Received 10 May 2023; revised 19 June 2023; accepted 8 July 2023. Date of publication 11 July 2023; date of current version 2 August 2023.

Digital Object Identifier 10.1109/OJCOMS.2023.3294445

Toward an Open, Intelligent, and End-to-End Architectural Framework for Network Slicing in 6G Communication Systems

MOHAMMAD ASIF HABIBI [1], BIN HAN [1] (Senior Member, IEEE), AMINA FELLAN [1], WEI JIANG [2] (Senior Member, IEEE), ADRIÁN GALLEGO SÁNCHEZ [3], IGNACIO LABRADOR PAVÓN [3], AMINA BOUBENDIR [4], AND HANS D. SCHOTTEN [1,2] (Member, IEEE)

[1] Division of Wireless Communications and Radio Navigation, Department of Electrical and Computer Engineering,
Rheinland-Pfälzische Technische Universität, 67663 Kaiserslautern, Germany

[2] Intelligent Networking Research Group, German Research Center for Artificial Intelligence, 67663 Kaiserslautern, Germany

[3] Research and Innovation Department, ATOS Spain, 28037 Madrid, Spain

[4] Network Architecture and Automation Department, Orange Labs Networks, 92130 Issy-les-Moulineaux, France

CORRESPONDING AUTHOR: M. A. HABIBI (e-mail: asif@eit.uni-kl.de)

This work was supported in part by the Horizon 2020 Research and Innovation Program of the European Union through Hexa-X under Grant 101015956, and in part by Hexa-X-II under Grant 101095759.

ABSTRACT The definition of the fundamental concepts and the design of the architectural framework for network slicing in fifth-generation communication systems have been successfully concluded; the standardization activities are almost over; and the commercial deployment has already commenced worldwide. To compete for digital supremacy and to be seen as front-runners in the international technological race, researchers from various regions and countries have begun exploring the technical requirements, envisioning potential applications, identifying innovative enablers, developing testbeds for the preliminary validation of several terrestrial and non-terrestrial technologies, and conceptualizing the architectural design for the next generation of mobile communication systems – the sixth-generation (6G) – aiming to connect the human, physical, and digital worlds with a high level of intelligence and openness for the 2030s. In support of such an ambitious vision, this article extends the end-to-end network slicing concepts, methods, solutions, and functioning architectures towards 6G. To this intent, the study first presents several decisive motivating trends behind such an extension of network slicing in order to make forthcoming mobile networks fully slicing-aware. Following that, the paper attempts to highlight the intelligentization of a number of key enabling technologies that will bring a renaissance to network slicing in the next decade. It then proposes a unified architectural framework and its principal building blocks in several layers, paving the way for the implementation of an open and intelligent network and network slicing in 6G. The proposed architectural solution harmonizes the most recent specifications of the relevant de jure and de facto standards development organizations in their applicable layers with the aim of architecting a pre-standard-compliant and preliminary framework for slicing the 6G network. Finally, the article is intended to spur interest and lay the groundwork for further investigations and subsequent research and development by highlighting a number of open research challenges and directions in this flourishing field.

INDEX TERMS 6G, automation, cloudification, end-to-end architecture, intelligentization, intelligent networks, intelligent network slice, management and orchestration, mobile communication systems, network architecture, network exposure, network slicing, network slice, openness, open network, open network slice, softwarization, standardization, standards, virtualization.

ACRONYMS

3GPP	Third Generation Partnership Project
3GPP-NSMS	3GPP-network slicing management system
4G	fourth generation
5G	fifth-generation
6G	sixth-generation

AI	artificial intelligence
API	application programming interface
BBF	Broadband Forum
C-MDAF	centralized-MDAF
CICD	continuous integration and continuous delivery
CISM	container infrastructure service management
CN	core network
CNF	cloud-native network function
CSMF	communication service management function
CT WG4	Core Network and Terminals WG 4
CU	centralized unit
DevOps	development and operations
DL	deep learning
DU	distributed unit
E2E	end-to-end
ELA	exposure level agreement
eMBB	enhanced mobile broadband
ENI	experiential networked intelligence
ETSI	European Telecommunications Standards Institute
FB	functional block
FCAPS	fault, configuration, accounting, performance, and security
FL	federated learning
gNB	next generation nodeB
GSMA	GSM Association
GST	Generic NS Template
HMTC	high-performance machine-type communications
IETF	Internet Engineering Task Force
IOWN	Innovative Optical and Wireless Network
ISG	industry specification group
ITU	International Telecommunication Union
KPI	key performance indicator
KVI	key value indicator
M&O	management and orchestration
MDAF	management data and analytics function
MDAS	management data analytics service
MEC	multi-access edge computing
MEF	Metro Ethernet Forum
MF	management function
ML	machine learning
MLA	management level agreement
mMTC	massive machine type communications
mULC	massive ultra-reliable low-latency communication
MVNO	mobile virtual network operator
NC	network controller
Near-RT RIC	near-real-time radio intelligent controller
NEF	network exposure function
NF	network function
NFD	network function descriptor
NFMF	network function management function
NFP	network forwarding path
NFV	network function virtualization
NFV-MANO	NFV-management and orchestration
NFVI	NFV infrastructure
NFVO	NFV orchestrator
NGMN	Next Generation Mobile Networks
NOC	network operations center
Non-RT RIC	non-real-time radio intelligent controller
NS	network slice
NSMF	NS management function
NSS	NS subnet
NSSAI	NS selection assistance information
NSSMF	NS subnet management function
NSST	NS subnet template
NST	NS template
NWDAF	network data and analytics function
O-RAN	Open-RAN
ONF	Open Networking Foundation
PLD	physical link descriptor
PM	physical machine
PNF	physical network function
PNFD	PNF descriptor
PoP	point of presence
QoE	quality of experience
QoS	quality of service
RAN	radio access network
RL	reinforcement learning
RU	radio unit
S-NSSAI	single NS selection assistance information
SCEF	service capability exposure function
SCF	Small Cell Forum
SD	slice differentiator
SDN	software-defined networking
SDO	standards developing organization
SDR	software-defined radio
SDS	software-defined storage
SDx	software-defined everything
SFC	service function chain
SLA	service level agreement
SMO	service management and orchestration
SST	slice/service type
T-NSSMF	transport-NSSMF
TIP	Telecom Infra Project
TMF	TeleManagement Forum
TN	transport network
TSG	Technical Specification Group
TSG SA2	TSG Service and System Aspects WG 2
TSG SA5	TSG Service and System Aspects WG 5
UE	user equipment
uLBC	ultra-reliable low-latency broadband communication
uMBB	ubiquitous mobile broadband
URLLC	ultra-reliable low latency communications
V2X	vehicle-to-everything
VIM	virtualized infrastructure manager

VL	virtual link
VLD	virtual link descriptor
VM	virtual machine
VNF	virtual network function
VNFD	VNF descriptor
VNFFG	VNF forwarding graph
VNFM	VNF manager
WG	working group
WIM	wide area networks infrastructure manager
ZSM	zero touch network and service management

I. INTRODUCTORY REMARKS

NETWORK slicing emerged as a revolutionary architecture solution aimed at logically partitioning the underlying infrastructure into customized and mutually isolated network slices (NSs) in order to enable the provisioning of heterogeneous services, the serving of myriad broadband consumers, and the hosting of innumerable vertical industries [1]. This evolution will continue with the adoption of a greater variety of diverging applications in the sixth-generation (6G), which will require an even wider range of requirements and competing criteria [2]. Such a transformation necessitates (a) innovative use-case-driven, context-aware, and situation-aware solutions; (b) the highest levels of automation, network programmability, disaggregation, openness, and interoperability; (c) data-driven and intent-based service management and orchestration (SMO) mechanisms and frameworks; and (d) predictive and proactive slicing approaches. Hence, a significant rethinking of the legacy slicing framework is required to optimize and adapt it to the heterogeneity of 6G services and use cases, the complexity of 6G applications, and the continuous technological trends and research efforts in this direction.

To make the above-mentioned challenging aspects of slicing in 6G a reality and enhance the prominent features and capabilities of existing and emerging 6G NSs, there is a clear and strong consensus among standards developing organizations (SDOs) [3], industry [4], and academia [5] that artificial intelligence (AI) can play a critical role in providing end-to-end (E2E) automation, full programmability, intelligent orchestration, zero-touch service management, and accurate self-optimization to various domains, layers, and network functions (NFs) of the slicing framework [6]. Specifically, advanced machine learning (ML) techniques – such as reinforcement learning (RL), federated learning (FL), and deep learning (DL) – can be considered the most powerful tools due to their ability to solve complex optimization problems and improve data privacy and security [2]. The primary objective of utilizing ML algorithms is to enable the slicing framework to deal with the enormous amount of data that the 6G ecosystem will generate and, therefore, to support making well-informed decisions. These algorithms collect quantitative and qualitative statistics from the core network (CN) down to the extreme edge of the 6G network, with the goal of enabling the slicing framework with descriptive, predictive, diagnostic, and prescriptive capabilities. Therefore, such cutting-edge intelligent tools and automation solutions provide network operators and industry owners with numerous advantages, including reduced deployment and operation costs, lower power consumption, smooth service management and resource orchestration, fewer seasonal on-site maintenance visits, faster network and service repair times, and autonomous optimization and configuration of NFs [2], [7].

Besides the E2E intelligentization of 6G networks, there has been a growing emphasis on defining open interfaces between the different disaggregated components, domains, and layers, allowing operators and service providers to deploy equipment from multiple vendors and stakeholders on a single infrastructure [8]. These disaggregated software and hardware components from various suppliers are expected to transparently and efficiently interoperate via open interfaces, enabling multiple 6G NSs to be provisioned to different tenants over the same network [7]. Openness, disaggregation, and interoperability in the 6G slicing framework are believed to: (a) encourage market competition and innovation; (b) accelerate update and upgrade cycles; (c) simplify the design and deployment of software and hardware components; (d) optimize operations and maintenance by means of a centralized abstraction layer, automation, and data-driven closed-loop control; and (e) facilitate open user, control, synchronization, and management planes [9]. These innovative capabilities will significantly drive down the total cost of ownership, achieve high network performance, boost the flexibility and scalability of 6G NSs, increase the reconfigurability and resiliency of the slicing framework, enable operators to select best-of-breed solutions, and enhance time-to-market for new features and services [8], [10]. Finally, open standardized interfaces, disaggregated components, and interoperability provide a path for small market players and businesses to enter the wireless industry, thereby enabling operators and service providers with a greater selection of suppliers and price points [10], [11].

A. REVIEW OF LITERATURE ON 6G NETWORK SLICING ARCHITECTURE

Despite the preceding advantages, defining open interfaces and utilizing ML algorithms in fifth-generation (5G) slicing ecosystem requires a wide range of optimizations, including a substantial reconstruction of its architectural framework, in order for such an ecosystem to fulfill the heterogeneous requirements of 6G services and applications. The existing slicing frameworks proposed by the Next Generation Mobile Networks (NGMN) Alliance in 2016 [12] and by us in 2017 [13] are believed to be inefficient and incapable of meeting the business, technical, standards, societal, and regulatory requirements of 6G in the 2030s [14]. Because these three-layer constructed frameworks lack, among other essential aspects, the adoption of open interfaces, the integration of intelligence, and the incorporation of automation.

TABLE 1. **The comparison of our proposed architectural framework for open, intelligent, and E2E network slicing in 6G to the most recent state-of-the-art architectural proposals.**

Reference	Comparison parameters					
	Open	**Intelligent**	**Std.-compliant**	**Slicing-aware**	**Year**	**Core objectives**
[7]	✗	✓	✗	✗	2019	Photonics-based cognitive 6G radio architecture
[17]	✗	✓	✗	✗	2019	AI-assisted architecture for green 6G network
[18]	✗	✓	✗	✗	2019	AI-empowered 6G wireless network architecture
[4]	✓	✓	✗	✓	2020	Beyond 5G and 6G mobile network architecture
[6]	✗	✓	✗	✗	2020	AI-enabled intelligent 6G network architecture
[19]	✓	✓	✗	✗	2020	Joint comm., comp., and cach. 6G architecture
[20]	✗	✓	✗	✗	2020	AI-enabled computing-power framework for 6G
[21]	✗	✓	✗	✗	2020	A self-sustained 6G RAN slicing framework
[22]	✗	✓	✗	✓	2021	AI-based slicing architecture for 6G network
[23]	✗	✓	✗	✗	2021	Generic inter-terrestrial 6G network architecture
[24]	✗	✓	✗	✗	2021	Joint comm., sens., and comp. 6G framework
[25]	✗	✓	✗	✗	2021	Native AI E2E 6G network architecture
[26]	✗	✓	✗	✗	2021	A cyber digital twin-based network architecture
[27]	✗	✓	✗	✗	2021	6G-enabled Network-in-a-Box framework
[28]	✗	✓	✗	✗	2021	AI-native and virtualized 6G RAN architecture
[29]	✗	✓	✗	✗	2021	Edge intelligence-empowered driving framework
[30]	✗	✓	✗	✗	2021	3C system architectural framework for 6G
[31]	✗	✓	✗	✗	2021	AI-empowered 6G network architecture
[32]	✗	✓	✗	✗	2021	Intelligent vehicular architecture for 6G
[33]	✗	✓	✗	✗	2021	AI-enabled conceptual architecture for 6G
[34]	✗	✓	✗	✓	2021	Intelligent architecture for 6G NS instances
[35]	✗	✓	✗	✓	2022	Intelligent 6G NS management and orchestration (M&O)
[36]	✗	✓	✗	✓	2022	Authentication framework for 6G NS instances
[37]	✗	✓	✗	✗	2022	Organic-like architecture for 6G CN
[38]	✗	✓	✗	✗	2022	Digital twin-empowered 6G architecture
[39]	✗	✓	✗	✓	2022	AI-native framework for 6G network slicing
[40]	✗	✓	✗	✗	2022	6G network architecture with pervasive AI
[41]	✗	✓	✗	✗	2022	Digital twin and AI-enabled 6G architecture
[14]	✗	✓	✗	✗	2022	E2E 6G architecture built on three layers
[16]	✓	✓	✗	✓	2022	SMO in 6G
[42]	✓	✓	✗	✓	2022	A service of services vision towards 6G
Our proposal	✓	✓	✓	✓	2023	Open and intelligent framework for 6G slicing

Additionally, they are incapable of considering 6G key performance indicators (KPIs) and key value indicators (KVIs), which can help ensure the qualitative and quantitative analysis of the 6G use cases as well as their expected performance and social impact. Therefore, the existing slicing frameworks necessitate significant architectural optimization in various layers and domains for efficiently slicing beyond 5G networks.

In the context of 6G, there are several studies that propose E2E or sub-network architectures. We provide a comparative analysis of these studies and a summary of their core contributions in Table 1. Despite the fact that these papers have introduced intelligence to the 6G and reported its implications on the 6G network, E2E openness and disaggregation of software and hardware components are either absent or not fully considered within their architectures. Notably, the majority of these studies neglected the slicing aspects, and the frameworks they proposed are neither fully nor partially slicing-aware.

In addition to some 6G architectures summarized in Table 1, which are proposed by individual authors, there are a number of projects devoted to the 6G research, such as Hexa-X, Hexa-X-II, 6G Flagship, etc. An up-to-date list of the ongoing 6G projects is detailed in [15]. We also provide a comparative analysis of some of these 6G architectures in Table 1. To our knowledge, Hexa-X is the only 6G project that focuses on researching network services and slices aspects and has also proposed a framework for 6G SMO [16]. However, the Hexa-X SMO framework is neutral and not aligned with any particular SDOs. Nevertheless, several potential alignments with a number of pertinent specifications have been provided in order to illustrate how the Hexa-X SMO framework could be implemented in accordance with certain relevant standards.

For clarity, we compare the above 6G architectures, proposed by research projects and individual authors, with respect to six aspects that can be clearly observed in the relevant columns of Table 1. These six columns are: (a) Open, which indicates whether the architecture is adopting openness; (b) Intelligent, which specifies whether the framework is empowered by ML algorithms; (c) Standard-compliant, which clarifies whether the architecture is in compliance with standards; (d) Slicing-aware, which points out whether the architecture assumes the slicing aspects for provisioning various NSs; (e) Year, which highlights the year of the publication of the relevant reference; and (f) Core Objectives, which provides a brief summary of the major contributions of the relevant publication. We also provide a distinction between our 6G architecture (which will be detailed in later sections) and the state-of-the-art 6G frameworks in Table 1 with regard to these parameters. The comparison of our framework to state-of-the-art solutions reveals that it is open, intelligent, E2E, standards-oriented, and slicing-aware, thereby demonstrating its potential to meet the diverse technical, societal, and business requirements of networks and use cases in the 6G era.

B. THE PRIMARY RESEARCH CHALLENGE

Notwithstanding the fact that the above studies and projects are the first attempts to define 6G functional architectures, which can also be considered benchmarks for our architecture, the corresponding authors and institutions of most of these deliverables (a) proposed high-level frameworks for integrating and harmonizing network components, layers, and domains without considering the latest contributions from SDOs; (b) disregarded the systematic integration of ML algorithms into network via standardized interfaces; (c) paid no or less attention to openness, disaggregation, and interoperability; and (d) fully or partially neglected the E2E slicing aspects of the underlying resources. In addition, there is a considerable need for a radical transformation (in terms of intelligentization and automation) of the mechanisms and procedures used to manage NSs and orchestrate their required resources, which have received less attention from the studies listed in Table 1 and beyond. Specifically, the increase in the overall complexity of network management caused by programmability and softwarization, the unprecedented operational agility required to enable emerging use cases and applications, and the extreme range of heterogeneous requirements (such as infinite bandwidth, ultra-high reliability, imperceptible latency, etc.) for provisioning various types of NSs, call for a significant transfiguration in the operations of the legacy slicing ecosystem. Therefore, an E2E framework designed for an open and intelligent slicing paradigm towards 6G, which is explicitly aligned with the most recent de jure and de facto specifications and requirements of the relevant SDOs in a fully autonomous manner, has not yet been defined.

C. OUR GOALS AND CONTRIBUTIONS

To contribute to filling this research gap, we propose an architectural solution by amalgamating the most recent frameworks, requirements, methods, and concepts from a number of SDOs that have been contributing to network slicing for several years in order to design an entirely new architecture for slicing in 6G. These SDOs are listed in alphabetic order as the Third Generation Partnership Project (3GPP), the Broadband Forum (BBF), the European Telecommunications Standards Institute (ETSI), the GSM Association (GSMA), the Internet Engineering Task Force (IETF), the Innovative Optical and Wireless Network (IOWN) Global Forum, the International Telecommunication Union (ITU), the Metro Ethernet Forum (MEF), the NGMN, the Open Networking Foundation (ONF), the Open-RAN (O-RAN) Alliance, the Small Cell Forum (SCF), the TeleManagement Forum (TMF), and the Telecom Infra Project (TIP). The proposed framework is amalgamated into several layers interconnected via standardized interfaces, with each layer performing a finite range of responsibilities by consuming certain standardized components, concepts, and solutions from particular SDOs. The objectives of such an amalgamation of the latest standards in a layer-based architecture are to: (a) eliminate vendor lock-in and facilitate multi-vendor and multi-stakeholder interoperability for slicing in 6G; (b) inject intelligent into the management, orchestration, operation, and maintenance of various open 6G NSs; and (c) integrate data-driven policies and rules that will enable the slicing framework with complete E2E automation capabilities, such as self-configuration, self-monitoring, self-scaling, self-healing, and self-optimization, for provisioning 6G NS instances. We believe that it is essential to design such an open, autonomous, and E2E architecture based on the most recent standards and that it should have a strong alignment with the state-of-the-art specifications of the relevant SDOs in order to outline a pre-standardization 6G slicing framework.

D. JUSTIFICATIONS FOR ALIGNING THE PROPOSED ARCHITECTURAL FRAMEWORK WITH THE RELEVANT SDOs

Contrary to a number of state-of-the-art 6G architectures (see Table 1), our framework is strongly aligned with the relevant specifications and recommendations of the corresponding SDOs. Unlike [16], which considers such an alignment "risky" due to the high diversity of SDOs and the rapid development of new specifications, we have a different position and we believe that this alignment will provide several benefits, including:

- *Interoperability:* Adhering to the most recent standards ensures that our framework is interoperable with other legacy and futuristic systems that also adhere to these standards, which is essential for (a) achieving seamless connectivity between different public and private networks; and (b) increasing the likelihood of other vendors adopting and integrating our framework (partially or completely) into their own products and services.
- *Robustness:* By incorporating the latest solutions and specifications of the relevant SDOs, our framework will have access to the latest and most robust standardized

technologies, protocols, and interfaces, enabling it to provide more reliable and high-performance services and offer greater resilience against failures or disruptions.

- *Future-proofing:* The 6G slicing framework is still in its infancy. Its specifications are undergoing constant development within the relevant SDOs. Aligning our framework with the most recent standards ensures its continued relevance and compatibility with future updates, thereby enhancing its long-term viability.
- *Regulatory compliance:* In some regions and countries, national and regional telecommunications regulatory bodies mandate compliance with particular specifications and standards. For example, by employing features such as E2E encryption and access control, our framework can help ensure compliance with data privacy regulations, such as the General Data Protection Regulation in the European Union. Aligning our framework with the latest standards thereby ensures compliance with applicable regulations and prevents legal or financial penalties. As they demonstrate a commitment to ethical and responsible business practices, these compliances can also help build trust with customers and other stakeholders.

On the basis of these advantages, we take into account the following considerations when designing our proposed architecture and aligning it with the relevant SDOs:

- We first define the scope of each layer, and then we assign the latest specifications and solutions from the relevant SDOs to execute the tasks that we believe each layer shall perform within our framework. This should not be confused with (or misunderstood as) a simple harmonization and amalgamation of the latest specifications.
- We recognize that certain 5G standards may not fulfill 6G requirements. Nonetheless, several SDOs are currently studying the possibility of developing 6G standards. It is believed that these SDOs will either improve the existing standards or develop new ones. For example, the joint framework of the ETSI industry specification group (ISG) NFV-management and orchestration (NFV-MANO) and the 3GPP-network slicing management system (3GPP-NSMS) manages the virtualized and physical parts of a 5G NS. To fulfill the M&O requirements of 6G and support cloud-native and micro-service concepts, it is anticipated that this joint framework will be equipped with novel automation, intelligent, and data-driven mechanisms, as well as a few new functional blocks (FBs) may be added on top of the legacy ones.
- We utilize the specifications of the SDOs, which have a long history of developing standards in specific network domains. We believe that most of these SDOs will maintain their current roles. For example, we employ the 3GPP Technical Specification Group (TSG) radio access network (RAN) and O-RAN Alliance specifications to design the 6G RAN domain. This is because we anticipate that both of these SDOs will continue to play an active role in the RAN domain as we move towards 6G. It is less likely that another SDO will emerge, develop standards, and supplant the 3GPP and O-RAN Alliance.

E. THE STRUCTURE OF THE ARTICLE

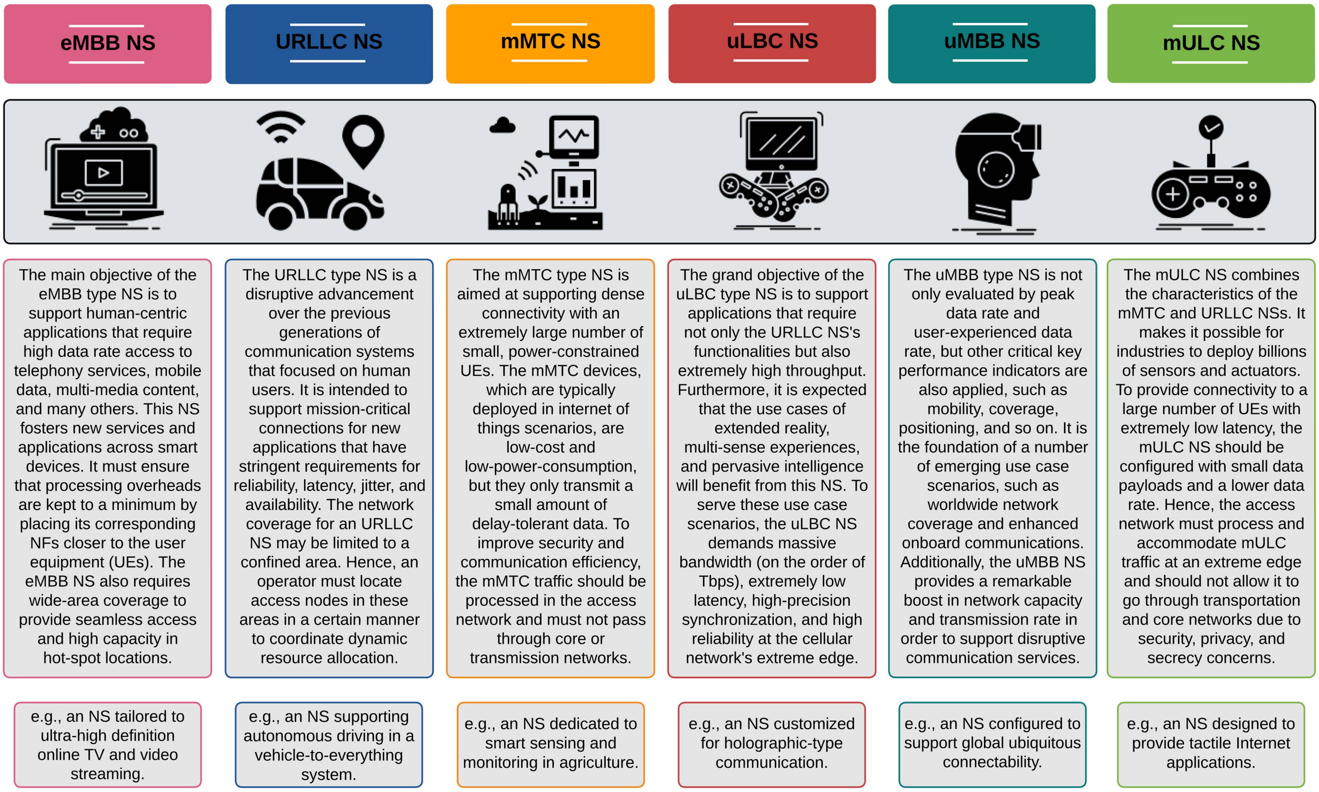
The following is a breakdown of how this article is structured. First, we present the primary motivation for extending the slicing architecture towards 6G in Section II. Then, we discuss a number of enabling technologies related to 6G slicing, including intelligent virtualization, intelligent softwarization, and intelligent cloudification, in Section III. Following that, we propose a standards-oriented, open, intelligent, and E2E framework for 6G slicing in Section IV. The proposed architecture is the primary contribution of this article; therefore, it deserves a comprehensive discussion and a complete analysis of all the domains of a 6G network. Additionally, we provide several research challenges and future directions that need to be addressed in order to provision various types of 6G NS instances in an efficient manner in Section V. Finally, we summarize the major conclusions of this article in Section VI.

II. THE MAIN MOTIVATION FOR EXTENDING NETWORK SLICING TOWARDS 6G

Before delving into technical details, it is essential to provide the reasons that motivated us to extend the legacy slicing architecture towards 6G and enable it with openness, intelligence, and automation. We believe that there are three reasons for reconstructing the slicing framework for 6G: (1) the likelihood of supporting a large number of heterogeneous use cases with a wide variety of applications; (2) the stringent requirements enabling these use cases; and (3) the increasing complexity resulting from the integration of multiple networked systems at various layers and domains, different types of software and hardware, and a large number of technologies. We discuss these motivating trends in the following.

A. DISRUPTIVE USE CASES AND POTENTIAL APPLICATIONS

The 5G slicing framework supports a diverse set of use cases and applications that have never been encountered by previous generations, such as automotive, healthcare, agriculture, etc. [43]. In the coming years, it is anticipated that the use cases enabled by the 5G slicing framework will pave the way for smart cities and the digitalization of service sectors, as well as have a revolutionary impact on the automation of vertical industries and intelligentization of manufacturing. The ITU categorizes 5G use cases and applications into three families, each of which can be supported by its corresponding type of NS [44]. Each NS is designed to meet the specific business and technical requirements of specific end-users or vertical industries in 5G and beyond networks.



eMBB NS

The main objective of the eMBB type NS is to support human-centric applications that require high data rate access to telephony services, mobile data, multi-media content, and many others. This NS fosters new services and applications across smart devices. It must ensure that processing overheads are kept to a minimum by placing its corresponding NFs closer to the user equipment (UEs). The eMBB NS also requires wide-area coverage to provide seamless access and high capacity in hot-spot locations.

e.g., an NS tailored to ultra-high definition online TV and video streaming.

URLLC NS

The URLLC type NS is a disruptive advancement over the previous generations of communication systems that focused on human users. It is intended to support mission-critical connections for new applications that have stringent requirements for reliability, latency, jitter, and availability. The network coverage for an URLLC NS may be limited to a confined area. Hence, an operator must locate access nodes in these areas in a certain manner to coordinate dynamic resource allocation.

e.g., an NS supporting autonomous driving in a vehicle-to-everything system.

mMTC NS

The mMTC type NS is aimed at supporting dense connectivity with an extremely large number of small, power-constrained UEs. The mMTC devices, which are typically deployed in internet of things scenarios, are low-cost and low-power-consumption, but they only transmit a small amount of delay-tolerant data. To improve security and communication efficiency, the mMTC traffic should be processed in the access network and must not pass through core or transmission networks.

e.g., an NS dedicated to smart sensing and monitoring in agriculture.

uLBC NS

The grand objective of the uLBC type NS is to support applications that require not only the URLLC NS's functionalities but also extremely high throughput. Furthermore, it is expected that the use cases of extended reality, multi-sense experiences, and pervasive intelligence will benefit from this NS. To serve these use case scenarios, the uLBC NS demands massive bandwidth (on the order of Tbps), extremely low latency, high-precision synchronization, and high reliability at the cellular network's extreme edge.

e.g., an NS customized for holographic-type communication.

uMBB NS

The uMBB type NS is not only evaluated by peak data rate and user-experienced data rate, but other critical key performance indicators are also applied, such as mobility, coverage, positioning, and so on. It is the foundation of a number of emerging use case scenarios, such as worldwide network coverage and enhanced onboard communications. Additionally, the uMBB NS provides a remarkable boost in network capacity and transmission rate in order to support disruptive communication services.

e.g., an NS configured to support global ubiquitous connectability.

mULC NS

The mULC NS combines the characteristics of the mMTC and URLLC NSs. It makes it possible for industries to deploy billions of sensors and actuators. To provide connectivity to a large number of UEs with extremely low latency, the mULC NS should be configured with small data payloads and a lower data rate. Hence, the access network must process and accommodate mULC traffic at an extreme edge and should not allow it to go through transportation and core networks due to security, privacy, and secrecy concerns.

e.g., an NS designed to provide tactile Internet applications.

FIGURE 2. The applications, functionalities, and supporting use cases of six types of NS instances proposed for an open, intelligent, and E2E network slicing framework in 6G.

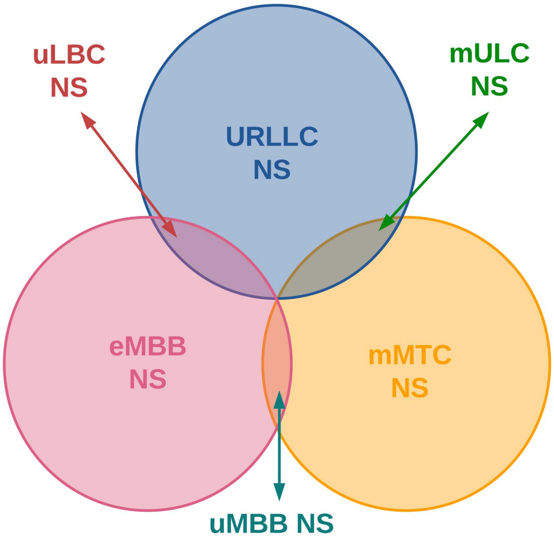


FIGURE 1. In addition to legacy 5G NSs, three enhanced emerging NSs are proposed for the 6G network slicing framework.

