Towards Cloud-Native RAN for 6G NTN-Based Connectivity in Aviation

Babak Mafakheri*, Farzad Veisi†, Mohamed Hafidi*, Tomaso de Cola‡, Leonardo Goratti*.
*Safran Passenger Innovations GmbH, Germany
† CNAM and Cedric Lab, France
‡ German Aerospace Center (DLR)

***Abstract*—The aviation sector is emerging as a key vertical in the development of 6G Non-Terrestrial Networks (NTN), driven by increasing demand for high-capacity and low-latency in-flight connectivity (IFC). In this work, we explore the potential of disaggregated and cloud-native Radio Access Network (RAN) architectures to support seamless gate-to-gate connectivity for commercial aircraft. We propose and analyze four deployment scenarios that distribute the functions of disaggregated gNB, the central unit (CU), the distributed unit (DU), and the radio unit (RU) across aircraft, satellite platforms, and ground infrastructure. Emphasizing latency budget considerations, we examine trade-offs between performance, scalability, and complexity across multi-orbit satellite constellations. We further argue that adopting cloud-native RAN principles, along with intelligent orchestration and AI-based adaptation, can enable flexible and resilient architectures for future 6G NTN-based aviation services.**

***Index Terms*—6G, NTN, In-Flight Connectivity, Satcom**

I. Introduction

In recent years, airlines and aircraft Original Equipment Manufacturers (OEMs) have shown a growing interest in providing ancillary services to passengers as a new revenue stream, including in-flight connectivity, in-flight entertainment, Bring-Your-Own-Device (BYOD) services, on-demand content, and advertisements. In this landscape, the commercial aviation sector has become an appealing vertical for future 6G Non-Terrestrial Networks (NTN).

The integration of NTN into the 3rd Generation Partnership Project (3GPP) technical specifications roadmap has been a key milestone in the evolution from 5G to 6G. Beginning with 5G-Advanced [1], NTN was formally included in 3GPP Releases 17 and 18, primarily targeting transparent satellite scenarios [2]. This has laid the foundation for more advanced developments under Release 19 and beyond, where the focus has shifted towards regenerative satellites equipped with onboard processing capabilities [3]. These advancements enable the deployment of computing and network functions in space, significantly extending the scope of NTN use cases. Industry initiatives and public funding programs have strongly supported this transition. In Europe, both the European Space Agency (ESA) and the European Commission (EC) have driven major R&D efforts. Specifically, the EC through the Smart Networks and Services Joint Undertaking (SNS JU), has recognised the important role that NTN can play in the transition from 5G to 6G.

In parallel, the aviation sector has emerged as one of the key verticals for 6G-NTN research. The Seamless Air Alliance has identified four critical use cases for aviation connectivity, with passengers' internet access as the first use case, leveraging NTN technologies [4], underscoring the growing industry demand for seamless and high-performance airborne communications. Recent studies have presented research testbeds demonstrating the feasibility of linking airborne platforms to terrestrial 5G core networks via Low Earth Orbit (LEO) satellite backhaul [5], while integrating edge computing to support advanced airborne services. Additionally, frameworks have been proposed for dynamic resource monitoring and distribution across heterogeneous aerial platforms, enabling cooperative networking and enhanced service delivery [6]. The article "6G for Connected Sky" [7] further contextualizes the need for 6G NTN integration across a range of aviation scenarios, from passenger inflight entertainment to aircraft and crew operations. Airplane connectivity was formally introduced in 3GPP during Service and System Aspects (SA) Working Group (WG)1 meeting #108 and captured in TR 22.887 (Release 20) [8]. Moreover, 3GPP TR 22.870 [9] has introduced two new aviation-specific use cases: onboard communications for Urban Air Mobility (UAM) vehicles and high-data-rate services for commercial aircraft in 6G.

Complementing these efforts, several EC flagship projects, such as 5G-STARDUST [10], 6G-NTN [11], and ETHER [12], have paved the way for novel NTN network architectures integrated into 5G and 6G ecosystems. Moreover, the AI@EDGE project has validated a complete aviation-certified edge-cloud infrastructure and 5G connectivity for In-Flight Connectivity (IFC) through LEO satellites in a laboratory environment [13]. Most recently, the EC-funded NexaSphere project, launched in January 2025, aims to explore next-gen 3D network paradigms harmonizing 6G, AI, and unified TN/NTN frameworks over a three-year horizon with a focus on aviation connectivity as one of its Proof of Concepts (PoCs) [14]. Moreover, research projects such as 6GSKY [7] are exploring adaptive network architectures combining satellite and Direct Air-to-Ground Communication (DA2GC) to ensure robust aerial connectivity. These combined advances, from standardization to experimental validation, signal a growing strategic focus on 6G NTN solutions for aviation. Thus, challenges related to channel modeling, mobility, satellite-based backhaul, integration of disaggregated Radio Access Network (RAN), etc, in the aviation field need to be extensively analyzed. In this context, our work explores different scenarios of disaggregated RAN architectures tailored for aircraft connectivity within the future 6G systems. We examine various deployment configurations of RAN components, including gNB-Centralized Unit (CU), gNB-Distributed Unit (DU), and gNB-Radio Unit (RU) spanning both onboard aircraft platforms and in-orbit satellite nodes, evaluating their trade-offs in terms of latency, Block Error Rate (BLER), and integration complexity across a multi-orbit satellite environment.

