMF needs to accept VHO requests directly sent by SGW-C; the AMFs of the involved TN and NTN networks need to exchange signaling via the SGW-S.

In the TN-to-NTN VHO case, since the UE leaves the TN coverage, we consider that there is no other TN cell to compare the signal level. Therefore, it would be unfair to decide on the VHO by comparing TN and NTN SINR levels together, as in the A3 event, since they have different characteristics; adopting a VHO condition like the B2 event is better. Then, **event E1** is based on the following condition on TN downlink SINR to decide the VHO from TN to NTN:

$$\begin{gathered} SINR_{TN} < SINR_{TN,threshold} + HOM_1 \\ \& \\ \text{This condition must hold for more than } TTT_1 \\ \text{with the pings via NTN reporting connectivity.} \end{gathered} \tag{1}$$

TTT_1 values can be in the set {0, 40, 64, 80, …, 5120} ms. HOM_1 values can be from 0 to 10 dBs.

Since there can time intervals during which the TN has a larger RTT than the NTN system, we need to switch the communication via the NTN path to support the best latency for MCX services. Therefore, we consider VHO **event E2** as:

$$\begin{gathered} RTT_{TN} > RTT_{NTN} + HOM_2 \\ \& \\ \text{This condition must hold for more than } TTT_2 \\ \text{with the pings via NTN reporting connectivity.} \end{gathered} \tag{2}$$

The TTT_2 interval is used to bridge short intervals during which RTT_{TN} is larger than RTT_{NTN}; in this case, there is no need for a VHO, as commented in Figure 7. HOM_2 can be on the order of 50 – 100 ms, and TTT_2 can be on the order of 20 s, but a suitable measurement campaign is needed to properly set these values.

When the VHO from TN to NTN is performed because (1) or (2) are satisfied, we need to use new higher HOM_1 or HOM_2 values (hysteresis) for NTN-back-to-TN VHO to prevent that oscillations a little bit longer than TTT_1 or TTT_2 can cause a back VHO (ping-pong effect) soon followed by a new TN-to-NTN VHO. The settings of all HOM and TTT values depend on the propagation conditions (i.e., adopted frequency, environmental conditions, path loss law, shadowing standard deviation, shadowing spatial correlation), the UE speed, the signaling scheme adopted for VHO and the characteristics of latency (i.e., RTT) of TN and NTN. The selection of these parameters needs to be optimized to minimize the probability of VHO failure and the probability of ping-pong VHOs.

D. VHO signaling

The VHO process starts when the SGW-C detects predefined E1 or E2 events. These events, based on SINR measurements and RTT measurements, prompt the SGW-C to initiate the VHO procedure as described in Figure 8. Following the VHO decision, the SGW-C sends a VHO request to the SWG-S that, in turn, propagates this command to the TN core network's AMF. The AMF identifies that a handover to NTN is required and coordinates signaling with the NTN AMF through the SGW-S. After the VHO Interruption Time (IT), the UE performs synchronization with NTN and RACH access. The SGW-S acts as the intermediary, facilitating seamless communication between the TN and NTN AMFs. Once RRC reconfiguration is completed through the NTN access network, data and control packet exchanges resume via the NTN, thus completing the VHO process. A signaling protocol between SGW-C and SGW-S is used to keep both synchronized from the routing standpoint (both using TN or NTN for certain traffic flows).

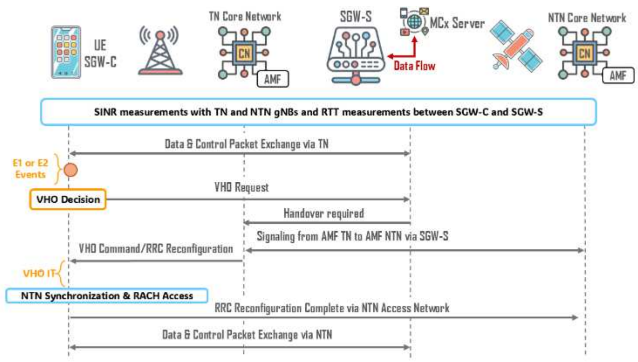

Fig. 8. Proposed signaling scheme for the TN-to-NTN VHO.