In the following, we provide theoretical definitions of the three types of NSs.

- *enhanced mobile broadband (eMBB) NS:* supports a stable high data rate connection with high reliability, moderate mobility, high device density, and a low packet error rate on the order of 10^{-3} [45]. The eMBB devices send large data payloads in the uplink direction to a given next generation nodeB (gNB) at a certain time [45].
- *ultra-reliable low latency communications (URLLC) NS:* supports data traffic with very high reliability (e.g., packet error rate around or lower than 10^{-5}) and very low latency transmission of small size data payloads. Due to these strict performance requirements, the URLLC traffic must be accommodated at the edge or extreme edge [45].
- *massive machine type communications (mMTC) NS:* provides connections to a large number of devices that are intermittently active and send a small size of data payloads to a gNB [45]. These devices are either stationary or move at a very low speed, requiring a low data rate and a low packet error rate on the order of 10^{-1}.

With the advent of 5G use cases and the evolution of current ones [14], there exist several unprecedented industrial, manufacturing, and end-user-related applications (such as holographic communication, extended reality, multi-sense experiences, digital twins, robotics and autonomous systems, etc. [2], [46], [47]), which have more extreme technical, environmental, business, and societal requirements than the 5G use cases. To efficiently accommodate such disruptive applications and use cases, the following three additional NSs are proposed for 6G, using a holistic and more reasonable methodology through the extension and integration of legacy

5G NSs. This is shown in Figure 1, where three systematically arranged circles are intersected in such a manner that their pairwise intersection forms three lens-shaped regions, each of which is assumed to represent a 6G NS.

- *ultra-reliable low-latency broadband communication (uLBC) NS:* combines existing URLLC services with eMBB capabilities. Per-user experienced data rate (up to Gbps), high reliability (99.99999%), and extremely low latency (less than 1ms) are challenging and conflicting KPIs that uLBC must balance well.
- *ubiquitous mobile broadband (uMBB) NS:* emphasizes on a dramatic extension of network coverage, far beyond what 5G can provide, via terrestrial networks. Ubiquitous connections with a high data rate and a network coverage of nearly 100% of the Earth's surface are targeted using a terrestrial-satellite-aerial integrated system by uMBB NS.
- *massive ultra-reliable low-latency communication (mULC) NS:* requires the provisioning of stringent URLLC services, as well as the support of a massive number of devices in a dense area (up to 10^7 per km^2). It is foreseen that highly-autonomous industrial production will benefit from mULC type of NS.

The applications, supporting use cases, and examples of the three proposed NSs, as well as 5G NSs (which are still applicable for 6G), are described in Figure 2. These use cases are expected to provide connectivity in the air/space (e.g., satellites, airplanes, space research, military applications and operations, etc.), on the ground (e.g., vehicles, high-speed trains, end-users, factories, etc.), and underwater (e.g., surface-underwater ships, underwater archaeology, underwater IoT devices, deep-sea tourism, etc.) [48], [49], [50]. There also exist a number of initiatives in the SDOs regarding the categorization of use cases for beyond 5G and 6G, for which customized NSs can be designed and instantiated. For example, the 3GPP has recently proposed two additional types of NSs [51], namely the vehicle-to-everything (V2X) and high-performance machine-type communications (HMTC). Such studies are needed both in research and standardization communities to continue to identify and categorize all types of 6G use cases and applications. It is also worth noting that the 6G service classes are not limited to those proposed in Figure 1. There may be additional service classes that can be achieved through the expansion of existing and futuristic service classes.

To provide connectivity to the above use cases in the deep sea and high altitude, it is anticipated that 6G network must meet a combination of a wide variety of requirements for an application, such as latency, throughput, reliability, device density, etc. [52]. The provisioning of the six types of NSs requires a significant reconstruction of the architecture, among other aspects, of the legacy slicing ecosystem. Hence, it is critical to first identify emerging use cases and applications and determine their required KPIs [53], [54]. Thereupon, it is essential to develop AI-driven, autonomous, and open air-sky-ground-underwater architectural solutions that are capable of routing, storing, and processing sensitive data at unprecedented levels of automation, openness, intelligence, sustainability, and trustworthiness [53]. Lastly, it is vital to define ML algorithms for performing all operations associated with a 6G NS in a time-, resource-, and energy-efficient manner.

B. STRINGENT TECHNICAL REQUIREMENTS

To design NSs for the above use cases, the 6G network needs to satisfy their heterogeneous requirements. There are numerous classifications of slicing requirements. Among them, the GSMA classification is widely adopted within the community. The GSMA categorizes the requirements of a 5G use case into three families of requirements [55]: performance, functional, and operational. The owner of an industry expresses these requirements through the use of a standardized interface, which are then converted by an operator into a format understood by the network and assumed as the characteristics of an NS. However, in order for a 6G network to meet the diverse requirements of 6G use cases and design their corresponding NSs, the requirements of the use cases described above are anticipated to exceed those defined by the GSMA [55] in terms of intensity and quantity. We believe that there could be two categories of requirements for an NS to be provisioned by the 6G network: existing requirements and futuristic requirements. We discuss the two categories below in detail and provide their estimated values in Table 2 [2], [46], [56].

The first category of requirements has already been implemented in 5G, such as throughput, mobility, latency, etc. These KPIs are expected to be improved to capture the stringent requirements of a 6G use case [46]. The rapid progress in wireless technologies, such as Terahertz communications and massive multiple input multiple output, will pave the way for high network performance. Thus, while some KPIs will be optimized by one order of magnitude as the evolution from fourth generation (4G) to 5G proceeds, others will be dramatically improved by several orders of magnitude, such as peak data rate, which is envisioned to increase from 20Gbps to 1Tbps [46], and connection density, which is expected to increase up to $10 - 100$ million devices per km^2 for massively connected machines [57].

The second category of requirements, which is driven by emerging technological enablers – most notably the explosion of ML algorithms and industrial and perception Internet [58] – considers a plethora of new requirements that go beyond deterministic performance measures, such as global roaming, quantum-resistant security, and sensing accuracy [56]. The purpose of implementing these requirements, which were not considered in previous generations, is intended to support not only the communication service of a 6G NS, but also other services, such as more precise localization and accurate sensing, which are inherent characteristics of a 6G NS.

TABLE 2. The qualitative and quantitative comparison of the existing and futuristic requirements of a 6G NS instance. It is critical to take into account these requirements when configuring the six types of 6G NS instances.

Category	Parameter	Estimated qualitative and quantitative values for 6G NS instances					
		eMBB	**URLLC**	**mMTC**	**uLBC**	**uMBB**	**mULC**
Existing requirements	Availability [%]	99.999	99.9999	99	99.9999	99.999	99.9999
	Reliability (frame error rate)	10^{-5}	10^{-7}	10^{-3}	10^{-7}	10^{-5}	10^{-7}
	Mobility [km/h]	10^3	250	$0-30$	250	10^3	100
	Latency	10 µsec	100 µsec	$1-10$ ms	100 µsec	10 µsec	100 µsec
	Security	Extremely high	Extremely high	Very high	Extremely high	Very high	Extremely high
	Priority	High	Very high	Medium	Very high	High	Very high
	Device density [UEs/km^2]	$\approx 10^5$	$\approx 10^3$	$\approx 10^7$	$\approx 10^4$	$\approx 10^6$	$\approx 10^7$
	Isolation	Very high	Extremely high	Medium	Very high	High	Very high
	Downlink per NS [Tbps/km^2]	≥ 1	≥ 0.5	≥ 0.4	≥ 1	≥ 0.75	≥ 0.5
	Uplink per NS [Tbps/km^2]	≥ 0.75	≥ 0.25	≥ 0.25	≥ 0.75	≥ 0.5	≥ 0.25
	Downlink per user [Gbps]	$\approx 10-100$	$\approx 1-10$	$\approx 0.1-1$	$\approx 10-100$	$\approx 1-100$	$\approx 0.1-1$
	Uplink per user [Gbps]	$\approx 1-10$	$\approx 0.1-1$	$\approx 0.01-0.1$	$\approx 1-10$	$\approx 1-10$	$\approx 0.1-1$
	Mobility interruption time [ms]	<5	<1	<50	<1	<5	<1
	Position accuracy	dc/m	cm/dm	cm – m	cm/dm	dc/m	cm/dm
	Energy efficiency	Very high	Very high	Extremely high	High	Very high	High
	Device battery life	High	High	Extremely high	Medium	Medium	Very high
	Coverage type	Wide area	Private	Wide area	Private	Global	Wide area
Futuristic requirements	Three dimensional coverage	High	Low	Low	Medium	High	Medium
	Quantum-resistant security	Medium	High	Medium	High	Medium	High
	Global roaming	Medium	Low	Low	Medium	Very high	Low
	Computation/memory efficiency	Medium	Medium	High	High	Medium	High
	AI/ML model convergence time	Medium	High	High	High	Medium	High
	Quality of decision	Low	High	Low	High	Medium	High
	Computational round-trip time	Medium	Very high	Low	High	Medium	High
	Periodicity	Low	High	Very high	High	High	Very high
	Determinism	High	Very high	Low	High	High	Very high
	Isochronicity	Low	High	Very high	Medium	High	High
	Versatility	Medium	High	Very high	High	High	Medium
	Sustainability	High	High	Very high	High	High	High
	Resilience	Very high	Very high	Low	Medium	High	High
	Situation-awareness	Very high	Very high	Medium	High	High	High
	Reconfigurability	Very high	High	High	High	High	High
	Localization accuracy	Very high	Very high	High	Medium	High	High
	Sensing accuracy	Medium	Medium	Medium	High	Medium	High

Note: (a) The values of some requirements are use-case-specific. Nevertheless, they can be considered for the majority of 6G use cases with varying degrees of intensity. (b) As of this writing, the specifications for 6G have not been officially defined. However, where applicable, we provide qualitative estimates based on the existing standardized requirements for 5G. (c) The values of these requirements are predicted by the authors based on information provided in Releases 16–18 of 3GPP and ITU. (d) The letters m, cm, and dm stand for meter, centimeter, and decimeter, respectively.

To fully embrace the vision of fulfilling both categories of requirements, it is critical to re-engineer the existing slicing architecture and equip it with intelligence and openness aimed at automating the operations and maintenance of various 6G NSs, as well as eliminating vendor lock-in across multiple domains of the 6G slicing framework. Once these requirements are satisfied, it is anticipated that the 6G slicing framework will be a more open, intelligent, dependable, and reliable architectural framework compared to its predecessor. In addition to communication data, the localization and sensing data that are frequently generated within the domain of a vertical industry are extremely sensitive [14]. This data can be used by the 6G network to provide localization and sensing services. Therefore, securing this type of information is an essential requirement that needs to be taken into consideration during the design phase of the 6G slicing framework [7].

C. THE INCREASING COMPLEXITY OF SYSTEMS INTEGRATION

The 5G slicing framework is composed of a variety of software and hardware from different vendors and SDOs. Such a complex and highly engineered framework is essential to meet the business demands and technical requirements of various use cases in a coherent and harmonized manner [56]. Nevertheless, the legacy slicing framework confronts tremendous challenges, including inadaptability to dynamic changes in an NS, inability to autonomously manage resources and network traffic, lack of interoperability among software and hardware components from different vendors, and human-caused errors in network management and service orchestration [1]. It is evident that the 6G slicing framework is likely to be even more heterogeneous and complex than its predecessor [46], expecting to support innumerable heterogeneous use cases and applications in the coming decade. As a result, conventional human-machine integration algorithmic solutions and proprietary lock-in architectural schemes are incapable of efficiently meeting these demands in terms of network resources, total cost, energy consumption, time to market, trustworthiness, sustainability, innovation, and other KPIs.

Among the above demands, energy efficiency is a critical issue that needs to be considered while designing and integrating complex 6G systems, as it directly impacts the operational cost and environmental impact of 6G. 6G is expected to be more complex and data-intensive than previous generations, which can result in increased energy consumption. Therefore, it is essential to develop energy-efficient solutions to reduce power consumption while improving system performance. One approach is to use advanced hardware and software technologies, such as energy-efficient transceivers and power management techniques. Furthermore, the use of renewable energy sources, such as solar and wind power, in the operation of 6G networks can also contribute to reducing their environmental impact.

The utilization of ML algorithms and openness in slicing framework are believed to be in the best interests of operator and tenant, as they enable the tight integration, full synchronization, intelligentization, and interoperability of multiple systems, components, and technologies aimed at efficiently overcoming the above challenges [2], [6]. To that end, a major rethinking of the various aspects of the legacy slicing framework is required to cope with the 6G use cases and enabling technologies. Such an autonomous, open, intelligent, and tightly integrated framework will be capable of adapting to dynamic changes associated with business objectives or environmental conditions of 6G NSs, which could have profound implications for network performance and tenant satisfaction. Therefore, human-dependent decision-making processes will be reduced, management and network complexity will be lowered, and the total cost of ownership will be decreased.

III. KEY ENABLING TECHNOLOGIES FOR REALIZING NETWORK SLICING IN 6G

In this section, we focus on three enabling technologies associated with the deployment of 6G slicing, namely intelligent virtualization, intelligent softwarization, and intelligent cloudification. There are several core, transport, cloud, and access technologies, as well as intent-based mechanisms and architectural solutions, that affect the slicing framework in direct or indirect ways. However, we concentrate only on the most pertinent enablers that have spawned 5G slicing and are believed to play a significant role in the slicing of 6G networks. Additionally, we elaborate on how intelligence and openness can enhance the performance of these enablers in 6G slicing.

A. INTELLIGENT VIRTUALIZATION

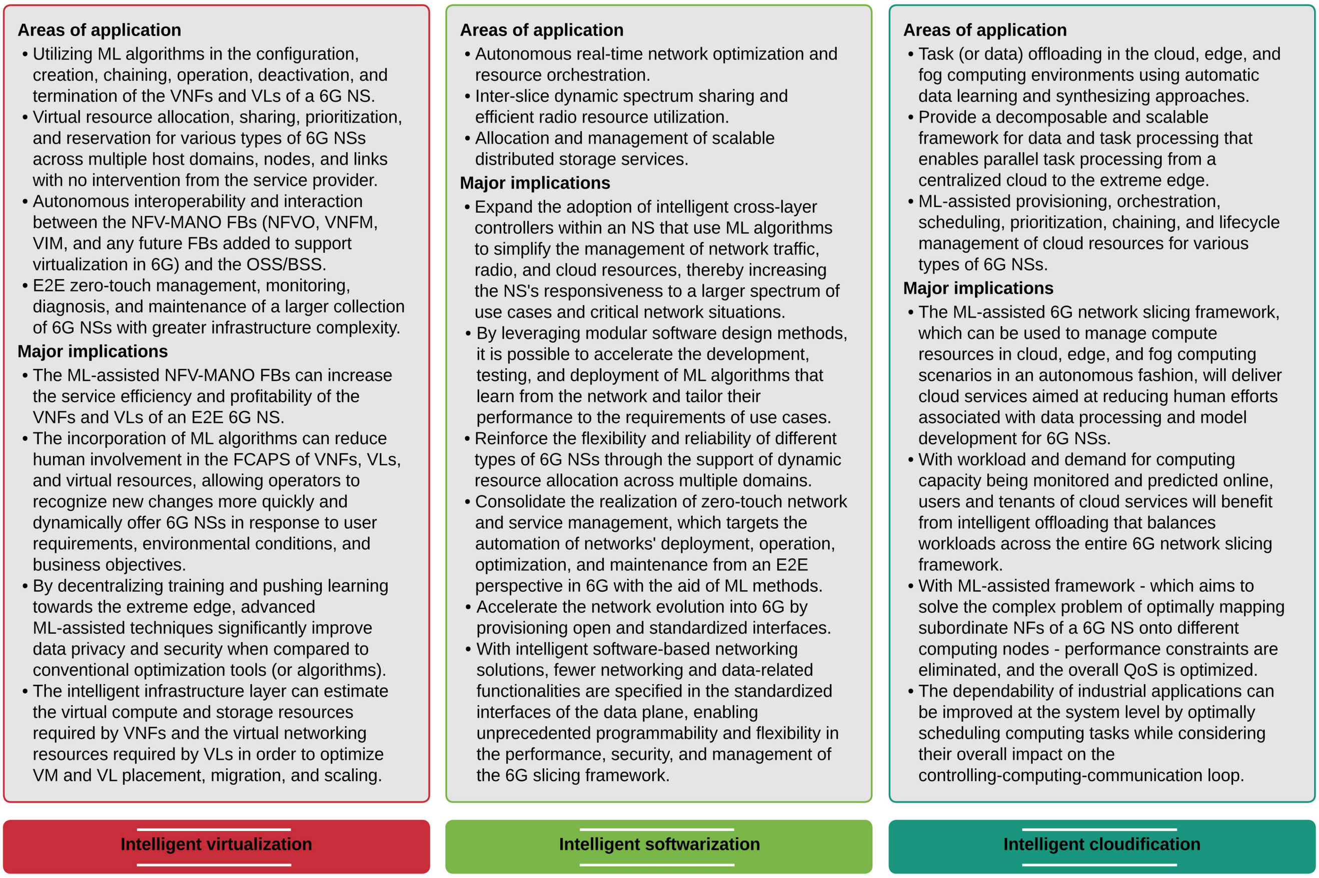
The network function virtualization (NFV) paradigm emerged as a critical enabler of 5G slicing and continues to be an integral component of the 6G slicing [5]. It enables the creation of virtual resources, enhances the deployment flexibility of NSs, and facilitates the integration of diverse communication, sensing, and localization services. Nevertheless, with the deployment of NFV in the early stages of 5G, a large number of divergent virtual network functions (VNFs) with novel engineering considerations such as an unprecedented level of sensitivity, criticality, versatility, priority, replicability, composability, and transferability have been identified. These heterogeneous VNFs call for the deployment of intelligent optimization algorithms capable of coping with network complexity and connection density at the core, transport, edge, and extreme edge domains of 6G networks [18].

One solution is the utilization of ML algorithms, which have been employed in various aspects of virtualized networks, most notably in network automation, resource orchestration, and service management [18], [59]. A large number of ML algorithms are supervised, semi-supervised, and unsupervised, which significantly improve operational efficiency, increase the effectiveness of high-volume data processing, and reduce total capital and operating expenditures. Nevertheless, their deployment in 5G has raised privacy, security, secrecy, and scalability concerns [2], [59]. Additionally, due to constant changes in network topologies, service requirements, and NS configurations, it is time-consuming to update the dataset and adapt these algorithms to a changing environment [18].

To this end, by incorporating RL, FL, and DL into NFV-enabled 6G networks and adopting openness in the underlying infrastructure, it is possible to optimize a wide range of network operations and maintenance tasks to achieve the required level of intelligence in resource scheduling, openness and innovation, network elasticity, and troubleshooting, while ensuring data privacy, network security, and user secrecy [59]. In Figure 3, we list several applications of these algorithms with the goal of designing an intelligent, credible, reliable, open, and sustainable NFV paradigm for 6G slicing. Notably, integrating ML has implications for a variety of operations in NFV-enabled networks, as also listed in Figure 3.

B. INTELLIGENT SOFTWARIZATION

A significant fraction of 5G technologies have softwarization as the core to their realization through software-defined everything (SDx) technologies. Softwarization refers to the realization of hardware-based functions using software in order to enable a flexible, dynamic, reconfigurable, and upgradable implementation of said functions [60]. Softwarization technology goes hand in hand with



Areas of application

- Utilizing ML algorithms in the configuration, creation, chaining, operation, deactivation, and termination of the VNFs and VLs of a 6G NS.
- Virtual resource allocation, sharing, prioritization, and reservation for various types of 6G NSs across multiple host domains, nodes, and links with no intervention from the service provider.
- Autonomous interoperability and interaction between the NFV-MANO FBs (NFVO, VNFM, VIM, and any future FBs added to support virtualization in 6G) and the OSS/BSS.
- E2E zero-touch management, monitoring, diagnosis, and maintenance of a larger collection of 6G NSs with greater infrastructure complexity.

Major implications

- The ML-assisted NFV-MANO FBs can increase the service efficiency and profitability of the VNFs and VLs of an E2E 6G NS.
- The incorporation of ML algorithms can reduce human involvement in the FCAPS of VNFs, VLs, and virtual resources, allowing operators to recognize new changes more quickly and dynamically offer 6G NSs in response to user requirements, environmental conditions, and business objectives.
- By decentralizing training and pushing learning towards the extreme edge, advanced ML-assisted techniques significantly improve data privacy and security when compared to conventional optimization tools (or algorithms).
- The intelligent infrastructure layer can estimate the virtual compute and storage resources required by VNFs and the virtual networking resources required by VLs in order to optimize VM and VL placement, migration, and scaling.

Intelligent virtualization

Areas of application

- Autonomous real-time network optimization and resource orchestration.
- Inter-slice dynamic spectrum sharing and efficient radio resource utilization.
- Allocation and management of scalable distributed storage services.

Major implications

- Expand the adoption of intelligent cross-layer controllers within an NS that use ML algorithms to simplify the management of network traffic, radio, and cloud resources, thereby increasing the NS's responsiveness to a larger spectrum of use cases and critical network situations.
- By leveraging modular software design methods, it is possible to accelerate the development, testing, and deployment of ML algorithms that learn from the network and tailor their performance to the requirements of use cases.
- Reinforce the flexibility and reliability of different types of 6G NSs through the support of dynamic resource allocation across multiple domains.
- Consolidate the realization of zero-touch network and service management, which targets the automation of networks' deployment, operation, optimization, and maintenance from an E2E perspective in 6G with the aid of ML methods.
- Accelerate the network evolution into 6G by provisioning open and standardized interfaces.
- With intelligent software-based networking solutions, fewer networking and data-related functionalities are specified in the standardized interfaces of the data plane, enabling unprecedented programmability and flexibility in the performance, security, and management of the 6G slicing framework.

Intelligent softwarization

Areas of application

- Task (or data) offloading in the cloud, edge, and fog computing environments using automatic data learning and synthesizing approaches.
- Provide a decomposable and scalable framework for data and task processing that enables parallel task processing from a centralized cloud to the extreme edge.
- ML-assisted provisioning, orchestration, scheduling, prioritization, chaining, and lifecycle management of cloud resources for various types of 6G NSs.

Major implications

- The ML-assisted 6G network slicing framework, which can be used to manage compute resources in cloud, edge, and fog computing scenarios in an autonomous fashion, will deliver cloud services aimed at reducing human efforts associated with data processing and model development for 6G NSs.
- With workload and demand for computing capacity being monitored and predicted online, users and tenants of cloud services will benefit from intelligent offloading that balances workloads across the entire 6G network slicing framework.
- With ML-assisted framework - which aims to solve the complex problem of optimally mapping subordinate NFs of a 6G NS onto different computing nodes - performance constraints are eliminated, and the overall QoS is optimized.
- The dependability of industrial applications can be improved at the system level by optimally scheduling computing tasks while considering their overall impact on the controlling-computing-communication loop.

Intelligent cloudification

FIGURE 3. The implications and application areas of incorporating automation and intelligence into the three key enabling technologies that facilitate the deployment of open and intelligent network slicing in the 6G.

virtualization technology, as only with the application of the latter is it possible to achieve the abstraction of hardware into virtual instances. Thus, allowing the utilization of software with virtual machines (VMs), containers, and/or unikernels on general-purpose hardware. SDx technologies include software-defined networking (SDN), software-defined storage (SDS), and software-defined radio (SDR). The SDN paradigm has been indispensable for the operation of 5G NSs by decoupling the network into application, control, and data forwarding planes [61]. SDRs allow for a more agile fronthaul transport network (TN) as well as more efficient spectrum utilization. SDS frameworks simplify the administration, provisioning, management, orchestration, and maintenance of cloud storage environments.

The common factor among the SDx solutions is the separation of the control layer and data-handling hardware [61]. This split enables 6G RAN to evolve into a cognitive network by leveraging ML algorithms and automation at the controller with a centralized view of the network in order to manage, maintain, orchestrate, and optimize the underlying infrastructure intelligently. SDx solutions coupled with ML algorithms will also help in the development of intelligent routing protocols with reduced latency and more reliable performance, as well as with the deployment of more secure and open application programming interfaces (APIs).

Given this intelligence and openness, the 6G slicing framework will continue to learn from the traffic and data it processes and transports, adapt to the ever-demanding and heterogeneous use-case requirements, and enhance its performance to meet the tenants' expected quality of service (QoS) and quality of experience (QoE) needs. Notably, ML algorithms can be integrated into various aspects of the softwarization in the 6G slicing framework, which are listed in Figure 3. In addition, the intelligentization of softwarization technology within the 6G slicing framework will also have a number of implications, which are also listed in Figure 3.

C. INTELLIGENT CLOUDIFICATION

The cloudification of network services enables mobile operators to develop an agile, flexible, and scalable slicing framework for 5G [9]. This trend, however, has recently been accelerated by pioneering efforts to integrate ML algorithms aimed at significantly improving network performance while enabling E2E automation, full openness, service innovation, and resource isolation. There is consensus that pervasive intelligence will be an indispensable component of 6G networks and that efficient cloudification is required to enable the majority of 6G use cases and applications [2], [6]. Consequently, 6G will be an intelligence-delivery and cloudified infrastructure rather than a solely information-transmitting network.

This new role for mobile networks necessitates a paradigm shift in the way they are evaluated. As controlling, computing, and communication converge more tightly than ever before in 6G, the QoS will be challenged not only by communication performance, but also by cloudification performance [62]. Moreover, various use cases and applications may have significantly different performance requirements for cloud services. It demands that we extend service heterogeneity from the data networking domain, where it was addressed by 5G, into the cloud computing domain. Therefore, use-case-specific heterogeneous management of cloud resources and tasks shall be implemented efficiently, agilely, and adaptably, as well as integrated into the 6G network slicing framework.

The forthcoming efforts towards intelligent and open cloudification in 6G shall begin with the definition of cloud services and their associated QoS requirements, followed by the gradual integration of open hardware and applications into the cloud domain. Then, it calls for novel methods to address the most significant computing problems of cloud services, such as ML-based resource allocation, load splitting, task scheduling, and service chain management [62]. We anticipate coupling these tasks with virtualization and softwarization technologies, in which slice-specific approaches shall be developed to orchestrate and manage cloud resources across aggregated data centers, i.e., local and regional, and distributed computing units, i.e., edge servers and fog nodes [62].

IV. THE PROPOSED OPEN, INTELLIGENT, AND E2E 6G NETWORK SLICING ARCHITECTURAL FRAMEWORK

Following a detailed discussion of the key enablers that intelligently empower slicing in 6G, we propose in this section an open, intelligent, and E2E framework that is capable of hosting a massive number of 6G NSs. Our architecture leverages ML, adopts openness, and utilizes the latest de facto and de jure standards to serve as a preliminary 6G architecture. The simplified and brief design of our framework is depicted in Figure 4. Inspired by the Telecommunication Management Network Functional Architecture of the ITU [63], it consists of eight layers: Application Layer, Design Layer, Management and Orchestration Layer, Network Slice Layer, Infrastructure Layer, Network Intelligence Layer, Network Exposure Layer, and Business Management Layer. This framework is anticipated to permit future additions or replacements of layers. These layers can interact directly or indirectly for the exchange of information or exposure of certain data via secure inter-layer APIs. The inter-layer interactions are not restricted to those shown in Figure 4. We present the minimally required inter-layer connections to demonstrate the behavior of our framework. Depending on the requirements of the architectural design of the operator and the solutions of the vendor, several inter-layer APIs can be added to or removed from this framework aimed at efficiently serving the 6G slicing ecosystem.

The proposed framework can be presented using three distinct perspectives (i.e., views): (i) Structural Perspective, discussing layers, managed objects, managing objects, and interconnection among these components via APIs; (ii) Functional Perspective, describing how these components can interact to provide the relevant behaviour and complex functionalities, as well as focuses on how our framework supports these interactions and processes; and (iii) Deployment Perspective, analyzing how the components from the Structural Perspective could be deployed in practice, considering the infrastructural resources, topological aspects, and the service level agreement (SLA). The objective is for these perspectives to collectively provide a coherent and comprehensive description of our framework, with each perspective conveying relevant information about distinct and more specific aspects, as well as to give a high-level yet exhaustive description of what a 6G slicing framework might look like. In this paper, we focus on the Structural and Functional Perspectives. We discuss the components of each layer in detail and focus on their behavior within the architecture and with external sources. The Deployment Perspective is beyond the scope. We leave it up to the operator to decide how to implement our architecture in the real world across a single or multiple administrative domains, each of which may belong to a different infrastructure provider.

In the rest of this section, the specially-constructed architectures, as well as the internal and external interconnections, of the eight layers are discussed in detail. Moreover, we illustrate the layer-specific architectures of these layers in a number of figures, which are derived from the extension of Figure 4.