The remainder of this article is structured as follows. Section II presents the overall scenario of in-flight connectivity within the 5G/6G NTN framework. Section III introduces four representative architectural models for implementing disaggregated, cloud-native RAN tailored to aviation use cases, followed by simulation results in Section IV. The advantages of adopting a cloud-native RAN approach are then discussed in Section V. Finally, Section VI provides technical insights into the findings and outlines potential directions for future research.

II. System Architecture

This section presents the system architecture designed to ensure seamless, flexible in-flight connectivity by integrating terrestrial and non-terrestrial networks using cloud-native and dRAN approaches.

A. Gate-to-Gate In-flight Multi-Connectivity Scenario

3GPP introduced in TR 22.887 the envisioned scenario where Internet access is provided onboard the aircraft, primarily for passenger entertainment, following aircraft communication service [8]. This connectivity service should be available gate-to-gate, meaning passengers can remain connected from the moment they board the aircraft until they disembark at their destination gate. To enable this, the IFC system must ensure consistent performance throughout all phases of the flight, from airplane taxi to landing. While until today aircraft connectivity has been primarily enabled by means of Geostationary Earth Orbit (GEO) satellites, Medium Earth Orbit (MEO) and new LEO constellations promise to revolutionize aircraft connectivity, open to multi-orbit strategies and in turn optimize the Quality of Experience (QoE) of the passengers.

To assess the use case in greater detail, this work assumes that the aircraft is capable of supporting both Terrestrial Networks (TN) and NTN connectivity, with access to multi-orbit satellite constellations (e.g., GEO, MEO, and LEO) provided by one or more 5G terrestrial or satellite operators. Two key assumptions are considered. First, part of the 5G gNB functionality is deployed onboard the aircraft, allowing passenger devices to connect directly to an onboard 5G cell. Second, passengers connect via the onboard Wi-Fi network, where the aircraft, equipped with a proper 5G modem and Very Small Aperture Terminal (VSAT), operates as a 5G UE connecting to the core network. Both assumptions aim to enable seamless integration with the broader 5G ecosystem and to ensure service continuity across different orbital segments, which are analyzed through simulation results in this study.

B. Cloud-native RAN

3GPP promotes dRAN to enable flexible, scalable, and cloud-native RAN architectures that support real-time applications, NTN scenarios, and multi-vendor interoperability. Disaggregated RAN can be achieved by deploying a distributed cloud-RAN, where virtualized DUs (vDUs), CUs (vCUs) and/or RUs are placed at different locations to replicate today's RAN architecture [15]. These functions run as Cloud-native Network Functions (CNFs) on Commercial Off-The-Shelf (COTS) servers, which are installed on cell sites and must meet typical baseband hardware requirements such as synchronization and communication reliability. Each server runs its own Kubernetes cluster, making its requirements distinct from centralized cloud RAN setups. To support this, vendors are developing specialized "RAN servers" for far-edge deployments. A key advantage of this model is centralized cloud-based management, enabling faster updates, automation, and flexibility using tools from major cloud vendors or open-source platforms. The F1 interface is established between the vDU and the vCU according to 3GPP TS 38.437 (NG-RAN; F1 Application Protocol) when they are not co-located. The F1 setup is always initiated by the vDU, which sends an F1 Setup Request message containing information such as the vDU ID, vDU name, served cell list, and other configuration details. The vCU responds with either a Setup Success or Setup Failure message. Furthermore, F1-based cell management operations, such as cell addition, reconfiguration and deletion, are performed following the structures defined in the same 3GPP specification.

For our aircraft scenario, we assume the gNB functional split follows the Midhaul Option 2, and Fronthaul 7-2 Evolution (7.2x) option, as supported by the O-RAN Alliance (depicted in Fig. 1). This split delegates more processing capabilities to the RU, while the layer 2 Medium Access Control (MAC) stays always within the vDU. As shown also in different scenarios proposed in this work (Sec. III) this split makes the cloudification of DU more flexible and supports inter-vendor operation between DU and RU.

Based on terrestrial cell splitting modes, lower-layer functional splits such as the split between the RU and DU, demand a strict latency as low as one-way propagation delay up to $250\mu s$ to ensure proper synchronization. However, such low-latency conditions are practically unachievable in NTNs due to the inherent propagation delays introduced by satellite links. Moreover, the latency budget between a UE and the DU in an NTN scenario is quite high compared to the terrestrial networks which impose synchronisation issues within layer 2. To address this, 3GPP TR 38.821 (Release 16) [16] proposes mechanisms to enable the use of lower-layer splits in NTN environments by enhancing tolerance to latency. Since this latency constraint arises primarily from the Hybrid Automatic Repeat Request (HARQ) process at the MAC layer of the

RAN, two key approaches are considered: either disabling the HARQ process completely in NR-NTN or optimizing HARQ for NR-NTN to better accommodate the larger delays introduced by satellite links.