E. VHO interruption time

Key Performance Indicators (KPIs) for VHO are VHO Interruption Time (IT), VHO Failure (HF), Radio Link Failure (RLF), and Ping-Pong (PP) rate. In what follows, we focus on the VHO IT for this preliminary study. 3GPP TS 38.133 [29] specifies how to measure the interruption time IT as the time between the end of the last TTI containing the RRC reconfiguration command implying handover on the old PDSCH and the time the UE starts transmission of the new PRACH, excluding the RRC procedure delay. When the UE receives an RRC message implying handover, the UE shall be ready to start the transmission on the new uplink PRACH channel within $D_{handover}$ seconds from the end of the last TTI containing the RRC command. $D_{handover}$ equals the maximum RRC procedure delay defined in TS 38.331 (i.e., the time for receiving and decoding the RRC Reconfiguration at the UE) plus the IT time explained before. Therefore, we have:

$$D_{handover} = T_{RRC\_procedure\_delay} + IT \text{ ms.} \tag{3}$$

Then, based on TS 38.133 [29], the VHO IT can be characterized as follows:

$$IT = T_{search} + T_{IU} + T_{processing} + T_{\Delta} + T_{margin} \text{ ms,} \tag{4}$$

where T_{search} is the time required to search for the target cell. In our VHO case, the UE can know the satellite's ephemerides and the satellite cell (i.e., satellite beam) it can connect with in advance by using a SIB. Thus, T_{search} is not needed. T_{IU} is the interruption uncertainty in acquiring the first available PRACH occasion in the new cell. This depends on the PRACH periodicity of the NTN system between 10 ms and 160 ms. $T_{processing}$ is the UE processing time; it can be up to 20 ms. T_{margin} is the SSB burst post-processing time, up to 2 ms.

T_Δ is the time for fine time tracking and acquiring full-timing information of the target cell. T_Δ is equal to the SMTC periodicity of the target cell.

We can also consider a sort of "enlarged interruption time," where we include the time to complete the PRACH access successfully; this requires at least an RTT time via NTN for the TN-to-NTN VHO under consideration. This enlarged interruption time is coherent with the use of timer T304. Therefore, a VHO is successful (and the VHO procedure is seamless) if the following condition is met:

$$T_{RRC\_procedure\_delay} + IT + \mathrm{PRACH_{time}} < \mathrm{T304}. \quad (5)$$

During IT, the data to/from the UE accumulate in the buffers with the risk of causing buffer congestion, packet losses, and jitter effects. VHO IT can be significantly reduced if multi-connectivity option is supported, which allows the UE to be simultaneously synchronized with TN and NTN.

VI. CONCLUSIONS

The 5G-GOVSATCOM project has the task of defining the architecture for the seamless provision of MCX services via TN and NTN, also allowing VHO to support mobility between them. A new building block has been adopted in the 5G architecture, the smart gateway that allows traffic engineering, hinders UE IP address changes during VHO, and supports the VHO via periodic measurement of TN and NTN RTT values and SINR conditions with terrestrial and satellite gNBs. The study carried out in this paper has allowed the definition of the 5G-GOVSATCOM architecture and the identification of VHO criteria. Further work is needed to study the optimized HOM and TTT settings, and the modifications needed to the 3GPP standard. Another interesting option to evaluate is the possibility of performing VHO by choosing from multiple GOVSATCOM satellite operators and taking satellite system outage conditions, resource availability, and RTT conditions into account. The demo phase of the 5G-GOVSATCOM project will be fundamental for evaluating the KPIs of the VHO process.

ACKNOWLEDGMENT

Funded by the European Union under the Grant Agreement 101129593 — 5G-GOVSATCOM. This work is also part of the project SOFIA PID2023-147305OB-C32 funded by MICIU/AEI/10.13039/501100011033 and by FEDER/UE. Views and opinions expressed are those of the authors only and do not necessarily reflect those of the European Union or EU Agency for the Space Programme (EUSPA). Neither the European Union nor the granting authority can be held responsible for them.

REFERENCES

[1] S. Morosi, *et al.*, "Terrestrial/Non-terrestrial Integrated Networks for Beyond 5G Communications," *Space Data Management*, A. Cortesi, Ed., vol. 141, Springer, Singapore, 2024, pp. 89-101.

[2] G. Masini, P. Reininger, M. El Jaafari, A. Vesely, N. Chuberre, B. Baudry, and J.-M. Houssin, "5G Meets Satellite: Non-terrestrial Network Architecture and 3GPP," *Int. J. Sat. Comms & Net*, Vol. 41, No. 3, pp. 249-261, 2023.

[3] M. M. Saad, M. A. Tariq, M. T. R. Khan, and D. Kim, "Non-Terrestrial Networks: An Overview of 3GPP Release 17 & 18," *IEEE Internet of Things Magazine*, Vol. 7, No. 1, pp. 20-26, Jan. 2024.

[4] EUSPA GOVSATCOM program website with URL [Online]: https://www.euspa.europa.eu/eu-space-programme/secure-satcom/govsatcom

[5] 5G-GOVSATCOM website URL [Online]: https://5g-govsatcom.eu/

[6] A. Warrier, L. Aljaburi, H. Whitworth, S. Al-Rubaye, and A. Tsourdos, "Future 6G Communications Powering Vertical Handover in Non-Terrestrial Networks," *IEEE Access*, Vol. 12, pp. 33016-33034, 2024.