A. APPLICATION LAYER

This layer contains a comprehensive and up-to-date list of 6G use cases and applications that need connectivity through the use of 6G slicing framework, including and beyond those discussed in Section II. Each use case or application may include one or multiple 6G services, such as communication, sensing, localization, intelligence, and others [64]. In addition, this layer could incorporate high-level KPIs and KVIs relevant to a use case or application. Defining the ranges and intensities of the KPIs and KVIs provide a more accurate qualitative and quantitative analysis of 6G use cases, as well as their expected performance and social impact. By including this information in Application Layer, operators can better align their services with the needs of their users and ensure that they are delivering high-quality experiences. Therefore, this layer shall also include an exhaustive list of 6G services (as well as their KVIs and KPIs) that can be utilized by industry proprietors. Furthermore, this layer determines the structure of a service (e.g., simple, composite, or nested) that needs to be provisioned via a 6G NS. This inclusive list of 6G applications and use cases, as well as a

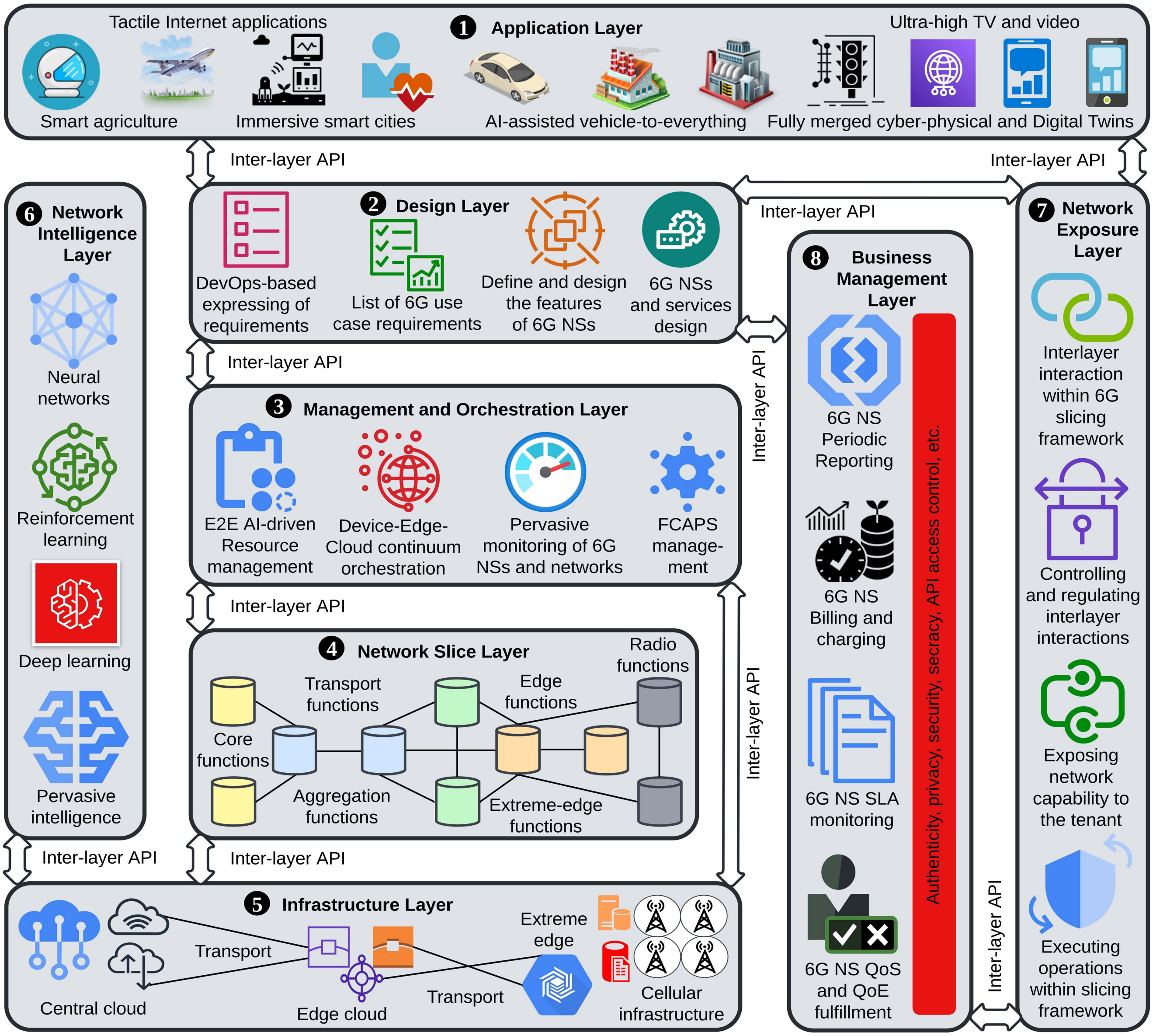
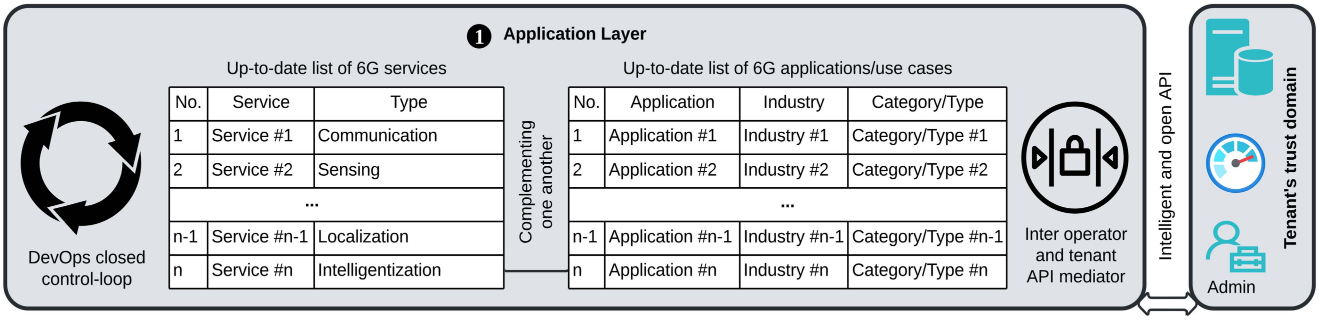

1 Application Layer
Tactile Internet applications
Ultra-high TV and video
Smart agriculture
Immersive smart cities
AI-assisted vehicle-to-everything
Fully merged cyber-physical and Digital Twins
Inter-layer API
Inter-layer API
2 Design Layer
DevOps-based expressing of requirements
List of 6G use case requirements
Define and design the features of 6G NSs
6G NSs and services design
Inter-layer API
Inter-layer API
3 Management and Orchestration Layer
E2E AI-driven Resource management
Device-Edge-Cloud continuum orchestration
Pervasive monitoring of 6G NSs and networks
FCAPS manage-ment
Inter-layer API
4 Network Slice Layer
Radio functions
Transport functions
Edge functions
Core functions
Aggregation functions
Extreme-edge functions
Inter-layer API
Inter-layer API
Inter-layer API
6 Network Intelligence Layer
Neural networks
Reinforcement learning
Deep learning
Pervasive intelligence
Inter-layer API
5 Infrastructure Layer
Central cloud
Transport
Edge cloud
Transport
Extreme edge
Cellular infrastructure
8 Business Management Layer
6G NS Periodic Reporting
6G NS Billing and charging
6G NS SLA monitoring
6G NS QoS and QoE fulfillment
Authenticity, privacy, security, secracy, API access control, etc.
Inter-layer API
7 Network Exposure Layer
Interlayer interaction within 6G slicing framework
Controlling and regulating interlayer interactions
Exposing network capability to the tenant
Executing operations within slicing framework

FIGURE 4. The simplified and brief design of the proposed framework for an open, intelligent, and E2E slicing in 6G. The layer-specific architectures, as well as their internal and external interconnections, derived from this figure are illustrated later.

list of their required services, can be stored in the domain of a service provider and made available to the tenant via a business-driven catalog, such as using an online self-service portal [65], as shown in Figure 5.

To date, various industry consortiums and SDOs identified and categorized several use cases, applications, and services for beyond 5G and 6G. During this study, we observed a significant lack of coordination and, at some points, conflict among the organizations focusing on the identification of the use cases and services. The majority of these institutions either identified the same use cases and services repeatedly or established varying ranges and values for the requirements of their customized NSs. For example, GSMA [55] and 3GPP [66] identified use cases and services for V2X systems. Both discovered the same use cases and services within the V2X industry. However, the requirements defined for them vary. This lack of coordination between these organizations leads to resource and time waste, and on the other hand, may confuse the service provider and tenants as to which standards should be followed during the design and provisioning of an NS, as well as the negotiation phase of SLA. Hence, it is essential for the organizations that contribute to slicing to jointly establish a unified platform aimed at uniformly defining, harmonizing, and classifying 6G use cases and applications, as well as consistently describing their technical, business, and regulatory requirements. This will allow service providers and tenant to refer to a standard list of use cases and services, as well as their respective


1 Application Layer
DevOps closed control-loop
Complementing one another
Inter operator and tenant API mediator
Intelligent and open API
Tenant's trust domain
Admin

Up-to-date list of 6G services

No.	Service	Type
1	Service #1	Communication
2	Service #2	Sensing
	...	
n-1	Service #n-1	Localization
n	Service #n	Intelligentization

Up-to-date list of 6G applications/use cases

No.	Application	Industry	Category/Type
1	Application #1	Industry #1	Category/Type #1
2	Application #2	Industry #2	Category/Type #2
	...		
n-1	Application #n-1	Industry #n-1	Category/Type #n-1
n	Application #n	Industry #n	Category/Type #n

FIGURE 5. The detailed architecture of the Application Layer of the proposed framework for network slicing in 6G.

requirements, to design and provision the corresponding 6G NSs timely, efficiently, and accurately.

The Application Layer leverages closed-loop to continuously automate updating 6G service and use case lists by combining development and operations (DevOps) techniques with a high level of automation [67]. The DevOps approaches, e.g., utilizing continuous integration and continuous delivery (CI/CD) pipelines [68], enable this layer to dynamically capture, develop, test, integrate, and update standard and non-standard 6G services and use cases. They also automatically renovate service and use case catalogs. This approach to developing and updating both lists, on the one hand, eliminates manual entry of service and use case details, which results in human-caused errors, and, on the other hand, provides service provider with the ability to offer tenants up-to-date services available in the market, lays out fast and reliable problem-solving techniques, and automates repetitive management tasks to a higher degree. The CI/CD methodology is an innovative feature of a telco-grade environment; however, its implementation in cellular networks is challenging. This is because, in a telco-grade environment, service development and catalog updates are typically collaborative efforts involving multiple vendors, tenants, and the development and operations teams of the service provider. These stakeholders have distinct corporate cultures and interests. The implementation of this methodology will likely necessitate addressing not only technical, but also cultural, procedural, economic, and societal challenges that need to be tackled by the Design Layer [14], [16].

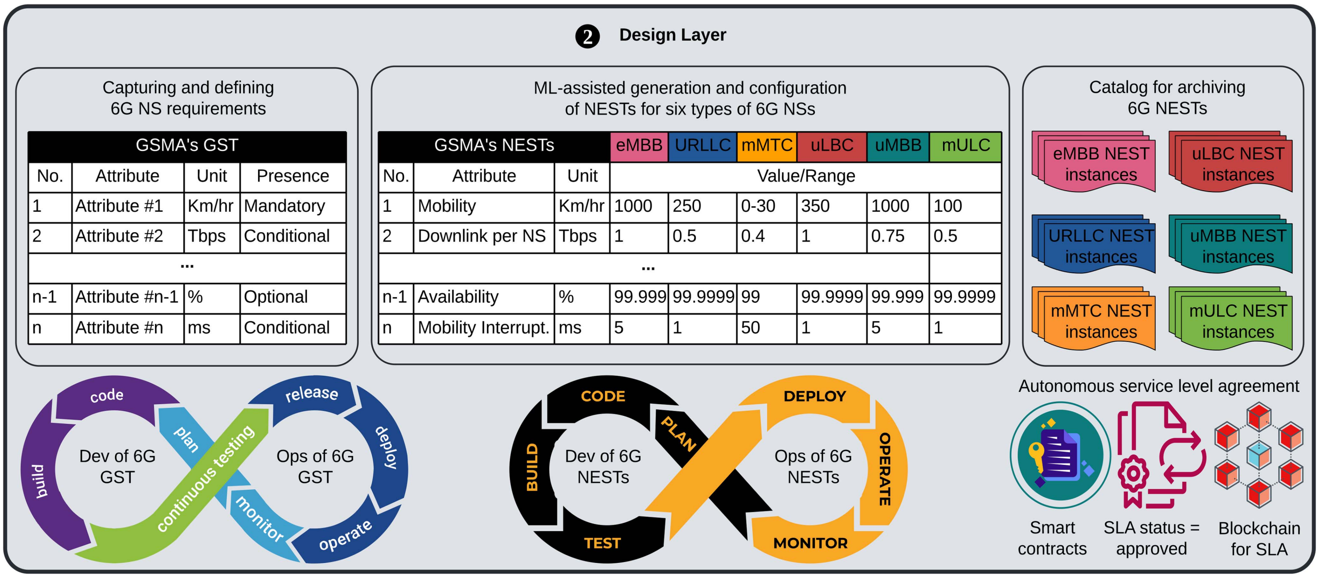
We assume that updated lists of 6G use cases and services exist, that their requirements have been precisely defined, and that the SDOs and consortiums have consensus on the values or ranges of these requirements. Based on these assumptions, as shown in Figure 5, the tenant uses an open, intelligent, and secure API [69] to access the catalog of use cases and then automatically notify the service provider about the selected use cases that best meet its requirements. To enable rapid, intelligent, zero-touch, and secure interoperability among the layers or with the tenants, the Application Layer can leverage the Open API suite of the TMF [70]. The TMF Open API is composed of a common language and design principles aimed at producing open interfaces to allow M&O M&O solutions, encompassing all aspects of provisioning services, from the formation of the consortium offering the service to its proposition, activation, operation, integration, maintenance, and monetization. In addition, the tenant must choose the appropriate services and vertical-specific applications for the selected use cases, as there may be homogeneous use cases and applications with distinct services [71], or multiple use cases within a single vertical industry [43]. Assuming the tenant selected the desired 6G use cases and services, this layer can interact with the Design Layer (discussed in Section IV-B) to design the respective 6G NSs, along with their associated flavors, instantiation levels, and geolocation parameters [72].

B. DESIGN LAYER

The Design Layer tackles the design of the characteristics and behavior of a 6G NS. This layer is mainly in charge of two design-related tasks that result in the layout of a complete logical network. First, it captures, defines, and translates the requirements of the selected use cases and services into the characteristics of their respective NS. Second, it determines the metrics related to the SLA. They are described below.

1) AUTOMATIC CAPTURING, DEFINING, AND TRANSLATING REQUIREMENTS RELATED TO THE 6G USE CASES AND SERVICES

The Design Layer captures, defines, and translates use case-specific and service-specific requirements (such as those outlined in Table 2 and beyond) of a vertical industry into a format that the service provider can comprehend. In addition, it determines the value and/or range of each requirement to accurately design its corresponding 6G NS. The requirements of a vertical industry include performance-related, operational-related, and functional-related requirements, as categorized by the GSMA [55]. This layer must also capture, define, and translate the geolocation and topological requirements. To that end, it must be capable of autonomously establishing a point of convergence between the tenant and the service provider in terms of their understanding of the requested use cases and services, as well as their requirements. Therefore, utilizing a standardized and unified template that contains an exhaustive and up-to-date list of all requirement attributes of vertical industries that can be



GSMA's GST			
No.	Attribute	Unit	Presence
1	Attribute #1	Km/hr	Mandatory
2	Attribute #2	Tbps	Conditional
...			
n-1	Attribute #n-1	%	Optional
n	Attribute #n	ms	Conditional

GSMA's NESTs			eMBB	URLLC	mMTC	uLBC	uMBB	mULC
No.	Attribute	Unit	Value/Range					
1	Mobility	Km/hr	1000	250	0-30	350	1000	100
2	Downlink per NS	Tbps	1	0.5	0.4	1	0.75	0.5
...								
n-1	Availability	%	99.999	99.9999	99	99.9999	99.999	99.9999
n	Mobility Interrupt.	ms	5	1	50	1	5	1

FIGURE 6. The detailed architecture of the Design Layer of the proposed architectural framework for network slicing in 6G.

employed for the design of 6G NSs is an efficient way of achieving this objective.

To construct a standard-compliant list of required attributes for 6G NSs, this layer can employ the Generic NS Template (GST) of the GSMA [73]. As shown in Figure 6, the GST is a universal blackprint for attributes. It was intended to define the mandatory, conditional, and optional attributes of 5G NSs. However, owning to its scope and effectiveness, the GST can be extended by adding requirements for 6G use cases and used to qualitatively and quantitatively determine a comprehensive set of attributes that characterize 6G NSs. Once a GST has been defined, it can be stored at the Design Layer and utilized by both the tenant and the service provider to design the behavior and characteristics of their desired 6G NSs.

Furthermore, this layer defines ranges and values for the attributes of GST in accordance with the requirements of a 6G NS. The service provider may authorize the tenant to configure the values or ranges of particular (or all) attributes of GST via a secure API. This authorization level and the number and diversity of attributes specified by the tenant can be defined in the attribute exposure policy. In this regard, the Design Layer can also use the NEtwork Slice Type (NEST) of GSMA [73]. As shown in Figure 6, the NEST is a filled-in version of the GST used to assign values or ranges to a finite set of attributes based on use case-specific and service-specific requirements of a vertical industry. The types, flavors, and instantiation levels of the 6G NS can be determined by defining the NEST. The outcome of designing a NEST is a fully-ordered NS whose resources must be managed and orchestrated by the 6G slicing framework. This NS order includes the necessary information for mapping the components of the designed 6G NS onto the CN, TN, and RAN infrastructure, which will be discussed in Section IV-E. This layer may generate different NESTs as a result of the varying requirements of the 6G NSs to be served. It must also be capable of designing standard and operator-specific NESTs for the 6G NSs shown in Figure 2 and beyond, and be able to archive these NESTs in its catalog for future use. Moreover, in addition to KPIs, the GST and NEST can also incorporate KVIs for the target 6G use case or application, and map them to the 6G NS design. The mapping of KVIs can be similar to that of mapping KPIs into a NS design.

Compiling a GST from a set of attributes, setting the values and ranges of the attributes in the NESTs, and assigning optional, conditional, and mandatory visibilities (i.e., statuses) to the attributes of the NESTs are challenging tasks that must be accomplished timely, efficiently, and automatically. The majority of service design, attribute selection, and determining information content-related tasks of a service and/or template catalog are performed manually [74]. We believe that automation and intelligentization will considerably improve the performance of catalogs with respect to the above tasks [75]. To that end, we propose a combination of automation and DevOps mechanism as an appropriate tool for achieving these objectives. As shown in Figure 6, through the use of DevOps automation, the Design Layer continuously captures up-to-date attribute requirements, adds the operator-understandable forms of the recently captured attributes to the GST and NESTs, assigns values and ranges to these attributes, and updates the overall statuses of the GST and NESTs automatically.

In addition to automating and integrating DevOps mechanisms, the Design Layer can use this method to continuously monitor, maintain, test, and update the internal structures, interfaces, codes, and software components of the 6G GST and NESTs. The DevOps task can be executed by a vendor in collaboration with a 6G service provider to design and keep up-to-date the service catalogs in a manner that must accurately capture the required attribute requirements, evaluate their consistency, and adapt to the requirements of the 6G service providers and tenants. Finally, adopting openness in the design of software components and internal structures of the service catalogs and description files can assist designers and developers within the open-source community, as well as the service provider, in designing and optimizing the number and diversity of attribute requirements, the scope of the information content, access control policies, attribute exposure policies, and the performance of GST and NEST [76].

2) AUTOMATIC DETERMINATION AND DYNAMIC NEGOTIATION OF SLA METRICS FOR THE REQUESTED 6G NS INSTANCE

The second essential duty of the Design Layer is to automatically determine and dynamically negotiate the metrics associated with autonomous SLA for the desired 6G NS, such as temporal, business, legal, and contractual metrics [77]. The determination of the SLA metrics must be performed after the NEST for the requested 6G services and use cases has been designed or selected from the service catalog. Consequently, the performance, functional, and operational requirements specified in the NEST can also be considered as metrics of the SLA. In the context of the SLA, the requirements of the NEST can be referred to as QoS metrics. These technical and non-technical metrics shall be determined transparently during the preparation phase of the future 6G SLA to satisfy auditability, trustworthiness, reliability, accountability, enforcement, and monitoring requirements of a smart, auto-executable, and immutable contractual mechanism [78], [79], [80].

Following the automatic and transparent determination of metrics, the service provider and tenant negotiate the SLA metrics and construct and sign the SLA after all terms and conditions have been successfully negotiated [81]. The Design Layer can use Blockchain and Smart Contract technologies to prepare, enforce, and monitor the SLA [79]. These technologies can empower the 6G SLA with cutting-edge capabilities, such as maintainability, auto-executability, intelligence, automation, and intent-based management [82]. This layer is also expected to include terms and conditions in the 6G SLA for imposing a penalty in the event of a minor, major, or critical incident affecting the ordered NS. These incidents should be monitored continuously in real-time with respect to a number of QoS, legal, business, and other metrics [77]. The monitoring mechanisms for the SLA metrics must be designed in a manner that is accessible to both parties and aimed at ensuring proper service configuration, management, and maintenance [83].

Finally, the scope of the Design Layer is limited until the SLA is signed, as shown in Figure 6. It is not responsible for enforcing and monitoring the metrics of the SLA. The Business Management Layer, which will be described in Section IV-H, encompasses the business aspects of the post-provisioning phase of the 6G SLA. Assuming the NEST has been defined or selected from a catalog and the SLA has been signed, the Design Layer then transfers the NEST of the 6G NS to the Management and Orchestration Layer, which will manage its components and orchestrate their resources for the duration of its existence in accordance with the SLA.

C. MANAGEMENT AND ORCHESTRATION LAYER

The Management and Orchestration Layer is a complex layer, which deserves a detailed analysis of the M&O procedures and entities. It automatically manages the managed objects (i.e., NFs and transport links) and intelligently orchestrates virtual and physical resources using managing objects across different 6G domains. M&O procedures can be performed in a centralized and/or decentralized manner using closed-loop automation, intelligent, and zero-touch solutions. Managing objects, also known as management entities or management functions (MFs), can be employed on an individual or group basis. This layer leverages a variety of managing objects to generate a large set of 6G M&O services, each of which represents a distinct capability, such as 6G NS provisioning, trace control, performance assurance and fulfillment, automated SLA monitoring, closed-loop automation, intent control, etc. Based on their scope, they can be classified into two categories: primary capabilities and complementary capabilities. The former category includes a collection of the most essential M&O capabilities, such as profiling NS requirements, SLA fulfillment, delivering desired QoS, etc. The latter category consists of a set of M&O capabilities that can complement the operation of the primary M&O capabilities, such as automation, intelligence, zero-trust security, etc.

To produce these capabilities, this layer can employ the latest standardized M&O entities. These de facto and de jure SDOs are the ETSI ISG NFV, the ETSI ISG zero touch network and service management (ZSM), the 3GPP TSG Service and System Aspects WG 5 (TSG SA5), the TMF, the TIP, the MEF, the IETF, the ONF, the BBF, the IOWN, and the O-RAN Alliance. The harmonization and exploitation of the M&O entities of these SDOs in this layer aimed at producing the above capabilities are depicted in Figure 7.

Depending on the designs of vendor and the requirements of operator, different archetypal flavors at the layer level or at the capability level can be considered. For example, a management entity may be assigned to provide all types of M&O capabilities for all managed objects in a layer, or a management entity may be tasked with providing a M&O capability to all layers, or other combinations. Although the Deployment Perspective is beyond the scope, we exclude consideration of the archetypal flavors and let operator chooses the flavor that best meets its requirements. However, we unify the latest standardized entities and assume that they deliver the most important M&O capabilities and the archetypal flavors. These entities, as well as the M&O capabilities, within the context of this layer, are described in the rest of this subsection.

1) ZERO-TOUCH, FULLY AUTOMATED, AND INTELLIGENT PROCEDURES FOR E2E M&O SERVICES

The foremost responsibility of the Management and Orchestration Layer is to automatically interpret the NESTs (provided by the Design Layer) into NS deployment descriptors. The NESTs interpretation must be performed such that 6G network can comprehend and enforce on the infrastructure to enable the instantiation of 6G NSs. To execute an autonomous and zero-touch interpretation, this layer can leverage a business-oriented and customer-facing management entity specified by the 3GPP TSG SA5 [84], namely the communication service management function (CSMF), as shown in Figure 7. The CSMF, which is specified for 5G slicing and used to translate service requirements into network-understandable requirements, can be upgraded with programmability, intelligence, and closed-loop automation to meet the complexity of 6G slicing [85]. In the interim, the internal structure and segments of the legacy CSMF shall also be substantially reconstructed and optimized to guarantee the correct interpretation of the 6G service requirements into their corresponding NS descriptors.

Before proceeding with a zero-touch interpretation, the CSMF must first receive the 6G NESTs from the Design Layer. Then, it translates the requirements specified in each NEST into a format that is comprehensible by the Management and Orchestration Layer, organizes the interpreted requirements in 6G NS descriptors, and archives NS descriptors for future use. Each descriptor must include information pertaining to physical network functions (PNFs) and VNFs, connectivity between the NFs, flavoring and customization, instantiation levels, the required virtual and physical resources, identification, etc. that characterize a 6G NS from CN to the extreme edge. For this end, this layer can use the NS template (NST), which is defined by the 3GPP TSG SA5 as an E2E descriptor for an NS [86]. Then, the CSMF is anticipated to maintain a comprehensive collection of pre-defined standard and non-standard NSTs in its catalog with the purpose of allowing effortless portability and ease of replication, as shown in Figure 7.

Once the NSTs have been defined or selected from a catalog, this layer initiates admission control, lifecycle management, resource reservation, and other preparation procedures for 6G NSs in an automated and intelligent manner. To that end, this layer can leverage the NS management function (NSMF), a centralized M&O entity specified by the 3GPP TSG SA5 [86], as depicted in Figure 7. Despite the fact that NSMF was integrated into 5G slicing, its utilization in 6G necessitates significant improvements, such as the inclusion of management capabilities of the extreme edge resources and managed objects. In addition, the internal features and structure of the NSMF should be updated to comply with 6G requirements regarding the extreme dynamicity and nomadicity of resources, particularly when considering the extreme edge domain and the characteristics of multi-domain environments in which different parts of a 6G NS should be provisioned and stitched, combining resources from multiple stakeholders. Furthermore, the legacy NSMF must be equipped with novel intelligence capabilities and cutting-edge zero-touch and closed-loop automation tools to confront the emerging E2E M&O challenges, as articulated and architected by the ETSI ISG ZSM [87].

To autonomously execute lifecycle management, the CSMF must first deliver NSTs to NSMF via a machine-consumable interface. The interoperability between CSMF and NSMF shall support a high degree of closed-loop automation, continuity, zero-trust security, and zero-touch synchronization, as depicted in Figure 7 [87]. Utilizing NSTs, the NSMF intelligently manages the 6G NSs and autonomously orchestrates their resources [88]. Upon receiving the NST, the NSMF can employ ML algorithms and leverage several means of automation (e.g., such as those specified by the ETSI ISG ZSM [87]) to examine from a resource, geographical, and temporal point of view and determine whether it is feasible or impractical to deploy the NSs. Thus, a decision as to whether to accept or reject the NSs shall be made within a microsecond interval. Assuming that the NSs have been accepted, the NSMF must extract information from the NSTs related to network and resource requirements without any human intervention. Using this information, the NSMF autonomously prepares the network environment as well as constructs and onboards the E2E NSs that meet the specified 6G service requirements.

To efficiently operate 6G NSs, we believe that a centralized M&O scheme where a single decision-making entity is managing the 6G slicing is challenging and risky. This is due to a large number of 6G NSs, significant overhead and delay, the potential for a single point of failure, and security risks. The centralized M&O approach, regardless of whether it leverages automation and intelligence, will most likely encounter these challenges. Hence, we are persuaded that decentralizing M&O operations and architecture, in which decentralized and federated entities are responsible for the M&O of 6G CN, TN, RAN, and extreme edge domains' resources, is an effective approach. To design such a decentralized management equipped with automation and intelligence, one possible solution that this layer can employ is the deployment of the NS subnet management function (NSSMF) [86], combined with cutting-edge zero-touch and automation capabilities, such as those defined by the ETSI ISG ZSM in [87]. Figure 7 shows such an autonomous architecture, in which the centralized entity (i.e., NSMF) is autonomously interacting with the decentralized management entities (i.e., NSSMFs) by means of zero-touch and closed-loop automation mechanisms.

The NSSMF is a domain-level M&O entity specified by 3GPP TSG SA5 that is tasked with managing a portion of an E2E NS [86]. This portion is called the NS subnet (NSS). Similar to the NSMF, the NSSMF must be upgraded with new internal alterations and capabilities to fulfill the zero-touch and fully autonomous management requirements of 6G slicing. In 5G, the 3GPP TSG SA5

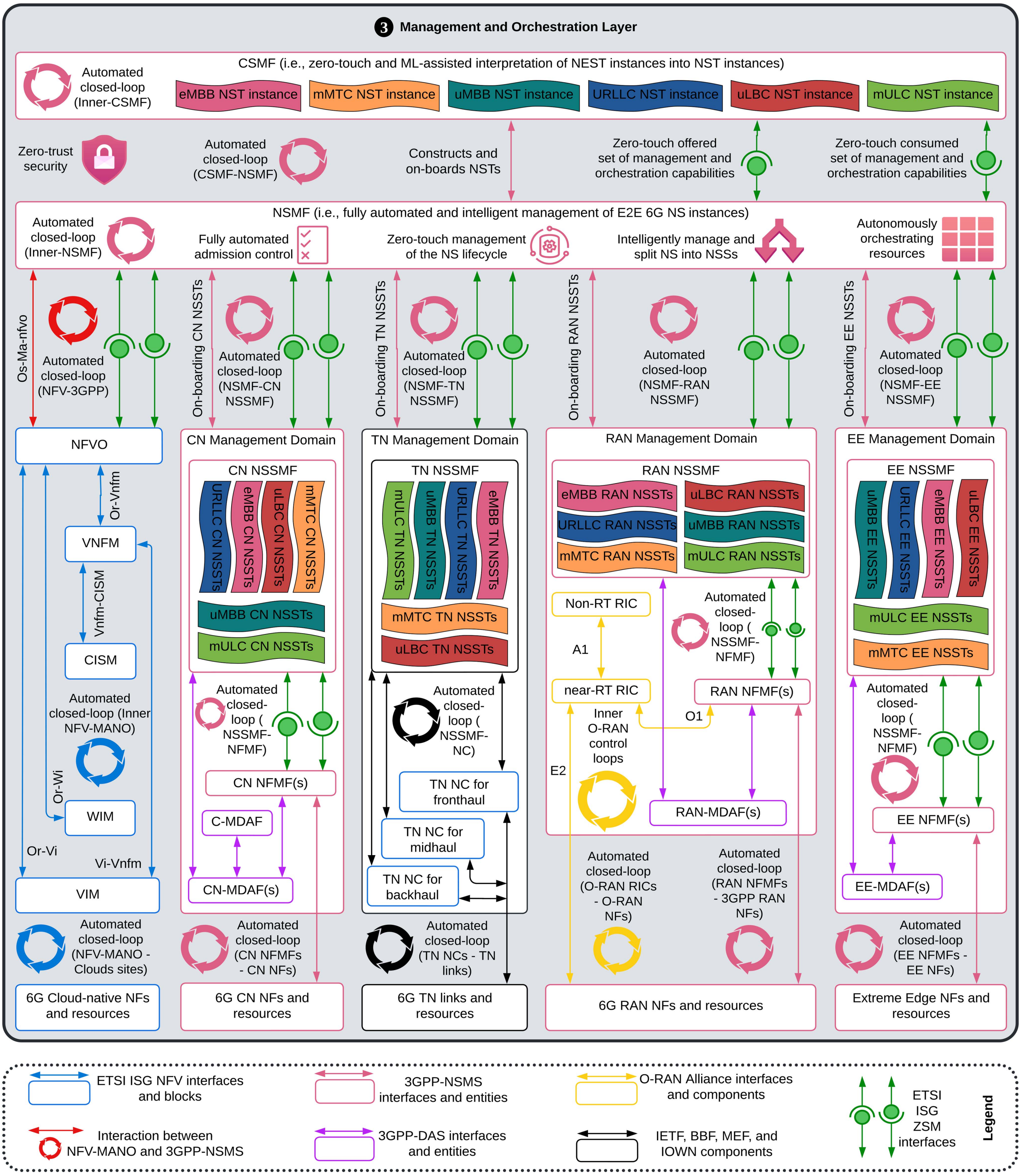

3 Management and Orchestration Layer
CSMF (i.e., zero-touch and ML-assisted interpretation of NEST instances into NST instances)
Automated closed-loop (Inner-CSMF)
eMBB NST instance
mMTC NST instance
uMBB NST instance
URLLC NST instance
uLBC NST instance
mULC NST instance
Zero-trust security
Automated closed-loop (CSMF-NSMF)
Constructs and on-boards NSTs
Zero-touch offered set of management and orchestration capabilities
Zero-touch consumed set of management and orchestration capabilities
NSMF (i.e., fully automated and intelligent management of E2E 6G NS instances)
Automated closed-loop (Inner-NSMF)
Fully automated admission control
Zero-touch management of the NS lifecycle
Intelligently manage and split NS into NSSs
Autonomously orchestrating resources
Os-Ma-nfvo
Automated closed-loop (NFV-3GPP)
On-boarding CN NSSTs
Automated closed-loop (NSMF-CN NSSMF)
On-boarding TN NSSTs
Automated closed-loop (NSMF-TN NSSMF)
On-boarding RAN NSSTs
Automated closed-loop (NSMF-RAN NSSMF)
On-boarding EE NSSTs
Automated closed-loop (NSMF-EE NSSMF)
NFVO
Or-Vnfm
VNFM
Vnfm-CISM
CISM
Automated closed-loop (Inner NFV-MANO)
Or-Wi
WIM
Or-Vi
Vi-Vnfm
VIM
Automated closed-loop (NFV-MANO - Clouds sites)
6G Cloud-native NFs and resources
CN Management Domain
CN NSSMF
URLLC CN NSSTs
eMBB CN NSSTs
uLBC CN NSSTs
mMTC CN NSSTs
uMBB CN NSSTs
mULC CN NSSTs
Automated closed-loop (NSSMF-NFMF)
CN NFMF(s)
C-MDAF
CN-MDAF(s)
Automated closed-loop (CN NFMFs - CN NFs)
6G CN NFs and resources
TN Management Domain
TN NSSMF
mULC TN NSSTs
uMBB TN NSSTs
URLLC TN NSSTs
eMBB TN NSSTs
mMTC TN NSSTs
uLBC TN NSSTs
Automated closed-loop (NSSMF-NC)
TN NC for fronthaul
TN NC for midhaul
TN NC for backhaul
Automated closed-loop (TN NCs - TN links)
6G TN links and resources
RAN Management Domain
RAN NSSMF
eMBB RAN NSSTs
uLBC RAN NSSTs
URLLC RAN NSSTs
uMBB RAN NSSTs
mMTC RAN NSSTs
mULC RAN NSSTs
Non-RT RIC
A1
near-RT RIC
Automated closed-loop (NSSMF-NFMF)
RAN NFMF(s)
O1
E2
Inner O-RAN control loops
RAN-MDAF(s)
Automated closed-loop (O-RAN RICs - O-RAN NFs)
Automated closed-loop (RAN NFMFs - 3GPP RAN NFs)
6G RAN NFs and resources
EE Management Domain
EE NSSMF
uMBB EE NSSTs
URLLC EE NSSTs
eMBB EE NSSTs
uLBC EE NSSTs
mULC EE NSSTs
mMTC EE NSSTs
Automated closed-loop (NSSMF-NFMF)
EE NFMF(s)
EE-MDAF(s)
Automated closed-loop (EE NFMFs - EE NFs)
Extreme Edge NFs and resources
ETSI ISG NFV interfaces and blocks
3GPP-NSMS interfaces and entities
O-RAN Alliance interfaces and components
ETSI ISG ZSM interfaces
Interaction between NFV-MANO and 3GPP-NSMS
3GPP-DAS interfaces and entities
IETF, BBF, MEF, and IOWN components
Legend

FIGURE 7. The detailed architecture of the Management and Orchestration Layer. Note that EE = extreme edge.

proposed three NSSMFs for CN, TN, and RAN. Each controls its corresponding NSS [86]. In 6G slicing, there will be resources at the extreme edge (e.g., resources belonging to the vertical industries, public service sectors, end-users, etc.) and resources from third-party facilities (e.g., in-factory networks, transportation hubs, hyperscaler cloud nodes, etc.). Therefore, in addition to the three NSSs, as shown in Figure 7, we propose an NSSMF-like MF at the extreme edge to manage the extreme edge portion of a 6G NS and its resources. As of this writing, the NSSMF-like MF for the extreme edge management domain has neither been proposed nor standardized. We believe that its incorporation into the 6G management system aimed at efficient resource management in this management domain is critical for the realization of an E2E 6G NS.