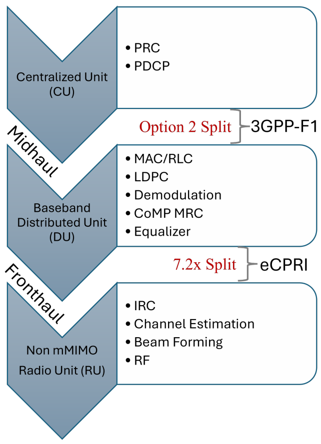


Fig. 1. 7.2x Lower Layer Split & Option 2 between DU/CU

III. Disaggregated RAN architecture for In-Flight Connectivity

The adoption of the cloud-native RAN paradigm enables flexible deployment, dynamic orchestration, and lifecycle management of RAN functionalities across distributed infrastructures. This flexibility is particularly advantageous in aviation environments, where connectivity requirements vary significantly based on flight phase, geographic location, and satellite availability. A cloud-native approach supports the runtime switching between different disaggregated RAN configurations, performance optimization, resource management, and enhancement of passengers' QoE.

In the following subsections, we describe four representative disaggregated RAN scenarios, reflecting different placements and distributions of RAN functions across the aircraft (as the Cell Site), satellites (as Far Edge-Cloud), and ground infrastructure (as Aggregated Edge-Cloud). Further, we rate the feasibility of each scenario by resorting to delay budget considerations as it will be later shown in Table I.

A. Transparent Payload with RU/vCU/vDU on the ground

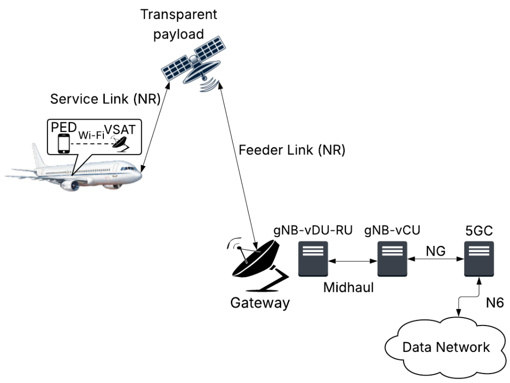
In this first scenario, shown in Fig. 2(a), all the disaggregated RAN components are deployed on the ground. With the airplane VSAT both service link and feeder link transport New Radio (NR). Fronthaul and midhaul are both through terrestrial network connecting edge-cloud data to the satellite gateway. In this case, satellites act purely as relay nodes without hosting any RAN functions. This scenario benefits from centralized processing, which simplifies network orchestration and makes resource sharing across aircraft efficient. Since the vDU is shared among multiple aircraft from a centralized cloud infrastructure, scaling and maintenance become also an easier task. However, this configuration is highly sensitive to the latency budget introduced by satellite propagation delays between the VSAT (acting as 5G-UE) and terrestrial vDU. For LEO satellites at approximately 900 km altitude, the one-way propagation delay from the aircraft to the ground (via satellite) is around 6 ms, which can even be higher at different elevation angles. When factoring in additional satellite-induced transmission and processing delays, the total one-way fronthaul latency can reach 15-20 ms [17]. Also, GEO satellites introduce one-way delays exceeding 280 ms [8]. The strict latency requirements imposed by layer two (MAC), in NTN considering 32 HARQ processes, even the (v)LEO-based constellation turns out to be difficult as the upper limit of latency budget can jumps to 32 ms round-trip-time (i.e., 16 ms one-way). Considering a higher HARQ process such as 64/128, the scenario can be considered borderline with sub-ideal latency even though a deeper throughput analysis will be needed. In GEO-based networks, the delay exceeds any existing threshold making synchronized operation infeasible unless the HARQ process is disabled as proposed in [16]. In summary, without disabling HARQ, this architecture can be considered borderline only for sub-ideal lower-latency satellite constellations such as (v)LEO with HARQ optimization.

B. Transparent payload with vCU on Ground and vDU and RU on the aircraft

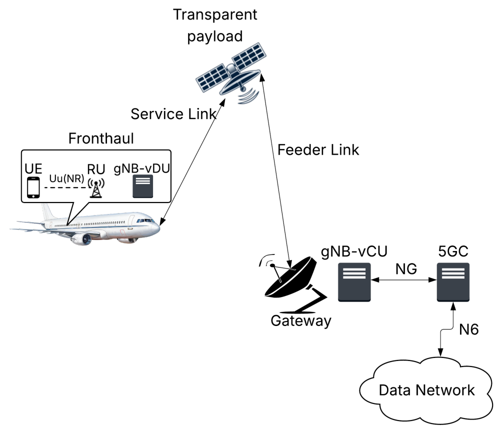
In the second scenario, both the RU and vDU are deployed on the aircraft, while the vCU remains on the ground, for example, co-located with the satellite gateway as shown in Fig. 2(b). An optical fiber link or equivalent high-speed connection can be used between the onboard RUs and the vDU (e.g., the server where the DU software is deployed). A vDU onboard can support multiple RUs, allowing multiple cells to be created within an aircraft environment with high user density. (e.g., large aircraft with multiple aisles). However, each aircraft requires its own dedicated vDU deployment. An important benefit of this configuration is the significant reduction in fronthaul latency requirement and also removing the latency budget problem of layer 2 (e.g., HARQ), since the fronthaul (between RU and vDU) is entirely on the aircraft and implemented over high-speed links (e.g., optical fiber). Also, the midhaul over the satellite is much more tolerant to latency, making this scenario far more resilient and stable under a 7-2x split, and ideal from the viewpoint of the fronthaul latency requirement. Moreover, deploying the vDU onboard gives each aircraft more autonomy and reduces its dependency on ground-side RAN orchestration for real-time decisions. However, this architecture does require deploying and managing the virtualized infrastructure onboard every aircraft, including the vDU, as well as hosting its own RU which increases the cost, power consumption, and maintenance complexity per aircraft. Also, the lack of shared vDU functionality nullifies potential efficiency gains that come from centralized processing for multiple aircraft.

