D2.4: Final report on ETHER network architecture, interfaces, and architecture evaluation

Version: v.1.0

Work package	WP 2
Task	Task 2.2
Due date	M18
Submission date	30/08/2024
Deliverable lead	OPL
Version	1.0
Editors	Lechosław Tomaszewski (OPL), Robert Kołakowski (OPL), Rafał Tępiński (OPL)
Reviewers	Konstantinos Ntontin (UL), Symeon Chatzinotas (UL), Andrés Cárdenas (I2CAT), Reza Mosahebfard (I2CAT), Agapi Mesodiakaki (AUTH)
Abstract	This deliverable presents the final ETHER architectural framework integrating terrestrial and non-terrestrial networks. The system design is based on the recent state-of-the-art solutions and requirements specific to the ETHER use cases and technical innovations. The proposed architectural framework covers in detail infrastructure, management and orchestration, end-to-end service and Artificial Intelligence layers, initially outlined in D2.1. The intra-layer and inter-layer interactions and interfaces are described together with the proposed implementation technologies. This deliverable also outlines the architecture evaluation regarding energy efficiency and total cost of ownership, in the context of the ETHER use cases.
Keywords	Architecture, network, interfaces, TN, NTN, integration, SDN, NFV, 5GS, 6GS, MANO, AI, LEO

www.ether-project.eu

Co-funded by the European Union

6GSNS

Grant Agreement No.: 101096526
Call: HORIZON-JU-SNS-2022

Topic: HORIZON-JU-SNS-2022-STREAM-B-01-03
Type of action: HORIZON-JU-RIA

Document Revision History

Version	Date	Description of change
V0.1	27/09/2023	Creation of the first draft
V0.2	30/04/2024	Change of document structure
V0.3	20/06/2024	Change of document structure
V0.4	16/07/2024	Accepted all changes. Added references.
V0.5	14/08/2024	Accepted all changes. Added references.
V0.6	19/08/2024	Accepted all changes. Prepared for the internal review.
V0.7	23/08/2024	Accepted changes. Addressing comments from the internal review.
V0.8	28/08/2024	Accepted changes. Addressing the comments from the final internal reviews.
V1.0	29/08/2024	Final formatting

List of authors

Organisation	Author
UL	Konstantinos Ntontin, Symeon Chatzinotas
AUTH	Agapi Mesodiakaki, Marios Gatzianas, Dimitrios Palianopoulos, George Kalfas, Charalampos Bratsoudis
CA	Hamada Alshaer
AVA	Simon Watts, George Fotiou
SIOT	Sergio Aguilar Romero, Shubham Gupta, Diogo Oliveira
UBW	Luís Santos, Hélio Simeão
I2CAT	Joan Adrià Ruiz-de-Azua, Andrés Cárdenas, Reza Mosahebfard, Jose Avila Acosta, Victor Montilla
NBC	Angelos Antonopoulos, Godfrey Kibalya
NCSRD	Fotis Lazarakis
LIU	Nikolaos Pappas
OPL	Lechosław Tomaszewski, Robert Kołakowski, Rafał Tępiński
MAR	-
NETAI	Marco Fiore, Alexis Duque

DISCLAIMER

ETHER (*sElf-evolving terrestrial/non-Terrestrial Hybrid nEtwoRks*) project has received funding from the **Smart Networks and Services Joint Undertaking (SNS JU)** under the European Union's **Horizon Europe research and innovation programme** under Grant Agreement No 101096526. This work has received funding from the **Swiss State Secretariat for Education, Research and Innovation (SERI)**.

COPYRIGHT NOTICE

Project co-funded by the European Commission in the Horizon Europe Programme		
Nature of the deliverable:	**R***	
Dissemination Level		
PU	*Public, fully open, e.g., web (Deliverables flagged as public will be automatically published in CORDIS project's page)*	✓
SEN	*Sensitive, limited under the conditions of the Grant Agreement*	
Classified R-UE/EU-R	*EU RESTRICTED under the Commission Decision No. 2015/444*	
Classified C-UE/EU-C	*EU CONFIDENTIAL under the Commission Decision No. 2015/444*	
Classified S-UE/EU-S	*EU SECRET under the Commission Decision No. 2015/444*	

* R: Document, report (excluding the periodic and final reports)

DEM: Demonstrator, pilot, prototype, plan designs

DEC: Websites, patents filing, press & media actions, videos, etc.

DATA: Data sets, microdata, etc.

DMP: Data management plan

ETHICS: Deliverables related to ethics issues.

SECURITY: Deliverables related to security issues

OTHER: Software, technical diagram, algorithms, models, etc.

EXECUTIVE SUMMARY

This document presents the final design of the ETHER architectural framework, aimed at integrating Terrestrial Networks (TNs) and Non-Terrestrial Networks (NTNs) to support the envisioned 6G services across the terrestrial, aerial and space strata.

First, the analysis of the current state of the art is presented, which focuses on trends and concepts related to the integration of TN-NTN, including envisioned 6G network evolution, key enablers that can contribute to the integrated TN-NTN ecosystem as well as concepts for integration from the perspective of relevant Standards Developing Organisations (SDOs), i.e., the 3rd Generation Partnership Project (3GPP), the International Telecommunication Union (ITU) or the Internet Engineering Task Force (IETF). This is followed by the description of ETHER use cases (UCs) that are relevant to the architecture definition. The remaining issues and challenges are also outlined to emphasise the gaps in currently adopted approaches. The challenges regarding TN-NTN communication constitute the rationale for the development of ETHER innovations. Afterwards, a brief description of novel mechanisms/technical innovations supporting TN-NTN integration, which are needed to facilitate the considered ETHER UCs, is provided. These innovations include direct handheld device access from Low Earth Orbit (LEO) satellites, unified waveform design, flexible payload, Artificial Intelligence (AI)-based predictive vertical/horizontal handover control mechanisms, data analytics, edge computing, and caching. This also includes semantics-aware data analytics and control, zero-touch orchestration and AI-driven End-to-End (E2E) optimisation. Collectively, these considerations constitute the motivation for the architectural design.

Next, the description of the final ETHER architecture is presented, starting with an outline of key design principles (generic, multi-domain, multi-provider, distributed, hierarchical, intent-based, zero-touch, scalable, service-based, modular, and open to extensions). This is followed by a high-level overview of the architectural layers, including Infrastructure, ETHER Management and Orchestration, E2E Service and AI layers. Subsequently, an in-detail presentation of each layer is provided, together with key targeted challenges and proposed solutions and innovations. The architecture description is concluded with a detailed description of inter- and intra-layer interactions as well as interfaces between architecture components, serving the role of an Interface Control Document for future ETHER architecture implementations.

The document is wrapped up with the ETHER architecture evaluation, which aims to analyse the contribution of the undertaken approach to key ETHER and 6G performance metrics, namely, energy- and cost-efficiency. To enable an E2E assessment, a complete network model is employed incorporating the power consumption of both computational nodes and X-haul. The simulation-driven tests, comprising various traffic distribution scenarios that correspond to the ETHER UCs, have demonstrated significant gains of the ETHER approach in terms of the aforementioned performance metrics while maintaining the required Quality of Service (QoS) for the network users.

Co-funded by the European Union 6GSNS

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

3D	3-Dimensional
3GPP	Third Generation Partnership Project
5G PPP	5G Public Private Partnership
5G	5th Generation of mobile networks
5GS	5G System
6G IA	6G Smart Networks and Services Industry Association
6G	6th Generation of mobile networks
6GS	6G System
AF	Application Function
AI	Artificial Intelligence
AIA	Authentication Information Answer
AIaaS	AI as a Service
AIOps	AI Operations
AIR	Authentication Information Request
AMF	Access and Mobility Management Function
AP	Application Plane
API	Application Programming Interface
ATM	Air Traffic Management
AUSF	Authentication Server Function
AV	Authentication Vector
B5G	Beyond 5G
BS	Base Station
BSS	Business Support System
CAPIF	Common Application Programming Interface Framework
CIoT	Cellular IoT
CM	Configuration Management
CN	Core Network

COTM	Communications On The Move
COTP	Communications On The Pause
CP	Control Plane
CP-OFDM	Cyclic Prefix OFD
CPU	Central Processing Unit
C-SGN	CIoT Serving Gateway Node
CSMF	Communication Service Management Function
CU	Central Unit
DevOps	Development Operations
DFT-s-OFDM	Discrete Fourier Transform-spread OFDM
DI	Domain Infrastructure
DL	DownLink
DMMF	Domain Mobility Management Function
DMOC	Domain M&O Component
DN	Data Network
DP	Data Plane
DU	Distributed Unit
E2E	End-to-End
E2EAO	E2E Application Orchestrator
E2ENO	E2E Network Orchestrator
E2ESL	E2E Service Layer
EASDF	Edge Application Server Discovery Function
ECI	Earth-Centred Inertial
ECM	EPS Connection Management
EGMF	Exposure Governance Management Function
EIR	Equipment Identity Registry
eMBB	enhanced Massive BroadBand
EMOC	E2E M&O Component

eMTC	enhanced Machine Type Communication
eNB	evolved Node B
ENI	Experiential Networked Intelligence
EPC	Evolved Packet Core
EPS	Evolved Packet System
ETSI	European Telecommunication Standards Institute
EU	European Union
E-UTRAN	Evolved Universal Terrestrial Radio Access Network
FCAPS	Fault, Configuration, Accounting, Performance, Security
FM	Fault Management
FPGA	Field Programmable Gate Array
GBR	Guaranteed Bit Rate
GEO	Geostationary Orbit
GIS	Geographical Information System
GISF	Geographical Information System Function
GMMF	Global Mobility Management Function
gNB	next-generation Node B
GNSS	Global Navigation Satellite System
GPU	Graphical Processing Unit
GS	Ground Station
GSM	Global System for Mobile communication
GSMA	GSM Association
GST	Generic Network Slice Template
GTO	Geostationary Transfer Orbit
GUTI	Global Unique Temporary ID
HAPS	High Altitude Platform System
HMTC	High performance Machine Type Communications
HSS	Home Subscriber Server

HTTP	Hypertext Transfer Protocol
ICD	Interface Control Document
IE	Information Element
IETF	Internet Engineering Task Force
IMSI	International Mobile Subscriber Identity
IMT	International Mobile Telecommunications
IoT	Internet of Things
IP	Internet Protocol
ISL	Inter-Satellite Link
ISM	In-Slice Management
ITU	International Telecommunication Union
JSON	JavaScript Object Notation
KPI	Key Performance Indicator
LAN	Local Area Network
LCM	Life Cycle Management
LD	Low-Density
LEO	Low Earth Orbit
LMMF	Local Mobility Management Function
M&O	Management and Orchestration (activities, processes)
MAcR	Multi-Access Rule
MANO	MANagement and Orchestration (architectural framework, software suite)
MAPE-K	Monitor-Analyse-Plan-Execute over a shared Knowledge
MAR	Mobile Autonomous Reporting
MDAF	Management Data Analytics Function
MEAO	MEC Application Orchestrator
MEC	Multi-Access Edge Computing
MEO	Medium Earth Orbit
MEP	MEC Platform

MEPM	MEC Platform Manager
mIoT	massive IoT
ML	Machine Learning
MLOps	Machine Learning Operations
MME	Mobility Management Entity
mMTC	massive Machine-Type Communications
MNO	Mobile Network Operator
MO	Mobile-Originated
MP	Management Plane
MPC	Model Predictive Control
MSDNO	Master SDNO
MT	Mobile-Terminated
NAS	Non-Access Stratum
NBI	NorthBound Interface
NB-IoT	Narrowband IoT
NEF	Network Exposure Function
NEST	NEtwork Slice Type
NF	Network Function
NFMF	Network Function Management Function
NFV	Network Functions Virtualisation
NFVO	NFV Orchestrator
NG	Next Generation
NGAP	NG Application Protocol
NG-RAN	Next Generation Radio Access Network
NGSO	Non-Geostationary Orbit
NI	Network Intelligence
NIF	Network Intelligence Function
NIF-C	Network Intelligence Function Component

NIO	Network Intelligence Orchestrator
NIP	Network Intelligence Plane
NIS	Network Intelligence Service
NR	New Radio
NRF	Network Repository Function
NSI	Network Slice Instance
NSMF	Network Slice Management Function
NSSAAF	Network Slice-Specific Authentication and Authorisation Function
NSSaaS	Network Slice Subnet as a Service
NSSF	Network Slice Selection Function
NSSMF	Network Slice Subnet Management Function
NTN	Non-Terrestrial Network
NWDAF	NetWork Data Analytics Function
OF	OpenFlow
OFDM	Orthogonal Frequency Division Multiplexing
O-RAN	Open RAN
OS	Operating System
OSM	Open Source MANO
OSS	Operations Support System
OTFS	Orthogonal Time-Frequency Space
P2P	Point-to-Point
PCF	Policy Control Function
PDI	Packet Detection Information
PDU	Protocol Data Unit
PFCP	Packet Forwarding Control Protocol
PLMN	Public Land Mobile Network
PM	Performance Management
PNF	Physical Network Function

QoE	Quality of Experience
QoS	Quality of Service
RAN	Radio Access Network
RAT	Radio Access Technology
RCP	Required Communication Performance
REST	REpresentational State Transfer
RF	Radio Frequency
RIC	Radio Intelligent Controller
RO	Resource Orchestrator
RT	Real-Time
SA	Stand-Alone
SBA	Service-Based Architecture
SBI	SouthBound Interface
SCP	Service Communication Proxy
SD	Slice Differentiator
SDN	Software-Defined Network
SDNC	Software-Defined Network Controller
SDNO	SDN Orchestrator
SDO	Standards Developing Organisation
SDR	Software-Defined Radio
SD-WAN	Software-Defined Wide Area Network
SEAF	SEcurity Anchor Function
SEAL	Service Enabler Architecture Layer
SF	Supplementary Function
SFC	Service Function Chain
SIB	System Information Block
SINR	Signal to Interference plus Noise Ratio
SMF	Session Management Function

SMO	Service Management and Orchestration
SNS JU	Smart Networks and Services Joint Undertaking
S-NSSAI	Single Network Slice Selection Assistance Information
SoC	System on Chip
SotA	State of the Art
SRI	Satellite Radio Interface
SST	Slice Service Type
SUCI	SUbscription Concealed Identifier
SUPI	SUbscription Permanent Identifier
TA	Tracking Area
TCO	Total Cost of Ownership
TLE	Two-Line Element
TN	Terrestrial Network
TR	Technical Report
TS	Technical Specification
TSN	Time-Sensitive Network
UART	Universal Asynchronous Receiver/Transmitter
UAV	Unmanned Aerial Vehicle
UC	Use Case
UDM	User Data Management
UDR	User Data Repository
UE	User Equipment
UL	UpLink
ULA	Update Location Answer
ULR	Update Location Request
UP	User Plane
UPF	User Plane Function
URE	User Reachability Entity

URLLC	Ultra-Reliable Low Latency Communications
V2X	Vehicle-to-Everything
VIM	Virtualised Infrastructure Manager
VM	Virtual Machine
VNF	Virtual Network Function
VNFM	VNF Manager
WAN	Wide Area Network
WG	Working Group
WIM	WAN Infrastructure Manager
WP	Work Package
XnAP	Xn Application Protocol
xNF	any type of Network Function
ZSM	Zero-touch network and Service Management

1 INTRODUCTION

1.1 OBJECTIVES

The primary objective of this deliverable is to define the final ETHER architecture supporting the use cases (UCs) to be demonstrated by WP5 within a federated, multi-domain and integrated Terrestrial Networks (TNs) and Non-Terrestrial Networks (NTNs) ecosystem encompassing terrestrial, aerial and space segments. This document concludes the joint effort of WP2, WP3 and WP4 to articulate a coherent 3-dimensional (3D) ETHER system architecture. It also serves as a fundamental guideline for the efficient implementation of 3D network services utilising new features, mechanisms and innovations proposed by the ETHER project. The cross-WP cooperation primarily involved the investigation of needs regarding dedicated architectural support (e.g., interfaces, exposure mechanisms, integration capabilities with external frameworks, new network mechanisms, etc.) for management and orchestration solutions developed in WP4, such as advanced orchestration frameworks, resource allocation algorithms, predictive analytics, etc., as well as handover mechanisms defined in WP3 (T3.4). This deliverable describes end-to-end (E2E) data, control and management planes using centralised or distributed computing resources. It specifies the finalised approach to cross-technology and cross-layer interfaces, thus also serving as the Interface Control Document (ICD) and provides the approach to the architecture evaluation and results.

1.2 SCOPE

The deliverable addresses the emerging challenges and opportunities presented by the seamless integration of NTNs with terrestrial infrastructure, focusing on the potential of an integrated TN-NTN architecture and interfaces ready for 6th Generation of mobile networks (6G). The proposal aims to leverage the well-established paradigms to accommodate standardised frameworks and achieve low latency, and zero-touch architecture that ensures efficient resource utilisation and seamless communications across diverse network elements.

1.3 DELIVERABLE STRUCTURE

This document is structured as follows:

- Section 2 presents concise explanations of terminologies used throughout the document to mitigate potential ambiguities arising from different naming conventions used by various Standards Developing Organisations (SDOs);
- Section 3 details the motivation driving the design of the integrated 3D ETHER architecture by outlining the evolution of 6G and key trends, architectural enablers, pertinent initiatives concerning the integration of TN-NTN (including standardisation), requirements of the ETHER UCs to be supported by the architecture, the pressing issues and challenges regarding the TN-NTN integration, and finally, the ETHER innovations (supported within the 3D ETHER architectural framework) targeting these challenges;
- Section 4 presents the final ETHER architecture, including the design principles and a detailed description of system layers – namely, Infrastructure Layer, ETHER Management and Orchestration (MANO) Layer, E2E Service Layer, and Artificial Intelligence (AI) Layer. This section also explores the interactions between layers, including the reference points and interfaces between architecture components (fulfilling the ICD role);

- Section 5 discusses the evaluation of ETHER architecture performed through simulations related to the ETHER UCs and analysis of the achieved gains in terms of key 6G performance metrics that align with the ETHER's objectives (energy consumption, cost-efficiency, communication quality);

- Section 6 summarises and concludes the document.

2 TERMINOLOGY AND CONVENTIONS

The ETHER project considers different systems and architectural approaches related to different frameworks. Given the diverse terminologies and potential ontological conflicts arising from different meanings assigned to the same terms, it is critical to address these inconsistencies to prevent reader's confusion. Additionally, there may arise conflicts of the same acronyms traditionally used in different domains for different terms.

Below are specific definitions for the concept of a service, as well as functional planes.

A **service** is defined and understood differently in different frameworks:

- European Telecommunication Standards Institute (ETSI) Network Function Virtualisation (NFV) network service is defined as a group of interconnected network functions building a communication network;
- 3rd Generation Partnership Project (3GPP) communication service is built on top of the 3GPP network;
- Service-Based Architecture (SBA) concept in which software instances, as producers, expose their services to other software instances acting as consumers;
- Application service runs on top of communication networks;
- ETHER E2E service is defined according to various contexts discussed in Section 4.

In this document, all references to services will be associated with an explicit statement of the context, such as “NFV network service”, “communication service”, “ETHER E2E service”, to avoid ambiguity. However, this requirement may be waived when the context within a section or a paragraph clearly defines the term.

Control Plane (CP) is utilised by the Software-Defined Networking (SDN) and 3GPP frameworks. In this document, all references to the CP will explicitly state the context, e.g., “5G System (5GS) CP” or “SDN CP”, to eliminate ambiguity. Exceptions to this rule may occur when the context provided by a section or paragraph clearly defines the term.

User Plane (UP) will be used according to the 3GPP concept of the user data forwarding plane and should not be confused with the **Data Plane** (DP) as per the SDN framework or various application architectures. In particular, in the case of SDN, the data plane can transport data of the 3GPP UP, 3GPP CP, Management Plane (MP), etc.

In terms of the 3D approach, the **space** stratum is considered equivalent to the **satellite** stratum.

In the 3GPP system, **mobility management** pertains to the mobility of User Equipment (UE). In the ETHER project, mobility management also refers to the mobility of the infrastructure, particularly non-geostationary satellite-based systems, leading to dynamic changes in the system topology.

“**Management and orchestration**” term is used in two contexts in this document. For architectural frameworks or software suites, the acronym MANO is used (e.g., ETSI NFV MANO, ETHER MANO). For activities or processes, the acronym M&O is used.

In this document, the Application Plane utilises the acronym AP, which should not be associated with “application protocol” occurring in, e.g., various 3GPP protocols.

In the 3GPP documents, the same acronym “MAR” is used for both Mobile Autonomous Reporting and Multi-Access Rule. Therefore, to provide necessary distinction, for the latter term, the acronym “MAcR” will be used in this document.

Reference point/interface designations are italicised in this document, excluding the designations of the reference points of the ETHER System, which are capitalised.

3 MOTIVATION

The rapid evolution of communication technologies has led to the wide deployment of carrier-grade 5G networks. Currently, ongoing research efforts are oriented towards defining the principles, targeted UCs and general capabilities of the 6G mobile networks. The design of 4G and 5G systems has been exclusively optimised for TNs. Only recently, with the 3GPP's Release 17, has the importance of NTNs been recognised to address the large coverage gaps across the globe in remote and rural areas. In particular, features in the 5G architectural design have been identified to support the integration of the satellite segment into TNs with minimal impact. However, optimal network performance can only be achieved through a unified approach, where the network design is optimised for both the TN and NTN components. It is widely acknowledged that 6G networks will aim for TN/NTN unification by design, facilitating deployment of 3D multi-stratum architectures. Such architectures will encompass the terrestrial, aerial, such as Unmanned Aerial Vehicles (UAVs) and High Altitude Platform Systems (HAPSs), as well as space nodes, particularly those in non-geostationary orbits – Medium Earth Orbit (MEO)/Low Earth Orbit (LEO) satellites. These nodes can function as Base Stations (BSs) and offer communication, processing, and potentially caching functionalities. The integration of TN-NTN platforms will enable the provisioning of ubiquitous coverage, reduced transmission delays, and enhanced service quality, paving the way for the 6G UCs. As existing research has mainly focused on tackling technical challenges individually, there is a general lack of comprehensive solutions for efficient management of such integrated networks and tackling its key technical issues. Consequently, there is a clear motivation to develop a complete, E2E, efficient, and programmable architecture for integrated TN-NTN platforms, thereby enabling the establishment of sustainable business frameworks.

The key objective of the ETHER project is to provide such a holistic approach for integrated TN-NTN systems. ETHER targets 100% network coverage, 99.99999% service continuity and 99.99999% reliability with three times higher energy efficiency and a 95% reduction in Total Cost of Ownership (TCO) compared to current terrestrial-only deployments. To meet the above demanding goals, the ETHER system architecture must not only act as the ecosystem facilitating the 3D network Integration but also provide support for the incorporation of cutting-edge technologies at each system level.

This section outlines the key concepts, requirements and challenges driving the architectural design choices and development of the ETHER framework. First, we summarise the most recent approaches to integrated 3D networks (Section 3.1). Afterwards, we present the UC-specific requirements (Section 3.2), derived through in-depth analysis conducted within T2.1 and reported in D2.2 [1]. Finally, we present the issues and challenges concerning the TN-NTN integration that can be addressed by ETHER system architecture and supported innovations (Section 3.3).

3.1 TN-NTN INTEGRATION TRENDS AND CONCEPTS

Although the integration of TNs and NTNs is still an emerging concept, it is already recognised by the entire ecosystem of the telco industry and is expected to create multiple new techno-business opportunities [2] and partnerships [3]. Numerous standardisation bodies, along with research and development projects, have already published a wide range of papers on the topic. These publications cover various aspects, including business justification, societal needs and UCs, particularly addressing technological gaps, challenges, and yet unsolved technical issues as well as architectural concepts and solutions. Deliverable D2.1 [4] outlined the initial ETHER system architecture and provided an exhaustive outlook on the current State of the Art

(SotA) related to 5G/6G networks, NTNs and key architectural approaches. Hereby, we summarise the vital concepts that have influenced the design of the ETHER system.

3.1.1 Evolution towards 6G

With the 5G system slowly maturing, activities related to the evolution from 5G to 6G are starting to take place. There is a general consensus that the 6G System (6GS) should contribute to addressing significant societal challenges identified by the United Nations Sustainable Development Goals. These include environmental sustainability, digital inclusion, reducing inequality and poverty, enhancing public safety and privacy, and lowering carbon footprint [5], [6], [7], [8]. To address the latter, 6GS is envisioned to embrace highly intelligent components that allow for the convergence of human, physical and digital worlds (i.e., connected intelligence) [6] as well as elevate system capabilities to reach sustainability targets, including overall system energy efficiency. Moreover, 6G aims to provide worldwide ubiquitous network access and seamless connectivity [5], [6], which is to be achieved in a sustainable manner, i.e., by integrating into the system design the non-conventional solutions such as non-terrestrial HAPSs or satellite-based services to deliver mobile services [9], or by leveraging existing 5G network components [10] to minimise infrastructural investments. The 6G flagship project funded by the European Union (EU), Hexa-X, also highlights additional challenges [6] that the future network is expected to solve. Given its distributed, dynamic and heterogenous character, referred to as "Network of Networks", 6G would need to integrate multiple divergent solutions into a unified E2E network. Furthermore, to meet the growing expectations of the vertical sector for novel UCs, the 6GS will need to provide extreme bitrates, extremely low latencies, and seemingly infinite capacity [6].

The main way to describe what current networks are missing and what future networks could offer are the UCs. The most commonly envisioned UCs include robots/cobots, holographic telepresence communication, immersive multi-modal communication for teleoperation, extended reality, remote healthcare, massive twinning (which intertwines the physical and digital worlds to improve system efficiency and enable flexible adoption of AI), applications related to network evolution (such as Trusted Native AI and AI as a Service – AIaaS, coverage expansion, and energy efficiency), enabling services (e.g., 3D positioning, localisation and tracking) or local trust zones (UCs within specialised sub-networks and networks of networks, that require extreme reliability, availability and resilience) [2], [7], [11]. Based on these considerations, key architectural features of 6G have been identified [10], which include:

- Adoption of SBA and a cloud-native approach [12].
- Multi-access convergence for ubiquitous coverage and new network functions supporting seamless access across disparate networks, encompassing mobility management, routing, security, policy control, charging and subscriber data management.
- Deep integration of communication, sensing and computing at the architecture level, including the operational management towards a fully converged and simplified network.
- Zero-trust-based system security, achieved by embedding security mechanisms into the architecture design and leveraging zero-trust mechanisms.
- Network complexity reduction by phasing out legacy technologies and introducing self-organising and self-contained modules, building new features and network functions.
- Increased elasticity through disaggregation and softwarisation, flexible function placement and standardised interfaces.

- Energy efficiency in each network segment, with a focus on Radio Access Network (RAN) [13] to facilitate “Green Future Networks”.

Although the existing visions of 6G are not fully aligned, they share many similarities, one of which is the goal of providing global service coverage via the integration of TNs and NTNs (using HAPSs or satellites). The ETHER system architecture seeks to integrate TNs with NTNs to bring the current SotA one step closer to realising 6G. Development of 6GS architecture, architectural enablers and frameworks are some of the goals of the Hexa-X II project [14], which is the follow-up to the previously mentioned Hexa-X project.

3.1.2 Key architectural enablers

In the process of 5GS evolution and maturing, various technological enablers have been developed and are now widely adopted in the IT and Telco industry. This section briefly summarises the key enabling technologies and paradigms that are now considered to play an essential role in shaping future networks, which will be leveraged by the ETHER concept.

Virtualisation – A key architectural concept for future networks is ETSI NFV, which separates software-based implementation of Network Functions (NFs) from the underlying hardware by adding a virtualisation layer. This separation results in decoupled infrastructure and software life cycles, increasing flexibility and agility of network management. In the context of NTNs, such as satellite networks, adopting NFV allows for the dynamic deployment of Virtual Network Functions (VNFs) on-board satellites. This facilitates the broadening of the NFs portfolio of satellite network operators (VNF-as-a-Service) and allows flexible system evolution with reduced upgrade costs. The NFV principles are widely adopted in Beyond 5G (B5G) architectures [15], [16], [17] due to its sustainability, flexibility and resilience [12]. These traits are especially important in the context of the dynamic nature of NTNs.

Management and orchestration – The NFV MANO architectural framework [18] streamlines the M&O of virtualised networks by clearly defining the roles of its three main components: the NFV Orchestrator (NFVO), the VNF Manager (VNFM), and the Virtualised Infrastructure Manager (VIM). NFVO is responsible for the orchestration and Life Cycle Management (LCM) of “NFV Network Services”, which consist of a constellation of VNFs forming a complete communication network. The VNFM handles the LCM of individual VNFs. Finally, the NFV Infrastructure (NFVI) resources, including compute, storage and network, are managed by the VIM. The NFV MANO framework is agnostic to VNFs and NFV Network Services, ensuring broad applicability across various network architectures. In addition, the business aspects of the “NFV Network Service” are managed by a separate entity, the Operations Support System/Business Support System (OSS/BSS). Several NFV MANO applications are described [18].

The basic NFV MANO framework has evolved significantly, incorporating several features to enhance functionalities in the context of NTN. These enhancements include support for managing Wide Area Network (WAN) infrastructure interconnectivity through the WAN Infrastructure Manager (WIM). The WIM is capable of managing distributed, multi-data centre deployment [19] (currently limited to static data centres, with potential future extensions to non-static nodes), containerisation and containers networking [19], [20], both generic and VNF provider-specific VNFMs [21], multiple infrastructure domains (interaction between NFVO/VNFM with multiple VIMs and a “Resource Orchestrator” [21]), multiple orchestration domains (interconnections of multiple MANO stacks) [21] and multi-tenancy [22].

However, with the paradigm shift towards 6G networks, the NFV-MANO concept should undergo substantial modifications to support highly heterogeneous integrated TN-NTN ecosystems in a flexible and dynamic manner (especially those involving aerial and satellite infrastructure), ensuring capabilities comparable to terrestrial counterparts.

Different research projects have investigated and proposed several promising solutions based on the ETSI NFV MANO concept. MonB5G proposed a novel zero-touch multi-domain hierarchical slice MANO platform that leverages AI-driven operations and addresses the scalability issues in network slice management and orchestration [23]. Hexa-X proposed microservices-based 6G MANO architecture [17] that includes intent-based services, support for AI Operations (AIOps)/Development Operations (DevOps), zero-touch management and private networks. Additionally, the open-source community is developing an ETSI-compliant software implementation, namely Open-Source MANO (OSM) [24]. It is important to note that zero-touch M&O are particularly important for the integrated TN-NTN ecosystem advanced by ETHER, due to the high complexity and dynamicity of large-scale integrated networks. Such an ecosystem necessitates a careful design of the individual functions and components of the MANO framework to realise the zero-touch vision.

Artificial Intelligence – The ever-increasing complexity of modern telecommunication networks gives rise to a plethora of new operational and management challenges. This is especially true for the TN-NTN ETHER ecosystem that amalgamates the three strata, terrestrial, aerial, and space. It is commonly agreed, that exploiting AI/Machine Learning (ML) solutions will not only allow to handle the network processes complexity and provide improved performance (compared to classical approaches), but is also an essential technology to facilitate the TN-NTN integration. Regarding the latter, numerous studies, present the potential of AI/ML applications for optimisation of handover management, resources (e.g., power/spectrum allocation, energy consumption), transmission delays or introduction of new radio technologies (e.g., mmWave communication) [25]. One of the key challenges of embedding AI/ML solutions into the networks is to automate E2E M&O processes. The Experiential Networked Intelligence (ENI) is defining a Cognitive Network Management architecture, which uses AI techniques and context-aware policies to adjust services to actual user needs, conditions, and business goals and enables automated and optimised service provisioning, operation, and assurance, as well as optimised slice management and resource orchestration [26], [27]. Moreover, the ENI system is based on an experiential architecture enabling the system evolution over time to improve decision-making. In addition, ENI has also launched proofs of concepts that aim to demonstrate how AI methods can be leveraged to assist network operation, including 5G.

ETSI Zero-touch network and Service Management (ZSM) framework targets the E2E closed-loop-based automation of network and service management that is optimised for data-driven AI/ML algorithms. While the reference architecture has already been proposed [28], there is ongoing work on realising zero-touch closed-loop automation, generic enablers and solutions for closed-loop automation, learning and cognitive capabilities, oversight level evaluation, autonomy, and operational confidence on the closed loops behaviour [29]. Moreover, the applicability of Network Digital Twin is deeply investigated, including existing, emerging and future UCs, principles and functionality needed to leverage the Network Digital Twin for zero-touch network and service management [30] and missing ZSM framework features to support Network Digital Twins. The security reference architecture for the ZSM framework, including requirements for supporting robustness and security of AI/ML models and closed-loop automation during the system's life cycle, have been defined [31].

Innovative approaches to making AI native to the network architectures have also been proposed by EU projects. Hexa-X strongly pursues AI integration into the network layer and introduces the framework for AIaaS, as well as functional AI system architecture addressing EU AI regulations, which considers specific entities for risk-related information, self-verification, record keeping, human oversight, redundancy, system management and decision making [32]. The DAEMON project works on innovative approaches to Network Intelligence (NI) design, including an E2E NI-native and zero-touch B5G architecture supporting the coordination of NI-assisted functionalities and targeting performance, sustainability and reliability. It also introduces Network Intelligence Plane (NIP) responsible for AI/ML models management, operations coordination and deconfliction [33].

Software-Defined Networking (SDN) – The highly dynamic network environment of NTNs, with constantly changing network topology, requires a flexible and programmable forwarding plane. These capabilities can be provided using the SDN paradigm, which introduces its CP and DP separation as well as facilitates application-driven control by exposing north-bound control interfaces. One architecture of the integrated NTN (in this case, a satellite network) with the TN with the use of SDN Controllers (SDNCs) is proposed [34], which presents the idea of a hierarchical transport layer with a master SDNC and a number of slave SDNCs to increase scalability of operations.

Open-RAN (O-RAN) – Constituting cloud-native virtualised RAN, the O-RAN architecture aims to deliver flexible beyond 3GPP RAN M&O, i.e., Fault, Configuration, Accounting, Performance, Security (FCAPS), RAN resources management, data management interfaces, etc., and RAN-related optimisations. The most important functional components introduced by the framework are the Non-Real-Time (Non-RT) Radio Intelligent Controller (RIC) responsible for large timescale RAN optimisation (e.g., policy configuration or ML model training) and the Near-Real-Time (Near-RT) RIC for near-real-time optimisation & control and data monitoring. The latter is mainly responsible for radio resources management and supporting xApps – generic 3rd party applications that can act as the vestibule into RAN as well as perform RAN-related operations and optimisation (e.g., interference coordination, dynamic spectrum and power management, etc.). While the TN operators have started to introduce open access solutions for RAN, as of today, the NTNs still rely on closed architectures featuring proprietary solutions. The existing technological heterogeneity poses a challenge to providing interoperability between NTN nodes effectively hindering the optimisation of integrated 3D TN-NTN systems. The inclusion of O-RAN into the NTN systems will allow to address the above issues, due to the introduced flexibility, programmability and standardised interfaces enabling seamless integration of RAN solutions from multiple vendors [35].

Multi-Access Edge Computing (MEC) – Edge computing has been identified as a key component of 5G networks to reduce the E2E latencies involved with task processing far from where the information is generated, such as the terrestrial cloud serves. This is even more important for ETHER due to the immense amount of information expected to be generated by the integrated network. The deployment of MEC in NTN, can offer numerous benefits such as caching/storage of content of terrestrial users, service provisioning in remote areas (without operating TN MEC), and computation offloading to overcome energy limitations (e.g., satellite-based MEC supporting UAVs) [36]. Especially relevant for ETHER is the role of the MEC orchestrator and the MEC Platform Manager (MEPM) [37], as the core components that have the complete overview of the system and manage the specific functionality and the applications running on each particular MEC host. Recommendations on the MEC System integration with other frameworks have been published, including a 5G System [38] or NFV MANO (i.e., MEC-in-NFV option) [37]. Moreover, MEC provides extensive application mobility support via functionalities of user context and/or application migration and synchronisation across MEC hosts or mobility related application programming interfaces (APIs) [39], which are of key importance to implement service continuity both in TN and NTN systems.

Enablers proposed by other research projects – Finally, several research projects have conducted activities to address the existing gaps and propose enablers or testbed facilities to contribute to TN-NTN integration. 5G-COMPLETE proposed a converged Computing/Storage/RAN infrastructure to merge the 5G New Radio (NR) fronthaul/midhaul/backhaul facilities into a common platform and transform the RAN into a low-power distributed computer [40]. QUANGO aimed to design a network of 12-Unit CubeSat LEO satellites offering combined capabilities for communication secured by quantum key distribution for 5G IoT NTN connection. Among other achievements, it delivered payload hardware (developed by SIOT and used a starting point for the flexible payload and the ETHER UC1), which provides 5G Internet of Things (IoT) NTN and Store and Forward functionalities. The SaT5G project developed a cost-efficient “plug and play” satcom solution for 5G to enable telcos and network vendors to accelerate 5G deployment [41]. EAGER aims to advance the existing NTN innovations (such as flexible payloads, inter-NTN links, waveform design or AI/ML) to prepare for next-generation satellite networks, targeting highly efficient and deeply integrated 3D satellite networks in B5G and 6G [42]. The 5GENESIS project created a testbed supporting 5G satellite backhaul with network slices [43]. 6G-SANDBOX combines digital and physical nodes to deliver fully configurable, manageable and controllable E2E networks for 6G research and technology validation. Among other activities, the project develops a satellite emulator to study hybrid access and dual connectivity between TN and NTN. Context-aware management is one of three architectural enablers recognised by the Hexa-X II project [14]. These enablers allow for dynamic adaptation to context in order to ensure the expected E2E QoS for the services and the expected Quality of Experience (QoE) for users.

3.1.3 TN-NTN use cases and requirements

The TN-NTN integration opens up opportunities to implement novel UCs. 5G Americas has proposed HAPS-based ones, such as greenfield coverage, white spot reduction, emergency communications and disaster recovery, IoT support, temporary coverage for events and tourist hotspots, fixed wireless access, connectivity for urban air mobility and drones, private networks, terrestrial site backhaul, and extended coverage over the sea [2]. Additionally, UCs for 3GPP service categories have been defined, including for enhanced Massive Broadband (eMBB) and massive Machine-Type Communications (mMTC) [5]. Furthermore, the Internet Engineering Task Force (IETF) has proposed satellite network deployment scenarios to be used for broadband internet access (single satellite as a relay, multiple satellite relays, and inter-satellite networking) or satellite network 5G backhaul [44].

The topic of 5G-NTN integration has been extensively studied by 3GPP. Starting with general NTN applications, i.e., coverage extension, IoT, disaster communication, global roaming, and broadcasting, 3GPP has identified very specific requirements [45], including:

- roaming between TN and NTN,
- broadcast and multicast with a satellite overlay,
- IoT with a satellite network, temporary use of a satellite component,
- optimal routing or steering over a satellite,
- satellite transborder service continuity,
- global satellite overlay,
- indirect connection through a 5G satellite access network,
- 5G fixed backhaul between 5G Radio Access Technology (RAT) – NR and the 5G Core Network (CN),
- 5G moving platform backhaul, 5G to premises (serving white spots),
- satellite connection of a remote service centre to an off-shore wind farm.

Within Release 19, new UCs have been proposed [46], including Store and Forward for a delay-tolerant/non-real-time IoT, drones connectivity or Local Area Network (LAN) over satellite access, information exchange between ships at sea, direct UE-satellite-UE communication for feeder link off-loading (including service continuity during NTN-TN handover), utilising satellite connectivity for data collection to aid TN planning, fleet management in the desert, and service differentiation for UEs via satellite access.

Alongside the specifications of the new technology, the requirements that standardisation organisation partnerships such as 3GPP must adhere to, are being defined. The service and network capability requirements, as well as UCs of fixed, mobile and satellite convergence for the International Mobile Telecommunications (IMT) 2020 (IMT-2020, i.e., 5G) networks and beyond, have been gathered in the International Telecommunication Union (ITU) Recommendation ITU-T Y.3200 [47]. The key requirements include:

- unified identity and unified subscription management, including unified authentication and authorisation,
- unified service provision and charging,
- service continuity during vertical handovers; support of best-effort and Guaranteed Bit Rate (GBR) QoS,
- support of various types of terminals,
- support for MEC, vertical, and roaming services,
- support of converged control, service, and management planes,
- support of cloud-based infrastructure enabled by NFV and SDN,
- support of AI/ML,
- support of unified mobility management, unified session management, unified connection management and policy control,
- unified exposure of converged network mechanisms,
- converged security,
- support for both satellite access and backhaul in the converged network.

The Recommendation ITU-T Y.3201 [48] proposes key design considerations, framework, enabling technologies, network function enhancements, procedures, and security considerations of fixed, mobile and satellite convergence for IMT-2020 networks and beyond. The framework includes the multi-access UE connected to a fixed, mobile or satellite access network (each connected to a land- or satellite-based core network joined using the interworking function). Land- or satellite-based service platforms and data networks are interconnected with the respective cores. While the land-based core implements the full functionality, the satellite-based core has to support only the lightweight implementations of user subscription management entity, network access and mobility control, session management and User Plane Function (UPF). Ephemeris data of satellites other than geostationary orbit (GEO) have to be supported by the mobility management and service continuity procedures.

3GPP has defined the functional and QoS-related requirements, such as delay limits for GEO/MEO/LEO options as well as for different usage scenarios, DownLink (DL)/UpLink (UL) data rates and area traffic capacities, overall user densities, user activity factors, expected UE speeds and types [49].

The IETF has presented problems and requirements related to the usage of satellites for Internet access [44] focusing on communication between satellites and Ground Stations (GSs), including Layer 2/3 capabilities exchange, protocols, security provisioning and packet processsing offloading from satellites. Several topics important for ETHER are considered, such as: i) deployment of UPFs on-board satellites, ii) support of tunnels between LEO satellite UPFs, iii) multi-hop Inter-Satellite Link (ISL) scenario, or iv) satellite-based routing protocols efficiency [50].

The generic 5GS architectural framework [51] provides several features contributing to the integrated TN-NTN platforms, including:

- A design based on network softwarisation and virtualisation principles, facilitating the inclusion of the 5GS management framework in the NFV MANO stack [52].
- SBA and CP programmability; manageable functions discovery and message brokering enable logical CP partitioning. The MP framework follows the SBA rules and mechanisms.
- UP programmability.
- Network slicing support (tailoring UPF, network slice selection/authentication/admission mechanisms, slice-specific CP functions, separation mechanisms of the SBA CP communication bus, UEs multi-slice attachment).
- Generic network and management data analytics mechanisms and application-specific ones to be implemented thanks to SBA extendibility, which also enables implementation of control loops, including AI/ML-based. The inclusion of the ETSI ZSM framework is also supported [52].
- CP/MP exposure mechanisms for integration with external systems and verticals, including Application Functions (AFs), Common Application Programming Interface Framework (CAPIF) [53], and Service Enabler Architecture Layer for Verticals (SEAL) [54].
- Support for satellite backhaul in scenarios with UPF deployed on-board satellite, where Access and Mobility Management Function (AMF) and Session Management Function (SMF) are capable of reporting the satellite backhaul category that can be used to trigger QoS monitoring by the Policy Control Function (PCF) [51].
- NTN-specific functionalities, including the support for identification and restriction of using NR satellite access, integration of NR satellite access into 5GS, support for discontinuous network coverage for satellite access, high latency, integrated access and backhaul, support for 5G satellite backhaul including edge computing via UPF deployed on satellites. It also includes local switch for UE-to-UE communications via UPF deployed on GEO satellites.

When considering satellite-NR integration, two architectures are typically considered. The first, based on “Transparent payload”, utilises satellites only as relays. The second, called “Regenerative”, assumes the deployment of next-generation Node B (gNB), i.e., 5G BS functions on-board satellites, enabling additional signal processing instead of simple retransmission. This further adds to the required flexibility of satellite nodes. The use of regenerative payload is essential if ISLs are employed. This setup also allows the satellite gateway on the ground to function as a router to the CN. One of the objectives of 3GPP's Release 19 is to provide support for regenerative payloads, with the initial studies and candidate solutions already proposed [55].

3.2 ETHER USE CASE-SPECIFIC REQUIREMENTS

The generic ETHER system architecture has to be capable of supporting the implementation of particular UCs, namely, flexible payload-enabled service provisioning to semantics-aware and delay-tolerant IoT applications (UC1), unified RAN for direct handheld device access at the Ka band (UC2), and support for air-space safety critical operations (UC3). The analysis conducted within task T2.1 and reported in D2.2 [1] resulted in the identification of the following general and UC-specific requirements shown in Table 3-1. The requirement identifiers are in the form *ETH-REQ-UCn-tt-xxx*. Where *n* is 1, 2 or 3 for the UC, and *x* is a sequential integer for each category and UC. In UC2 and UC3 *tt* is FN for functional requirements whereas it is NF for non-functional requirements. In UC1, *tt* refers to the specific aspects; so, FP = Flexible Payload & service orchestration, DT = Delay Tolerant IoT, and SE = Semantics-aware information handling.

Table 3-1: Identified UC-specific requirements

Identifier	Requirement	Description
ETH-REQ-UC1-FP-01	Payload Field Programmable Gate Array (FPGA) resources management	Manage and deploy in a dynamic and autonomous way FPGA resources considering specific context.
ETH-REQ-UC1-FP-02	Payload FPGA resources availability	Control available resources and its percentage of use.
ETH-REQ-UC1-FP-03	Payload FPGA services deployment	Ensure that services are deployed correctly using virtualisation techniques plus containers.
ETH-REQ-UC1-FP-04	Payload FPGA resources sharing	Ensure proper multiplexing for resource sharing, considering interfaces like Universal Asynchronous Receiver/Transmitter (UART), Ethernet, etc., and other hardware resources (memory, buffers, analogue-to-digital converters).
ETH-REQ-UC1-FP-05	Payload system performance metrics	Extract metrics of the system when enabled to monitor different parameters of the system: power consumption, Central Processing Unit (CPU), disk and memory usage.
ETH-REQ-UC1-DT-01	Intermittent – scheduled contacts	Agreement to establish a contact at a particular time.
ETH-REQ-UC1-DT-02	Intermittent – opportunistic contacts	Contacts are not scheduled but present themselves unexpectedly.
ETH-REQ-UC1-DT-03	Intermittent – predicted contacts	Predicted contacts have no fixed schedule, but instead are predictions of likely contact times and durations based on a history of previous observed contacts or some other information (satellite ephemerides) and ML ETHER.
ETH-REQ-UC1-DT-04	Congestion and flow control	Device messages arrive at destination, with support for vertical and horizontal handover, flow congestion and flow control, message retransmissions.