For efficient slice descriptor management, the NST can also be divided into several parts, each called the NS subnet template (NSST) [86]. The NSST can be assigned to an NSS. In addition to the three conventional NSSTs, we propose a NSST-like descriptor, as illustrated in Figure 7, for the extreme edge management domain. Once such a decentralized and hierarchical M&O framework has been constructed, the NSMF shall pass the NSSTs to the CN, TN, RAN, and extreme edge NSSMFs. Moreover, the NSMF can act as a Cross-domain Integration Fabric, as defined by the ETSI ISG ZSM [89], to enable interoperation between the NSSMFs, employing various means of automation such as policy-driven, intent-based, and ML-based [87], [90]. Each NSSMF employs a NSST to manage its resources for instantiating its NSS and reports to the NSMF all activities relevant to that particular NSS. Furthermore, each NSSMF is accountable for (a) lifecycle management, (b) fault, configuration, accounting, performance, and security (FCAPS), and (c) all operations associated with the NSS it manages. More details on the scope of the four NSSMFs are provided below.

2) INTELLIGENT M&O PROCEDURES FOR THE 6G CN

The 6G CN is envisaged to be more open, modular, cloud-native, slicing-aware, and scalable than 5G CN. Hence, the M&O of 6G CN NSSs will also be more intelligent, zero-touch, and intent-based to accommodate the unprecedented levels of heterogeneity and complexity. To design an autonomous CN management domain, we believe that further decentralization of 6G slicing would bring numerous advantages, such as improved security, quicker NSS design and provisioning times, optimized resource management, etc. The initial step towards this goal would be to establish NF-level or sub- and micro NF-level MFs. To that end, this layer can employ network function management function (NFMF), specified by 3GPP TSG SA5 [86], as a management entity of NF(s). Depending on the deployment flavors, the NFMF is expected to autonomously execute the M&O of a particular type or all NFs of an NSS employing network function descriptors (NFDs). Once the NFMFs have been designed, the CN NSSMF operates as a Domain Integration Fabric, as specified by the ETSI ISG ZSM [87], to permit interaction with the NSMF and inter-operation between the CN NFMFs by means of closed-loop automation, as shown in Figure 7.

We are convinced that 6G CN will be fully virtualized, cloud-native, and service-based architected. Hence, the NFs of a 6G CN NSS would be delivered as cloud-native network functions (CNFs) and/or VNFs. The CN domain in 5G consists of a variety of NFs, such as access and mobility, policy control, NS selection, etc. There may also exist a number of CN NFs in 6G that will be AI-driven, cloud-native, and slicing-aware, which will probably be specified by the 3GPP. Such NFs will require virtual and cloud resources. Since the M&O of VNFs and CNFs fall within the scope of ETSI NFV ISG, we anticipate that NFV ISG will continue to play an active role in moving forward the standardization efforts towards the 6G. This can be accomplished in 6G by either introducing new FBs or optimizing existing ones in the NFV-MANO framework.

To accomplish this, the interaction between NFV orchestrator (NFVO) and NSMF (or NSSMF) through closed-loop automation is a prerequisite, as shown in Figure 7. Once the NFVO receives information from the NSMF regarding the instantiation of the VNFs, it creates and manages the VNF descriptors (VNFDs) for the duration of the CN NSS. The NFVO then instructs the VNF manager (VNFM) to interact with each NFMF to manage performance and fault-related tasks. The VNFM also manages the lifecycle of VNFs. The NFVO requests the virtualized infrastructure manager (VIM) to configure and allocate the virtual compute and storage resources for VNFs. This layer also considers container infrastructure service management (CISM) when maintaining containerized workloads [91]. Regardless of whether containerization or virtualization technologies are employed, the VIM and CISM can use closed-loop automation to manage, orchestrate, and scale the managed resources of the 6G CN NSS. In 6G slicing, the legacy or any FBs that may be introduced in the future, could also be equipped with closed-loop automation or other autonomous capabilities to improve the performance of the interoperability of the NFV-MANO FBs.

To integrate intelligence into the CN management domain, this layer can employ management data and analytics function (MDAF) [92], which is defined by 3GPP TSG SA5 as a management data analytics service (MDAS) to automate and intelligently enhance M&O in 5G and beyond. The MDAF utilizes management data gathered from 3GPP-NSMS to generate centralized and/or domain-specific analytics. At the 6G CN management domain, the CN-MDAF gathers management data from CN NFMFs and NSSMF (or futuristic MFs), analyzes this data, and makes recommendations or predictions. Using closed-loop automation, the CN NFMFs and NSSMF receive a portion of NFs management data from the NSMF and part of data related to resource management from the managed NFs and resources. Finally, the CN-MDAF is linked to the centralized-MDAF (C-MDAF), which collects data from the domain-level MDAF (i.e., CN MDAF) and delivers centralized analytics to the management domain (i.e., CN management domain), as shown in Figure 7.

3) INTELLIGENT M&O PROCEDURES FOR 6G THE RAN

To design an autonomous M&O platform for 6G RAN slicing, it is vital to determine virtual and physical NFs of a RAN NSS. The majority of latest studies [9] and specifications [93] partition the radio protocol stack into three units: the centralized unit (CU), the distributed unit (DU), and the radio unit (RU). They consider CU and DU as VNFs and RU as a PNF. As of this writing, the CU is fully virtualized;

however, the DU requires further research and is likely to be provided in virtualized form [9], [94]. Due to increasing interest in extending virtualization and cloudification towards the edge, we foresee that, certain RU functionalities may also be virtualized in long term. However, the complete virtualization of RU is nearly unfeasible due to the impossibility of virtualizing some physical layer components, such as receiving and transmitting antennas. Therefore, it is anticipated that the 6G RAN slicing will rely heavily (but not fully) on intelligent virtualization, softwarization, and cloudification.

Based on the above projection, this layer can assign the MFs specified by 3GPP TSG SA5 and ETSI ISG NFV to conduct M&O at the domain- and NF-levels of slicing the 6G RAN, as shown in Figure 7. The RAN NSSMF manages managing- and managed-objects in the 6G RAN management domain by leveraging the RAN NSST. The NFMFs can be tasked by NSSMF to manage the NFs of a RAN NSS. Notably, the NFMF directly manages the PNF of a RAN NSS by exploiting the PNF descriptor (PNFD) and the VNF through the utilization of the NFV-MANO FBs by leveraging the VNFD [9]. The M&O of a VNF and its virtual resources of a RAN NSS are performed in a similar fashion by the NFV-MANO FBs that manage the VNFs of the CN NSS. Likewise, the interaction between the RAN NSSMF and the NFVO can be established either directly or through the NSMF. The interactions among RAN MFs shall support closed-loop automation. The RAN NSSMF, NFMFs, NFV-MANO FBs, RAN NSST, VNFDs, and PNFDs shall also be upgraded with novel capabilities to meet the requirements of a virtualized, cloudified, programmable, and slicing-aware 6G RAN [95].

The second capability that influences the 6G RAN slicing management is the adaptation of open interfaces and open software and hardware. The O-RAN Alliance and the OpenRAN Project Group of TIP are the two industry consortiums responsible for standardizing openness in RAN. The focus of TIP is on legacy RANs and their integration by leveraging open interfaces [96]. However, the O-RAN Alliance has been focused on 5G and beyond RAN [93]. Hence, among other factors, the specifications of the O-RAN Alliance have garnered considerable interest. On the basis of its current accomplishments, it is believed that the O-RAN Alliance will continue to play its role in advancing the openness of 6G RAN and making it completely open, disaggregated, data-driven, and intelligent. Consequently, we utilize the O-RAN Alliance's SMO framework [93] to manage and orchestrate the open components of 6G RAN. The SMO is an intelligent platform for the managing- and managed-objects of the Open RAN. It can be viewed as Open RAN‘s Domain Fabric within the context of the ETSI ISG ZSM. The latest SMO has been confronted with three issues relating to interoperability, zero-trust security, and performance. Therefore, additional research is required to develop intelligent solutions for these challenges and equip the future edition of the SMO with novel capabilities suitable for the M&O of open 6G RAN slicing.

To manage managed objects and configure management capabilities in 6G RAN management domain, the SMO framework uses the near-real-time radio intelligent controller (Near-RT RIC) and non-real-time radio intelligent controller (Non-RT RIC) in a centralized location and closed to the user equipments (UEs), respectively [93]. Using closed-loop automation, the Near-RT RIC provides M&O capabilities with an automation loop lasting less than one-second [97], whereas the Non-RT RIC has an automation loop lasting longer than one-second [98]. The SMO framework can also interact with the slice MFs to enable Open RAN slicing. Depending on the deployment flavors, the RAN NSSMF and NFMF can be placed within or beyond the SMO framework [99]. Lastly, we anticipate that managed objects may not always be fully open in the 6G RAN domain. There might be a scenario in which Open- and Closed-RANs coexist. Hence, the SMO capabilities should be extended to the point where they can autonomously support the M&O of hybrid managed objects in a complex multi-vendor and multi-technology 6G RAN.

The third capability is the integration of intelligence and automation into the 6G RAN management domain. To enable them, the framework can utilize ZSM and MDAF. Figure 7 illustrates that the ZSM-enabled interoperability is facilitated between the RAN NSSMF and NFMF by means of closed-loop automation. It enables zero-touch operation and interaction between the two MFs, as well as their interoperability with the NSMF. The RAN MDAF is capable of interacting with the NSSMF and NFMF [92]. By leveraging current and historical data pertaining to RAN M&O, network events and status, performance measurements, etc., the MDAF generates statistics, recommendations, and predictions for the management of 6G RAN NSSs. The output of the MDAF can be used by the RAN management domain to enable autonomous root cause analysis, intelligent courses of action, zero-touch maintenance, automated scaling, etc. The RAN MDAF is a part of the E2E MDAS. Therefore, it shall also be connected via a standardized interface to the C-MDAF to enable a complete analytical view of the 6G network management [100].

4) INTELLIGENT M&O PROCEDURES FOR THE 6G THE TN

The 6G TN management domain manages the transport links and orchestrates their networking resources to enable the 6G slicing framework with intelligent, open, flexible, programmable, time-sensitive, mission-critical, and cost-efficient transportation infrastructure. In 5G slicing, each TN NSS typically consists of backhaul, midhaul, and fronthaul [101]. This three-link composed topology has also recently been considered in the specification of the MEF [102], O-RAN Alliance [103], and TIP [104] for designing an E2E TN NSS. We anticipate that this topology will continue to exist for 6G TN NSS. Based on this, the TN domain can apply automation and intelligence to the three links aimed at ensuring E2E QoS for numerous

KPIs, including latency, throughput, reliability, resiliency, etc. This management domain must provide autonomous, dynamic, and functional split-aware interconnection between the physical, virtual, and cloud-native NFs that may exist in the CN and RAN NSSs. To enable autonomous, zero-touch, and intelligent TN networks, this domain shall leverage transport-aware domain- and link-level slicing management entities that can manage wired and wireless transport slicing.

The TN management domain can also use NSSMF. Beyond this, it cannot leverage the 3GPP TSG SA5 MFs and MDAS. Because the M&O, as well as the intelligentization, of a TN NSS fall beyond the scope of the 3GPP [9]. However, both the NSMF and TN NSSMF must autonomously interact via a standardized interface using closed-loop automation to realize an E2E 6G NS, as shown in Figure 7. The TN NSSMF is conceptually equivalent to (a) the TN Domain Integration Fabric of the ETSI ISG ZSM [87]; (b) the NS Controller of the IETF Traffic Engineering and Architecture Signalling working group (WG) [105]; (c) the transport-NSSMF (T-NSSMF) of the BBF [106]; and (d) the SDN-based Domain Controller of the ONF [107]. The objective of the TN NSSMF is to manage the E2E inter-working between the 6G RAN and CN and orchestrate the networking resources over the transport links. It receives the TN NSST from NSMF, in which the resource and lifecycle requirements are defined to be managed and orchestrated by the 6G TN management domain. Once the NSMF delivers the appropriate NSST to the TN NSSMF, it has no control over the TN NSS.

For link-level management, the TN management domain can use the wide area networks infrastructure manager (WIM) and network controller (NC), which are defined by the ETSI ISG NFV [108]. The WIM and the NC are conceptually equivalent to the Transport Slice Orchestrator and the Transport Network Controller specified by the IETF in [109], respectively. The WIM manages the multi-site connectivity of TN NSSs over the backhaul, midhaul, and fronthaul. The NC manages either the E2E transport links of a single TN NSS or a single type of transport link shared by multiple TN NSSs. The WIM and NC establish a hierarchy in which the NC controls the underlying resources directly and reports to the WIM. The NC uses physical link descriptors (PLDs) and virtual link descriptors (VLDs) to deploy and manage the physical and virtual links (VLs), as well as their corresponding networking resources, of a TN NSS, respectively. Both the PLD and VLD are derived from the TN NSST and provided to the WIM for the management of the networking resources of a 6G TN NSS. We anticipate that the descriptors and MFs of the TN management domain will require significant modifications to fulfill 6G transport slicing requirements.

To adopt openness and integrate intelligence into the transport management domain, WG 9 of the O-RAN Alliance has also utilized the management entities of some of the said SDOs and harmonized them to design an open and intelligent management domain for TN NSSs [103]. The O-RAN Alliance has not yet specified how these entities would be exploited to manage the transport domain inside the Open RAN. Future releases of WG 9 will study this topic in more detail for beyond 5G and 6G slicing frameworks. Nonetheless, we are convinced that the architecture of the futuristic TN management domain of the O-RAN Alliance would closely resemble our proposal depicted in Figure 7. Finally, the IOWN Global Forum introduced an E2E Transport Orchestrator [110], which can also be leveraged by the transport management domain of our architecture to manage the 6G TN NSSs over future wireless and optical transport infrastructure. The IOWN Global Forum is currently standardizing the most advanced transport technologies for the 6G. The slicing aspects have not yet been exhaustively studied by this forum during the course of our research. However, it is anticipated that the IOWN Global Forum will extend the proposed orchestrator to the M&O of 6G transport slicing.

5) INTELLIGENT M&O PROCEDURES FOR THE 6G EXTREME EDGE

In 6G slicing, there will be managed resources and objects that are beyond the public network domain and will be located close to end-users or within the premises of industries and enterprises. For example, managed objects and resources exist within industrial enterprises, intelligent agricultural farms, and automobile factories. This domain is called the “extreme edge domain,” “far edge domain,” or “beyond the edge domain” [16]. It targets new business opportunities, enhances end-user and tenant experiences, provides tenants with a set of management responsibilities, and ensures the privacy and security of tenant data for those who do not desire to share their data or information on public networks. The extreme edge managed objects will be hosted by extreme edge clouds and orchestrated by the extreme edge management domain. The resources located in this domain may exhibit a significant degree of asynchrony, necessitating specialized M&O procedures. The integration of the extreme edge management domain into the public network management domain is a challenge of paramount importance due to the large quantity and heterogeneity of devices and managed resources. The extreme edge management domain might be managed by the service provider, the owner of an industry, or jointly by the service provider and the tenant. In the case where the tenant has control over its M&O, it is crucial that the service provider and tenant define the management level so that both parties understand their responsibilities regarding M&O of 6G NSs.

The management level of the extreme edge domain can be defined such that the tenant has either complete or limited control over managing- and managed-resources. To reach a win-win agreement over the M&O of the extreme edge domain, we propose the concept of management level agreement (MLA). The MLA could be a separate, legally enforceable contract signed by both parties, or it could be an essential element of the SLA. It must define a full or partial set of capabilities, services, and features for the M&O of the extreme edge domain. The MLA grants a tenant the required autonomy and the distinguishing feature of developing, deploying, and orchestrating its own applications, services, resources, and policies related to a 6G NS through negotiation and enforcement with the service provider. During negotiation, the tenant may not be permitted to undertake certain M&O procedures due to MLA constraints. The negotiation between over the MLA, including the access rights and levels, the degree (e.g., zero, partial, or full) of autonomy, and M&O capabilities, can be defined and performed via a standardized interface.

No SDO has thus far specified a management system for the M&O of the extreme edge domain. To address this challenge, there may be two possibilities. First, the SDOs focusing on the standardization of cloud and edge computing may take the initiative to propose and standardize this management domain. Second, the SDOs related to telecommunications may extend their legacy M&O frameworks towards the extreme edge management domain. In either scenario, the extreme edge management domain must interoperate with the E2E M&O entity (i.e., NSMF). This management domain shall utilize innovative MFs to manage the asynchronous and error-prone nature of end-user devices and objects. For the purpose of completeness in our framework, we assume the harmonized architecture of the 3GPP TSG SA5 and ETSI ISG NFV as the management system and intelligentization framework of the extreme edge resources. Hence, the extreme edge management domain can leverage an NSSMF-like MF for the M&O of this domain. It can also make use of the NFMF-like MF for the M&O of the PNFs and VNFs.

The extreme edge NSSMF shall interact with the NSMF via a northbound interface and with the NFMF via a southbound interface. With the integration of the extreme edge management domain into the E2E M&O domain, the so-called “device-edge-cloud continuum M&O” concept can be accomplished in the proposed framework [16]. In the device-edge-cloud continuum M&O, the Management and Orchestration Layer can provide continuous orchestration from devices (in the extreme edge domain) through the edge data centers and up to the centralized data centers, considering all tiered data centers and network resources in between. This E2E M&O system should be unified, federated, decentralized, intelligent, open, and self-sufficient. And the NSMF shall play the role of a central M&O function that has complete knowledge of topology and available resources in all four domains. Finally, the adoption of ML algorithms can significantly enhance the efficiency of extreme-edge M&O strategies, as it can help to automatically discover patterns in the dynamic and variable trends of managing- and managed-resources availability. Since the interoperability between the MDAS, 3GPP-NSMS, and NFV-MANO has been discussed previously, we skip such a discussion and assume that a similar interaction applies to the intelligentization of the extreme edge management domain.

D. NETWORK SLICE LAYER

The Network Slice Layer defines and archives the logical components of an E2E 6G NS. It is situated between the Management and Orchestration Layer and the Infrastructure Layer. This is due to the requirement that the logical components of a 6G NS can be directly managed by the Management and Orchestration Layer and hosted by the Infrastructure Layer. The logical components primarily consist of logical NFs and logical transport links connecting them to realize a complete, customized, and isolated logical 6G network. The logical components can be provided in virtual, cloud-native, or physical forms, and they can be deployed at various points across the underlying infrastructure. They can be temporarily or permanently deactivated or activated based on the topological requirements of a 6G NS. Due to their requirements, some logical components may require fewer resources for certain 6G NSs, whereas the same logical components may require more resources for other types of 6G NSs. Furthermore, the logical components can be deployed in shared, dedicated, or hybrid fashions. Hence, this layer shall define the logical components that can be shared by 6G NSs and those that can be deployed in dedicated or hybrid manners. Additionally, the logical components in each 6G NS are subject to a certain traffic originating, transmitting, receiving, and terminating chain, as well as a customized network topology, which must be defined, exploited, and optimized by this layer.

To utilize a standardized entity for sequentially chaining the logical components of a 6G NS, this layer can use the service function chain (SFC), which is specified by the IETF WG SFC [111]. The SFC is deployed in 5G slicing [112]. We believe that, due to its capabilities and implications on service delivery, it will continue to exist in 6G. Nevertheless, the SFC will also require application-level and architectural-level optimization to fulfill the chaining and sequencing requirements of the logical components of a 6G NS. In our framework, a SFC can specify an ordered set and connected sequence of available NFs and transport links. It can also determine the ordering constraints that must be imposed on network traffic of logical components (i.e., packets, frames, and/or flows) in accordance with the requirements of a 6G NS in the upstream and downstream directions. The Network Slice Layer can employ SFC for each of the instantiated 6G NSs to define the behavior, assist in the development of the business model, and enable the customization of the required and contracted service delivery. The SFCs can be designed and archived for future usage in the Network Slice Layer repository.

The design of a SFC for a 6G NS shall be in compliance with the deployment guidelines and architectural principles of the IETF, specified in [111] and [113], respectively. During the preparation and provisioning of a 6G NS, the SFC can be selected from the repository and utilized for the corresponding NS. However, during the lifetime of a 6G NS, a static SFC deployment model may not be able to fulfill the requirements of an NS in a environment that is subject to frequent changes. Additionally, deploying a static SFC, which is tightly coupled to the underlying network topology and physical resources, reduces (or, in some cases, eliminates) the ability of a service provider to introduce new logical components to the running SFC, requires a predefined network topology, and is not suited for complex network configurations [114], [115]. In order to overcome these challenges, the SFC must support dynamic behaviors and actual network conditions during the operation of a 6G NS by adding, removing, reprogramming, and modifying NFs, transport links, and service nodes in a flexible, optimal, and autonomous manner [113].

Each SFC can stretch over numerous domains. The NFs and transport links in each domain may be delivered in physical, virtual, or cloud-native forms. To define an E2E SFC for a 6G NS, there are two design and deployment options.

- First, the Network Slice Layer can define a unified SFC that can be applicable to all network domains, NFs, and transport links. Since the standardization of different domains falls under the purview of separate SDOs, the cooperation of the relevant SDOs is a prerequisite for the design of a complete 6G SFC.
- Second, this layer can design domain-or technology-specific micro SFCs and merge them to produce an E2E SFC. For example, four micro SFCs, each for a domain, can be designed and merged. To accomplish this, the Network Slice Layer can employ the VNF forwarding graph (VNFFG), specified by the ETSI ISG NFV in [116], to design the order and ordering constraints of the cloud-native NFs and VNFs at 6G CN and RAN domains. For the transport domain, this layer can leverage the network forwarding path (NFP) to define the path(s) taken by actual traffic. For the extreme edge domain NFs chaining, we presume that an SDO will define a standardized entity in the future. Finally, these four functions and pathways chaining entities can be combined to construct a complete SFC for an E2E 6G NS.

Each option has its own advantages and challenges. Because the Deployment Perspective is beyond the scope, we let the network operator and tenant decide which of the proposed SFC design and deployment methods best meet their requirements.

To define the components of a SFC for an E2E 6G NS, we separately look into the details of the logical NFs and logical transport interfaces the four subnets as we defined them in the Management and Orchestration Layer.

- Figure 8 illustrates that the CN NSS consists of several NFs, including legacy NFs such as mobility management, NS selection, policy control, etc., as well as novel NFs that might be introduced in the future to augment the capabilities of 6G CN slicing. Since several decades, the 3GPP TSG Service and System Aspects WG 2 (TSG SA2) and 3GPP Core Network and Terminals WG 4 (CT WG4) have played a significant role in the standardization of NFs, architectures, and protocols for the CN domain [117], [118]. We anticipate that both of them will standardize slicing-aware 6G CN NFs, architectures, and function chains to define various 6G NSSs.
- Figure 8 also depicts that a RAN NSS consists of several logical NFs, such as radio resource control, medium access control, radio frequency, etc. Based on existing functional split options, the legacy RAN NSS NFs can be divided into three components, namely the CU, DU, and RU, with each component containing a subset of the above NFs. Traditionally, the 3GPP TSG SAs and the 3GPP TSG RANs have been standardizing the RAN NSS-related aspects [117], [119]. We believe that both of these TSGs, along with O-RAN Alliance, TIP, and SCF, will extend the standardization of RAN slicing towards 6G by introducing novel radio NFs, split options, protocols, etc. for the future design of 6G RAN NSSs.
- As depicted in Figure 8, the TN NSS can contain up to three interfaces: NG, F1, and Fx. In some cases, this NSS can also include the logical interface between the access network and the end-user, the air interface. The SDOs related to the TN domain, which were mentioned previously, will continue to play a critical role in standardizing the logical interfaces as well as introducing their capabilities for the 6G transport slicing.
- Finally, the extreme edge NSS may also include several logical components. This NSS has not yet been standardized. However, its standardization is currently a hot topic in numerous 6G projects and discussions around the world [16]. We anticipate that this NSS will also consist of multiple logical NFs and be deployed sequentially.

To propose a baseline architecture for the Network Slice Layer, we consider the latest specifications of the relevant SDOs. We anticipate that the state-of-the-art logical components of an NS will be upgraded and that several novel logical components will be introduced in 6G slicing that can be integrated into this layer. Some logical components may be deployed in the CN, while others may be moved to the edge or even the extreme edge to satisfy the extreme requirements (described in Table 2 and beyond) of various 6G NSSs. For example, the Network Slice Layer shall relocate (or design) the NFs (which are traditionally deployed in the CN) to the edge of a 6G network, aimed at satisfying the performance requirements of the URLLC devices. In addition, this layer should be capable of scaling any future NFs and providing logical connectivity among them for any network domain within an NS. This layer should also be

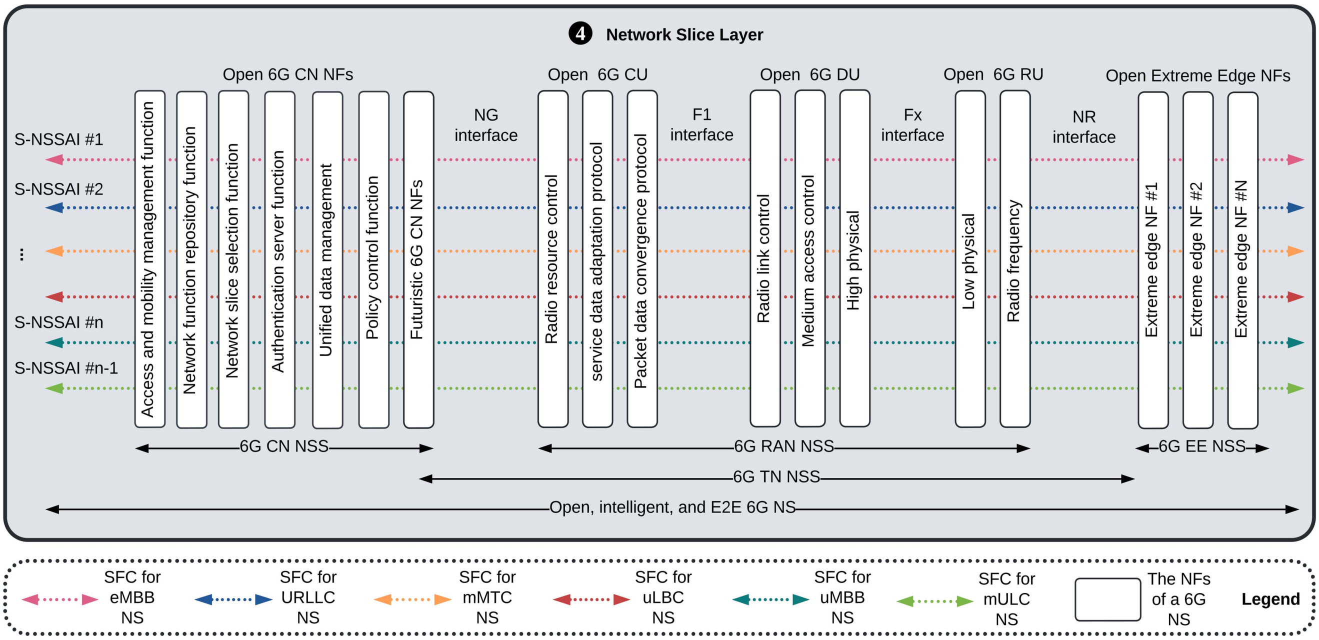


FIGURE 8. The Network Slice Layer architecture of the proposed framework.

able to expose the design of logical components, enabling tenants and developers to access, create, and modify new logical components, as well as innovate to meet the evolving requirements of vertical industries and use cases. These features and capabilities must be clearly stated in the descriptors of the relevant SFC of a 6G NS. The SFC must also define the optimal placement of NFs and transport links onto the underlying physical components with respect to the requirements of a 6G NS. For example, latency-sensitive and bandwidth-devouring NFs can be deployed close to the end-user devices, whereas latency-tolerant and low-bandwidth NFs can be placed in a core or a centralized data center [120]. It is worth noting that the legacy SFCs, VNFFGs, NFPs, and their respective templates must also be upgraded to fulfill the capabilities that the Network Slice Layer shall support.