[7] 3GPP RAN2-WG webpage with URL [Available]: https://www.3gpp.org/3gpp-groups/radio-access-networks-ran/ran-wg2

[8] Thales, "Status Report for WI NR NTN (Non-Terrestrial Networks) Enhancements," RP-2301193GPP TSG RAN #99, March 20-23, 2023.

[9] 3GPP SA WG6, 3GPP TS 23 280, "Common Functional Architecture to Support Mission Critical Services: Stage 2," pp. 309, 2024.

[10] A. Santiago, *et al.*, "Analysis of Mission Critical Services Radio Access Networks Capacity Limitations over 5G", *IEEE Access Journal*, Vol. 12, pp. 6191-6203, 2024.

[11] 3GPP TS 23.289, "5G: Mission Critical Services Over 5G System: Stage 2," Release 19, 2025.

[12] A. Warrier, *et al.*, "Future 6G Communications Powering Vertical Handover in Non-Terrestrial Networks", *IEEE Access*, Vo. 12, pp. 33016 - 33034, February 2024.

[13] G. Giuliano and E. Innocenti, "Machine Learning Techniques for Non-Terrestrial Networks," *Electronics*, vol. 12, no. 3, p. 652, 2023.

[14] A. F. Ashour and M. M. Fouda, "AI-Based Approaches for Handover Optimization in 5G New Radio and 6G Wireless Networks," *2023 Int. Conference on Computer Science, Information Technology and Engineering (ICCoSITE)*, Jakarta, Indonesia, 2023, pp. 336-341.

[15] S. Plastras, *et al.*, "Non-Terrestrial Networks for Energy-Efficient Connectivity of Remote IoT Devices in the 6G Era: A Survey," *Sensors (Basel)*, vol. 24, no. 4, p. 1227, Feb. 2024.

[16] Thales Alenia Space, "NTN in 3GPP Roadmap and Release 17-18," Eager Project, May 2022. [Online]. Available: www.eagerproject.eu

[17] N. Chuberre, "Overview on NTN Enhancements in Release 18," Thales Alenia Space, 2022. [Online]. Available: www.eagerproject.eu

[18] K. Liolis, *et al.*, "Use Cases and Scenarios of 5G Integrated Satellite-Terrestrial Networks for Enhanced Mobile Broadband: The SaT5G approach," *Int. J. Sat. Comms & Net*, Vol. 37, No. 2, pp. 91–112, 2019.

[19] H. Saito, *et al.*, "High Bit-rate Communications in X Band for Small Earth Observation Satellites - Result of 505 Mbps Demonstration and Plan for 2 Gbps Link," *Engineering Physics*, 2016.

[20] J. Sedin, L. Feltrin, and X. Lin, "Throughput and Capacity Evaluation of 5G New Radio Non-Terrestrial Networks with LEO Satellites," *Proc. of 2020 IEEE Global Communications Conference*, Dec. 2020.

[21] 3GPP TS 22.281, "Mission Critical (MC) Video," Release 16.

[22] 3GPP TS 22.280, "Mission Critical (MC) Common Requirements," version 16.0.0 Release 16.

[23] G. Araniti, A. Iera, S. Pizzi, and F. Rinaldi, "Toward 6G Non-Terrestrial Networks," *IEEE Networks*, Jan/Feb. 2022.

[24] M. Hosseinian, *et al.*, "Review of 5G NTN Standards Development and Technical Challenges for Satellite Integration With the 5G Network," *IEEE Aerospace and Elect. Syst. Mag.*, Vol. 36, No. 8, August 2021.

[25] "FS_5GSAT_ARCH_Ph3 Architecture Assumption Updates – UE-satellite-UE communications," SA WG2 Meeting #S2-160 S2-2313366, 13-17 November 2023, Chicago, USA.

[26] K. Ito and N. Izuka, "Proposal of Client-Server Based Vertical Handover Scheme Using Virtual Routers for Edge Computing in Local 5G Networks and WLANs," *Proc. of IEEE 13th Annual Computing and Communication Workshop and Conference (CCWC)*, Las Vegas, NV, USA, 2023, pp. 0999-1004.

[27] 3GPP Handover Events Description at the URL [Available]: https://telecompedia.net/handover-events-in-lte/

[28] 3GPP TS 36.331, "Evolved Universal Terrestrial Radio Access (E-UTRA); Radio Resource Control (RRC); Protocol specification," Release 18, V18.3.1 (2024-09).

[29] 3GPP 38.133 "New Radio (NR); Requirements for Support of Radio Resource Management," V18.7.0 (2024-09).

[30] OOKLA Italy speed tests for both mobile and fixed access at the URL [Available]: https://www.speedtest.net/global-index/italy#mobile

[31] D. Laniewski, *et al.*, "Starlink on the Road: A First Look at Mobile Starlink Performance in Central Europe," arXiv:2403.13497v1 [cs.NI] 20 Mar 2024.