Identifier	Requirement	Description
ETH-REQ-UC1-DT-05	High latency, low data rate	Support for high latency at low data rates for Delay-Tolerant IoT Applications.
ETH-REQ-UC1-DT-06	Connection discontinuity	Support of a Low-Density (LD) LEO constellation with servi-ce link and feeder link discontinuity.
ETH-REQ-UC1-DT-07	Store and Forward	Support for Store and Forward on LD LEO constellation, to solve connection discontinuity, and support of Store and Forward over vertical and horizontal handovers according to ML ETHER.
ETH-REQ-UC1-DT-08	Traffic model mobile autonomous reporting (MAR)	Handle MAR exception reports (notify sporadic events), MAR periodic reporting (regular transmission).
ETH-REQ-UC1-DT-09	Mobility management	Dynamic organisation of tracking areas, or broadcasting of ephemerides to end devices to assist them in using network and power resources efficiently.
ETH-REQ-UC1-DT-10	Support for different services	Support of multi-radio applications based in Narrowband IoT (NB-IoT) using ML in ETHER and orchestrated by the ETHER MANO.
ETH-REQ-UC1-SE-01	Sample processing	Influence the whole information chain from the point we generate the information, encoding, transmitting, and recei-ving. Furthermore, the utilisation of information to achieve a certain goal, for example datasets (or partial datasets) to train an ML algorithm.
ETH-REQ-UC1-SE-02	Joint sample and transmit	Influence the whole information chain from the point we generate the information, encoding, transmitting, and recei-ving. Furthermore, the utilisation of information to achieve a certain goal, for example datasets (or partial datasets) to train an ML algorithm.
ETH-REQ-UC1-SE-03	Support for E2E information handling beyond the sample and transmit	Influence the whole information chain from the point we generate the information, encoding, transmitting, and recei-ving. Furthermore, the utilisation of information to achieve a certain goal, for example datasets (or partial datasets) to train an ML algorithm.
ETH-REQ-UC1-SE-04	Content caching	Need to find the criteria for reusable traffic to effectively cache the freshest and also valuable information proactive-ly.
ETH-REQ-UC2-FN-01	Migrate TN to NTN	Migration of the communication of a user from a terrestrial RAT to a non-terrestrial one and vice versa when it is nee-ded.
ETH-REQ-UC2-FN-02	Vertical handover	The vertical handover process can involve either the AMF of the open air-interface core that will be used (*N2* handover) or the *Xn* interface between the BSs that belong to the different RATs.
ETH-REQ-UC2-FN-03	Waveform	Choice of a suitable waveform in the cases of communication with either stationary or fast-moving platforms.
ETH-REQ-UC2-FN-04	LEO Swarm	LEO satellites work in swarm formations and in a distributed way, transmit the same signal towards a user on the ground.
ETH-REQ-UC2-FN-05	UE beam steering	Designed handheld device antenna enables electronic beam-steering, based on which the trajectory of a moving platform acting as a BS, such as a LEO satellite, can be followed.

Identifier	Requirement	Description
ETH-REQ-UC2-NF-01	Vertical handover	Seamless vertical handover process to the users.
ETH-REQ-UC2-NF-02	Broadband	Broadband communication to the users by either terrestrial or non-terrestrial means.
ETH-REQ-UC2-NF-03	Coverage	Complete user coverage in urban and remote/rural areas.
ETH-REQ-UC3-FN-01	RAN in TN, HAPS, and satellite (SAT)	Enables the 3D network layers to communicate with the aircraft user equipment, with a possibility to link them through a unified RAN framework.
ETH-REQ-UC3-FN-02	Open 5G CN	Required to communicate with the core network resource allocation and user mobility and management gateway functions through specific exposed APIs.
ETH-REQ-UC3-FN-03	Channel emulation	Required for satellite channel, aircraft UE channel, and all the channels of deployed BSs in the 3 strata including terrestrial, HAPSs, and satellite ones.
ETH-REQ-UC3-FN-04	Network resource monitoring	Contributes towards E2E service communications and management with guaranteed QoS, adaptive scheduling for computation offloading and contents caching; in this context resources means information on links, ports, power and spectrum.
ETH-REQ-UC3-FN-05	Multilink functionality	In-flight operations are justified by the required communication performance (RCP) and connectivity resilience related to the impact of safety and the efficiency of Air Traffic Management (ATM) operations; standards based and/or open-source solutions preferred.
ETH-REQ-UC3-FN-06	Network orchestrator	3D network links and edge resources orchestration, employing predictive data analytics associated with traffic monitoring, traffic prioritisation, resource allocation per flight phase and meeting the expected E2E network performance.
ETH-REQ-UC3-FN-07	3D unified SDN management	Manages the relationship between the different controllers running in the 3D network layers. The emphasis vision for the centralised controllers is focused on improving the SDN CP efficiency by minimising the signalling required prior to aircraft communication transmission and improving DP in the different integrated layers.
ETH-REQ-UC3-FN-08	Service performance determinism	Ensures E2E performance guarantee perspective, from end-device to the last application function, which require evolving time-sensitive network (TSN) capabilities and integrating model predictive control (MPC) solutions for dynamic ATM applications.
ETH-REQ-UC3-NF-01	RCP requirements	Where the capacity provided by the ETHER network architecture should be efficiently allocated to support scenarios requesting reliable, resilient, low latency and high data rates connectivity with integrity, considering high aircraft traffic density. The scenarios are intended to meet the RCPs requirements of the different ATM communication services delivered in the different flight phases.

Identifier	Requirement	Description
ETH-REQ-UC3-NF-02	3D RAN low latency	Where the overall service latency depends on the delay at the different 5G RAN radio interfaces deployed in the different 3D network layers, the transmission within the 5G RANs deployed in terrestrial, HAPSs and satellites, the transmission to a server, which may be external to the whole 5G network, and the data processing. Some of these elements depend directly on the 5G RAN itself, whereas for others, the impact can be reduced by suitable interconnections between the 5G RAN and external services by allowing, for example, local services hosting on network edges.
ETH-REQ-UC3-NF-03	Handover reliability and delay	Considers delay resulting in the horizontal handover process during the process of transferring data from one cell to another inside the same access core network (i.e., intra-system handover). Or, vertical handover, which involves handover between different access technologies when they are available, but the objective remains the same: transparently guarantee the session continuity from a final aircraft users' point of view.
ETH-REQ-UC3-NF-04	3D network programmability	Turns the 3D network layers from connectivity platform to service enablement platforms by applying SBA patterns (REST-based HTTP APIs) across all layers.
ETH-REQ-UC3-NF-05	3D network connected intelligence	Contributes to building up a data and connectivity infrastructure supporting cooperation of trusted AI functions from different network layers.
ETH-REQ-UC3-NF-06	3D network resources optimisation	Contributes to optimised network resource provisioning that leverages AI-enabled data analytics and ensures E2E QoS requirements that support highly critical avionics services in an optimal and flexible manner.

Moreover, based on the analysis of other recommendations in the field – 5G Public Private Partnership (5G PPP)/ITU/Next G Alliance, a set of additional general requirements have been identified [1], with two tightly connected to the ETHER architecture design.

Identifier	Requirement	Description
ETH-REQ-GEN-FN-01	Service continuity	Fixed, mobile and satellite converged network shall support service continuity during handover between different access networks.
ETH-REQ-GEN-FN-02	Minimum service	Fixed, mobile and satellite converged networks shall support best-effort QoS for supported services and applications.

While all of the above requirements cannot be fulfilled by the architecture itself, the final system design has to be flexible and capable of incorporating innovative mechanisms that will address the remaining issues.

3.3 CHALLENGES CONCERNING THE TN-NTN INTEGRATION AND ETHER SOLUTIONS

3.3.1 TN-NTN integration challenges

The integration of TN with NTNs primarily aims to bridge a digital division gap caused by a lack of coverage in remote/rural areas in which terrestrial infrastructure cannot be deployed due to physical infeasibility, economic reasons or energy consumption. As the ultimate 6G goal is to provide ubiquitous network services in a sustainable manner, the optimal exploitation of the air and space assets to establish seamless solutions for different scenarios is vital. While significant progress has already been made by the academia and standardisation groups, there is still a plethora of challenges and gaps regarding the integration of TN-NTN networks. The primary ones that will be tackled by the ETHER system are listed below:

[C1] Hardware/software dependency – satellite systems are currently being developed in a custom manner and comprise various dedicated hardware and software conceived by satellite manufacturers based on the mission requirements. This approach largely limits the genericity and flexibility in terms of reconfiguration as well as compatibility issues of services/applications among different satellite systems. To this end, there is a strong need for decoupling hardware and software, e.g., via the adoption of the NFV concept, flexible payload solutions, etc., to enable their optimal usage in a seamless and flexible manner to facilitate NTN evolution.

[C2] Heterogeneity of network assets – as of today, both TNs and NTNs are built with a variety of components and technologies, leading to extreme heterogeneity and increased complexity in terms of management. While the processes of virtualisation and network softwarisation enable the flexible deployment of network components atop a unified resource layer, there remains the issue of non-homogenous M&O across administrative/technological domains (including both application and network orchestration) as well as inter-connection of different network segments (heterogenous transport network). As the network-wide technological convergence is improbable (especially in the NTN case), there is a need for the adoption of a common MANO approach, which allows for generic domain APIs exposure, as well as inclusion of solutions to handle domain-specific problems such as, e.g., infrastructure mobility. Moreover, the E2E system architecture should be capable of supporting multi-vendor and multi-stakeholder deployments.

[C3] Infrastructure mobility – time-variant network topology poses multiple challenges for network organisation and M&O. First, efficient, seamless vertical and horizontal handover mechanisms should be implemented to maintain service continuity. The MANO system has to be capable of handling resources in a flexible manner, including dynamic function placement and migration, topology awareness and prediction to act prior to important mobility events (e.g., losing inter-satellite visibility and ISL). In this context, the SDN approach can play a pivotal role due to its elasticity and logically centralised network view, which allows global network control. Due to the SDN CP scalability problem, there is a need for distributed SDN architecture, which aggravates the problem of optimal placement of SDNCs. Additionally, due to fast topology changes (e.g., in fast-moving platforms, LEO, MEO), the impact of CP latencies between the SDNC and switches should be considered to enable efficient operation with minimal management overhead. This is especially important in the context of path reconfigurations needed for seamless handovers.

[C4] Delay variation and Doppler shift – the high speed of space/airborne RAN nodes relative to the Earth (LEO, MEO) and variable distance between the UE and serving node introduces variant delay and Doppler shift. The commonly used 5G Orthogonal Frequency Division Multiplexing (OFDM) waveform suffers high performance degradation in such conditions. While the newly introduced modulation methods, such as Orthogonal Time-Frequency Space (OTFS), offer improved performance [56], the design of the unified waveform optimal for both TN and NTN communication remains a challenge.

[C5] Direct handheld access for terrestrial UEs – while direct handheld device access from NTN platforms is feasible, there exists a significant gap between the performance achieved by TNs in millimetre-wave (mmWave) bands and satellite ones (congested spectrum in S-band). While there is the provisioning of broadband services through mmWave transmissions from satellites to very small aperture terminals, such provisioning is very challenging for handheld devices due to their very low antenna gain. This has dictated the need to consider ways to increase the satellite antenna gain without resorting to bulky monolithic structures that would skyrocket the launching costs. An innovative approach is to consider collaborative transmission from satellite swarms so as to increase the antenna gain from space by creating a large virtual array, particularly from LEO satellites.

[C6] 3GPP system deficiencies – 3GPP identified several issues and gaps in the current 3GPP system definition related to mobility management for different coverage areas (large and moving), delay in satellite, QoS enforcement for both satellite access and backhaul, RAN mobility, multi-connectivity (satellite backhaul and hybrid satellite/terrestrial backhaul), content distribution close to the edge, and regulatory services for satellites overlapping multiple countries [57]. Moreover, the impact of latency (due to operation altitude), satellite/HAPS cell coverage (larger than usual cell, differential delay, possible multi-country coverage), different propagation channel models than included in the link power budget considerations or radio unit performance associated with satellite/HAPS payload performance, remains to be addressed. Additionally, 3GPP recognises the need for special treatment of NTN regarding mobility management procedures (e.g., AMF (re-)selection by gNB to ensure that UE connects to AMF serving the country in which UE is located) or contextual information (providing gNB with the regular/on-demand ephemeris information or NTN Gateways location) [58].

[C7] E2E latency of integrated network – as the links between TN and NTN nodes are characterised by typically much larger distances, which translates to increased latency, efficient mechanisms and schemes for interaction with NTNs are necessary to ensure as few transmissions as possible, which further can translate to decreased E2E latency.

[C8] Lack of efficient handover control mechanisms for integrated TN-NTN – both vertical and horizontal handovers in the integrated TN-NTN have higher complexity compared to handovers within typically static TNs. New handover mechanisms need to be developed to account for the dynamic characteristics of NTN nodes of integrated network topology, which can cause issues such as unsuccessful or too frequent handovers.

[C9] E2E MANO of integrated networks – to efficiently manage the integrated network's resources, a MANO architecture needs to be developed, which allows to interconnect network functions deployed in different vertical domains into service chains.

[C10] Optimisation of E2E network performance – as specific services can have different QoS requirements, the MANO system should employ optimisation algorithms, which will improve the overall network performance.

3.3.2 ETHER solutions for the TN-NTN integration

To address the aforementioned challenges (C1-C10) associated with TN-NTN integration, ETHER introduces eight main technical innovations (T-1 – T-8) extending the current SotA. The integrated architecture (T-1), being the foremost of these innovations, serves as the foundation for deploying the subsequent innovations. The final ETHER reference architecture is detailed in Section 4 of this deliverable.

3.3.2.1 Direct handheld device access at the Ka band from LEO satellites distributed LEO swarms (T-2)

Towards the seamless integration of terrestrial with NTNs and achieving broadband direct communication to handheld devices, either from terrestrial or LEO satellite-based gNBs, ETHER radically considers the deploying swarm formations of LEO satellites. These swarms can cooperatively and coherently transmit signals to ground terminal or receive from them. An illustrative scenario, depicted in Figure 3-1, shows a ground user with a handheld device is initially connected to a ground gNB receiving broadband communication in the Ka band. As the user moves, the connection might deteriorate due to obstacles, such as a building, or due to increased distance from the gNB. Additionally, traffic load of the gNB may necessitate migrating the user to another gNB. In such scenarios, to ensure uninterrupted connectivity, the communication with the ground user can be handed over to LEO satellites. If a LEO satellite is almost vertical to the user, under ideal weather conditions, a single LEO satellite might suffice for broadband downlink communication with the ground UE, as initial studies have shown [59]. However, in adverse weather conditions or when satellites are at low elevation angles relative to the users, small LEO satellites flying in swarm formations will collaboratively act as virtual arrays by transmitting the same signal to the ground user, enabling coherent combining at the receiver. Hence, the ETHER architecture envisions that LEO satellites equipped with regenerative payloads and flying in formation will be orchestrated as needed to form virtual arrays and collaborative transmission to or reception from a ground handheld device [60]. This architecture requires fast ISLs for dealing with common impairments associated with distributive collaboration, such as time, carrier frequency, and phase desynchronisation. The latter can arise from minor perturbation in satellite positions, potentially disturbing the beamforming/combining operations if not adequately tracked. This innovation is crucial for the ETHER UC1, which requires broadband communication from space to handheld devices, and targets the C5 challenge.

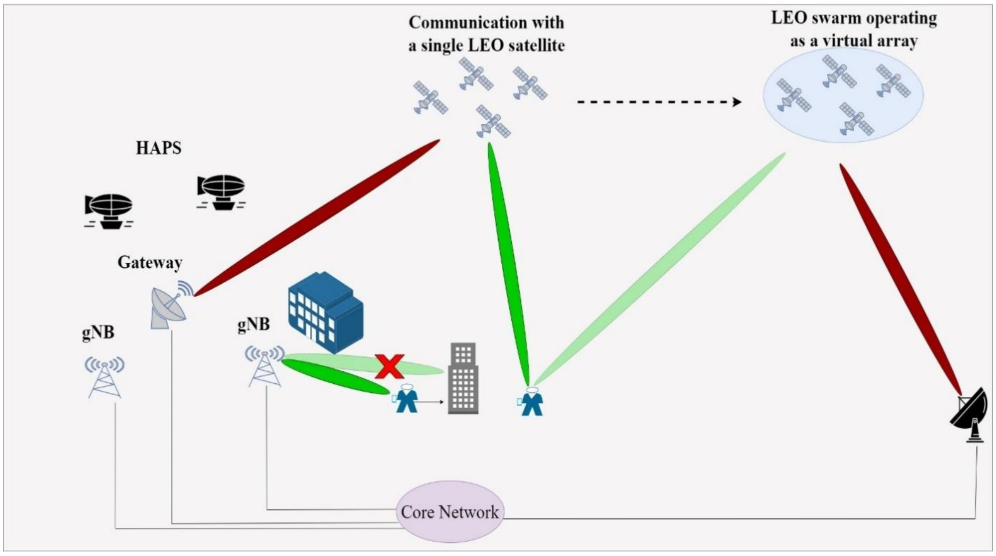

Figure 3-1: Distributed transmission from a swarm of LEO satellites

3.3.2.2 Unified waveform design (T-3)

The commonly used standardised 5G OFDM waveform can induce high performance degradation since it is very sensitive to Doppler frequency shifts that can arise from the communication of ground terminals with the fast-moving LEO and MEO satellites. This is why the newly introduced OTFS modulation has been brought forward as a waveform that is insensitive to such shifts and suitable for communication with both stationary and fast-moving platforms, such as LEO satellites [56]. Hence, the first step in ETHER would be to identify potential scenarios incorporating communication with fast-moving flying nodes, such as LEO satellites where OTFS is much more beneficial than OFDM. As 5G NR already supports mechanism of waveform switching, the key limiting factor for OTFS usage is mainly related to the selection of pilot sequences and estimation of Doppler shift, which requires further research. The particular innovation is of importance for all the ETHER UCs since in all of them communication with LEO satellites is envisioned. It will be showcased though in the context of the ETHER UC2 demonstration activity. This innovation targets the C4 challenge.

3.3.2.3 Flexible payloads (T-4)

The adoption of Software-Defined Radio (SDR) platforms enables the definition of satellite payloads that can flexibly adapt to operators' services. The flexible payload concept is based on three primary mechanisms, each representing a different level of flexibility, studied in the project:

- Flexibility level 1: Involves hardware reconfiguration (either full or partial) to add extra resources to the system for current services and adaptability for future services.
- Flexibility level 2: Focuses on software reprogramming and Operating System (OS) abstraction by applying virtualisation techniques. This level enhances flexibility and allows resources management from a software perspective.
- Flexibility level 3: Pertains to service-based deployment from outside the satellite, facilitated by an orchestrator and leveraging the capabilities offered by Levels 1 and 2.

A general diagram of the envisioned flexible payload concept is depicted in Figure 3-2.

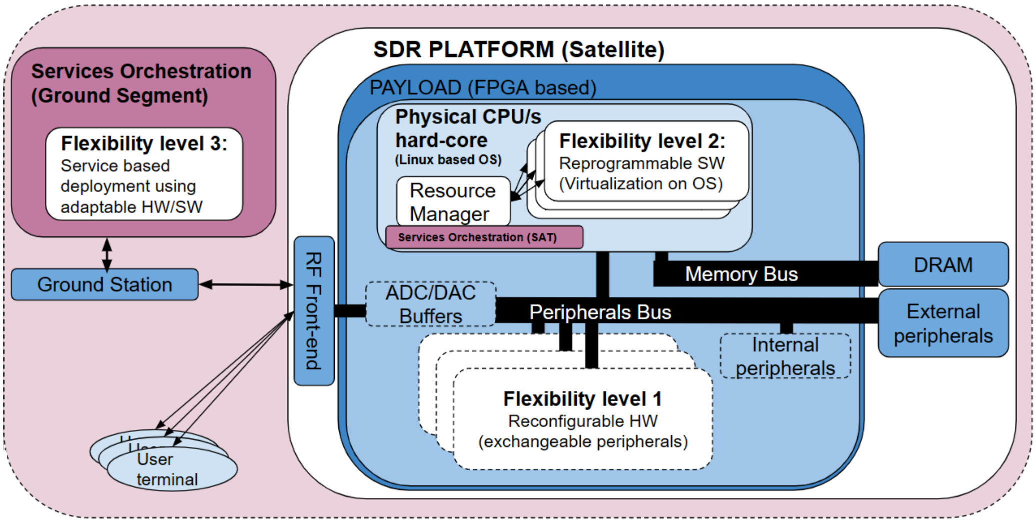

Figure 3-2: Representation of the virtualisation levels achieved in the flexible payload design

Although flexible payloads play a crucial role in all the three ETHER UCs, their functionality will be showcased in the ETHER UC1 demonstration activity. This innovation targets C1 and C2 challenges.

3.3.2.4 Data analytics, edge computing, and caching, including semantics-aware data analytics and control (T-5)

ETHER will utilise the semantics of information in NTNs combined with edge computing and caching to increase the efficiency further and reduce the E2E latency without affecting the amount of conveyed information. In particular, semantics refer to the importance and relevance rather than the meaning of information. More specifically, by leveraging semantics in NTNs, ETHER will allow to generate and transmit only a small fraction of data without affecting the conveyed information.

This can be achieved by defining metrics that capture the innate and contextual attributes of information and could include the cost of actuation error when we target remote control and actuation of devices as well as the age, version, and value of information that can capture the timeliness and the importance of information in status updating systems. Furthermore, adapting those metrics to capture the unique characteristics of NTNs is crucial. Thereby, the development of caching schemes and cooperative computing techniques will be enabled that will outperform the current schemes by achieving lower latency and higher energy efficiency. For example, incorporating the version age of information in caching schemes can allow for less frequent updating, leading to less power required for the operation and transmissions and, thus, higher energy efficiency. By reducing the transmissions, we reduce the delays inside the network, thus further reducing the E2E latency.

These metrics will be further utilised in data analytics schemes where the current approaches neglect the importance and relevance of information, thus making them cumbersome to be applied directly in NTNs.

The semantics-handling of information and caching will be showcased in the ETHER UC1, whereas data analytics and edge computing will be demonstrated within the ETHER UC3 activities. The information semantics innovation targets challenges C3, C7, and C10.

3.3.2.5 Horizontal/vertical handovers including AI-based handover control mechanisms (T-6)

In such a heterogeneous 3D architecture, efficient horizontal/vertical handover mechanisms should be developed to achieve a good trade-off between complexity and performance. ETHER advances common handover implementations, which leverage mainly signal quality metrics obtained by the UE, considering additional parameters, such as time availability and network sustainability. For example, handover requests within satellites may need to go through the Store and Forward mechanism and be processed only when a connection with the ground segment is available.

ETHER will consider more elaborate criteria for performing horizontal and vertical handovers in hybrid TN-NTN platforms by leveraging AI-based algorithms to make the process autonomous. This allows to deal with several parameters that need to be considered when migrating to another gNB. For example, even if *Xn* interfaces among the terrestrial, aerial, and space gNBs are always available, the *Xn*-type handover (direct handover leveraging the *Xn* logical interfaces among the gNBs without the need to go through AMF of the core network) might not be always the best solution. This is due to the serving gNBs that might have limited context information about how the network will evolve over time in contrast to gathered information in the core network, which will affect the decision for handovers. If, for example, a LEO satellite can provide better quality over a serving terrestrial one for a particular service, but the link is available for a short amount of time due to its fast movement, such a handover should be avoided.

Another important parameter is the future traffic that the target gNBs need to facilitate, together with their energy consumption demands arising from the user migrations. The latter is especially important for grid-less platforms that can act as target gNBs, such as LEO satellites and HAPS. Owing to the aforementioned considerations, ETHER will design novel ML-based algorithms for horizontal/vertical handovers in the integrated network that will balance a variety of criteria related to the evolvement of the network, such as latency, rate, traffic, while targeting 70% higher energy efficiency than the SotA work [58]. Towards this, federated learning will also be leveraged for proactive handovers due to avoiding the need for exchanging a large amount of data among the nodes, which is beneficial in terms of latency and energy consumption. Proposed horizontal/vertical handover strategy targets challenges C3, C6, and C8.

Horizontal handovers will be showcased in the ETHER UC1, whereas vertical ones will be demonstrated in the ETHER UC2 demonstration activities.

3.3.2.6 Automated MANO for the integrated network (T-7)

Satellites mobility management

One of the main challenges of satellite communication networks includes managing the dynamic links between the ground station and SAT, which are crucial for transmitting application images and managing tasks but are limited by satellite mobility and intermittent connectivity. Additionally, resource management within satellites is a complex task due to fixed, predefined hardware capabilities and the inability for satellites to dynamically compile or adjust applications, constraining operational flexibility. Coordination via the GS-SAT link allows for some remote management of satellite configurations; however, issues such as failure reporting and self-healing capabilities are limited until the next GS contact.

To tackle the abovementioned challenges, In the ETHER project, the satellite communications architecture integrates into the NFVI within the ETSI framework, as shown in Figure 3-3. This integration enhances the efficiency and reliability by employing a distributed, multi-node cluster approach managed by a GS, where SATs function as cluster nodes with distributed master systems to enhance availability and resilience, using ISLs for improved connectivity and latency management. Furthermore, resource management is decentralised, allowing efficient resource utilisation across the cluster, with a federation system that keeps masters restricted to one cluster, enabling robust routing capabilities. The system is highly autonomous, capable of self-healing and maintaining continuous service during node failures or disconnections from the GS. Finally, the considered approach utilises a VIM to manage satellite mobility and network connectivity, where the containerisation is facilitated by Kubernetes acting as the VIM.

Kubernetes is chosen over other orchestration tools, such as Docker and Openshift, for its robust scalability, extensive community support, and proven ability to manage complex container tasks. Its ability to handle complex container management tasks, such as automatic deployment, scaling, and management of containerised applications, makes it ideally suited for the dynamic and distributed nature of integrated TN-NTN systems. Kubernetes orchestrates containers efficiently across satellites, utilising features such as ReplicaSets for stability and autoscaling to adjust services based on satellite positions. This advanced setup ensures uninterrupted service and aligns with 5G technology, making ETHER a resilient and dynamic infrastructure solution for space networks.

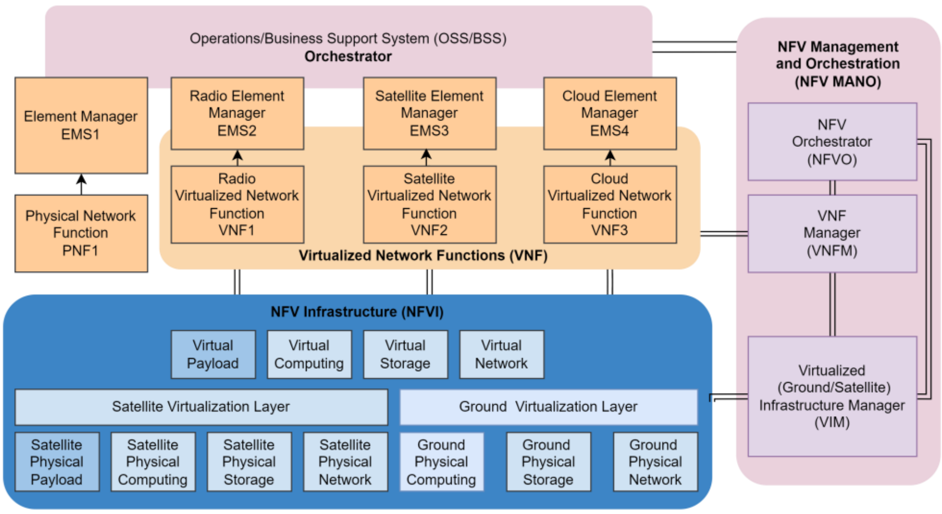

Figure 3-3: Architecture representation based on ETSI architecture with satellite systems integration

ETHER WAN Infrastructure Manager

SDN has revolutionised network management by separating the CP from the forwarding plane, facilitating centralised, programmable control that enhances dynamic network management and reduces the need for physical access (cf. Figure 3-4). This transformation supports innovative NFV-based UCs, improving network agility and innovation through automation and API integration. Building on SDN's role, the importance of the northbound interface (NBI) and southbound interface (SBI) is critical. NBI allows SDN applications to program the network and request services, while SBI enables the controller to manage network devices, supporting functionalities such as automated network provisioning and dynamic bandwidth allocation. Additionally, the dynamicity, flexibility and efficient network resources management provided by SDN is crucial in transport networks (also known as WANs), which comprise the physical infrastructure including routers, switches, and both wired and wireless connections. These networks, particularly in integrated TN-NTN contexts, serve as the backbone interconnecting ground and space segments, ensuring seamless communication. Applying SDN to WAN is a well-studied technology called Software-Defined Wide Area Network (SD-WAN), where several commercial vendors have proposed programable solutions to handle the connectivity between domains [61].

On the other hand, the integration of SDN with NFV in multi-domain environments has led to the introduction of the concept of WIM, which has been defined by ETSI [62] (cf. Figure 3-4). WIM manages the E2E orchestration of resources and services (VNFs) across multiple domains and technologies. The SDN-based WIM provides a unified interface and CP for network operators to configure, monitor, and manage the transport network elements. Notably, WIM, by utilising SDN and NFV, can efficiently allocate transport resources by leveraging orchestrated VNFs. This capability allows network operators to dynamically connect VNFs with the necessary transport network resources. As a result, this integration fosters enhanced service agility, boosts resource utilisation, and leads to optimised network performance.

Recent studies have explored placing SDNCs on non-terrestrial platforms, assessing various metrics like communication latency, controller coverage, potential link and node failures, and deployment costs [63]. Placement in hybrid 3D networks, involving GEO, MEO, and LEO satellites in addition to terrestrial locations, presents challenges due to the diverse and mobile nature of potential host nodes (cf. [63], Table 4). This strategy exploits satellite advantages like wide coverage and low latency, optimising SDNC numbers and placement. There are two main placement strategies [64]: i) a static approach, which is cost-efficient and based on offline metrics, and ii) a dynamic approach that adjusts controller deployment across network layers in real-time based on data demands, enhancing performance at a higher cost and complexity.

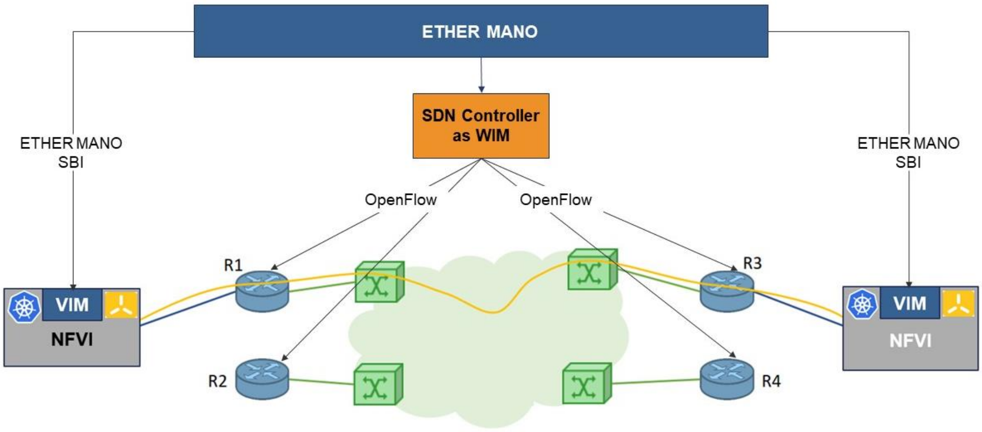

Figure 3-4: SDN as WIM controller in transport networks

In the context of ETHER, a multi-layer SDNC setup is possibility assumed. In particular, SDNCs are envisioned to be placed in terrestrial, aerial, and space layers. In addition, in a hierarchical approach, a chief controller in each layer controls the corresponding slave controllers, and a master controller on the ground integrates the data across these layers. ETHER aims to investigate the optimal SDNC placement under both static and dynamic approaches, based on various service-related metrics. Incorporating the aerial stratum as an additional location for SDNC placement introduces greater complexity to the placement problems. This complexity underscores the necessity for developing low-complexity algorithmic solutions, especially for dynamic placement strategies. Furthermore, it is crucial to jointly optimise the placement of SDNCs along with the design of the aerial and satellite infrastructure, rather than merely adjusting SDNC positions to fit pre-existing infrastructures. Optimising SDNCs independently of infrastructure design would yield suboptimal results compared to a holistic approach that considers both elements simultaneously.

MEC and application orchestration

Edge computing, combined with satellite communications, enables faster, more reliable application services by reducing latency and enhancing data security. According to the ETHER proposal, satellites integrated with edge resources function as MEC devices. This integration requires not just network orchestration but also sophisticated application orchestration for end-users. To this end, ETHER proposes a zero-touch orchestration framework for the LCM of application at all layers of the edge infrastructure by utilising AI/optimisation algorithms. As shown in Figure 3-5, the key components of this framework include AI-driven analytics for application placement decisions and geographically distributed MEC Platforms (MEPs) that support Kubernetes clusters, ensuring adherence to targeted QoS and QoE. These functionalities will be demonstrated in ETHER UC1, showcasing the capabilities of ETHER MANO.

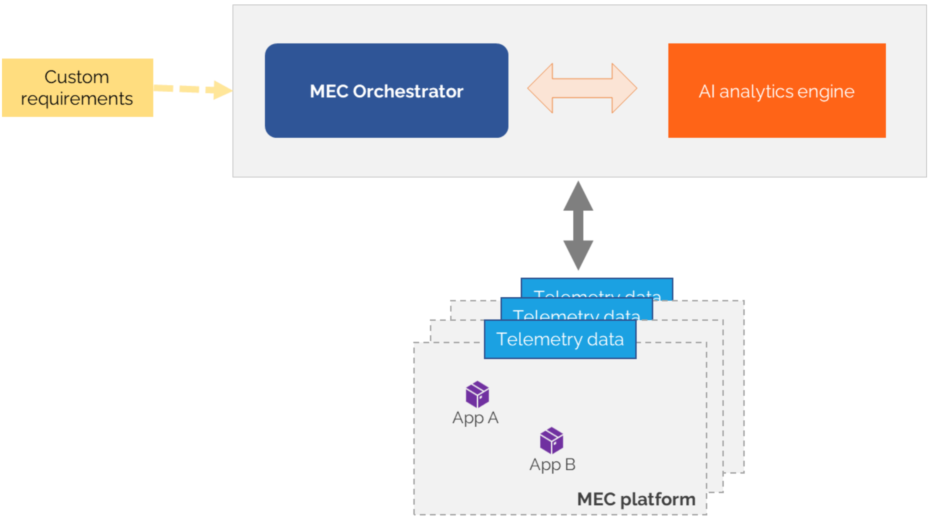

Figure 3-5: MEC and application orchestration

The **applications** are the end-user orchestrated MEC applications. The functionality of the ETHER MANO will be showcased in the demonstration activity of ETHER UC1. Considering the envisioned capabilities, the ETHER MANO innovation targets challenges C3 and C9.

3.3.2.7 AI-driven E2E network performance optimisation (T-8)

From an E2E viewpoint, the mechanisms for the management of the RAN and the solutions for zero-touch orchestration of the resources (across the virtualised network and cloud continuum) must be integrated. Such an integration is paramount to enable a comprehensive and effective coordination of network slices, from the RAN to the core segments, for different kinds of 6G services and accounting for heterogeneous terrestrial, aerial and satellite radio access domains.

The E2E MANO framework builds on the architectural specifications already presented above for RAN and core and does not introduce further requirements from an architectural viewpoint. Yet, it introduces a need for original mechanisms capable of meeting the following targets:

- Real-time network monitoring and Key Performance Indicators (KPIs) prediction.
- Online (for run-time self-configuration) and offline (for long-timescale self-configuration) E2E cross-layer optimisation procedures that build on the monitoring and prediction tools above. These mechanisms shall ensure that:
  - The QoS/QoE requirements of each service are met in a transparent way with respect to the radio access domain the user equipment is in (e.g., by prioritising core network traffic to compensate for different radio access latency and meet E2E delay targets);
  - High energy and cost efficiency are achieved (e.g., by solving a joint network, computational and storage resource allocation problem based on the deployment scenario).

Given the complexity of the tasks above, it is expected that data-driven models based on SotA AI paradigms, tailored to the ETHER context, need to be developed to meet the requirements effectively.

The AI-driven E2E network performance optimisation will be showcased in the demonstration activity of the ETHER UC3 and targets C10 challenge.

4 FINAL ETHER REFERENCE ARCHITECTURE

4.1 ARCHITECTURE DESIGN PRINCIPLES

The ETHER architecture is expected to provide the integrated TN-NTN ecosystem capable of accommodating cross-layer 5G and beyond services. To this end and to provide a compatible solution in the long-term perspective, it is essential to consider the main 6G visions, trends and paradigms in the overall architecture design (as described in Section 3.1). The ETHER architecture will also need to support and efficiently integrate the technical innovations (cf. Section 3.3.2), some of which are built on the foundation of SotA solutions. Moreover, the planned UCs can also impose specific requirements on the architecture shape, e.g., in terms of openness, genericity or exposure capabilities to integrate with frameworks external to ETHER. Finally, the business ecosystem characteristics need to be reflected as well to enable the creation of new business opportunities in the integrated TN-NTN ecosystem. The typical modern business environment is a multi-stakeholder ecosystem, resulting in the following general requirements to be considered in the ETHER architecture design: multiple administrative domains (both in terms of distributed ownership of underlying resources or potential unbundling of resources ownership as well as the role of resource management, including, for example, the role of a consolidated resource manager or infrastructure broker), heterogeneous domain-specific technologies (commodity or specific hardware), etc. The undertaken approach to architecture shape will focus on opening new business opportunities and creating new business roles to maximise potential benefits from the TN-NTN integration (which will be further analysed in WP6).

The following principal features have been taken into account when designing the ETHER architecture:

- Generic – ability to accommodate different frameworks, e.g., ETSI NFV/MEC, SDN, 3GPP 5GS, O-RAN and being capable of supporting the network evolution.
- Multi-domain – both in terms of 3D access (terrestrial, aerial, or space stratum) and in terms of functional split (RAN, edge, core, and transport).
- Multi-provider – ability to accommodate different providers' solutions, provided that they follow standardised mechanisms or can be adapted to them.
- Hierarchical – following a hierarchical approach to the scope of responsibility or concern, in particular, the communication service-network-subnetwork-function-resource hierarchy, to functional split, M&O.
- Distributed – ensuring the placement of processing functions as close as possible to data generation points, avoiding unnecessary data transfer back and forth, etc.
- Zero-touch (automated, AI-based network/service management & orchestration, also capable of supporting "intent-based" operation) – minimising the necessity of manual interactions in operations and management.
- Scalable – able to expand with a service/traffic demand growth.

- Service-based – focused on discrete functionality units instead of a monolithic design, where the service producer/consumer relation is followed, applying the publish-subscribe, request-response mechanisms, and REST API-based approaches.

- Modular (associated with the preceding two) – based on the functional split into modules to support their independent creation, testing, parametrisation, modification, replacement, exchange, interconnection, logical separation, etc.

- Openness to future extensions and trends (e.g., micro-service mesh).

4.2 OVERALL ARCHITECTURE AND ARCHITECTURAL LAYERS

The high-level ETHER architecture conceptualisation is presented in Figure 4-1. The ETHER system is immersed in a shared infrastructure that includes both physical and virtual infrastructure resources. The 3D ETHER system aspect is represented by terrestrial, aerial, and satellite strata through which softwarised Application, Communication, and Steering Planes are laid out. While the Application and Communication Planes directly and visibly serve the user, the Steering Plane ensures the seamless operation of these former using control mechanisms. The continuity of the mutually cooperating planes through all the strata is ensured by resources (especially transport ones) and mechanisms of interconnection and mediation. Two additional parallel planes encompass orthogonally three of the abovementioned planes as well as the infrastructure expanse: i) the Management and Orchestration Plane, and ii) the AI Plane. The role of the former is to provide management of the ETHER system and orchestration of resources and other planes. The latter supports not only the former for automation and autonomy but also exposes its services to applications, communication, and steering mechanisms. This is how the idea of embedding AI as a native element of the ETHER system is expressed. The Application Plane (AP) is an integral part – not external as it used to be before – of the ETHER system using common resources and mechanisms of steering, management, and orchestration powered by common AI. In this manner, the maximum flattening of the system is achieved to augment its performance efficiency.

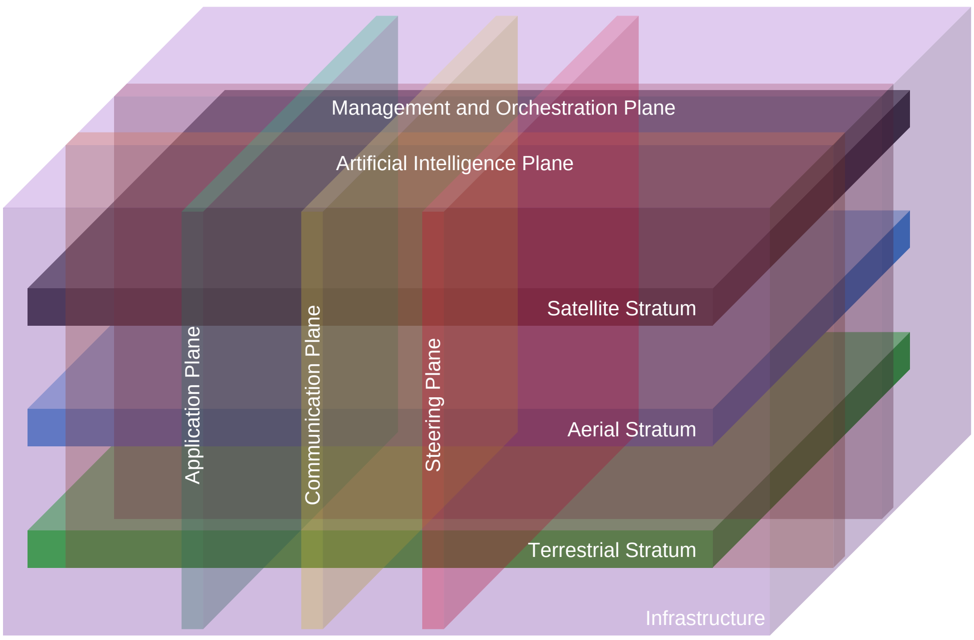

Figure 4-1: High-level ETHER architecture conceptualisation

The overall ETHER system architecture, based on the presented conceptualisation is shown in Figure 4-2.

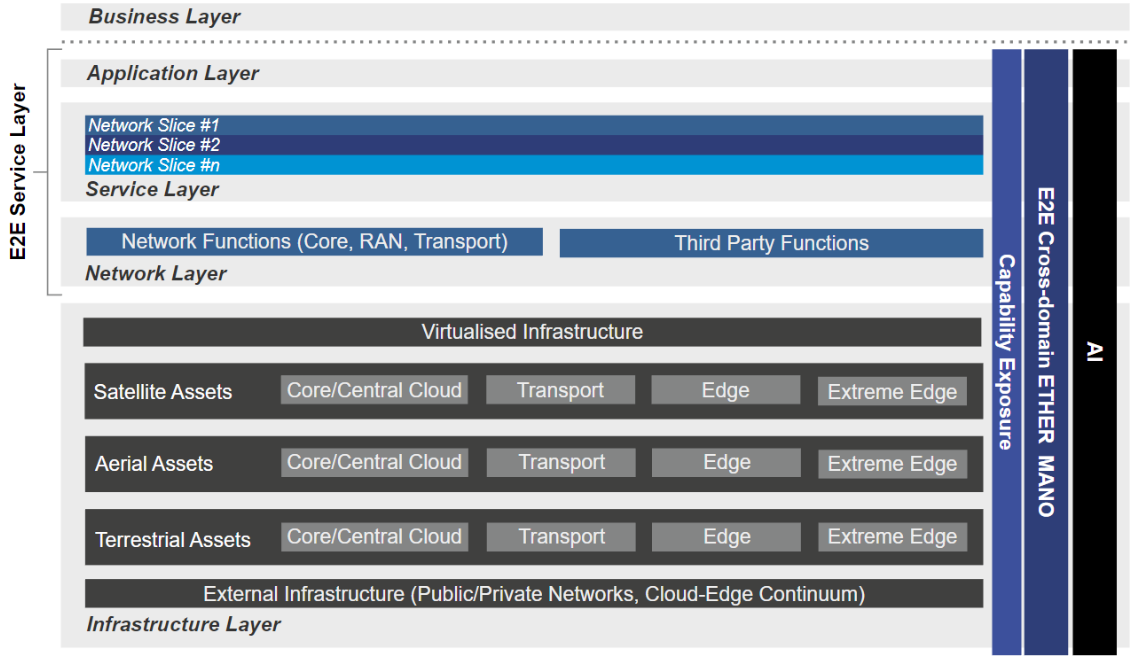

Figure 4-2: Overall ETHER system architecture

The overall ETHER system architecture is composed of:

- *Infrastructure Layer* – which includes the TN and NTN assets (both satellite and aerial components) serving as Core/Central Cloud, Transport, Edge and Extreme Edge infrastructure, external infrastructure such as Public/Private Networks, or Cloud-Edge Continuum resources offered by the providers external to the ETHER system, and Virtualised Infrastructure (obtained by virtualisation of the above physical assets or provided by external entities). This layer can be seen as an NFVI layer of ETSI NFV MANO extended by non-virtualised resources.

- *Network Layer* – composed of NFs (e.g., 3GPP CN, RAN, transport-related functions) and third-party functions understood as any functions external to the ETHER scope but supporting the network operation. This layer can be seen as a VNF layer of ETSI with dedicated Physical Network Function (PNF) support functions.

- *Service Layer* – constituted by the Network Slice Instances (NSIs) composed of NFs residing in the Network Layer with the relevant slice management and mechanisms for exposure to the upper layers.

- *Application Layer* – containing applications that exploit the capabilities provided by the NS (or several NSs).

- *Business Layer* – constituted by the relevant business actors – Mobile Network Operators (MNOs), verticals – interacting with the MANO framework. The business model aspects are out of the scope of this document and will be investigated by WP6.

The above-mentioned layers are coordinated on multiple levels by the E2E cross-domain ETHER MANO solution (handling the M&O processes of the ETHER system), leveraged by the capabilities provided by the internal AI applications (e.g., AI-based M&O enablers) or external to the ETHER system, e.g., exposed to the framework via AIaaS mechanisms. The following sections will provide an overview of each of the ETHER system layers and their internal components. The intra- and inter-layer interactions are described in Section 4.7.

4.3 INFRASTRUCTURE LAYER

ETHER envisions building a 6G integrated 3D multi-layered network infrastructure comprised of NTN assets represented by space, HAPSs, aerial, and TN. As a result, a 3D multi-layered horizontally and vertically integrated TN-NTN network has emerged de facto as a global network to provide communication services to users active in the different integrated 3D network layers. The users include flying objects like aircraft, electric vehicles, UAVs and space IoT devices attached to low- or high-altitude platforms and terrestrial objects like ground users and IoT devices. The designed 3D network provides global broadband coverage to remote and rural areas that cannot be covered by the existing terrestrial communication network infrastructure. This refers to the deployment cost and high energy consumption of the different terrestrial infrastructure components.

ETHER 3D network infrastructure deployed in the three strata (cf. Figure 4-1) is described as follows:

Terrestrial stratum involves a multitude of heterogeneous radio, wireless, cellular, ground satellite, and small cells operating at different frequency bands, primarily sub-6 GHz and mmWave. Until 5G, the antennas of these BS nodes have been downwards tilted to serve the ground users. However, the ETHER project follows the changes considered in the 6G network concept by either equipping them with a large number of antennas to enable 3D beamforming or with uptilted antennas. This would allow, in turn, to better interconnect (integrate) terrestrial BSs or users with other aerial and space users, including flying BSs installed on-board satellite spaceship, aircraft, HAPS or UAVs. The availability of such access networks in the different strata provides a resilient communication network to aircraft, which supports their critical-time operations and safety. It also enables the flying objects (e.g., aircraft, UAVs) to select the air-to-ground links according to their communication performance requirements. Obviously, network scenarios such as traffic or user load balancing or traffic re-routing can be realised to better serve aerial and space users. The TN infrastructure also includes different gateways that connect the aerial and space platforms to the terrestrial core network through feeder links.