Finally, each 6G NS must be uniquely identified by an identifier across all network domains. The identifier must be carried with the network traffic to enable isolation and help the Network Slice Layer in differentiating the network traffic of a 6G NS. To that end, this layer uses the 3GPP-specified NS selection assistance information (NSSAI) [121], where each 6G NS can be allocated with its single NS selection assistance information (S-NSSAI), as shown in Figure 8. Each S-NSSAI is provided with standardized optional and mandatory identification information, called the slice differentiator (SD) and slice/service type (SST), respectively. The 3GPP has specified that the SST can provide up to 2^8 values and the SD can provide up to 2^{24} values [9]. It can be foreseen that the number of standardized and operator-specific NSs in 5G slicing framework is lower than the number of NSs expected to be defined in 6G. Therefore, to provide a unique S-NSSAI for each 6G NS in 6G framework, the length of the SST and SD can be increased by the 3GPP.

E. INFRASTRUCTURE LAYER

The Infrastructure Layer represents the virtual, cloud-native, and physical resources that host the logical components of 6G NSs. It is responsible for configuring, allocating, optimizing, and terminating the resources of a 6G NS across virtualized, cloudified, and physical network sites and transport links in an intelligent and optimal manner. To deploy an E2E 6G NS, this layer shall enable the internal connectivity between the infrastructural resources of an operator, as well as their connectivity with third-party resources and external systems. The configuration and allocation of the underlying 6G processing, storage, and connectivity resources can be performed (in a centralized or distributed manner) by an infrastructure manager. The infrastructure manager can use ML to reduce energy consumption, enhance performance, improve QoS, etc. In addition to the aforementioned resources, this layer can consist of electrical power resources and systems that must be intelligently planned, exploited, and maintained with the goal of providing energy-efficient and sustainable services in 6G.

Figure 9 shows that Infrastructure Layer is composed of two major parts: Public Network Domain and Vertical Premises.

- The Public Network Domain part is a public 6G network that provides 6G services to tenants by utilizing cutting-edge infrastructure, technologies, resources, and solutions. In the Public Network Domain, this layer includes network data processing sites, transport links, and storage resources that can be deployed across

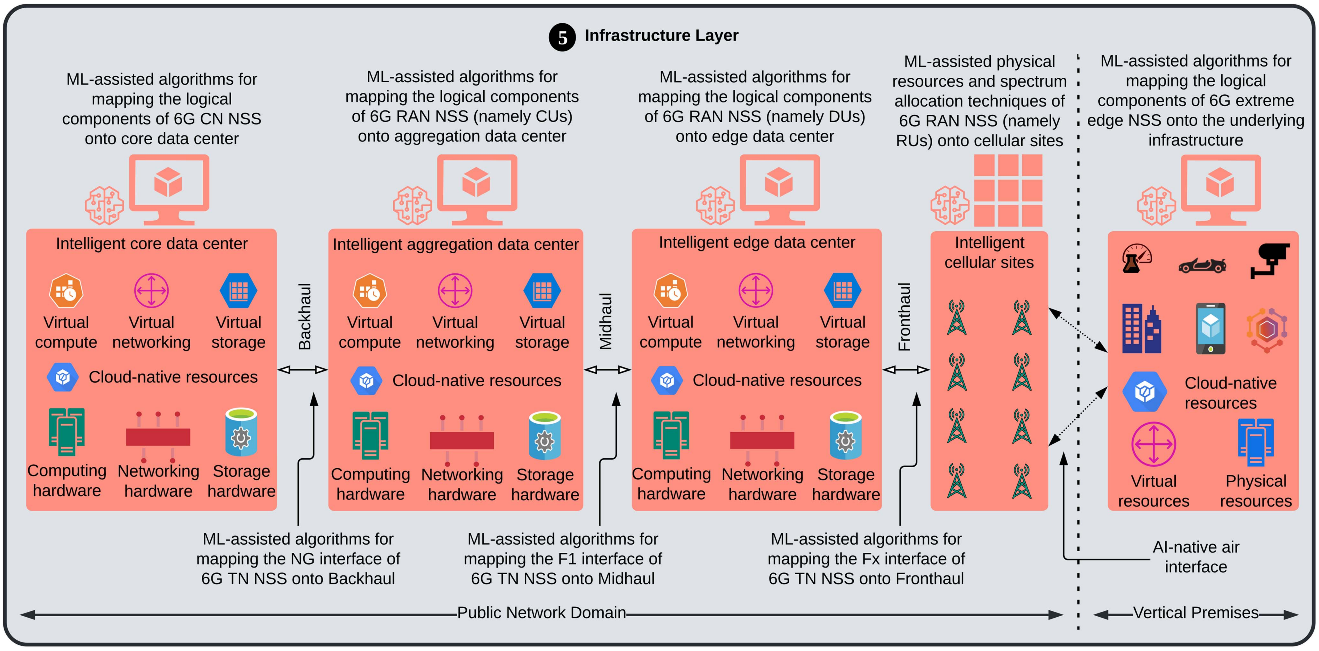

5 Infrastructure Layer
ML-assisted algorithms for mapping the logical components of 6G CN NSS onto core data center
ML-assisted algorithms for mapping the logical components of 6G RAN NSS (namely CUs) onto aggregation data center
ML-assisted algorithms for mapping the logical components of 6G RAN NSS (namely DUs) onto edge data center
ML-assisted physical resources and spectrum allocation techniques of 6G RAN NSS (namely RUs) onto cellular sites
ML-assisted algorithms for mapping the logical components of 6G extreme edge NSS onto the underlying infrastructure
Intelligent core data center
Intelligent aggregation data center
Intelligent edge data center
Intelligent cellular sites
Virtual compute
Virtual networking
Virtual storage
Cloud-native resources
Computing hardware
Networking hardware
Storage hardware
Backhaul
Midhaul
Fronthaul
Virtual resources
Physical resources
ML-assisted algorithms for mapping the NG interface of 6G TN NSS onto Backhaul
ML-assisted algorithms for mapping the F1 interface of 6G TN NSS onto Midhaul
ML-assisted algorithms for mapping the Fx interface of 6G TN NSS onto Fronthaul
AI-native air interface
Public Network Domain
Vertical Premises

FIGURE 9. The architecture of the Infrastructure Layer of the proposed open and intelligent framework for slicing in 6G.

the Infrastructure Layer using a variety of legacy and standardized topologies, such as bus, tree, mesh, star, hybrid, etc. Taking a tree topology into consideration, the Public Network Domain consists of intelligent core data center, aggregation data center, edge data center, and cellular sites. The intelligent data centers and cellular sites are interconnected through backhaul, midhaul, and fronthaul transport links, respectively (see Figure 9). The data centers host the VNFs and cloudified NFs of the CN NSS and RAN NSS, while the cellular sites host the PNFs of a RAN NSS. This hierarchical topology has been widely adopted in the literature due to its simplicity, modularity, flexibility, and easy manageability [9]. However, it poses a number of challenges in terms of resilience, dependency, and scalability. Therefore, the hierarchical topology may serve as a benchmark for evaluating alternative topologies for this layer, such as recursive, random, hybrid, etc. The underlying transport links intelligently host the logical TN interfaces of a 6G TN NSS.

- The Vertical Premises part is a private (i.e., non-public) network domain that can be owned by the owner of an industry and is used to host specific subnets of 6G NSs by employing cutting-edge private communication and industry-specific resources and technologies. In the Vertical Premises, this layer is composed of tenant-owned processing, storage, and networking resources, in addition to end-user and IoT devices. Similar to the Public Network Domain, these resources can be delivered and deployed in virtual, cloud-native, and physical forms.

In the rest of this subsection, we will examine in depth the components and connectivity of both parts, respectively.

1) PUBLIC NETWORK DOMAIN

Figure 9 depicts that the Public Network Domain consists of three intelligent data centers: core data center, aggregation data center, and edge data center. These cloud data sites are distributed geographically across the central, regional, and edge locations of a 6G network, respectively [9]. Each data center consists of cloud-native resources, virtual resources, general-purpose hardware, and software components. The cloud site manager can utilize these resources to host, manage, and execute the cloudified and virtualized functionalities of 6G NSs. The manager of each data center employs ML algorithms to manage, orchestrate, allocate, and optimize the cloudified, virtualized, and physical resources. Utilizing intelligent softwarization, cloudification, and virtualization technologies, the physical resources of a cloud data center can be abstracted and partitioned into virtualized and cloudified environments, such as VMs, containers, unikernels, or other ultra-lightweight and single-purpose cloud and virtual environments that might be introduced in the future. The cloud site manager in a data center employs intelligent algorithms to map the virtual and cloud-native NFs of an NS onto the aforementioned cloudified and virtualized environments.

There are several SDOs that are involved in the standardization of various aspects of cloud and edge computing data centers. In this article, we focus on the SDOs that are related to the telecom industry, specifically those standardizing slicing-aware cloud and edge computing data centers for mobile networks. Based on this, we observed that the ETSI ISG NFV [122], ETSI ISG multi-access edge computing (MEC) [123], and O-RAN Alliance WG6 [124] are currently the most influential bodies and believed to play a significant role in the standardization of 6G cloudified and virtualized infrastructure in the future. In these SDOs, the notions NFV infrastructure (NFVI)-point of presence (PoP), MEC Application Host, and Open-Cloud are used as synonyms for the intelligent cloud data center, respectively. Each of these SDOs has specific interests and scope in the standardization of the underlying infrastructure of a specific domain. In the following, we elaborate on each intelligent cloud data center and how the relevant specifications of the above SDOs can be employed to design a 6G-enabled, slicing-aware, open, and intelligent cloud infrastructure for 6G slicing ecosystems.

The intelligent core data center is located in the CN domain, as illustrated in Figure 9. It hosts the logical components of a 6G CN NSS. These components, which can be sequentially chained together to form a complete CN NSS, are mapped onto their respective virtualized or cloud-native environments. The goal of the mapping process is to deploy the components of a 6G CN NSS onto suitable virtualized or cloudified environments that are capable of fulfilling their cloud-native and virtual resource requirements. To map a 6G CN NSS, the infrastructure manager can employ ML algorithms, improving the configuration and allocation of the underlying resources of the intelligent core data center. To guarantee resource isolation among the 6G CN NSSs, each cloudified or virtualized environment can be limited to hosting a single type of CN NF. For instance, a lightweight VM hosting the access and mobility MF can be configured exclusively to deploy this particular CN NF throughout the lifetime of the respective CN NSS. We believe that such a customization of the mapping process can assist the infrastructure manager in identifying the root cause and resolving the failure, as well as facilitate a resource-sharing mechanism that allows virtualized or cloudified environments to be shared between multiple homogeneous 6G CN NSSs, thereby enhancing resource utilization. Once the mapping process of the CN NFs is complete, the infrastructure manager monitors the physical, cloud, and virtual resources and scales up and down or in and out the required resources of the mapped CN NFs according to the dynamic changes and environmental conditions of the 6G CN NSS.

The intelligent aggregation and edge data centers are deployed in the regional cloud and edge cloud, respectively (see Figure 9). Both data centers host the components of the CU and DU of a 6G RAN NSS, which are chained together in an ordered fashion and provided in a virtualized or cloudified manner. The infrastructure managers in both data centers abstract virtual or cloud-native resources from the general-purpose hardware units, map the logical components of the CU and the DU onto the abstracted resources, and autonomously scale the allocated storage and compute resources throughout the duration of the RAN NSS. The infrastructure managers perform these tasks intelligently in both aggregation and edge data centers. Both data centers shall allocate the compute and storage resources of a RAN NSS in complete isolation from the NFs of other RAN NSSs operating over a shared infrastructure. To enhance the efficiency of resource allocation and facilitate optimal M&O of the resources of the intelligent data centers, the infrastructure manager can use standardized description files for the virtual resources of the CU and DU. The aspects related to intelligent data centers (hosting both CN and RAN NSSs) in mobile networks are mainly within the scope of ETSI ISG NFV, ETSI ISG MEC, and O-RAN Alliance WG6. We anticipate that these SDOs will continue standardizing the architectural frameworks, resource abstraction and allocation, intelligentization and automation, and several aspects of data centers for the 6G networks.

Figure 9 also depicts that Public Network Domain is composed of cellular sites, each consisting of physical resources. The cellular sites can be deployed to host the logical components of the RU for the 6G RAN NSSs using physical resource descriptors, as depicted in Figure 9. To design an AI-driven, open, sustainable, and competitive physical layer for the 6G slicing framework, novel AI-native interfaces and transmission technologies shall be designed and deployed in the RU. This includes exploring and employing new waveform schemes, multi-antenna methods, spectrum-sharing mechanisms, channel coding solutions, etc. These innovative physical layer technologies and solutions could be massive multiple-input multiple-output, reconfigurable intelligent surfaces, non-orthogonal multiple access, etc. The 3GPP TSG RAN [119], the SCF [125], and the O-RAN Alliance [93] are the leading SDOs that are engaged in the standardization of the above aspects of the physical layer for 5G and beyond. We anticipate that these SDOs will continue to produce relevant specifications for 6G that can be utilized by the Infrastructure Layer.

Finally, the Public Network Domain comprises transport links that carry network traffic of a 6G NS over a highly reliable and high-capacity transmission medium in the upstream and downstream directions. They connect the CN and RAN NSSs, which are hosted at intelligent data centers and cellular sites. Figure 9 illustrates that TN can consist of up to three links: backhaul, midhaul, and fronthaul. They host the three logical transport interfaces of a TN NSS, respectively. The underlying TN can be made up of reliable physical links (in the case of wired communication). Each physical link can be virtualized into several VLs, which are composed of virtual networking resources. Once the virtualization process of the physical links is complete, the logical interfaces of the TN NSS are intelligently mapped onto their respective VLs. The required networking resources of a TN NSS are described in the description templates of the virtual and physical links. The infrastructure manager of the TN domain employs these descriptors to allocate, manage, and orchestrate the virtual and physical networking resources of the ordered 6G TN NSS and the entire transport links of a 6G network.

2) VERTICAL PREMISES

Figure 9 shows that the Vertical Premises (or non-Public Network Domain) can be owned by the owner of a vertical industry, a third-party facility, and a mobile virtual network operator (MVNO). This part includes the extreme edge domain resources, which can consist of storage, networking, and computing. The extreme edge domain can be deployed in virtualized, cloudified, and physical forms. This part also comprises end-user and IoT devices. For example, personal devices, automotive devices, IoT gateways, gaming devices, tenant-owned data centers, etc. These tenant-owned resources shall optimally and intelligently host the VNFs and PNFs, as well as the transport links, of an extreme edge NSS, as we defined in Network Slice Layer and showed in Figure 8. To efficiently perform the mapping process of the extreme edge NFs and transport links onto the underlying non-public resources, this layer can utilize the ML algorithms, similar to the mapping algorithms used by the infrastructure manager of the Public Network Domain. The extreme edge domain resources and devices may necessitate the adoption of resource allocation and configuration mechanisms tailored to their unique constraints, such as lower power capability, limited resources that are typically shared with user-controlled applications, high-number and heterogeneity of devices, volatile behavior, mobility patterns, etc. Similar to the NFs and transport links of the remaining NSSs, the NFs and transport links of an extreme edge NSS can also be managed by an infrastructure manager designed for this domain, utilizing dedicated description files.

The extreme edge domain must be distinguished from the regular domains, which are typically subject to stringently controlled conditions. Extreme edge devices, on the other hand, fall under the purview of the end-user and tenant, which means they are not necessarily subject to routine maintenance (they may be prone to error) and may be unexpectedly moved or even turned off and on (they could behave asynchronously regarding their status). This applies not only to personal devices but also to devices in corporate environments (e.g., vertical industry environments) that fall outside the purview of the operator. Additionally, the extreme edge environments will contain a heterogeneous collection of devices and data protocols, including devices with limited computing and storage capabilities. It will also be enormous in scale, even surpassing the human scale in terms of routine operations. It would be necessary to align with providers of extreme edge devices so that they provide the necessary software development kit and well-defined APIs.

F. NETWORK INTELLIGENCE LAYER

This layer collects data from the Infrastructure Layer of the proposed 6G slicing framework on a periodic or as-needed basis, analyzes the collected data using AI and ML algorithms, and provides intelligent solutions, real-time predictions, and accurate recommendations for the purpose of empowering the operations and processes of the underlying Infrastructure Layer. The Network Intelligence Layer is designed specifically for the intelligentization and automation of the Infrastructure Layer, as the remaining layers of the proposed framework have their own relevant standardized intelligent schemes and tools designed by the corresponding SDOs. For example, the Management and Orchestration Layer utilizes, among others, the ETSI ISG ZSM and 3GPP TSG SA5 intelligent and automation frameworks to automate and intelligentize its operations and processes. Assuming that some layers may require historical data or may need to import a trained AI/ML model from the Network Intelligence Layer, the proposed framework shall support these interactions. Consequently, the scope of the Network Intelligence Layer may go beyond the intelligentization of the Infrastructure Layer.

The integration of AI into the Infrastructure Layer can bring several benefits [2], [9], such as improved flexibility and adaptability, enhanced security and privacy, better efficiency and performance, accurate localization and positioning, and an intelligent network through ML and edge computing. The primary features and capabilities of this layer include data collection and analysis, intelligent solutions, real-time predictions, and accurate recommendations [2], [9], [126].

To perform the above tasks and deliver novel intelligence and automation capabilities, the Network Intelligence Layer uses two standardized frameworks: the 3GPP network data and analytics function (NWDAF) and the ETSI ISG experiential networked intelligence (ENI). The NWDAF framework is applicable to 3GPP-defined domains and components. The ENI framework is designed in such a manner that it can be applied to any domain of a 6G network. The latest frameworks of the 3GPP NWDAF and the ETSI ISG ENI are illustrated in part (a) and part (b) of Figure 10, respectively. In the following, we discuss their applicability to the intelligentization and automation of the Infrastructure Layer.

1) UTILIZING THE 3GPP NWDAF FRAMEWORK

The NWDAF has been specified by the 3GPP TSG SA2 for the integration of intelligence and automation into 3GPP-defined networks [127]. The NWDAF was defined for the intelligentization of 5G and can be extended towards 6G. The latest framework of the NWDAF can be deployed in a centralized or distributed manner [128], where each of the deployment options can have its own capabilities and limitations. The Network Intelligence Layer can deploy NWDAF in the CN domain and RAN domain of the Infrastructure Layer. In both domains, the NWDAF can collect a large amount of historical and real-time network-related data from the corresponding service nodes, applications, and NFs through the use of ML data collection techniques. It then extracts relevant information from the collected data, performs network analytics, produces insights through the use of a standardized architecture, and takes well-informed decisions with the goal of enhancing the functionality and performance of the 6G RAN and CN NSSs [127]. This standard mechanism for collecting, analyzing, and exposing data related to the Infrastructure Layer enables the proposed framework to manage, automate, and optimize the operations

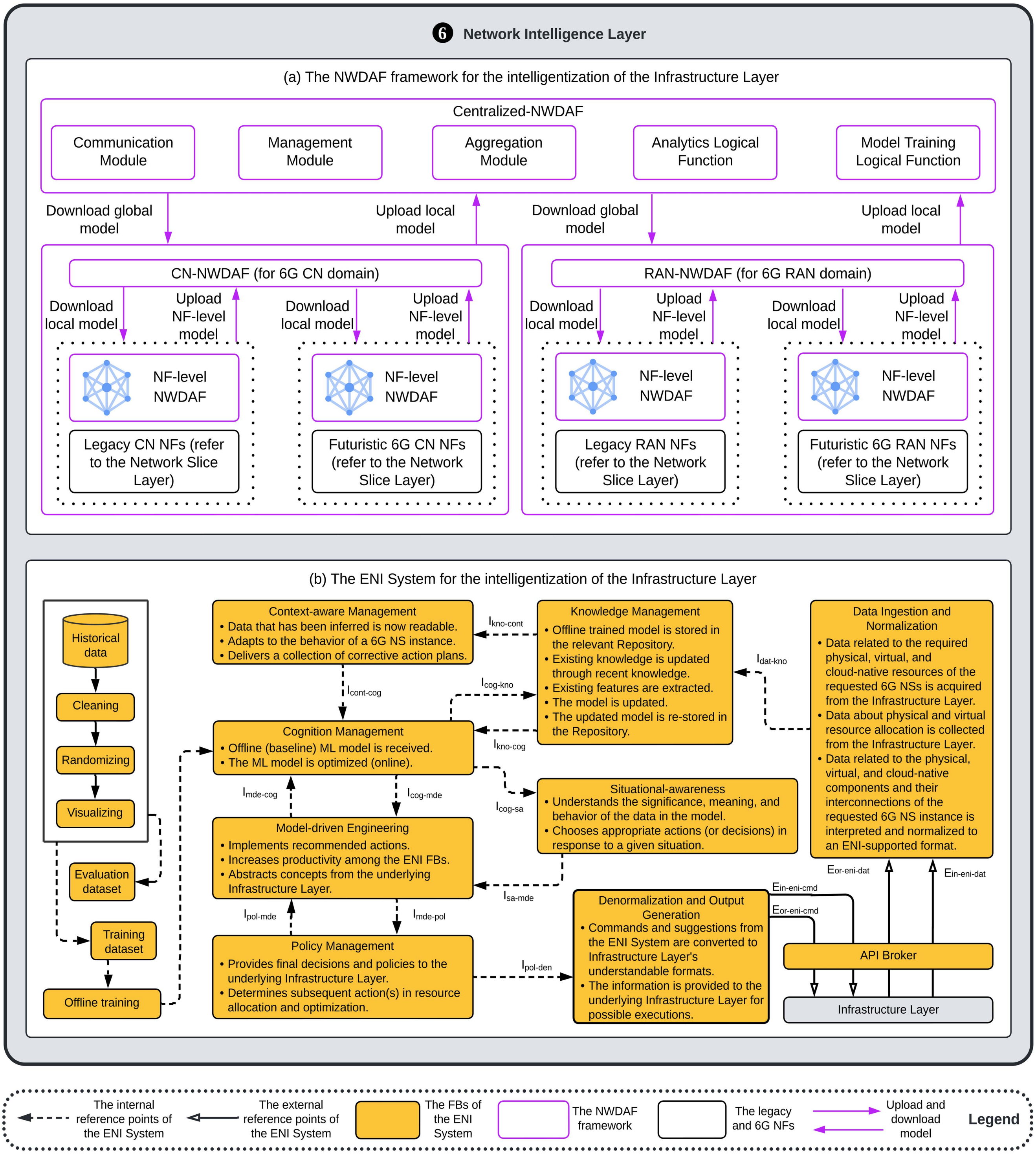

6 Network Intelligence Layer
(a) The NWDAF framework for the intelligentization of the Infrastructure Layer
Centralized-NWDAF
Communication Module
Management Module
Aggregation Module
Analytics Logical Function
Model Training Logical Function
Download global model
Upload local model
Download global model
Upload local model
CN-NWDAF (for 6G CN domain)
RAN-NWDAF (for 6G RAN domain)
Download local model
Upload NF-level model
Download local model
Upload NF-level model
Download local model
Upload NF-level model
Download local model
Upload NF-level model
NF-level NWDAF
NF-level NWDAF
NF-level NWDAF
NF-level NWDAF
Legacy CN NFs (refer to the Network Slice Layer)
Futuristic 6G CN NFs (refer to the Network Slice Layer)
Legacy RAN NFs (refer to the Network Slice Layer)
Futuristic 6G RAN NFs (refer to the Network Slice Layer)
(b) The ENI System for the intelligentization of the Infrastructure Layer
Historical data
Cleaning
Randomizing
Visualizing
Evaluation dataset
Training dataset
Offline training
Context-aware Management
• Data that has been inferred is now readable.
• Adapts to the behavior of a 6G NS instance.
• Delivers a collection of corrective action plans.
Knowledge Management
• Offline trained model is stored in the relevant Repository.
• Existing knowledge is updated through recent knowledge.
• Existing features are extracted.
• The model is updated.
• The updated model is re-stored in the Repository.
Data Ingestion and Normalization
• Data related to the required physical, virtual, and cloud-native resources of the requested 6G NSs is acquired from the Infrastructure Layer.
• Data about physical and virtual resource allocation is collected from the Infrastructure Layer.
• Data related to the physical, virtual, and cloud-native components and their interconnections of the requested 6G NS instance is interpreted and normalized to an ENI-supported format.
Cognition Management
• Offline (baseline) ML model is received.
• The ML model is optimized (online).
Situational-awareness
• Understands the significance, meaning, and behavior of the data in the model.
• Chooses appropriate actions (or decisions) in response to a given situation.
Model-driven Engineering
• Implements recommended actions.
• Increases productivity among the ENI FBs.
• Abstracts concepts from the underlying Infrastructure Layer.
Policy Management
• Provides final decisions and policies to the underlying Infrastructure Layer.
• Determines subsequent action(s) in resource allocation and optimization.
Denormalization and Output Generation
• Commands and suggestions from the ENI System are converted to the Infrastructure Layer's understandable formats.
• The information is provided to the underlying Infrastructure Layer for possible executions.
API Broker
Infrastructure Layer
$I_{kno-cont}$
$I_{cont-cog}$
$I_{cog-kno}$
$I_{kno-cog}$
$I_{dat-kno}$
$I_{mde-cog}$
$I_{cog-mde}$
I_{cog-sa}
I_{sa-mde}
$I_{pol-mde}$
$I_{mde-pol}$
$I_{pol-den}$
$E_{or-eni-dat}$
$E_{in-eni-dat}$
$E_{in-eni-cmd}$
$E_{or-eni-cmd}$
The internal reference points of the ENI System
The external reference points of the ENI System
The FBs of the ENI System
The NWDAF framework
The legacy and 6G NFs
Upload and download model
Legend

FIGURE 10. The architecture of the Network Intelligence Layer of the proposed open and intelligent 6G slicing framework.

and processes of 6G NSs more efficiently in comparison to the state-of-the-art algorithmic solutions.

Figure 10 illustrates that the NWDAF framework can consist of a Centralized-NWDAF, two domain-specific NWDAFs (i.e., the RAN NWDAF and the CN NWDAF), and a number of NF-level NWDAFs. We believe that deploying a centralized NWDAF is challenging due to data privacy and security, requiring additional network resources and costs to collect and process a large amount of data in Centralized-NWDAF, limited flexibility in terms of customization, etc. Hence, a decentralized intelligence architecture (proposed in Figure 10) for the NWDAF is expected to overcome most of the said challenges. In each domain of the NWDAF architecture, the NF-level NWDAFs are deployed close to the CN and RAN NFs and are used to construct customized NF-level models using data analytics collected from the corresponding NFs. The CN and RAN NWDAFs derive data analytics from several NF-level NWDAFs within their domains with the intention of developing models customized for their respective domains, namely models for 6G RAN and CN NSs. The Centralized-NWDAF collects the domain-specific models from the distributed NWDAFs and constructs a global model, namely the model for a 6G NS. These three various hierarchical NWDAFs employ 3GPP standard interfaces to download and upload the relevant models and are suitable for the deployment of FL algorithms in the Network Intelligence Layer.

The Centralized-NWDAF is assumed to be deployed in a central cloud and is composed of at least three major modules and two main logical functions (see Figure 10) [127], [128]. The Analytics Logical Function executes inferences and produces analysis information (such as predictions, recommendations, statistics, etc.). The Model Training Logical Function is used to train the ML model (i.e., mainly global model) and provides the trained model to the Analytics Logical Function for the execution of inferences and generation of analysis information. The Communication Module is used to connect the Centralized-NWDAF with the CN- and RAN-NWDAFs via 3GPP-defined interfaces. The Aggregation Module can collect local models from domain-specific NWDAFs to generate a global model, aggregate a global model into several local models, and provide local models back to domain-specific NWDAFs. The Management Module is employed to manage the inferences, aggregation, and analysis processes within the Centralized-NWDAF. The Centralized-NWDAF is connected with CN- and RAN-NWDAFs via a 3GPP-defined interface to upload local models and download global model. Depending on the design and deployment of the Network Intelligence Layer, the two domain-level NWDAFs can also consist of some of the modules and logical functions discussed above and be responsible for executing the above tasks in their domains.

The CN-NWDAF is assumed to be deployed in the CN domain. To perform analytics collection, the CN-NWDAF is connected with its associated NF-level NWDAFs by means of standardized interfaces for uploading NF-level models and downloading local model (i,e, CN NSS model). Each NF (legacy and futuristic 6G NFs) in the CN domain has its own customized NF-level NWDAF, which can collect and analyze data from the corresponding CN NF and deliver the trained NF-level model to the CN-NWDAF. We anticipate that each NF-level NWDAF can only be provided with the learning module. Once the CN-NWDAF receives the trained models from all CN NF-level NWDAFs, it merges the NF-level models to generate a local model at the 6G CN NSS level. The NF-level NWDAFs can download the CN NSS model to enhance the performance of the NF-level models. The upload of NF-level models and the download of the CN NSS model between both entities can be performed on the basis of needs or in a specific time interval. Figure 10 illustrates that each NF-level NWDAF in the CN domain is responsible for the analytics of the NF to which it is assigned and connected. Hence, the model trained in a NF-level NWDAF can be accessed only by the CN-NWDAF and this specific NF-level NWDAF. This customization feature of NWDAF is believed to improve the privacy and security of training, uploading, downloading, and managing the NF-level models and local models across the CN domain in the Infrastructure Layer.

The RAN-NWDAF is assumed to be deployed at the RAN domain of a 6G network and establishes a connection with all NF-level NWDAFs in the 6G RAN domain via 3GPP-standardized interfaces. The NF-level NWDAFs in the RAN domain collect analytics related to the 6G RAN domain from the CU, DU, and RU, as well as upload the NF-level trained models to the RAN NWDAF and download the local model (i.e., RAN NSS model) from the RAN NWDAF based on needs or time-interval. In addition to the legacy RAN NFs, there will be 6G radio NFs that may also require NF-level NWDAFs. Similar to the CN domain, each NF in the RAN domain has its customized NF-level NWDAF, collecting data and training model related to this specific NF. The NF-specific model in the RAN domain is only accessible by the corresponding NF-level NWDAF and RAN-NWDAF, thereby improving the security and privacy of RAN intelligentization in the 6G slicing framework.

Figure 10 illustrates that in 6G CN and RAN domains, the data is collected and analyzed at the NFs of both domains, and only the trained models (i.e., NF-level, domain-level, and global models) are shared among the entities of the proposed NWDAF framework. In each network domain, each NF provides a specific type of data analytics to its corresponding domain-level NWDAF. The domain-level NWDAF collects the required data analytics from all (or certain) NFs in order to deliver meaningful analytics at the corresponding domain level. Likewise, each domain-level NWDAF in Network Intelligence Layer provides specific analytics at the domain level which is then delivered to the Centralized-NWDAF, where a number of domain-level models are trained to generate a more accurate global model. This is similar to the FL algorithms, in which the data collection and storing at a centralized repository is avoided and only the trained models are shared among the corresponding components. Hence, applying FL algorithms within the context of the proposed NWDAF framework of the Intelligence Layer can be efficient in terms of security, privacy, delay, and several other performance metrics.