C. Regenerative payload with vCU on ground and vDU/RU on LEO

In this third scenario, the aircraft act as a 5G-UE hosting its VSAT (similar to scenario A), the vCU is deployed on the aggregated edge-cloud on the ground, and the vDU-RU

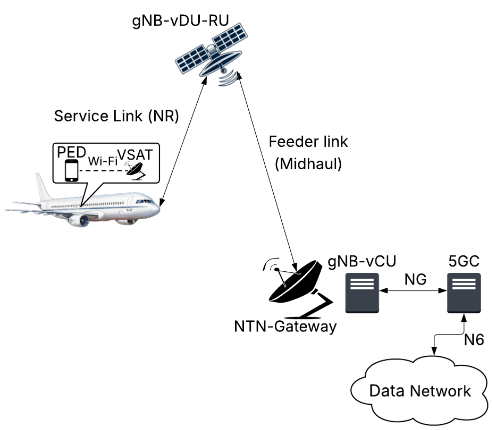


(a) Scenario A

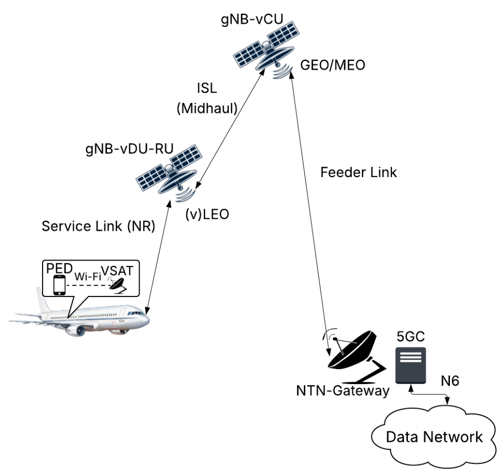


(b) Scenario B

Fig. 2. Transparent Payload



(a) Scenario C



(b) Scenario D

Fig. 3. Regenerative Payload

is hosted on the (v)LEO satellite itself, leveraging on the regenerative payload as illustrated in Fig. 3(a). This scenario balances performance and flexibility by distributing the gNB across the ground, satellite, and airborne nodes. The latency between UE (e.g., airplane) and DU is approximately equal to one-way propagation delay from airplane to v(LEO) satellites, ranging from 1 ms to 4 ms (e.g., vLEO & LEO placed in 300km & 1200km, respectively), making the latency budget requirement of layer 2 near-ideal and the split even more feasible compared to Scenario A. Some thorough analysis is required to find an optimized HARQ process considering a service link latency budget of 1-4 ms. Hosting the vDU on the satellite allows for shared processing among multiple aircraft, reducing the aircraft's onboard computational burden and improving scalability. However, the limited computing resources and power constraints on (v)LEO satellites impose challenges in supporting multiple vDUs/RUs on a scale [18]. The orchestration and lifecycle management of virtualized network functions in orbit is still in its infancy and requires robust fault management and redundancy. The reliance on continuous feeder link availability for the midhaul (satellite to ground) also introduces a point of failure and may constrain the reliability under adverse weather conditions or network congestion.

D. Regenerative payload with vCU on GEO/MEO, vDU/RU on LEO

Fig. 3(b) is the fourth and final scenario, in which the vDU-RU remain located on (v)LEO satellites, and the vCU is moved to higher-orbit satellites such as GEO or MEO. In this architecture, the F1 interface is established between LEO and GEO/MEO satellites, potentially leveraging Free Space Optics (FSO) for high-throughput, low-latency inter-satellite communication links. The feeder link connects the vCU back to the terrestrial network, acting as the main backhaul. From the viewpoint of the latency requirement of layer 2, the upper limit on the delay remains near-ideal as in Scenario C.

TABLE I

SUMMARY OF BENEFITS AND DRAWBACKS OF DISAGGREGATED RAN DEPLOYMENT SCENARIOS FOR IN-FLIGHT CONNECTIVITY

Scenario	Description	Benefits	Drawbacks
A. Transparent payload, all vCU/vDU/RU on ground	All RAN functions deployed on the ground; aircraft acts as 5G UE via VSAT.	• Centralized orchestration and maintenance • Efficient resource sharing across multiple aircraft • Lower onboard complexity and power needs	• High fronthaul latency due to satellite link • HARQ timing constraints make even LEO challenging • GEO infeasible without HARQ disabling
B. Transparent payload, vCU on ground; vDU+RU on aircraft	RU and vDU deployed onboard; vCU on ground near gateway; aircraft acts as gNB. It is not a full NTN scenario, using satellite as midhaul.	• Removes HARQ/fronthaul latency issues • Resilient under 7.2x split • Aircraft autonomy and stable performance	• Higher cost and power consumption onboard • Per-aircraft infrastructure deployment • Loss of centralization benefits
C. Regenerative payload, vCU on ground; vDU+RU on LEO	vDU and RU hosted on LEO satellite, vCU remains on ground; aircraft acts as 5G UE via VSAT.	• Low latency between aircraft and DU • Shared satellite-side DU supports multiple aircraft • Reduces onboard processing load	• Limited satellite compute/power capacity • Complex orchestration in orbit • Reliance on feeder link stability
D. Regenerative payload, vCU on GEO/MEO; vDU+RU on LEO	vDU and RU on LEO; vCU hosted on higher-orbit satellite; ; aircraft acts as 5G UE via VSAT.	• Enables multi-orbit architectures • Can leverage optical ISLs for efficient midhaul • High decentralization and flexibility	• Most complex in orchestration and synchronization • High cost of in-orbit virtualization • Handover and maintenance challenges