Aerial stratum comprises platforms such as commercial airplanes (aircraft), UAVs, and HAPSs that fly at altitudes of 8-11 km, up to 1 km, and 17-20 km, respectively. In 6G networks, the aerial platforms are envisioned to be boarded with 5G gNBs. These turn them into flying BSs that can serve users within the same network stratum, as well as other users within the strata above and below theirs. The advantage of operating gNBs on-board HAPS stems from their much lower altitude relative to the ground than the satellites installed in the different orbits. Also, the quasi-stationary nature of HAPSs eliminates the need for frequent handovers, like the case with the non-Geostationary Orbits (NGSOs) when they are about to lose visibility with ground terminals. Furthermore, their altitude and size allow them to be equipped with sizeable antennas that can offer high gains. This could enable direct handheld device access even at mmWaves (e.g., Ka band). Moreover, as their service coverage gets wider, relying on their much larger footprint compared to the terrestrial gNBs, it would better prevent the frequent handovers of high-mobility users such as aircraft or UAVs, in contrast with fast-moving ground network mobility, such as trains, where terrestrial cells may fail in providing ubiquitous services. However, what differentiates this stratum more is that the gNBs on-board the platforms deployed within this stratum can act as a relay node enabling the expansion of the visibility of NGSOs to better serve ground users and those in the other network strata. The UAVs have been used for different purposes, including military, cargo, and rescue operations. They can also act as aerial gNBs in the case of natural disasters and high-traffic demand events like sport events. Airplanes (aircraft) require resilient and reliable networks not only to support critical operations or safety, but also to provide continuous broadband connections for passengers on board. This has been challenging and costly with TN networks and can only be expected during the take-off and landing times. However, when the airplanes spend the majority of their flight time at high altitudes, the ETHER integrated 3D network can be deployed to support the required broadband coverage, which can be provided by means of satellites like LEO, NGSO, and HAPS.

Space stratum is the upper stratum in the 6G ETHER 3D network architecture. It comprises satellites operating at different altitudes in orbit around the Earth. Satellite communications have been around for sixty years; it has become integrated and is increasingly unified as part of the NTN within the 3GPP framework.

Satellite trade-offs

When designing the satellite to provide satellite communications services (satcom), there are, of course, many trade-offs that need to be considered. Factors such as the launch mass, the amount of power that can be generated and the excess heat that can be radiated, and the availability of electronics suitable for the increased levels of radiation above the atmosphere all need to be considered. A related trade-off is for the satellite to be either bent-pipe or regenerative. Bent-pipe satellites have been the predominant GEO satellites for many years as most of the available power budget is used to efficiently receive and retransmit the same radio signal; whereas regenerative satellites demodulate the signal, may process the data in some way and then retransmit it. This does improve the E2E link budget as there is no additive noise from the uplink and downlink transmissions however it also signifies that less power is available for transmission. Two benefits of regenerative satellites supporting NTN are that they facilitate ISLs and allow some aspects of the 5G CN to be instantiated in space thereby improving CP response times. Regenerative payloads are directly connected with the notion of flexible payloads, described subsequently in the ETHER MANO Layer. Flexible Payload is based on the principle of programming the logical resources of hardware boards (e.g., FPGAs) and applies software virtualisation mechanisms on the base OS. Thus, through Flexible Payloads equipping NTN nodes, such as satellites and HAPS, resources such as power, bandwidth, and beams, can be adjusted through software on the ground (SDR-based) in accordance with the application demands. The capabilities of each of the corresponding NTN nodes, such as size and generated power, will dictate the FPGA capabilities and the corresponding Flexible Payload functionalities that can be realised.

Characterisation by orbit

Satellites are usually characterised by orbit type as follows:

- **GEO** – geostationary orbit where the satellites orbit above the equator so that they appear stationary in the sky from the ground; only a few satellites are needed to provide near global coverage (only the polar regions are excluded), but the propagation delay is too high for some applications and the link budget tends to require higher gain antennas for the user equipment;
- **LEO** – low earth orbit satellites are much closer to the ground, which improves the link budget performance and reduces the propagation delays at the cost of the satellite moving rapidly across the sky when viewed from the ground and needing many satellites in the constellation to provide continuous coverage;
- **MEO** – medium earth orbit satellites can be found between LEO and GEO; they provide global coverage with far fewer satellites than LEO with an intermediate propagation delay and link budget performance;

- **HEO** – highly elliptical earth orbit satellites are a specialist application where a few satellites are launched into the same orbit that drops low (and fast at perigee) before raising to a much higher (and slower at apogee) part of the ellipse. If these orbits are also synchronised to the Earth's rotation, they can provide GEO-like performance in the polar regions when the satellites move relatively slowly across the sky.

As a result of all these trade-offs, different orbits have different UCs (cf. Appendix 1). ETHER focuses on direct-to-device UCs via satellites in a LEO constellation.

ETHER 3D edge computing and storage continuum is another capability that would majorly distinguish the designed 3D network in the sense it would become more responsive to service demands type and volume and match it with the requested infrastructure availability in an efficient manner relying on real-time predictive data analytics. Noting that a key difference between the 3GPP's Release 19 and earlier releases regarding the integration of TNs with NTNs is that Release 19 will include the regenerative payload feature for NTN platforms. This means that the nodes will be able to encode and decode the received information and act as aerial and space gNBs. This, in turn, will enable them to act as edge computing and storage units, which can heavily alleviate the large amount of data that the cloud TNs need to process and store. Hence, the ETHER 3D network is a TN-NTN distributed edge computing and storage network that allows such tasks to be performed in terrestrial, aerial, and space nodes. This can be possible by the existence of inter-space-air-terrestrial links for fast data routing among nodes. The federation of the multi-edge computing deployed in the different 3D network strata will be of paramount importance to support the full single pilot operations. It will also allow the running of machine learning and artificial intelligence algorithms, training, and testing on the fly to better ensure aircraft operations and navigation.

The above infrastructure contains diverse resources, including radio channels, wireless, cellular and satellite air interface and spectrum, 5G core network functions, as well as added memory, storage, processing, and communication payload to HAPSs and satellites, in addition to the content cache and edge computing distributed in the different network strata. These diverse resources provide opportunities for creating and deploying massive virtualisation assets to support 3D network slices that can meet the emerging requirements of E2E space, aerial and ground applications. It helps space, aerial and ground network operators to immediately open their physical network infrastructure platforms to the concurrent deployment of multiple logical self-contained networks, virtualised and orchestrated according to their specific E2E service requirements. The created network slices are temporarily owned by tenants who have control over multiple layers, i.e., the physical layer, the virtualisation layer, and the service layer in 3D, of a unified 5G infrastructure, while they are also verticals. That is, they integrate the 5G infrastructure vertically on ground, aerial and space networks. The availability of this vertical market multiplies the monetisation opportunities of the network infrastructure as (i) new players, such as the space industry and military, may come into play, and (ii) a higher infrastructure capacity utilisation can be achieved by admitting network slice requests and exploiting multiplexing gains. With network slicing, different services, such as space IoT, safety-critical aircraft operations connectivity, UAV connectivity, and mobile broadband, can be provided by different network slice instances. Each of these instances consists of a set of virtual network functions that run on the same infrastructure with a tailored orchestration. In this way, heterogeneous requirements can be provided on the same infrastructure, as different network slice instances can be orchestrated and configured separately according to their specific requirements, e.g., in terms of network QoS. Additionally, this is performed in a cost-efficient manner as the different network slice tenants share the same physical infrastructure.

Additionally, to these connectivity services, the capability to deploy virtual assets in NTN infrastructure enables the exploration of new commercial opportunities from other domains. For instance, the possibility of deploying software-based payloads allows the execution of algorithms or applications associated with Earth Observation UCs [63] or radio frequency (RF) interferences monitoring [65]. This capability thus opens the opportunity to explore the re-usage of operative non-terrestrial infrastructure. This approach also contributes to the recycling of spacecraft and aircraft. Specifically, satellites may become decommissioned if the original mission has been achieved. The possibility of deploying new applications in this hardware, without considering any hardware dependency, enables the recycling of these satellites for new purposes. Thus, the usage of this novel architecture, which enables automatic deployment of applications on top of non-terrestrial infrastructure, would promote the apparition of new commercial opportunities.

To harness the power of such virtualised infrastructure, an intelligent architecture of upcoming 5G networks calls for an efficient management framework that provides a uniform and coherent orchestration of various resources across the multiple layers of the 5G ecosystems deployed in the different 3D network strata. NFV and their MANO systems offer themselves as effective approaches, aiming at decreasing the cost and complexity of implementing and deploying novel services, maintaining running services, and managing available resources in the ETHER 3D network infrastructure.

4.4 ETHER MANO LAYER

As future networks strive towards notarisation, network functions virtualisation becomes the leading trend, with 5G CN using SBA, and the envisioned 6G MANO [17] also being service-based (service-based management architecture). It is evident that the orchestration of network services requires special attention.

In future 6G networks, orchestration processes can be split into four basic categories [17]:

- E2E seamless integration processes, which refer to the orchestration of services using available infrastructure as a common pool of resources. The main challenge of these processes is the increasing heterogeneity of infrastructure resources.
- Programmable processes, which refer to network programmability, focusing on the deployment of enablers for management and troubleshooting of network devices.
- Automation processes, which are a direct consequence of network programmability. With the rapidly increasing sophistication of networks, manual processes are no longer a viable option.
- Data-driven processes, which mainly refer to the optimisation of networks using data originating from multiple network levels (slice, service, infrastructure, etc.). The main example of a data-driven process is AI/ML-driven orchestration, which is a necessity to achieve zero-touch automation.

The main focus of ETHER MANO will revolve around E2E seamless integration processes, such as the orchestration of E2E slices/services and MEC Applications, which can be done in at least two main ways [17]:

- Using centralised orchestration being able to orchestrate resources of different subnetworks directly,
- Use multiple orchestration domains with a dedicated orchestrator.

This chapter will present approaches to orchestration as well as the initial proposed architecture of ETHER MANO. The Hexa-X project, as a flagship EU-funded 6G project, proposed a 6G MANO with key features including: a microservices-based approach, intent-based services, AIOps/DevOps friendly, the possibility of private networks and zero-touch management. Hexa-X also proposed a MANO architecture specifically tailored to future 6G networks and features decoupling of management from orchestration. However, ETHER relies on standardised MANO solutions tailored to existing 5G network, and aims to extend them by considering gaps, possible solutions and novel features of MANO recognised by Hexa-X. These extensions will be possible through the creation of a generic MANO framework capable of hosting diverse solutions. The most notable standardised MANO framework compatible with the 3GPP management system is ETSI NFV MANO.

4.4.1 ETHER MANO generic architecture

One of the potential solutions to tackle the scalability issue can be introduced by the inclusion of hierarchical ETHER MANO, with an administrative domain split. This is the second of the previously mentioned E2E orchestration approaches. In this approach, each ETHER MANO Domain is self-managed and responsible for the LCM of its VNFs and services. E2E ETHER MANO facilitates the creation of E2E services across multiple domains. At the Administrative Domain Level, there can be multiple domain types with different administrative and technological properties (e.g., TN/NTN Cloud, RAN, Edge, Transport, domain combinations, etc.).

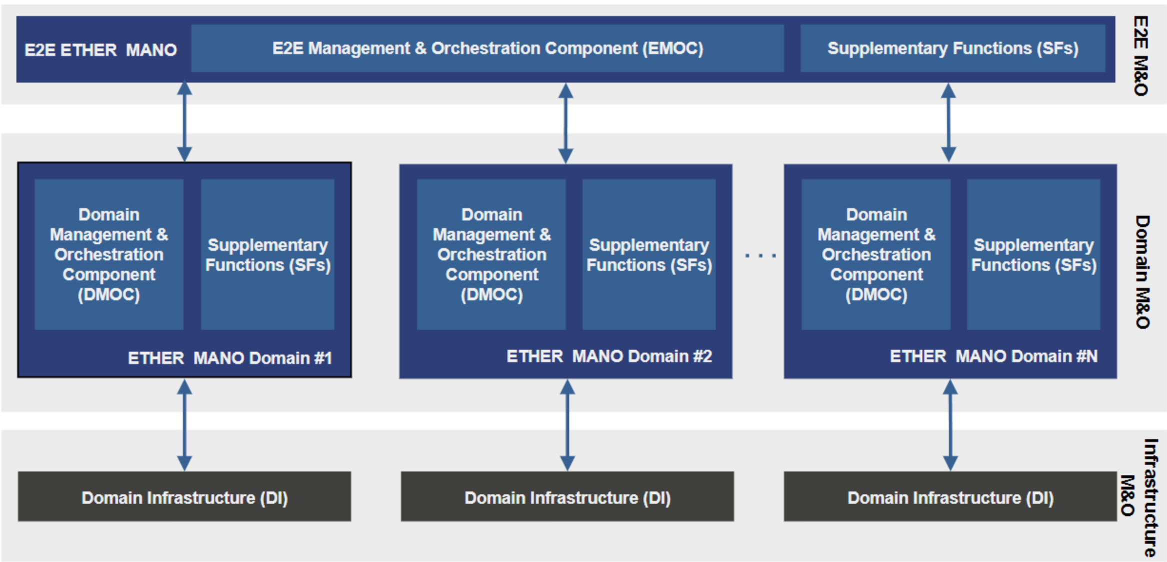

Figure 4-3: Generic ETHER MANO architecture

The M&O in ETHER MANO is split into three levels of hierarchy (cf. Figure 4-3):

- **E2E level** – composed of the E2E Management & Orchestration Component (EMOC), which has the global visibility of the available technological and administrative domains together with interconnections between them (i.e., the spatial distribution and topology of domains under the umbrella of ETHER MANO). Additionally, EMOC is responsible for E2E M&O and the required Supplementary Functions (SFs) that support E2E network operations and capabilities exposure.

- **Domain level** – composed of multiple self-contained and isolated domains, each composed of a Domain Management & Orchestration Component (DMOC) and additional functions supporting domain-specific operations and capabilities exposure.

- **Infrastructure level** – composed of multiple Domain Infrastructure (DI) entities constituting an abstracted view of the underlying infrastructure (physical, virtual, or combination thereof).

Considering the key ETHER MANO functionalities and mechanisms, the main capabilities of EMOC and DMOC have been identified. It is assumed that in order to maintain the separation of concerns between M&O and M&O-related processes (e.g., business logic, secure capability exposure) both on the domain and global levels, the EMOC/DMOC are focused on the M&O, while the latter are implemented in the form of a service mesh denoted as SFs (e.g., microservices).

The key functions of EMOC include:

- **E2E M&O processes,** including LCM of E2E slices/services, E2E slice/service template partitioning to sub-slices, selection of target domains for the respective sub-slices and sending intent-based requests to the respective DMOCs. The selection of domains and partitioning can be driven by multiple factors, including the capabilities offered by the specific domain as well as business aspects (e.g., price for the deployment of slice/sub-slice components),

- **E2E M&O optimisation** by exploiting relevant functions of the orchestrator, e.g., via Monitor-Analyse-Plan-Execute over a shared Knowledge (MAPE-K) control loops [66]),

- **Handling the dynamic registration** and connectivity establishment to new administrative/technological domains to allow intent-based communication. It is assumed that dedicated SFs can be deployed to assist the connection process.

The key DMOC functionalities are:

- **Domain M&O processes**, which include the LCM of domain slice/sub-slice/service implemented using service templates indicated by the DMOC or requested via domain-level business interfaces. They also involve capabilities exposure mechanisms (e.g., the domain M&O APIs, NS/VNFs catalogue).

- **Domain-level M&O processes optimisation**, e.g., VNF placement decisions, VNF migration, resource scaling, domain M&O FCAPS.

- **Resource orchestration** via direct interaction with NFVI or SFs mediating the communication with external NFVI providers (e.g., handling the negotiation processes, or interconnecting with cloud-edge continuum).

To allow the proper separation of concerns, the EMOC/DMOC components are focused on the resource aspects, while the slice/service-specific management is performed via dedicated functions implementing the In-Slice Management (ISM) concept (VNFs deployed in the form of SFs attached to the VNFs constituting the slice) [67]. To this end, the following functionalities of ETHER MANO are implemented by the loosely coupled SFs operating both on the E2E level and domain levels:

- OSS/BSS system compliant with the 3GPP approach;

- ETHER MANO exposure mechanisms towards business stakeholders allowing for, e.g., issuing high-level requests for E2E/domain slices and services, such as Generic Network Slice Template (GST)/NEtwork Slice Type (NEST) [68]. These mechanisms also involve their translation to E2E templates, and LCM of E2E slice/service.

- Support for AIaaS, including maintenance of AI models catalogue for AI-driven M&O, exposure of AI models used within the framework, interconnection with external AI models consumption, and anonymised data sharing to trusted entities.
- Security-related mechanisms leveraging zero-trust approach handling the authorisation and authentication of entities interacting with the ETHER MANO framework.
- Zero-touch management of E2E slices/services exploiting the ISM concept on the domain level with hierarchical components managing the entire sub-service chain.
- Interconnection with NFVI providers including interactions with cloud continuum brokers and NFVI providers.
- Support for adopting autonomic control loops to automate ETHER MANO processes.

It is assumed that the above list of functionalities is non-exhaustive. Hence any function extending the framework can be implemented as a dedicated SF.

4.4.2 Generic architecture instantiation using existing management and orchestration frameworks

The ETHER MANO approach described in Section 4.4.1 allows the integration of multiple different M&O stacks under the umbrella of technology-agnostic EMOC. ETHER MANO is a generic framework, and therefore, existing MANO solutions should be able to fit inside its boundaries. Hereby, we outline the proposal of instantiation of the ETHER MANO framework with technological domain separation using other existing M&O frameworks. On the domain level, ETHER architecture considers administrative and technological domains such as Cloud, Edge (e.g., MEC-based edge), RAN and Transport residing in the Terrestrial (T), Aerial (A) and Space (S) Strata. An example of this architecture instantiation is presented in Figure 4-4.

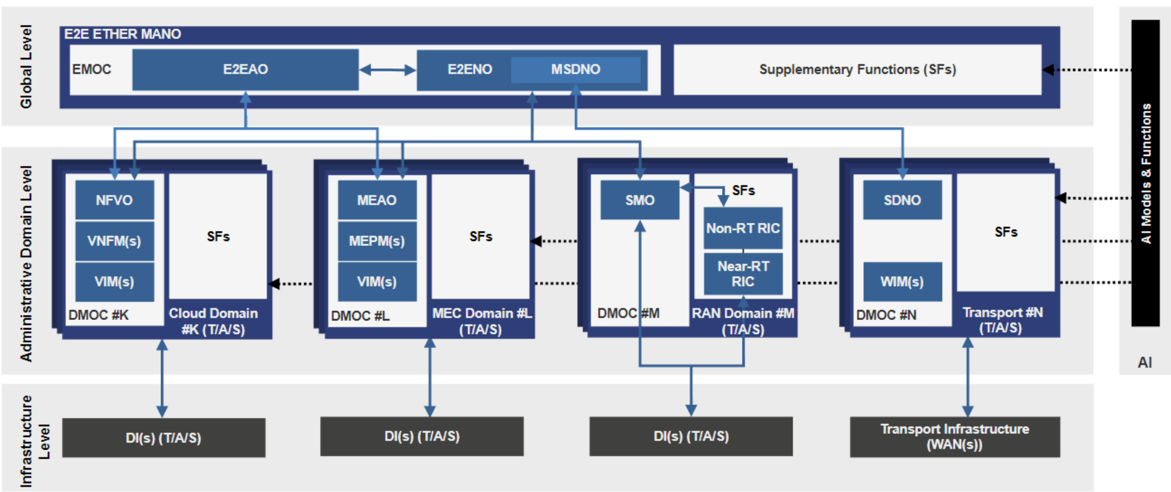

Figure 4-4: High-level view of ETHER MANO with cloud, MEC, RAN and transport domain separation

In this example, the cloud domains are handled by ETSI NFV MANO orchestrators and edge domains by MEC orchestrators, implementing the full technological stacks as standardised by ETSI. To perform RAN management open-source frameworks such as Open RAN [69] or proprietary controllers can be used. Specifically, the different functions would be deployed in the Cloud domain, composing thus the O-Cloud of O-RAN. Additionally, the SMO (Service M&O) and both RICs (Radio Intelligent Controllers) would conform a dedicated RAN domain that directly interacts with the disaggregated gNB functions previously deployed and using the

O-RAN interfaces. O-RAN management system is explained in detail in [70] and summarised in Section 3.1.2. The interconnection of the domains is performed via multiple Transport Domains, each consisting of two main blocks: SDN Orchestrator (SDNO), and WIM. This approach is an extension of the SD-WAN concept. The SDNO block is responsible for sending domain-level SDN orchestration requests, which are supposed to enable the connectivity of services within a single transport domain. WIM is responsible for the translation of SDNO requests into SDN configuration requirements for SDNCs it manages. It has to be noted that one SDNO can contact multiple WIMs (i.e., operate in the multi-vendor and multi-provider WAN environment) to obtain the path at the domain level (cf. Figure 4-5), however, it is generally expected that stitching of the path (i.e., composing the E2E path from a number of domain paths) will be performed on the E2E level. It is essential to emphasise that SDNCs extend beyond transport network segments, encompassing computational domains, the 5G core network (as detailed in Section 5.4.3), and access networks. Within this framework, the EMOC part associated with the network (discussed in details below) plays a crucial role in managing network connectivity to ensure seamless E2E communications. With comprehensive visibility of all SDNCs, it has the capability to dynamically adjust network topologies to meet service demands.

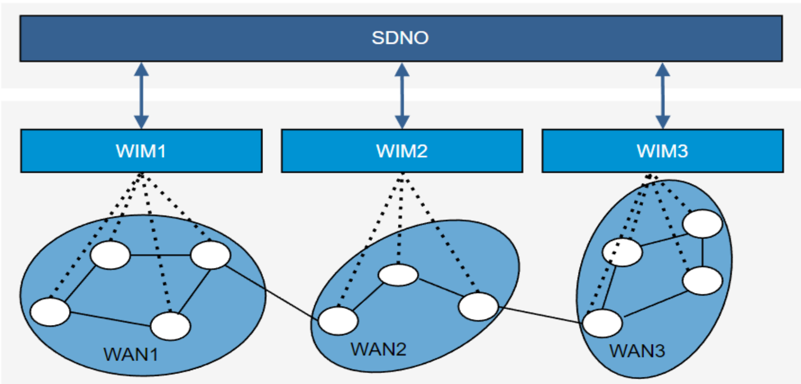

Figure 4-5: SDNO interactions with multiple WIMs (WAN providers) within a single Transport Domain

EMOC in E2E ETHER MANO consists of two functional blocks:

- E2E Application Orchestrator (E2EAO), which overlooks Cloud and MEC domains in order to orchestrate MEC Applications.

- E2E Network Orchestrator (E2ENO), which overlooks all orchestration domains and facilitates the creation of E2E network slices/services. E2ENO also hosts Master SDNO (MSDNO), which is responsible for E2E SDN Orchestration, which includes coordination of SDNOs to create cross-domain virtual links for E2E slices/services out of domain-level virtual links provided by relevant Transport Domains.

The proposed ETHER MANO architecture is scalable and generic and provides a strong separation of concerns. This genericity allows the adaptation of different frameworks, such as ETSI NFV for Cloud domains, ETSI MEC for MEC domains and O-RAN for RAN domains. It allows for management on different levels of hierarchy. The presence of intra-domain SFs allows for other concepts like In-Slice Management or AI-driven M&O to be used if needed. Specific domains can be tailored to work either with specific vertical infrastructural levels (satellite/aerial/terrestrial) or any combination of them at the same time. The adoption of virtualisation allows for high network flexibility, which is especially important in the context of satellite and aerial infrastructure. To enable full E2E integration within the ETHER MANO framework, it is imperative to specify inter-level as well as intra-level interfaces, which is within the scope of future work of WP2 and WP4.

4.4.3 Cross-domain management and orchestration in 3D networks

ETSI NFV, as outlined in [71], establishes methodologies for managing the life cycle of VNFs distributed across multiple points of presence or NFVI nodes. It presents various architectural options for managing different VIMs and VNFMs through one or multiple NFVOs. The interaction between NFVOs enables the execution of distributed network services across administrative domains composed of one or multiple VIMs managed by a single NFVO. This orchestration-level collaboration facilitates coverage of increasingly expansive geographical areas.

On the other hand, in document [72], ETSI introduces the concept of multi-site connectivity management, aiming to guarantee the links between distributed VNFs. To achieve this objective, ETSI NFV proposes the use of a WIM to manage the links between NFVI nodes. This empowers service providers to guarantee service performance by ensuring their KPIs through the control of network elements in the transport network. WIM exposes APIs to program the transport network on demand. Consequently, an NFVO can leverage WIM during the instantiation and configuration of a network service composed of VNFs distributed across different NFVIs. In this manner, E2E connectivity is assured, and traffic between network services can be handled independently or even isolated (applying the concept of network slicing).

While these solutions focus on scenarios, where both NFVI resources (managed by VIMs) and transport network elements are static nodes with a relatively stable topology, ETHER MANO must possess the capability to manage the life cycle of infrastructure resources and network services deployed across the three strata: ground, air, and space. The air and space strata, by their nature, encompass mobile devices such as drones, airplanes, helicopters, satellites, and so forth.

Since ETHER MANO is built upon the ETSI NFV architectural framework, managing non-static resources presents challenges that have not been addressed by industry bodies or SotA literature. The following section delves into the challenges and solutions proposed by ETHER to ensure seamless connectivity for network services deployed across the three strata (3D networks).

4.4.4 Key challenges in ETHER MANO Layer and proposed solutions

Challenges

ETHER MANO faces several challenges in managing and orchestrating infrastructure and network service domains distributed across static and non-static domains. The complexity primarily lies in managing non-static domains, mainly LEO and MEO satellites, due to their mobility capabilities. There are infrastructures that follow predictable movements, such as satellites and airplanes, and others that perform random movements, such as drones and helicopters. The ability to maintain communication with this type of infrastructure and with ground-based infrastructures can result in limited connectivity due to the visibility zones of the satellites. Connectivity intermittence can be a determining factor for services that require high availability.

On the other hand, the resources available in HAPS are limited, which requires efficient allocation of resources to the services running on them. For example, managing the resources of LEO satellites for executing regenerative payloads involves the application of mechanisms that allow reconfiguring the onboard hardware of the satellite. Similarly, in medium-altitude HAPSs such as airplanes or drones, dynamic resource management requires mechanisms that allow resources to be assigned or withdrawn flexibly and dynamically according to service requirements and network topology. To achieve this objective, it is necessary to keep the current location of the HAPS updated. The unified resource coordination in the ETHER MANO system allows efficient orchestration of virtualised services and E2E connectivity through 3D networks.

This implies that operators or algorithms, reactively or proactively, can make changes to service configurations and network topology updates, especially in integrated TN-NTN scenarios.

Therefore, comprehensive management of TN-NTN scenarios involves applying hierarchical models for the orchestration of different domains (e.g., satellite, ground, air) and network segments (e.g., access network, transport, and clouds). However, the limited connectivity between the different strata makes it difficult to perform dynamic fault recovery or reconfiguration operations for VNFs. It means preventive measures, which can be mechanisms or functionalities deployed on the HAPSs that report relevant availability and location information to autonomous optimisation mechanisms, are necessary.

Solutions

To address the challenges described above, ETHER MANO proposes specific solutions at both the infrastructure and M&O levels. Specifically, to solve the hardware reconfiguration in NTNs, such as satellites and HAPSs, to enable the deployment of reconfigurable virtualised services, ETHER proposes the concept of Flexible Payload, which is based on the principle of programming the logical resources of hardware boards (e.g., FPGAs) and applying software virtualisation mechanisms on the base OS. This implies that a satellite becomes an NFVI node that can be managed by a VIM.

On the other hand, to enable comprehensive and seamless management of these NFVIs through a VIM, ETHER proposes functionalities that allow tracking the availability of the nodes that make up the cluster. These solutions are described in more detail below.

Flexible Payload as the virtualisation layer of the NFVI

The flexible payload is a hardware and software-based framework prepared to run over a generic satellite payload platform via the execution of a custom Linux OS. It adopts the SDR concept by managing the data path between the main system processor and any available Radio Frequency transceiver. To materialise the design, the framework relies on FPGA System-on-Chip (SoC) technology, which lately offers enough resources to deploy an entire payload system in a single integrated circuit (IC) and in combination with a high-speed transceiver (above GHz).

Given the abundant computational resources currently offered by these platforms, the framework can deploy infrastructure in up to three different areas:

1. Hardware-level deployments are achieved by programming the available logical resources (cells) of the FPGA SoC. Although the processors integrated into modern FPGAs are the same for each family of products, the number of logical cells available per product varies and it is important to manage their occupancy and distribution when deploying specific services. The flexible payload framework aims to solve this problem by in-flight reprogramming the empty logic cell areas to accommodate additional hardware services.

2. Software-level deployments use a software virtualisation mechanism called containerisation. The idea behind this is to separate services in different containers (guests) that can be easily launched via system commands. This offers an abstraction layer between the host's OS and the service OS. However, they share the same kernel functions.

3. Orchestration from GSs or HAPS or other satellites within a constellation via ISL. The framework integrates a Kubernetes distribution as a demonstrative tool (and because it is the de facto open source tool for orchestrating containers) to enable service orchestration via an RF link from the ground or from an ISL link to other satellites.

The use of the Flexible Payload framework in devices oriented to the aerial stratum could be also an option but its inclusion is strictly limited to the computing architecture and capacity of the provided hardware. Unlike in space, the aerial hardware platforms are diverse and can present very different computing profiles. For example, drones tend to incorporate low-end, low-power processors and do not run OS. More heavy platforms such as HAPS can integrate more complex computing architectures depending on service requirements (mission goals). They can also have different energy consumption profiles. In this way, if a FPGA fits, all the features of the Flexible Payload can be preserved, but if a simple processor is used instead then hardware reconfiguration would not be possible. Furthermore, if no OS is running, then no virtualisation mechanism will be available, and the software virtualisation (containerisation) will not be implemented. Therefore, it is not possible to generalise the use of the Flexible Payload solution in the aerial layer.

The Flexible Payload framework can integrate easily into the NFVI architecture as a virtualisation layer. The envisioned role of the framework is to enable the virtualisation of commercial payload services into general-purpose satellite payload platforms and provide, at the same time, certain flexibility of deployment based on premises such as geographic area, resources allocation, computer power, battery life, etc.

Because services usually require additional network resources, both the high-level applications and the service required infrastructure components can be implemented through the flexible payload framework. This vision matches with NFV, which aims to virtualise all physical network resources via available Virtual Machines (VMs) – or containers in the specific case – management programs. The framework allows the network to grow without adding more physical hardware devices but via mapping virtual services to either software or logic cells of the payload platform which creates new infrastructure for the network. Orchestration capabilities from local FPGA-based satellite nodes add remote deployment flexibility.

VIM functionalities across space, ground and air stratum

VIM is a critical component in managing the NFVI across terrestrial, aerial and space stratum. The NFVI encompasses computing, storage, and network resources either onboard non-terrestrial or terrestrial nodes, including FPGAs that enable hardware reconfiguration. VIM can be based on either VMs or containers. However, due to the resource-constrained environment of satellite systems and drones, containers are the preferred choice for virtualisation. Containers offer several advantages over VMs in resource-constrained environments. Furthermore, they are more lightweight and require fewer resources to run, making them more efficient for deploying applications in environments with limited computing power and memory. Additionally, containers can be easily scaled up or down on demand, which is particularly important in satellite communications, where resources may be limited or variable. Kubernetes is a de facto open-source container orchestration platform that automates the deployment, scaling, and management of containerised applications. However, for platforms with constrained resources, a lightweight version of Kubernetes is more suitable, known as k3s. It serves as a suitable VIM tool for various non-terrestrial nodes (e.g., UAV, HAPS, aircrafts, satellites) deployment scenarios, impacting factors like autonomy, performance, and scalability. Its compact nature makes it an ideal choice for resource-constrained platforms, allowing for efficient management of containerised applications without compromising on performance.

The strategic placement of VIM functionalities with k3s enables seamless integration with 5G and 6G mobile networks. This integration facilitates a hybrid terrestrial/non-terrestrial architecture, where VIM orchestrates resources across both networks, creating a unified infrastructure for service delivery. This architecture leverages the strengths of both TNs (high capacity, low latency) and NTNs (broad coverage, global reach). The k3s deployment scenarios impact factors like autonomy, performance, and scalability. In a single-node cluster, k3s is embedded directly within the non-terrestrial node, simplifying management as all resources are localised. k3s orchestrates containerised applications and manages resource allocation for onboard functionalities. It can also directly control FPGAs for hardware reconfiguration tasks specific to mission requirements. This setup is ideal for missions requiring high autonomy and minimal ground intervention.

For multi-node clusters, k3s is distributed among the terrestrial, air and space stratum components. The ground-based k3s instance handles complex management tasks like service deployment and orchestration across the cluster. The air-based and space-based k3s instance focuses on real-time resource allocation and payload reconfiguration, including managing containerised applications and dynamically reconfiguring FPGAs based on network demands to optimise performance. Multi-node clusters with high availability build upon the multi-node cluster concept by distributing k3s master instances across multiple nodes for redundancy and fault tolerance. This ensures continuous operation even if a node fails. k3s facilitates collaboration between nodes to reconfigure FPGAs across the cluster, maintaining optimal performance.

Finally, cluster federation involves dynamic and federated clusters of satellites that maintain service continuity over geographical areas. k3s acts as the federated VIM, managing resource allocation and service orchestration across the entire network. It ensures efficient resource utilisation and coordinates FPGA reconfiguration across the distributed NTN to meet varying demands.

By strategically placing k3s functionalities between space, air and ground, is achieved a balance between autonomy, performance, and scalability. This placement ensures that NTN can adapt to varying operational demands for seamless service delivery. k3s plays a vital role in managing resources, orchestrating containerised applications, and collaborating on FPGA reconfiguration for optimised performance and adaptability. Additionally, the integration with 5G/6G networks unlocks new possibilities for hybrid network architectures and dynamic service provisioning to meet the evolving needs of next-generation communication systems.

Nevertheless, coordinating these technologies (flexible payload and VIM management) through the MANO framework is a crucial aspect that is not contemplated by ETSI NFV. For this reason, ETHER MANO proposes an innovative infrastructure mobility management model that allows the orchestration of the life cycle of the network and infrastructure substrates. The next section provides more details about the mobility management model's components, roles and functionalities, which are integrated with the MANO framework.

4.4.5 Infrastructure mobility management model

The management of infrastructure mobility is one of the most complex aspects that ETHER MANO addresses during the project's execution. Neither the ETSI NFV nor the ETSI MEC standards define functionalities in their architectures that allow the LCM of non-static NFVIs through one or multiple VIMs. Therefore, the LCM of VNFs that need to be deployed in these domains becomes impossible using the current architectural frameworks defined by ETSI and the existing literature.

The dynamic and geographically distributed nature of NTNs presents significant challenges for traditional MANO systems [73]. Existing MANO frameworks, primarily designed for TNs, struggle to effectively manage the location of mobile nodes, optimise resource allocation across vast geographical areas, and ensure seamless service continuity [74]. Additionally, traditional MANO frameworks such as OSM and OpenBaton, while providing a robust foundation for orchestrating network functions, have limited capabilities when adapted to the unique characteristics of NTNs [75].

A Geographical Information System (GIS) offers a powerful toolkit for managing and analysing spatial data. In the context of NTNs, GIS plays a crucial role in planning, optimisation, and management of trajectories and geographical dispersion. NTN operators leverage geospatial data to plan satellite orbits, simulate communication scenarios, and optimise trajectories with special focus in the mobility of the nodes [76]. Tools like STK [77], provide essential geospatial capabilities for NTN planning, by enabling precise modelling of satellite trajectories and facilitates informed decision-making within the complex Earth's orbit environment [78]. Integrating mobility management capabilities from GIS into the ETHER MANO enables optimisation of resource allocation, prediction of coverage patterns, and enhancement of overall network performance [79]. This integration facilitates more intelligent and efficient management of NTNs, paving the way for the realisation of 6G network capabilities. Consequently, ETHER introduces an innovative infrastructure mobility management model, which incorporates the integration of SFs, including geostationary satellite information systems and location tracking of NFVI nodes deployed in ground, air, and space domains.

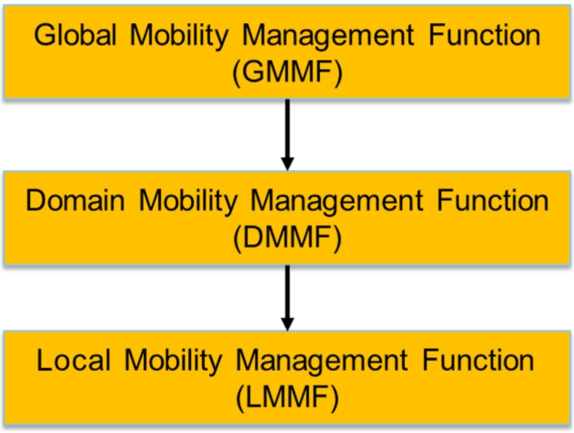

Figure 4-6: Infrastructure mobility management model

Figure 4-6 describes the infrastructure mobility management model. It is composed of three functions displayed hierarchically: Global Mobility Management Function (GMMF), Domain Mobility Management Function (DMMF), and Local Mobility Management Function (LMMF). Below is a description of the functionalities and roles of each component of the proposed model.

Global Mobility Management Function (GMMF)

The Global Mobility Management Function (GMMF) serves as the primary point of contact for the mobility management framework, particularly concerning infrastructure mobility. One of its key roles is to register and discover available domains by interfacing with the Domain Mobility Management Functions (DMMFs). Through this interaction, the GMMF ensures a comprehensive understanding of the network's landscape, facilitating efficient resource allocation and management.

Furthermore, the GMMF identifies the domains over which the physical infrastructure moves within the target area by communicating with the DMMFs. This identification process is crucial for maintaining seamless connectivity and service delivery as it allows the GMMF to continually track and update the domains that remain within the target area. If there is a change in the node's location, the DMMF promptly notifies the GMMF of this modification, ensuring that all necessary adjustments are made to sustain optimal network performance and reliability.

Domain Mobility Management Function (DMMF)

The DMMF holds several critical responsibilities within the mobility management framework. Primarily, it is responsible for the registration and discovery of available domains by establishing connections with LMMFs. Domain part refers to each network segment such as cloud, edge, transport, and RAN. This connectivity allows the DMMF to maintain an updated and comprehensive registry of all accessible domains, which is essential for efficient network operation and resource management.

In addition, the DMMF plays a vital role in identifying the domains that the physical infrastructure traverses within the target area, based on their mobility patterns, through continuous communication with the LMMFs. This identification process is essential for maintaining seamless connectivity and service continuity, as it allows the DMMF to monitor and update the domains that persist within the target area dynamically. Furthermore, should there be any changes in a node location, the DMMF is responsible for promptly notifying the GMMF of such modifications. This ensures that all necessary adjustments are made to maintain optimal network performance and reliability.

Local Mobility Management Function (LMMF)

The LMMF is responsible for several critical tasks in managing the mobility of physical infrastructure (e.g., UAVs, satellites, HAPs, etc., acting as edge, transport and RAN elements). Firstly, the LMMF is tasked with the registration and discovery of the location of physical infrastructure components, such as nodes. This involves maintaining and updating their locations according to their respective mobility patterns. Additionally, the LMMF must be capable of providing the current location of the physical infrastructure to external functions and modules at any given time. Furthermore, when possible, the LMMF is also responsible for estimating the future locations of the physical infrastructure based on mobility patterns.

To illustrate these responsibilities, consider the example of satellite management. The LMMF registers a satellite as a node with an ECI (Earth-Centred Inertial) position and orbit TLEs (Two-Line Elements). The LMMF propagates the satellite's orbit to compute updated positions, which are then refined with new TLEs and/or Global Navigation Satellite System (GNSS) measurements. This updated satellite position is provided through a standard API. The orbit propagation is performed using an orbit model, such as SGP4 [80], ensuring precise and reliable location updates.

Geographical Information System Function (GISF)

The GIS Function (GISF) holds several important responsibilities within the context of network management. Primarily, the GISF is tasked with mapping the required target areas to precise geographical coordinates. This mapping process ensures accurate spatial representation, which is crucial for various operational and strategic purposes.

Additionally, the GISF provides these geographical coordinates to E2E orchestrators, enabling them to efficiently manage and optimise network resources across the mapped areas. Furthermore, the GISF connects to external and online geographical information systems, facilitating the integration and utilisation of comprehensive geographical data from diverse sources. This connectivity enhances the accuracy and relevance of the geographical information provided to the orchestrators and other network management functions.

4.4.6 3GPP 5G System management and orchestration

The approach of 3GPP to MP is complementary to the ETSI NFV vision (cf. [52], Annex A.4), as it is shown in Figure 4-7. The following hierarchy of the 3GPP management functions:

Communications Service Management Function (CSMF) – Network Slice Management Function (NSMF) – Network Slice Subnet Management Function (NSSMF) – Network Function Management Function (NFMF) is mapped so that NFMF is the equivalent of NE according to ETSI NFV, and NSSMF is the equivalent of OSS/BSS to ETSI NFV. In this way, 3GPP supports the integration of many technological and administrative domains, in particular, the exposition of "Network Slice Subnet as a Service" (NSSaaS) by an operator to another operator.

The communications service and network management vision apply the SBA approach to the 5GS MP similarly to 5GS CP with all relevant concepts (Producer-Consumer, Request-Reply, Subscribe-Notify). Additionally, the equivalents of the 5G CP Network Exposure Function (NEF) and Network Data Analytics Function (NWDAF) are Exposure Governance Management Function (EGMF) and Management Data Analytics Function (MDAF) in the 5GS MP.

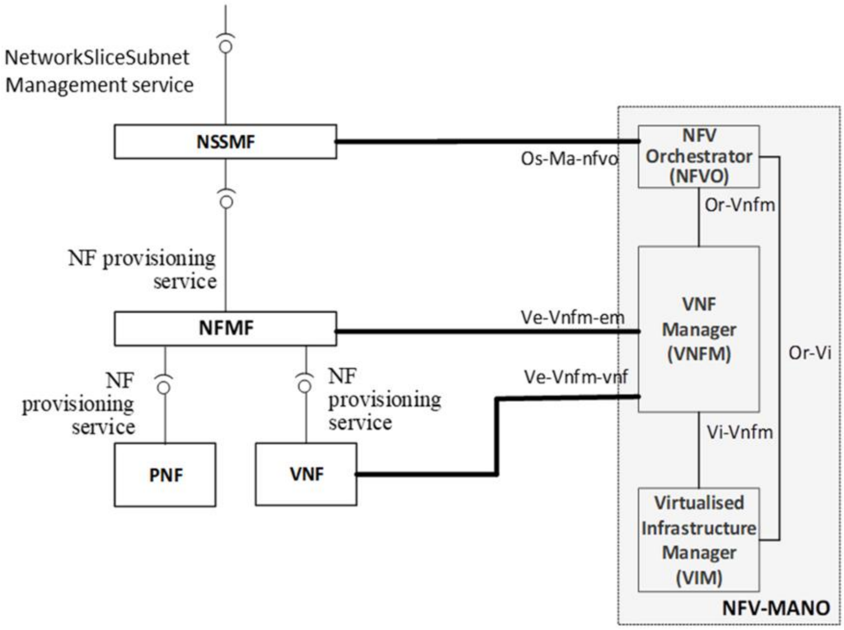

Figure 4-7: Integration of the 3GPP network and service management framework and ETSI NFV MANO framework (image reproduced from [52])

According to Figure 4-7, NSSMF and NFMF are consumers of the ETSI NFV MANO reference points *Os-Ma-nfvo* and *Ve-Vnfm-em* interfaces, respectively.

The NSSMF can consume the VNF LCM and network service LCM services provided by NFV MANO. It is also a producer of network slice subnet-related management services. NSSMF is responsible for managing network slice subnets, which are subsets of a network slice. It consumes VNF LCM and network service LCM services to orchestrate the life cycle of VNFs and network services within the network slice subnet.

The NFMF is capable of application-level management of VNFs and PNFs. It acts as a producer of the NF Provisioning service, which includes Configuration Management (CM), Fault Management (FM), and Performance Management (PM). NFMF also consumes the NF Provisioning service produced by VNFs and PNFs. NFMF is deployed for managing VNFs and PNFs at the application level. It handles provisioning, configuration management, fault detection and resolution, and performance monitoring tasks. It produces the NF Provisioning service, enabling other entities to consume these management services. NSSMF focuses on network slice subnet management, consuming VNF and network service LCM services, while NFMF is involved in the application-level management of VNFs and PNFs, both producing and consuming the NF Provisioning service for CM, FM, and PM.

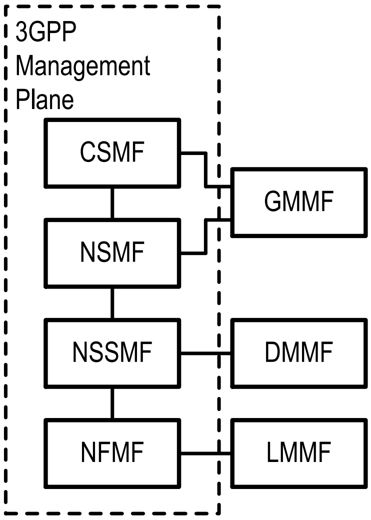

Figure 4-8: 3GPP 5G System MP in the ETHER framework

The ETHER framework extends the generic 3GPP MP vision according to the concept illustrated in Figure 4-8. The 3GPP MP functions are interfaced with corresponding xMMF instances to receive the supplementary information used to associate functional instances of the 3GPP framework with their location in the 3D space. As a consequence, it is possible to feed the 5G network control and management algorithms with data potentially important for, e.g., mechanisms for handling dynamic network topology, UE handovers, etc. Like the ETHER xMMF hierarchy, the 3GPP MP hierarchy is implemented in the area of SFs of the ETHER MANO and stretched over the ETHER MANO hierarchy.

4.4.7 Extensions to initial ETHER MANO architecture

The management of 3D networks, a key focus of the ETHER project, presents unique challenges. Unlike traditional TNs, 3D networks comprise network segments (domains) distributed across three distinct strata: ground, air, and space. This distributed nature necessitates a distinct infrastructure mobility management model that operates within a hierarchical and distributed architectural framework. Figure 4-9 provides a visual representation of this framework, highlighting the distribution of functionalities across the different domains and strata of the network.

The proposed model defines three primary domains within the 3D network:

- Cloud domain: This domain encompasses both edge and public/private computational resources, providing the processing power and storage necessary for network functions.

- RAN domain: This domain primarily consists of HAPSs or satellites functioning as network access points or BSs. HAPSs offer wider coverage compared to terrestrial access points due to their aerial positioning.

- Transport domain: This domain encompasses all network elements deployed across the three strata, including ground-based infrastructure, aerial links, and satellite communication infrastructure, facilitating data transmission between the various domains within the 3D network.

While the infrastructure mobility management model primarily focuses on managing the mobility of network elements within the air and space strata (e.g., drones, airplanes, satellites), it is also adaptable to include the ground stratum. This extensibility is particularly relevant in scenarios where seamless connectivity between TN and NTN segments is required.