2) UTILIZING THE ETSI ISG ENI FRAMEWORK

The Network Intelligence Layer of the proposed architecture can also employ the ETSI ISG-defined ENI framework for integrating numerous types of cutting-edge ML algorithms, context-aware policies, closed-loop schemes, and metadata-driven mechanisms into various network domains, transport domains, and operations of a 6G network [129]. The grand objective of such an AI-based framework is to adjust the operations and processes, including resource scheduling, optimization, and orchestration, of the Infrastructure Layer according to the frequent changes in end-user requirements, business objectives, and 6G network operations. In our proposed framework, the ENI System can collect a large scale of data from network domains of the Infrastructure Layer, understand their configuration and operational status in real-time, and employ advanced ML algorithms to enable intelligent service deployment, resource management, monitoring, maintenance, predictions, and other operations within the said layer [9]. To perform these tasks, we extend the ENI framework and integrate it into the Network Intelligence Layer to automate and intelligentize the operations and processes of the Infrastructure Layer. The ENI framework consists of several input- and output-facing FBs as well as API brokers that are used to collect and process the data and provide well-informed actionable decisions to the Infrastructure Layer [129]. The extended ENI framework in Network Intelligence Layer is depicted in part (b) of Figure 10. It can train and deploy a relevant ML model in both online and offline scenarios.

In the case of online training, the input-facing FBs acquire data from various network domains of the Infrastructure Layer via an open and intelligent API broker, ingest the input data, normalize the collected data into a format the ENI System can understand, and distribute it to its internal entities. The internal entities of the ENI System process the data, employ the relevant ML algorithms to make real-time predictions, accurate recommendations, and intelligent solutions, and provide them to its output-facing FBs. The outputs of a ML algorithm are then denormalized and translated into a format the Infrastructure Layer can understand via an API broker. Then, the Infrastructure Layer can automate its operations and processes with respect to the deployed 6G NSs using the predictions, recommendations, solutions, etc., acquired from the Network Intelligence Layer. The ENI System in the Network Intelligence Layer and network domains of the Infrastructure Layer can be connected via ETSI ISG ENI-defined external interfaces. Moreover, the ENI System is also connected to a set of applications, which are required to participate in the automation and intelligentization of the Infrastructure Layer. The ENI System is controlled by an administrator, located in the operator's domain. For simplicity, these interconnections are not shown in Figure 10. Interested readers can refer to [9], [129] and the references therein for a detailed description of the ENI framework and its applicability in public networks.

The Data Ingestion is an input-facing FB in the ENI framework, which ingests structured, semi-structured, and unstructured data – provided by streaming, batch, and on-demand mechanisms – related to various network domains of the Infrastructure Layer. This FB performs filtering, correlation, cleansing, anonymization and pseudonymization, augmentation, and labeling operations in the Network Intelligence Layer on raw data collected from the Infrastructure Layer. Once the data collection and ingestion processes are completed, the ingested data is then sent to the Normalization FB to interpret and normalize it into a single, common, and unified format that is understandable for further processing by the internal entities of the ENI framework. The normalization of ingested data related to the Infrastructure Layer into a standard format, on the one hand, enables Network Intelligence Layer to quickly learn about the optimal parameters of each of the nodes located in the Infrastructure Layer, on the other hand, reduces complex computational problems between the ENI System and the network domains of the Infrastructure Layer.

Subsequently, the normalized data is sent to internal entities, especially Knowledge Management FB, which is used to filter the normalized data. The filtered data must then be subjected to additional processing by multiple internal entities. These internal entities (see Fig. 10) are initially in charge of producing recommendations, commands, and knowledge and are used to automate the operations of the Infrastructure Layer with respect to 6G NS instances. In addition, they use existing knowledge and/or add new knowledge to enable the ENI framework to adapt its behavior according to the dynamic changes in the Infrastructure Layer. The internal entities are Context-aware Management, Cognition Management, Model-driven Engineering, Situational-awareness, and Policy Management FBs. To avoid redundant information and for the sake of simplicity, we provide a brief description of their functionalities and roles within the context of the ENI framework in Figure 10. More detail on them can be found in [129]. Once the commands and recommendations are generated by the internal entities, they are delivered to the Denormalization and Output Generation FB, which is considered an output-facing FB in the ENI framework.

In the case of offline training, the ENI framework utilizes corresponding historical data, which is collected from the Infrastructure Layer and archived in repositories of the Network Intelligence Layer. The historical data is first cleaned, randomized, and visualized using a variety of data cleaning, randomization, and visualization techniques by the Network Intelligence Layer and delivered to the ENI framework. The ENI framework then partitions the historical data into two distinct sets: the training data set and the evaluation data set. The former data set is used to train the relevant ML model. The latter data set is employed to assess the performance objectives of the selected ML algorithms. Assuming that the model has been successfully trained, the ENI framework provides the offline-trained model to its internal entities, which undergo similar processes as the online model discussed above. The internal entities then provide recommendations and predictions produced by the offline-trained model to the Denormalization and Output Generation FB to deliver them to the corresponding network domain in the Infrastructure Layer.

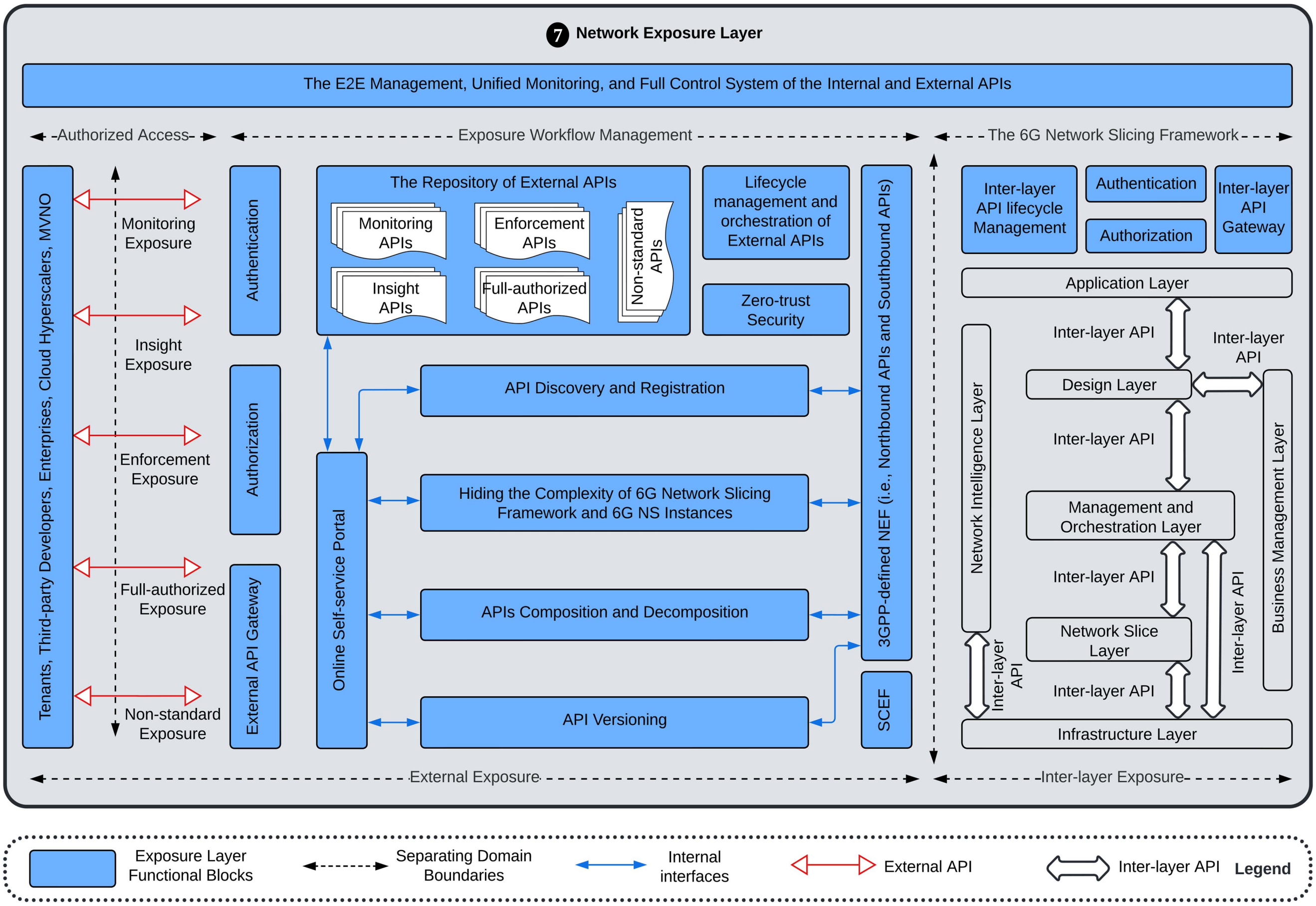

7 Network Exposure Layer
The E2E Management, Unified Monitoring, and Full Control System of the Internal and External APIs
Authorized Access
Exposure Workflow Management
The 6G Network Slicing Framework
Tenants, Third-party Developers, Enterprises, Cloud Hyperscalers, MVNO
Monitoring Exposure
Insight Exposure
Enforcement Exposure
Full-authorized Exposure
Non-standard Exposure
Authentication
Authorization
External API Gateway
The Repository of External APIs
Monitoring APIs
Enforcement APIs
Insight APIs
Full-authorized APIs
Non-standard APIs
Lifecycle management and orchestration of External APIs
Zero-trust Security
3GPP-defined NEF (i.e., Northbound APIs and Southbound APIs)
SCEF
Online Self-service Portal
API Discovery and Registration
Hiding the Complexity of 6G Network Slicing Framework and 6G NS Instances
APIs Composition and Decomposition
API Versioning
Inter-layer API lifecycle Management
Authentication
Authorization
Inter-layer API Gateway
Application Layer
Inter-layer API
Inter-layer API
Design Layer
Inter-layer API
Management and Orchestration Layer
Inter-layer API
Network Slice Layer
Inter-layer API
Infrastructure Layer
Network Intelligence Layer
Business Management Layer
Inter-layer API
Inter-layer API
External Exposure
Inter-layer Exposure
Exposure Layer Functional Blocks
Separating Domain Boundaries
Internal interfaces
External API
Inter-layer API
Legend

FIGURE 11. The architecture of the Network Exposure Layer of the proposed open and intelligent framework for 6G slicing.

During both online and offline training of an algorithm, the Denormalization FB executes the inverse functionalities of the Normalization FB. It specifically means that – once it receives the processed data from the internal entities – it is used to translate the processed data coming from the ENI's entities into the format(s) that is(are) understandable by the Infrastructure Layer. The translated data is then forwarded to the Output Generation FB to deliver it to the corresponding network domain of the Infrastructure Layer. The Infrastructure Layer utilizes the commands and recommendations of the ENI System to automate operations related to 6G NSs. The Denormalization FB communicates with the Infrastructure Layer through an API broker in order to translate the APIs of the ENI System to the APIs of the Infrastructure Layer (see Fig. 10). The API broker shall be compliant with all internal and external interfaces of the ENI framework.

G. NETWORK EXPOSURE LAYER

The Network Exposure Layer is an essential layer of the proposed slicing framework, offering a number of standardized and secure interfaces for exposing novel 6G capabilities and slice-specific services to external applications, tenants, third-party developers, and MVNOs. The Network Exposure Layer can expose, control, and manage a large set of well-defined, secure, standardized, and controllable capabilities of APIs across the proposed framework. These features enable the Network Exposure Layer to regulate effective communication and interaction between all layers of the proposed slicing framework, as well as authorize the tenants and third-party developers with a certain level of exposure in order to grant access to and perform specific control, management, and orchestration tasks related to 6G services in a 6G network. In the proposed framework, the Network Exposure Layer enables the tenant and third-party developers to manage the E2E lifecycle of 6G NSs that are tailored to meet the unique requirements of their particular use cases. Hence, this layer must act as a gateway between the 6G network and the external ecosystem for effective exposure, efficient management, and seamless programming of exposed services and NFs of a 6G NS. The Network Exposure Layer shall also enable operators to expose and monetize their network assets while providing a secure and reliable environment for third-party developers and tenants to build and deploy their own 6G applications and services on top of the existing ones in a 6G network. Network and service exposure was regarded as a capability of a communication network in previous generations [130]; however, in beyond 5G and 6G slicing-aware networks, we anticipate that it shall be considered a requirement for managing and innovating on heterogeneous NSs. In addition, we believe that the network exposure and relevant APIs in 6G networks will be ultra-programmable, easily usable, incredibly cost-effective, more open, and extremely intelligent compared to the state-of-the-art network exposing mechanisms and service APIs.

Network exposure encompasses not only access to NFs, data, nodes, and services within the proposed 6G slicing framework, but also the ability for tenants and third-party developers to execute a set of well-defined operations associated with their respective 6G NSs in a 6G network. As stated previously, the exposed NFs, services, and nodes inside the 6G slicing framework can be made accessible through a set of open and secure APIs. The external applications (i.e., those associated with third-party developers and tenants) and the NFs and services within the 6G slicing framework need to communicate to exchange exposure-related information in a secure manner. To accomplish this, the Network Exposure Layer shall provide the necessary signaling to enable the exchange of such information between both sides. Although the design and deployment of the Network Exposure Layer may differ from one operator to another, we believe that it would be efficient for both the tenant and network operator that each 6G NS be offered with its own customized API(s). The tenants or third-party developers can perform either all or some of the following tasks using the offered APIs [131]: monitoring (i.e., to detect critical, major, and minor incidents associated with a 6G NS), insight (i.e., to analyze the collected data and acquire the desired results), and enforcement (i.e., to execute well-informed operations in the 6G network). Each of these three tasks can be performed through either a task-customized API that can be combined and provided to the tenant by the operator, or a single API capable of executing them in a secure, controlled, and flexible manner [130], [132]. The authorization of the tenants and third-party developers to perform the aforementioned three tasks shall be clearly defined in the SLA [81], [133].

Within the context of the Network Exposure Layer, a network operator must carefully determine which features of the 6G network slicing framework should be exposed to tenants and third-party developers and which should remain hidden. While it is important to provide access to 6G network resources and services, it is equally important to ensure that sensitive or confidential information is not exposed to unauthorized tenants and developers. There are several features and metrics that a network operator must hide from third-party developers and tenants to protect the security of the proposed architectural framework and the privacy of all tenants who access the 6G network slicing framework. They consist of: network topology, security policies, user and tenant data, SLAs, authentication and authorization mechanisms, network and service repositories, network and service management systems, billing and charging databases, and other sensitive information and data [130], [131], [132]. In general, network operators must find a balance between exposing network resources and services to third-party developers and safeguarding the security and privacy of the 6G network. By judiciously regulating which features are hidden from third-party developers, network operators may ensure that their network stays secure and reliable while fostering innovation and ecosystem growth. While hiding sensitive information and data (both network- and user-related), the network operator shall ease the design, deployment, and operation of the APIs. In order to accomplish this, the network operators can use a combination of a number of novel strategies, including standardized APIs, virtualization and abstraction, robust testing and monitoring, and secure authentication and authorization.

In order for the network operator to manage a large number of APIs and control the access of each to the 6G slicing framework, it is critical that a comprehensive API management and control system be incorporated into the Network Exposure Layer architecture. The management and control capabilities of the Network Exposure Layer require a combination of technical and non-technical measures to ensure that APIs are deployed efficiently, operating correctly, and meeting the requirements of tenants and third-party developers. The API management and control system shall have the capabilities to discover and register each API, archive the templates of APIs that shall contain up-to-date information regarding their functionalities and parameters, provide real-time monitoring of API performance, facilitate the onboarding of third-party developers, streamline the API management process, etc. It shall also be in charge of authenticating and authorizing the third-party developers accessing the 6G NS and the components included in the 6G NS, thereby preventing possible attacks and ensuring that the third-party developers access only the subscribed components in a 6G network. By establishing such a comprehensive management and control system in the Network Exposure Layer, network operators are able to effectively manage a large number of APIs while still maintaining the security and reliability of the 6G NSs.

To effectively perform the aforementioned management and control tasks, the Network Exposure Layer can utilize the 3GPP-defined network exposure function (NEF) and a number of novel components. The architecture of the Network Exposure Layer is shown in Figure 11. The NEF can provide a standardized and secure way to expose the capabilities of the Network Exposure Layer and the elements of various layers of the proposed framework to third-party developers and tenants. In 6G, the NEF and all components of the proposed layer-specific architecture (shown in Figure 11) can be optimized to fulfill the exposure capability of the network slicing framework and 6G network in general, similar to the evolution of service capability exposure function (SCEF) to NEF from 4G to 5G [130]. In addition to gaining access to network elements, the NEF and the rest of the components of this layer shall allow third-party developers and tenants to execute operations within the 6G slicing framework in a manner that conceals the complexity of the underlying infrastructure. Hence, the NEF and the architecture of the Network Exposure Layer shall enable a tenant to configure, monitor, control, and modify specific policies, QoS parameters, M&O resources, and the status of connected devices to the corresponding NS.

1) THE UNIFIED FRAMEWORK OF THE NETWORK EXPOSURE LAYER

The architectural framework of the Network Exposure Layer is illustrated in Figure 11. As can be seen at the extreme bottom side, this architecture is divided into two major parts: the inter-layer exposure part and the external exposure part. Internal exposure is referred to the capability of the Network Exposure Layer that enables one layer to gain access to other layers of the proposed framework. Through the use of external exposure capability, the Network Exposure Layer allows the third-party developers, tenants, enterprises, and others to obtain authorized access to the proposed 6G slicing framework. Both parts of the Network Exposure Layer are managed, monitored, and controlled by an E2E API management, monitoring, and control system. In the rest of this section, we discuss both parts and their functioning architecture and management system in more detail.

In the inter-layer exposure part, the Network Exposure Layer allows each layer to expose its data, capabilities, functionalities, applications, etc. to certain (or all) layers via a secure, standardized, and intelligent inter-layer API. The inter-layer APIs, which are also known as network APIs, are essential for the optimal design and deployment of the proposed framework, as they allow different layers to be developed and maintained independently while ensuring seamless interaction between them. The inter-layer APIs also provide a method for enforcing security and data validation checks, as well as ensuring that modifications to one layer do not negatively impact the functionality of the layers above or below it. For instance, the Management and Orchestration Layer may need to communicate with the Design Layer to retrieve data (i.e., the templates and description files) pertaining to the design of a 6G NS, as well as communicate with the Network Slice Layer to translate design-related parameters into network-understandable languages and metrics. Such an inter-layer API would specify the set of functions and parameters that the Management and Orchestration Layer can use to send requests to the Design Layer and Network Slice Layer, as well as the format of the responses that both layers should return. In Figure 11, we show a minimum required number of inter-layer APIs in the proposed framework. Depending on the requirements and design, as well as the number of 6G NSs, the number of the inter-layer APIs will decrease or increase. Since the Deployment Perspective is beyond the scope of this article, we let the network operator decide the interconnection between these layers.

Figure 11 illustrates that the internal exposure part of the Network Exposure Layer consists of a number of FBs, such as authentication FB, authorization FB, inter-layer API gateway, and inter-layer API lifecycle management FB. In the internal exposure part, the authentication FB verifies the identity of the software component or application of a layer that is making a request to the inter-layer API. It also ensures that only authorized software components or applications from a layer can access and utilize the inter-layer APIs provided by the proposed framework [134]. Without authentication, any software component or application could potentially access the inter-layer APIs provided by other layers, which could lead to unauthorized access to sensitive data or systems. By requiring authentication, the inter-layer APIs can ensure that only authorized components can access its functionality, helping to maintain the security and integrity of the 6G slicing framework. In the internal exposure part, the authorization FB is responsible for determining whether the software component or application that is making a request to the inter-layer API has the necessary permissions to access the requested resource or perform the requested action [134]. Authentication and authorization are the two critical FBs utilized by the internal exposure part to safeguard inter-layer APIs and data. Authentication verifies the identity of software components or applications, whereas authorization governs their access permissions [134]. Figure 11 illustrates that the internal exposure part also includes an API gateway that can serve as an entry point for requests from one layer to another. In the proposed slicing framework, it can handle tasks such as protocol translation, load balancing, security enforcement, and service discovery between layers. Finally, the internal exposure part includes the inter-layer lifecycle management FB that is responsible for the creation, deployment, deactivation, and decommissioning of an inter-layer API in the proposed framework.

The external exposure part of the Network Exposure Layer is responsible for exposing the functionalities, capabilities, data, applications, and other relevant properties to external parties of the 6G network via secure and standard APIs. These types of external APIs are also called service APIs. The parties that can access the 6G slicing framework can be tenants, third-party developers, enterprises, cloud Hyperscalers, MVNOs, and others. The external parties can use these APIs for a variety of purposes: monitoring, insight, enforcement, and even for non-standard exposure capabilities. This part is divided into two sub-parts: authorized access and exposure workflow management. The former is in charge of determining the external parties that access the proposed 6G slicing framework. The latter, similar to the internal exposure part, includes a number of FBs that are connected via internal standardized interfaces to perform the management and execution of external APIs. These FBs in the exposure workflow management subpart are shown in Figure 11 and described in the following.

The authentication FB is responsible for verifying the identity of the third-party developer or application that is making a request to the API. It ensures that only authorized third-party developers or applications can access the resources provided by the API. This FB typically implements user authentication, token-based authentication, OAuth-based authentication, digital certificates, and other authentication mechanisms to ensure the identity of the third-party developers and applications accessing the API [134], [135]. We believe that access to any layer should be restricted on the basis of business requirements and the agreement between the third-party developer and the network operator. By implementing a layer-by-layer access control mechanism, the third-party developer should not be granted the right to enter every layer by default. Access to any particular layer has to be granted only if there is a specific need to access that specific layer. Within the context of external exposure, the authorization FB is responsible for determining whether a third-party developer or tenant making a request to an external API has the required permissions to access the requested resource or execute the requested action within the proposed framework. This FB carries out the authorization on the basis of identity, tenant or third-party roles, access control lists, or other relevant attributes [135]. It shall also implement a set of rules defining which third-party developers or tenants are authorized to access which resources or perform which actions in the proposed framework. External API gateway provides an abstraction layer between the tenants and the proposed slicing framework, handling tasks such as request routing, protocol translation, load balancing, and security enforcement. The API gateway makes it easier to access, manage, and control a large number of external APIs. It can be deployed as a standalone component or a long with authentication and authorization FBs as a large entry point of the Network Exposure Layer.

The 6G network operators may provide an online self-service developer portal to assist in the onboarding of authorized third-party developers and tenants and streamline the external API administration process. This portal should provide a user-friendly interface for developers and tenants to discover and request access to external APIs, access API documentation, and track API usage. The Network Exposure Layer additionally includes a repository for archiving a large number of external API template files. With an online self-service portal, third-party developers and tenants can access this repository and select a suitable API. To help third-party developers understand how to use APIs, network operators may provide extensive documentation that details the functionality, parameters, and usage instructions for each API. This documentation should be kept up-to-date and readily accessible to developers in this repository. Figure 11 illustrates that Network Exposure Layer also includes API discovery and registration FB that is used by tenants to discover and register with the corresponding API and by the network operator to manage the registration and discovery operations of the APIs. This can be accomplished by developing an API registration and discovery system that provides a central location for developers to locate and request access to APIs.

Once the tenant or third-party developer is authorized to access and select an API, the Network Exposure Layer is then responsible for providing the proper access and execution roles related to the 6G NS in the proposed framework. However, during the utilization phase of the API, this layer shall hide the underlying network and its complexity from the third-party developers and expose only the necessary functionality required by the developer. These tasks are executed by a FB shown in Figure 11. In addition to technical measures, the Network Exposure Layer may also choose to implement contractual and legal measures to protect the network and its infrastructure from unauthorized parties. The Network Exposure Layer also consists of an API composition and decomposition FB, as shown in Figure 11. The term “composition” refers to the process of combining multiple APIs to create higher-level functionality or services. In other words, composition allows developers and tenants to create more complex applications by combining smaller, more specialized APIs. This approach can help to simplify the development process by breaking down complex tasks into smaller, more manageable components. Decomposition, on the other hand, refers to the process of breaking down a large, monolithic API into smaller, more modular components. This approach can help to improve the maintainability and scalability of the API by making it easier to add or remove functionality as needed. The Network Exposure Layer also includes an API versioning FB. It ensures that changes to APIs do not disrupt third-party developers, and network operators must implement versioning to manage API changes. This can include maintaining backward compatibility, using semantic versioning to indicate the level of changes, and providing adequate notice and support for API changes.

The Network Exposure Layer can perform the above tasks using 3GPP-defined NEF and SCEF [131], [136]. Both FBs are defined by the 3GPP for 4G and 5G network exposure, respectively [136]. We anticipate that these FBs will be upgraded with novel capabilities and features to fulfill the network and service exposure requirements in 6G, providing both northbound and southbound APIs services. In addition, the external exposure part consists of security FB, ensuring the security and integrity of an API and its associated resources. The security FB performs a number of critical security and privacy functions, such as encryption, rate limiting, auditing and logging, etc. Another critical FB of the external exposure part is the API lifecycle M&O FB, which is playing an essential role in developing and maintaining an external API. This FB manages an API from creation to deprecation. It includes various stages such as design, development, testing, deployment, monitoring, versioning, and retirement. Each stage requires careful planning and execution to ensure that the API is reliable, scalable, and secure. Furthermore, this FB is also in charge of managing multiple APIs in a way that they can work together seamlessly. It includes managing the connections between APIs, managing API requests, and responses, and handling errors and exceptions. API orchestration tools are used to automate these tasks and to ensure that the APIs work efficiently and effectively.

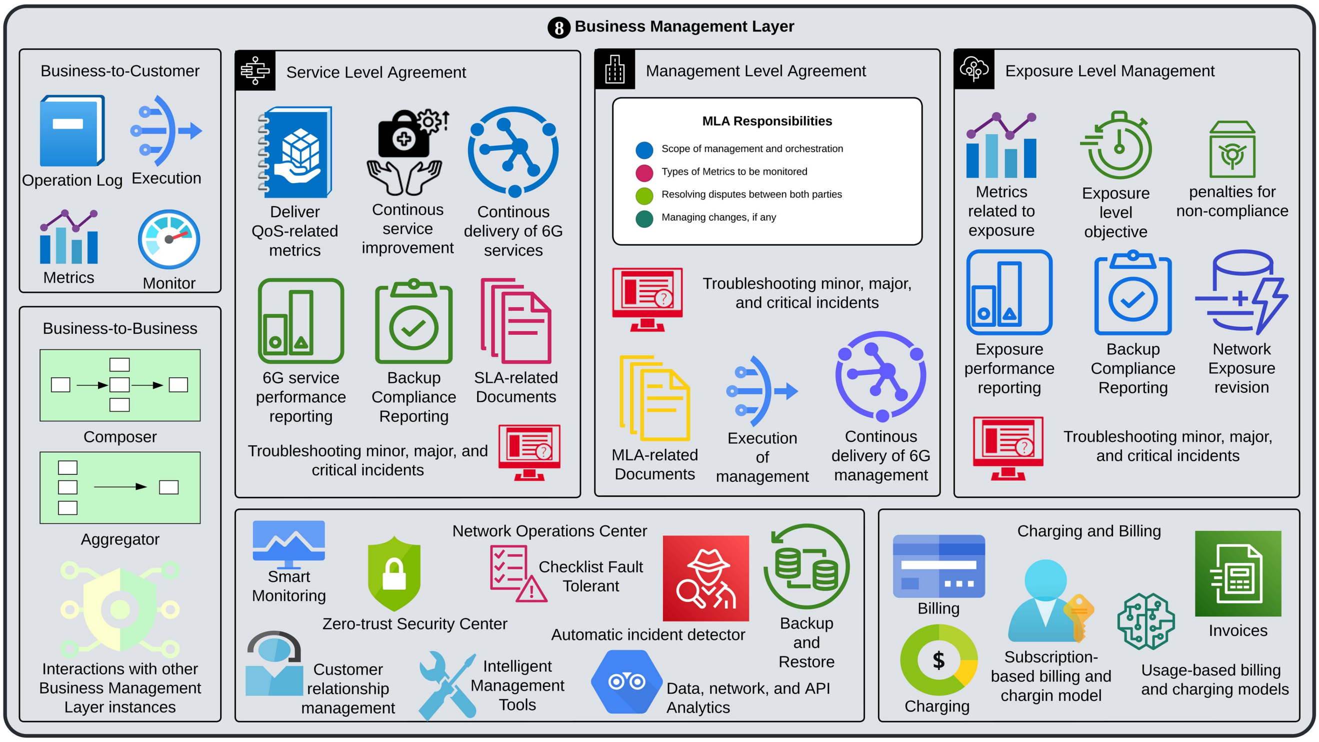

8 Business Management Layer
Business-to-Customer
Operation Log
Execution
Metrics
Monitor
Business-to-Business
Composer
Aggregator
Interactions with other Business Management Layer instances
Service Level Agreement
Deliver QoS-related metrics
Continous service improvement
Continous delivery of 6G services
6G service performance reporting
Backup Compliance Reporting
SLA-related Documents
Troubleshooting minor, major, and critical incidents
Management Level Agreement
MLA Responsibilities
Scope of management and orchestration
Types of Metrics to be monitored
Resolving disputes between both parties
Managing changes, if any
Troubleshooting minor, major, and critical incidents
MLA-related Documents
Execution of management
Continous delivery of 6G management
Exposure Level Management
Metrics related to exposure
Exposure level objective
penalties for non-compliance
Exposure performance reporting
Backup Compliance Reporting
Network Exposure revision
Troubleshooting minor, major, and critical incidents
Network Operations Center
Smart Monitoring
Checklist Fault Tolerant
Zero-trust Security Center
Automatic incident detector
Backup and Restore
Customer relationship management
Intelligent Management Tools
Data, network, and API Analytics
Charging and Billing
Billing
Charging
Subscription-based billing and chargin model
Invoices
Usage-based billing and charging models

FIGURE 12. The architecture of the Business Management Layer of the proposed open and intelligent 6G slicing framework.

Figure 11 demonstrates that both the internal and external exposure parts are managed, controlled, and monitored by an E2E management, monitoring, and control system. This system monitors the entire Network Exposure Layer, provides each tenant with a certain level of controlling and managing rights of the 6G NSs, and manages the internal and external APIs throughout their lifetime. Finally, and as we discussed, the network and service exposure in the 6G slicing framework can be extremely complicated and, meanwhile, critical, requiring careful consideration during the design, deployment, and termination of the relevant APIs. There could be technical, business, and legal requirements that are needed to be negotiated and agreed upon between the third-party developers and the network operator before the design and preparation of an API. Traditionally, these aspects are considered metrics of an SLA and are negotiated between the network operator and third-party developers. In addition to metrics related to network exposure, there are several metrics related to other aspects of a 6G NS, including network performance, lifetime, billing and charging, monitoring, and many others. Taking into account all the metrics that are needed to be negotiated between both parties in a single SLA could be challenging, mostly in terms of the lengthy negotiation time. Hence, we propose a novel concept of exposure level agreement (ELA) within the context of the proposed framework. The ELA could be a similar document and may have the same structure as the SLA; however, it shall contain only the metrics related to the service and network exposure. We believe that such a customized document will help both parties efficiently negotiate network exposure and agree upon their relevant metrics. We will discuss the ELA in Section IV-H.