The main advantage of this setup lies in its extreme decentralization, which can benefit from multi-orbit constellations (e.g., upcoming IRIS^2 [19]). If supported by high-throughput optical inter-satellite links (e.g., FSO), the F1 interface between vCU and vDU can meet midhaul functional split requirements (e.g., option 2). However, this scenario is also the most complex and demanding in terms of orchestration, synchronization, and payload capability. Hosting both vCU and vDU functions on satellites requires a multi-layer satellite network with advanced virtualization, real-time coordination between orbits, and resilient inter-satellite links. The challenges in timing synchronization, compute availability, and fault tolerance are significant. Moreover, handover complexity increases when switching across satellite beams or constellations, and maintenance or updates of in-orbit functions can be very costly.

IV. SIMULATION RESULTS

To evaluate the proposed architectures, we performed system-level simulations using a Monte Carlo Python-based framework. Real flight trajectory data from a long-haul Singapore–London route were integrated, including aircraft latitude, longitude, altitude, heading, roll, and pitch (reported in [20]), sampled every 30 sec. The satellite constellation followed the IRIS^2 initiative [19], with 264 LEO satellites at 1200 km and 18 MEO satellites at 8800 km. At each step, distances to visible satellites were computed, and the serving satellite was chosen based on the highest Signal to Noise Ratio (SNR). For performance evaluation, scenario B served as baseline since no HARQ issues arise when UE and vDU are co-located onboard. Scenario D was excluded, as its results are identical to scenario C due to the same UE–vDU placement.

From the resulting slant range, the required HARQ processes were derived as follows: The round-trip time, considering the real positions of serving satellite and airplane and the elevation angle, is given by $RTT = 2 \cdot T_{\text{one-way}} + T_{\text{proc}}$ where the $T_{\text{one-way}} = \frac{d}{c}$, and $T_{proc} = 0.5ms$ is the HARQ processing time, d the slant range (km) (UE-DU) and c the speed of light (km/ms).

Following 3GPP TS38.821, the required number of HARQ process can be found as:

$$N_{\text{proc,req}} = \left\lceil \frac{RTT}{T_{\text{slot}}} \right\rceil \tag{1}$$

Where $T_{\text{slot}} = \frac{1}{2^\mu}\text{ms}$, and $\mu = \log_2\left(\frac{\text{SCS}}{15\ \text{kHz}}\right)$

Finally, the simulation considers the final number of HARQ process as: $N_{\text{proc}} = \min(N_{\text{proc,req}},\, N_{\text{cap}})$. In our setup, the sub-carrier spacing (SCS) is set to $60kHz$, suitable for Ka-band operation.

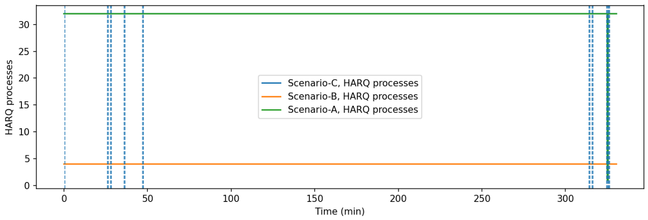
Fig. 4 illustrates the evolution of HARQ processes for scenarios A, B, and C with HARQ caps of 32, 64, and 128. The dashed vertical line marks the orbit switch from LEO to MEO. As shown in Fig. 4(a), with $N_{cap} = 32$, neither LEO nor MEO delays can be compensated, as the HARQ process remains in starvation (see Eq. (IV)). Increasing the cap to 64, as shown in Fig. 4(b) is sufficient for LEO (1200 km) in scenario C but remains inadequate for scenario A and for MEO (8800 km). With a further increase to 128, scenario A can also compensate for the delay, while the cap still appears insufficient for MEO.

To further investigate the impact of HARQ process limits on link reliability, we extended the simulator to analyze the BLER in scenarios A, B and C. Scenario B serves as a baseline reference without HARQ timing constraints, while Scenario C highlights the effect of round-trip latency under different HARQ caps. We calculated the per-round BLER as:

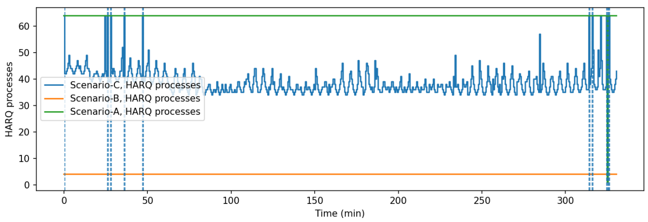
$$\text{BLER}_i(\text{SNR}) = \max\left(\frac{1}{1 + e^{\left(k\,(\text{SNR}+g_i-\text{SNR}_{50})\right)}}\ ,\ \text{BLER}_{\min} \right), \tag{2}$$

$i = 1, \ldots, R_{\text{eff}}$, SNR_{50} is MCS-specific SNR when BLER=0.5, and k the SNR-BLER slope [21]. g_i combining gain (dB) in round i, and $BLER_{min} = 10^{-4}$.