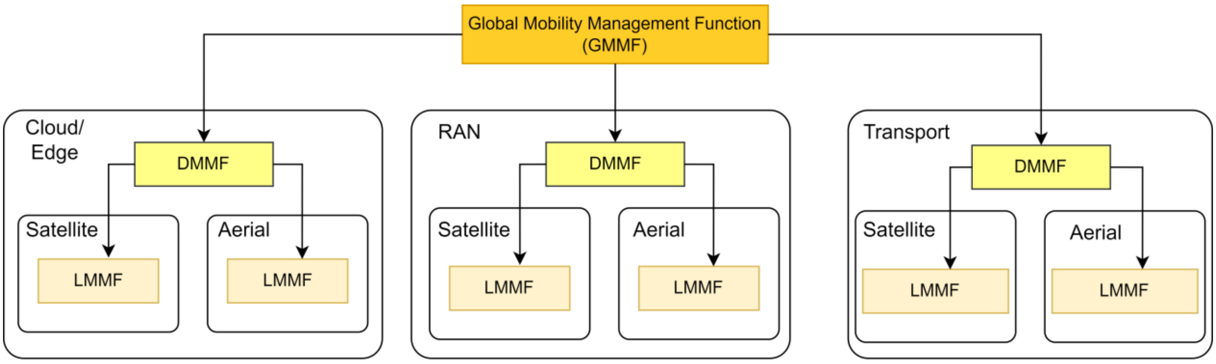

Figure 4-9: Infrastructure mobility management model applied in ETHER 3D Networks

The effective management of 3D networks necessitates the integration of all ETHER M&O functionalities with the infrastructure mobility management model. Document D2.1 [4] establishes a foundational version of the ETHER architecture, consisting of three primary levels:

- Infrastructure level: This layer encompasses the deployment of all networks and computing elements, including both terrestrial and non-terrestrial infrastructure. Notably, network service instances are deployed on these infrastructure elements.

- Administrative domain level: This layer defines the M&O functions specific to each network segment. For instance, the ETSI NFV architectural framework [62] governs the LCM of computational resources and VNFs in cloud domains. Similarly, the ETSI MEC architectural framework [37] manages the life cycle of computing systems and their associated applications within edge domains. Likewise, RAN domains leverage dedicated orchestration and control systems for network resources, while SDN orchestrators manage transport networks through WIM controllers.

- Global level: This centralised level oversees the overall M&O of the ETHER system. Centralised orchestrators are defined for both functionalities: E2E management of virtualised network services (i.e., E2EAO) and their interconnection across network segments (i.e., E2ENO).

The integrated NT-NTN approach aimed by ETHER seeks that the M&O capabilities seamlessly interact with the infrastructure mobility management model to enable efficient network operation within the three-dimensional network environment.

To enrich ETHER MANO capabilities, the functionalities of the infrastructure mobility management model are seamlessly integrated. Figure 4-10 depicts this architectural extension. The infrastructure mobility management functions, represented by yellow boxes, interact with ETHER MANO functions at both the global and administrative domain levels. At the global level, the GMMF interacts with the E2EAO and E2ENO. This interaction facilitates the exposure of the abstracted location information for all infrastructures deployed across the different domains. Conversely, at the administrative domain level, LMMFs interact with each domain's specific infrastructure manager. In essence, each domain houses several LMMFs corresponding to the number of infrastructure controllers present. These LMMFs connect to a single DMMF, which aggregates the location information for all infrastructures within the domain. It is important to note that each domain encompasses three distinct strata: ground, air, and space. While Figure 4-10 solely focuses on the connectivity of the infrastructure mobility management model functions, a detailed description of all ETHER MANO system component interfaces is provided in Section 4.7.

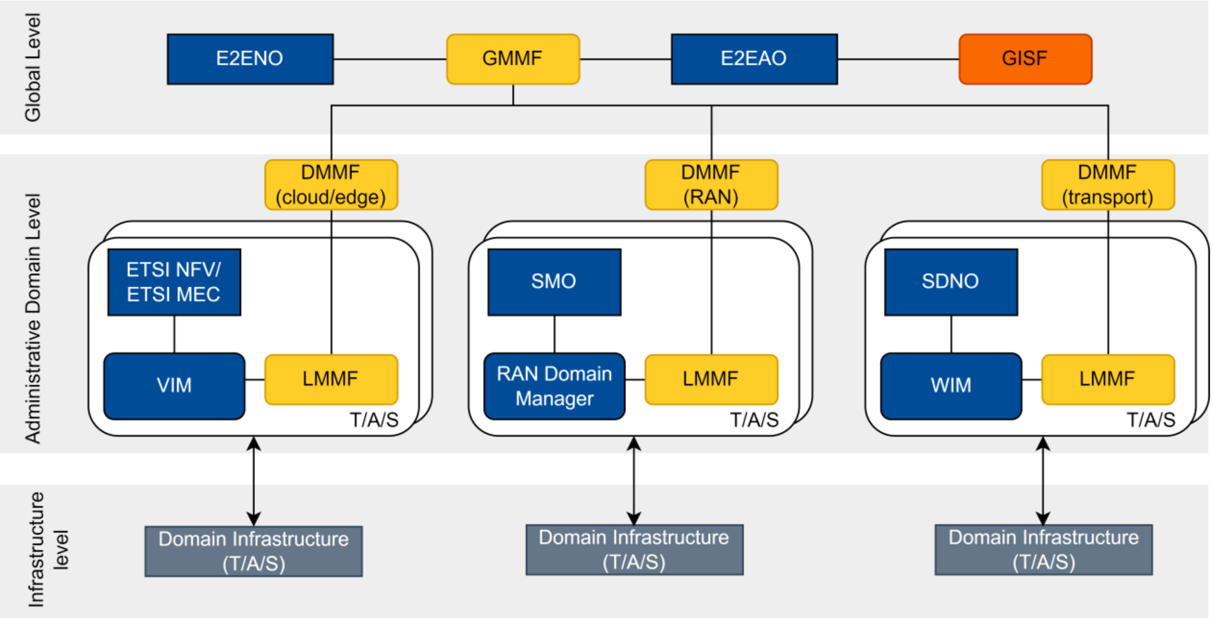

Figure 4-10: Infrastructure mobility management model integrated within ETHER MANO architecture

4.5 ETHER E2E SERVICE LAYER

The ETHER E2E Service Layer (E2ESL) implements the Application, Communication, and Steering strata as designed in the overall ETHER architecture (cf. Section 4.2). The following subsections present the key features of the 3GPP 5GS Stand-Alone (SA) architecture that are crucial for the ETHER System architecture, the analysis of the ETHER requirements associated with E2ESL, and the outline of E2ESL with embedded 5GS SA.

4.5.1 3GPP 5GS SA architecture and its key features

Generic 5GS architecture framework

The evolution of the mobile network towards the 5G has followed the opportunities created both by the softwarisation of the network and the separation of its hardware and software layers. The mobile network is traditionally organised in two planes, which reflect the fundamental activities of the network: user data transfer (in UP) and network control mechanisms (in CP). In a 5G network, these planes are spatially separated, i.e., they may be located optimally, according to the spatial distribution of user data traffic and control traffic. Furthermore, the softwarised CP NFs and their interactions have been reorganised into SBA, utilising well-known web mechanisms (REST API, HTTP/2, JSON format). The diagram of the generic vision of the 5GS SA architecture [51] with the indication of the most important functions is shown in Figure 4-11.

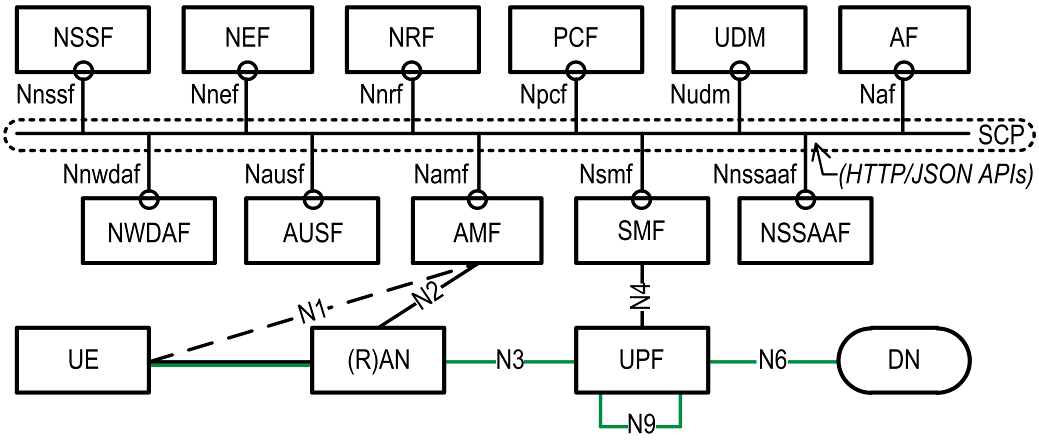

Figure 4-11: Generic 5GS architecture

The user terminal (i.e., UE) communicates with the core part of the mobile network via the access network (usually RAN). The user data follow the path of UP (depicted as green), where the external Data Network (DN) based on the Internet Protocol (IP), e.g., the public Internet, is reached via the UPF, which represents a chain of various data processing functions, e.g., packets inspection and classification, selective forking/rerouting, firewalling, marking/enrichment, validation, etc. Unlike in the previous mobile network generations, in which UP was just a data tunnel anchored at one point, the UPF can now be flexibly arranged according to the service class specificity and then further tailored and parametrised following the individual UC's needs. The signalling traffic between UE and the Core Network (represented as *N1*) is forwarded by RAN to the AMF, which acts as a CP-side termination point of UE-CP signalling and coordinates all UE interactions with CN. RAN is controlled by CN via the interface *N2*.

The CP functions, according to the SBA paradigm, expose CP services as their "producers" and utilise them as "consumers" (both Request-Reply and Subscribe-Notify mechanisms are possible). The mutual interaction and communication of CP NFs is supported by the Service Communication Proxy (SCP) as a CP message broker and Network Repository Function (NRF) for CP services discovery. The User Data Management (UDM) function with User Data Repository (UDR) behind manages user subscription data to be demanded for UE requests validation (network attach, data session establishing, QoS, etc.). The user data session-related requests from UE are forwarded by AMF to dedicated UPF-specific SMF supported by PCF. Also, the UE mobility-related network activities (e.g., UE handovers between BSs, UE tracking) are coordinated by AMF with SMF and RAN for UP flow continuity. The process of UE authentication is supported by the Authentication Server Function (AUSF). The service of analytics of network events and performance data, collected from all NFs, is produced by the NWDAF and served to all subscribing consumer functions.

To provide services to users outside the area of their own home network, the roaming implementation architecture of the 5GS framework has been defined. From the users' point of view, two basic strategies of UP traffic handling are distinguished: home-routed (the user's UP chain is terminated in the DN far away from the current location of the user) and local breakout (the UPF is terminated in the DN nearby). Each strategy might be applied according to the specificity of the user data channel UC (for more details, cf. [51]).

5G network exposure

The role of the NEF is to act as a CP gateway for non-native or untrusted CP NFs. They will consume the exposed CP services via NEF. NEF is also designed to integrate external higher architecture level systems, e.g., business or operation support layer systems, as well as vertical environments, with mobile network control mechanisms. In particular, NEF supports the following functionality:

- Secure exposure of capabilities, events, charging information and network analytics;
- Secure exposure/supplying the information of/for NWDAF at the 5G network CP border;
- Secure provision of information from external systems to the 5G network;
- Format adaptation, translation of information, including masking of network/user-sensitive information.

In the case of NEF used for external CP exposure to the northbound systems and environments (i.e., higher level systems), the CAPIF defined by 3GPP [53] may be used that standardises common aspects applicable to any northbound service APIs of the 5GS (discovery and publishing of service APIs, authorisation, logging, charging, API management). Additionally, SEAL [54] has been defined, which specifies the functional architecture and the procedures, information flows and APIs for each SEAL service to support vertical applications over the 5GS. The application layer architecture, procedures, and information flows necessary for enabling edge applications deployment over 3GPP networks (EDGEAPP) have also been specified [81].

Mechanism of Application Function

AF architectural entity may be considered as the "embassy" of the application layer within the mobile network CP. It provides means for the UC-specific applications to interact with the CP mechanisms; as AFs are aware of the quality situation in UP, usually the interaction with PCF is indicated for reconfiguring the UPF by altering the traffic routing, requesting the bandwidth or usage thresholds according to the application's requirements, as well as subscribing for session events' notifications. However, AF may potentially consume all CP services to the extent granted by the 5G network operator. For trusted AFs (e.g., MNO's own) the interaction with CP NFs may be set directly, while for untrusted AFs – via NEF.

Slicing in 5GS

The support of network slicing by the 5GS is illustrated in Figure 4-12. The slicing-aware UE, according to some application's needs, may request establishing the data session through a service type- or UC-specific UPF chain controlled by its dedicated SMF, accompanied by PCF, to a specific DN (e.g., virtual private network). NSI selection process is supported by a special functional entity, Network Slice Selection Function (NSSF), based on user subscription data in UDM. It is assumed that UE may be concurrently attached to up to 8 different NSIs [58]. NSSF determines availability of a requested NSI in a specific area and indicates a substituting NSI in case of unavailability. In some specific situations, the UE admission to the NSI may be associated with a special authentication and authorisation procedure, e.g., with the participation of a third party. To that end, the Network Slice-Specific Authentication and Authorisation Function (NSSAAF) may be involved.

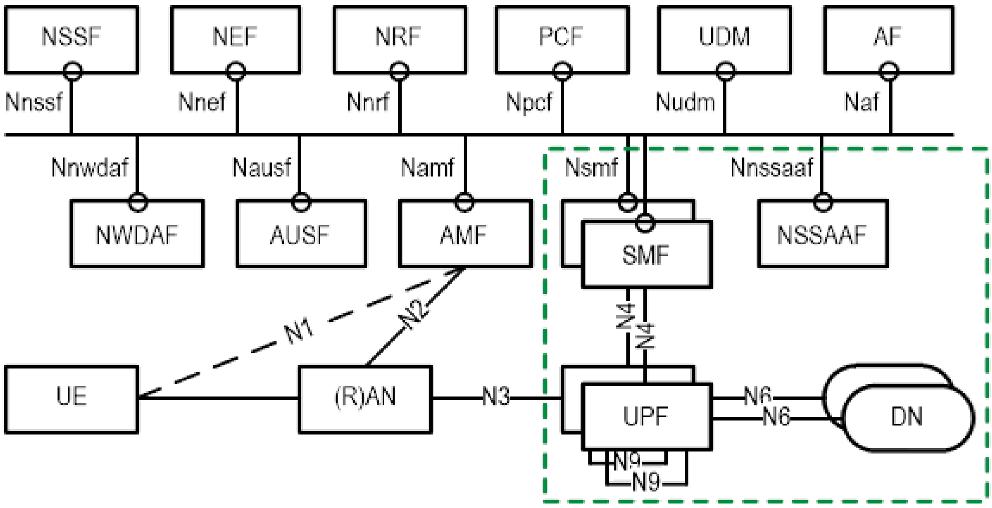

Figure 4-12: Network slicing in 5GS

The entire signalling exchange between UE and AMF, to distinguish the context of specific NSI, is labelled with the unique Single-Network Slice Selection Assistance Information (S-NSSAI) [51], which is composed of Slice Service Type (SST) – 8 bits, and optional Slice Differentiator (SD) – 24 bits; according to MNO's naming convention may encode properties of specific network slice template, purpose, usage, tenant etc. So far, 3GPP has defined 5 SSTs: eMBB; Ultra-Reliable Low Latency Communications (URLLC); massive IoT (mIoT), as defined by ITU-R IMT-2020 [82], and further, gradually coming extensions: Vehicle-to-Everything (V2X) and HMTC (high performance MTC). Generic network slice templates, based on a defined list of attributes, are described by the GSM Association (GSMA) [68].

The area surrounded by a green dashed line in Figure 4-12 shows the group of NSI-specific functions (primarily UPF-SMF pairs for UP slicing and optional NSSAAF). There are, however, other ways of slicing implementation in CP.

To satisfy the specific needs of a certain UC or application, NWDAF may implement additional analytic algorithms, e.g., a mechanism of detection and early warning of possible communication loss if the UE trajectory is maintained. The NWDAF functionality may be split and distributed within CP, and multiple, differently specialised NWDAFs with the scope related to supported NFs classes, collection and analytics mechanisms can be deployed per area or span level (global, shared by several NSIs or slice-specific – per individual NSI). The implementation of AF may also be either NSI-specific or AF can be shared by a group of NSIs if a specific AF has a wider scope of applicability than a single tenant's UC. NEF may implement specific services for a UC/NSI application.

Finally, network slicing in CP is supported by the CP SBA mechanisms presented earlier (NRF, NEF, SCP). While NRF supports NF service discovery and maintains NF profiles with available NF instances, their supported services and health status, these features may refer to NF instance span level. In particular, these are the authorisation of CP NFs services discovery and access at the per-producer/per-consumer level, which enables logical partitioning of CP into "common/dedicated slices".

Network slicing is also envisioned in RAN, aiming to better utilise the available physical resources, which are limited by the spectrum width. 5G slices overcome the QoS issue in 4G, which does not have the ability to perform E2E traffic isolation. While 4G QoS mechanisms can discriminate between different types of traffic but cannot discriminate and differentially treat the same type of traffic coming from different sources, 5GS, due to slicing, is able to do both at a deeper granularity (per user session). 5G network slicing also provides isolation between slices, thanks to radio partitioning in RAN. The allocation of radio resources to specific slices (scheduling) is achieved at two levels: radio spectrum allocation per slice (inter-slice) and specific scheduling within each slice (intra-slice). When resources are exhausted in a cell or in a slice, the pre-emption will ensure that each slice from a prioritised pool can have the minimal number of resources that were provisioned to it and that services with the highest priority are supported within a slice. There are mechanisms of congestion control – if a requested slice is overloaded, a new UE will be connected to a default slice.

RAN also supports the abovementioned mechanism of S-NSSAI. RAN slices are coupled to AMF instance based on S-NSSAI in the initial attach procedure, which may be provided either by UE or 5G CN. UE admission includes protection against unauthorised access to a slice. RAN Slicing supports service continuity over mobility; when a slice indicated in the handover request is not supported in the next cell, the service is handed over to a slice with the most similar QoS.

Role of SMF

The fundamental role of SMF is to provide the CP services that control UPF. The interaction between the SMF and UPF in 5GS involves coordination and communication to ensure proper management and handling of user data during sessions. Here are the key aspects of the SMF and UPF interaction:

- Session establishment – When a Protocol Data Unit (PDU) session is established, SMF determines the appropriate UPF for the session. It selects and assigns a specific UPF instance based on policy rules, QoS requirements, and other factors.

- Data routing – Once the session is established, SMF communicates with the assigned UPF to provide the necessary information for data routing. This includes sharing information about the PDU Session ID, QoS parameters, and other session-related details.

- Policy enforcement – SMF communicates policy rules and enforcement instructions to the UPF. It instructs UPF on how to handle user data packets, enforce traffic rules, apply QoS policies, and ensure appropriate security measures.

- User data forwarding – UPF performs the actual data forwarding and processing tasks based on the instructions received from SMF. It routes user data packets to the appropriate destination, applies policy-based traffic shaping or filtering, and manages data flows according to the defined QoS parameters.

- Dynamic control – SMF and UPF maintain ongoing communication to dynamically adjust the data handling based on changing conditions or events. This may involve updating policy rules, modifying QoS parameters, or responding to session-related events triggered by SMF.

- Session termination – When a PDU session is terminated, SMF instructs the assigned UPF to release the associated resources and terminate the data forwarding for that session.

The SMF and UPF interaction is crucial for ensuring proper management, control, and processing of user data during sessions in the 5G network, especially the continuity of user data flow (cf. the description of SMF involvement in handover procedures). Through their coordination, SMF defines policies and rules, while UPF implements them and performs the necessary data forwarding and traffic management tasks.

Mechanisms of handover

In the context of 5GS, there are two fundamental ways to manage handovers. The first way, known as "*Xn*-based" handover, utilises the *Xn* interface between the source and target Next Generation Radio Access Network (NG-RAN) BSs to manage the handover process. In this scenario, the *Xn* interface serves as the primary means of communication for coordinating the handover. The second case is referred to as the "*N2*-based" handover, where the *N2* interface between the NG-RAN BS and AMF is utilised to manage the handover. Here, the *N2* interface plays a crucial role in facilitating the communication and coordination between the NG-RAN BS and the AMF during the handover process. Below is a high-level description of NF roles in both approaches, highlighting the differences and similarities:

- *Xn*-based handover – the procedure is initiated and coordinated by the source BS in direct interaction with the target BS. Source BS is able to determine the target BS itself. 5GS CN (via the gateway, AMF) is informed about the actions taken and the need to switch the UP *N2* path. SMF manages the preparation and execution of UPF path reconfiguration to provide its continuity before and after the handover. After the successful handover, the target BS takes the responsibility of the UE serving, and the source BS releases its resources.

- *N2*-based handover – the procedure is initiated by the source BS, which is unable to determine the target BS within the same AMF area. It requests its source AMF to take the coordination role. The source AMF selects the target AMF and the latter requests SMF to prepare the UP switching. After the successful path preparation, the target BS is requested for handover preparation, and then the source AMF is notified about the completeness of the preparation. When the execution of handover is triggered by the source AMF, both source BS (directly) and target BS (via the target AMF) are activated to proceed. After the successful UE synchronisation to the new target BS, SMF is requested to coordinate the path switching execution in UPF. Then, as the target BS and AMF have taken the responsibility of the UE serving, the source AMF and BS are released.

The basic principles of the 5GS architecture are the maximum possible flattening of the communication network and the delegation of decision-making powers of the CP as close to the UE as possible in order to shorten and accelerate the decision-making path. Therefore, an *Xn* handover is preferred where the coordination takes place at the NG-RAN level, and if this is not possible then an *N2* handover coordinated by CN takes place. However, regardless of the coordination point of the handover procedure, the responsibility for controlling the preparation and execution of the UPF path for UE each time lies on SMF.

Scalability

Scalability is often referred to as the property of the network to keep its performance independent of its size. In the hitherto reality, especially related to the existence of a universal general-purpose mobile network (up to 4G), the only way to keep pace with the increase in service and traffic demand was the continuous, gradual expansion of network capacity through the installation of more or bigger network devices to be then gradually filled up until the next expansion cycle driven by their saturation. In view of the paradigm shift in favour of flexible delivery of individualised solutions and the growing importance of energy efficiency, such a development model cannot be sustained, and it also does not have to be thanks to new technologies. The need for scalability also has a strong business dimension: it is related to how easily a UC-supporting network solution can be upscaled or downscaled following the dynamically changing demand ("pay as you go" model) without overspending at one extremity or performance threat at the other.

The scalability of the 5G network is brought at different levels:

- Network softwarisation – decoupling of hardware and software enables functions portability to follow the demand in space; software load balancing mechanisms allow for flexible reaction (running more or less application instances) to the load;

- Network virtualisation/containerisation – the underlying hardware can be upscaled/downscaled without impact on upper layers; its life cycle does not impact the life cycle of the virtualised resources or applications consuming them;

- SBA of the 5G CN and high granularity of the functional decomposition – the softwarised functions can be run on demand to serve a specific demand (spatial, type of service- or network slice-related, etc.); mechanisms of NEF, NRF and SCP additionally support logical partitioning of the entire network;

- Network slicing – isolation and individual processing of a traffic fraction in a way tailored to UC requirements.

The mechanisms of scaling control, acting at different levels, have to be coordinated by the management layer to avoid their mutual competition leading to instability. A very important aspect of scaling is the dynamics – while the timescale of “old fashioned” hardware-based network extension was months, the virtualised infrastructure level scaling timescale is tens of seconds up to minutes. The unanswered challenge is still a dynamic slicing, i.e., delivery on demand or life cycle operations of the E2E network slice in a second’s timescale.

Integration of MEC and 5G

The 5GS can be integrated with the MEC technology. MEP communicates with a Data Plane (i.e., the underlying mobile network used for forwarding the MEC application traffic) via *Mp2* interfaces. In particular, through this way the MEP requests the network to redirect the traffic from the mobile terminal to the MEC application located at the Edge host. This is also the interface through which MEP can acquire the information to be further exposed to applications through MEP APIs related to bandwidth, radio interface, user location and identity.

From the point of view of 5GS, MEP is seen as a special case of AF within the 5GS CP, consuming the CP services. Consequently, the *Mp2* interface is equal to *Naf* according to the 5GS framework convention (cf. Figure 4-11). If MEP is trusted, it can communicate directly with other CP NFs. In a general case, it will be “gatewayed” by NEF. The Edge Application Server, where MEC applications reside, is seen within some local DN, different to the central one, through which the general user’s traffic is forwarded. Additionally, to support the mutual understanding between both frameworks, 5GS and MEC, the new CP NF has been introduced: Edge Application Server Discovery Function (EASDF) [83]. It should be noted that the cooperation of both has to be also assured, e.g., for coordination of terminal handover between BSs and MEC client at mobile terminal handover between MEC application instances located at different hosts. Additionally, integration with MEC provides an additional factor to increase the scalability of 5GS.

4.5.2 Key issues for the ETHER E2ESL and proposed solutions

Below, the 3GPP 5GS features have been validated against the requirements and basic considerations described in Section 3.2.

Genericity is supported by embracing standardisation, employing a modular architecture, using APIs, virtualisation, and supporting multi-domain operation, in the ETHER core network. This flexibility allows the network to accommodate different frameworks and technologies, making it adaptable to evolving industry requirements and advancements.

Multi-domain capability of the ETHER core network facilitates the integration, coordination, and seamless communication between different access (terrestrial, aerial, space) and functional domains (RAN, edge, core, transport). Integration of the ETHER core network with ETSI NFV MANO supports the multi-domain paradigm, facilitating the coordination, orchestration, and management of resources, services, and functions across diverse administrative or functional domains. Network slice subnet and NSSaaS are also essential components in the multi-domain paradigm of ETHER network. They enable the creation, coordination, and dynamic composition of network slices that span across multiple domains, allowing for collaboration among different stakeholders in the ETHER ecosystem.

Multi-provider functionality is fulfilled as a logical implication of the multi-domain paradigm. Resources provided by different service providers, operators, enterprises, or regulatory operators being virtualised are used to implement services within different domains. These domains refer to distinct entities or organisations that have administrative control over specific parts of the network, enabling collaboration and interoperability between different stakeholders.

Modularity is achieved by representing the individual network functions in the ETHER core network that are provided as independent services. Each network function service encapsulates a specific functionality, such as mobility management, session management, authentication, or policy control. By decomposing network functions into standalone services, the SBA achieves modularity, allowing functions to be developed, deployed, and updated independently.

Zero Touch Management and **Intent-based** principle in the ETHER core network can be achieved by service providers and network operators which consume provided APIs. NEF, NRF, and NWDAF are the functions to provide data for the exposed APIs. APIs form the foundation for building an automated, flexible, and efficient ETHER core network that operates with minimal human intervention and delivers mechanisms to continuously monitor the network and evaluate its performance against the defined intent.

Hierarchical approach refers to management organisation and is supported by the management functions stack (cf. the issue “Cross-layer management and orchestration according to 3GPP” in Section 4.5.1). CP and UP are designed as flat as possible to avoid unnecessary processing delays. However, the hierarchical approach to management by 3GPP enables fast-acting local mechanisms, while upwards the hierarchy, slower mechanisms can be used, but at a higher level of abstraction.

Distribution implies the dispersal of network functions, services, or resources across different locations or entities, aiming to optimise performance, resilience, and resource proximity. The 5GS SBA architecture with underlying virtualisation technology allows network functions, services, or resources to be distributed across different geographical locations or data centres to reduce latency, improve performance, ensure fault tolerance and resilience in case of failures, or enhance availability.

Scalability in 5GS is achieved both at the level of its architecture as Network Function Services (NFSs) are designed to be modular and horizontally scalable, enabling the addition or removal of instances on demand, and vertical scaling mechanisms supported by the underlying virtualisation technology.

Service-based principle is inherent due to the SBA architecture of 5GS CP, including the publish-subscribe, and request-response model, as well as the consumption of REST API-based interfaces and service discovery supported by NRF.

Openness to future extensions and trends embraces flexibility, modularity, the SBA principle of 5GS, virtualisation technologies and cloud-native principles enabling the seamless integration of new technologies and services (e.g., micro-service mesh) as they become available.

The validation above refers to the UP and CP of the ETHER E2ESL. For AP, compatibility will be implementation-specific.

In Table 4-1 below, the evaluation of the ETHER requirements supported by 5GS is provided, including the potential solutions, as well as for cases of gaps. The ETHER “definition of work” in the Grant Agreement states that “*Design of the ETHER architecture will be consistent with the management architecture for integrated satellite NR-RAT proposed by 3GPP in Release 17*”. However, the evaluation below versus the 5GS architecture (cf. TS 23.501, relevant versions) also includes the currently existing 3GPP solutions in Release 18 as well as visions of the future scope of Release 19.

Table 4-1: Evaluation of the ETHER requirements with impact on the ETHER E2ESL

Identifier	Requirement	Key issue for E2ESL architecture	Potential solution
ETH-REQ-UC1-DT-01	Intermittent – scheduled contacts	UE should be notified by about a contact schedule.	From LEO/MEO system OSS to UE through AP.
ETH-REQ-UC1-DT-02	Intermittent – opportunistic contacts	UE should detect the presence of LEO/MEO satellite and request a contact.	Standard 5GS mechanism; presence sampling period vs. battery saving to be determined.
ETH-REQ-UC1-DT-04	Congestion and flow control	1. There should be included mechanisms of flow control in UE-satellite and satellite-ground links to avoid congestion and data loss (including reception validation and retransmission). 2. Mobility management mechanisms should include redirection of DL data during the handover procedure. 3. In case of LD LEO constellation, “no interruption” feature may not be possible, but the next approaching satellite to serve the specific UE is known due to the defined orbit, so the DL data for the UE can be forwarded from the source satellite (losing a contact with UE) to the target satellite (about to catch a contact with UE) in advance together with UE context.	1. Flow control through AP to keep the UE genericity. 2. Standard handover procedures include necessary DL data redirections. 3. No such feature in 3GPP Release 17, development is needed (update of UPF, SMF, AMF, gNB functions).
ETH-REQ-UC1-DT-06	Connection discontinuity	1. Service link (UE-satellite) should be resistant to temporary interruptions (loss of contact). 2. ETHER E2ESL should be immune to temporary system partitioning due to loss of feeder link (satellite-ground).	1. Flow control/AP mechanisms and gNB presence detection by UE (standard 5GS mechanism). 2. 5GS may be distributed over the terrestrial and non-terrestrial strata.

Identifier	Requirement	Key issue for E2ESL architecture	Potential solution
ETH-REQ-UC1-DT-07	Store and Forward	ETHER E2ESL should support Store and Forward functionality.	No such mechanism supported in 3GPP Release 17 (Store and Forward mechanism in 3GPP Release 19 requirements study; update of UPF, SMF, AMF functions). Development is needed in these areas.
ETH-REQ-UC1-DT-09	Mobility management	1. Integration of NTN with TN needs dynamic reallocation of gNBs in motion (LEO/MEO) between Tracking Areas (TAs) to ensure permanent mapping of geographical areas and TAs. Satellites moving over successive countries should be dynamically connected to TNs in these countries and the relevant TAs. 2. Cf. ETH-REQ-UC1-DT-01.	1. Such mechanism is supported in 3GPP Release 17 (Tracking Area handling for NR satellite access), which ensures that each TA is Earth-stationary even if the radio cells are moving across the Earth's surface. 2. From LEO/MEO system OSS to UE through AP or CP (*N1* interface, either access or non-access stratum).
ETH-REQ-UC1-DT-10	Support for different services	ETHER E2ESL should support integration of different RATs with the same CN (e.g., 4G and 5G).	The integration of different 3GPP RATs with the same CN implies the non-SA 5GS architecture where network slicing is not supported, according to the 3GPP approach. The requirement contradicts with the general network slicing support expectation as well as ETH-REQ-UC3-FN-01.
ETH-REQ-UC1-SE-01	Sample processing	ETHER E2ESL should support both access to data contained in all its planes and ability to influence them back.	CP provides a generic mechanism with NEF to be used with all 5GS CP databases. Other impact on other CP NFs may require their functional upgrade. Sampling and return impact on UP data needs inclusion of additional processing functions to UPF, e.g., deep packet inspection, buffering, etc. In the case of AP, the means will be UC- and implementation-specific. These mechanisms will be exposed to other ETHER layers.
ETH-REQ-UC1-SE-02	Joint sample and transmit	Cf. ETH-REQ-UC1-SE-01	Cf. ETH-REQ-UC1-SE-01
ETH-REQ-UC1-SE-03	Support for E2E information handling beyond the sample and transmit	Cf. ETH-REQ-UC1-SE-01	Cf. ETH-REQ-UC1-SE-01
ETH-REQ-UC1-SE-04	Content caching	Cf. ETH-REQ-UC1-SE-01	Cf. ETH-REQ-UC1-SE-01
ETH-REQ-UC2-FN-01	Migrate TN to NTN	E2ESL should support handovers between 3D strata.	3GPP Release 17 supports *Xn* and *N2* handovers; no need of new CP functions identified; however, exposure of AMF and gNB for impact of external logic (AI-driven) may need additional development.
ETH-REQ-UC2-FN-02	Vertical handover	Cf. ETH-REQ-UC2-FN-01	Cf. ETH-REQ-UC2-FN-01

Identifier	Requirement	Key issue for E2ESL architecture	Potential solution
ETH-REQ-UC2-NF-01	Vertical handover	ETHER E2ESL should support handovers that are autonomous and imperceptible for the end user.	3GPP Release 17 supports *Xn* and *N2* handovers without involvement of the end user; no need of new CP functions identified; however, exposure of AMF and gNB for impact of external logic (AI-driven) may need additional development, as AI-driven handover prediction may be crucial for providing near to zero interruption time during vertical handovers.
ETH-REQ-UC2-NF-02	Broadband	ETHER E2ESL should support broadband communication service and TN-NTN handovers.	3GPP Release 17 supports *Xn* and *N2* handovers (cf. ETH-REQ-UC2-FN-01); in case of NTN sufficient gNB capacity has to be provided.
ETH-REQ-UC2-NF-03	Coverage	ETHER E2ESL should provide 100% coverage.	No direct impact on E2ESL; for feasibility, TN/NTN integration is needed.
ETH-REQ-UC3-FN-01	RAN in TN, HAPS, and SAT	ETHER E2ESL should provide a unified RAT to be exploited by unified UEs.	Unified RAT in 3D strata needed.
ETH-REQ-UC3-FN-02	Open 5G CN	ETHER E2ESL should provide APIs for trusted external systems.	3GPP Release 17 supports NEF/AF functionality in 5G CP; in case of RAN, OSS should expose NBI or O-RAN with NBI and x-App/r-App may be exploited.
ETH-REQ-UC3-FN-05	Multilink functionality	ETHER E2ESL should provide connectivity resilience (e.g., redundant connectivity links).	3GPP Release 17 supports dual connectivity/multi-connectivity; its compatibility with NTN integration should be validated; however, no need of new CP functions identified. A compatible UE and traffic control software has to be used.
ETH-REQ-UC3-NF-03	Handover reliability and delay	ETHER E2ESL should provide seamless vertical/horizontal handover mechanisms that include proactive approach in minimising the interruption time.	If the target gNB can be selected, the session continuity is provided by a principle in 3GPP Release 17. For the other aspects, cf. ETH-REQ-UC2-FN-01 and ETH-REQ-UC2-NF-01.
ETH-REQ-UC3-NF-04	3D network programmability	ETHER E2ESL should provide REST APIs for interconnection with other layers.	In 3GPP Release 17, both internal CP bus and APIs exposed through NEF REST-based. gNB-RAN OSS interface – seems to be left by 3GPP for implementation; O-RAN interfaces are subject to O-RAN Alliance standardisation (if no REST API at specific reference point is available, a mediation module development may be necessary).
GENERAL-FN-01	Detection of UE NTN gNBs attach capability	ETHER E2ESL should detect the UE capability to be served by NTN to	In 3GPP Release 17, no such mechanism is defined. 5G-Equipment Identity Registry

Identifier	Requirement	Key issue for E2ESL architecture	Potential solution
		prevent infeasible mobility management.	(EIR) could be involved during the Registration/Attach procedure to label a registered UE as NTN-capable. gNB and AMF should be updated for compatibility.
GENERAL-FN-02	Detection of UE non-stationary NTN gNBs attach capability	ETHER E2ESL should detect the UE capability to be served by the non-stationary NTN gNBs (e.g., Doppler shift-proof waveforms) to prevent infeasible mobility management.	In 3GPP Release 17, no such mechanism is defined. 5G-EIR could be involved during the Registration/Attach procedure to label a registered UE as NTN-capable. gNB and AMF should be updated for compatibility.
GENERAL-FN-03	Detection of TN-preferred UEs	ETHER E2ESL should detect the UEs with preference of being served by TN access, however able to be served by NTN access.	In 3GPP Release 17, no such mechanism is defined. 5G-EIR could be involved during the Registration/Attach procedure to label a registered UE as NTN-capable. RAT preference mechanism is not supported, gNB and AMF should be updated for compatibility.
GENERAL-FN-04	Detection of NTN subscription	ETHER E2ESL should be aware of UE NTN subscription to prevent mobility management not allowed by subscription.	In 3GPP Release 17, mechanisms of forbidden area and service area restrictions can be used for NR satellite access. The mobility restrictions can be enforced based on NTN RAT type identifiers defined by 3GPP: NR(LEO), NR(MEO), NR(GEO), and NR(OTHERSAT).
GENERAL-FN-05	SDN-based transport	ETHER E2ESL should be able to control the UPF continuity in an SDN-based transport.	SMF should interact with the SDN CP via the *N4* interface [84]. The need of SMF update for compatibility needs further validation.

4.5.3 ETHER E2ESL architecture

The ETHER E2ESL comprises AP, UP, and CP, which implement the Application, Communication, and Steering strata, respectively (cf. Figure 4-1). UP and CP are implemented with the 3GPP 5G+ network, according to the ETHER assumptions based on the analysis of the 3GPP 5GS framework architecture and its cross-verification with the ETHER requirements presented in Sections 4.5.1 and 4.5.2, respectively. All these planes are orchestrated and managed by the ETHER MANO Layer supported by the ETHER AI Layer. It should be noted that the 5G+ network CP is not an exclusive implementation of Steering Stratum. SDN CP, outside of the perimeter of 3GPP 5GS, also implements it.

The roles of all planes composing E2ESL are as follows:

- AP hosts the application-level services. It is localised on the UE side (the UE application or application implemented in a local system that includes UE) and on the application platform, which is, in general, a distributed solution, hosting a distributed application server. The application platform is embedded in the ETHER system's shared resources. Moreover, it is assumed that AP can consume the services exposed by the ETHER AI Layer, e.g., exploit AI models for predictive analytics to optimise the applications' performance. The detailed AP architecture will, however, depend on the specificity of the solution.

- UP provides connectivity service for AP with predefined QoS guarantees (latency, bandwidth, reliability, etc.). The UP operations include application-specific data forwarding and processing (e.g., deep packet inspection, traffic filtering, parental control, etc.), also (and especially) in terms of relative mobility of terminals and networks, since both UE mobility and radio access node mobility are supported in the ETHER system.

- CP provides all necessary mechanisms to provide the UP services for the communication services user. These control mechanisms run autonomously or can be exposed through a CP API.

The simplified ETHER E2ESL architecture is presented in Figure 4-13. The AP interactions are served by the 5G+ UP composed of a service-specific UPF, which is a chain of atomic virtualised functions distributed in the virtualised infrastructure (not shown in the picture except for SDN-based transport providing the transport layer for the functions composing the UPF chain, so the 5G+ UPF connections can go down to the SDN-based transport and back several times).

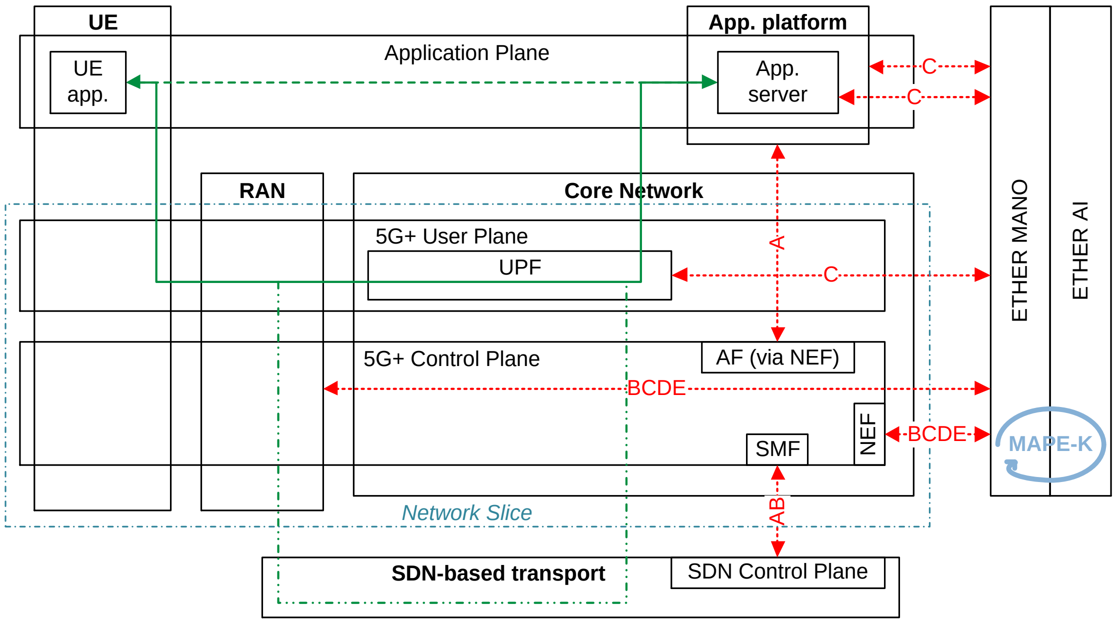

Figure 4-13: ETHER E2ESL architecture and support of specific ETHER mechanisms

Selected ETHER E2ESL entities are indicated in Figure 4-13 to emphasise their role in the support of fundamental ETHER mechanisms, namely:

A. MEC & caching: the application platform (e.g., MEC), integrated as AF, requests via NEF UE-related data from the CP services or redirection of specific UE traffic fraction to the application server. The latter request is further processed by SMF and sent to the SDN CP in SD-WAN through the *N4* interface [84] (*N4*-SDN CP NBI mediator to be potentially developed). In the case of the caching mechanism, it can be embedded in UP only or also use AP. For the latter case, an interaction with SDN CP may also be used. The SDN CP interfaces can be accessed either directly or via MSDNO/SDNO APIs. The deployment configuration and potential benefits of both approaches will be investigated within WP3 activities.

B. Seamless vertical/horizontal handover: the AI-based handover logic controls upgraded AMFs via NEF and upgraded gNBs in RAN (cf. proposed solution for ETH-REQ-UC2-NF-01 in Table 4-1). The details of the handover control solution will be defined by WP3.

C. Semantic-aware data analytics: the interaction of semantic data analytics is provided with all planes: both application server and platform in AP, dedicated entities in the UPF chain (to be further studied by WP3), and relevant core CP functions via NEF or directly from gNBs in RAN (here, the data may also be acquired via RAN OSS). At this stage, no need for interaction via SMF with SD-WAN CP in the context of UP analytics has been identified.

D. AI-based predictive analysis: the data for analytics are exposed by core CP functions via NEF or directly from gNBs in RAN (here, the data may also be acquired via RAN OSS). Defining the details of the analytics solutions will be performed within WP3 activities.

E. AI-enabled E2E network performance optimisation: the data for analytics are exposed by core CP functions via NEF or directly from gNBs in RAN (here, the data may also be acquired via RAN OSS), and the control feedback may go exactly the same way (RAN OSS mechanisms will probably be used for actuation in RAN). The decision about either direct interaction with core CP NFs or via NSSMF/NFMF (cf. Figure 4-7) is implementation-dependent (to be further investigated by WP4).

It should be emphasised that the mutual relations/interactions of the ETHER MANO and AI layers will be further investigated by WP4. The ETHER E2ESL exposes data and mechanisms that can be used by both layers.

Interaction between CP of E2ESL and ETHER MANO Layer – these specific interactions will follow the 3GPP approach presented in Figure 4-7, i.e., utilisation of the interfaces standardised by the ETSI NFV group at the ETSI NFV MANO reference points *Os-Ma-nfvo*, *Ve-Vnfm-em*, and *Ve-Vnfm-em*. The issue of whether these interfaces need additional modifications is for further study in cooperation with WP3 and WP4.

4.5.4 Distributed 3GPP architecture for supporting NTN-based IoT applications

In recent years, advancements in IoT technologies have opened up new markets and enabled new UCs. However, these services are currently only viable in cities and densely populated areas, as it is where cellular network coverage is available. Unfortunately, rural and offshore regions, encompassing 85% of Earth's surface, lack cellular network coverage. This limitation hinders technological and social progress in rural communities, inhibits new business opportunities, and especially impacts the potential of massive IoT applications.

To fully harness the potential of new technologies in massive IoT achieving global connectivity is crucial. NTNs that are based on satellites are vital in delivering the required connectivity. While existing global connectivity solutions are not cost-efficient, efforts within the 3GPP have aimed to standardise cellular NTN based on established terrestrial cellular technologies like NR, enhanced Machine Type Communication (eMTC) and NB-IoT. Nevertheless, in Release 17, 3GPP has focused on transparent payload architectures, requiring complex megaconstellations.

Delay-tolerant IoT applications, such as smart agriculture, livestock monitoring, asset tracking, and environmental monitoring, using NB-IoT can benefit from the deployment of LD LEO satellite constellations. These constellations offer small complexity and regenerative payloads, allowing operations within the satellite and diverse services directed by a MANO system. This approach significantly reduces costs and enhances interoperability. However, it presents challenges, as using an LD satellite constellation introduces service link discontinuities.

Moreover, the feeder link connecting the LD satellite constellation and GSs is accessible only at a limited number of locations. Note that a LD satellite constellation yields revisits times in the magnitude of hours and visibility windows in the magnitude of minutes [85]. This justifies the adoption of a regenerative payload architecture that supports a Store and Forward mechanism. This mechanism is indispensable for both UP and CP and will enable ETHER's flexible payload from UC1 to provide NB-IoT coverage to delay-tolerant applications. Implementing these modifications necessitates adapting the 3GPP standard procedures [86].

Although NB-IoT was introduced in the 3GPP's Release 13 and co-existed with LTE networks, it is considered a 5G technology by the 3GPP because it meets the 5G requirements for mMTC. The 3GPP has ensured that NB-IoT continues to evolve as part of the 5G specifications, enabling it to coexist with other 5G technologies like eMBB and URLLC.

In ongoing discussions for 3GPP Release 19 specifications, NB-IoT is evolving by leveraging existing EPC architecture and by exploring options offering support to a complete 5G core [87], [88]. This provides mobile operators with backward compatibility and allows for a gradual transition towards a complete 5G network.

Enabling NB-IoT services from an LD satellite constellation entails a series of challenges. Sustaining a functional service link independently of GS connectivity is a primary requirement. This entails enabling discontinuous NB-IoT backhauling to ensure satellite service availability even when not connected to GSs via the feeder link. Additionally, to make the satellite constellation deployments cost-efficient, it is essential to establish standard 3GPP interfaces that allow multiple service providers (e.g., MNO) to utilise the same LEO constellation and extend their service coverage (roaming). To tackle these challenges, there is a growing focus on developing a distributed 3GPP architecture. This architecture, which is currently under standardisation in Release 19, is built upon regenerative payload and the "Store and Forward" principle, as proposed in the ETHER's UC1.

The main challenges posed by LD LEO constellations with discontinuous service and feeder link arise from the presumption of uninterrupted connectivity among the core components of the mobile network. Procedures such as Attach/Detach, (periodic) Tracking Area Update, data transmission/reception, Service Request, and Paging require modification. This need for adaptation is largely due to signalling timers that oversee mobility and sessions. These timers exist on both the UE and network side, with the signalling timer (in the order of minutes) for Non-Access Stratum (NAS) procedures being particularly critical. Thus, these procedures must be completed within a single UE-satellite visibility period.