H. BUSINESS MANAGEMENT LAYER

The Business Management Layer is the last layer of the proposed architecture and is where the tenant and network operator manage and control all aspects related to the business and service management of a 6G NS. This layer specifically focuses on the management of contractual agreements, negotiations, complaint handling, invoicing, and all aspects related to service handling of a 6G NS provided to a tenant (or available to potential new tenants) in single or multiple administrative domains. The Business Management Layer is also responsible for developing and implementing business plans and strategies to ensure the long-term success of the 6G network. This would involve identifying market trends and opportunities, developing new services and applications, and working with stakeholders to ensure that the 6G network meets their needs. It also provides the tenant with a certain level of management capabilities, orchestration resources, monitoring tools, and execution authority through the Network Exposure Layer to design, manage, control, perform, monitor, and terminate a wide range of predefined tasks and a number of components related to a 6G NS. Aspects and tasks that are in the scope of this layer, such as QoS and QoE fulfillment, FCAPS reporting, billing, charging, and many others, are directly experienced by the tenant and stakeholders. Hence, it is important for the network operator to take measures to ensure that the tenants and third-parties are satisfied with the level of service they receive. These measures include, but are not limited to, ensuring SLAs are met, providing fair and transparent billing and charging, allowing flexibility in terms of service delivery and customization, addressing any incident in the network promptly, and delivering high-quality, reliable, and flexible 6G services. We illustrate the architecture of the Business Management Layer in Figure 12 and discuss its role in the proposed framework in the rest of this subsection.

The foremost important responsibility of the Business Management Layer is handling negotiations, agreements, contracts, and other business and legal processes related to a 6G NS between the tenants and the network operator (i.e., business-to-consumer situations), as well as between network operators and service providers (i.e., business-to-business situations), as illustrated in Figure 12. To avoid misunderstanding, it is essential to note that the Business Management Layer within our proposed framework plays a very minor role (or at times none at all) in the design of a 6G NS, the determination of metrics that are needed to be included in its corresponding SLA, and the definition of pre-deployment business and legal related matters. The aspects related to the design of an NS and its SLA are within the scope of the Design Layer (see Section IV-B). In a situation where the deployment of an NS is a matter between the tenant and network operator, this layer may involve delivering the agreed-upon level of performance, reliability, availability, and other various types of metrics for the tenant's specific NS (see Figure 12). In addition, this layer may also facilitate specific management-level and network-exposure-level capabilities to allow a tenant and a third-party to execute certain tasks within the proposed framework [137]. In a scenario where the provisioning of an NS crosses the boundaries of a network operator (such as in global coverage or public-private network infrastructure services), it is essential to equip the Business Management Layer of a network operator with the necessary capabilities to aggregate and compose resources, features, and capabilities from different Business Management Layer instances belonging to different 6G network operators [138].

One of the essential components of the Business Management Layer is the network operations center (NOC). As shown in Figure 12, the NOC is responsible for continuous monitoring of different segments, domains, and layers of the proposed framework and ensuring that all segments, domains, and layers are operating properly. The NOC is equipped with a wide range of automatic detection, smart monitoring, and intelligent management tools, including data-driven network management software, zero-touch performance monitoring tools, and zero-trust security tools [138], [139], [140]. These innovative tools and novel solutions enable the NOC to automatically detect and predict probable issues, failures, bottlenecks, and inefficient configurations, triggering and coordinating dynamic responses in the short and medium term. One of the primary responsibilities of the NOC is to ensure that each 6G NS is optimized in terms of performance and efficiency and fulfills the business-related requirements (among others) of the tenant. This involves monitoring network utilization, identifying bottlenecks, and making changes to the underlying network configuration to resolve any performance issues through the use of intent-based mechanisms, intelligent algorithms, and data- or model-driven solutions [141]. In order for the Business Management Layer to perform intent-based continuous monitoring of the operations of the proposed slicing framework, we anticipate that the NOC shall execute the following tasks in an intelligent and efficient manner: (a) collect and store raw data from different layers; (b) provide the necessary APIs to the tenants to monitor custom and user-defined metrics; and finally (c) deliver intelligent and ML monitoring and execution solutions to third-parties as well as to the rest of the layers of the network slicing architecture (see Figure 12).

The Business Management Layer is also responsible for billing and charging aspects of the proposed framework, as illustrated in Figure 12. This layer shall employ intelligent mechanisms and innovative approaches to billing and charging to support the diverse range of applications and services that will be enabled by 6G NSs. Billing and charging aspects of a 6G NS would depend on a variety of factors, including the type and level of service provided, the amount of data transferred, and the QoS guarantees offered [140]. One approach to billing and charging in 6G could be to use a usage-based model where tenants are charged based on the amount of data transferred, the number of devices connected, or the duration of the connection. Another approach could be to use a subscription-based model, where tenants pay a fixed fee for a defined level of service over a specific time period. Overall, the intricacies of billing and charging for a 6G NS will ultimately depend on the implementation and business models adopted by network operators and service providers [139], [140]. However, a key challenge for 6G slicing will be guaranteeing fair and accurate charging for the use of network resources. This would necessitate the development of new charging mechanisms and protocols that can accurately track and account for network resource usage across different 6G NSs and tenants.

Another key responsibility of this layer is the customer relationship management tasks. As shown in Figure 12, we assume customer relationship management is part of the NOC of the Business Management Layer. These tasks involve tracking customer usage and providing customer support. The grand objective of the Business Management Layer shall always be building and maintaining strong relationships with tenants and MVNOs, by understanding their needs and providing them with high-quality services and support. We believe that excellent customer relationship management is crucial for establishing solid customer relationships and ensuring long-term success in the 6G industry. By emphasizing personalized services, high-quality support, customer engagement, and data analytics, the Business Management

Layer may establish a customer-centric culture that provides extraordinary value to customers and encourages their loyalty and retention. Finally, the Business Management Layer also works closely with the marketing and sales teams to develop and implement marketing campaigns and sales strategies. To accomplish these objectives, the Business Management Layer may take some key steps to develop effective strategies, such as:

- conducting market research to identify potential 6G use cases and their specific requirements and priorities. This might involve surveying existing stakeholders, analyzing market trends and competitor activity, and gathering insights from industry experts and analysts.
- developing a compelling value proposition for the 6G network slicing framework that emphasizes the unique benefits and advantages it offers compared to other solutions in the market. This could include features such as improved performance, flexibility, customization, cost-effectiveness, and many others.
- defining specific target segments based on market research, and developing tailored marketing and sales strategies for each segment. For instance, different segments may have different needs and preferences related to performance, security, or pricing, and the Business Management Layer should develop messages and offers that resonate with each segment.

We believe that the above aspects and tasks (i.e., service-related, management-related, and exposure-related) within the scope of the Business Management Layer can be arranged into three types of agreements, signed between the network operator and tenant, and continuously monitored by the said layer. These three types of agreements are SLA, MLA, and ELA, which we introduced in the previous sections. The pre-deployment phases of the three types of agreements can be jointly performed mainly by the Design Layer, Management and Orchestration Layer, Network Exposure Layer, and Business Management Layer. However, during the operation, maintenance, and termination phases, the tenants and third-parties are directly connected through an intelligent and secure APIs to the Business Management Layer to monitor, manage, control, and execute certain pre-defined operations related to their corresponding 6G NSs within the proposed framework. In the rest of this subsection, we discuss how the Business Management Layer can assist tenants, third-parties, and network operators in accomplishing these goals.

1) THE 6G SLA DELIVERY, MANAGEMENT, AND IMPROVEMENT

Determining the performance metrics, defining the lifecycle management, identifying legal and regulatory requirements, specifying the types of 6G services, reaching an agreement on a business model, and numerous other aspects pertaining to the 6G services provided by a 6G NS to a third-party can be arranged in the SLA and be signed by both parties in the Design Layer of the proposed framework. However, once the SLA is signed and the 6G NS is in the operation phase, the Business Management Layer then takes over responsibility for all aspects of the 6G services, including technical, legal, business, etc., and is in direct contact with the tenant and the Business Management Layers of other network operators. In the Business Management Layer, as shown in Figure 12, the SLA should specify the QoS-related aspects of 6G services that the 6G network operator will offer the tenant during the lifetime of the ordered 6G NS. These factors may include network availability, network performance, security, privacy, service customization, flexibility, etc.

We believe that the SLA in the Business Management Layer shall be mainly focusing on service delivery, service management, service improvement, and service reporting aspects of a 6G NS (see Figure 12). These aspects, which are expected to be thoroughly defined in the SLA, shall be delivered by the Business Management Layer to the third-parties. Regarding service delivery, this layer is expected to deliver the 6G services specified in the SLA to the tenant and third parties in a timely, effective, and efficient manner. The service delivery can include ensuring that the 6G NS is available, performing according to the agreed-upon standards, and providing the expected level of tenant support. With respect to service management, the Business Management Layer is responsible for managing the 6G services provided to the tenant, including monitoring the performance, identifying and resolving any issues or incidents, and ensuring that the service meets the needs and requirements of the tenant throughout the lifetime of the ordered 6G NS. During the operation phase of the SLA, various types of incidents may happen to a 6G NS, as illustrated in Figure 12. Based on the severity of the incidents, we can arrange them into three categories [81]: minor incidents, major incidents, and critical incidents. If an incident occurs and the network operator successfully resolves it, the Business Management Layer is responsible for calculating the penalty on the basis of SLA and returning the cost of damage to the tenants and third-parties. In respect of service improvement, the Business Management Layer is responsible for continuously improving the 6G services provided to the tenant. This includes monitoring tenant feedback and usage patterns, identifying opportunities for improvement, and implementing changes to enhance the quality and value of the 6G service. Finally, this layer provides the tenants with regular reports on the performance of the 6G NSs and the 6G services provided to them. This includes providing metrics and KPIs that measure network availability, performance, and tenant satisfaction, as well as any other relevant information that the tenant may need to manage their business and connectivity.

2) THE ROLE OF MLA IN BUSINESS MANAGEMENT LAYER

In Section IV-C, we introduced the concept of MLA, which is a legally binding agreement at the M&O level and is signed between the service provider and tenant. Once the metrics, degree of autonomy, and access rights and levels of tenants have been defined in the MLA, the Business Management Layer then assumes responsibility for providing the tenant with the necessary M&O capabilities. The MLA serves as a critical document that defines the scope of the M&O capabilities, establishes expectations for performance and tenant support, and provides a framework for managing the relationship between the network operator and the tenant with respect to managing and orchestrating the ordered 6G NS. By outlining these aspects, the MLA helps to ensure that both parties have a shared understanding of their roles and responsibilities and can work together effectively to deliver high-quality M&O capabilities. In either case, whether as part of the SLA or as a separate formal agreement, the MLA shall specify and include only the aspects related to the M&O processes and operations that the tenant is expected to execute in all four management domains of the Management and Orchestration Layer.

We believe that, among many others, the MLA in the Business Management Layer shall typically cover four key aspects, as shown in Figure 12. First, it shall define the scope of the M&O capabilities to be provided by the network operator, as well as their levels of autonomy, standards, and support that the tenant can expect. Second, it shall specify the types of reports and metrics that the network operator will provide to the tenant in order to monitor the performance of the provided M&O capabilities. Third, it shall describe the procedures and processes for resolving disputes between the network operator and the tenant regarding the M&O capabilities. This includes escalation procedures, resolution timelines, and any other pertinent information. Finally, it shall outline the procedures and processes for managing changes to the M&O capabilities provided to the tenant. This includes the process for submitting change requests, timelines for review and approval, and any other relevant details.

3) THE ROLE OF ELA IN BUSINESS MANAGEMENT LAYER

In Section IV-G, we introduced the concept of ELA, which is used to define the level of network and service exposure required for a tenant to access and modify a particular service or application belonging to the 6G NS. Similar to SLA and MLA, the ELA is also a legal agreement between the tenant and the operator that includes metrics related to the network and service exposure aspects of a 6G NS. It can be a separate legal document or be a part of the SLA. The ELA must also include a set of metrics or exposure level objectives that define the minimum levels of network and service exposure performance and effectiveness that must be provided to support the service or application running on the 6G NS. Among others, these metrics may include API response time, API availability, API throughput, exposure activation time, and API exposure quality. Upon signing the ELA, the Business Management Layer is then responsible for continuously monitoring the performance and effectiveness of these metrics, as well as identifying areas where improvements can be made to optimize the performance of network exposure and support a wide range of applications and services for the 6G NS.

Figure 12 shows that this layer consists of several components that interact with each other, with the tenants, and with the remaining layers to enable network and service exposure solutions. The components of the ELA shall be designed to ensure that the service provider has the necessary network resources, capabilities, and solutions to support an intelligent and open network and service exposure while also providing the operator with clear guidelines for meeting the agreed-upon performance metrics and exposure level objectives. By defining and enforcing the ELA, both parties can work together to ensure that high-quality network and service exposure is provided to the tenants and third-parties. These components may vary depending on the specific service or application being exposed, but they are mainly exposure metrics, exposure level objectives, penalties for non-compliance, reporting and monitoring, and negotiation and revision processes.

V. RESEARCH CHALLENGES AND FUTURE DIRECTIONS

This section identifies a number of research challenges associated with the extension of network slicing architecture towards intelligent and open 6G. These challenges require substantial research efforts to meet the increasing requirements of futuristic networks and to enable emerging use cases and applications. According to their purpose, these challenges are classified into four categories, which are shown in Figure 13 and described below. They are: (a) Technological Challenges, which contain issues relating to three enabling technologies for intelligent slicing; (b) Infrastructural Challenges, which include issues pertaining to the underlying infrastructure that hosts 6G NSs; (c) Functional Challenges, which highlight challenges associated with network functionalities and capabilities; and (d) Operational Challenges, which consist of issues related to network operations and procedures.

A. TECHNOLOGICAL RESEARCH CHALLENGES

Extending AI-inspired virtualization to the edge and extreme edge: Notwithstanding the unprecedented interest in virtualizing edge and extreme edge domains through the utilization of ML algorithms, there are several challenges (which are enumerated in Figure 13) that must be confronted to realize nearly E2E virtualization in 6G slicing framework. While doing so, there is a critical need to model and deploy groundbreaking ML algorithms for zero-touch operations of the virtual components tailored to various types of 6G NSs at the edge or extreme edge of a 6G network, such as chaining, mapping, scaling (in and out or up and down), migration, etc.

Intelligent softwarization: SDx solutions were central to transitioning into 5G and will continue to prove essential well beyond 5G. Consequently, several challenges related to intelligent softwarization need to be addressed for the advancement of slicing in 6G, as listed in Figure 13. For instance, optimization of SDN controller placement that takes scalability and network KPIs into account, automated prediction and mitigation of control plane link failures, and integration and cross-coordination between the various SDx framework controllers within the RAN [61] are among the most challenging aspects for enabling intelligent softwarization in 6G slicing.

Intelligent cloud and edge computing: Intelligent task-offloading has been an important topic since the advent of cloud and edge computing, primarily the trade-off between power consumption and latency. In 6G, where various NSs are ubiquitously available and dependable, a diverse set of metrics will be considered when deciding between local and cloud computing at various hierarchical data centers [62]. For example, physical machines (PMs) in central data centers are typically more reliable and safer than edge nodes, owing to their extensive resource redundancy and security. They do, however, have higher latency and may introduce additional outage risks, e.g., from the backhaul [142]. Furthermore, since the PM has significantly greater access to user data in the central data center than at the network edge, ML algorithms can typically benefit from a more complete data set. However, they will suffer from decreased agility in responding to changes in local data patterns. This complex trade-off between multiple performance metrics necessitates an intelligent slice-specific task management approach that offloads computing tasks to the most appropriate computing unit based on the use case and flexibly schedules pending tasks based on their QoS requirements and the computing unit's load, to maximize overall utility while being aware of fairness. Along with the challenges outlined in Figure 13, reliable isolation of virtual compute and storage resources is required to protect end-user data against leakage and to ensure network performance independence between different 6G NSs.

B. INFRASTRUCTURAL RESEARCH CHALLENGES

Integrating advanced AI/ML algorithms into 6G slicing framework: Given the robustness, efficiency, and agility that automation and intelligence bring to 6G slicing, integrating AI/ML algorithms into various layers of the proposed framework is not a simple task [6]. To successfully adopt AI/ML algorithms, there are several research challenges (which are listed in Figure 13) that must yet be overcome. Additionally, we believe that a strong collaboration between regulatory authorities (national, regional, and international) for AI/ML algorithms and SDOs for 6G is essential to creating the required legal and ethical frameworks for dynamic interoperability among all layers of the proposed 6G slicing architecture.

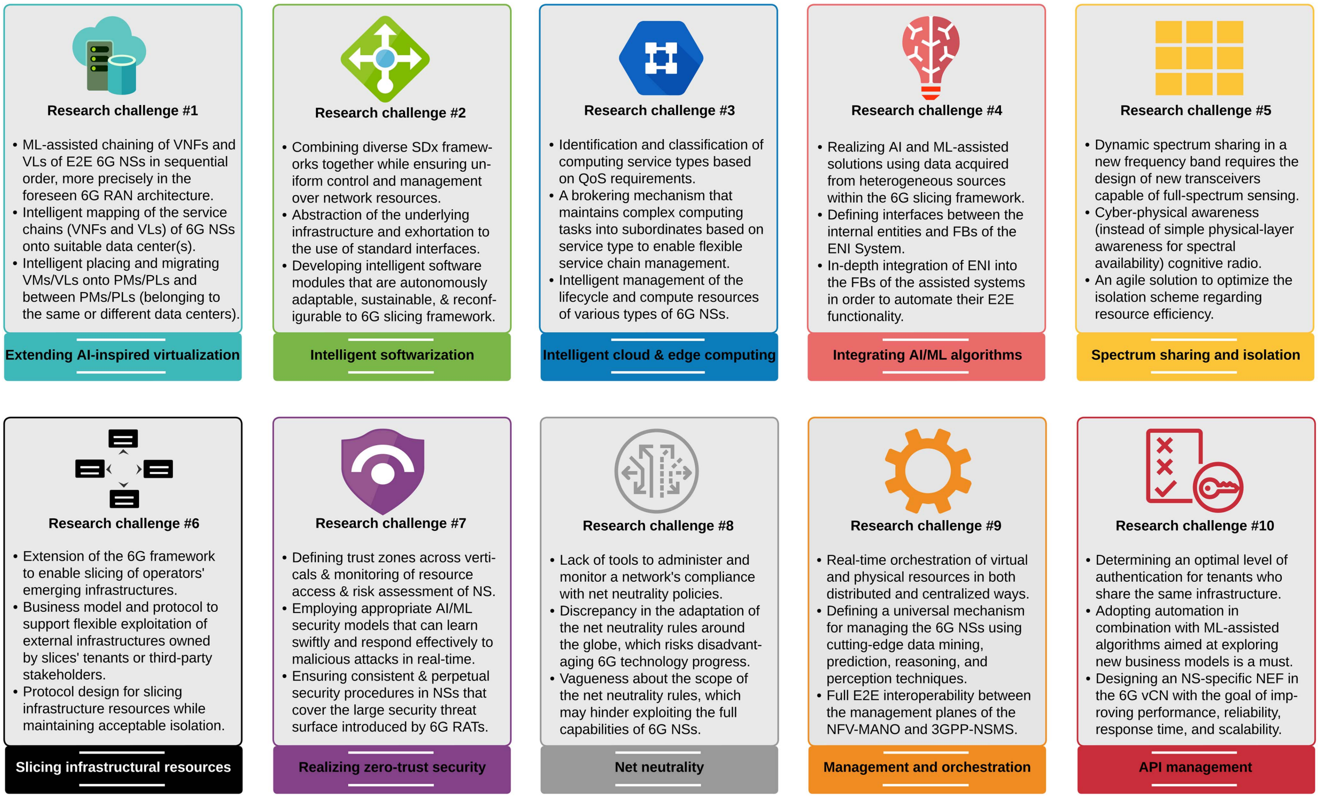
Spectrum sharing and isolation of physical resources: Network slicing is known to improve flexibility and resource isolation [1]. There is, however, a fundamental trade-off between the two benefits, particularly in the case of RAN physical resource slicing, such as spectrum sharing. Isolation in slicing framework can occur on a variety of levels [143], which we list here in descending order of degree: physical isolation, physical resource splitting, logical capacity delimitation, logical prioritization, and simple logical isolation. The greater the isolation level, the better the QoS of prioritized NSs can be guaranteed, but at the expense of spectral flexibility and utilization efficiency. Furthermore, even within the 6G NSs, multiple isolation approaches can be applied to different layers of its protocol stack. Finding an optimal isolation scheme that maximizes resource efficiency while still meeting the isolation requirement is a critical challenge in the legacy slicing framework. With the emerging use cases and applications introduced by 6G, as well as several challenges listed in Figure 13, this task may become even more complex to overcome.

Slicing infrastructural resources of emerging types: Most existing slicing approaches focus on radio and virtual resource slicing, leaving a huge research gap for E2E slicing, i.e., to efficiently slice infrastructural resources in the physical layer of a RAN protocol stack. Differing from physical resource blocks such as power, storage, and compute resources, which are relatively easy to split and assign to NSs, bare-metal resources such as antennas are extremely difficult to share flexibly among NSs while guaranteeing satisfactory isolation [143]. This and several other issues listed in Figure 13 can be amplified further by emerging 6G infrastructures and enabling technologies, such as high-altitude platforms, unmanned aerial vehicles, intelligent reflecting surfaces, etc. [5].

C. FUNCTIONAL RESEARCH CHALLENGES

Realizing zero-trust security paradigm: The security architecture for 5G slicing defined in [144] – including the management security and procedures for NSs – is aligned with the principles of the zero-trust security paradigm, which requires constant verification and scrutiny of virtually all elements accessing and leveraging an NS. By default, any tenant, device, service, or application is given the least privileges and needs to be explicitly authorized by the security management entity. However, as intelligent virtualization will be extended to the edge and extreme edge of 6G networks, and wireless NFs are expected to be deployed on multi-tenanted clouds in the 6G RAN, the futuristic slicing framework is likely to fall victim to high-impact cyberattacks due to (a) vulnerabilities resulting from the exploitation of heterogeneous wireless systems and technologies, (b) a failure to adhere to sound and secure practices, and (c) a lack of economic incentive to implement high-security solutions. Hence, several research challenges, which are listed in Figure 13, must be tackled to adopt the E2E zero-trust security paradigm in the 6G slicing framework.

Net neutrality: Numerous countries and regions around the world have defined and adopted regulations and guidelines with strict policies that ensure open and indiscriminate access to the Internet for all end-users. Net neutrality regulations apply mainly to traffic carried by public broadband



Research challenge #1

- ML-assisted chaining of VNFs and VLs of E2E 6G NSs in sequential order, more precisely in the foreseen 6G RAN architecture.
- Intelligent mapping of the service chains (VNFs and VLs) of 6G NSs onto suitable data center(s).
- Intelligent placing and migrating VMs/VLs onto PMs/PLs and between PMs/PLs (belonging to the same or different data centers).

Extending AI-inspired virtualization

Research challenge #2

- Combining diverse SDx frameworks together while ensuring uniform control and management over network resources.
- Abstraction of the underlying infrastructure and exhortation to the use of standard interfaces.
- Developing intelligent software modules that are autonomously adaptable, sustainable, & reconfigurable to 6G slicing framework.

Intelligent softwarization

Research challenge #3

- Identification and classification of computing service types based on QoS requirements.
- A brokering mechanism that maintains complex computing tasks into subordinates based on service type to enable flexible service chain management.
- Intelligent management of the lifecycle and compute resources of various types of 6G NSs.

Intelligent cloud & edge computing

Research challenge #4

- Realizing AI and ML-assisted solutions using data acquired from heterogeneous sources within the 6G slicing framework.
- Defining interfaces between the internal entities and FBs of the ENI System.
- In-depth integration of ENI into the FBs of the assisted systems in order to automate their E2E functionality.

Integrating AI/ML algorithms

Research challenge #5

- Dynamic spectrum sharing in a new frequency band requires the design of new transceivers capable of full-spectrum sensing.
- Cyber-physical awareness (instead of simple physical-layer awareness for spectral availability) cognitive radio.
- An agile solution to optimize the isolation scheme regarding resource efficiency.

Spectrum sharing and isolation

Research challenge #6

- Extension of the 6G framework to enable slicing of operators' emerging infrastructures.
- Business model and protocol to support flexible exploitation of external infrastructures owned by slices' tenants or third-party stakeholders.
- Protocol design for slicing infrastructure resources while maintaining acceptable isolation.

Slicing infrastructural resources

Research challenge #7

- Defining trust zones across verticals & monitoring of resource access & risk assessment of NS.
- Employing appropriate AI/ML security models that can learn swiftly and respond effectively to malicious attacks in real-time.
- Ensuring consistent & perpetual security procedures in NSs that cover the large security threat surface introduced by 6G RATs.

Realizing zero-trust security

Research challenge #8

- Lack of tools to administer and monitor a network's compliance with net neutrality policies.
- Discrepancy in the adaptation of the net neutrality rules around the globe, which risks disadvantaging 6G technology progress.
- Vagueness about the scope of the net neutrality rules, which may hinder exploiting the full capabilities of 6G NSs.

Net neutrality

Research challenge #9

- Real-time orchestration of virtual and physical resources in both distributed and centralized ways.
- Defining a universal mechanism for managing the 6G NSs using cutting-edge data mining, prediction, reasoning, and perception processes.
- Full E2E interoperability between the management planes of the NFV-MANO and 3GPP-NSMS.

Management and orchestration

Research challenge #10

- Determining an optimal level of authentication for tenants who share the same infrastructure.
- Adopting automation in combination with ML-assisted algorithms aimed at exploring new business models is a must.
- Designing an NS-specific NEF in the 6G vCN with the goal of improving performance, reliability, response time, and scalability.

API management

FIGURE 13. Several research challenges that must be addressed to extend the network slicing towards open and intelligent 6G.

Internet providers; other forms of traffic falling under "specialized services," such as private networks, are exempt. Also, net neutrality regulations permit "reasonable network management" measures. Depending on the scope of an NS, adhering to regulations without exploiting their ambiguities and loopholes may prove challenging in practice. A narrow interpretation of net neutrality regulations, however, risks hindering the advancement of 6G technologies. Some issues related to net neutrality in the context of 6G slicing are listed in Figure 13.

D. OPERATIONAL RESEARCH CHALLENGES

M&O: The 6G NS spans multiple administrative domains and network layers. The M&O plane in each domain and layer has a finite set of responsibilities for service orchestration and resource management [145]. While NFV-MANO and 3GPP-NSMS promise dynamic service orchestration and resource scaling throughout the NS lifetime, many tasks in both management systems are still executed manually by the tenant or service provider. Therefore, full dynamism, E2E automation, openness, and intelligent scaling would be critical features of the 6G slicing framework aimed at efficiently managing 6G NSs and orchestrating their required virtual and physical resources. To that end, we identified three challenges related to automated network management and intelligent service orchestration within the slicing framework (see Figure 13).

API management: Enabling tenants to design, create, deploy, monitor, and terminate their own NSs and services via a set of secure and open APIs is a unique capability that is expected to be provided by the NEF in legacy slicing framework. However, facilitating robust, scalable, secure, open, and tenant-friendly access to exposed applications and services of an NS while hiding network topology and protecting tenants' privacy is still challenging. With the enabling of 6G use cases and applications, this issue is anticipated to become more sophisticated and challenging. To modernize service exposure in 6G slicing framework, there is a critical need for intelligent and open solutions for identity management of homogeneous and heterogeneous NSs, network and service programmability, monitoring and maintenance of NS-specific events, and several other research challenges shown in Figure 13.

VI. CONCLUDING REMARKS

In this article, we extended the E2E architectural framework, concepts, and methods of network slicing towards open and intelligent 6G. To this end, we provided a number of compelling reasons and trends that motivated us to explore such an extension in order to enable 6G networks to be fully slicing-aware in the forthcoming decade. Then, we identified three enabling technologies that are essential to the realization of 6G slicing, and we explored how intelligence and openness can enhance them to empower the futuristic slicing framework, which is intended to host innumerable heterogeneous use cases and applications. For the purpose of provisioning customized and mutually isolated intelligent and open 6G NSs, we proposed a preliminary and standards-oriented slicing architecture that unifies the most recent de jure and de facto standards and harmonizes them on multiple layers. The proposed architecture embraces openness and employs ML algorithms, such as RL, DL, FL, etc., to eliminate vendor lock-in and empower the 6G slicing framework with advanced intelligence and automation capabilities in order to effectively manage workloads and intelligently make better-informed decisions. Finally, we identified several issues arising from the extension of open and intelligent slicing towards 6G.

ACKNOWLEDGMENT

The authors are grateful to the anonymous reviewers for their insightful comments and valuable remarks, which significantly improved the quality of this article, as well as to the Area Editor and Associate Editor for coordinating the peer-review process.

REFERENCES

[1] A. Ksentini and N. Nikaein, "Toward enforcing network slicing on RAN: Flexibility and resources abstraction," *IEEE Commun. Mag.*, vol. 55, no. 6, pp. 102–108, Jun. 2017.

[2] J. Du, C. Jiang, J. Wang, Y. Ren, and M. Debbah, "Machine learning for 6G wireless networks: Carrying forward enhanced bandwidth, massive access, and ultrareliable/low-latency service," *IEEE Veh. Technol. Mag.*, vol. 15, no. 4, pp. 122–134, Dec. 2020.

[3] E. Pateromichelakis et al., "End-to-end data analytics framework for 5G architecture," *IEEE Access*, vol. 7, pp. 40295–40312, 2019.

[4] V. Ziegler, H. Viswanathan, H. Flinck, M. Hoffmann, V. Räisänen, and K. Hätönen, "6G architecture to connect the worlds," *IEEE Access*, vol. 8, pp. 173508–173520, 2020.

[5] W. Jiang, B. Han, M. A. Habibi, and H. D. Schotten, "The road towards 6G: A comprehensive survey," *IEEE Open J. Commun. Soc.*, vol. 2, pp. 334–366, 2021.

[6] H. Yang, A. Alphones, Z. Xiong, D. Niyato, J. Zhao, and K. Wu, "Artificial-intelligence-enabled intelligent 6G networks," *IEEE Netw.*, vol. 34, no. 6, pp. 272–280, Nov./Dec. 2020.

[7] B. Zong, C. Fan, X. Wang, X. Duan, B. Wang, and J. Wang, "6G technologies: Key drivers, core requirements, system architectures, and enabling technologies," *IEEE Veh. Technol. Mag.*, vol. 14, no. 3, pp. 18–27, Sep. 2019.

[8] S. D'Oro, L. Bonati, M. Polese, and T. Melodia, "OrchestRAN: Network automation through orchestrated intelligence in the O-RAN," in *Proc. INFOCOM*, 2022, pp. 270–279.

[9] M. A. Habibi, F. Z. Yousaf, and H. D. Schotten, "Mapping the VNFs and VLs of a RAN slice onto intelligent PoPs in beyond 5G mobile networks," *IEEE Open J. Commun. Soc.*, vol. 3, pp. 670–704, 2022.

[10] B. Brik, K. Boutiba, and A. Ksentini, "Deep learning for B5G open radio access network: Evolution, survey, case studies, and challenges," *IEEE Open J. Commun. Soc.*, vol. 3, pp. 228–250, 2022.

[11] L. Wang, Z. Lu, H. Shao, M. Wan, and X. Wen, "Open wireless network architecture in radio access network," in *Proc. IEEE VTC Fall*, 2013, pp. 1–5.

[12] "Description of network slicing concept, version 1.0," NGMN-Alliance, Frankfurt, Germany, Jan. 13, 2016.

[13] M. A. Habibi, B. Han, and H. D. Schotten, "Network slicing in 5G mobile communication: Architecture, profit modeling, and challenges," in *Proc. ISWCS*, Bologna, Italy, 2017, pp. 1–6.

[14] B. M. Khorsandi et al., "Deliverable D1.3: Targets and requirements for 6G–initial E2E architecture," Hexa-X, Heidelberg, Germany, Feb. 28, 2022.

[15] W. Jiang and H. D. Schotten, "The KICK-OFF of 6G research worldwide: An overview," in *Proc. ICCC*, 2021, pp. 2274–2279.