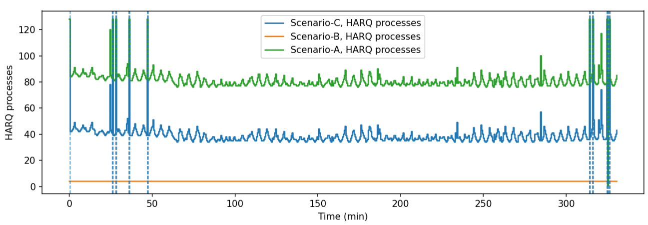
The logistic approximation in (2) is a common practice in link-to-system modeling. 3GPP TR 36.814 requires using BLER–SINR mappings from link-level studies, which are typically tabulated. For efficient system-level simulation, these curves are often fitted with smooth analytical functions. Following [22], we use a logistic form parameterized by the SNR at 50% BLER and a slope factor, which captures the characteristic S-shaped BLER–SINR behavior.



(a) HARQ Process CAP of 32



(b) HARQ Process CAP of 64



(c) HARQ Process CAP of 128

Fig. 4. HARQ starvation observed at different cap values

The BLER is then a product over all the rounds where a Transport Block (TB) fails only if all rounds fail.

$$\mathrm{BLER}_{\mathrm{HARQ}}(\mathrm{SNR}) \;=\; \prod_{i=1}^{R_{\mathrm{eff}}} \mathrm{BLER}_i(\mathrm{SNR}) \tag{3}$$

$$R_{\mathrm{eff}} \;=\; \begin{cases} R_{\mathrm{tim}}, & \text{Scenario B (local EU–DU)} \\ \max\Big(1,\; R_{\mathrm{tim}} \cdot U\Big), & \text{Other Scenarios (latency-limited)} \end{cases}$$

$$R_{\mathrm{tim}} \;=\; \max\left(1,\; \frac{T_{\mathrm{budget}}}{RTT}\right), \tag{4}$$

R_{tim} indicates how many HARQ retransmission opportunities can fit inside the allowed feedback window, given the propagation delay and processing time.

$$T_{\mathrm{budget}} = k_1.T_{\mathrm{slot}} + m,\; with\; k_1 = 24slots\; and\; m = 4ms$$

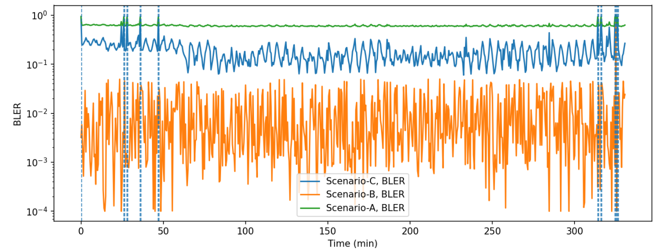
U is the HARQ pipeline utilization (1= no starvation)

$$U \;=\; \frac{N_{\mathrm{proc}}}{N_{\mathrm{proc,req}}} \;=\; \frac{\min(N_{\mathrm{proc,req}},\, N_{\mathrm{cap}})}{N_{\mathrm{proc,req}}} \in (0,1]$$

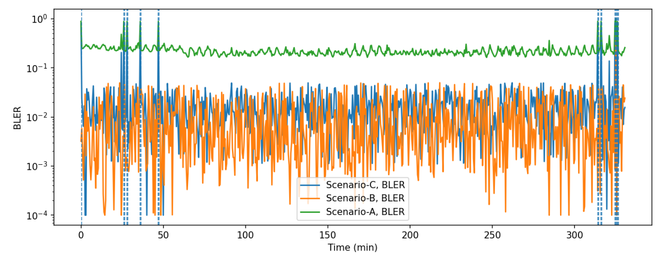
$$\mathrm{BLER}_{\mathrm{eff}} \;=\; 1 - \big(1 - \mathrm{BLER}_{\mathrm{HARQ}}\big)\,\big(1 - p_{\mathrm{timeout}}\big) \tag{5}$$

$$p_{\mathrm{timeout}} \;=\; (1-U)\,w,$$

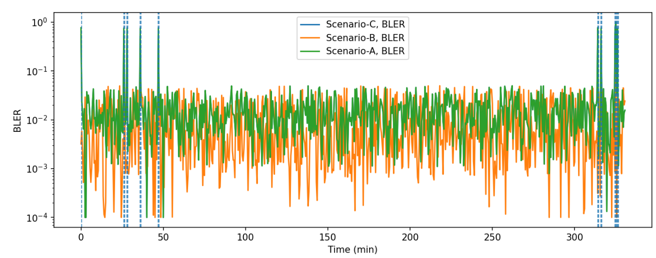
with w weight for the late-ACK penalty.