The following subsections present distributed 3GPP architecture solutions in ongoing discussions in Release 19 that aim to address these challenges.

4.5.4.1 Evolved Packet Core (EPC) distributed architecture

The challenges to the Evolved Packet System (EPS) arising from discontinuous service and feeder link using Store and Forward can be summarised as follows:

- Completing NAS signalling within a few minutes in a single UE-satellite visibility window.

- Ensuring security and subscription information are available on the satellite to execute NAS procedures.
- Conducting EPS Connection Management (ECM) procedures, service requests and paging.
- Dynamic organisation of TAs.
- Broadcasting of ephemerides to end devices.
- Enabling UE context sharing for NAS signalling support via multiple satellites in a constellation.

To address the challenges described above, a distributed architecture compatible with the 3GPP has been proposed as a solution [89]. This concept, depicted in Figure 4-14, is classified as distributed due to the arrangement of its components in distinct layers, with the NTN serving as the novel key element.

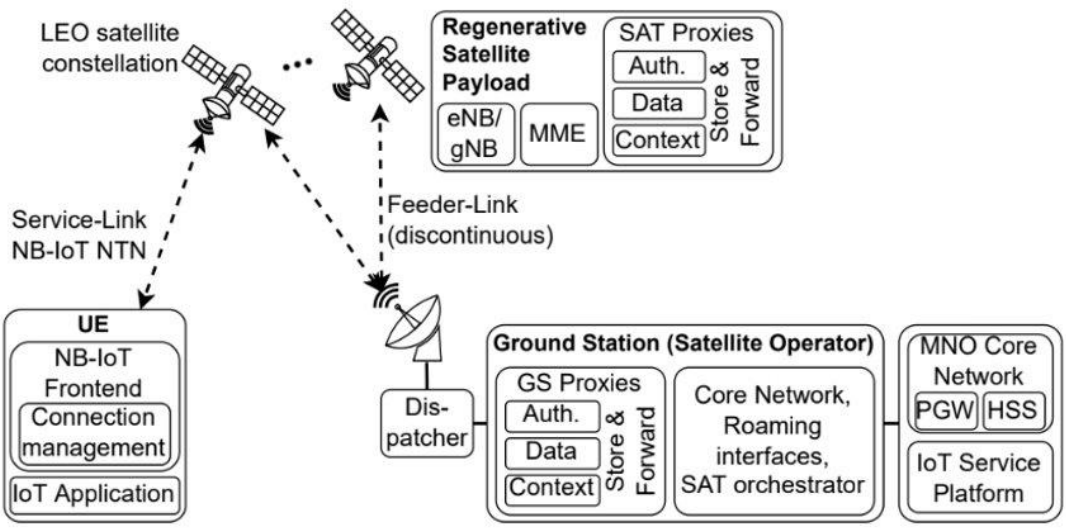

Figure 4-14: Evolved Packet Core distributed 3GPP architecture

In a distributed 3GPP architecture [89], components of the Evolved Packet Core (EPC), including the eNB/gNB and Mobility Management Entity (MME), reside within the regenerative satellite payload. The eNB/gNB is responsible for radio access, while the MME facilitates the timely completion of the NAS signalling process within a single UE-satellite visibility window. Additionally, the regenerative satellite payload incorporates novel proxies that support the Store and Forward functionality. These proxies are categorised into three distinct types: the Authentication proxy, which manages the attach procedure across different visibility windows; the User Data proxy, tasked with overseeing Mobile-Originated (MO) and Mobile-Terminated (MT) traffic; and the User Context proxy, entrusted with the dissemination of UE context among satellites.

Likewise, the GS incorporates analogous Store and Forward mechanism, implemented through proxies that effectively establish connections to the proxies aboard the satellite, enabling the relay of information. Other elements of the EPC are hosted on the ground and provide standard roaming interfaces to other MNOs, and connectivity to the IoT Service Platform.

The UE contains the NB-IoT Frontend to connect to the LEO satellite constellation, alongside the connection management that coordinates the IoT device states (idle, connected, sleep) with the information received about the satellite constellation ephemeris. The IoT application is a delay-tolerant application interchanging data messages.

4.5.4.2 5G and beyond architecture

In the NG-RAN, new interfaces and protocols are being integrated to support NTNs. An NTN platform can function as either a space mirror or a gNB (next-generation NodeB) in space, leading to two possible satellite-based NG-RAN architectures: transparent and regenerative. In regenerative architectures, the NTN platform may implement partial or full gNB functionality, depending on the functional split of the gNB. Additionally, NTNs can be classified by access type: direct access, where the NTN terminal is directly served by the platform (satellite access architecture), or relay-like architecture, where communication occurs via a relay node [90].

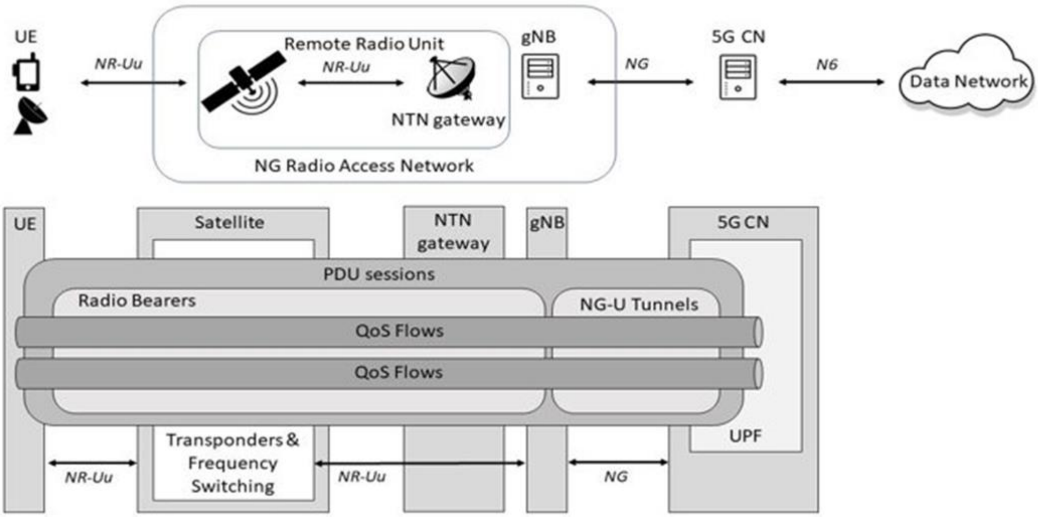

Figure 4-15: Transparent-based payload satellite [90]

There are different satellite-based architectures for NTN in 5G communication [90]. The transparent satellite architecture relays NR signals between the NTN gateway and terminal, using the same Satellite Radio Interface (SRI) for feeder and service links, with multiple satellites potentially connected to a single ground-based gNB (shown in Figure 4-15). The regenerative satellite architecture (shown in Figure 4-16) features on-board processing to handle NR signals directly, with the *NR-Uu* interface on the service link and the *NG* interface connecting the NTN platform to the 5G CN via the NTN gateway. Additionally, ISLs are the transport links between NTN platforms.

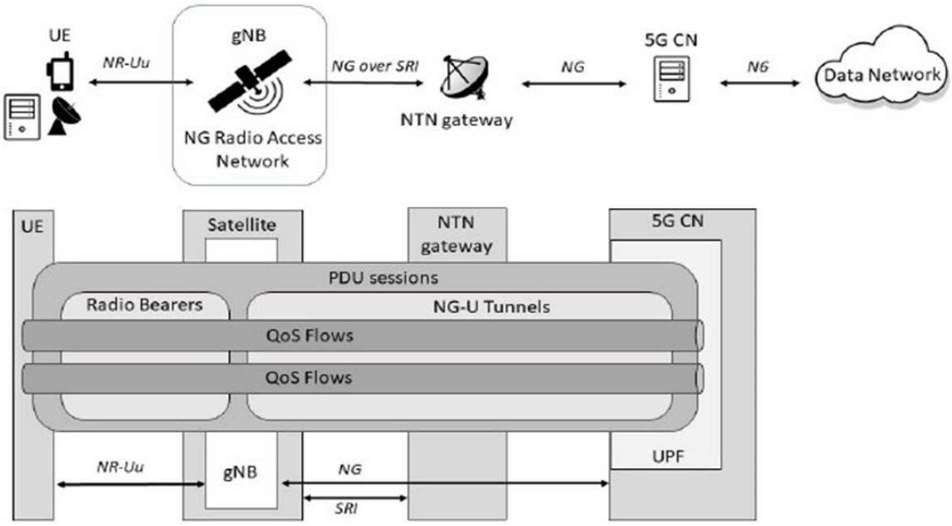

Figure 4-16: Regenerative-based payload satellite [90]

The NG-RAN architecture includes a gNB comprising a central unit (gNB-CU) and one or more distributed units (gNB-DU). A “5G NR friendly” NTN architecture (shown in Figure 4-17), featuring a regenerative satellite, involves the gNB-CU on the ground connecting via the *F1* interface over SRI to the NTN platform, which serves as a gNB-DU. The radio interface *NR-Uu* links the NTN terminal and the gNB-DU on the satellite, while the *NG* interface connects the ground-based gNB-CU to the 5GC. Multiple gNB-DUs on different NTN platforms can be linked to the same ground-based gNB-CU.

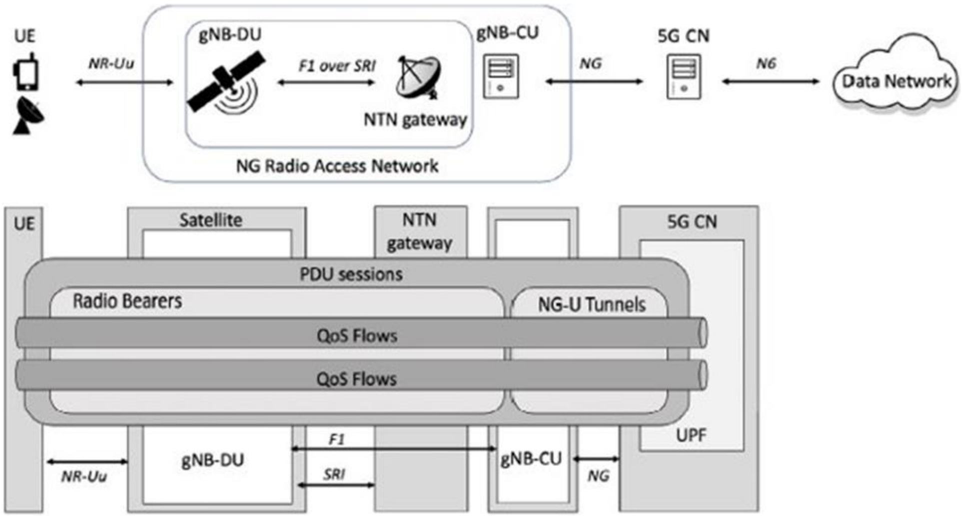

Figure 4-17: Regenerative satellite-based gNB-DU [90]

Various relay-like architectures are possible for NTN in 5G communication [90]. Figure 4-18(a) illustrates the access network forwarding the NR signal to the NTN terminal via a relay node from a transparent payload-based satellite. Figure 4-18(b) and Figure 4-18(c) depict a regenerative payload-based satellite, incorporating parts or the complete gNB, with the relay node forwarding the NR signal received from the satellite to the NTN terminal.

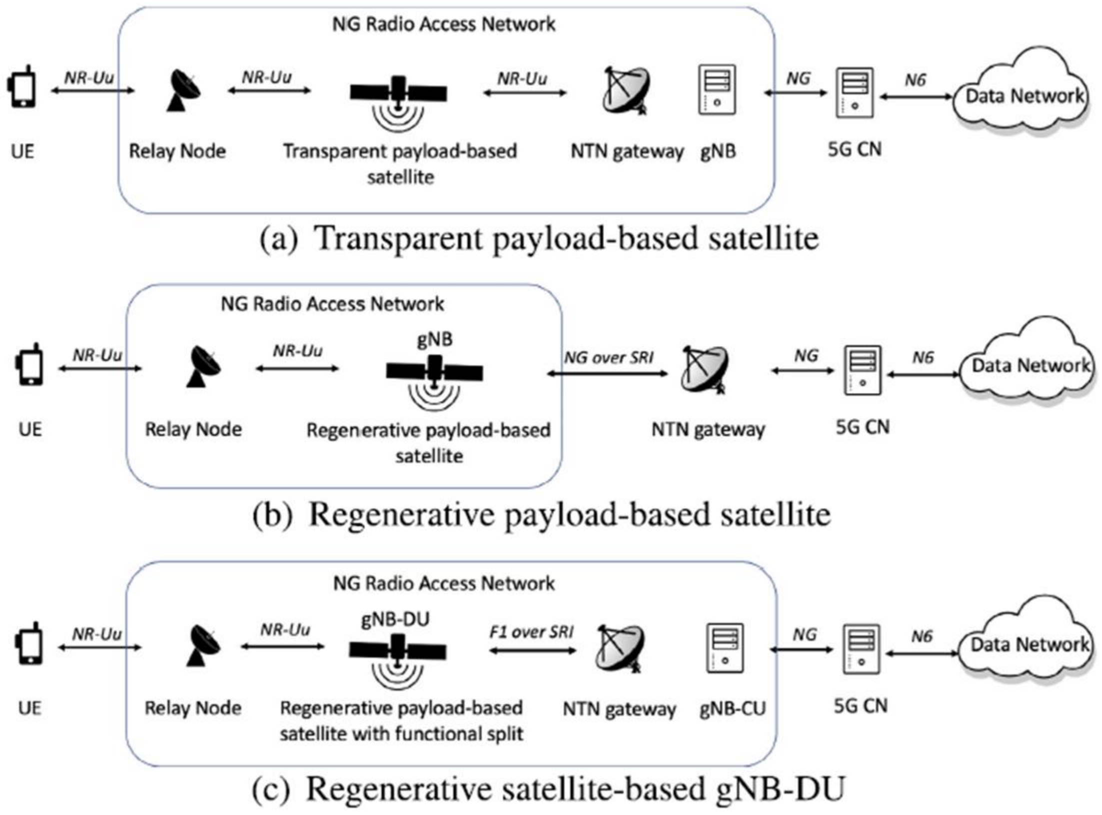

Figure 4-18: Relay-like architecture [90]

In Figure 4-19, we illustrate the distributed 5G CN between the satellite and the ground for Store and Forward Satellite operation. This architecture includes three main entities within the 5G CN: AMF, UPF, and SMF. The external interfaces of the 5G CN align with these combined entities' interactions with other network components. Figure 4-19 demonstrates a possible distribution of core network functions between the satellite (5G CN-SAT) and the ground (5G CN-GND). The interface between 5G CN-SAT and 5G CN-GND, referred to as the Store and Forward interface in Figure 4-14, may vary based on implementation specifics. The 5G CN functions onboard the satellite include AMF functions for handling the *N2/N3* interface with the onboard gNB and terminating NAS protocol signalling to/from UEs via the onboard gNB.

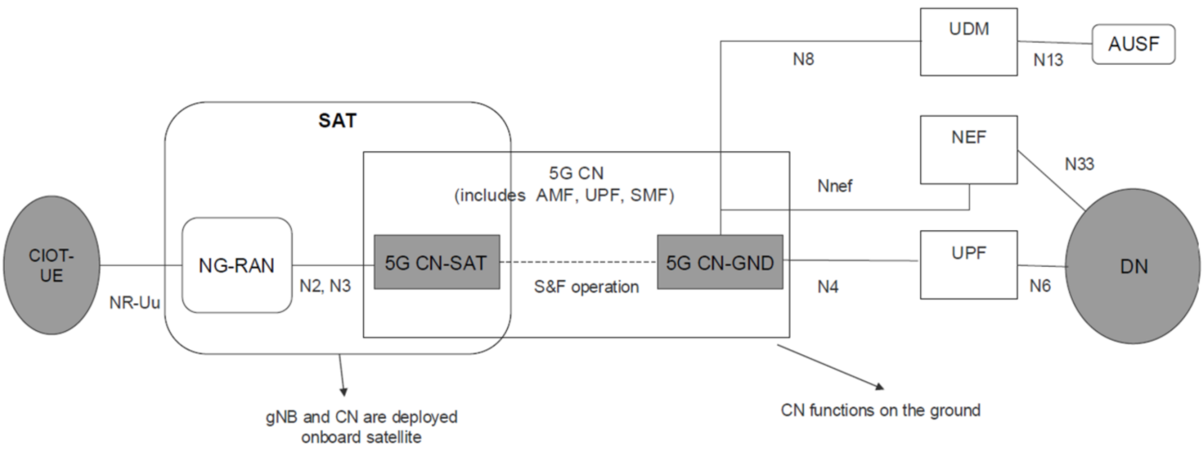

Figure 4-19: Proposed distribution of 5G CN functions between the satellite and the ground for Store and Forward Satellite operation

To support Store and Forward Satellite operations, the proposed solution relies on specific architectural assumptions. The functions of the 5G CN are divided between the satellite (5G CN-SAT) and the ground (5G CN-GND). Termination endpoints for *N2*/*N3* and NAS protocols are located on the satellite as part of the 5G CN-SAT, while endpoints for *N8*, *N4*, *N6*, and *Nnef* interfaces remain on the ground as part of the 5G CN-GND. This functional split also applies to multi-satellite scenarios, where a single instance of 5G CN-GND can interact with multiple 5G CN-SAT instances on various satellites. The interface between 5G CN-SAT and 5G CN-GND is implementation-specific, allowing flexibility based on specific constellation and network configurations.

One possible implementation using the Store and Forward functionality for the 5G core, is the split of core functionalities, where the regenerative payload of the satellites is equipped with gNB radio access capabilities and additional necessary components to complete data and signalling exchanges within the same satellite visibility window. This is possible by positioning the AMF on the satellites, which enables the completion of NAS procedures and interactions within a limited timeframe when a satellite is visible to a UE, as shown in Figure 4-20.

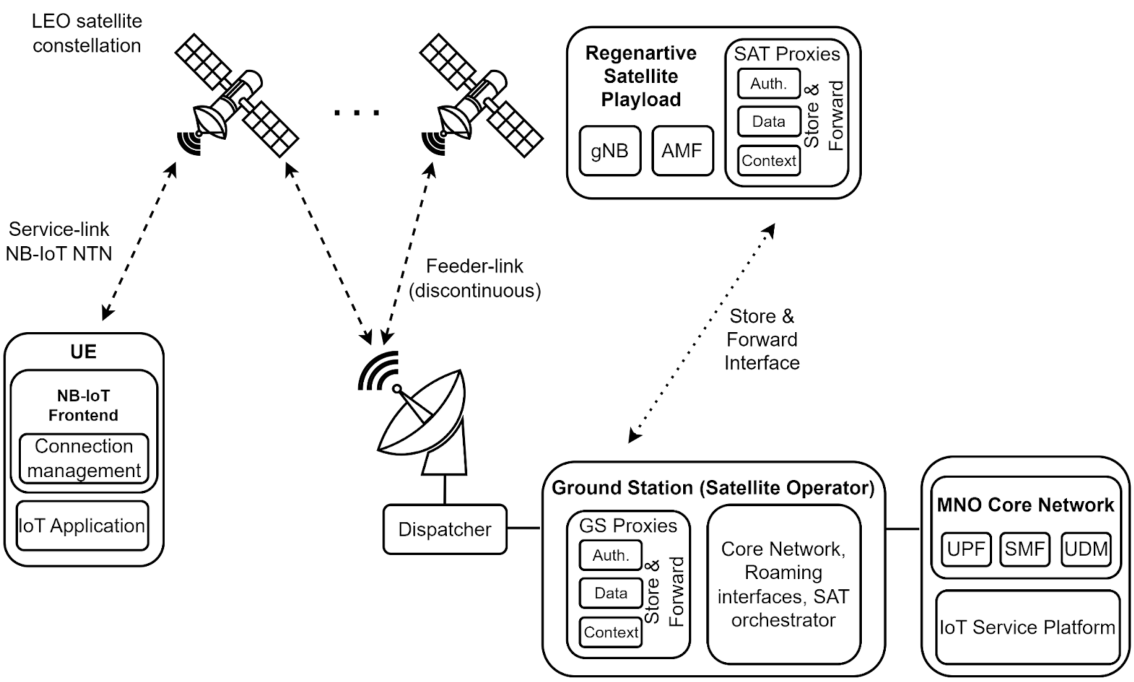

Figure 4-20: Distributed 5G CN

4.5.5 Store and Forward mechanism

The ongoing work in the 3GPP's Release 19 for integrating satellite components into the 5G architecture [55] presents various solutions for the Store and Forward mechanism implementation, also considering both the EPC and the 5G CN.

4.5.5.1 EPC support

Integrating NB-IoT with NTNs takes advantage of existing 4G infrastructure to provide widespread and cost-efficient IoT connectivity. This approach is crucial for ensuring the rapid deployment of IoT services in areas without TN coverage.

4.5.5.1.1 Advertising capabilities and Store and Forward attachment procedure

For Store and Forward operation, the core network needs to notify the UEs when the network is operating in Store and Forward Satellite mode to prevent them from requesting unsupported services. A satellite cell may even switch between Store and Forward mode and normal mode based on network policies (e.g., in areas where the satellite can connect to the ground network simultaneously, it may remain in Store and Forward mode or switch to normal mode).

The attach procedure using Store and Forward mechanism involves at least four steps, illustrated in Figure 4-21. The signalling message exchange and the functional entities involved during these four steps are depicted in Figure 4-22.

Step 1: During the initial UE-satellite contact, the UE enters cell coverage and transmits an Initial Attach message to the satellite MME. The MME queries the onboard entity responsible for Store and Forward functionality, the *S6a* Auth Proxy SAT in Figure 4-22. This proxy stores the International Mobile Subscriber Identity (IMSI) for the Authentication Information Request (AIR) and also for the Update Location Request (ULR) to be resolved when in contact with the GS. As no information is available for the UE onboard the satellite, the MME sends an Attach Reject message to the UE. A proposed solution is the use of a new rejection code that indicates that the Attach Request is being processed and that the response will be available in the next satellite pass [89].

Step 2: The second step encompasses the AIR and ULR, after the *S6a* Auth Proxy SAT and *S6a* Auth Proxy GS proxies have exchanged information using the feeder link. This implies that the authentication information and update location is performed in a single step and stored in the proxies.

Step 3: The vectors saved in the *S6a* Auth Proxy GS are subsequently forwarded to the *S6a* Auth Proxy SAT on the satellite. This proxy is now prepared to authenticate the UE. This satellite can be the same one involved in the first step or another satellite within the constellation. Both steps 2 and 3 can occur within the same contact between the satellite and the GS.

Step 4: The UE comes into satellite cell coverage and resends the Attach Request message. The onboard MME sends the AIR to the *S6a* Auth Proxy SAT which responds with the Authentication Information Answer (AIA) message. This allows the UE and MME to proceed the mutual authentication, resulting in the UE being successfully registered with the network. Additionally, the MME resolves the ULR message with the *S6a* Auth Proxy SAT. The proxy utilises information from the Update Location Answer (ULA) to allocate an IP address to the UE and establish an EPC bearer, as illustrated in Figure 4-22.

Co-funded by the European Union

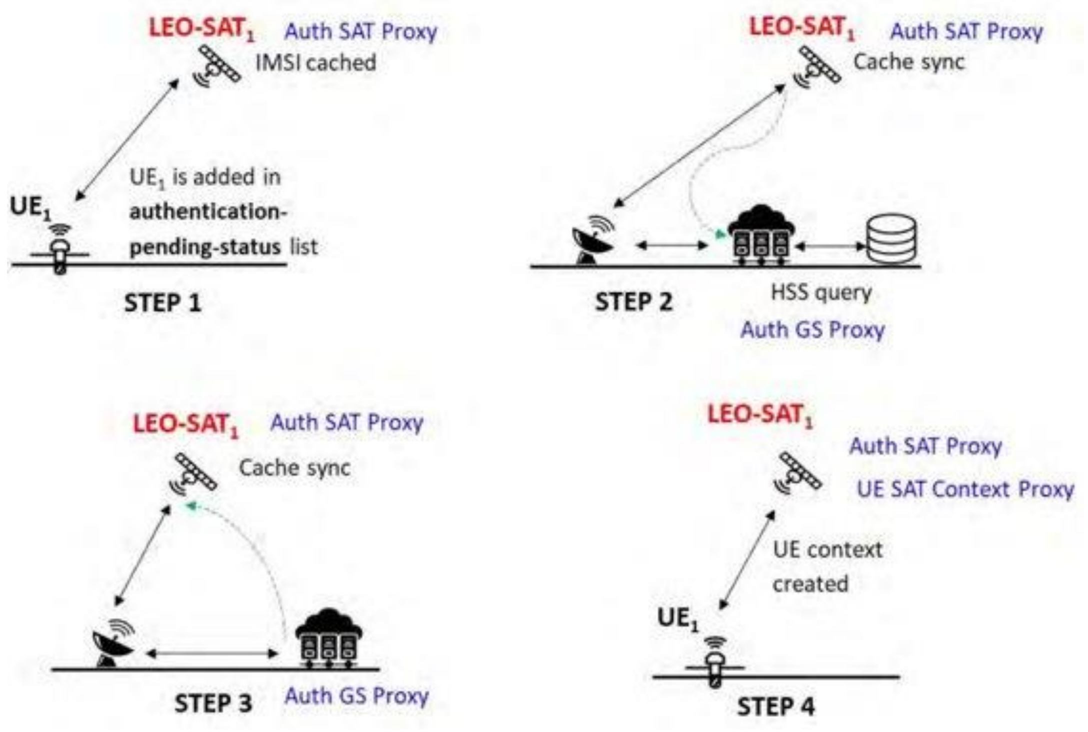

Figure 4-21: Steps required for the UE attach procedure with service and feeder link discontinuity and Store and Forward mechanism with one satellite

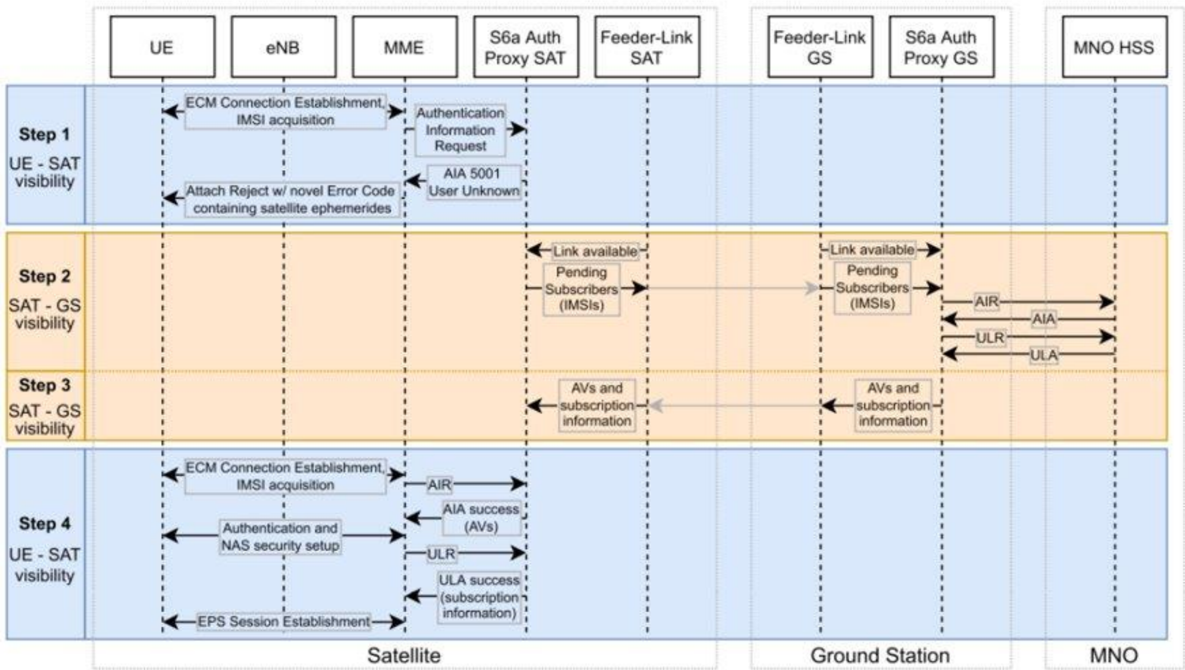

Figure 4-22: Signalling between UE, satellite and GSs presenting the four steps required for attachment procedure with service and feeder link discontinuity and Store and Forward mechanism with one satellite

4.5.5.1.2 Attachment procedure in Store and Forward multi-satellite deployment

Ensuring that the Store and Forward operation mode can effectively support multi-satellite deployment scenarios by enabling a UE to be potentially served by multiple satellites is essential. Multi-satellite support is critical for reducing the latencies in data transfer delivery, given that the revisit time of a single satellite may be around 12 hours or longer for typical orbits [55]. In this context, the distributed architecture previously described is also applicable to a multi-satellite scenario, where a single instance of GS proxies on the ground could interact with multiple instances of SAT proxies on board the satellites (cf. Figure 4-20).

A Store and Forward core network deployed over multiple satellites must be capable of initiating or suspending procedures in one satellite and resuming or terminating them in another satellite on the same Public Mobile Land Network (PLMN) in order to benefit from reduced latencies [55]. For that, a solution considers the following additional elements:

- Means for the UE to advertise the network that it supports Store and Forward.
- New reject causes and backoff timers for terminating/suspending the Non-access stratum (NAS) procedures and assisting the UE on subsequent re-attempts in the same or other satellite cells of the same PLMN.
- Means for the networks to communicate to the UE in which satellites user data transfer is allowed.

The attachment procedure for a multi-satellite scenario is illustrated in Figure 4-23, considering three different satellites (SAT#i, SAT#j, SAT#k) and a Home Subscriber Server (HSS) with subscription information on the ground [55]. This example shows that the attach procedure can start with one satellite and finish with another, and user data transmission can occur through any of the three satellites.

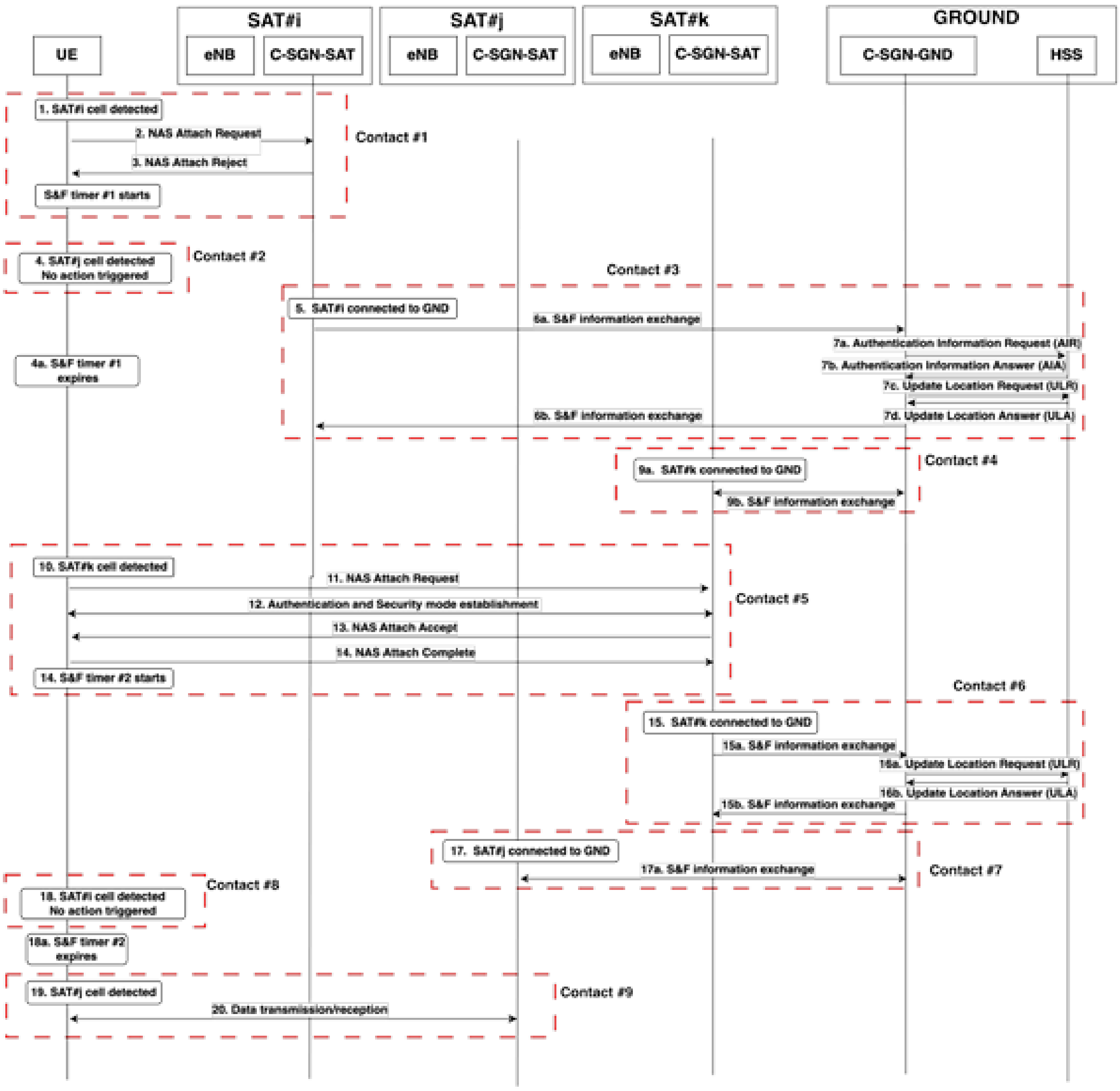

Figure 4-23: Evolved Universal Terrestrial Radio Access Network (E-UTRAN) Initial Attach procedure under Store and Forward operation in a multi-satellite deployment and HSS on the ground [55]

The attaching procedure includes the following contacts:

Contact #1: SAT#i comes into the communication range of the UE. The UE detects the satellite cell, acquires necessary information (e.g., the PLMN and TLEs of other satellites of the constellation), and learns that the satellite operates in Store and Forward mode. The UE then initiates an attach procedure, indicating its capability to handle Store and Forward operations. However, SAT#i rejects the attach request due to the lack of subscription data on board the satellite and the absence of an active feeder link for data retrieval, issuing an Attach Reject message including a new reject cause and a timer (backoff) to indicate to the UE that the attach procedure needs to be suspended and when it can re-attempt, respectively.

Contact #2: SAT#j comes within the communication range of the UE. SAT#j is assumed to be in Store and Forward operation mode and to have no subscription information on board either to be able to resume/complete the attach registration of the UE. The satellite cell is detected by the UE, which identifies that this satellite cell belongs to the same PLMN, and that SAT#j is operating in Store and Forward mode. The UE does not trigger an attach attempt because

SAT#j operates in Store and Forward mode and the timer value (S&F Timer#1) has not yet expired.

Contact #3: SAT#i connects to the ground network via the feeder link, Cellular IoT (CIoT) Serving Gateway Node (C-SGN) entities: satellite C-SGN-SAT and ground C-SGN-GND exchange information about the initial attach attempt detected during Contact #1. C-SGN-GND communicates with the HSS of the home PLMN of the UE over the *S6a* interface to retrieve authentication information using the UE's IMSI. Additionally, C-SGN-GND may trigger an ULR to pre-fetch subscription data for future satellite passes, obtaining necessary subscriber information, including E-UTRAN Authentication Vectors (AVs).

Contact #4: SAT#k gets connected to the ground network via the feeder link. The assumption is that SAT#k is expected to fly over the area where the UE that attempted to attach is located. E-UTRAN AVs for that subscriber/UE are also uploaded to the C-SGN-SAT of SAT#k.

Contact #5: SAT#k comes within the communication range of the UE. The UE detects the satellite cell and re-tries the initial attach procedure on SAT#k, given that the Store and Forward Timer#1 has already expired. The authentication and security setup are successfully completed, as SAT#k has the necessary information about the UE/IMSI. The network attach is accepted, creating a new UE/MME context. A timer value (Store and Forward Timer#2) is provided to the UE to indicate when it can attempt data transmission/reception in any satellite, preventing premature data transfer before the network completes the location update procedure and synchronises the new context across other satellites of the constellation.

Contact #6: SAT#k gets connected to the ground network via the feeder link. The C-SGN-GND learns about the new UE/MME context created on SAT#k.

Contact #7: SAT#j gets connected to the ground network via the feeder link. The new UE/MME context is uploaded into SAT#j.

Contact #8: SAT#i comes within the communication range of the UE. The UE may detect the satellite cell, but it does not trigger any user data transmission because the Store and Forward Timer#2 has not yet expired.

Contact #9: SAT#j comes within the communication range of the UE, and user data transmission can be performed.

4.5.5.1.3 Data transfer in Store and Forward satellite operation mode

After the successful registration with the satellite PLMN and the establishment of a valid UE/MME context across all or a subset of the constellation's satellites, data transfer can occur between the UE and any of these satellites when they are within reach. Figure 4-24 illustrates a data transfer mechanism for MO data. For MT data, a similar approach applies, with the only difference being the trigger of the paging procedure by the CN-SGN-SAT when the UE is expected to be within the coverage of the flying satellite. The MO data is transferred in two steps: (1) a contact between UE and SAT#i, and (2) a contact between SAT#i and ground.

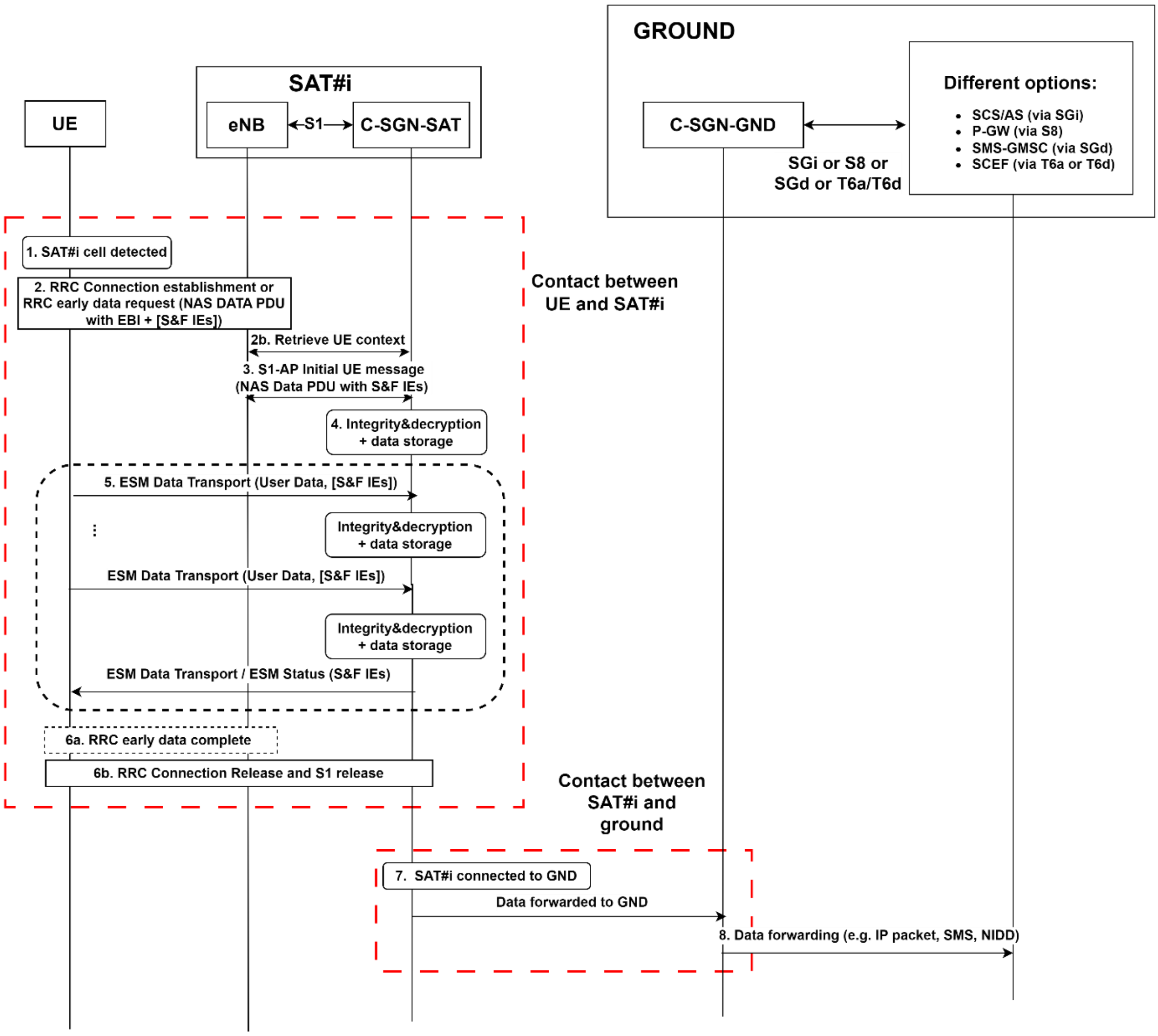

Figure 4-24: MO user data transfer [55]

Contact between UE and SAT#i: The UE detects that a SAT#i is reachable and verifies that Store and Forward Satellite mode is supported and the “Store and Forward CP data transfer“ procedure. The UE either establishes an RRC connection or sends an RRCEarlyDataRequest message, which includes an integrity-protected NAS PDU that carries the EPS Bearer ID and encrypted Uplink data. The eNodeB, based on configuration, may retrieve the EPS negotiated QoS profile from the MME and additional parameters (e.g., UE Radio Capabilities) for relaying the NAS PDU to the MME via an S1-AP Initial UE message. The MME checks the integrity of the message, decrypts, and stores the data. Additional user data may be exchanged via the ESM data transport mechanism, which includes specific Store and Forward Information Elements (IEs) to manage data quotas and inform about expected delivery times. IEs may indicate data transmission allowances, priority levels, and expected delivery times for buffered data. To complete the data transfer, an RRC connection and *S1* release are triggered or an RRC Early Data Complete is sent.

Contact between SAT#i and ground: SAT#i gets connected to the ground network via the feeder link and forwards the stored user data to the C-SGN-GND. Finally, downloaded user data is forwarded to the destination via proper interfaces.

As part of the ongoing work by the 3GPP, several challenges have been undergone analysis across various Working Groups (WGs), with efforts focused on the methods to transfer the ephemerides to end devices [91], [92], [93], enhancements to paging procedures aimed at reducing the paging load [94], and refining the Store and Forward operations [46], [95], [96], [97], [98], [99], [100], [101]. However, certain challenges persist, such as the sharing of UE context between satellites, mobility management, the dynamic organisation of moving tracking areas and the broadcasting of ephemeris data to UEs to aid them in utilising network and power resources optimally.

4.5.5.1.4 Impact on the ETHER E2ESL when using Store and Forward operation

Using Store and Forward in the network does not impact any entities or interfaces, specifically of the EPS architecture. However, some modifications are required to support Store and Forward operations at the E2ESL, especially in a multi-satellite constellation scenario [55].

The network needs to advertise its support of Store and Forward capabilities or the activation of Store and Forward operation mode. To achieve this, new messages are required, which can be new IEs in the System Information Block (SIB), e.g., SIB31 or SIB32, or it can also be new IEs in NAS messaging. Moreover, the UE must advertise the support of Store and Forward, which can be achieved by a UE network capability flag indicating Store and Forward support.

Regarding the NAS procedures spanning several satellites passes and potentially different satellites, two extensions have been identified. Firstly, new reject causes and backoff timers for terminating/suspending NAS procedures and assisting the UE on how to proceed with subsequent re-attempts in future passes of the same or different satellite cells. Secondly, a new IEs for the network to inform the UE of the satellites in which user data transfer is allowed after completing network registration (e.g., a list of satellite identifiers for which the registration is valid).

In terms of the user data transfer, more specifically on the CP CIoT EPS optimisation, the data transfer procedure may be extended with new IEs in the ESM Data Transport messages and/or new causes in the ESM Status messages to handle Store and Forward data transfer aspects such as Store and Forward data quotas, delivery priority levels and providing information on expected delivery times.

Furthermore, some extensions could be also needed from the E2ESL, such as a subscription parameter related to Store and Forward satellite operation in the HSS and the option to pre-fetch the necessary data for Store and Forward operation before UE authentication. Also, the awareness of the Store and Forward satellite operation to the Service Capability Server/Application Server entities via *T8* procedures.

4.5.5.2 5G and beyond support

The 3GPP's Release 19 discusses future developments for 5G and beyond, where support for Store and Forward functionality would be provided by a complete 5G architecture, incorporating the components of the 5G CN. Figure 4-25 illustrates a possible 5GS architecture, where a regenerative payload is considered and the AMF at the satellite is responsible for the onboard Store and Forward feature [55].

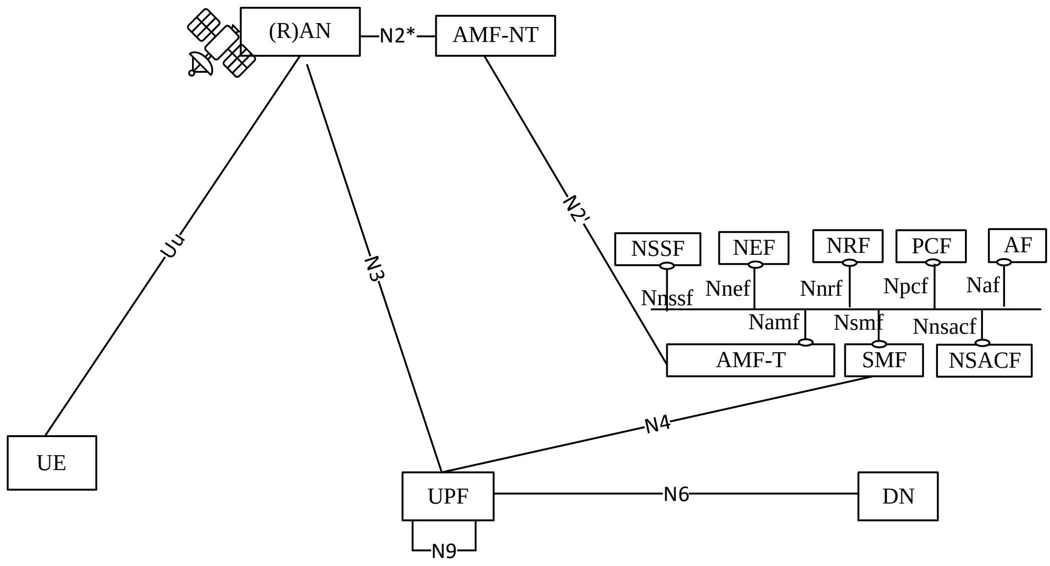

Figure 4-25: AMF split architecture for Store and Forward operation for 5G architecture [55]

The AMF role is divided between terrestrial (AMF-T) and non-terrestrial (AMF-NT) AMF entities. The buffering of MO and MT data occurs on the NF in the orbit and on the NF on the ground, with the peer AMF entities hiding the Store and Forward functionality from other NFs. The Store and Forward mechanism supports small data transmission when the satellite cell intermittently loses GS connectivity, with its necessity determined during registration, mobility, or periodic updates [55].

4.5.5.2.1 Registration procedure

It is expected that the target UE would be under direct coverage at least occasionally, allowing registration and mobility updates via TNs or satellite cells with GS connectivity. In this solution, the need to use Store and Forward operation can be detected earlier at registration, mobility update or periodic update time. The registration procedure is illustrated by Figure 4-26 and can be described by the 25 steps below [55]:

- **Step 1**: The UE sends an Access Network message with parameters and a Registration Request, including the type of registration (initial, mobility update, periodic update, or emergency). If the UE has a valid 5G NAS security context and needs to non-cleartext information, it includes a NAS message container with the registration request. If there is no need for non-cleartext information or if the security context is invalid, the UE sends the registration request without the NAS message container. The complete registration request, including both cleartext and non-cleartext IEs, is included in the NAS message container as part of the security mode complete message.