[16] I. L. Pavon et al., "Deliverable D6.2–design of service management and orchestration functionalities," Hexa-X, Heidelberg, Germany, Apr. 29, 2022.

[17] T. Huang, W. Yang, J. Wu, J. Ma, X. Zhang, and D. Zhang, "A survey on green 6G network: Architecture and technologies," *IEEE Access*, vol. 7, pp. 175758–175768, 2019.

[18] K. B. Letaief, W. Chen, Y. Shi, J. Zhang, and Y.-J. Angela Zhang, "The roadmap to 6G: AI empowered wireless networks," *IEEE Commun. Mag.*, vol. 57, no. 8, pp. 84–90, Aug. 2019.

[19] Y. Zhou et al., "Service-aware 6G: An intelligent and open network based on the convergence of communication, computing and caching," *Dig. Commun. Netw.*, vol. 6, no. 3, pp. 253–260, 2020.

[20] X. Wang, X. Ren, C. Qiu, Y. Cao, T. Taleb, and V. C. M. Leung, "Net-in-AI: A computing-power networking framework with adaptability, flexibility, and profitability for ubiquitous AI," *IEEE Netw.*, vol. 35, no. 1, pp. 280–288, Jan./Feb. 2021.

[21] J. Mei, X. Wang, and K. Zheng, "An intelligent self-sustained RAN slicing framework for diverse service provisioning in 5G-beyond and 6G networks," *Intell. Conver. Netw.*, vol. 1, no. 3, pp. 281–294, 2020.

[22] A. Dogra, R. K. Jha, and S. Jain, "A survey on beyond 5G network with the advent of 6G: Architecture and emerging technologies," *IEEE Access*, vol. 9, pp. 67512–67547, 2021.

[23] P. P. Ray, "A perspective on 6G: Requirement, technology, enablers, challenges and future road map," *J. Syst. Architect.*, vol. 118, Sep. 2021, Art. no. 102180.

[24] Z. Feng, Z. Wei, X. Chen, H. Yang, Q. Zhang, and P. Zhang, "Joint communication, sensing, and computation enabled 6G intelligent machine system," *IEEE Netw.*, vol. 35, no. 6, pp. 34–42, Nov./Dec. 2021.

[25] J. Wu et al., "Toward native artificial intelligence in 6G networks: System design, architectures, and paradigms," Accessed: Apr. 18, 2023. https://arxiv.org/pdf/2103.02823.pdf.

[26] B. Tan, Y. Qian, H. Lu, D. Hu, Y. Xu, and J. Wu, "Toward a future network architecture for intelligence services: A Cyber digital twin-based approach," *IEEE Netw.*, vol. 36, no. 1, pp. 98–104, Jan./Feb. 2022.

[27] P. P. Ray, N. Kumar, and M. Guizani, "A vision on 6G-enabled NIB: Requirements, technologies, deployments, and prospects," *IEEE Wireless Commun.*, vol. 28, no. 4, pp. 120–127, Aug. 2021.

[28] P. Carbone, G. Dán, J. Gross, B. Göransson, and M. Petrova, "NeuroRAN: Rethinking Virtualization for AI-native radio access networks in 6G." [Online]. Available: https://arxiv.org/pdf/2104.08111.pdf.

[29] B. Yang et al., "Edge intelligence for autonomous driving in 6G wireless system: Design challenges and solutions," *IEEE Wireless Commun.*, vol. 28, no. 2, pp. 40–47, Apr. 2021.

[30] S. Shahzadi, M. Iqbal, and N. R. Chaudhry, "6G vision: Toward future collaborative cognitive communication (3C) systems," *IEEE Commun. Stand. Mag.*, vol. 5, no. 2, pp. 60–67, Jun. 2021.

[31] S. Shen, C. Yu, K. Zhang, J. Ni, and S. Ci, "Adaptive and dynamic security in AI-empowered 6G: From an energy efficiency perspective," *IEEE Commun. Stand. Mag.*, vol. 5, no. 3, pp. 80–88, Sep. 2021.

[32] W. Qi, Q. Li, Q. Song, L. Guo, and A. Jamalipour, "Extensive edge intelligence for future vehicular networks in 6G," *IEEE Wireless Commun.*, vol. 28, no. 4, pp. 128–135, Aug. 2021.

[33] X. Shen, J. Gao, W. Wu, M. Li, C. Zhou, and W. Zhuang, "Holistic network Virtualization and pervasive network intelligence for 6G," *IEEE Commun. Surveys Tuts.*, vol. 24, no. 1, pp. 1–30, 1st Quart., 2022.

[34] H. Chergui et al., "Zero-touch AI-driven distributed management for energy-efficient 6G massive network slicing," *IEEE Netw.*, vol. 35, no. 6, pp. 43–49, Nov./Dec. 2021.

[35] H. Chergui, A. Ksentini, L. Blanco, and C. Verikoukis, "Toward zero-touch management and orchestration of massive deployment of network slices in 6G," *IEEE Wireless Commun.*, vol. 29, no. 1, pp. 86–93, Feb. 2022.

[36] Z. Ren, X. Li, Q. Jiang, Y. Wang, J. Ma, and C. Miao, "Network slicing in 6G: An authentication framework for unattended terminals," *IEEE Netw.*, vol. 37, no. 1, pp. 78–86, Jan./Feb. 2023.

[37] M. Corici, E. Troudt, and T. Magedanz, "An organic 6G core network architecture," in *Proc. ICIN*, 2022, pp. 1–7.

[38] M. Tariq, F. Naeem, and H. V. Poor, "Toward experience-driven traffic management and orchestration in digital-twin-enabled 6G networks," 2022. [Online]. Available: https://arxiv.org/pdf/2201.04259.pdf.

[39] W. Wu et al., "AI-native network slicing for 6G networks," *IEEE Wireless Commun.*, vol. 29, no. 1, pp. 96–103, Feb. 2022.

[40] Y. Yang et al., "6G network AI architecture for everyone-centric customized services," *IEEE Netw.*, early access, Jul. 25, 2022, doi: 10.1109/MNET.124.2200241.

[41] G. Liu, Na Li, J. Deng, Y. Wang, J. Sun, and Y. Huang, "The SOLIDS 6G mobile network architecture: Driving forces, features, and functional topology," *Engineering*, vol. 8, pp. 42–59, Jan. 2022.

[42] T. Taleb et al., "6G system architecture: A service of services vision," *ITU J. Future Evol. Technol.*, vol. 3, no. 3, pp. 710–743, Dec. 2022.

[43] M. A. Habibi, B. Han, F. Z. Yousaf, and H. D. Schotten, "How should network slice instances be provided to multiple use cases of a single vertical industry?" *IEEE Commun. Stand. Mag.*, vol. 4, no. 3, pp. 53–61, Sep. 2020.

[44] "IMT vision-framework and overall objectives of the future development of IMT for 2020 and beyond," ITU, Geneva, Switzerland, Sep. 29, 2015.

[45] P. Popovski, K. F. Trillingsgaard, O. Simeone, and G. Durisi, "5G wireless network slicing for eMBB, URLLC, and mMTC: A communication-theoretic view," *IEEE Access*, vol. 6, pp. 55765–55779, 2018.

[46] H. Tataria, M. Shafi, A. F. Molisch, M. Dohler, H. Sjöland, and F. Tufvesson, "6G wireless systems: Vision, requirements, challenges, insights, and opportunities," *Proc. IEEE*, vol. 109, no. 7, pp. 1166–1199, Jul. 2021.

[47] B. Han et al., "Digital twins for industry 4.0 in the 6G era." Accessed: Feb. 26, 2023. https://arxiv.org/abs/2210.08970.

[48] M. Adhikari, A. Hazra, V. G. Menon, B. K. Chaurasia, and S. Mumtaz, "A Roadmap of next-generation wireless technology for 6G-enabled vehicular networks," *IEEE Internet Things Mag.*, vol. 4, no. 4, pp. 79–85, Dec. 2021.

[49] C. D. Alwis et al., "Survey on 6G frontiers: Trends, applications, requirements, technologies and future research," *IEEE Open J. Commun. Soc.*, vol. 2, pp. 836–886, 2021.

[50] M. Naimi, M. A. Habibi, and H. D. Schotten, "Platoon–assisted vehicular cloud in VANET: Vision and challenges," ESCC, Paris, France, Sep. 2019.

[51] "Management and orchestration; management capabilities (release 17)," 3GPP, Sophia Antipolis, France, 3GPP Rep. TS 128 537 V17.2.0, Jun. 2022.

[52] S. Sambhwani et al., "Transitioning to 6G part 1: Radio technologies," *IEEE Wireless Commun.*, vol. 29, no. 1, pp. 6–8, Feb. 2022.

[53] M. Elsayed and M. Erol-Kantarci, "AI-enabled future wireless networks: Challenges, opportunities, and open issues," *IEEE Veh. Technol. Mag.*, vol. 14, no. 3, pp. 70–77, Sep. 2019.

[54] B. Han, W. Jiang, M. A. Habibi, and H. D. Schotten, "An abstracted survey on 6G: Drivers, requirements, efforts, and enablers," in *Proc. Workshop Next Gener. Netw. Appl.*, 2020, pp. 1–6.

[55] "Network slicing use cases requirements," GSMA, London, U.K., Aug. 19, 2018.

[56] M. Giordani, M. Polese, M. Mezzavilla, S. Rangan, and M. Zorzi, "Toward 6G networks: Use cases and technologies," *IEEE Commun. Mag.*, vol. 58, no. 3, pp. 55–61, Mar. 2020.

[57] K. David and H. Berndt, "6G vision and requirements: Is there any need for beyond 5G?" *IEEE Veh. Technol. Mag.*, vol. 13, no. 3, pp. 72–80, Sep. 2018.

[58] B. Ji et al., "Several key technologies for 6G: Challenges & opportunities," *IEEE Commun. Stand. Mag.*, vol. 5, no. 2, pp. 44–51, Jun. 2021.

[59] M. Bagaa, T. Taleb, J. Riekki, and J. Song, "Collaborative cross system AI: Toward 5G system and beyond," *IEEE Netw.*, vol. 35, no. 4, pp. 286–294, Jul./Aug. 2021.

[60] M. Condoluci and T. Mahmoodi, "Softwarization and virtualization in 5G mobile networks: Benefits, trends and challenges," *Comput. Netw.*, vol. 146, pp. 65–84, Dec. 2018.

[61] R. F. Moyano, D. Fernández, N. Merayo, C. M. Lentisco, and A. Cárdenas, "NFV and SDN-based differentiated traffic treatment for residential networks," *IEEE Access*, vol. 8, pp. 34038–34055, 2020.

[62] Y. Xiao, G. Shi, Y. Li, W. Saad, and H. V. Poor, "Toward self-learning edge intelligence in 6G," *IEEE Commun. Mag.*, vol. 58, no. 12, pp. 34–40, Dec. 2020.

[63] "M.3010; principles for a telecommunications management network," ITU, Geneva, Switzerland, Feb. 4, 2000.

[64] C. D. Lima et al., "Convergent communication, sensing and Localization in 6G systems: An overview of technologies, opportunities and challenges," *IEEE Access*, vol. 9, pp. 26902–26925, 2021.

[65] X. Zhou, R. Li, T. Chen, and H. Zhang, "Network slicing as a service: Enabling enterprises' own software-defined cellular networks," *IEEE Commun. Mag.*, vol. 54, no. 7, pp. 146–153, Jul. 2016.

[66] "Service requirements for the 5G systems; stage 1 (version 19.0.0 release 19)," 3GPP, Sophia Antipolis, France, 3GPP, Rep. TS 22.261 version 15.9.0, Sep. 23, 2022.

[67] C. Ebert, G. Gallardo, J. Hernantes, and N. Serrano, "DevOps," *IEEE Soft.*, vol. 33, no. 3, pp. 94–100, May/Jun. 2016.

[68] S. Garg and S. Garg, "Automated cloud infrastructure, continuous integration and continuous delivery using docker with robust container security," in *Proc. IEEE CMIPR*, 2019, pp. 467–470.

[69] A. M. Sanchez, A.-S. Charismiadis, D. Tsolkas, D. A. Guillen, and J. G. Rodrigo, "Offering the 3GPP common API framework as microservice to vertical industries," in *Proc. EuCNC 6G Summit*, 2022, pp. 363–368.

[70] "Open digital architecture," TMF, London, U.K., Feb. 2022.

[71] A. Cárdenas, D. Fernández, C. M. Lentisco, R. F. Moyano, and L. Bellido, "Enhancing a 5G network slicing management model to improve the support of mobile virtual network operators," *IEEE Access*, vol. 9, pp. 131382–131399, 2021.

[72] S. Zhang, "An overview of network slicing for 5G," *IEEE Wireless Commun.*, vol. 26, no. 3, pp. 111–117, Jun. 2019.

[73] "From vertical industry requirements to network slice characteristics," GSMA, London, U.K.: Jan. 10, 2020.

[74] F. Schorr and L. Hvam, "The use of design-science to define information content requirements for IT service catalogs," in *Proc. IEEE IEEM*, 2018, pp. 497–501.

[75] N. Nazarzadeoghaz, F. Khendek, and M. Toeroe, "Automated design of network services from network service requirements," in *Proc. ICIN*, 2020, pp. 63–70.

[76] M. Muszynski, S. Lugtigheid, F. Castor, and S. Brinkkemper, "A study on the software architecture documentation practices and maturity in open-source software development," in *Proc. IEEE ICSA*, 2022, pp. 47–57.

[77] X. Li et al., "Automated service provisioning and hierarchical SLA management in 5G systems," *IEEE Trans. Netw. Service Manag.*, vol. 18, no. 4, pp. 4669–4684, Dec. 2021.

[78] H. Zhou, C. de Laat, and Z, Zhao, "Trustworthy cloud service level agreement enforcement with blockchain based smart contract," in *Proc. IEEE CloudCom*, 2018, pp. 255–260.

[79] F. Javed and J. Mangues-Bafalluy, "Demo: Blockchain-based inter-provider agreements for 6G networks," in *Proc. CNSM*, 2022, pp. 364–366.

[80] B. Veith, D. Krummacker, and H. D. Schotten, "The road to trustworthy 6G: A survey on trust anchor technologies," *IEEE Open J. Commun. Soc.*, vol. 4, pp. 581–595, 2023.

[81] M. A. Habibi, B. Han, M. Nasimi, and H. D. Schotten, "The structure of service level agreement of slice-based 5G network," in *Proc. IEEE PIMRC*, Bologna, Italy, Sep. 2018, pp. 1–6.

[82] Ramneek and S. Pack, "A credible service level agreement enforcement framework for 5G edge," in *Proc. INFOCOM*, 2021, pp. 1–2.

[83] H. Yang, Y. Hu, R. Liu, K. Yao. X. Duan, and T. Sun, "A technical research towards 5G SLA: System definition, sense and assurance," in *Proc. IEEE ICCT*, 2021, pp. 462–471.

[84] "5G; management and orchestration; architecture framework (version 16.4.0 release 16)," 3GPP, Sophia Antipolis, France, 3GPP Rep. TS 28.533, Aug. 30, 2020.

[85] M. A. Habibi et al., "Enabling network and service programmability in 6G mobile communication systems," in *Proc. IEEE FNWF*, 2022, pp. 320–327.

[86] "Telecommunication management; study on management and orchestration of network slicing for next generation network (version 15.1.0 release 15)," 3GPP, Sophia Antipolis, France, 3GPP Rep. TR 28.801, Jan. 2018.

[87] "ZSM; end-to-end management and orchestration of network slicing (version 1.1.1)," ETSI, Sophia Antipolis, France, Rep. ETSI GS ZSM 003, Jun. 8, 2021.

[88] "Management and orchestration; concepts, use cases and requirements (version 17.3.0 release 17)," 3GPP, Sophia Antipolis, France, 3GPP Rep. TS 28.530, Sep. 30, 2022.

[89] "ZSM; reference architecture (version 1.1.1)," ETSI, Sophia Antipolis, Rep. ETSI GS ZSM 002, France, Aug. 13, 2019.

[90] "ZSM; means of automation (version 1.1.1)," ETSI, Sophia Antipolis, France, Rep. ETSI GR ZSM 005, May 19, 2020.

[91] "NFV release 3; architecture; report on the enhancements of the NFV architecture towards "cloud-native" and "PaaS" (version 3.3.1)," ETSI, Sophia Antipolis, France, Rep. ETSI GR NFV-IFA 029, Nov. 2019.

[92] "Management and orchestration; management data Analytics (MDA) (release 17)," 3GPP, Sophia Antipolis, France, 3GPP Rep. TS 28.104 Sep. 27, 2022.

[93] "O-RAN.WG1; technical specifications; O-RAN architecture description," ORAN, Alfter, Germany, Oct. 21, 2022.

[94] M. A. Habibi, M. Nasimi, B. Han, and H. D. Schotten, "A comprehensive survey of RAN architectures toward 5G mobile communication system," *IEEE Access*, vol. 7, pp. 70371–70421, 2019.

[95] M. A. Habibi, B. Han, M. Nasimi, N. P. Kuruvatti, A. Fellan, and H. D. Schotten, "Towards a fully Virtualized, Cloudified, and slicing-aware RAN for 6G mobile networks," in *6G Mobile Wireless Networks* (Computer Communications and Networks). Cham, Switzerland: Springer, 2021, pp. 327–358.

[96] "Automation and optimisation for OpenRAN," TIP White Paper, Boston, MA, USA, Jan. 2022.

[97] "Near-real-time RAN intelligent controller and E2 interface; near-RT RIC architecture," ORAN, Alfter, Germany, Rep. WG3 V03.00, Oct, 2022.

[98] "Non-RT RIC and A1 interface; non-RT RIC architecture," ORAN, Alfter, Germany, Rep. WG2 V02.01, Oct. 2022.

[99] "Slicing arch," ORAN, Alfter, Germany, Rep. WG1 V08.00, Oct. 2022.

[100] "Management and orchestration; study on enhancement of MDA (release 17)," 3GPP, Sophia Antipolis, France, 3GPP Rep. TR 28.809, Mar. 25, 2021.

[101] S. Bhattacharjee et al., "Network slicing for TSN-based transport networks," *IEEE Access*, vol. 9, pp. 62788–62809, 2021.

[102] "Subscriber network slice service and attributes," MEF, Los Angeles, CA, USA, Rep. 84, Jun. 2021.

[103] "Technical specifications; management interfaces for transport network elements," ORAN, Alfter, Germany, Rep. WG9 V04.00 21, Jul. 2022.

[104] "TIP OOPT MUST optical Whitepaper; target architecture: Disaggregated open optical networks," Telecom Infra Project, Wakefield, MA, USA, Jul. 2021.

[105] "IETF definition of transport slice; draft-nsdt-teas-transport-slice-definition-04," IETF, Fremont, CA, USA, Sep. 2020.

[106] "Mobile-transport network slice instance management interfaces," BBF, Fremont, CA, USA, BBF Rep. 522, Jun. 25, 2022.

[107] "Applying SDO architecture to 5G slicing," ONF, Palo Alto, CA, USA, Rep. 526, Apr. 14, 2016.

[108] "ETSI GR NFV-IFA 022; NFV release 3; management and orchestration; report on management and connectivity for multi-site services (version 3.1.1)," ETSI, Sophia Antipolis, France, Apr. 5, 2018.

[109] "Framework for IETF network slices; draft-ietf-teas-ietf-network-slices-16," IETF, Fremont, CA, USA, Oct. 2022.

[110] "Open all-Photonic network functional architecture (version 1.0)," IOWN global forum, Wakefield, MA, USA, Jan. 27, 2022.

[111] "Service function chaining use cases in mobile networks; draft-ietf-sfc-use-case-mobility-09," IETF, Fremont, CA, USA, Jul. 2019.

[112] H. Hantouti, N. Benamar, and T. Taleb, "Service function chaining in 5G & beyond networks: Challenges and open research issues," *IEEE Netw.*, vol. 34, no. 4, pp. 320–327, Jul./Aug. 2020.

[113] "Service function chaining architecture," IETF, Fremont, CA, USA, Oct. 2015.

[114] H. Hantouti, N. Benamar, T. Taleb, and A. Laghrissi, "Traffic steering for service function chaining," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 1, pp. 487–507, 1st Quart., 2019.

[115] A. M. Medhat, T. Taleb, A. Elmangoush, G. A. Carella, S. Covaci, and T. Magedanz, "Service function chaining in next generation networks: State of the art and research challenges," *IEEE Commun. Mag.*, vol. 55, no. 2, pp. 216–223, Feb. 2017.

[116] "ETSI GS NFV-IFA 014; NFV release 4; management and orchestration; network service templates specification (version 4.2.1)," ETSI, Sophia Antipolis, France, Jun. 2022.

[117] "Study on enhancement of support for edge computing in 5GC (release 17)," 3GPP, Sophia Antipolis, France, 3GPP Rep. TR 23.748, Dec. 2020.

[118] "Study on user plane protocol in 5GC (release 16)," 3GPP, Sophia Antipolis, France, 3GPP Rep. TR 29.892, Sep. 2019.

[119] "NG-RAN; architecture description (release 17)," 3GPP, Sophia Antipolis, France, 3GPP Rep. TS 38.401, Jan. 6, 2023.

[120] M. Nasimi, M. A. Habibi, B. Han, and H. D. Schotten, Edge-Assisted Congestion Control Mechanism for 5G Network Usin, "Edge-assisted congestion control mechanism for 5G network using software-defined net," in *Proc. ISWCS*, 2018, pp. 1–5.

[121] "System architecture for the 5G system (5GS); stage 2 (release 18)," 21 3GPP, Sophia Antipolis, France, 3GPP Rep. TS 23.501, Dec. 2022.

[122] "ETSI GS NFV-INF 001; NFV; infrastructure overview," ETSI, Sophia Antipolis, France, Dec. 2018.

[123] "ETSI GR MEC 031; MEC 5G integration," ETSI, Sophia Antipolis, France, Oct. 2020.

[124] "O-RAN; technical report; O-RAN working group 6 (Cloudification and orchestration) cloud architecture and deployment scenarios for O-RAN Virtualized RAN," ORAN, Alfter, Germany, Jul. 2022.

[125] "Small cell forum; 5G use cases; URLLC and slicing in 5G small cell networks," SCF, Bradenton, FL, USA, Feb. 2018.

[126] M. K. Shehzad, L. Rose, M. Majid Butt, I. Z. Kovács, M. Assaad, and M. Guizani, "Artificial intelligence for 6G networks: Technology advancement and Standardization," *IEEE Veh. Technol. Mag.*, vol. 17, no. 3, pp. 16–25, Sep. 2022.

[127] "5G; architecture enhancements for 5G system (5GS) to support network data Analytics services (version 16.4.0 release 16)," 3GPP, Sophia Antipolis, France, 3GPP Rep. TS 23.288, 15 June 2022.

[128] Y. Jeon, H. Jeong, S. Seo, T. Kim, H. Ko, and S. Pack, "A distributed NWDAF architecture for federated learning in 5G," in *Proc. IEEE ICCE*, 2022, pp. 1–2.

[129] "ETSI GS ENI 005; experiential networked intelligence (ENI); system architecture," v1.1.1, ETSI, Sophia Antipolis, France, Sep. 2019.

[130] D. Fragkos et al., "NEFSim: An open experimentation framework Utilizing 3GPP's exposure services," in *Proc. EuCNC 6G Summit)*, 2022, pp. 303–308.

[131] "Technical specification group core network and terminals; 5G system; network exposure function southbound services; stage 3 (release 18)," 3GPP, Sophia Antipolis, France, 3GPP Rep. TS 29.591, Dec. 16, 2022.

[132] L. Lin, B. Zhu, Q. Wang, L. Xu, and J. Mu, "A novel 5G core network capability exposure method for telecom operator," in *Proc. IEEE ISPA/BDCloud/SocialCom/SustainCom*, 2020, pp. 1450–1454.

[133] P. Li and Y. Xing, "Capability exposure Vitalizes 5G network," in *Proc. IWCMC*, 2021, pp. 874–878.

[134] P. Porambage, G. Gür, D. P. Moya Osorio, M. Liyanage, A. Gurtov, and M. Ylianttila, "The roadmap to 6G security and privacy," *IEEE Open J. Commun. Soc.*, vol. 2, pp. 1094–1122, 2021.

[135] S. A. Chaudhry et al., "A lightweight authentication scheme for 6G-IoT enabled maritime transport system," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 2, pp. 2401–2410, Feb. 2023.

[136] "Technical specification group core network and terminals; 5G system; network exposure function northbound APIs; stage 3 (release 18)," 3GPP, Sophia Antipolis, France, 3GPP Rep. TS 29.522, Dec. 16, 2022.

[137] S. S. Yrjölä, P. Ahokangas, and M. Matinmikko-Blue, "Value creation and capture from technology innovation in the 6G era," *IEEE Access*, vol. 10, pp. 16299–16319, 2022.

[138] M. Giordani, M. Polese, A. Laya, E. Bertin, and M. Zorzi, *6G Drivers for B2B Market*. Hoboken, NJ, USA: Wiley, 2022, pp. 9–22.

[139] S. S. Yrjölä, M. Matinmikko-Blue, and P. Ahokangas, "How could 6G transform engineering platforms towards Ecosy. business models?" in *Proc. 6G SUMMIT*, 2020, pp. 1–5.

[140] G. M. Karam, M. Gruber, I. Adam, F. Boutigny, Y. Miche, and S. Mukherjee, "The evolution of networks and management in a 6G world: An inventor's view," *IEEE Trans. Netw. Serv. Manag.*, vol. 19, no. 4, pp. 5395–5407, Dec. 2022.

[141] Nidhi, A. Mihovska, A. Kumar, and R. Prasad, "Business opportunities for beyond 5G and 6G networks," in *Proc. WPMC*, 2022, pp. 543–548.

[142] B. Han, S. Wong, C. Mannweiler, M. R. Crippa, and H. D. Schotten, "Context-awareness enhances 5G multi-access edge computing reliability," *IEEE Access*, vol. 7, pp. 21290–21299, 2019.

[143] D. Marabissi and R. Fantacci, "Highly flexible RAN slicing approach to manage isolation, priority, efficiency," *IEEE Access*, vol. 7, pp. 97130–97142, 2019.

[144] "5G; security architecture and procedures for 5G system (version 15.4.0 release 15)," 3GPP, Sophia Antipolis, France, 3GPP Rep. TS 33.501, Jun. 17, 2022.

[145] M. A. Habibi et al., "The architectural design of service management and orchestration in 6G communication systems," in *Proc. INFOCOM*, New York, NY, USA, May 2023, pp. 1–2.

MOHAMMAD ASIF HABIBI received the B.Sc. degree in telecommunication engineering from Kabul University, Afghanistan, in 2011, and the M.Sc. degree in systems engineering and informatics from the Czech University of Life Sciences, Czech Republic, in 2016. He is currently pursuing the Ph.D. degree with the Division of Wireless Communications and Radio Navigation, Rheinland-Pfälzische Technische Universität (previously known as Technische Universität Kaiserslautern), Germany, where he has been working as a Research Fellow since January 2017. From 2011 to 2014, he worked as a Radio Access Network Engineer with Huawei. His main research interests include network slicing, network function virtualization, resource allocation, machine learning, and radio access network architecture.

BIN HAN (Senior Member, IEEE) received the B.E. degree from Shanghai Jiao Tong University in 2009, the M.Sc. degree from Technische Universität Darmstadt in 2012, and the Ph.D. (Dr.-Ing.) degree from Karlsruher Institute für Technologie in 2016. He joined Rheinland-Pfälzische Technische Universität (previously known as Technische Universität Kaiserslautern) in 2016, where he is currently a Senior Lecturer with the Division of Wireless Communications and Radio Navigation. Researching in the area of wireless communication and networking, he has authored over around 50 research papers and book chapters, and participated in multiple EU collaborative research projects. He serves as an Editor for *Network*, the TPC Chair of the Workshop on Next Generation Networks and Applications, and a TPC Member of GLOBECOM, EuCNC, and European Wireless.

AMINA FELLAN received the B.Sc. degree in electrical engineering from the Ajman University of Science and Technology, UAE, in 2011, and the M.Sc. degree in communication engineering from RWTH Aachen University, Germany, in 2015. She is currently working as a Research Assistant with the Rheinland-Pfälzische Technische Universität (previously known as Technische Universität Kaiserslautern), researching software-defined radios, networks, within the context of industrial networks in 5G and beyond.

WEI JIANG (Senior Member, IEEE) received the Ph.D. degree in computer science from the Beijing University of Posts and Telecommunications in 2008. From 2008 to 2012, he was with the 2012 Laboratory, Huawei Technologies. From 2012 to 2015, he was with the Institute of Digital Signal Processing, University of Duisburg–Essen, Germany. Since 2015, he has been a Senior Researcher with the German Research Center for Artificial Intelligence (DFKI), which is the biggest European AI research institution and is the birthplace of "Industry 4.0" strategy. Meanwhile, he is a Senior Lecturer with Rheinland-Pfälzische Technische Universität (previously known as Technische Universität Kaiserslautern), Germany. He is the author of three book chapters and over 80 conference and journal papers and holds around 30 granted patents. He currently serves as an Editor for IEEE ACCESS and is a Moderator for IEEE TechRxiv.

ADRIÁN GALLEGO SÁNCHEZ received the B.Sc. degree in telematic engineering and the M.Sc. degree in telecommunications engineering with a specialization in defense and cybersecurity from the University of Alcalá de Henares, Spain, in 2017 and 2019, respectively. From 2018 to 2021, he was a Researcher with the NETCOM Research Group, University Carlos III of Madrid, where he was involved in 5GPPP H2020 ICT-17 European 5G projects. Since 2021, he has been a Research Specialist in the IT and Defense industries and has collaborated on several NATO NIAGs and Horizon Europe B5G/6G projects. He is currently working with the Research and Innovation Department, ATOS Spain. His main research interests include B5G/6G mobile networks, intelligent networks, network slicing, and AI.

IGNACIO LABRADOR PAVÓN received the degree in industrial engineering from the Polytechnic University of Madrid, Spain, in 1998. He has been in the IT and electronics industry for more than 30 years, especially in the mobile telecommunications sector, developing value-added services for different mobile network operators. During this time, he worked in multiple technical areas playing different roles. He currently works as a Research and Innovation Engineer with ATOS Spain. His research focuses on mobile networks M&O. He has participated in multiple EU collaborative projects. His most recent assignment is in the Hexa-X project, leading the work package related to intelligent orchestration and service management for future B5G/6G networks.

AMINA BOUBENDIR received the master's degree in network architecture and cybersecurity and the Ph.D. degree in computer science and networks from Télécom Paris in 2013 and 2016, respectively. She is a Researcher and a Project Manager with Orange Labs Networks, France. Her research is in the area of network design and management and particularly on network softwarization and automation. She is a member of the Orange Expert community on "Networks of Future."

HANS D. SCHOTTEN (Member, IEEE) received the Diploma and Ph.D. degrees in electrical engineering from the Aachen University of Technology, Germany, in 1990 and 1997, respectively. Since August 2007, he has been a Full Professor and the Head of the Division of Wireless Communications and Radio Navigation, Rheinland-Pfälzische Technische Universität (previously known as Technische Universität Kaiserslautern). Since 2012, he has also been a Scientific Director of the German Research Center for Artificial Intelligence, heading the Intelligent Networks Department. He was a Senior Researcher, the Project Manager, and the Head of the research groups with the Aachen University of Technology, Ericsson Corporate Research, and Qualcomm Corporate R&D. During his time at Qualcomm, he has also been the Director for Technical Standards and the Coordinator of Qualcomm's activities in European research programs.