(a) Block Error Rate - HARQ process 32



(b) Block Error Rate - HARQ process 64



(c) Block Error Rate - HARQ process 128

Fig. 5. Comparison of BLER in Scenario A, B and C with HARQ processes of 32, 64 and 128

As shown in Fig. 5, when the HARQ process cap is limited to 32, the BLER increases sharply, ranging from approximately 0.1 to 0.8 in LEO and approaching nearly 100% loss in MEO; Fig. 5(a). Increasing the cap to 64 significantly improves performance in LEO with a regenerative payload, where the vDU is hosted onboard the LEO satellite (e.g., Scenario C), resulting in BLER values comparable to the baseline Scenario B. However, with 64 processes, the long RTT of a transparent LEO or MEO cannot be compensated, and BLER remains high; Fig. 5(b). A further increase of the HARQ process cap to 128 brings BLER in transparent LEO closer to the baseline Scenario B, but still shows no improvement for MEO; Fig. 5(c).

We then conducted a more detailed analysis of the regenerative LEO case at 1200 km altitude (Scenarios C and D), as summarized in Fig.6. The results show that increasing the HARQ cap from 32 to 64 reduces the average BLER from 0.176 to 0.017, representing an order-of-magnitude improvement. Beyond 64 processes (e.g., 128 or 285), no further significant BLER gains are observed, indicating a saturation effect. A similar trend was observed for Scenario A, where convergence occurred at a HARQ cap of 128.

However, its corresponding plot is omitted here since it closely mirrors the behavior already illustrated in Fig.6.

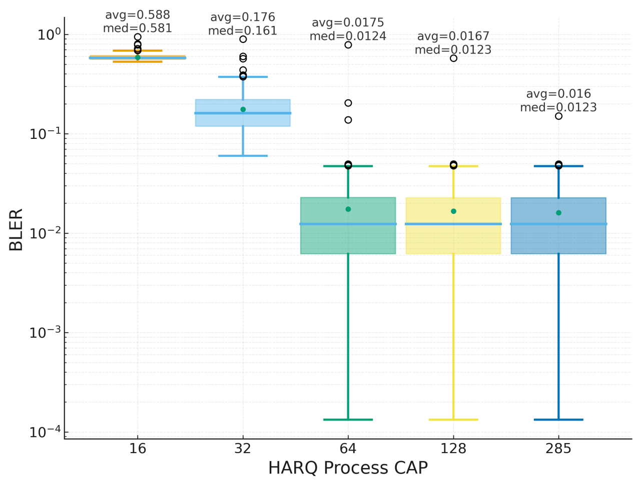


Fig. 6. BLER observation at different HARQ cap for scenario C & D

V. Cloud-Native dRAN for Adaptive DU Placement

A key contribution of this work is to highlight the role of cloud-native dRAN in enabling flexible deployment of the DU across different NTN/TN nodes. Traditional RAN architectures assume a fixed location for RU/DU/CU functions, which in high-speed mobility such as NTN aviation scenarios can lead to severe performance degradation due to Hybrid-ARQ (HARQ) timing constraints. Our results demonstrate that when the DU is placed on a satellite, the large round-trip delay between the UEs (e.g., passengers cellphones onboard airplane & onboard satellite DU) limits the number of HARQ processes that can be supported, especially on higher altitude satellites. In contrast, co-locating RU and DU onboard the aircraft eliminates this bottleneck, helping to remove the negative effects of HARQ process on BLER and ultimately on achieved throughput. Cloud-native RAN principles (containerization, micro-services, orchestration) allow the DU to be dynamically instantiated in different locations depending on network conditions. This approach enables a new level of adaptability for NTN aviation, namely:

Low-latency LEO case: When the aircraft is served by a nearby LEO with sufficient latency budget, the DU can remain on the satellite, reducing onboard energy consumption while sustaining the target BLER by only adjusting the HARQ process.

High-latency MEO or GEO: When HARQ pipeline starvation threatens BLER, the DU can be seamlessly instantiated onboard the aircraft. The onboard RU then connects locally to the DU, and only higher-layer functions remain satellite-anchored.

This cloud-native approach can potentially help to optimize error-rate, throughput and energy. The DU is only moved onboard when necessary to prevent the aircraft from continuously hosting high-power processing when link conditions do not demand it. The cloud-native adaptability provides a novel design lever for NTN aviation: instead of a static choice between transparent and regenerative architectures, the RAN can be flexibly reconfigured per-orbit and per-user condition. Such elasticity is also aligned with 6G system design objectives of resiliency, efficiency, and sustainability.

VI. Discussion and Future Directions

Building on the analysis of the four disaggregated RAN implementation scenarios for in-flight connectivity, each approach shows distinct trade-offs in latency, scalability, computational load, and architectural complexity, as summarized in Table I. Integrating cloud-native RAN into a multi-layer satcom architecture adds flexibility, enabling dynamic network adaptation and optimized resource allocation across airplanes, satellites, and ground infrastructure. Cloud-native RAN enables the orchestration of network functionsin a modular, service-based manner, supporting seamlesstransitions between different deployment models as opera-tional conditions change. This adaptability can be beneficial in the dynamic and constrained environment of in-flight connectivity, where latency, mobility, and coverage continuity must be tightly managed. With modular RAN components deployable near UEs, at the Far-Edge, or in the aggregated edge-cloud, operators can select the optimal setup for airborne connectivity. Building on this flexibility, future research should focus on further enhancing cloud-native RAN effectiveness in these dynamic environments.