- **Step 2:** The RAN sends an *N2* message to the AMF-NT-1. This *N2* message includes *N2* parameters and the registration request as described in step 1. If the AMF-NT-1 does not have connectivity to the ground upon receiving the registration request, it will store the registration request message.

- **Step 3:** The AMF-NT-1 sends a DL NAS transport message to the UE via the RAN. This message includes a temporary UE identifier. It is assumed that all AMF-NTs in the satellite constellation have unique AMF IDs, ensuring non-colliding Global Unique Temporary ID

(GUTI) allocations. The UE will store this temporary identifier (interim GUTI) for future transactions. The AMF may also provide the UE with a validity timer for this temporary identifier.

- **Step 4:** When the satellite containing RAN-1 and AMF-NT-1 moves away from the UE and reconnects to a GS, the AMF-NT-1 forwards the stored registration request message, along with the Interim GUTI, to AMF-T.

- **Step 5:** The AMF-T may decide to initiate UE authentication by invoking an AUSF. In this case, the AMF-T selects an AUSF based on the Subscription Permanent Identifier (SUPI) or Subscription Concealed Identifier (SUCI). If the AMF-T is configured to support Emergency Registration for unauthenticated SUPIs, and the UE has indicated an Emergency Registration type, the AMF-T can either skip the authentication or proceed with the registration procedure even if the authentication fails.

- **Step 6:** If authentication is required, the AMF-T requests it from the AUSF. Upon receiving the request from the AMF-T, the AUSF executes the UE's authentication procedure, selects a UDM entity and retrieves the authentication data from it.

- **Step 7:** When the AUSF returns the authentication data to AMF-T, the AMF-T stores it associated with the interim GUTI received from AMF-NT onboard the satellite. The AMF-T then determines the next probable satellite that can reach the UE's location

- **Step 8:** The AMF-T sends a *Namf_N1N2MessageTransfer* to the next probable satellite's AMF-NT-2. This message contains the authentication request NAS message and the interim GUTI. Additionally, the AMF-T provides the last known location of the UE.

- **Step 9**: When the UE moves into the area covered by the new cell (RAN-2) or the AMF-NT-2 pages the UE using the interim GUTI, the UE connects to RAN-2. Upon connection, RAN-2 sends an initial UE message containing the service request.

- **Step 10:** The AMF-NT-2 sends an authentication request message to the UE from step 9.

- **Step 11:** The UE sends the authentication response message to the AMF-NT-2. As the AMF-NT-2 does not have connectivity to the ground when receiving an authentication response, it will store the authentication response message.

- **Step 12:** When AMF-NT-2 moves away from the UE and reconnects to a GS, it forwards the authentication response to the AMF-T.

- **Step 13:** The AMF-T validates the authentication response and then sends an authentication request to AUSF.

- **Step 14:** The AUSF sends an authentication response with the K_{SEAF}, i.e., key for Serving Authentication Function (SEAF), to the AMF-T.

- **Step 15:** Since the NAS security context does not exist, the AMF-T initiates the NAS security by sending a *Namf_N1N2MessageTransfer* to AMF-NT-3. This message contains the Security Mode Command for the UE, the registration request message received in step 4, and the interim GUTI. The AMF-NT-3 stores this information until it reaches the UE's coverage area. The AMF-T selects a suitable AMF-NT-3 based on the criteria discussed in step 8 and provides the last known location of the UE.

- **Step 16:** The AMF-NT-3 tries to reach the UE via paging using the interim GUTI, or the UE attempts to establish an RRC connection using a service request with the interim GUTI. Once the connection is made, the AMF-NT-3 sends the NAS security mode command to the UE.

- **Steps 17-18:** The UE enables the security and acknowledges it to AMF-NT-3.

- **Step 19:** Upon receiving the security mode command acknowledgement from the UE, the AMF-NT-3 stores it until it regains connectivity with the GS. Once reconnected, the AMF-NT-3 sends the security mode command acknowledgement to the AMF-T.

- **Step 20:** The AMF-T performs necessary procedures for registration and, at this stage, may determine whether the UE is allowed to be served in the Store and Forward scenario or not.

- **Step 21:** The AMF-T then selects a suitable candidate satellite to serve the UE next, based on the criteria described in step 8. The AMF-T sends the registration accept message to the AMF-NT-4 using the *N1N2MessageTransfer*. The AMF-NT-4 stores the registration accept message. Additionally, AMF-T provides the last known location of the UE.

- **Steps 22-23:** When AMF-NT-4 reaches the UE serving area, it follows the steps mentioned in step 10. AMF-NT-4 sends the registration accept message and shares the TMSI and security keys with RAN-4.

- **Steps 24-25:** The UE may respond with a registration complete message to the AMF-NT-4. The AMF-NT-4 will store this message and forward it to AMF-T once it regains ground connectivity.

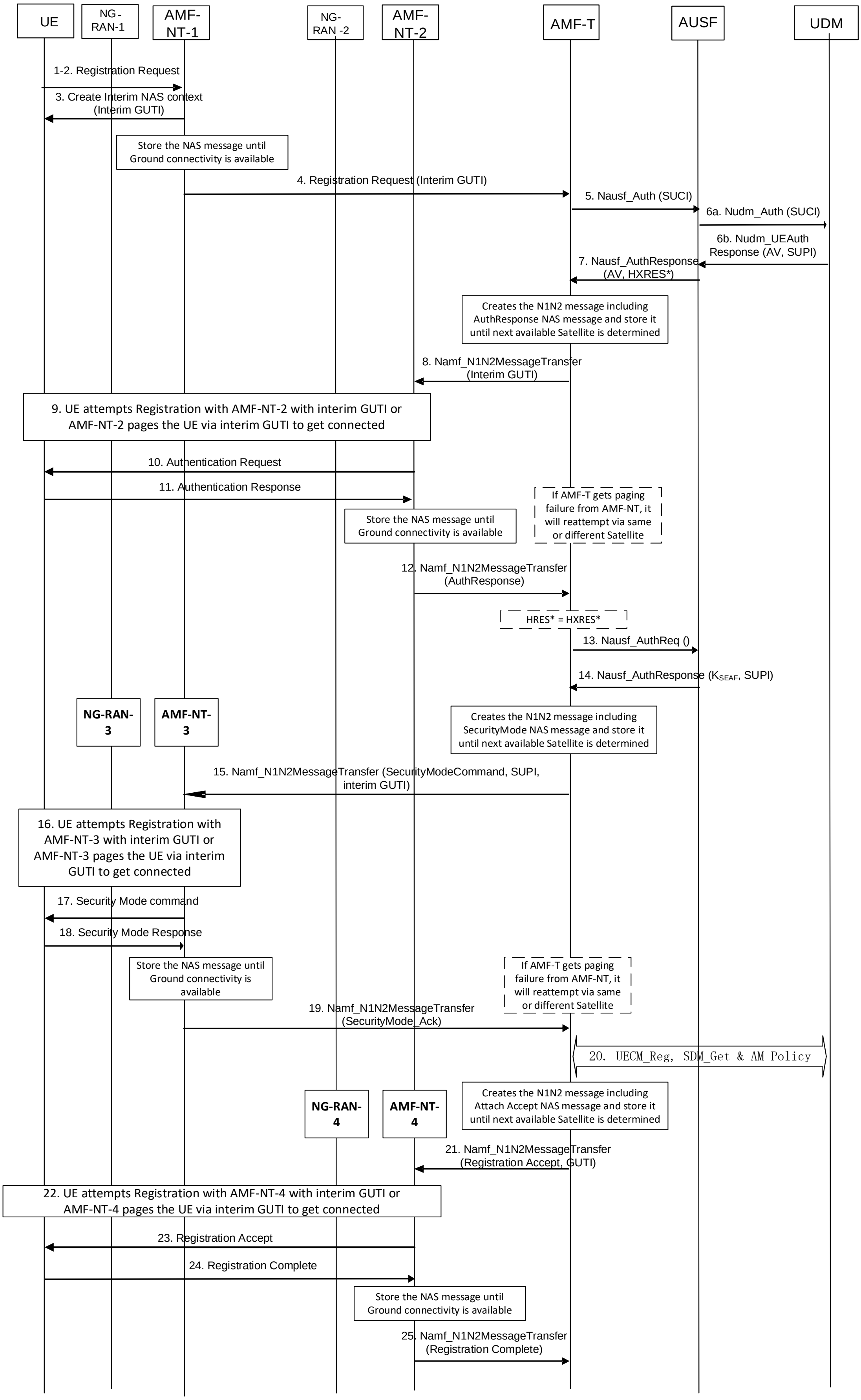

Figure 4-26: Registration procedure Store and Forward in the 5GS architecture

4.5.5.2.2 UE reachability estimation

This architecture further proposes the introduction of a new NF – User Reachability Entity (URE). A ground-based function that estimates which NTN elements can reach a UE based on:

- UE location (and optionally trajectory and velocity).
- Non-terrestrial NFs (i.e., satellites) location, velocity, and ephemeris, including ground-satellite links and radio connectivity.

For static UEs, the URE uses satellite ephemeris to create a precise reachability schedule. For moving UEs, it considers trajectory and velocity, potentially including reachability probabilities.

A terrestrial NF can query the URE for the earliest "UE Reachability Time", "UE Reachability Duration", and candidate satellites to set NAS procedure timers, determine message timings, and assign NAS timer values.

The URE's response may include multiple alternatives for reachability times and durations via different satellites. Additional information like SAT-ID and the IDs of NTNFs serving the UE during the specified reachability time and duration may also be included based on network configuration and policies [55].

4.5.5.2.3 Impacts on services, entities and interfaces

The Store and Forward mode of operation impacts various network elements significantly. For the UE, for example, the NAS must handle larger delays and incorporate new mobility management capabilities and interim GUTI allocation. The AMF-T needs to buffer and store *N2* messages, determine appropriate transmission timings, and manage the secure transmission and reception of NAS messages with the AMF-NT. Similarly, the AMF-NT must handle buffering and determine transmission timings once connected to a GS.

The URE is a newly proposed network function that estimates UE reachability and identifies the relevant NTN elements (i.e., satellites), providing information on when and how the UE can be reached. The NRF must include a new URE profile with supported constellation IDs. These are some required changes to ensure efficient and reliable Store and Forward operation in a 5GS architecture [55].

4.5.6 Control of SDN-based UPF

According to the 3GPP 5G vision, the interaction of CN CP (i.e., SMF) with SDN-based UPF will be performed over the *N4* interface, implementing Packet Forwarding Control Protocol (PFCP) [84]. The SDN CP will be seen by SMF as the control execution enabler of the fundamental packet routing & forwarding functionality of UPF [51], logically embedded within UPF. The detailed architecture of the SDN CP is discussed in Section 4.4. Hereby, it is assumed that the UPF can be implemented as a set of dedicated virtual/containerised/physical NFs, i.e., virtualised or physical SDN-enabled Network Elements, SDNCs and necessary support functions (in the form of SDN applications) needed to terminate *N4* interface and support 3GPP functionalities. It is assumed that ETHER MANO can deploy the UPFs via dynamic orchestration of virtualised components (SDNC, switches) or reuse the already existing ones (e.g., SDNCs belonging to ETHER MANO Transport domains and physical SDN switches), ensuring appropriate resource allocation and control privileges over the CP/DP devices. In both cases, the E2E path setup process and session configuration are conducted via the *N4* interface following standard 3GPP procedures.

The PFCP controls packet processing within the UPF for each Protocol Data Unit (PDU) session through PFCP sessions. For each incoming packet, UPF identifies the relevant PFCP session and the Packet Detection Rule (PDR). Upon PDR hit, the associated rules are applied to the packets belonging to the PDU session. The 3GPP introduces a set of actions which enable enforcing packet forwarding behaviour – Forwarding Action Rule (FAR), packet buffering and retention in UPF – Buffering Action Rule (BAR), QoS for packets belonging to the QoS-flow – QoS Enforcement Rule (QER), handling multi-access PDU sessions – Multi-Access Rule (MAcR) and measurements of UE/QoS flow metrics and events – Usage Reporting Rule (URR) [84].

In order to implement the SDN-based UPF, the proper translation of PFCP and the SDN CP protocol has to be ensured. Today, the most popular and widespread SDN CP protocol is OpenFlow (OF) [102]. While the mobile UP design embodies some of the SDN and OF principles, such as control and user plane separation, support for programmatic control or a packet processing approach (pipeline consecutively executing instructions over the packet), there exist misalignments between the OF protocol and PFCP, which make the implementation of PFCP using OF primitives problematic (cf. Table 4-2). The key deficiencies include the lack of the OF primitives enabling the configuration of hardware and so, the implementation of QER and BAR. Regarding the foremost, the OF protocol only supports flow mapping to the preconfigured in-switch queues. To modify the low-level QoS features, hardware-dependent configuration protocols and solutions have to be used, such as, e.g., OF-Config [103], OVSDB [104], etc. Therefore, in order to enforce QoS in an E2E manner, there arises the necessity of unifying the DP fabric (out of the underlying heterogeneous hardware) in terms of supported QoS features and their configuration. Moreover, OF does not support in-DP delay measurements, which makes QoS enforcement problematic. The lack of hardware configuration capabilities also does not allow buffering control and, hence, the implementation of BAR. The remaining rules specified by the 3GPP can be implemented by using OF-primitives and include PDR (matching packets to the flow entries in flow tables), FAR (associating actions with flow entries, e.g., specifying flow output ports, queues, etc.), URR and SRR (provisioning of statistics on a per-flow basis, configuration of the OF switch event notifications, relevant SDN application-level components to handle metrics acquisition from the DP).

Table 4-2: Comparison of PFCP and OF protocol

PFCP context elements	Scope of Rule	Related OF functionalities/primitives	Missing features
Packet Detection Rules (PDR)	Each contains one or more **Packet Detection Information (PDI)** fields against which, the incoming packets are matched (e.g., IP, port information). Each PDR is associated with a set of rules to be applied on match (FAR, QER, SRR, MAcR, URR). Each PDR has at least one associated FAR and zero or more other rules.	**Flow Table** – a list of Flow Entries **Flow Entry** – traffic flow processing instructions triggered upon match on the predefined criterion (e.g., IP, MAC, MPLS labels, etc.)	-
Forwarding Action Rules (FAR)	Specifies instructions related to the processing of the packets by: **Apply Action** parameter (forward, duplicate, drop or buffer DL/UL packet, notification to CP function; multicast handling) **Forwarding, buffering and/or duplicating parameters** to be used by the UPF	**Flow Entry** – traffic flow processing instructions triggered upon match on the predefined criterion (e.g., IP, MAC, MPLS labels, etc.) **Action –** operations done to the traffic such as flow redirection to SDNC, output	-

PFCP context elements	Scope of Rule	Related OF functionalities/primitives	Missing features
		to port(s)/queue(s)/table(s), or flow drop	
QoS Enforcement Rules (QER)	Specifies QoS Enforcement of the traffic, including: **Gating Control** – control passing traffic through to the desired endpoint **QoS Control** including Maximum Bit Rate (MBR), GBR or Packet Rate enforcement) **DL flow level marking** for application detection **UL flow level marking** for supporting 5G CN reflective QoS	**Flow Priorities** defining precedence of application of entries to the arriving flow **Meters** enabling dropping flow traffic exceeding the threshold at switch ingress **Actions**, e.g., setting meters, setting output queue/ports/tables, setting fields, e.g., VLAN IDs, or MPLS labels, etc. **Egress Queues** mapping for traffic flow shaping and packet handling routines are supported but cannot be configured using OF protocol (alternative hardware protocols have to be used)	**Packet error rate** per flow (available per port)
Session Reporting Rules (SRR)	Detection and reporting of events for a PFCP session, PFCP session not related to any PDRs, or the ones not related to traffic usage measurements: **Change of availability** of 3GPP/non-3GPP access for multi-access PDU session **QoS Monitoring and Reporting** per QoS Flow **Traffic Parameter Measurement Report** per QoS Flow	**Statistics** – port, flow, table, group, or queue level **DP-level** – simple events notification to the SDNC, e.g., related to flow entry removal due to expiry or idle timeout	Mechanisms for measuring delay (link/switch/queue)
Buffering Action Rules (BAR)	Controls the buffering behaviour of the UPF for all the FARs of the PFCP session set with an Apply Action parameter requesting the packets to be buffered and associated with this BAR. The following parameters are considered: **DL Buffering Duration** **DL Buffering Suggested Packet Count**	**Support for buffering** of newly arrived flows during handling of *ofp_packet_in* message **Specification of *n_buffers*** – a maximum number of packets buffered at once in the physical or virtual ingress queue (switch-dependent)	Buffering configuration and management capabilities
Multi-Access Rules (MAcR)	Instructions to the UPF on how to forward the packets matching the PDR over 3GPP and non-3GPP accesses	**Flow Table** – set of Flow entries **Flow Entry** – processing rules based on packet header metadata and predefined match instructions	-
Usage Reporting Rules (URR)	Measures network resources usage in terms of traffic data volume, duration and UP-specific events (based on the provisioned configuration). Instructs the UPF to send a usage report to the CP functions on threshold crossings, periodically, on event detection or on CP specific request.	Behaviour can be implemented in the form of SDNC applications using OF primitives such as asynchronous messages and statistics collection.	-

Therefore, to allow seamless integration with the ETHER MANO stack, there would be a need for implementing the mediator entity, which would allow termination of the *N4* interface and translation of PFCP commands into OF primitives (and vice-versa), e.g., for the purpose of QoS enforcement of the UPF chain (i.e., including the transport network segments interconnecting UPFs on *N9* reference points). The deployment of such mediator entities can be done as SFs at the E2ENO and Transport domain levels. Overall, the SDN-based UPF will allow addressing the challenge of NTN infrastructure mobility by introducing higher flexibility of control over the DP and the ability to quickly adjust the E2E path depending on the current network conditions.

4.5.7 Dynamic topology management

5GS provides fundamental mechanisms for dynamic topology management needed by the ETHER framework, both due to the high mobility of satellite nodes and distributed CN adapting itself to the nodes' mobility, traffic distribution, etc. They can be distinguished separately for CN and RAN:

- Dynamic CN topology can be managed using the CP mechanisms of NRF and SCP, including the discovery mechanism (through SBI) of functions earlier registered in NRF that can be used by any CN CP NF. If NRF is not used, the 3GPP MP (cf. Figure 4-8) can also directly manage the CP NFs configuration through the NFMF-NF interface. It should be noted that NRF does not track the status of registered NFs, so the 3GPP MP should provide management of the updates.

- Dynamic RAN topology management includes mutual awareness of neighbour BSs (Xn Application Protocol – XnAP, gNB-gNB *Xn* interface [105]) and BSs and CN (NG Application Protocol – NGAP, gNB-AMF *N2* interface [106]). Both application protocols provide necessary mechanisms for the creation, updating, and removal of point-to-point (P2P) *Xn* and *N2* links, respectively.

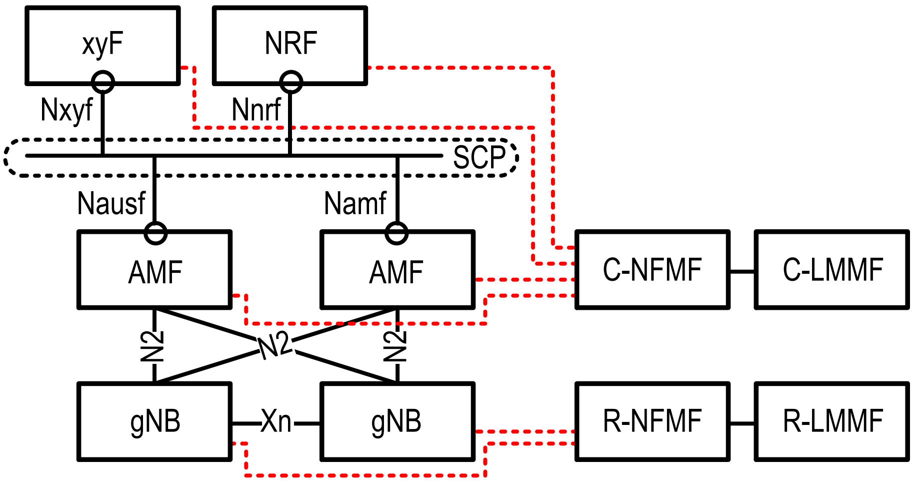

Figure 4-27: Architecture of the ETHER dynamic topology management concept

Figure 4-27 presents the architecture of the dynamic topology management concept. NFMF instances (R-NFMF in the RAN part and C-NFMF in the CN part) control the gNB instances and AMF instances configuration respectively via red dotted line interfaces, and also can have an ability to trigger the XnAP (R-NFMF instances only) or NGAP (R-NFMF or C-NFMF instances) link creation procedures when the IP connectivity between respective functions is active, then update and removal procedures according to the underlying resources mobility patterns (e.g., satellites' trajectories, their visibility windows for specific parts of the TN). For this purpose, NFMF instances are connected to the corresponding (C-/R-)LMMF instances. Alternatively, they can provide the resources' mobility-related information to the managed CN NF instances. Hierarchical coordination will be provided at higher levels of the 3GPP management stack – by NSSMF and NSMF instances (cf. Figure 4-8). C-NFMF instances have the ability of controlling NRF instances and any other CN NF instances (indicated generally as xyF) that are on board of mobile resources, according to the CN distribution within terrestrial and non-terrestrial parts. In general, multiple instances of NFMFs are previsioned within the ETHER framework, both NF type-dedicated and associated with various types of NF. Additionally, the ETHER framework assumes the freedom to implement the interaction between the NF and its NFMF: completely separated entities, complete nesting of the NFMF instance inside the NF instance, and mixed – embedding of the management agent in the NF, which will cooperate with an external NFMF (cf. Embedded Element Manager, [67]). Similarly to the 3GPP approach to the 5GS management, the ETHER framework does not create a standardised approach and specification here but leaves this issue to the implementation.

The mobile network configuration management activities related to dynamic topology changes are associated with two aspects: who is the leader and who is the executor.

The leadership aspect options:

- Direct remote triggering of NF instances – NFMF to force the signalling procedures (weakness: temporary discontinuity of management channels due to, e.g., loss of a feeder link).
- Propagation of resources/NF instances mobility patterns to NF instances or embedded NFMF agents therein – local following the list and semi-autonomous operation, implying the local increase of computational load.

The options of signalling procedures execution responsibility (especially the P2P procedures of XnAP and NGAP):

- Mobile NF instance to initiate – as it can possess the information regarding its position versus the terrestrial counterparts directly in real-time, it can determine proper moments for the execution of the signalling procedures (the computational effort on the side of mobile nodes, which implies concerns about non-terrestrial computational and power resources capacity as well as power consumption issues).
- Stationary (terrestrial) NF to initiate – it predicts the position of the counterpart mobile NF (based on the mobility pattern or broadcasted information with inherent delays) and initiates the signalling procedures (off-loading of non-terrestrial nodes' computational and power resources).

The decision regarding the options of both aspects is left to the WP3/WP4 considerations and simulation of their performance. However, the ETHER framework design supports all the above options. It should also be noted that the extension of CN CP SBI to RAN, as proposed for 6GS [107], can provide a unified interface and protocol for dynamic topology management both in RAN and CN.

4.5.8 AI-based handover control mechanisms

According to the 3GPP standardisation, both in the case of *Xn* handover procedure (cf. [58], clause 9.2.3.2.1) and *N2* handover procedure (cf. [108], clause 4.9.1.3.2), the decision to trigger the handover procedure is made by the source (UE-serving) gNB. Additionally, the source (UE-serving) AMF selects the target AMF (cf. [108], clause 4.9.1.3.2) corresponding to the target gNB. The handover decisions are based on the UE measurement reports periodically sent to the serving gNB (cf. [58], clause 9.2.4).

The AI-based control mechanisms targeting handover management optimisation falls into the perimeter of a wider concept of user-centric “context awareness”, i.e., adaptation of UE-related network control mechanisms to the UE situational context [109]. Generally, the full UE situational context information may not be available within the serving gNB where the handover decision is to be made, if the information exceeds the standardised exchange with CN using UE context management procedures of NGAP (cf. [106], clauses 8.3 and 9.2.2). Hence, this information (i.e., relevant data) should be provided to the decision engine. At the moment of the hereby document delivery, the data sources for AI-based handover control mechanisms have not been determined by the ETHER WP3. Hence, the ETHER framework will support the following optional scenarios (cf. Figure 4-28):

- Local handover decision at the source gNB (as currently defined by the 3GPP standards).
- Handover decision engine externalised outside gNB or consulted “on the fly” with an external functional instance (gNB delegates the decision to an external functional entity or fetches the decision-affecting factors, and then follows the handover procedure, according to its 3GPP definition).

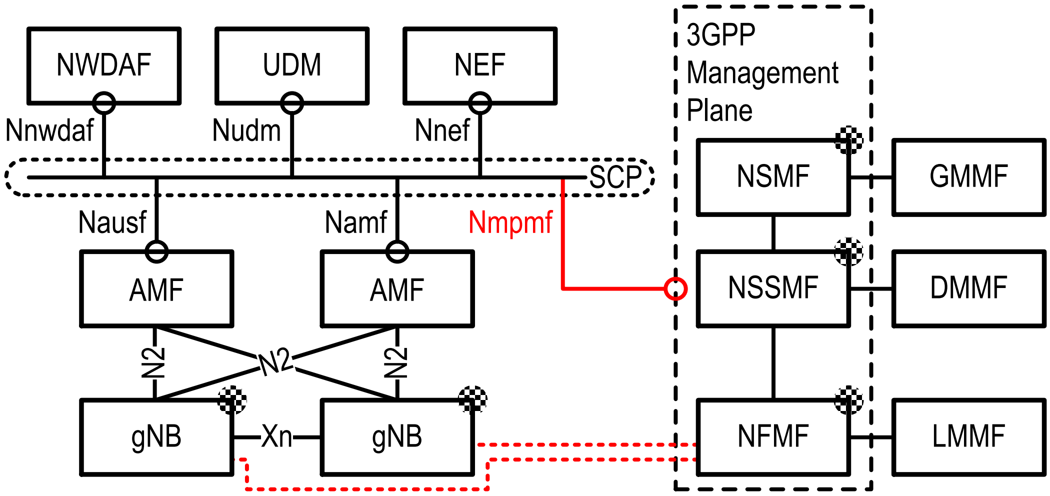

Figure 4-28: Architecture of the ETHER AI-based handover control mechanisms support (shaded spots indicate potential optional locations of handover control engine)

The former scenario assumes that the additional information will be provided to gNB through its NFMF, which has access to the ETHER LMMF. The information originating from the CN CP (potentially, e.g., from UDM or dedicated NWDAF) will be delivered via the NFMF-CP interface using the existing SBI/NEF/SCP mechanisms. This approach will not satisfy only the non-real-time UE situational context needs (NFMF to proxy the CN CP information access), but this way, the NGAP definition will not be affected. However, the gNB software will need to be updated, which in O-RAN-based implementation may be eased by the use of “plug-in” mechanisms of rApps/xApps. It should also be noted that the extension of CN CP SBI to RAN, as proposed for 6GS [107], can provide a unified multilateral and flexible interface for requesting/accessing any CN CP information by gNB instances.

The latter scenario assumes that the external (complete or partial) handover logic will be located in the 3GPP MP (cf. Figure 4-8) and accessible via the gNB-NFMF interface. The exact location (NFMF/NSSMF/NSMF) will be related to the level (perimeter) of the handover coordination decided by the ETHER WP3. However, this scenario provides potentially lower execution speed than the former one.

In both scenarios, the 3GPP MP entities access to CN CP SBI is assumed (indicated globally as a general Management Plane Management Function *Nmpmf* interface, implementation-specific, as there is no 3GPP specification for this bridge between the CP bus and MP bus).

4.5.9 Semantics-aware analytics

According to the information received from WP3, the semantics-aware analytics engine is planned to be located in AP, but no details of the data sources are specified at this stage of the project. Therefore, it is assumed that the support of this mechanism will be through the generic CN CP mechanisms (NEF/AF or NEF alone), as indicated in Figure 4-13 (mechanism A). Additionally, for specific purposes, the dedicated NWDAF may be applied. RAN-specific CP information unavailable in CN CP can be also fetched via gNB NFMF. For that sake, the semantics-aware analytics engine will need an interface to the 3GPP MP.

4.5.10 Direct handheld access

The 5G NR supports two waveforms, namely Cyclic Prefix OFDM (CP-OFDM) and Discrete Fourier Transform-spread OFDM (DFT-s-OFDM). CP-OFDM is only for the downlink, whereas both can be selected for the uplink. Regarding the DFT-S-OFDM, it may be selected based on the channel conditions in the serving cell. Then, the gNB selects the waveform to be used for the physical uplink shared channel and instructs the UE accordingly. In the cases of NTNs, both waveforms have been adopted. However, the performance might be highly deteriorated in scenarios with high Doppler shifts that may arise as the access point of a UE switches from a terrestrial gNB to an NTN one. That is why one of the technical innovations of ETHER is related to the study of OTFS as an alternative to OFDM-based waveforms in NTN-related scenarios of high mobility of the NTN platforms, such as LEO satellites. In such a case, through measurements of the Doppler shift by the UE and reporting to the gNB, a threshold-based decision can be taken for switching to OTFS or not (based on the Doppler shift). Such measurements can be performed under the assumption that the UEs are equipped with a GNSS receiver, so that they know their position, and the satellites transmit their ephemeris data, which contain their position and velocity. In cases, though, where the GNSS signals are weak and the position of the UE cannot be estimated with accuracy, there might be significant residual Doppler effects in the compensation process for OFDM-based waveforms. That is why operating in the delay-Doppler domain, used in OTFS, is advantageous because the channel becomes sparser and varies on a much larger time fashion than in the time-frequency domain. There are several open areas for investigation regarding the introduction of a new waveform, such as OTFS. These include the design of synchronisation algorithms, random access protocols, and reference symbols for the Doppler shift estimation.

Finally, regarding distributed simultaneous transmission from multiple satellites, which is another important feature studied in ETHER for achieving direct handheld access, the concept is similar to the standardised coordinated multipoint (CoMP) aspect of LTE-Advanced, which allows joint transmission to a UE from several distributed antennas. In our case, the different satellites take the role of the distributed flying antennas, for which synchronisation in time, frequency, and phase needs to be done.

4.6 ETHER AI LAYER

According to the general vision of the overall ETHER architecture (cf. Section 4.1), the ETHER AI Layer is envisioned as a separate plane that is transversal to existing ones and is natively integrated into the ETHER system. The role of the ETHER AI Layer is to expose AI-related services to the ETHER MANO, including all the associated applications, communication, and steering mechanisms. This calls for an architectural module that can develop synergies in terms of data management, and that can handle the interaction with Machine Learning Operations (MLOps) platforms, managing the complete life cycle of both complex NI instances and atomic NI functions, the maintenance of catalogues of AI models that ease de-composition and orchestration.

The ETHER AI Layer thus aligns with the emerging vision for 6G networks that targets E2E integration of AI into the network architecture so as to enable zero-touch management. Such a pervasive incorporation of AI into the mobile network requires a coordination of the many AI-based algorithms running across schedulers, controllers, and orchestrators, ensuring their conflict-free and synergic operation. The ETHER project pushes such a concept to NTN domains, whose dedicated AI shall also be included in the E2E AI-driven operation.

4.6.1 Reference design for the ETHER AI Layer

Current frameworks by main SDOs are far from supporting a native NI integration: for instance, 3GPP Release 19 will not include centralised AI/ML control, which will thus not be considered by the SDO until 2026 [110]. The main efforts in terms of the definition of an AI-native 6G architecture come instead from scientific research, where projects and initiatives dedicated to this topic are proposing early designs.

ETHER adopts a pragmatic approach to the definition of its AI Layer, recognising that a number of focused actions are already introducing dedicated layers or planes for AI coordination and management in beyond 5G networks. Such previous and ongoing actions are identifying the precise requirements for an AI Layer as well as offering initial blueprints of the organisation and functionalities of such a new layer. Considering that the AI Layer is an important component of the ETHER architecture but not at the core of the contributions of the project, it would be inconsistent for this action to duplicate effort from much more focused projects and define yet another architectural model for AI management that would necessarily be less elaborated and detailed than other proposals that are already available in the literature.

Instead, the strategy ETHER embraces, is that of building on existing proposals for AI layers, integrating them into the ETHER architecture and making sure that the AI-driven functions developed in the project fully align with the target AI layer design and operation.

Specifically, ETHER identified one proposal for an AI Layer as the most promising for inclusion in the ETHER architecture. This is NIP introduced by the DAEMON project funded by the European Commission under the call ICT-52 of the Horizon 2020 framework programme [111]. NIP has, in fact, been adopted by 5G PPP as part of its architectural view for 6G under the terminology of NI Stratum [112], as seen in Figure 4-29, which has then been inherited by the Architecture WG of SNS JU and 6G Smart Networks and Services Industry Association (6G IA).

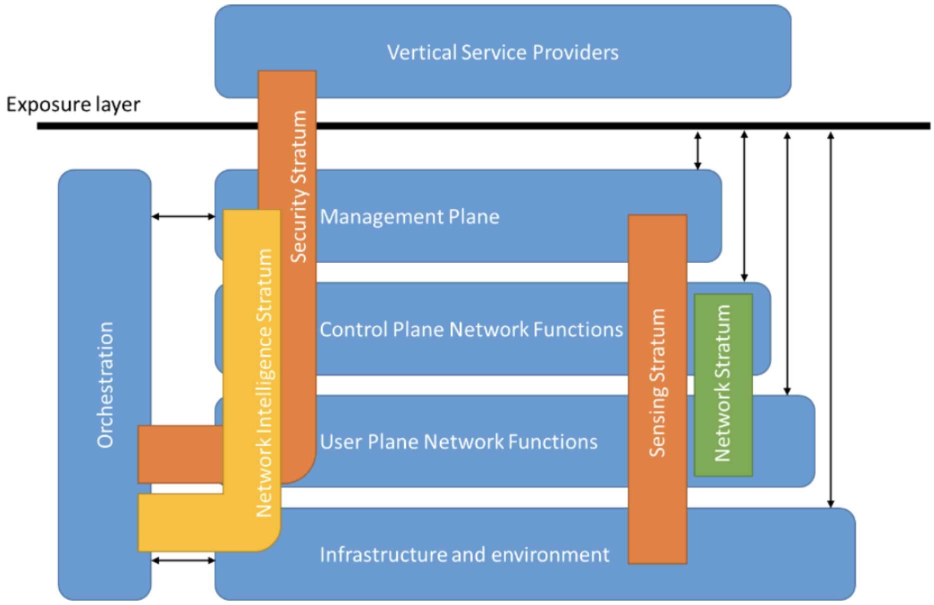

Figure 4-29: Architectural view for 6G by 5G PPP (image reproduced from [112])

In the following, we review and summarise the operation of the NI Stratum as it is originally proposed by the DAEMON project, limited to the structures and functionalities that are relevant to the solutions developed in ETHER. Through the execution of the action, we will ensure that the AI models and overall architectural definitions produced by ETHER are aligned with the representations and interfaces outlined next, so as to guarantee that the AI Stratum can be seamlessly deployed as part of the E2E architecture developed in ETHER.

4.6.2 AI model representation

The DAEMON NI Stratum commends that all AI models deployed in the mobile network are aligned to a single general representation. This is paramount for the NI Stratum – hence the ETHER AI Layer that builds on it – to be then capable of interpreting and managing such models.

More precisely, AI models are defined along three levels of complexity, as follows.

- NI Functions (NIFs) are decision-making functionalities for deployment in a controller, NFVO, or individual NF. NIFs feature well-defined interfaces and behaviours, and basically correspond to one individual AI model that operates in autonomy and serves a specific functionality. For instance, in the context of the ETHER architecture, an AI-driven satellite link quality predictor is a NIF, and an AI model deciding whether users shall connect through the terrestrial or non-terrestrial access networks is another NIF.

- NI Services (NISs) are an assembly of NIFs with a specific objective, often associated with a particular set of targeted KPIs. Essentially, when multiple AI models concur to the same network management task, they are composed into a NIS. For instance, in the context of the ETHER architecture, a solution for the automated and anticipatory scheduling of users across terrestrial and non-terrestrial access networks is a NIS composed of, e.g., the predictor and user association NIFs mentioned at the point above, since the forecast of the link quality is necessary to take the user association decision in a proactive way.

- NIF Components (NIF-Cs) are atomic elements describing one specific sub-operation required to perform a NI task. Therefore, a NIF-C is a fundamental and atomic operation that is required to execute an AI model. For instance, in the context of the ETHER architecture, the collection of real-time information about the status of a satellite link is a NIF-C needed by the NIF for satellite link quality prediction in order to gather the input necessary to perform the target forecast.

Given the three levels of disaggregation (or aggregation) of AI models above, one NIF can be essentially understood as a composition of NIF-Cs. This gives the possibility of introducing a standard representation for NIFs that hinges upon the well-known MAPE-K loop for autonomic and self-adaptive systems [66]. Specifically, the composition of NIF-Cs into individual NIFs is organised according to an extended Network MAPE-K (N-MAPE-K) model tailored to the mobile network environment. The N-MAPE-K model extends the original MAPE-K reference by adding original training and closed control loops that a NIF may implement, e.g., in the context of the ETHER AI-driven functions [113].

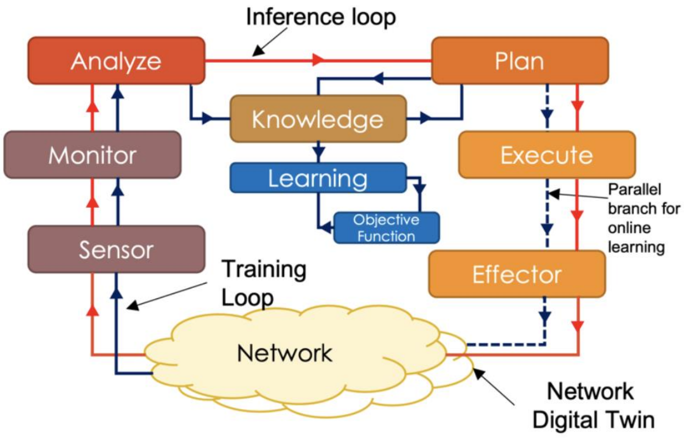

Figure 4-30: N-MAPE-K representation for AI models (image reproduced from [112])

The N-MAPE-K representation is illustrated in Figure 4-30. Each NIF-C maps to one block, and a specific NIF may be constituted of all blocks in the figure or a subset of them. For instance, in the context of the ETHER architecture, the satellite link quality prediction NIF mentioned above may be composed of Sensor/Monitor (to capture real-time information about the link quality), Analyse (to aggregate such information over time to the desired temporal granularity), Plan (to derive future link quality estimates) and Execute (to pass the output prediction to the downstream NIFs that rely on it) NIF-Cs. The same NIF also uses the Know-ledge (to store historical link quality data) and Learning (to train the AI-based predictor) NIF-Cs during the training phase of the data-driven model.

4.6.3 AI Layer organisation

Abiding by the pragmatic approach introduced above, we align the organisation of the ETHER AI Layer to that of the DAEMON NI Stratum, which is summarised in Figure 4-31. The DAEMON NI Stratum – hence the ETHER AI Layer – is managing NISs, NIFs and NIF-Cs, following a similar concept to ETSI NFV MANO, through the following elements.

- The NIF-C Manager takes care of the LCM of NIF-Cs, including onboarding, instantiation, termination, scaling, and state retrieval. For instance, in 3GPP 5G SBA, it interacts with NEF to fetch data (Source), NRF to provide control decisions (Sink), and NWDAF for infe-rence and storage (Analyse-Plan-Knowledge). Or, in O-RAN, it interacts with xApps/rApps as data providers, intelligence holders, or inference consumers.

- The NIF Manager oversees the life cycle of the NIF, including the creation of NIF-Cs, the configuration of the model meta-parameters, or the monitoring of the quality of the decisions.
- The NI Orchestrator (NIO) is responsible for overseeing the LCM of the NIS by effectively coordinating the NIFs that constitute each of them. In doing so, the NIO serves a number of key purposes when it comes to the efficient coexistence of AI models across network domains and planes, including:
  - avoiding conflicts, such as those generated by the presence of different NI algorithms that aim at configuring the same network functions or resources but operated at diverse timescales or based on diverse input;
  - leveraging synergies, such as those arising from knowledge sharing, where the knowledge learned by a NIF can support or enhance the decisions of other NIFs, possibly across domains;
  - performing model selection, cataloguing, and re-training so as to ensure that the deployed NIFs match the specific hardware, software and environmental characteristics of network functions by exchanging info with the MANO system to select the appropriate model for inference within a NIF.

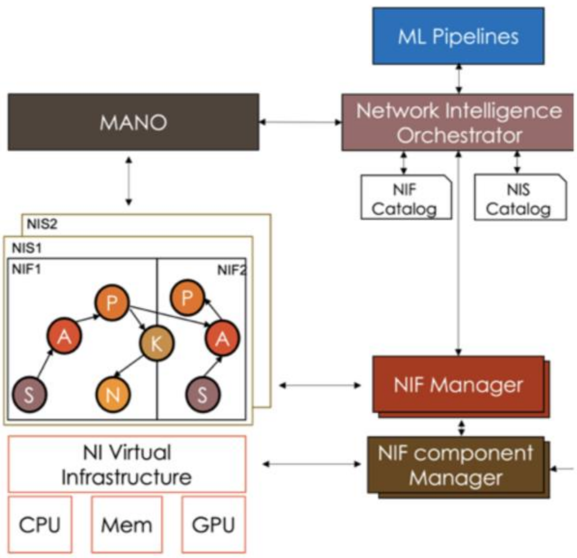

Figure 4-31: Organisation of the AI Layer (image reproduced from [33])

The reference design for the ETHER AI Layer also specifies a number of interfaces between the NIO above and all relevant network elements to AI operations. Figure 4-32 illustrates such interfaces by extending the previous view. Among all interfaces specified by the DAEMON NI Stratum [114], those of highest importance to the ETHER architecture are the following:

- *Nio-Mano* creates real-time synchronisation with MANO frameworks. It allows direct handling of available system resources relevant for AI execution, including computing power, storage, and network characteristics (both in the radio access and core) state and health of network slices. The computing power can include CPU, Graphical Processing Unit (GPU) as well as FPGA entities. The interface also allows the NIO to dynamically adapt NI decisions, possibly abiding by the specific NI requirements of a vertical service provider.

- *Nio-Ext* allows the NIO to communicate with external orchestrators and controllers, including emerging standards that are putting AI operations upfront into the network architectural design, such as O-RAN.

- 3GPP 5G CN.

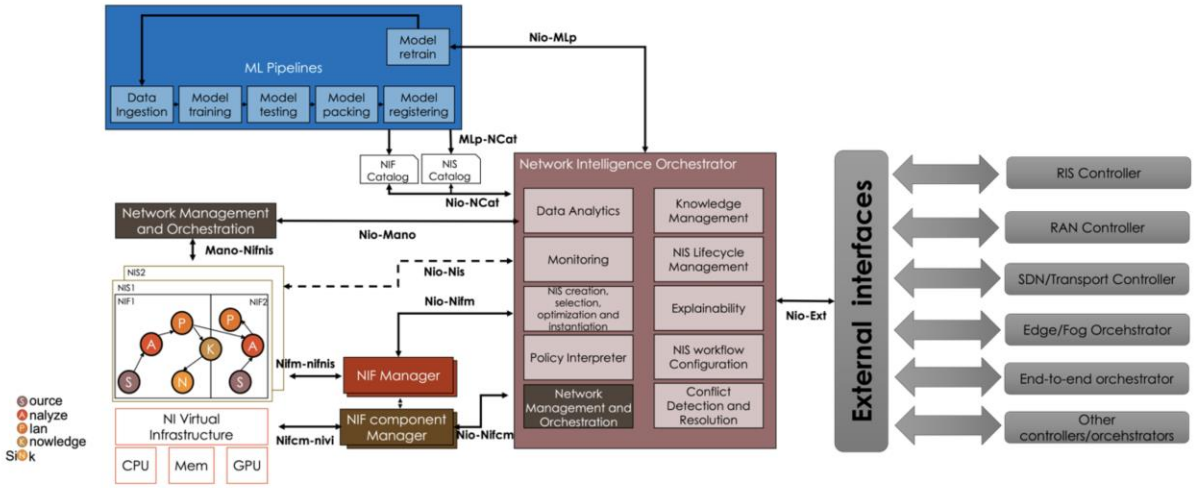

Figure 4-32: Extended view of the AI Layer, including NIO interfaces (image reproduced from [114])

Ultimately, the reference design for the ETHER AI Layer above enables reusing a full set of procedures that are defined for the DAEMON NI Stratum and are aimed at regulating the operation of AI models deployed in the network, including, for instance, LCM (i.e., NIS creation, instantiation, update, termination, etc.) or coordination (e.g., conflict resolution, knowledge sharing, etc.).

4.6.4 ETHER alignment to the reference design for the AI Layer

The ETHER project will ensure full adherence to the reference design of an AI Layer provided by the NI Stratum operation as defined above via the following steps:

- All AI models developed in ETHER will be represented as NISs (for complex AI-driven operations and functionalities) or NIFs (for monolithic AI models).

- Each NIF developed in ETHER will be decomposed and modelled as a set of NIF-Cs, which will then be mapped to the N-MAPE-K representation.

- The local resources to be used by each NIF-C will be identified so that they can be communicated to the NI Stratum NIO via *Nio-Mano*. Conflicts and/or synergies will then be identified, and suitable coordination mechanisms defined.

4.7 INTRA- AND INTER-LAYER INTERACTIONS

After describing the design principles for the ETHER reference structure (cf. Section 4.1) and defining what layers are involved in the system architecture (outlined in Section 4.2 and further developed in the following sections), the next goal should be defining the interfaces between the architecture elements that operate in each of these layers. In D2.1 [4], some proposed inter-layer interactions were brought up, predicting what issues would come up with the ETHER goals and services that are envisioned in an integrated TN-NTN ecosystem. They were:

- SBA and REST API for inter-layer interactions, as a general principle for all ETHER-specific interfaces, to simplify inter-layer communications between network elements. These protocols should be followed unless an industrially standardised interface for a specific reference point already exists.

- Interactions between the CP of E2ESL and ETHER MANO, as the interfaces defined in the 3GPP approach might require additional modifications for the ETHER architecture.

- Cross-layer control loops, to ensure compliance of the interaction chains between architecture layers, like E2ESL, ETHER MANO, ETHER AI Layer, or even higher-level management hierarchy control loops, within the ETHER project UCs.

- Considerations regarding the management and optimisation of horizontal/vertical handover, as well as managing policies and enforcing them in the executive ETHER architecture entities, such as gNB and AMF.

- Dynamic RAN topology considerations, since providing constant awareness of the changing RAN topology towards the BSs and AMF instances is crucial to maintain proper UE mobility management, as well as maintaining consistency with global management from the point of view of the overall ETHER system architecture.

For this final report on the ETHER network architecture, an outline of the architecture map was finalised, as illustrated by Figure 4-33.

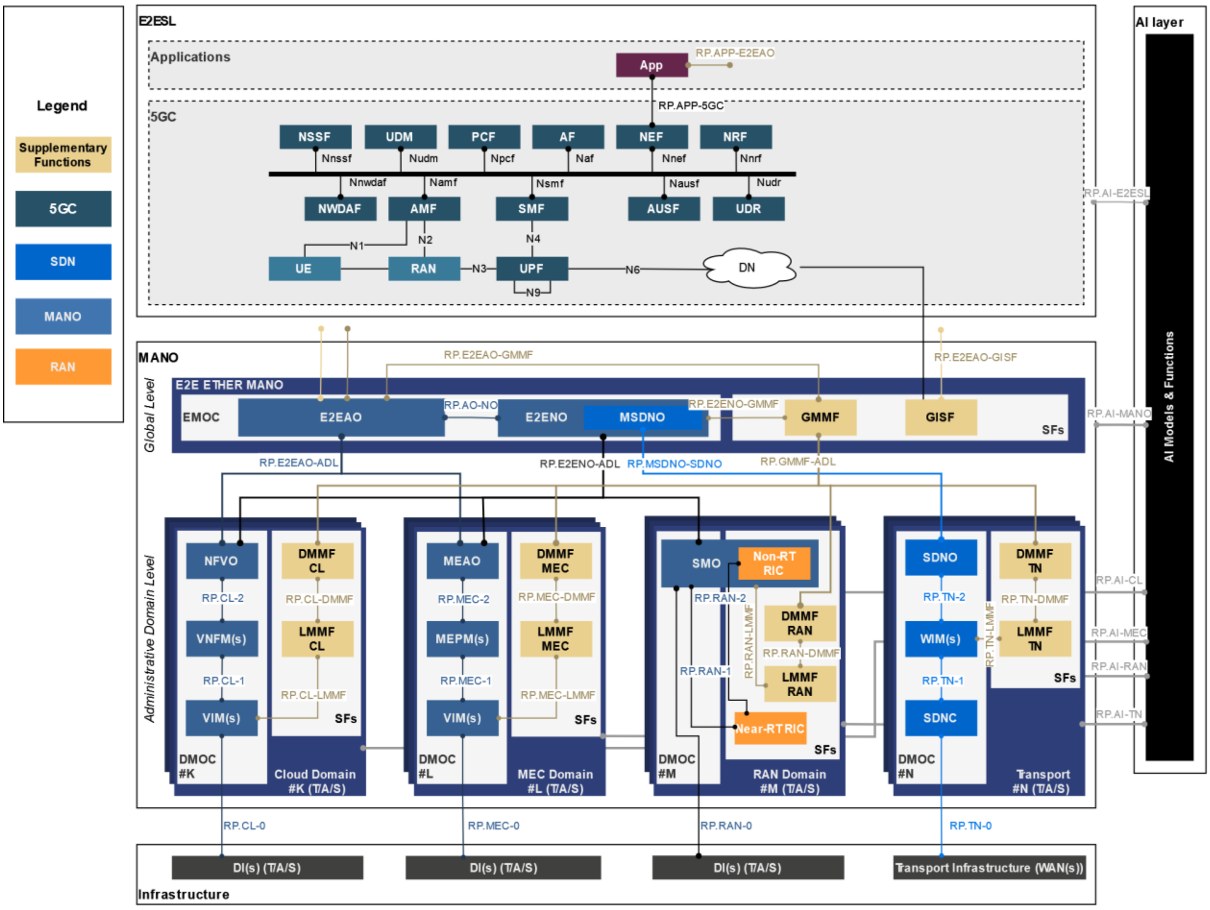

Figure 4-33: Interface map of the ETHER architecture

This is a key step in this document in order to further substantiate the vision for the ETHER project. By presenting a concrete plan for the ETHER architecture, these interfacing issues that were previously raised can be addressed by specifying the interactions between the elements of the network. It also serves as a way to gather all of the knowledge and work done in the previous sections.

While the ETHER architecture layers have been covered previously in this document, the reference points that link network elements together have not been thoroughly specified. This is especially relevant for the purposes of solving the aforementioned interfacing issues but also, on a more important level, to assure that the ETHER architecture is able to provide on its capabilities, being able to not only innovate on the current 5G services and technologies in the TN-NTN sphere, but also evolve into the coming 6GSs.

Accordingly, this specification is divided into two parts: covering both the interfaces that occur between layers, that is, network elements belonging to different layers having a reference point connecting them, and also the interfaces between network elements belonging to the same architecture layer. This is done in order to maintain compliance between the projected network structure and the predicted requirements.

4.7.1 Inter-layer reference points

4.7.1.1 E2ESL – AI Layer

RP.AI-E2ESL: This reference point between 5G CN within E2ESL and the AI Layer exposes AI-related services to the CP/UP network functions and AP entities, including all the associated applications, communication, and steering mechanisms.

4.7.1.2 E2ESL – ETHER MANO

RP.APP-E2EAO: This reference point establishes the connection between the application server and the E2EAO. It facilitates the deployment of applications with particular compute characteristics, such as CPU and memory, to a specific node or infrastructure. It also manages the life cycle of the application. This management includes tasks like creating the application instance on the designated infrastructure, updating the application or its specific parameters, and deleting the application when it is no longer required.

4.7.1.3 ETHER MANO – AI Layer

RP.AI-MANO: This reference point between E2E ETHER MANO and the AI Layer exposes AI-related services to the E2E ETHER MANO, including all the associated applications, communication, and steering mechanisms.

RP.AI-CL: This reference point allows the AI Layer to communicate with the Cloud network domain orchestrators and controllers, putting AI operations upfront onto the Cloud networks in the ETHER structure.

RP.AI-MEC: This reference point allows the AI Layer to communicate with the MEC network domain orchestrators and controllers, putting AI operations upfront onto the MEC networks in the ETHER structure.

RP.AI-RAN: This reference point allows the AI Layer to communicate with the RAN domain orchestrators and controllers, putting AI operations upfront onto the RAN networks in the ETHER structure.

RP.AI-TN: This reference point allows the AI Layer to communicate with the Transport network domain orchestrators and controllers, putting AI operations upfront onto the Transport networks in the ETHER structure.

4.7.1.4 ETHER MANO – Infrastructure Layer

RP.CL-0: This reference point establishes the communication between the VIM(s) and the Cloud DI(s) and will be described as the *Nf-Vi* interface [115] specified by ETSI.

RP.MEC-0: This reference point establishes the communication between the VIM(s) and the MEC DI(s), and will be described as the *Mm7* interface [37] specified by ETSI.

RP.RAN-0: This reference point establishes the communication between the SMO and the RAN DI(s) and will be described as the *O2* interface [116] specified by the O-RAN Alliance group.

RP.TN-0: This reference point establishes the communication between the SDNC and the Transport Infrastructure (WAN(s)) and will be described by the OF protocol [102] specified by the Open Network Foundation, part of the family of The Linux Foundation Projects.

4.7.2 Intra-layer reference points

4.7.2.1 ETHER MANO

RP.E2EAO-GISF: This reference point establishes the communication between the E2EAO and the GISF. It exposes geographical data from target areas to be used for spatial representation and other strategic purposes. It will be developed using standard protocols established by the Open Geospatial Consortium, such as WFS, WMS, and WCS, defined in [117].

RP.E2EAO-GMMF: This reference point establishes the communication between the E2EAO and the GMMF. It exposes mobility management data collected by the GMMF from all the infrastructures across the different domains. It will be developed using standard protocols established by the Open Geospatial Consortium, such as WFS, WMS, and WCS, defined in [117].

RP.AO-NO: This reference point enables the E2ENO to expose the UE applications data contents to the E2EAO through routing paths meeting their QoS requirements. The E2EAO enforces the process of scaling up and down the multi-edge computing resources according to the volume and sensitivity of application contents. This reference point also exposes UE applications' contents to the active domains in the ETHER structure and helps enable E2E network data routing and management across all domains.

RP.E2EAO-ADL: This reference point establishes the communication between the E2EAO, NFVO and MEC Application Orchestrator (MEAO). It is tasked with providing and orchestrating Applications hosted by the ETHER structure from E2EAO to both the Cloud and MEC domains. It will be described as the *Mm1* interface [37] and *Os-Ma-nfvo* interface [118] specified by ETSI.

RP.E2ENO-GMMF: This reference point establishes the communication between the E2ENO and the GMMF. It exposes mobility management data collected by the

GMMF from all the infrastructures across the different domains. It will be developed using standard protocols established by the Open Geospatial Consortium, such as WFS, WMS, and WCS defined in [117].

RP.E2ENO-ADL: This reference point establishes the communication between the E2ENO, NFVO, MEAO and SMO. It is tasked with enabling the creation and management of new network slices, and providing and orchestrating services hosted by the ETHER structure towards the Cloud, MEC, and RAN domains. It will be described as the *Mm1* interface [37] and *Os-Ma-nfvo* interface [118] specified by ETSI.

RP.MSDNO-SDNO: This reference point establishes the communication between the MSDNO and the SDNO, and will be done by applying ONF-oriented SDN specification to define SDNO northbound interface as a transport API (TAPI) [119].

RP.GMMF-ADL: This reference point established the communication between the GMMF and the respective DMMF from the Cloud, MEC, RAN and Transport Network domains. It is tasked with discovering and registering available domains and updating the GMMF of any changes. It will be developed using standard protocols established by the Open Geospatial Consortium, such as WFS, WMS, and WCS, defined in [117].

RP.CL-2: This reference point establishes the communication between the NFVO and the VNFM(s) in the Cloud Domain and will be described as the *Or-Vnfm* interface [120] specified by ETSI.

RP.CL-1: This reference point establishes the communication between the VNFM(s) and the VIM(s) in the Cloud Domain and will be described as the *Vi-Vnfm* interface [121] specified by ETSI.

RP.CL-DMMF: This reference point establishes the communication between the DMMF and the LMMF of the Cloud domain. It is tasked with discovering LMMF instances on the Cloud domain and reporting mobility data from the LMMF to the DMMF through a continuous communication channel. It will be developed using standard protocols established by the Open Geospatial Consortium, such as WFS, WMS, and WCS, defined in [117].

RP.CL-LMMF: This reference point establishes the communication between the LMMF and the VIM(s) of the Cloud Domain. It is tasked with the registration of the local physical components of the Cloud Domain and registering mobility patterns of said physical components as mobility data, through frequent position updates. It will be developed using standard protocols established by the Open Geospatial Consortium, such as WFS, WMS, and WCS, defined in [117].

RP.MEC-2: This reference point establishes the communication between the MEAO and the MEPM(s) in the MEC Domain and will be described as the *Mm3* interface [37] specified by ETSI.

RP.MEC-1: This reference point establishes the communication between the MEPM(s) and the VIM(s) in the MEC Domain and will be described as the *Mm6* interface [37] specified by ETSI.

RP.MEC-DMMF: This reference point establishes the communication between the DMMF and the LMMF of the MEC domain. It is tasked with discovering LMMF instances on the MEC domain and reporting mobility data from the LMMF to the DMMF through a continuous communication channel. It will be developed using standard protocols established by the Open Geospatial Consortium, such as WFS, WMS, and WCS, defined in [117].

RP.MEC-LMMF: This reference point establishes the communication between the LMMF and the VIM(s) of the MEC Domain. It is tasked with the registration of the local physical components of the MEC Domain and registering mobility patterns of said physical components as mobility data, through frequent position updates. It will be developed using standard protocols established by the Open Geospatial Consortium, such as WFS, WMS, and WCS, defined in [117].

RP.RAN-2: This reference point establishes the communication between the Non-RT RIC and the Near-RT RIC in the RAN Domain and will be described as the *A1* interface [116] specified by the O-RAN Alliance group.

RP.RAN-1: This reference point establishes the communication between the SMO and the Near-RT RIC in the RAN Domain and will be described as the *O1* interface [116] specified by the O-RAN Alliance group.

RP.RAN-DMMF: This reference point establishes the communication between the DMMF and the LMMF of the RAN domain. It is tasked with discovering LMMF instances on the RAN domain and reporting mobility data from the LMMF to the DMMF through a continuous communication channel. It will be developed using standard protocols established by the Open Geospatial Consortium, such as WFS, WMS, and WCS, defined in [117].

RP.RAN-LMMF: This reference point establishes the communication between the LMMF and the SMO of the RAN Domain. It is tasked with the registration of the local physical components of the RAN Domain and registering mobility patterns of said physical components as mobility data, through frequent position updates. It will be developed using standard protocols established by the Open Geospatial Consortium, such as WFS, WMS, and WCS, defined in [117].

RP.TN-2: This reference point establishes the connection between the SDNO and the WIM(s) of the Transport Network Domain. The communication between them will follow the ONF-oriented SDN specification to define its northbound transport API (TAPI) [119].

RP.TN-1: This reference point establishes the connection between the WIM(s) and the SDNC of the Transport Network Domain. The communication between them will follow the ONF-oriented SDN specification to define its northbound transport API (TAPI) [119].

RP.TN-DMMF: This reference point establishes the communication between the DMMF and the LMMF of the Transport Network domain. It is tasked with discovering LMMF instances on the Transport Network domain and reporting mobility data from the LMMF to the DMMF through a continuous communication channel. It will be developed using standard protocols established by the Open Geospatial Consortium, such as WFS, WMS, and WCS, defined in [117].

RP.TN-LMMF: This reference point establishes the communication between the LMMF and the WIM(s) of the Transport Network Domain. It is tasked with the registration of the local physical components of the Transport Network Domain and registering mobility patterns of said physical components as mobility data, through frequent position updates. It will be developed using standard protocols established by the Open Geospatial Consortium, such as WFS, WMS, and WCS, defined in [117].

4.7.2.2 E2ESL

RP.APP-5GC: This reference point is used for communication between the AP and 5GS CP, namely the Application server and NEF as an entry point to the 5GS CP mechanisms and data. The communication will depend on the application context specificity and include both general mechanisms exposed by standard 5G CN NFs via NEF and application-specific AF optionally embedded in the 5G CN CP. The exchange will be developed based on the 3GPP-standardised mechanisms like CAPIF [53], SEAL [54], or EDGE-APP [81].

4.7.3 Intra- and inter-layer reference points lists

Below are listed the inter-layer (cf. Table 4-3) and intra-layer (Table 4-4) reference points of the ETHER framework, including their end-points identification and specification/technology details, including the information about exploiting already existing standard or necessity of a new definition.

Table 4-3: Inter-layer interface list

Layer	RP ID	Termination Point 1	Termination Point 2	Specification	Technology
E2ESL/AI Layer	RP.AI-E2ESL	5GC	AI Layer	This reference point between E2ESL and the AI Layer exposes AI-related services to the E2E SL, including all the associated applications, communication, and steering mechanisms.	REST API, gRPC
E2ESL/ETHER MANO	RP.APP-E2EAO	App	E2EAO	This reference point between the application server and the E2EAO facilitates the deployment of applications with particular compute characteristics to a specific node or infrastructure, and manages the life cycle of the application.	REST API, gRPC

Layer	RP ID	Termination Point 1	Termination Point 2	Specification	Technology
ETHER MANO/AI Layer	RP.AI-MANO	E2E SFs	AI Layer	This reference point between E2E ETHER MANO and the AI Layer exposes AI-related services to the E2E ETHER MANO, including all the associated applications, communication, and steering mechanisms.	REST API, gRPC
ETHER MANO/AI Layer	RP.AI-TN	Transport Domain SFs	AI Layer	This reference point allows the AI Layer to communicate with the Transport network domain orchestrators and controllers, putting AI operations upfront onto the Transport networks in the ETHER structure.	REST API, gRPC
ETHER MANO/AI Layer	RP.AI-RAN	RAN Domain SFs	AI Layer	This reference point allows the AI Layer to communicate with the RAN domain orchestrators and controllers, putting AI operations upfront onto the RAN networks in the ETHER structure.	REST API, gRPC
ETHER MANO/AI Layer	RP.AI-MEC	MEC Domain SFs	AI Layer	This reference point allows the AI Layer to communicate with the MEC network domain orchestrators and controllers, putting AI operations upfront onto the MEC networks in the ETHER structure.	REST API, gRPC
ETHER MANO/AI Layer	RP.AI-CL	Cloud Domain SFs	AI Layer	This reference point allows the AI Layer to communicate with the Cloud network domain orchestrators and controllers, putting AI operations upfront onto the Cloud networks in the ETHER structure.	REST API, gRPC
ETHER MANO/Infrastructure	RP.CL-0	VIM(s)	Domain Infrastructure (T/A/S)	ETSI GS NFV-IFA 019	*Nf-Vi*
ETHER MANO/Infrastructure	RP.MEC-0	VIM(s)	Domain Infrastructure (T/A/S)	ETSI GS MEC 003	*Mm7*
ETHER MANO/Infrastructure	RP.RAN-0	SMO	Domain Infrastructure (T/A/S)	O-RAN.WG1.OAD-R003-v12.00	*O2*
ETHER MANO/Infrastructure	RP.TN-0	SDNC	Transport Infrastructure (WAN(s))	ONF TS-001	OF protocol

Table 4-4: Intra-layer interface list

Type	ID	Termination Point 1	Termination Point 2	Specification	Technology
ETHER MANO	RP.E2EAO-GISF	E2EAO	GISF	OGC Web Services Common Specification	WMS, WFS, WCS
ETHER MANO	RP.E2EAO-GMMF	E2EAO	GMMF	OGC Web Services Common Specification	WMS, WFS, WCS
ETHER MANO	RP.AO-NO	E2EAO	E2ENO	OF-based routing specification	REST API
ETHER MANO	RP.E2EAO-ADL	E2EAO	NFVO + MEAO	ETSI GS NFV-IFA 013, ETSI GS MEC 003	ETSI NFV – *Os-Ma-nfvo*; ETSI MEC – *Mm1*
ETHER MANO	RP.E2ENO-GMMF	E2ENO	GMMF	OGC Web Services Common Specification	WMS, WFS, WCS
ETHER MANO	RP.E2ENO-ADL	E2ENO	NFVO + MEAO + SMO	ETSI GS NFV-IFA 013, ETSI GS MEC 003	ETSI NFV – *Os-Ma-nfvo*; ETSI MEC – *Mm1*
ETHER MANO	RP.MSDNO-SDNO	MSDNO	SDNO	ONF SDN specification	REST API (TAPI)
ETHER MANO	RP.GMMF-ADL	GMMF	DMMF CL + DMMF MEC + DMMF RAN + DMMF TN	OGC Web Services Common Specification	WMS, WFS, WCS
ETHER MANO	RP.CL-2	NFVO	VNFM(s)	ETSI GS NFV-IFA 007	*Or-Vnfm*
ETHER MANO	RP.CL-1	VNFM(s)	VIM(s)	ETSI GS NFV-IFA 006	*Vi-Vnfm*
ETHER MANO	RP.CL-DMMF	DMMF Cloud	LMMF	OGC Web Services Common Specification	WMS, WFS, WCS
ETHER MANO	RP.CL-LMMF	LMMF	VIM(s)	OGC Web Services Common Specification	WMS, WFS, WCS
ETHER MANO	RP.MEC-2	MEAO	MEPM(s)	ETSI GS MEC 003	*Mm3*
ETHER MANO	RP.MEC-1	MEPM(s)	VIM(s)	ETSI GS MEC 003	*Mm6*
ETHER MANO	RP.MEC-DMMF	DMMF MEC	LMMF	OGC Web Services Common Specification	WMS, WFS, WCS
ETHER MANO	RP.MEC-LMMF	LMMF	VIM(s)	OGC Web Services Common Specification	WMS, WFS, WCS
ETHER MANO	RP.RAN-2	SMO	Near-RT RIC	O-RAN.WG1.OAD-R003-v11.00	*A1*
ETHER MANO	RP.RAN-1	Near-RT RIC	Non-RT RIC	O-RAN.WG1.OAD-R003-v11.00	*O1*
ETHER MANO	RP.DMMF RAN	DMMF RAN	LMMF RAN	OGC Web Services Common Specification	WMS, WFS, WCS
ETHER MANO	RP.LMMF RAN	LMMF RAN	SMO	OGC Web Services Common Specification	WMS, WFS, WCS
ETHER MANO	RP.TN-2	SDNO	WIM(s)	ONF SDN specification	REST API (TAPI)

Type	ID	Termination Point 1	Termination Point 2	Specification	Technology
ETHER MANO	RP.TN-1	WIM(s)	SDNC	ONF SDN Specification	REST API (TAPI)
ETHER MANO	RP.TN-DMMF	DMMF transport	LMMF	OGC Web Services Common Specification	WMS, WFS, WCS
ETHER MANO	RP.TN-LMMF	LMMF	WIM(s)	OGC Web Services Common Specification	WMS, WFS, WCS
E2ESL	RP.APP-5GC	App	NEF	3GPP TS 23.222, 3GPP TS 23.434, 3GPP TS 23.558	CAPIF, SEAL, EDGEAPP

5 ARCHITECTURE EVALUATION

In this section, the ETHER architecture, described in Section 4, is evaluated in terms of key 6G performance metrics, i.e., energy- and cost-efficiency. These metrics closely map ETHER's goal, which is to maintain the energy consumption and the TCO of the next-generation mobile networks as low as possible, while satisfying the UEs' QoS. To that end, in this section, the proposed 3D ETHER architecture, incorporating its aforementioned components and their requirements, is evaluated under different traffic distribution scenarios, shedding light on its performance in terms of energy- and cost-efficiency. To that end, first, the new employed models are analysed and then the simulation scenario and results are presented, while interesting insights are gained and analysed. Please note that further evaluation of the ETHER components is expected to be performed in the respective tasks of WP3 and WP4, as the corresponding work progresses, and are, thus, to be reported in the WP3's and WP4's upcoming deliverables, i.e., D3.1, D3.2, D4.1 and D4.2.

5.1 EMPLOYED MODELS

The power model for the access network is presented in detail in [122] where a similar model has been configured, but studies only the access network for each type of BS. The inclusion of X-haul links as well as the computational nodes to the previous model, makes it a complete network model, and the additional model parameters are listed below.

Table 5-1: Notation table of the parameters used in the employed models

Notation	Definition
P_{pm}	Power consumption of the physical machine
$p_{idle_{pm}}$	Idle power consumption of the physical machine
$p_{max_{pm}}$	Maximum power consumption of the physical machine
θ_{pm}^{CPU}	Percentage of the CPU load usage of the physical machine
P_{switch}	Fibre X-haul power consumption
$p_{idle_{switch}}$	Idle power consumption of the switch
P_i	Power consumption of port *i* of the switch
P_{BH}	Wireless X-haul power consumption
N_{TRX_l}	Number of transceiver chains of the wireless X-haul link *l*
P_{0_l}	Idle power consumption of the wireless X-haul link *l*
Δ_{p_l}	Slope of the load-dependent power consumption of the wireless X-haul link *l*
P_{out_l}	Output power of the RF transceiver of the wireless X-haul link *l*
$SINR_l^{trg}$	Signal to Interference plus Noise Ratio (SINR) target for successful communication between the X-haul nodes of the wireless X-haul link *l*
β_l	Total losses of the system, subtracting its gains for the wireless X-haul link *l*
L_{TX_l}	Transmitter losses of the wireless X-haul link *l*
L_{RX_l}	Receiver losses of the wireless X-haul link *l*
PL_l	Total path loss of the wireless X-haul link *l (*includes the free space path loss and the atmospheric loss)
LM	Link margin for the wireless X-haul links

Notation	Definition
N_{TH}	Thermal noise for the wireless X-haul links
NF_{BH}	Noise figure for the wireless X-haul link receiver
G_{TX_l}	Transmitter's antenna gain of the wireless X-haul link *l*
G_{RX_l}	Receiver's antenna gain of the wireless X-haul link *l*

5.1.1 Computational node power consumption model

To calculate the power consumption (P_{pm}) for the physical machines, in Watts, it holds:

$$P_{pm} = p_{idle_{pm}} + \left(p_{max_{pm}} - p_{idle_{pm}}\right) \theta_{pm}^{CPU}, \tag{1}$$

where $p_{idle_{pm}}$ is the power consumption of each physical machine, while idle, $p_{max_{pm}}$ is the maximum possible power consumption and the last part, θ_{pm}^{CPU} is the percentage of the CPU load usage.

5.1.2 X-haul power model

The X-haul power model is divided into the fibre and the wireless X-haul network power consumption.

To calculate the fibre X-haul network power consumption (P_{switch}), in Watts, it holds:

$$P_{switch} = p_{idle_{switch}} + \sum_{i \in ports} P_i, \tag{2}$$

where $p_{idle_{switch}}$ is the power that each switch consumes when idle and P_i is the power consumption of each port of the switch multiplied by the total number of active ports.

For the wireless X-haul power consumption network (P_{BH}), the calculation is more complex. The main equation is:

$$P_{BH} = \sum_{l \in L} N_{TRX_l}\left(P_{0_l} + \Delta_{p_l} P_{out_l}\right), \tag{3}$$

where N_{TRX_l} is the number of transceiver chains, P_{0_l} is the power consumed in the wireless X-haul link while idle, Δ_{p_l} is the slope of the load-dependent wireless power consumption and P_{out_l} is the output power of the RF transceiver. The output of the equation is measured in Watts. To calculate P_{out_l}:

$$P_{out_l} = SINR_l^{trg} + \beta_l, \tag{4}$$

Where $SINR_l^{trg}$ is the necessary Signal to Interference plus Noise Ratio (SINR) for successful communication between the X-haul nodes, which is explained thoroughly in [122], and β_l is the total losses of the system, subtracting its gains, which is calculated, in dBm, as follows:

$$\beta_l = L_{TX_l} + L_{RX_l} + PL_l + LM + N_{TH} + NF_{BH} - G_{TX_l} - G_{RX_l}, \quad (5)$$

where L_{TX_l} and L_{RX_l} are the losses in the transmitter and the receiver, respectively, PL_l is the total path loss, which includes the free space path loss and the atmospheric loss, LM is the link margin for the wireless X-haul links, N_{TH} stands for the thermal noise, NF_{BH} is the noise figure of the X-haul receiver and G_{TX_l} and G_{RX_l} are the transmitter's and receiver's antenna gains.

5.2 PERFORMANCE EVALUATION

5.2.1 Studied algorithms

For the performance evaluation, the proposed network deployment algorithm, focuses on maximising the network energy efficiency in accordance with the ETHER's KPIs, whereas the reference SotA algorithm aims at reducing the total E2E delay of the network. The two algorithms study the joint any type of Network Function (xNF) placement, traffic routing and user association, taking into account both the access and X-haul networks.

The two algorithms have been split into two steps each. For the proposed algorithm in the initial step, a weighted graph is constructed and all potential paths from the source to the destination are evaluated based on power consumption. Each path includes all feasible wireless and fibre X-haul transport links, as well as the access network link between the serving BS and the UE. The path with the lowest power consumption, or the shortest-weighted path, is selected to meet the UE demands, provided that the capacity and delay constraints are not violated. If a violation occurs, it chooses the next available shortest path without any constraint violations. If no such path exists, it blocks the UE.

Once a path is chosen, it proceeds to the next step, which is xNF placement. For placing each xNF of the requested Service Function Chain (SFC) in the available nodes specified by the chosen path, the nodes are sorted by a parameter denoted by Ω. This parameter includes the node's closeness centrality, maximum computational capacity, and CPU load. The CPU load is categorised into four values: a) 1 (high priority) when the node has sufficient computational capacity and can host the xNF without initiating a new instance, b) 0.5 (medium priority) when the node is active and has sufficient capacity but requires a new instance initiation, c) 0.1 (low priority) when the node is idle and needs to be activated and be able to host the potential xNF and d) (no priority) when the node cannot host the xNF. After sorting, it selects the highest-ranked node and places the xNF, ensuring computational constraints are met. If no suitable node is found, it returns to the first step to select the next shortest path, repeating the process until all xNFs are placed or no more paths are available, at which point the UE is blocked.

The proposed heuristic algorithm is compared with a SotA approach, which has been appropriately adapted for the scenarios under study to ensure a fair comparison. The SotA first conducts the xNF placement based on the highest betweenness centrality and then handles traffic routing with the goal of minimising delay, ensuring that the E2E latency constraint is not breached. Since the SotA does not consider the user association problem, the default criterion has been employed where UEs connect to the BS from which they receive the strongest signal power.

5.2.2 Simulation scenario

Extensive simulations have been executed in MATLAB 2023b. The simulations had 100 samples per user state and consisted of 10 different scenarios where the BSs were changing posi-

Co-funded by the European Union 6GSNS

tion in each one of them. Each scenario had 10 different snapshots where BSs were stable, and the users were positioning differently in each one of them. To simulate the various network loads, the tests were run for 10, 20, 30 and 40 users, meaning that the traffic in each case was increased, as well as the network demands, making it more complex and demanding at each step.

The 3D network under study consists of 1 macrocell (gNB), 2 small cell clusters each with 4 small cells (SCs), numbering a total of 9 terrestrial BSs, 1 HAPS in an altitude of 20 km, multiple LEO satellites, specifically 2 from Iridium, 28 from Eutelsat/OneWeb and 104 from Starlink at 770 km, 1200 km and 550 km of altitude on average, respectively. Two GEO satellites are also considered, HellasSat 3 and 4, at the altitude of 35786 km, being the regular altitude of GEO satellites.

For the results, a wide range of different SFCs have been used that simulate both everyday casual usage and a few commercial scenarios, such as the Tracking, Telemetry and Command (TT&C) by satellite operators to control and observe their fleet. In Table 5-2, every SFC and its details are presented, along with their percentage of share in the available users. Please note that these SFC types include the proposed ETHER UCs, i.e., the TT&C applications include the air space safety UC (ETHER UC3), the IoT applications the delay-tolerant applications of UC1 and the rest of the SFC types the direct handheld device access of the ETHER UC2.

Table 5-2: Studied SFC types

SFC Type	Rate (Mbps)	Latency (ms)	Share (%)
Web	0.6-1	500	15
VoIP	0.384-0.64	100	15
Streaming	5-24	100	30
Gaming	0.24-0.5	60	10
Ultra RT AI/ML	15-25	1	10
IoT Applications	0.1-0.5	400	10
TT&C Applications	1-5	250	10

The extra parameters used for the X-haul network are presented in Table 5-3, where TN indicates terrestrial fibre X-haul network, TN-NTN indicates mmWave X-haul links between terrestrial and non-terrestrial nodes, HTS indicates HAPS to satellite nodes mmWave X-haul links. ISLs are used exclusively between satellite nodes to communicate and operate in the optical spectrum, utilising the Free Space Optical technology. The use of optical spectrum constitutes the most preferred way of the newer and future satellites to communicate, since it provides notable higher capacity than RFs.

Table 5-3: Simulation parameter values

Parameter	TN	TN-NTN	HTS	ISL
f (GHz)	60	28	28	200000
BW (GHz)	0.2	0.6	0.6	20
N_{TRX_i}	32	32	64	128
P_{0_i} **(W)**	3.9			
Δ_{p_i}	5	4	4	10

For the TCO calculation, the TCO prices from [122] have been used, along with some new values to calculate the additional X-haul power consumption cost and machine purchase cost, which are presented in Table 5-4.

Table 5-4: CAPEX per computational node based on its type (MEC, Fog, Cloud) and number of cores

Computational nodes	Number of cores	CAPEX (€)
MEC	8	1362
	10	2208
Fog	24	8569
	32	10231
Cloud	48	13352
	64	16933

As a potential cost figure for the MEC nodes, the NVIDIA Jetson Xavier [123] and Orin [124] have been selected, both in their industrial editions, for the 8-core and the 10-core, respectively. For the Fog and Cloud Computing nodes, RAX QT24-12E9 has been chosen [125], differently configured each time to fit the occasion and the core number, setting the X-haul CAPEX. Each of the BSs with computing capabilities, is equipped with a physical machine. In particular, the terrestrial BSs, the HAPS and the LEO satellites are equipped with MEC capabilities, the 1st aggregation layer terrestrial nodes and the MEO satellites with fog computing capabilities, and the 2nd aggregation layer terrestrial nodes as well as the GEO satellites with cloud computing capabilities. To calculate the CAPEX, when a xNF is deployed in one of the computing nodes, and it initiates its physical machine, the cost of that machine is added for the specific scenario. If additional xNFs are also deployed in the same physical machines, the CAPEX cost will not change, since it has already been purchased, but instead the power consumption (OPEX) will be impacted (specifically it will be increased).

To calculate the power consumption cost (which constitutes the main part of the OPEX of the computational nodes), the average kilowatt-per-hour price has been set to fit the average price in Europe (€0.2008 per kWh for the second half of 2023, according to Eurostat [126]). This considers the power consumption from mmWave X-haul links, fibre X-haul links, physical machines (the computational nodes) and switches that utilised for each user and multiplied by the total number of hours of a year to set the total OPEX. Finally, the CAPEX (purchase cost) and the OPEX (power consumption cost) are summed up, setting the TCO.

5.2.3 Results

The average energy efficiency of the algorithms for the different traffic loads is depicted in Figure 5-1. The proposed algorithm achieves 78% higher energy efficiency than the SotA for high traffic load conditions, i.e., for 40 users, minimising the total energy consumption of the network, without violating the required user needs.

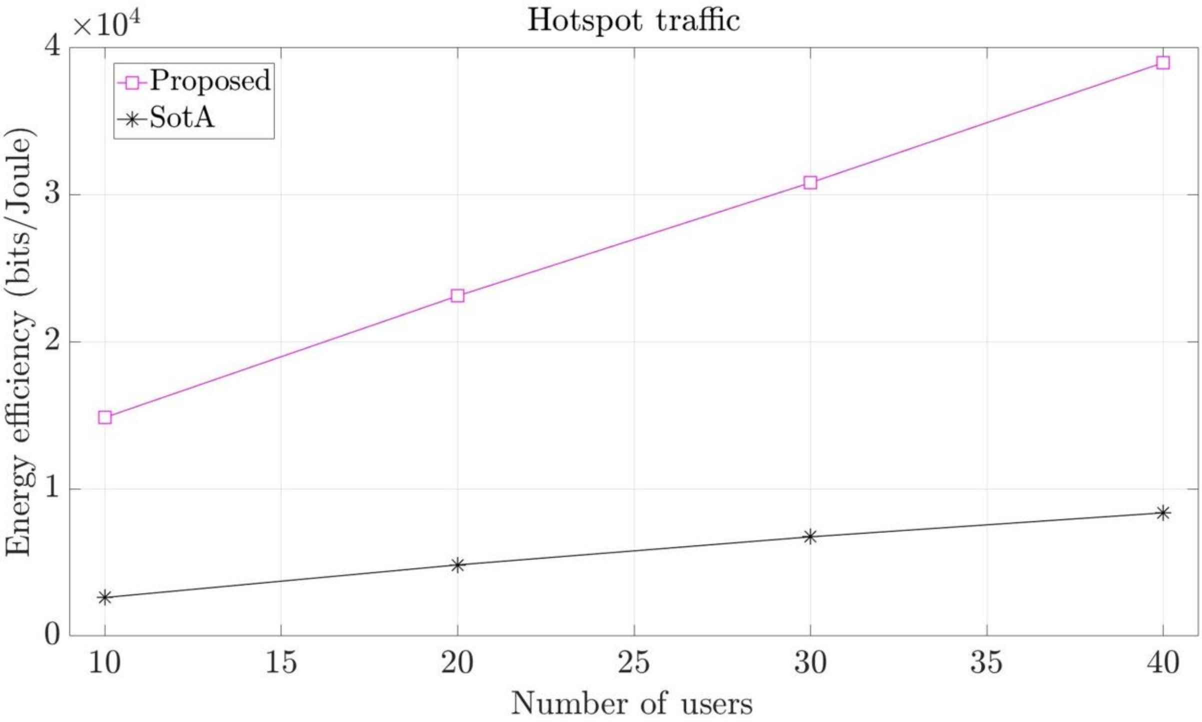

Figure 5-1: Energy efficiency versus different traffic load conditions for both the proposed and SotA offline resource allocation algorithms under the ETHER 3D architecture

At each traffic load the proposed algorithm is shown to be more energy efficient that the SotA, at the expense, however, of a little higher execution time because of its higher overall complexity, as shown in Figure 5-2. Still, unlike online resource allocation problems, where network orchestration decisions need to be made in real-time, for network deployment problems, where an offline network optimisation is required, the high energy efficiency gains at the expense of a little higher complexity, is a price worth paying for network deployment optimisation.

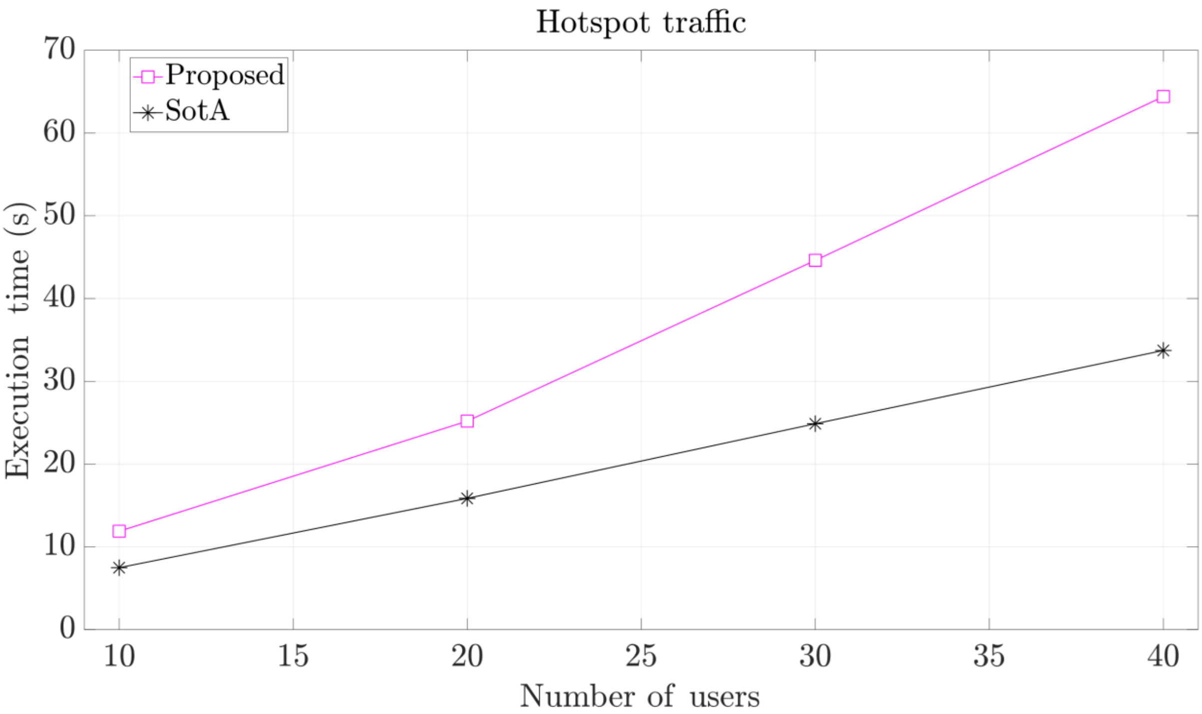

Figure 5-2: Execution time versus different traffic load conditions for both the proposed and SotA offline resource allocation algorithms under the ETHER 3D architecture

An important note is that for both algorithms the user acceptance was 100%, meaning that no user was unsatisfied/blocked, as shown in Figure 5-3.

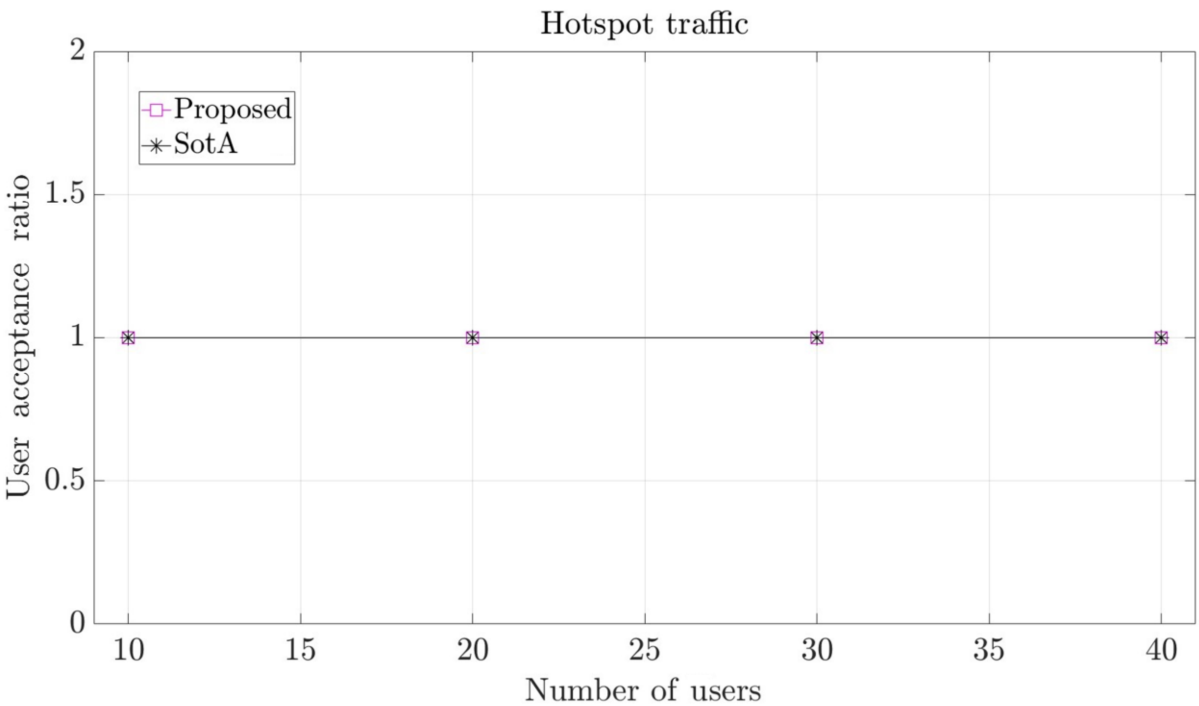

Figure 5-3: User acceptance ratio versus different traffic load conditions for both the proposed and SotA offline resource allocation algorithms under the ETHER 3D architecture

In the following figure, the power breakdown of each algorithm is analysed for the different BSs and backhaul links. As shown, the proposed algorithm succeeds to maintain a notably lower power consumption compared to SotA because of its higher flexibility. In particular, in the SotA, the serving BSs are already determined based on the best SINR, and then the xNF placement and traffic routing are performed. Hence, the main difference between the compared algorithms is that the SotA prefers to mainly deploy the macro cell BS (gNB), which is remarkably more power-hungry than the rest of the available nodes, thus increasing its total consumption and making it unfavourable for commercial deployment. It is also worth noting that the power consumption of the proposed deployment solution scales better than the SotA with increasing load, as shown in Figure 5-4, thus making it a suitable choice for highly complex traffic-demanding 6G network deployment problems.

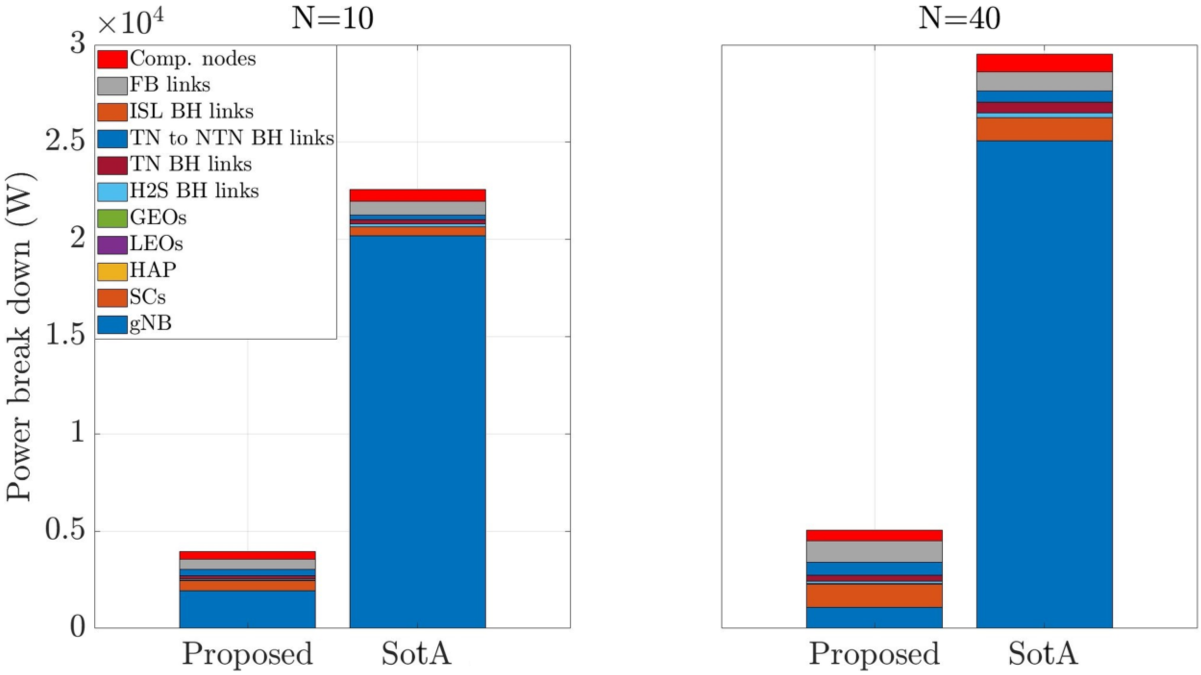

Figure 5-4: Power breakdown for different traffic load conditions (low and high traffic) for both the proposed and SotA offline resource allocation algorithms under the ETHER 3D architecture [Comp.: Computing, FB: fibre, BH: Backhaul, H2S: HAP to Satellite, SC: Small Cell]

As depicted in Figure 5-5, the proposed algorithm achieves to maintain a significantly lower average TCO than the SotA, achieving up to 94% lower overall TCO in the scenarios with 40 users (meaning in capacity dependent scenarios) and up to 23% in lower traffic scenarios (10 users). This can be explained as the proposed algorithm deploys fewer physical machines and BSs to achieve its goal, thus, requiring less power to be consumed and fewer machines to be purchased.

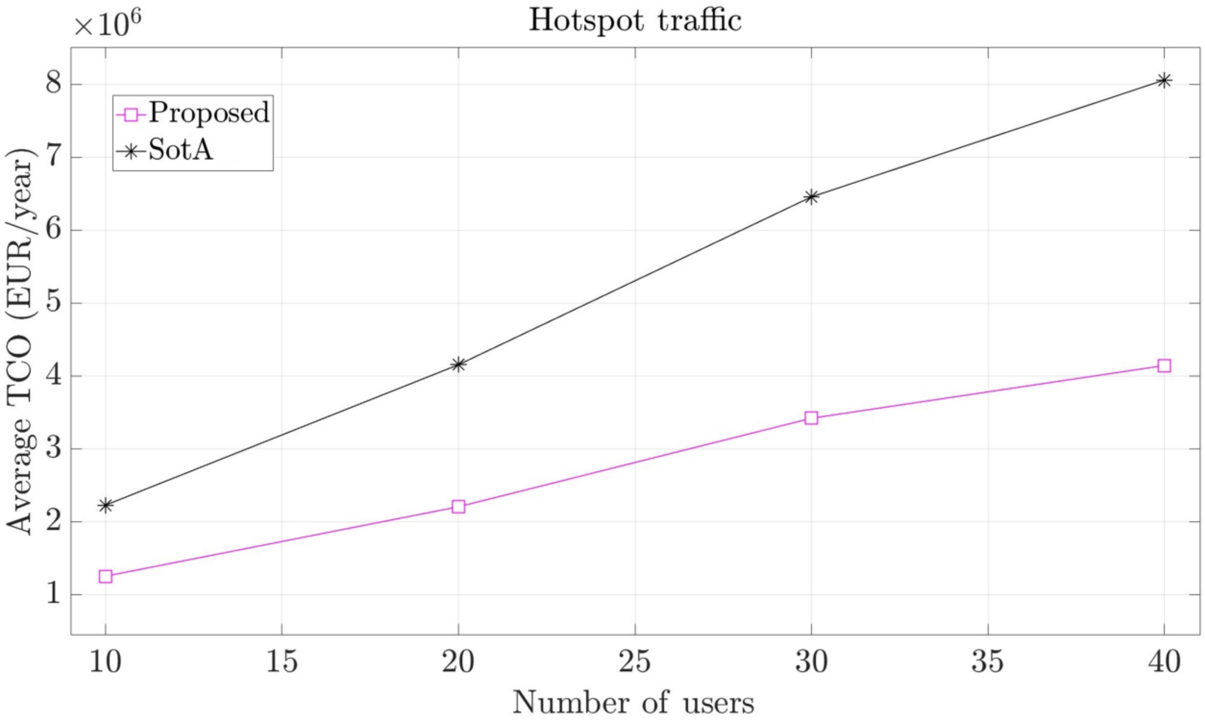

Figure 5-5: Average TCO versus different traffic load conditions for both the proposed and SotA offline resource allocation algorithms under the ETHER 3D architecture

As a result, the ETHER 3D architecture is shown to achieve very high energy- and cost-efficiency performance, especially when combined with the offline resource allocation algorithms developed within the project and analysed in Section 5.2.1. In parallel, the performance gains between the proposed and the SotA solution dictate the necessity to focus on developing sustainable networks, which will allow the emerging networks to be affordable, allowing more users with demanding services in the network, while maintaining the total TCO as low as possible.