AI-driven prediction for optimized deployment: Leveraging AI-driven prediction models offers a promising path for further enhancements. By analyzing real-time flight data, traffic patterns, satellite availability, and link performance, AI can support the intelligent selection of functional splits and optimal architectural deployment on a per-flight or even per-phase-of-flight basis. Such predictive capabilities can help maximize Quality of Experience (QoE) for passengers while minimizing operational costs for service providers.

Enhancing onboard computing for efficient NTN cloud-RAN: Future research should also explore how onboard computing capabilities can be advanced to address the inherent limitations of LEO satellites, which face constraints such as limited power budgets, reliance on solar energy, and thermal management in the vacuum of space [18]. While technologies like the EDGE-1100 Payload Processor and FPGAs are expanding onboard processing potential, Cloud RAN remains limited by the need for scalable and distributed architectures that can effectively handle complex RAN tasks under energy and environmental constraints. Research must focus on balancing onboard and ground-based processing to achieve optimal performance, considering that while ground systems offer slightly better performance (6-8%), onboard computing can be more energy-efficient (3-9% less energy per 100 images processed) [17]. Future work should thus define strategies for dynamic task distribution and scaling, ensuring robust, energy-aware, and fault-tolerant cloud-native RAN architectures for NTN.

ACKNOWLEDGMENT

This work was funded by the European Union under the Horizon Europe programme through the SNS JU (Grant Agreement No. 101192912, **NexaSphere Project**). Views expressed are those of the authors and do not necessarily reflect those of the EU or the SNS JU.

REFERENCES

[1] X. Lin *et al.*, "On The Path to 6G: Embracing the next wave of low earth orbit satellite access," *IEEE Communications Magazine*, vol. 59, no. 12, pp. 36–42, 2021.

[2] A. Guidotti *et al.*, "Role and evolution of Non-Terrestrial Networks towards 6G systems," *IEEE Access*, 2024.

[3] "3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Study on integration of satellite components in the 5G architecture; Phase 3," 3GPP TR 23.700-29, Tech. Rep., 2024.

[4] 3GPP Non Terrestrial Networks (NTN) for aviation connectivity. 3GPP. Access on April, 2025.

[5] B. Mafakheri *et al.*, "Edge intelligence in 5G and beyond aeronautical network with LEO satellite backhaul," in *2023 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit)*. IEEE, 2023, pp. 579–584.

[6] A. Papa *et al.*, "Enabling 6G Applications in the Sky: Aeronautical federation framework," *IEEE Network*, 2023.

[7] M. Ozger *et al.*, "6G for Connected Sky: A vision for integrating Terrestrial and Non-Terrestrial Networks," in *2023 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit)*. IEEE, 2023, pp. 711–716.

[8] "3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Study on Satellite Access - Phase 4, v1.1.0," 3GPP TR 22.887, Tech. Rep., 2025.

[9] "3rd Generation Partnership Project; Technical Specification Group TSG SA; Study on 6G Use Cases and Service Requirements; Stage 1," 3GPP TR 22.870, Tech. Rep., 2025.

[10] M. Corici *et al.*, "Transforming 5G Mega-Constellation Communications: A Self-Organized Network Architecture Perspective," *IEEE access*, 2025.

[11] H. Shahid *et al.*, "Emerging advancements in 6G NTN Radio Access Technologies: An Overview," in *2024 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit)*. IEEE, 2024, pp. 593–598.

[12] K. Ntontin *et al.*, "A vision, survey, and roadmap toward space communications in the 6G and beyond era," *Proceedings of the IEEE*, 2025.

[13] Final Validation and Use Case Benchmarking. Access on April, 2025.

[14] NexaSphere PoC 1: Next-Generation In-Flight Connectivity (IFC). Access on June, 2025.

[15] F. Veisi *et al.*, "Non-Terrestrial Disaggregated RAN: Prospective Architectures and Performance Insights," *IEEE Network*, 2025.

[16] "3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Solutions for NR to support non-terrestrial networks (NTN)," 3GPP TR 38.821, Tech. Rep., 2023.

[17] Z. Lai *et al.*, "{StarryNet}: Empowering Researchers to Evaluate Futuristic Integrated Space and Terrestrial Networks," in *20th USENIX Symposium on Networked Systems Design and Implementation (NSDI 23)*, 2023.

[18] R. Xing *et al.*, "Deciphering the Enigma of Satellite Computing with COTS Devices: Measurement and Analysis," in *Proceedings of the 30th Annual International Conference on Mobile Computing and Networking*, 2024.

[19] IRIS^2: the new eu secure satellite constellation. Access on June, 2025.

[20] C. Yan and B. Mafakheri, "Satellite connectivity prediction for fast-moving platforms," *arXiv preprint arXiv:2508.00877*, 2025.

[21] L. Méndez-Monsanto *et al.*, "Bler-snr curves for 5g nr mcs under awgn channel with optimum quantization," in *2024 IEEE 100th Vehicular Technology Conference (VTC2024-Fall)*. IEEE, 2024.

[22] K. Brueninghaus *et al.*, "Link performance models for system level simulations of broadband radio access systems," in *2005 IEEE 16th international symposium on personal, indoor and mobile radio communications*, vol. 4. IEEE, 2005, pp. 2306–2311.