6 SUMMARY AND CONCLUSIONS

In this deliverable, the final functional ETHER network architecture, which enables the integration of terrestrial, aerial and space strata has been presented. The architecture design has been driven by the following main factors: i) the envisioned 6G architecture, paradigms, expected capabilities, and targets; ii) the most recent SotA approaches to the integrated 3D network architectures; iii) ETHER-specific requirements originating from the UCs that will be implemented in WP5 (established within T2.1 activities and presented in D2.2); iv) key issues and challenges concerning TN-NTN integration as perceived by the SDOs and the academia; v) required system-level support for ETHER technical innovations. To this end, the final ETHER architecture is composed of 4 layers, namely, Infrastructure, E2E Service, MANO and AI layers.

The Infrastructure Layer is constituted by the 3D multi-layered network infrastructure comprised of TN and NTN assets (space, HAPSs, aerial) integrated both horizontally and vertically. The Infrastructure Layer is further decomposed into following strata: i) Terrestrial – radio, wireless, cellular, ground satellite, and small cells that operate at various frequency bands, but mostly sub-6 GHz and mmWave); ii) Aerial – commercial airplanes (aircraft), UAVs, and HAPS that fly at altitudes of >1 km, 8-11 km, and 17-20 km; iii) Space – LEO, MEO, GEO satellites. The considered user devices also fall under this category and include both flying objects (aircraft, UAVs, space IoT devices) and terrestrial objects like ground users or IoT devices.

The ETHER E2ESL is constituted by the Application, Communication, and Steering strata and generally embed the components of 3GPP 5GS SA architecture enhanced with extensions facilitating the system operation in the integrated TN-NTN ecosystem. This layer embeds the following key features and mechanisms: i) Store and Forward to allow conducting network procedures to address the NTN topology sparsity and limited activity time of ISLs; ii) distributed 3GPP architecture for NTNs; iii) dynamic topology management to address the gNB mobility in the NTN stratum; iv) SDN-based UPF for the flexible integration with network-wide SDN-based approach to MANO; v) AI-based handovers for service continuity; vi) semantic-aware analytics for energy and resources saving; vii) and direct handheld access to allow multi-RAT connectivity for UEs. As the standardisation of 6G has not yet started, the base communication system considered by ETHER is 5G Advanced. Therefore, the verification of support by the 5G components and its openness to extensions has been conducted, which proved that innate 5G capabilities such as SBA, generic AF or CP exposure mechanisms via NEF allow for tight integration with external systems such as MEC, ETHER MANO components, SDN-based transport or AI Layer components.

The ETHER MANO Layer features a multi-level hierarchical and distributed multi-domain system composed of generic self-contained domains (ETHER MANO Domains) each containing DMOC and domain-specific SFs implementing functionalities needed by the individual domains and following a microservice-based approach (such as OSS/BSS functionality, components allowing integration with AI Layer, optimisation of selection of infrastructure providers, etc.). The E2E M&O is accomplished by the EMOC component assisted with SFs, which coordinates the behaviour of each domain (the umbrella MANO for the federation of domains). The generic MANO architecture proposal is followed by the example of instantiation using the well-established standardised solutions (ETSI NFV MANO, ETSI MEC, O-RAN), which are planned to be the foundation for the ETHER UCs demonstrations. Moreover, the concept of a lightweight Transport Network domains encapsulating the SDN-based WIM environment is proposed, allowing for fast data path reconfigurations across the data-centres and strata (i.e., terrestrial, aerial, satellite). The overall architecture can be hence seen as the 3D mesh of domains (with terrestrial, aerial and satellite variants of each administrative/technological domain) of specific types interconnected by the set of Transport Domains (the domain gateways visible from the E2E level).

To address the AI-nativeness of the next-generation networks, the ETHER architecture incorporates the AI Layer, which is based on the NIP concept introduced by the EU project DAEMON. The approach defines a multi-level AI MANO stack (NIF-C Manager, NIF Manager and NI Orchestrator), follows a similar concept to ETSI NFV MANO and is synergic to the ETHER MANO Layer. The N-MAPE-K model, specifically tailored to the mobile network environment, is adopted that extends the MAPE-K concept by addition of original training and closed control loops that a NIF may implement. Each AI-driven ETHER component adheres to the NI-specific requirements.

Moreover, the intra- and inter-layer interactions have been defined (e.g., CP and MP to manage the vertical handover operations or perform E2E network optimisation) together with the reference points and the interfaces between components (including novel and the ones already standardised by the SDOs). The in-detail specification will serve as the ICD for the development of the UCs in WP5 and can serve for general architecture instantiation by interested parties.

The proposed final ETHER architecture contributes to multiple challenges posed by the next-generation networks. First, the 6G is envisioned as a highly distributed and modular ecosystem constituting the “Network of Networks”, composed of multiple self-contained, multi-access and multi-vendor environments, which are AI- and cloud-native by design. The ETHER architecture addresses this challenge by defining a hierarchical multi-domain MANO architecture, in which several stakeholders can participate by offering domain-level resources by exposing DMOC interfaces to the top-level EMOC entity owned and operated by the MNO. The flexible domain definition also contributes to the issues of overall system complexity reduction. The network components and operations disaggregation and softwarisation and domain differentiation (technological, administrative and stratum-levels) allows for domain resources abstractions (i.e., treating a domain as a self-contained environment with *plug-and-play* DMOC component).

One of the key challenges concerning the integration of TN and NTN is caused by the high dynamicity of the network assets, especially LEO satellites. The final ETHER architecture addresses this issue by proposing hierarchical mobility management functions, namely GMMF, DMMF and LMMF (allowing E2E, domain and infrastructure-level mobility handling), which are integrated with standardised 3GPP Management stack and its relevant counterparts, i.e., CSMF, NSMF, NSSMF, and NFMF. Moreover, the novel use of GISF is proposed, which provides geographical coordinates to E2E orchestrators for efficient network resources management and optimisation across the mapped areas as well as establish the interconnect point for external and online GIS systems for increased comprehension and accuracy.

The ETHER architecture design allows for a cascade of new possibilities. Technical innovations of ETHER work in tandem with the proposed integrated architecture (T-1). Integrated architecture employing virtualisation and MANO (T-7) allows for reconfiguration of software-defined-radios onboard satellites. This is reconfigurability is considered as flexible payload (T-4), and in turn, allows for construction of virtual arrays by the LEO satellite swarms, which helps in achieving the direct handheld device access at the Ka band (T-2). Unified waveform design (T-3), which proposes new OTFS modulation is further focused on improving the stability of the connection in high-doppler environments such as LEO channels. 5G NR already contains a waveform switching mechanism, which justifies research in this field. Data analytics, edge computing, and caching (T-5) are possible thanks to inclusion of MEC into the architecture. Introduction of AI Layer further helps incorporating efficient horizontal/vertical handover mechanisms (T-6) as well as E2E network performance optimisation (T-8). The above innovations address the issues and challenges concerning TN-NTN integration described in this document.

Finally, the architecture evaluation is conducted with regards to the key 6G performance metrics that map to the ETHER's project targets (i.e., energy- and cost-efficiency, UEs' QoS maintenance). The undertaken modelling approach considers an E2E network, which includes not only the power consumed by computational nodes but also energy consumed by X-haul devices. The conducted simulations, comprised of different traffic distribution scenarios and corresponding to the ETHER UCs, have shown significant gains in terms of TCO and energy consumption reduction. The presented evaluation shows promising results in the context of WP5 demonstrations, paving the way for the future ETHER-related developments.

7 REFERENCES

[1] ETHER, “Deliverable D2.2 Use cases and KPIs/KVIs”, 2023.

[2] GSMA, “High Altitude Platform Systems”, GSM Association, White Paper, ver. 2.0, February 2022. [Online]. Available: https://www.gsma.com/futurenetworks/wp-content/uploads/2022/02/HAPS-Towers-in-the-skies-draft-v-2.1-clean.pdf. [Accessed August 2024].

[3] 5G Americas, “Update on 5G Non-terrestrial networks”, July 2023. [Online]. Available: https://www.5gamericas.org/wp-content/uploads/2023/07/Update-on-5G-Non-terrestrial-Networks-Id.pdf. [Accessed August 2024].

[4] ETHER, “Deliverable D2.1 Initial report on ETHER network architecture, interfaces and architecture evaluation”, 2023.

[5] NGMN, “6G Drivers and vision”, 19 April 2021. [Online]. Available: https://www.ngmn.org/wp-content/uploads/NGMN-6G-Drivers-and-Vision-V1.0_final.pdf. [Accessed August 2024].

[6] M. A. Uusitalo, P. Rugeland, M. R. Boldi, E. C. Strinati, P. Demestichas, M. Ericson, G. P. Fettweis, M. C. Filippou, A. Gati, M.-H. Hamon, M. Hoffmann, M. Latva-Aho, A. Pärssinen, Björn Richerzhagen, Hans Schotten and Tommy Svensson, “6G Vision, Value, Use Cases and Technologies”, *IEEE Access,* vol. 9, pp. 16004-16020, 2021.

[7] ITU-R, “Framework and overall objectives of the future development of IMT for 2030 and beyond”, M.2160, 11 2023. [Online]. Available: https://www.itu.int/rec/R-REC-M.2160/en. [Accessed August 2024].

[8] Hexa-X II, “Deliverable D1.3 Environmental and social view on 6G”, March 2024. [Online]. Available: https://hexa-x-ii.eu/wp-content/uploads/2024/03/Hexa-X-II_D1.3_v1.00_GA_approved.pdf. [Accessed August 2024].

[9] NGMN, “Non-Terrestrial Networks Position Paper”, 9 December 2019. [Online]. Available: https://www.ngmn.org/publications/ntn-position-paper.html. [Accessed July 2023].

[10] NGMN, “6G Use cases and analysis”, 22 February 2022. [Online]. Available: https://www.ngmn.org/publications/6g-use-cases-and-analysis.html. [Accessed August 2024].

[11] Hexa-X II, “Deliverable D1.2 6G Use Cases and Requirements”, December 2023. [Online]. Available: https://hexa-x-ii.eu/wp-content/uploads/2024/01/Hexa-X-II_D1.2.pdf. [Accessed August 2024].

[12] NGMN, “Cloud Native Manifesto”, September 2023. [Online]. Available: https://www.ngmn.org/wp-content/uploads/NGMN_Cloud_Native_Manifesto.pdf. [Accessed August 2024].

[13] NGMN, “Green Future Networks: Metering in Virtualised RAN Infrastructure”, April 2024. [Online]. Available: https://www.ngmn.org/wp-content/uploads/GFN_Metering-Virt.-RAN-Infrastructure_V1.0.pdf. [Accessed August 2024].

[14] Hexa-X II, “Deliverable D3.3 Initial analysis of architectural enablers and framework”, April 2024. [Online]. Available: https://hexa-x-ii.eu/wp-content/uploads/2024/04/Hexa-X-II_D3.3_v1.0.pdf. [Accessed August 2024].

[15] 5GPPP, “View on 5G Architecture”, October 2021. [Online]. Available: https://5g-ppp.eu/wp-content/uploads/2021/11/Architecture-WP-V4.0-final.pdf. [Accessed August 2024].

[16] 5G-COMPLETE, “Deliverable D2.3: Final report on 5G-COMPLETE network architecture and interfaces”, July 2022. [Online]. Available: https://5gcomplete.eu/wp-content/uploads/2023/05/D2_3.pdf. [Accessed August 2024].

[17] Hexa-X, “Deliverable D6.2: Design of service management and orchestration functionalities”, April 2022. [Online]. Available: https://hexa-x.eu/wp-content/uploads/2022/05/Hexa-X_D6.2_V1.1.pdf. [Accessed August 2024].

[18] ETSI ISG NFV, “Network Functions Virtualisation (NFV); Use Cases”, European Telecommunications Standards Institute, ETSI GS NFV 001 V1.3.1, March 2021. [Online]. Available: https://www.etsi.org/deliver/etsi_gr/NFV/001_099/001/01.03.01_60/gr_NFV001v010301p.pdf.

[19] ETSI ISG NFV, “Network Functions Virtualisation (NFV) Release 4; Management and Orchestration; Architectural Framework Specification”, European Telecommunications Standards Institute, ETSI GS NFV 006 V4.5.1, May 2024. [Online]. Available: https://www.etsi.org/deliver/etsi_gs/NFV/001_099/006/04.05.01_60/gs_nfv006v040501p.pdf.

[20] ETSI ISG NFV, “Network Functions Virtualisation (NFV) Release 5; Architectural Framework; Report on enhanced container networking”, European Telecommunications Standards Institute, ETSI GR NFV-IFA 043 V5.1.1, May 2024. [Online]. Available: https://docbox.etsi.org/ISG/NFV/open/Publications_pdf/Specs-Reports/NFV-IFA%20043v5.1.1%20-%20GR%20-%20enhanced_container_networking_report.pdf.

[21] ETSI ISG NFV, “Network Functions Virtualisation (NFV); Management and Orchestration; Report on Architectural Options”, European Telecommunications Standards Institute, ETSI GS NFV-IFA 009 V1.1.1,

July 2016. [Online]. Available: https://www.etsi.org/deliver/etsi_gs/nfv-ifa/001_099/009/01.01.01_60/gs_nfv-ifa009v010101p.pdf.

[22] ETSI ISG NFV, "Network Functions Virtualisation (NFV) Release 5; Evolution and Ecosystem; Report on Multi-tenancy in NFV", European Telecommunications Standards Institute, ETSI GR NFV-EVE 018 V5.1.1, May 2024. [Online]. Available: https://www.etsi.org/deliver/etsi_gr/NFV-EVE/001_099/018/05.01.01_60/gr_NFV-EVE018v050101p.pdf.

[23] MonB5G, "Deliverable D2.4: Final release of the MonB5G architecture (including security)", October 2021. [Online]. Available: https://www.monb5g.eu/wp-content/uploads/2021/11/D2.4_MonB5G-Architecture.pdf.

[24] ETSI, "Open Source MANO", [Online]. Available: https://osm.etsi.org/. [Accessed August 2024].

[25] M. Khalid, J. Ali and R. Byeong-hee, "Artificial Intelligence and Machine Learning Technologies for Integration of Terrestrial in Non-Terrestrial Networks", *IEEE Internet of Things Magazine,* vol. 7, no. 1, pp. 28-33, 2024.

[26] ETSI, "ETSI Experiential Networked Intelligence", European Telecommunications Standards Institute, [Online]. Available: https://www.etsi.org/technologies/experiential-networked-intelligence. [Accessed August 2024].

[27] ETSI ISG ENI, "Experiential Networked Intelligence (ENI); System Architecture", European Telecommunications Standards Institute, ETSI GS ENI 005 V2.1.1, December 2021. [Online]. Available: https://www.etsi.org/deliver/etsi_gs/ENI/001_099/005/02.01.01_60/gs_ENI005v020101p.pdf.

[28] ETSI ISG ZSM, "Zero-touch network and Service Management (ZSM); Reference Architecture", European Telecommunications Standards Institute, ETSI GS ZSM 002 V1.1.1, August 2019. [Online]. Available: https://www.etsi.org/deliver/etsi_gs/ZSM/001_099/002/01.01.01_60/gs_ZSM002v010101p.pdf.

[29] ETSI ISG ZSM, "Zero-touch network and Service Management (ZSM); Closed-Loop Automation; Part 3: Advanced topics", European Telecommunications Standards Institute, ETSI GR ZSM 009-3 V1.1.1, August 2023. [Online]. Available: https://www.etsi.org/deliver/etsi_gr/ZSM/001_099/00903/01.01.01_60/gr_ZSM00903v010101p.pdf.

[30] ETSI ISG ZSM, "Zero-touch network and Service Management (ZSM); Network Digital Twin", European Telecommunications Standards Institute, ETSI GR ZSM 015 V1.1.1, February 2024. [Online]. Available: https://www.etsi.org/deliver/etsi_gr/ZSM/001_099/015/01.01.01_60/gr_ZSM015v010101p.pdf.

[31] ETSI ISG ZSM, "Zero-touch Network and Service Management (ZSM); ZSM security aspects", European Telecommunications Standards Institute, ETSI GS ZSM 014 V1.1.1, March 2024. [Online]. Available: https://www.etsi.org/deliver/etsi_gs/ZSM/001_099/014/01.01.01_60/gs_ZSM014v010101p.pdf.

[32] Hexa-X, "Deliverable D5.3 Final 6G architectural enablers and technological solutions", April 2023. [Online]. Available: https://hexa-x.eu/wp-content/uploads/2023/05/Hexa-X_D5.3_v1.0.pdf. [Accessed August 2024].

[33] M. Camelo, M. Gramaglia, P. Soto, L. Fuentes, J. Ballesteros, A. Bazco-Nogueras, G. Garcia-Aviles, S. Latré, A. Garcia-Saavedra and M. Fiore, "DAEMON: A Network Intelligence Plane for 6G Networks", in *2022 IEEE Globecom Workshops (GC Wkshps)*, Rio de Janeiro, 2022.

[34] S. Yuan, M. Peng, Y. Sun and X. Liu, "Software defined intelligent satellite-terrestrial integrated networks: Insights and challenges", *Digital Communications and Networks,* vol. 9, no. 6, pp. 1331-1339, 2023.

[35] R. Campana, C. Amatetti and A. Vanelli-Coralli, "O-RAN based Non-Terrestrial Networks: Trends and Challenges", in *2023 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit)*, 2023, pp. 264-269.

[36] M. M. Azari, S. Solanki, S. Chatzinotas, O. Kodheli, H. Sallouha, A. Colpaert, J. F. Mendoza Montoya, S. Pollin, A. Haqiqatnejad, A. Mostaani, E. Lagunas and B. Ottersten, "Evolution of Non-Terrestrial Networks From 5G to 6G: A Survey", *IEEE Communications Surveys & Tutorials,* vol. 24, no. 4, pp. 2633-2672, 2022.

[37] ETSI ISG MEC, "Multi-access Edge Computing (MEC); Framework and Reference Architecture", European Telecommunications Standards Institute, ETSI GS MEC 003 V3.2.1, April 2024. [Online]. Available: https://www.etsi.org/deliver/etsi_gs/MEC/001_099/003/03.02.01_60/gs_mec003v030201p.pdf.

[38] ETSI ISG MEC, "Multi-access Edge Computing (MEC) 5G Integration", European Telecommunications Standards Institute, ETSI GR MEC 031 V2.1.1, October 2020. [Online]. Available: https://www.etsi.org/deliver/etsi_gr/MEC/001_099/031/02.01.01_60/gr_MEC031v020101p.pdf.

[39] ETSI ISG MEC, "Multi-access Edge Computing (MEC); Application Mobility Service API", European Telecommunications Standards Institute, ETSI GS MEC 021 V3.1.1, February 2024. [Online]. Available: https://www.etsi.org/deliver/etsi_gs/MEC/001_099/021/03.01.01_60/gs_MEC021v030101p.pdf.

[40] 5G-COMPLETE, "A unified network, Computational and stOrage resource Management framework targeting end-to-end Performance optimization for secure 5G muLti-tEchnology and multi-Tenancy Environments", [Online]. Available: https://5gcomplete.eu/. [Accessed August 2024].

[41] SaT5G, "Satellite and Terrestrial Network for 5G", [Online]. Available: https://5g-ppp.eu/sat5g/. [Accessed August 2024].

[42] EAGER, “Innovative Technologies & Techniques for SatCom Beyond 5G”, [Online]. Available: https://www.eagerproject.eu/. [Accessed August 2024].

[43] 5GENESIS, “5th Generation End-to-end Network, Experimentation, System Integration, and Showcasing”, [Online]. Available: https://5genesis.eu/. [Accessed August 2024].

[44] L. Han, R. Li, A. Retana, M. Chen, L. Su, T. Jiang and N. Wang, “Problems and Requirements of Satellite Constellation for Internet”, IETF Internet Draft draft-lhan-problems-requirements-satellite-net, 2023. [Online]. Available: https://datatracker.ietf.org/doc/draft-lhan-problems-requirements-satellite-net/.

[45] 3GPP, “Study on using satellite access in 5G”, 3rd Generation Partnership Project, Technical Report TR 22.822, ver. 16.0.0, July 2018. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3372.

[46] 3GPP, “Study on satellite access – Phase 3”, 3rd Generation Partnership Project, Technical Report TR 22.865, ver. 19.2.0, December 2023. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=4089.

[47] ITU-T, “Fixed, mobile and satellite convergence – Requirements for IMT-2020 networks and beyond”, International Telecommunication Union – Telecommunication Standardization Sector, Recommendation Y.3200, February 2022. [Online]. Available: https://www.itu.int/rec/T-REC-Y.3200-202202-I.

[48] ITU-T, “Fixed, mobile and satellite convergence – Framework for IMT-2020 networks and beyond”, International Telecommunication Union – Telecommunication Standardization Sector, Recommendation Y.3201, January 2023. [Online]. Available: https://www.itu.int/rec/T-REC-Y.3201-202301-I.

[49] 3GPP, “Service requirements for the 5G system”, 3rd Generation Partnership Project, Technical Specification TS 22.261, ver. 19.7.0, June 2024. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3107.

[50] H. Dongxu, X. Min, F. Zhou and D. Yuan, “Satellite Network Routing Use Cases”, IETF Internet Draft draft-hou-tvr-satellite-network-usecases, 2023. [Online]. Available: https://datatracker.ietf.org/doc/draft-hou-tvr-satellite-network-usecases/. [Accessed August 2024].

[51] 3GPP, “System architecture for the 5G System (5GS)”, 3rd Generation Partnership Project, Technical Specification TS 23.501, ver. 19.0.0, June 2024. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3144.

[52] 3GPP, “Management and orchestration; Architecture framework”, 3rd Generation Partnership Project, Technical Specification TS 28.533, ver. 18.2.2, July 2024. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3416.

[53] 3GPP, “Common API Framework for 3GPP Northbound APIs”, 3rd Generation Partnership Project, Technical Specification TS 23.222, ver. 19.2.0, July 2024. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3337.

[54] 3GPP, “Service Enabler Architecture Layer for Verticals (SEAL); Functional architecture and information flows”, 3rd Generation Partnership Project, Technical Specification TS 23.434, ver. 19.2.0, June 2024. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3587.

[55] 3GPP, “Study on Integration of satellite components in the 5G architecture; Phase 3”, 3rd Generation Partnership Project, TR 23.700-29, ver. 19.0.0, June 2024. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=4212.

[56] J. Shi, Z. Li, J. Hu, Z. Tie, S. Li, W. Liang and Z. Ding, “OTFS enabled LEO Satellite Communications: A Promising Solution to Severe Doppler Effects”, *IEEE Network,* vol. 38, no. 1, pp. 203-209, January 2024.

[57] 3GPP, “Study on architecture aspects for using satellite access in 5G”, 3rd Generation Partnership Project, Technical Report TR 23.737, ver. 17.2.0, March 2021. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3485.

[58] 3GPP, “NR; NR and NG-RAN Overall description; Stage-2”, 3rd Generation Partnership Project, Technical Specification TS 38.300, ver. 18.2.0, July 2024. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3191.

[59] P.-D. Arapoglou, S. Cioni and E. R. A. Ginesi, “Direct Access to 5G New Radio User Equipment from NGSO Satellites in Millimeter Waves”, in *2020 10th Advanced Satellite Multimedia Systems Conference and the 16th Signal Processing for Space Communications Workshop (ASMS/SPSC)*, Graz, 2020.

[60] D. Tuzi, T. Delamotte and A. Knopp, “Satellite Swarm-Based Antenna Arrays for 6G Direct-to-Cell Connectivity”, *IEEE Access,* vol. 11, pp. 36907-36928, 2023.

[61] K. G. Yalda, D. J. Hamad and N. Ţăpuş, “A survey on Software-defined Wide Area Network (SD-WAN) architectures”, in *2022 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*, Ankara, 2022.

[62] ETSI ISG NFV, “Network Functions Virtualisation (NFV); Management and Orchestration; Report on Management and Orchestration Framework”, European Telecommunications Standards Institute, ETSI GR

NFV-MAN 001 V1.2.1, December 2021. [Online]. Available: https://www.etsi.org/deliver/etsi_gr/NFV-MAN/001_099/001/01.02.01_60/gr_NFV-MAN001v010201p.pdf.

[63] W. Jiang, "Software defined satellite networks: a survey", *Digital Communications and Networks,* vol. 9, no. 6, pp. 1243-1264, 2023.

[64] S. Wu, X. Chen, L. Yang, C. Fan and Y. Zhao, "Dynamic and static controller placement in Software-Defined Satellite Networking", *Acta Astronautica,* vol. 152, pp. 49-58, November 2018.

[65] M. Troglia Gamba, B. D. Polidori, A. Minetto, F. Dovis, E. Banfi and F. Dominic, "GNSS Radio Frequency Interference Monitoring from LEO Satellites: An In-Laboratory Prototype", *Sensors,* vol. 24, no. 2, p. 508, 2024.

[66] IBM, "An architectural blueprint for autonomic computing", IBM Autonomic Computing White Paper, Fourth Edition, June 2006.

[67] S. Kukliński and L. Tomaszewski, "DASMO: A scalable approach to network slices management and orchestration", in *NOMS 2018 – 2018 IEEE/IFIP Network Operations and Management Symposium*, Taipei, 2018.

[68] GSMA, "Generic Network Slice Template", GSM Association, Official Document NG.116, ver. 9.0, April 2023. [Online]. Available: https://www.gsma.com/newsroom/wp-content/uploads//NG.116-v9.0-1.pdf.

[69] O-RAN Alliance, "O-RAN", [Online]. Available: https://www.o-ran.org/. [Accessed August 2024].

[70] O. Adamuz-Hinojosa, P. Munoz, J. Ordonez-Lucena, J. J. Ramos-Munoz and J. M. Lopez-Soler, "Harmonizing 3GPP and NFV Description Models: Providing Customized RAN Slices in 5G Networks", *IEEE Vehicular Technology Magazine,* vol. 14, no. 4, pp. 64-75, 2019.

[71] ETSI ISG NFV, "Network Functions Virtualisation (NFV) Release 3; Management and Orchestration; Report on architecture options to support multiple administrative domains", European Telecommunications Standards Institute, ETSI GR NFV-IFA 028 V3.1.1, January 2018. [Online]. Available: https://www.etsi.org/deliver/etsi_gr/NFV-IFA/001_099/028/03.01.01_60/gr_nfv-ifa028v030101p.pdf.

[72] ETSI ISG NFV, "Network Functions Virtualisation (NFV) Release 3; Management and Orchestration; Interface and Information Model Specification for Multi-Site Connectivity Services", European Telecommunications Standards Institute, ETSI GS NFV-IFA 032 V3.2.1, April 2019. [Online]. Available: https://www.etsi.org/deliver/etsi_gs/NFV-IFA/001_099/032/03.02.01_60/gs_NFV-IFA032v030201p.pdf.

[73] P. Lin, Z. Zhang and L. Liu, "Research on Space-Air-Ground Integrated Network Application", *2024 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB),* pp. 1-6, 2024.

[74] K. Antevski, C. J. Bernardos, L. Cominardi, A. de la Oliva and A. Mourad, "On the integration of NFV and MEC technologies: architecture analysis and benefits for edge robotics", *Comput. Networks,* vol. 175, p. 107274, 2020.

[75] N. Slamnik-Kriještorac and J. M. Marquez-Barja, "Demo Abstract: Assessing MANO Performance based on VIM Platforms within MEC Context", *IEEE INFOCOM 2020 – IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS),* pp. 1338-1339, 2020.

[76] H. Biyang, J. Xiao, X. Hu, G. Zhang and Q. Zhou, "DoDAF and STK Based Modeling and Simulation of MAV/UAV Cooperative Detection Architecture", *2022 IEEE 17th Conference on Industrial Electronics and Applications (ICIEA),* pp. 755-759, 2022.

[77] Agi, "Systems Tool Kit (STK)", [Online]. Available: https://www.agi.com/products/stk.

[78] Z. Zhikai, G. Shushi, Z. Qinyu and X. Jiayin, "Big data application simulation platform design for onboard distributed processing of LEO mega-constellation networks", *China Communications,* vol. 21, pp. 334-345, 2024.

[79] D. Yong, J. Song, W. Dayang, B. Yang, X. Huichen and Z. Jincheng, "A STK-Based Constellation Architecture Implementation for 5G Low-Orbit Satellites", *2022 IEEE 4th International Conference on Power, Intelligent Computing and Systems (ICPICS),* pp. 602-606, 2022.

[80] F. R. Hoots and R. L. Roehrich, "Models for Propagation of NORAD Element Sets. Spacetrack Report #3 U.S. Air Force: Aerospace Defense Command", Air Force: Aerospace Defense Command, 1980.

[81] 3GPP, "Architecture for enabling Edge Applications", 3rd Generation Partnership Project, Technical Specification TS 23.558, ver. 19.2.0, 30 June 2024. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3723.

[82] ITU-R, "IMT Vision – Framework and overall objectives of the future development of IMT for 2020 and beyond", International Telecommunication Union – Radiocommunication Sector, Recommendation M.2083, September 2015. [Online]. Available: https://www.itu.int/rec/R-REC-M.2083.

[83] 3GPP, "5G System Enhancements for Edge Computing; Stage 2", 3rd Generation Partnership Project, Technical Standard TS 23.548, ver. 18.6.0, 26 June 2024. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3856.

[84] 3GPP, “Interface between the Control Plane and the User Plane nodes”, 3rd Generation Partnership Project, Technical Standard TS 29.244, ver. 18.6.0, June 2024. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3111.

[85] M. Guadalupi, J. Ferrer, R. A. Ferrús Ferré, I. Llorens Aymerich, A. González Muiño, A. Mañero Contreras, R. Brandborg Sørensen, H. Krogh Moeller, A. M. Calveras Augé, T. Kellermann and M. S. Afaqui, "Designing a 3GPP NB-IoT NTN service for CubeSats in low density constellations”, in *Proceedings of the 72nd International Astronautical Congress (IAC)*, Dubai, UAE, 2021.

[86] 3GPP, “Discussion on the service link discontinuity and affected procedures for NB-IoT NTN”, 3rd Generation Partnership Project, R2-2101248, 5 February 2021. [Online]. Available: https://www.3gpp.org/ftp/TSG_RAN/WG2_RL2/TSGR2_113-e/Docs/R2-2101248.zip.

[87] GSMA, “Mobile IoT in the 5G future: NB-IoT and LTE-M in the context of 5G”, April 2018. [Online]. Available: https://www.gsma.com/solutions-and-impact/technologies/internet-of-things/wp-content/uploads/2018/05/GSMA-5G-Mobile-IoT.pdf. [Accessed August 2024].

[88] Ericsson, “Cellular IoT in the 5G era”, Ericsson White Paper, February 2020. [Online]. Available: https://wcm.ericsson.net/48ff1f/assets/local/reports-papers/white-papers/Cellular_IoT_in_5G_whitepaper_AW.pdf. [Accessed August 2024].

[89] T. Kellermann, R. Pueyo Centelles, D. Camps-Mur, R. Ferrús, M. Guadalupi and A. C. Augé, “Novel Architecture for Cellular IoT in Future Non-Terrestrial Networks: Store and Forward Adaptations for Enabling Discontinuous Feeder Link Operation”, *IEEE Access,* vol. 10, pp. 68922-68936, 2022.

[90] F. Rinaldi, H.-L. Maattanen, J. Torsner, S. Pizzi, S. Andreev, A. Iera and Y. Koucheryavy, "Non-Terrestrial Networks in 5G & Beyond: A Survey”, *IEEE Access,* vol. 8, pp. 165178-165200, 2020.

[91] 3GPP, “ASN.1 proposal for satellite assistance information for prediction of discontinuous coverage”, 3rd Generation Partnership Project, May 2022. [Online]. Available: https://www.3gpp.org/ftp/tsg_ran/WG2_RL2/TSGR2_118-e/Docs/R2-2206115.zip.

[92] 3GPP, “On satellite pass predictions for UE wake-up management under discontinuous coverage”, 3rd Generation Partnership Project, May 2021. [Online]. Available: https://www.3gpp.org/ftp/tsg_ran/WG2_RL2/TSGR2_114-e/Docs/R2-2104863.zip.

[93] 3GPP, “On satellite ephemeris information types for discontinuous coverage in IoT-NTN”, 3rd Generation Partnership Project, January 2021. [Online]. Available: https://www.3gpp.org/ftp/tsg_ran/wg2_rl2/TSGR2_116bis-e/Docs/R2-2201017.zip. [Accessed 24 August 2023].

[94] 3GPP, “Solution for paging enhancement under discontinuous coverage based on definition and management of UE-specific Dynamic Tracking Areas”, 3rd Generation Partnership Project, August 2022. [Online]. Available: https://www.3gpp.org/ftp/tsg_sa/WG2_Arch/TSGS2_152E_Electronic_2022-08/INBOX/S2-2207184.zip. [Accessed 24 August 2023].

[95] 3GPP, “Complementing discontinuous coverage with minimum support for discontinuous feeder link operation”, 3rd Generation Partnership Project, March 2023. [Online]. Available: https://www.3gpp.org/ftp/tsg_ran/WG2_RL2/TSGR2_121/Docs/R2-2301886.zip.

[96] 3GPP, “New SID on satellite access Phase 3”, 3rd Generation Partnership Project, May 2023. [Online]. Available: https://www.3gpp.org/ftp/tsg_sa/WG1_Serv/TSGS1_98e_EM_May2022/Docs/S1-221167.zip.

[97] 3GPP, “Description of store and forward operation”, 3rd Generation Partnership Project, August 2022. [Online]. Available: https://www.3gpp.org/ftp/tsg_sa/WG1_Serv/TSGS1_99e_EM_Aug2022/Docs/S1-222336.zip.

[98] 3GPP, “Use case store and forward – MO”, 3rd Generation Partnership Project, August 2023. [Online]. Available: https://www.3gpp.org/ftp/tsg_sa/WG1_Serv/TSGS1_99e_EM_Aug2022/Docs/S1-222334.zip.

[99] 3GPP, “Use case store and forward – MT”, 3rd Generation Partnership Project, August 2023. [Online]. Available: https://www.3gpp.org/ftp/tsg_sa/WG1_Serv/TSGS1_99e_EM_Aug2022/Docs/S1-222335.zip.

[100] 3GPP, “Consideration on RAN1/2/3 led NTN topics for Release 19 (Annex)”, 3rd Generation Partnership Project, June 2023. [Online]. Available: https://www.3gpp.org/ftp/TSG_RAN/TSG_RAN/TSGR_AHs/2023_06_RAN_Rel19_WS/Docs/RWS-230048.zip.

[101] 3GPP, “Satellite Access: SA Rel-19 topics & priorities”, 3rd Generation Partnership Project, June 2023. [Online]. Available: https://www.3gpp.org/ftp/tsg_sa/TSG_SA/Workshops/2023-06-13_Rel-19_WorkShop/Docs/SWS-230063.zip.

[102] Open Networking Foundation, “OpenFlow Switch Specification”, ONF TS-025, ver. 1.5.1, March 2015. [Online]. Available: https://opennetworking.org/wp-content/uploads/2014/10/openflow-switch-v1.5.1.pdf.

[103] ONF, “OF-CONFIG 1.2: OpenFlow Management and Configuration Protocol; ONF TS-016”, July 2017. [Online]. Available: https://opennetworking.org/wp-content/uploads/2017/07/of-config-1.2.pdf. [Accessed 3 July 2024].

[104] B. Pfaff and B. Davie, "The Open vSwitch Database Management Protocol", December 2013. [Online]. Available: https://datatracker.ietf.org/doc/html/rfc7047. [Accessed August 2024].

[105] 3GPP, "NG-RAN; Xn Application Protocol (XnAP)", 3rd Generation Partnership Project, Technical Standard TS 38.423, ver. 18.2.0, July 2024. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3228.

[106] 3GPP, "NG-RAN; NG Application Protocol (NGAP)", 3rd Generation Partnership Project, Technical Standard TS 38.413, ver. 18.2.0, 3 July 2024. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3223.

[107] 5G PPP Architecture Working Group, "The 6G Architecture Landscape – European perspective", Feb 2023. [Online]. Available: https://doi.org/10.5281/zenodo.7313232. [Accessed Jun 2024].

[108] 3GPP, "Procedures for the 5G System (5GS)", 3rd Generation Partnership Project, Technical Specification TS 23.502, ver. 19.0.0, June 2024. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3145.

[109] D. Koshnichrova, A. Mihovska, P. Koleva and V. Poulkov, "Context-Aware Communication Networks With Users in the Loop-A Short Survey", in *2022 IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom)*, Sofia, Bulgaria, 2022.

[110] J. Chuang, "AI In Telcom & RAN – Standard Evolution", 17 March 2024. [Online]. Available: https://www.linkedin.com/pulse/ai-telcom-ran-standard-evolution-jessica-chuang-0pi4c/. [Accessed June 2024].

[111] "DAEMON – Network intelligence for aDAptive and sElf-Learning MObile Networks", [Online]. Available: https://github.com/h2020daemon/main. [Accessed June 2024].

[112] M. K. Bahare, A. Gavras, M. Gramaglia, J. Cosmas, X. Li, Ö. Bulakci, A. Rahman, A. Kostopoulos, A. Mesodiakaki, D. Tsolkas, M. Ericson, M. Boldi, M. Uusitalo, M. Ghoraishi and P. Rugeland, "The 6G Architecture Landscape – European perspective", 5G PPP, February 2023. [Online]. Available: https://doi.org/10.5281/zenodo.7313231. [Accessed June 2024].

[113] M. Camelo, L. Cominardi, M. Gramaglia, M. Fiore, A. Garcia-Saavedra, L. Fuentes, D. De Vleeschauwer, P. Soto-Arenas, N. Slamnik-Krijestorac, J. Ballesteros, C.-Y. Chang, G. Baldoni, P. Hellinckx, J. M. Marquez-Barja, P. Hellinckx and S. Latré, "Requirements and Specifications for the Orchestration of Network Intelligence in 6G", in *IEEE 19th Annual Consumer Communications & Networking Conference (CCNC)*, Las Vegas, 2022.

[114] L. E. Chatzieleftheriou, M. Gramaglia, M. Fiore, N. Slamnik-Kriještorac, M. Camelo, P. Soto, E. Kosmatos, A. Garcia-Saavedra and M. Gucciardo, "Network Intelligence in Action: the DAEMON Perspective", in *European Conference on Networks and Communications & 6G Summit*, Antwerp, 2024.

[115] ETSI ISG NFV, "Network Functions Virtualisation (NFV); Acceleration Technologies; Acceleration Resource Management Interface Specification; Release 3", European Telecommunications Standards Institute, ETSI GS NFV-IFA 019 V3.1.1, July 2017. [Online]. Available: https://www.etsi.org/deliver/etsi_gs/nfv-ifa/001_099/019/03.01.01_60/gs_nfv-ifa019v030101p.pdf.

[116] O-RAN Alliance, "O-RAN Work Group 1 (Use Cases and Overall Architecture); O-RAN Architecture Description", O-RAN Specifications; Technical Specification, O-RAN.WG1.OAD-R003-v12.00. [Online]. Available: https://orandownloadsweb.azurewebsites.net/download?id=641. [Accessed August 2024].

[117] Open Geospatial Consortium Inc., "OGC Web Services Common Standard", Open Geospatial Consortium Inc., OGC 06-121r9 V2.0, 07 April 2010. [Online]. Available: https://portal.ogc.org/files/?artifact_id=38867.

[118] ETSI ISG NFV, "Network Functions Virtualisation (NFV) Release 3; Management and Orchestration; Os-Ma-nfvo reference point – Interface and Information Model Specification", European Telecommunications Standards Insitute, ETSI GS NFV-IFA 013 V3.7.1, November 2022. [Online]. Available: https://www.etsi.org/deliver/etsi_gs/NFV-IFA/001_099/013/03.07.01_60/gs_NFV-IFA013v030701p.pdf.

[119] ONF, "Transport API (TAPI) 2.0; Overview", Open Networking Foundation; Draft V0.0, August 2017. [Online]. Available: https://opennetworking.org/wp-content/uploads/2017/08/TAPI-2-WP_DRAFT.pdf.

[120] ETSI ISG NFV, "Network Functions Virtualisation (NFV) Release 5; Management and Orchestration; Or-Vnfm reference point – Interface and Information Model Specification", European Telecommunications Standards Institute, ETSI GS NFV-IFA 007 V5.1.1, June 2024. [Online]. Available: https://www.etsi.org/deliver/etsi_gs/NFV-IFA/001_099/007/05.01.01_60/gs_NFV-IFA007v050101p.pdf.

[121] ETSI ISG NFV, "Network Functions Virtualisation (NFV) Release 5; Management and Orchestration; Vi-Vnfm reference point – Interface and Information Model Specification", European Telecommunications Standards Institute, ETSI GS NFV-IFA 006 V5.1.1, June 2024. [Online]. Available: https://www.etsi.org/deliver/etsi_gs/NFV-IFA/001_099/006/05.01.01_60/gs_NFV-IFA006v050101p.pdf.

[122] ETHER, "Deliverable D2.3 Technoeconomic analysis and business plan", 2024.

[123] NVIDIA, "NVIDIA Jetson Xavier", [Online]. Available: https://www.nvidia.com/en-us/autonomous-machines/embedded-systems/jetson-xavier-series/. [Accessed August 2024].

[124] NVIDIA, "NVIDIA Jetson Orin", [Online]. Available: https://www.nvidia.com/en-us/autonomous-machines/embedded-systems/jetson-orin/. [Accessed August 2024].

[125] Thinkmate, "RAX QT24-12E9", [Online]. Available: https://www.thinkmate.com/system/rax-qt24-12e9. [Accessed August 2024].

[126] Eurostat, "Electricity price statistics", [Online]. Available: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Electricity_price_statistics#Electricity_prices_for_non-household_consumers. [Accessed August 2024].

APPENDIX 1 – SATELLITE TRADE-OFFS PROVIDING DIFFERENT SERVICES

As introduced in Section 4.3, the different orbits have different characteristics and therefore are often deployed for different UCs. An overview of this is given below

	LEO	MEO	GEO	HEO
Orbital altitude	500-2000 km (note 1)	2000-35786 km	~35786 km	Varies, e.g., from around 400 km (perigee) to 40000 km (apogee)
Round trip propagation delay (note 2)	20-40 ms	~110 ms	480-500 ms	Similar to GEO in operational part of orbit
Typical data response time	20-40 ms	150-180 ms	600-650 ms	As above
Constellation size for near global coverage	Hundreds or thousands	10 to 20 (note 7)	3 to 6 (note 8)	Not available, typically 3 satellites per orbit for 24×7 service
Typical beam size	Small	Medium	Medium/small to large	Medium to large
Regional coverage implications	Global or near global visibility (note 3)	Can be global or restricted to nearer the equator	Regional coverage often implemented	By definition, service covering one part of a polar region
Spacecraft size	Small to large	Medium	Medium to very large	Large
Launch costs (note 4)	$3M per 1000 kg	N/a	$8M per 1000 kg in geostationary transfer orbit (GTO), cf. note 5	Similar to GEO
Typical spacecraft lifetime	5 to 7 years	10 years plus	15 years plus	Unknown
End-of-life	De-orbited to burn up in the atmosphere	Re-orbited to a graveyard orbit	Re-orbited to the GEO graveyard orbit	Unknown
Gateway implications	Large number of gateways needed, depends on geographic coverage needed and lower if ISLs used (note 9)	Uses regional gateways	Uses small number of regional gateways	Unknown, probably uses small number of regional gateways
User equipment	Varies from low cost low bit rate systems to high performance systems. The latter require 100 W or more to operate	Use motorised reflectors or electrically steerable antennas, need to be able to access two satellites at same time for continuous service	Static and communications-on-the-pause (COTP) sites can use passive reflectors reducing power consumption (<35 W for consumer <70 W for professional) Communications-on-the-move	Use motorised reflectors or electrically steerable antennas, need to be able to access two satellites at same time for continuous service

	LEO	MEO	GEO	HEO
			(COTM) use motorised or electrically steerable antennas	
Installation	Needs good visibility of much of the sky	Needs good visibility of a fair part of the sky	Needs good visibility in a specific direction (possibly fairly low elevation)	Needs good visibility of some of the sky
Typical uses to date – telecoms	Low bit rate data Internet access	High speed point to point links (remote island, cruise liners etc)	Data networks Internet access Resilient connections News gathering Military	None in orbit at time of writing
Companies offering satcom services (examples)	Globalstar Iridium Starlink OneWeb (Eutelsat) AST SpaceMobile	O3B (SES)	Intelsat SES Eutelsat Avanti Viasat Yahsat Echostar	Inmarsat planning launch
Typical uses to date – non-telecoms	Earth observation (EO), Experiments	GNSS, Weather observations	TV broadcast, Weather observations	Satellite radio
Illustrative potential NTN / 5G and beyond UCs (note 6)	Direct to smartphone (limited speeds) and other UE IoT services Sensing services Positioning, navigation, and timing services Backhaul and relay service to moving "platforms"	High-speed backhaul to remote and moving BSs, for both primary and resilient connections	Backhaul to remote locations (especially power limited sites) and moving BSs, Resilient connections, Multicast broadcast service distribution	Similar to GEO, providing service in polar regions

Notes:

1. There is a variant called vLEO (very low Earth orbit), which requires specialist designs to work below this level probably requiring special materials, propulsion, and aerodynamic design to cope with the drag caused by the faint atmosphere below about 500 km.
2. Defined here as UE to a single satellite (no intersatellite link) to gateway and return.
3. User terminal and gateway both need visibility of the satellite to establish E2E communications unless the system has ISLs, this means that gateways need to be located around the globe to use the system everywhere.
4. Based on Falcon 9 list price in 2022.
5. GTO is a HEO with a low perigee and an apogee around GEO, in this case the launch vehicle launches the satellite into a GTO; then the satellite circularises its orbit at geostationary altitude.
6. These are illustrative to show core strengths of each orbit.
7. May not include polar coverage to maximise capacity towards demand.
8. This excludes polar regions (roughly more than +/- 75° latitude) as the satellite is too low an elevation or indeed below the horizon.
9. The use of intersatellite links requires the use of regenerative satellites.

Co-funded by the European Union